

“© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”

A stepwise based fuzzy regression procedure for developing customer preference models in new product development

Kit Yan Chan, *Member IEEE*, Hak Keung Lam, *Senior Member IEEE*, Tharam S. Dillon, *Life Fellow IEEE*, Sai Ho Ling, *Senior Member IEEE*,

Abstract — Fuzzy regression methods have commonly been used to develop consumer preferences models which correlate the engineering characteristics with consumer preferences regarding a new product; the consumer preference models provide a platform whereby product developers can decide the engineering characteristics in order to satisfy consumer preferences prior to developing the products. Recent research shows that these fuzzy regression methods are commonly used to model customer preferences. However, these approaches have a common limitation in that they do not investigate the appropriate polynomial structure which includes significant regressors with only significant engineering characteristics; also, they cannot generate interaction or high-order regressors in the models. The inclusion of insignificant regressors is not an effective approach when developing the models. Exclusion of significant regressors may affect the generalization capability of the consumer preference models.

In this paper, a novel fuzzy modelling method is proposed, namely fuzzy stepwise regression (F-SR), in order to develop a customer preference model which is structured with an appropriate polynomial which includes only significant regressors. Based on the appropriate polynomial structure, the fuzzy coefficients are determined using the fuzzy least square regression. The developed fuzzy regression model attempts to obtain a better generalization capability using a smaller number of regressors. The effectiveness of the F-SR is evaluated based on two design problems, namely a tea maker design and a solder paste dispenser design. Results show that better generalization capabilities can be obtained compared with the fuzzy regression methods commonly-used for new product development. Also, smaller-scale consumer preference models with fewer engineering characteristics can be obtained. Hence, a simpler and more effective product development platform can be provided.

Index Terms— fuzzy regression, stepwise regression, new product development, consumer preferences, engineering characteristics, customer satisfaction, fuzzy least square regression

I. INTRODUCTION

Increasingly challenging competitiveness in the consumer market has dramatically changed the technologies of product

planning processes from production-centralized to consumer-driven [28]. The product planning process needs to be performed at the outset of new product development and mainly involves three major tasks: identification of customers and markets to be targeted, specification of new products to be developed, and determination of appropriate engineering characteristics of the new products. This paper focuses on the third task which is aimed to determine the appropriate engineering characteristics of new products in order to satisfy consumer preferences of the product which is given high priority in new product development. In fact, the role of product developers has been extended to include both the sounding out of customers and determining the engineering characteristics of a product that will satisfy consumer preferences [19]. Therefore, it is necessary to convert engineering characteristics into consumer preferences to ensure the design quality of new products. When the correlation between consumer preferences and engineering characteristics is specified, a successful new product is more likely to be developed in the market [3].

Traditionally, the correlations between consumer preferences and engineering characteristics can be illustrated by quality function deployment (QFD) [15, 19] which utilizes a matrix, namely houses of quality (HOQ), to relate customer preferences to engineering characteristics. Target values of engineering characteristics that are housed at the bottom of a HOQ, provide definitive and quantitative specifications for new products. However, nonlinearities usually exist between engineering characteristics and consumer preferences; moreover, an evaluation of the degrees of consumer preferences is fuzzy as it is normally accomplished in a subjective or heuristic manner. Therefore, the development of the HOQ associated with engineering characteristics is a complex decision-making process. Alternatively, the empirical customer preference models have commonly been developed in order to illustrate the relationship between customer preferences and engineering characteristics [3]. These customer preference models are developed using numerical data collected from experiments or consumer surveys which investigate customer preferences with respect to engineering characteristics.

To evaluate customers' perceptual preferences, fuzzy modelling methods including the Takagi-Sugeno model [34], fuzzy inference model [31], fuzzy logic model [32], have been commonly used. However, these methods can generate only implicit customer preference models which cannot provide explicit information. These methods are not widely utilized by engineers, as they reveal no explicit reasons for product design. Analytical information such as significances, sensitivities, and interactions with engineering characteristics cannot be indicated by the models. To overcome the limitation of these

Manuscript received xxxx; accepted xxxx. Date of publication xxxx; date of current version xxxx.

Kit Yan Chan is with the Department of Electrical and Computer Engineering, Curtin University, Australia. Kit Yan Chan is the corresponding author. Phone: 61-8-9266 9269; fax: 61-8-9266 7548; e-mail: kit.chan@curtin.edu.au.

Hak Keung Lam is with the Department of Informatics, King's College London, U.K. (e-mail: hak-keung.lam@kcl.ac.uk).

Tharam S. Dillon is with the Department of Computer Science and Computer Engineering, Latrobe University, Australia (e-mail: tharam.dillon7@gmail.com).

Sai Ho Ling is with the School of Electrical, Mechanical and Mechatronic Systems, University of Technology, Sydney, Australia (e-mail: Steve.Ling@uts.edu.au).

fuzzy modelling methods, multivariate regression methods can be used to generate polynomial models for customer preferences [16, 17, 18, 52]. Using these polynomial models, explicit information can be derived. These methods are based on the assumption that deviations between real observations and model estimates are random errors. In other words, the models can be generated only if the collected observations are distributed randomly to the models; also, the collected observations must be crisp. However, discrepancies between observations and estimates can be caused by vague knowledge or imprecise evaluations of consumer preferences. Hence, errors in the models can be caused by fuzziness, but these models are unable to capture the fuzziness of consumer preferences.

Recently, fuzzy regression has become more commonly used for developing consumer preference models, as the developed models are explicit; analytical information can be indicated for the new product development [26]; and the developed models can address the fuzziness in consumer preferences. Kim et al. [26] have first applied fuzzy linear regression to develop consumer preference models, whereby fuzzy coefficients with symmetric triangular membership are used to address the fuzziness of customer preferences, which is due to cognitive sources including human estimations in evaluating the customer preferences, and imprecise evaluations of consumer preferences. This method has been applied to many product or service designs [24, 45, 46]. Advanced versions of fuzzy linear regression were integrated with asymmetric triangular fuzzy coefficients in order to increase the flexibility when addressing fuzziness of customer preferences [5, 6, 13]. Also, fuzzy linear least-squares regression has been developed in order to address the randomness in customer preference observations [27]. This randomness is generally due to non-cognitive sources including the physical randomness of engineering characteristics, statistical uncertainty caused by limited information in estimating the customer preferences, model uncertainties caused by the assumptions in simplifying the predicative models, variability influenced by product quantities, and instrumentation errors when measuring the consumer preferences [11, 12]. More recently, heuristic methods have been integrated with TS-fuzzy regression [2, 22], whereby the genetic programming [40] and chaos method [44] are used to generate the polynomial structure involved with interaction and high-order terms; based on the polynomial structure, TS-fuzzy regression [48] is used to determine the fuzzy coefficients.

Even if the aforementioned methods are able to address fully or partially the characteristics of customer preferences, they have a common limitation: they do not analyse the significance of each engineering characteristic prior to developing the models. In fact, some engineering characteristics may not be significant to consumers. Using all specified engineering characteristics which include insignificant or uncorrelated ones, is not the most effective means of developing the models. Rawlings [43] also mentioned that the determination of significant regressors is critical when developing an appropriate polynomial model, as the regressors in product development consist of each individual engineering characteristic, interaction between engineering characteristics, and high-order engineering characteristics. The inclusion or exclusion of insignificant regressors or significant regressors may affect the generalization capability of the polynomial models. The

inclusion of insignificant regressors could lead to the modelling of spurious behaviours of customer preferences. The exclusion of significant regressors could lead to the model not recognizing important behaviours relating to customer preferences. Apart from this, an optimal polynomial structure, which includes the significant high-order or interaction regressors, is also closely correlated to the generalization capability of the model [23]. Therefore, it is desirable to develop a methodical process for determining the appropriate structure of the polynomial models which consist only of significant regressors. With an appropriate structure, the polynomial model can illustrate an appropriate correlation between significant engineering characteristics and consumer preferences.

In this paper, a novel modelling method, namely fuzzy stepwise regression (F-SR) which incorporates the mechanisms of fuzzy least square regression [4] and stepwise regression [43], is proposed, in order to develop the customer preference model which consists only of significant regressors. Also, F-SR can be used to generate models which include the significant high-order or interaction regressors. The developed fuzzy regression model attempts to obtain a better generalization capability using a smaller number of regressors. The F-SR uses the mechanism of stepwise regression [23] to identify the significant regressors which has commonly been used for modelling many systems such as power systems [53], transformer design [14], electric machine design [47], vehicle design [29], air-pollution monitors [9] etc. The F-SR first initializes an "empty" polynomial model with no regressor. It adds the significant regressors one by one to the consumer preference model, until no improvement can be obtained by adding another regressor. Finally, a polynomial structure with only the significant regressors is generated. Based on the final polynomial structure, the fuzzy coefficients are determined based on the fuzzy least square regression [4] by means of which the randomness caused by the quantitative evaluations and measures of consumer preferences [39] can be addressed. Hence, the resulting model consists of significant regressors only, but is able to address high-order or interaction terms for engineering characteristics.

Apart from new product development, the F-SR bridges the research gap in the existing fuzzy regression methods [1, 5, 10, 25, 31, 38, 49, 50] which do not address the issue of selecting significant regressors with high-order or interaction terms for modelling. The effectiveness of the F-SR is evaluated based on a case study of a tea maker design and a solder paste dispenser design. The results obtained by the F-SR are compared with those obtained by the commonly-used methods for new product development. Better generalization capabilities with more effective usage of engineering characteristics can be obtained.

The rest of this paper is organized as follows: Section II presents the formulation of the customer preference models correlated with the engineering characteristics. Section III presents the mechanism of the proposed F-SR. Section IV describes how the F-SR is applied to the development of customer preference models for tea makers. In Section V, a further evaluation of the F-SR is conducted by modelling a solder paste dispenser. Finally, a conclusion and discussion of possible future research are given in Section VI.

II. CONSUMER PREFERENCE MODELS

In the planning process of a new product, the following consumer preference model described in (1) is essential in order to represent the functional relationship between consumer preferences and engineering characteristics for a new product [3]:

$$y_i = f_i(x_1, x_2, \dots, x_m), \quad i = 1, 2, \dots, n \quad (1)$$

where y_i is the i -th consumer preference indicating the i -th degree of customer satisfaction to purchase the product; x_j with $j = 1, 2, \dots, m$ is the j -th engineering characteristic which affects one or more of the consumer preferences; n and m are the numbers of consumer preferences and engineering characteristics respectively defined by the product designers in the new product planning; and f_i represents the functional relationship between the m engineering characteristics to the i -th consumer preference, y_i .

The consumer preference model described in (1) can be represented by the following fuzzy polynomial regression in the Kolmogorov-Gabor form [2, 22], in order to address interactions between engineering characteristics and fuzziness on consumer preference.

$$\tilde{y}_i = \tilde{B}_0^i z_0 + \tilde{B}_1^i z_1 + \tilde{B}_2^i z_2 + \dots + \tilde{B}_k^i z_k + \dots + \tilde{B}_{N_R}^i z_{N_R}, \quad (2)$$

where \tilde{y}_i is the fuzzy estimate of the i -th customer preference which is given by the fuzzy number, $\tilde{y}_i = (\tilde{y}_i^C, \tilde{y}_i^R, \tilde{y}_i^L)$, with the center, \tilde{y}_i^C , right spread, \tilde{y}_i^R , and left spread, \tilde{y}_i^L . $(1+N_R)$ is the number of regressors of the consumer preference model represented in (2) and N_R can be represented in (3) based on the Binomial Theorem:

$$N_R = 2^m - 1; \quad (3)$$

the regressors namely z_0, z_1, \dots, z_{N_R} in (2) are aligned with the linear, interaction and the high-order terms of the engineering characteristic, x_j , given in (1). They are given by:

$$z_0 = 1, \quad (4a)$$

$$z_1 = x_1, \quad z_2 = x_2, \dots, \quad z_m = x_m, \quad (4b)$$

$$z_{m+1} = x_1 \cdot x_1, \quad z_{m+2} = x_1 \cdot x_2, \quad z_{m+3} = x_1 \cdot x_3, \dots, \\ z_{m^2} = x_m \cdot x_m, \dots, \quad (4c)$$

$$z_k = x_{p(1)} \cdot x_{p(2)} \cdot \dots \cdot x_{p(d)} \quad \text{with } p(1), p(2), \dots, p(d) \in \{1, 2, \dots, m\}, \\ m^2 + m + 1 \leq k < N_{NR} \quad \text{and } 3 \leq d < m, \dots, \quad (4d)$$

$$\text{and } z_{N_{NR}} = x_1 \cdot x_2 \cdot \dots \cdot x_m; \quad (4e)$$

and the fuzzy coefficients in (2) are given as

$$\tilde{B}_0^i = (b_0^{i,C}, b_0^{i,L}, b_0^{i,R}),$$

$$\tilde{B}_1^i = (b_1^{i,C}, b_1^{i,L}, b_1^{i,R}), \quad \tilde{B}_2^i = (b_2^{i,C}, b_2^{i,L}, b_2^{i,R}), \dots,$$

$$\tilde{B}_m^i = (b_m^{i,C}, b_m^{i,L}, b_m^{i,R}),$$

$$\tilde{B}_{m+1}^i = (b_{m+1}^{i,C}, b_{m+1}^{i,L}, b_{m+1}^{i,R}), \quad \tilde{B}_{m+2}^i = (b_{m+2}^{i,C}, b_{m+2}^{i,L}, b_{m+2}^{i,R}), \dots,$$

$$\tilde{B}_{m+m^2}^i = (b_{m+m^2}^{i,C}, b_{m+m^2}^{i,L}, b_{m+m^2}^{i,R}), \dots,$$

$$\tilde{B}_k^i = (b_k^{i,C}, b_k^{i,L}, b_k^{i,R}) \quad \text{with } (m^2 + m + 1) < k < N_{NR}$$

and $3 \leq d \leq m, \dots,$

$$\text{and } \tilde{B}_{N_R}^i = (b_{N_R}^{i,C}, b_{N_R}^{i,L}, b_{N_R}^{i,R}).$$

However, some regressors in (2) may be redundant. The inclusion of all regressors in (2) is an ineffective means of modelling the customer preferences, as some of the regressors are not significantly correlated to \tilde{y}_i . When both significant and insignificant regressors are included in the customer preference model, the generalization capability of the model may not be better than one that includes only significant regressors; also, an over-fitted model is likely to be produced. Also, the removal of insignificant regressors helps to produce simpler structure of customer preference models which could be interpreted easier by users. In order to represent only the significant regressors, (2) is formulated by (5) as:

$$\tilde{y}_i = \tilde{B}_0^i + \tilde{B}_1^i z_{I^M(1)} + \tilde{B}_2^i z_{I^M(2)} + \tilde{B}_3^i z_{I^M(3)} + \dots + \tilde{B}_{N_{Term}}^i z_{I^M(N_{Term})}, \quad (5)$$

where N_{Term} (i.e.: $N_{Term} \leq N_R$) is the number of significant regressors; and $I^M(j)$ with $j = 1, 2, \dots, N_{Term}$ are the indexes of significant regressors which are significantly correlated to the consumer preference, \tilde{y}_i , and they are given by:

$$I^M(j) \in \{1, 2, \dots, N_R\}, \quad \text{but } I^M(j) \neq I^M(k)$$

$$\text{with } j \neq k \quad \text{and } j, k = 1, 2, \dots, N_{Term}. \quad (6)$$

Section III proposes an approach, namely fuzzy stepwise regression (F-SR), which incorporates the mechanisms of fuzzy least square regression [4] and stepwise regression [43], in order to develop the customer preference model formulated in (5) which contains only significant regressors to the consumer preference. Also, the customer preference model is capable of addressing two common characteristics of consumer preferences, namely (a) interactions and high order terms for engineering characteristics, and (b) uncertainties caused by the quantitative evaluations and measures of consumer preferences.

III. FUZZY STEPWISE REGRESSION FOR CONSUMER PREFERENCE MODELS

The F-SR is proposed as a means of finding the number of significant regressors, N_{Term} , and the indexes of the significant regressors $I^M(j) \in \{1, 2, \dots, N_R\}$ with $j = 1, 2, \dots, N_{Term}$, when the original data set for each engineering characteristic, namely D_o^i , is available to represent the relationship between the i -th customer preference, \tilde{y}_i , and the m engineering characteristics.

D_o^i is represented as:

$$D_o^i = \{ \tilde{y}_i(k), x_1(k), x_2(k), \dots, x_m(k) \mid k = 1, 2, \dots, N_D \} \quad (7)$$

where N_D is the number of observations; $x_j(k)$ is the k -th data with respect to the j -th engineering characteristic; $\tilde{y}_i(k) = (\tilde{y}_i^C(k), \tilde{y}_i^R(k), \tilde{y}_i^L(k))$ is the k -th observation with respect to the i -th customer preference. $\tilde{y}_i^C(k)$, $\tilde{y}_i^R(k)$ and $\tilde{y}_i^L(k)$ are the mean, right spread and left spread of the k -th observation which is determined by repeating some trials with the same setting of engineering characteristics. They can be determined based on (18), (19) and (20).

Before performing the F-SR, the full data set for each regressor, namely D_F^i , is generated based on D_o^i , where D_F^i represents the relationship between the regressors, z_j with $j=1,2,\dots,N_R$, and \tilde{y}_i . D_F^i is given as:

$$D_F^i = \left\{ \left(\tilde{y}_i^C(k), \tilde{y}_i^R(k), \tilde{y}_i^L(k) \right), z_0(k), z_1(k), \dots, z_{N_R}(k) \mid k=1,2,\dots,N_D \right\} \quad (8)$$

where N_R is given by (3):

$$\begin{aligned} z_0(k) &= 1; \\ z_1(k) &= x_1(k), z_2(k) = x_2(k), \dots, z_m(k) = x_m(k); \\ z_{m+1}(k) &= x_1(k) \cdot x_1(k), z_{m+2}(k) = x_1(k) \cdot x_2(k), \dots, \\ z_{m+m^2}(k) &= x_m(k) \cdot x_m(k); \dots, \\ z_j(k) &= x_{p(1)}(k) \cdot x_{p(2)}(k) \cdot \dots \cdot x_{p(d)}(k) \text{ with} \\ p(1), p(2), \dots, p(d) &\in \{1, 2, \dots, m\}, 3 \leq d < m \text{ and} \\ (m^2 + m + 1) &\leq j < N_{NR}, \dots; \text{ and} \\ z_{N_{NR}}(k) &= x_1(k) \cdot x_2(k) \cdot \dots \cdot x_m(k). \end{aligned}$$

Based on D_F^i , the F-SR uses the mechanism of stepwise regression [43] to identify the significant regressors, where the stepwise regression is integrated with the approaches of backward elimination and forward selection. The two approaches attempt to determine whether each of the regressors can be included in the model. The backward elimination starts with the general model which includes all regressors. It eliminates the regressor which is most insignificant to the model, and then it fits the reduced model of which the most insignificant regressor is excluded. The procedure terminates until no more regressor can be removed from the model. However, once a regressor has been removed from the model, it cannot be included again. The effectiveness of backward elimination is restricted by the order of regressors which are removed from the model. The forward selection is a reversed version of the backward elimination. It starts with an 'empty' model with no regressor, and adds regressors one by one until the model cannot be significantly improved by adding another regressor. However, once a regressor has been included in the model, it cannot be removed. The effectiveness of forward selection is also restricted by the order of regressors which are added to the model.

The stepwise regression is an improved version of both approaches. It permits re-examination of the regressors incorporated in the model and also those removed from the model in previous steps. A regressor that entered at an early stage may become superfluous at a later stage, because of its relationship with other regressors which are added in the model. In stepwise regression, this superfluous regressor can be removed. Also, some regressors which have been removed in previous steps can be selected again if they are significant on the current model. The stepwise regression considers more relevant models than those considered by the forward selection and the backward elimination. It increases the chance of obtaining the better model. Therefore, the F-SR is proposed based on the stepwise regression in this research, in order to

select the significant regressors for the customer preference models.

A. Fuzzy stepwise regression algorithm (F-SR)

The F-SR is given in Algorithm 1. In *Step 0.1*, the F-SR first initializes a parameter, namely k , and two sets, namely I_k^M and I_k^R , where k is the iteration number of the F-SR, I_k^M is the index set of the regressors which exist in the model, and I_k^R is the index set of the regressors which have not being picked up. At $k=0$, I_0^M is initialized as an empty set since no regressor index is included, and I_0^R is initialized as $I_0^R = [1, 2, \dots, N_R]$ since no regressor has been picked up. In *Step 0.2 to Step 0.4*, the F-SR determines the first regressor index in I_1^M when the iteration is at $k=1$. Here the strategy of determining the first regressor is based on [23, 43], where the one with the largest F -value is selected as the first regressor. The F-SR generates N_R models relating the customer preference \tilde{y}_i to each of the regressor z_j with $j \in I_0^R$. Based on the N_R models, the F -value of each regressor is computed with respect to each model. The index of the regressor with the largest F -value is entered as the first index for I_1^M . This index it is removed from I_0^R in order to generate I_1^R .

After the first index has been determined, the F-SR performs the procedures of forward selection and backward elimination sequentially, where the forward selection is based on *Step k.1 to Step k.4* and the backward elimination is based on *Step k.5 to Step k.8*. In the forward selection, the F -value of each of the regressors indexed with I_k^R are computed and the regressor with the largest F -value is selected. In order to determine whether this regressor is significant to improve the prediction of the model indexed with I_k^M , the F -test is performed on this regressor associated with the other regressors already included in I_k^M . If the F -test indicates that this regressor is significant to improve the prediction of the model, the index of this regressor is entered in order to generate I_{k+1}^M . Also it is removed in order to generate I_{k+1}^R . Then the iteration number, k , is incremented by 1.

After the forward selection has been performed, the backward elimination is carried out. The elimination process attempts to remove the regressors which become insignificant after the new regressor has been entered by the forward selection. Hence, it removes the insignificant regressors indexed with I_{k-1}^M , where I_{k-1}^M is the regressor index set of the previous model before new regressor has been entered by the forward selection. It first computes the F -value of each regressor indexed with I_{k-1}^M . The regressor with the smallest F -value is selected to perform the F -test associated with the model which consists of the regressors indexed with I_k^M . Hence, this model consists of the regressors which are indexed with I_{k-1}^M and it also consists of the new regressor which has been newly entered by the forward regression. If the F -test indicates that this regressor with the smallest F -value is not significant, this regressor is removed from both I_k^M and I_{k-1}^R .

Then the model is refitted with the remaining regressors indexed with I_k^M . Similarly the elimination process is repeated until all regressors indexed with I_{k-1}^M are significant.

The two processes of forward selection and backward elimination continue until no more regressor can be entered into I_k^M or removed from I_k^R .

Algorithm 1: Stepwise fuzzy regression (F-SR)

Input: D_F^i which is the full data set for all regressors;

Output: I_k^M which is the index set for the significant regressors

// Step 0.1 Initialize the iteration number and the index sets

Begin

Step 0.1: Initialize k , I_k^M and I_k^R by setting $k = 0$, setting I_0^M as an empty set and setting $I_0^R = [1, 2, \dots, N_R]$.

// k is the iteration number, I_k^M is the index set of the regressors which exist in the model, and I_k^R is the index set of the regressors which have not being picked up.

// Step 0.2 to Step 0.4: Determine the 1-st index for the fuzzy regression model

Step 0.2: Generate N_R models,

$$\tilde{y}_i = \tilde{B}_0^i + \tilde{B}_1^i z_j \quad \text{with } j \in I_0^R \quad (9)$$

where the fuzzy coefficients on (9) can be determined based on the fuzzy least square regression discussed in Section III.C.

Step 0.3: Determine the most significant regressor z_j , where

$F(j', I_0^M) > F(j, I_0^M)$ for all $j, j' \in I_0^R$ but $j \neq j'$. $F(j, I_0^M)$ is the F -value with respect to z_j and is computed based on Section III.B.

Step 0.4: Add j' to I_0^M , in order to generate I_1^M (i.e. $I_1^M = [I_0^M, j']$). Remove j' from I_0^R in order to generate I_1^R (i.e.

$$I_1^R = I_0^R \setminus j'). \text{ Set } k=k+1.$$

// Step k.1 to Step k.4: Forward selection

Step k.1: Compute the F -values, $F(j', I_k^M)$, for all $j' \in I_k^R$, based on Section III.A.

//where the number of elements in I_k^R is $|I_k^R|$

Step k.2: Determine the largest $F(j^*, I_k^M)$, where

$$F(j^*, I_k^M) > F(j', I_k^M) \text{ with } j^*, j' \in I_k^R \text{ but } j^* \neq j'.$$

Step k.3: If $F(j^*, I_k^M) > F_{1, N_D - |I_k^M|, 1-\alpha}$, then add j^* into I_k^M and remove j^* from I_k^R , where $F_{1, N_D - |I_k^M|, 1-\alpha}$ is the critical value for the F -test. $|I_k^M|$ is the number of elements in I_k^M .

Step k.4: Set $k = k + 1$.

// Step k.5 to Step k.8: Backward elimination.

Step k.5: Compute the F -values, $F(j', I_k^M \setminus j')$ with $j' \in I_{k-1}^M$, based on Section III.A.

Step k.6: Determine the smallest, $F(j^*, I_k^M \setminus j^*)$, where

$$F(j^*, I_k^M \setminus j^*) < F(j', I_k^M \setminus j') \text{ with } j^*, j' \in I_{k-1}^M \text{ but } j^* \neq j'.$$

Step k.7: If $F(j^*, I_k^M \setminus j^*) < F_{1, N_D - (|I_k^M| - 1), 1-\alpha}$, then remove j^* from I_k^M and I_{k-1}^M , and add j^* into I_k^R , where $F_{1, N_D - (|I_k^M| - 1), 1-\alpha}$ is the critical value for the F -test.

Step k.8: Repeat Step k.5 to Step k.7 until all $F(j', I_k^M \setminus j)$ with all $j' \in I_{k-1}^M$ is larger than $F_{1, N_D - (|I_k^M| - 1), 1-\alpha}$. Otherwise, **Goto Step k.9.**

// Step k.9: Determination whether the process can be terminated.

Step k.9: Compute $F(j', I_k^M)$, for all $j' \in I_k^R$, based on Section III.B.

If all $F(j', I_k^M) < F_{1, N_D - |I_k^M|, 1-\alpha}$,

then return I_k^M as the indexes of the fuzzy regression model as no more regressors can be entered or removed. Hence, the fuzzy regression model can be represented as:

$$\tilde{y}_i = \tilde{B}_0^i + \tilde{B}_1^i z_{I^M(1)} + \tilde{B}_2^i z_{I^M(2)} + \dots + \tilde{B}_{N_{term}}^i z_{I^M(N_{term})}$$

where N_{term} is the number of regressors of the fuzzy regression model and is given as $N_{term} = |I^M|$.

Otherwise, Goto Step k.2 to continue the process.

End

B. Computation of the F-value and the F-test

The F -test can be used to evaluates whether or not the addition of the regressor, z_j , improves the prediction of \tilde{y}_i in the fuzzy regression model (10) given that the regressors, $z_{I(1)}$, $z_{I(2)}$, \dots , and $z_{I(N)}$, indexed with $I = [I(1), I(2), \dots, I(N)]$ are already in the model:

$$\tilde{y}_i = \tilde{B}_0^i + \tilde{B}_1^i \cdot z_{I(1)} + \tilde{B}_2^i \cdot z_{I(2)} + \dots + \tilde{B}_N^i \cdot z_{I(N)} \quad (10)$$

Based on [56], equation (11) computes the F -value, $F(j, I)$, of z_j :

$$F(j, I) = \frac{RSS([I, j]) - RSS(I)}{\frac{RSS([I, j])}{N_D - (N + 1)}} \quad (11)$$

where $F(j, I)$ concerns z_j given that $z_{I(1)}$, $z_{I(2)}$, \dots , and $z_{I(N)}$ are already in the model; $RSS(I)$ is the residual sum of square due to regression in model (10) of which $z_{I(1)}$, $z_{I(2)}$, \dots , and $z_{I(N)}$ are included and z_j is not included in the model; N_D is the number of observations given in (7). The fuzzy coefficients in (10) are estimated based on the fuzzy least square regression discussed in Section III.C. $RSS(I)$ is estimated by

$$RSS(I) = \sum_{k=1}^{N_D} (\tilde{y}_i(k) - \hat{\tilde{y}}_i(k))^2 \quad (12)$$

where $\hat{\tilde{y}}_i(k)$ is the estimate generated by the model (10) and $\tilde{y}_i(k)$ is the actual data.

$RSS(I)$ is the corresponding residual sum of square when z_j is added in the model (10) and the fuzzy coefficients are updated. As the model estimate and the actual data are involved with the center, the left spread, and the right spread, it is not sufficient to use only the center value to be the estimate for the customer preference, as the right and left spreads are also important in describing the characteristics of the model estimates and the actual data. Therefore, it is necessary to map

the three values including the center, the left spread, and the right spread into a real number. Here, the weighted fuzzy arithmetic given in [41] was adopted to derive the crisp values of the corresponding model estimates, where the crisp value, \tilde{A}^* , of a number involved with center, left spread and right spread, $\tilde{A} = (a^c, a^L, a^R)$, is determined by (13). This crisp value estimation has also been used for fuzzy regression defuzzification [4,54].

$$\tilde{A}^* = a^c + \frac{1}{6} \cdot (a^R - a^L) \quad (13)$$

In the F -test, the null hypothesis is defined as:

H₀: the fuzzy coefficient of z_j is zero given that

$z_{I(1)}, z_{I(2)}, \dots$, and $z_{I(N)}$ and z_j are in the model,

where we assumed that the sample size, N_D , is large enough.

If $F(j, I)$ in (11) is less than $F_{1, N_D - N, 1 - \alpha}$, **H₀** cannot be rejected, where $F_{1, N_D - N, 1 - \alpha}$ is the critical value for the F -test and α is the significance level. Hence, z_j is not significant in improving the prediction of \tilde{y}_i to model (10). Including z_j may lead to finding insignificant effects for relevant regressors. Hence, z_j is not required to be included in (10), as z_j is no help in predicting \tilde{y}_i . The elimination of z_j enables one to reduce the set of possible regressors to an economical set of significant regressors. If $F(j, I)$ in (11) is larger than $F_{1, N_D - N, 1 - \alpha}$, **H₀** can be rejected, since z_j significantly improves the prediction of \tilde{y}_i to the model (10). Hence, z_j can be added to (10).

This F -test implies that the probability of type 1 error (i.e. probability of incorrectly rejecting a correct hypothesis) is lower than the significant level, α of each test. Therefore, in order to reduce the probability in generating a high type 1 error, a low α of which $\alpha \in [0.01\%..5\%]$ is usually used.

C. Fuzzy least square regression

When the indexes of the N regressors, $I(j)$ with $j=1, 2, \dots, N$, are available, the fuzzy least square regression can be used to determine the fuzzy coefficients, $\tilde{B}_{I(j)}^i = (b_{I(j)}^c, b_{I(j)}^L, b_{I(j)}^R)$ with $I(j)$. The fuzzy coefficients are determined by minimizing the total sum of residual errors, $Error$, given in (14) based on the fuzzy least square regression [4].

$$\begin{aligned} Error &= \sum_{k=1}^{N_D} (\hat{y}_i(k) - \tilde{y}_i(k))^2 \\ &= \sum_{k=1}^{N_D} ((\hat{y}_i^c(k), \hat{y}_i^L(k), \hat{y}_i^R(k)) - (\tilde{y}_i^c(k), \tilde{y}_i^L(k), \tilde{y}_i^R(k)))^2 \end{aligned} \quad (14)$$

where $Error$ evaluates the sum of the differences between the actual data, $\tilde{y}_i(k) = (\tilde{y}_i^c(k), \tilde{y}_i^L(k), \tilde{y}_i^R(k))$ with $k=1, 2, \dots, N_D$, and the estimate, $\hat{y}_i(k) = (\hat{y}_i^c(k), \hat{y}_i^L(k), \hat{y}_i^R(k))$ with $k=1, 2, \dots, N_D$ which are generated by the consumer preference model.

By substituting $\hat{y}_i(k)$ with the model outcome,

$\sum_{j=1}^N (b_{I(j)}^{i,C}, b_{I(j)}^{i,L}, b_{I(j)}^{i,R}) \cdot z_{I(j)}(k)$, (14) can be rewritten as:

$$Error = \sum_{k=1}^{N_D} \left(\sum_{j=1}^N (b_{I(j)}^{i,C}, b_{I(j)}^{i,L}, b_{I(j)}^{i,R}) \cdot z_{I(j)}(k) - (\tilde{y}_i^c(k), \tilde{y}_i^L(k), \tilde{y}_i^R(k)) \right)^2 \quad (15)$$

(15) can be elaborated as (16), based on the weight fuzzy arithmetic operations for asymmetric triangular membership functions [46].

$$\begin{aligned} Error &= \frac{1}{3} \sum_{k=1}^{N_D} \left(\sum_{j=1}^N b_{I(j)}^{i,C} z_{I(j)}(k) - \tilde{y}_i^C(k) \right) \left[\left(\sum_{j=1}^N b_{I(j)}^{i,R} z_{I(j)}(k) - \tilde{y}_i^R(k) \right) - \right. \\ &\quad \left. \left(\sum_{j=1}^N b_{I(j)}^{i,L} z_{I(j)}(k) - \tilde{y}_i^L(k) \right) \right] + \frac{1}{12} \sum_{k=1}^{N_D} \left[\left(\sum_{j=1}^N b_{I(j)}^{i,R} z_{I(j)}(k) - \tilde{y}_i^R(k) \right)^2 + \right. \\ &\quad \left. \left(\sum_{j=1}^N b_{I(j)}^{i,L} z_{I(j)}(k) - \tilde{y}_i^L(k) \right)^2 \right] + \sum_{k=1}^{N_D} \left(\sum_{j=1}^N b_{I(j)}^{i,C} z_{I(j)}(k) - \tilde{y}_i^C(k) \right)^2 \end{aligned} \quad (16)$$

Minimization of $Error$ can be performed, by deriving (16) with respect to each element of $b_{I(j)}^{i,C}$, $b_{I(j)}^{i,R}$, and $b_{I(j)}^{i,L}$, with $j=1, 2, \dots, N$ and then solving the derivatives which are set to zeros. By defining the three elements $b_{I(j)}^{i,*}$ with $i^* \in [C, R, L]$, the derivatives of (16) with respect to $b_{I(j)}^{i,*}$ are given as:

$$\begin{aligned} \frac{\partial(Error)}{\partial b_{I(1)}^{i,*}} &= \left(\sum_{k=1}^{N_D} (z_{I(1)}(k))^2 \right) \cdot b_{I(1)}^{i,*} + \left(\sum_{k=1}^{N_D} z_{I(1)}(k) \cdot z_{I(2)}(k) \right) \cdot b_{I(2)}^{i,*} + \\ &\quad \left(\sum_{k=1}^{N_D} z_{I(1)}(k) \cdot z_{I(3)}(k) \right) \cdot b_{I(3)}^{i,*} + \dots + \left(\sum_{k=1}^{N_D} z_{I(1)}(k) \cdot z_{I(N)}(k) \right) \cdot b_{I(N)}^{i,*} \\ &\quad - \sum_{k=1}^{N_D} (z_{I(1)}(k) \cdot \tilde{y}_i^*(k)) = 0, \text{ with } j=1; \end{aligned} \quad (17a)$$

$$\begin{aligned} \frac{\partial(Error)}{\partial b_{I(2)}^{i,*}} &= \left(\sum_{k=1}^{N_D} z_{I(1)}(k) \cdot z_{I(2)}(k) \right) b_{I(1)}^{i,*} + \left(\sum_{k=1}^{N_D} (z_{I(2)}(k))^2 \right) \cdot b_{I(2)}^{i,*} + \\ &\quad \left(\sum_{k=1}^{N_D} z_{I(2)}(k) \cdot z_{I(3)}(k) \right) \cdot b_{I(3)}^{i,*} + \dots + \left(\sum_{k=1}^{N_D} z_{I(2)}(k) \cdot z_{I(N)}(k) \right) \cdot b_{I(N)}^{i,*} - \\ &\quad \sum_{k=1}^{N_D} (z_{I(2)}(k) \cdot \tilde{y}_i^*(k)) = 0, \text{ with } j=2; \end{aligned} \quad (17b)$$

$$\begin{aligned} & \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \frac{\partial(Error)}{\partial b_{I(N)}^{i,*}} &= \left(\sum_{k=1}^{N_D} z_{I(1)}(k) \cdot z_{I(N)}(k) \right) b_{I(1)}^{i,*} + \left(\sum_{k=1}^{N_D} z_{I(2)}(k) \cdot z_{I(N)}(k) \right) \cdot b_{I(2)}^{i,*} + \\ &\quad \left(\sum_{k=1}^{N_D} z_{I(3)}(k) \cdot z_{I(N)}(k) \right) \cdot b_{I(3)}^{i,*} + \dots + \left(\sum_{k=1}^{N_D} (z_{I(N)}(k))^2 \right) \cdot b_{I(N)}^{i,*} - \\ &\quad \sum_{k=1}^{N_D} (z_{I(N)}(k) \cdot \tilde{y}_i^*(k)) = 0, \text{ with } j=N; \end{aligned} \quad (17c)$$

where $\tilde{y}_i^*(k)$ with $'*'\in['C','R','L']$ represents the center $\tilde{y}_i^C(k)$, right spread, $\tilde{y}_i^R(k)$, and left spread $\tilde{y}_i^L(k)$ respectively.

In (17a), (17b), and (17c), the unknowns are all $b_{l(j)}^{i,*}$ with $j=1,2,\dots,N$, and those involved with $\tilde{y}_i^*(k)$ and $z_{l(j)}(k)$ are all constant values. Therefore, the equation set represented in (17) is identical to the formulation of the least square regression with N unknowns and N equations. Using the least square regression, the centers of fuzzy coefficients in (17), $b_{l(j)}^{i,*}$ with $'*'\in['C']$ and $j=1,2,\dots,N$, can be obtained based on the centers of the collected data, $\tilde{y}_i^*(k)$ with $'*'\in['C']$ and $k=1,2,\dots,N_D$. Similarly, the left spreads, $b_{l(j)}^{i,*}$ with $'*'\in['L']$, and the right spreads, $b_{l(j)}^{i,*}$ with $'*'\in['R']$, can be obtained respectively, based on left spreads of the collected data, $\tilde{y}_i^*(k)$ with $'*'\in['L']$, and right spreads of the collected data, $\tilde{y}_i^*(k)$ with $'*'\in['R']$. Hence, the fuzzy coefficients, $(b_{l(j)}^{i,C}, b_{l(j)}^{i,L}, b_{l(j)}^{i,R})$ with $j=1,2,\dots,N$, can be determined.

IV. EVALUATION OF THE FUZZY STEPWISE REGRESSION

The effectiveness of the F-SR is evaluated using the case study of a tea maker design. The results obtained by the F-SR are compared with those obtained by the commonly-used fuzzy regression methods for developing customer preference models.

A. A Case study of tea maker design

In the tea maker design, two customer preferences, namely catechin content, y_1 , and tea concentration, y_2 , are essential in order to satisfy tea drinkers. y_1 is a type of antioxidant found in great abundance in the leaves of the tea plant. Its health benefits have been under close examination, due to tea consumption being associated with health and longevity in many ancient cultures. y_2 elicits three affective streams from tea drinkers, namely the rating of tea in terms of aroma, texture and overall taste.

To brew a tea, the following five steps are involved in the tea makers; the schematic diagram of conceptual tea makers is given in Figure 1 in the appendix. In the five steps, five engineering characteristics, namely x_1, x_2, x_3, x_4 and x_5 which are correlated to y_1 , and y_2 , are identified. They are discussed in Step 1 to Step 5.

Step 1 Water boiling: 2.5 litres of water are poured into Container 2, and are boiled to 98 degrees Celsius.

Step 2 Tea loading and water reheating: Tea with 70 grams is poured into the tea infuser and is then placed into Container 1. The original water temperature decreases as heat is lost by immersing with the tea infuser. The water is reheated, where the reheating temperature is the first engineering characteristic, x_1 .

Step 3 First brewing: The tea chemical contents are released by dropping the tea infuser into the reheat water for certain times. In each drop, the tea infuser is immersed in the water for 10 seconds and then 10 seconds elapse before the next drop. The number of drops is the second engineering characteristic, x_2 .

Step 4 Tea dipping: To release more chemical contents, the tea is immersed in the water through dipping. The immersing time is the third engineering characteristic, x_3 .

Step 5 Second brewing: The second brewing attempts to release more chemical contents into the water. Same as the first brewing, the tea infuser is immersed with certain drops. In each drop, the tea infuser is immersed with certain time and with 10 seconds elapsing before the next drop. The number of drops and the immersion time are the fourth engineering characteristic, x_4 , and the fifth engineering characteristic, x_5 , respectively.

Therefore, the 5 engineering characteristics which are significant for customer preferences are given as: reheating temperature (x_1), number of drops for the first brewing (x_2), dipping time (x_3), number of drops for the second brewing (x_4), and immersion time in the second brewing (x_5). The above five steps are detailed in [57].

B. Experimental observations for tea maker design

One major task when developing a new product is to study the effects of the setting of the engineering characteristics on the two customer preferences of the tea makers, y_1 , and y_2 . Experiments were conducted by setting the five engineering characteristics, x_1, x_2, x_3, x_4 , and x_5 , with different values in order to study, y_1 , and y_2 . In design of experiments (DOE), the number of levels of the factors in design of experiments (DOE) can affect efficiency of the models. When a linear model is developed, two levels are used since two points define a line. When a model involved with quadratic regressors is developed, three levels are used since the two extremes and the midpoint can illustrate the quadratic. For a cubic model, four levels are used since two extremes and two points in between can illustrate the cubic. Extra levels allow the examination of complicated nonlinear functions, but more experimental time is required. Here, the ranges of the engineering characteristics are given and are quantized into four levels as illustrated in Table I.

Table 1: Experimental ranges and settings of the five engineering characteristics for the tea maker design

Engineering characteristics	Reheating temperature (degrees Celsius) x_1	Number of drops in the first brewing cycle x_2	Dipping time (minutes) x_3	Number of drops in the second brewing cycle x_4	Immersion time in the second brewing cycle (seconds) x_5
Experimental ranges	93-99	1-4	8.5-10	2-5	10-40
Level 1	93	1	8.5	2	10
Level 2	95	2	9	3	20
Level 3	97	3	9.5	4	30
Level 4	99	4	10	5	40

Table 2: The orthogonal array, $L_{16}(4^5)$, used for the tea maker design

Experiments	x_1	x_2	x_3	x_4	x_5
1	1	1	1	1	1
2	1	2	2	2	2
3	1	3	3	3	3
4	1	4	4	4	4
5	2	1	2	3	4
6	2	2	1	4	3
7	2	3	4	1	2
8	2	4	3	2	1
9	3	1	3	4	2
10	3	2	4	3	1
11	3	3	1	2	4
12	3	4	2	1	3
13	4	1	4	2	3
14	4	2	3	1	4
15	4	3	2	4	1
16	4	4	1	3	2

As there are five engineering characteristics and each of them is quantized by four levels, 1024 (or 4^5) experiments need to be conducted when a full factorial design is used. If 3 hours are required for setting up and performing each experiment involved with tea brewing and all measures, 3072 hours (or 384 days if people works 8 hours per day) are required for the full factorial design which is too time-consuming. Therefore, the orthogonal array, namely $L_{16}(4^5)$ illustrated in Table 2, was used for the experimental design to study the effects of the five engineering characteristics with four levels. Based on $L_{16}(4^5)$ involved with 4 levels, cubic models can be developed. The 16 configurations of the experimental trials are shown in Table 2. As an example, the first experiment is conducted based on the 5 engineering characteristics with a level 1 setting. The fifth experiment is conducted based on the setting of x_1 with level 1, x_2 with level 2, x_3 with level 2, x_4 with level 3, and x_5 with level 4. As only 16 experiments are required to study the main effects of the five engineering characteristics, only 6 days are required for designing the tea maker. Also, 1008 (or 1024-16) experiments can be saved compared with the full factorial design approach.

In order to study the randomness of the experimental observations for the two customer preferences, the experiments configured with $L_{16}(4^5)$ were repeated twice; the experimental data for catechin content, and tea concentration are shown in Table 3 and Table 4 respectively, where $y_{i,j}(k)$ is the experimental data of the j -th trial with respect to the k -th configuration of $L_{16}(4^5)$, and the experimental data for catechin content and tea concentration are represented as $y_{1,j}(k)$ and $y_{2,j}(k)$ respectively. To measure catechin contents, a method based on colour development in the vanillin-hydrochloric acid (V-HCl) was used. In the method, two reactions are involved with colour development from the catechin contents. First, the catechin compound reacts with the vanillin content in an acidic condition. Then, HCl is used to catalyse the reaction to produce more vanillin adducts, and the vanillin adduct is generated to react with cyclohexanol to develop red colour precipitant. Generally, the reaction takes three hours. A spectrometer with the UV light of 500nm wavelength is used to measure the red colour precipitant. Based on the amount of red precipitant, catechin contents in tea can be indicated. To measure the tea

concentration, we used an automatic digital refractometer which is suitable for measuring water soluble samples such as green tea, black tea, oolong tea and low sugar beverages. Each experiment is required about three hours.

Table 3: Experimental data and observations for the catechin content

The k -th experiment	Experimental data		Observations		
	First trial $(y_{1,1}(k))$	Second trial $(y_{1,2}(k))$	Mean $(\tilde{y}_1^C(k))$	Left spread $(\tilde{y}_1^L(k))$	Right spread $(\tilde{y}_1^R(k))$
1	1.530	1.683	1.6065	0.0765	0.0765
2	1.607	1.735	1.671	0.064	0.064
3	1.481	1.563	1.522	0.041	0.041
4	1.628	1.503	1.5655	0.0625	0.0625
5	1.362	1.57	1.466	0.104	0.104
6	1.566	1.623	1.5945	0.0285	0.0285
7	1.542	1.829	1.6855	0.1435	0.1435
8	1.497	1.666	1.5815	0.0845	0.0845
9	1.766	1.608	1.687	0.079	0.079
10	1.602	1.686	1.644	0.042	0.042
11	1.680	1.796	1.738	0.058	0.058
12	1.660	1.793	1.7265	0.0665	0.0665
13	1.709	1.664	1.6865	0.0225	0.0225
14	1.707	1.934	1.8205	0.1135	0.1135
15	1.443	1.717	1.58	0.137	0.137
16	1.844	1.922	1.883	0.039	0.039

Table 4: Experimental data and observations for the tea concentration results

The k -th experiment	Experimental data		Observations		
	First trial $(y_{1,1}(k))$	Second trial $(y_{1,2}(k))$	Mean $(\tilde{y}_1^C(k))$	Left spread $(\tilde{y}_1^L(k))$	Right spread $(\tilde{y}_1^R(k))$
1	1.545	1.566	1.5555	0.0105	0.0105
2	1.676	1.77	1.723	0.047	0.047
3	1.886	1.443	1.6645	0.2215	0.2215
4	1.85	1.587	1.7185	0.1315	0.1315
5	1.792	1.77	1.781	0.011	0.011
6	1.898	1.58	1.739	0.159	0.159
7	1.887	1.779	1.833	0.054	0.054
8	1.857	1.812	1.8345	0.0225	0.0225
9	1.934	1.767	1.8505	0.0835	0.0835
10	2.047	1.647	1.847	0.2000	0.2000
11	1.916	1.777	1.8465	0.0695	0.0695
12	1.963	1.866	1.9145	0.0485	0.0485
13	2.062	2.099	2.0805	0.0185	0.0185
14	1.778	2.103	1.9405	0.1625	0.1625
15	1.812	1.907	1.8595	0.0475	0.0475
16	2.009	2.005	2.007	0.002	0.002

The k -th observation, $(\tilde{y}_i^C(k), \tilde{y}_i^L(k), \tilde{y}_i^R(k))$, attempts to address the randomness for the k -th experimental condition, where $(\tilde{y}_i^C(k), \tilde{y}_i^L(k), \tilde{y}_i^R(k))$ are generated based on (18), (19) and (20) and $(\tilde{y}_i^C(k), \tilde{y}_i^L(k), \tilde{y}_i^R(k))$ with $i=1,2$ represents the observations for catechin content and tea concentration respectively. Here the number of trials, N_i , is set to be 2, since two replicated trials were conducted for each experimental condition. Two trials were collected for this design problem, as the experimental time is lengthy and three hours are required for each experiment. In

$(\tilde{y}_i^C(k), \tilde{y}_i^L(k), \tilde{y}_i^R(k))$, $\tilde{y}_i^C(k)$ addresses the mean for the k -th experimental condition. $\tilde{y}_i^R(k)$ and $\tilde{y}_i^L(k)$ are the right and left spreads for the k -th experimental condition respectively.

$$\tilde{y}_i^C(k) = \frac{1}{N_i} \sum_{j=1}^{N_i} y_{i,j}(k), \quad k = 1, 2, \dots, 16; \quad (18)$$

$$\tilde{y}_i^L(k) = \tilde{y}_i^C(k) - \min_{j=1,2,\dots,N_i} (y_{i,j}(k)), \quad k = 1, 2, \dots, 16; \quad (19)$$

$$\text{and } \tilde{y}_i^R(k) = \max_{j=1,2,\dots,N_i} (y_{i,j}(k)) - \tilde{y}_i^C(k), \quad k = 1, 2, \dots, 16. \quad (20)$$

C. Development of customer preference models

The customer preference models regarding catechin content and tea concentration can be evaluated by investigating the mean absolute errors, which indicate the difference between the real observations and the estimates of the customer preference models. The mean absolute errors are defined by e_{MAE}^i , with $i=1, 2$, where e_{MAE}^1 , and e_{MAE}^2 represent the errors for catechin content, and tea concentration:

$$e_{MAE}^i = \frac{1}{P} \sum_{k=1}^P \left| \frac{\tilde{y}_i^C(I(k)) - \tilde{f}_i^*(\bar{x}(I(k)))}{\tilde{y}_i^C(I(k))} \right| \times 100\% \quad (21)$$

with

$$\bar{x}(I(k)) = [x_1(I(k)), x_2(I(k)), x_3(I(k)), x_4(I(k)), x_5(I(k))];$$

P is the number of experiments which have been conducted; $x_1(I(k))$, $x_2(I(k))$, $x_3(I(k))$, $x_4(I(k))$ and $x_5(I(k))$ are

parameter values for the $I(k)$ -th experimental configuration of the orthogonal array, $L_{16}(4^5)$, where $I(k) \in [1, 2, \dots, N_D]$,

but $I(k) \neq I(j)$ with $k \neq j$ and $j, k = 1, 2, \dots, N_D$. Here $N_D = 16$, as 16 experiments have been conducted;

$\tilde{f}_i^*(\bar{x}(I(k)))$ and $\tilde{y}_i^C(I(k))$ are the $I(k)$ -th model estimate and the mean of the $I(k)$ -th real observation respectively.

Determination of the crisp value of \tilde{f}_i^* can be referred to (13).

Apart from using the F-SR, four effective fuzzy regression approaches for generating customer preference models were employed for comparison purposes:

a) **Tanaka and Watada's fuzzy regression (TS-FR)** [48]

can generate fuzzy linear polynomial models based on a small amount or uncompleted set of experimental observations. The approaches of T-FR have commonly been used to develop the customer preference models [26, 6, 7, 24, 13, 20, 45, 46].

b) **Peters' fuzzy regression (Peters-FR)** [35] is an extension

of T-FR. Based on P-FR, the estimated interval on the generated fuzzy linear model is influenced by all experimental observations and the generated fuzzy linear model is robust in the presence of outliers. P-FR has been used to develop customer preference models for affective design of mobile phones [2].

c) **Hybrid fuzzy least square regression (H-FLSR)** [4] can

be used to generate fuzzy linear polynomial models of which the uncertainties caused by fuzzy and random natures of the experimental observations can be addressed.

H-FLSR has been used to develop customer preference models based on experimental observations which are fuzzy and crisp [27,57].

d) **Genetic programming-based fuzzy regression (GP-FR)**

[2] can generate explicit models in fuzzy polynomial forms involved with high order and interaction terms. In the GP-FR, the structures of the fuzzy polynomials are generated by evolutionary operations, and the fuzzy coefficients of the polynomials are determined by the fuzzy regression method. The GP-FR has been used to generate fuzzy regression models [2]. The parameters used in the GP-FR are identical to the ones used in [2]: population size = 100; probability of crossover = 0.5; probability of mutation = 0.5; and number of generations = 500.

The five fuzzy regression methods, F-SR, TS-FR, Peters-FR, H-FLSR, and GP-FR, were implemented using Matlab for this tea maker design. Using the 16 pieces of experimental data shown in Tables 3 and 4, the customer preference models for catechin content and tea concentration are developed based on the proposed F-SR as (22) and (23) respectively, where the significance level with $\alpha = 0.05$ is used as it is commonly used for the hypothesis tests. The customer preference model for catechin content is developed as:

$$\begin{aligned} \tilde{y}_1 = & (1.627, 0.009, 0.009) + (0.180, 0.003, 0.003) \cdot x_1 \cdot x_5 \\ & - (0.131, 0.002, 0.002) \cdot x_4 \cdot x_5 \end{aligned} \quad (22)$$

where the training error involved with the 16 pieces of experimental data with respect to e_{MAE}^1 was found to be 2.729%. The customer preference model for tea concentration is developed as:

$$\begin{aligned} \tilde{y}_2 = & (1.638, 0.051, 0.051) + (0.320, 0.029, 0.029) \cdot x_1 \\ & + (0.088, 0.005, 0.005) \cdot x_3 - (0.099, 0.002, 0.002) \cdot x_1 \cdot x_3 \end{aligned} \quad (23)$$

where the training error was found to be 2.051%.

The interaction regressors in (22) and (23) can be explained by a commonly used approach namely the interaction plots [36] which give a clear picture of the magnitudes of interactions between engineering characteristics. If the lines on the interaction plot are parallel, no interaction exists between engineering characteristics. If the lines cross, strong interaction occurs between engineering characteristics. Figures 2(a) and 2(b) show the interaction plot between x_1 and x_3 and that between x_4 and x_5 with respect to the catechin content respectively. As the line crosses exist, strong interaction occurs between x_1 and x_5 , and also occurs between x_4 and x_5 . Hence, it explains why the catechin content model in (22) consists of the regressor interacted with x_1 and x_5 and that interacted with x_4 and x_5 . Figure 2(c) shows the interaction plots between x_1 and x_3 with respect to the tea concentration. It shows that strong interaction occurs, as the line crosses exist. Hence, it explains why the regressor interacted with x_1 and x_3 exists for the tea concentration model in (23). The analysed interaction further validates the sustainability of the models developed by the proposed F-SR.

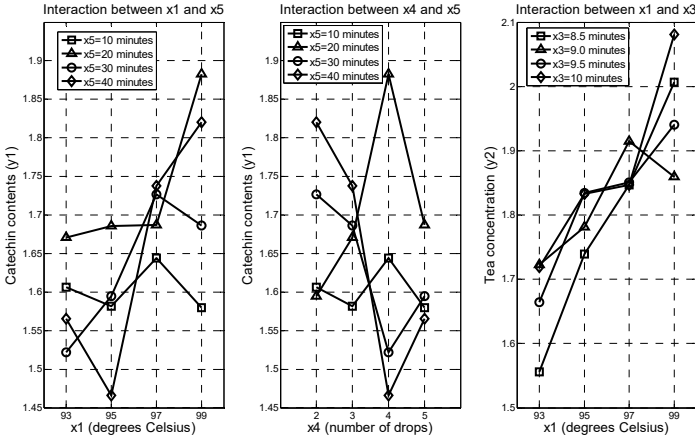


Figure 2 (a) Interaction plot between x_1 and x_5 with respect to catechin content y_1
Figure 2 (b) Interaction plot between x_4 and x_5 with respect to catechin content y_1
Figure 2 (c) Interaction plot between x_1 and x_3 with respect to tea concentration y_2

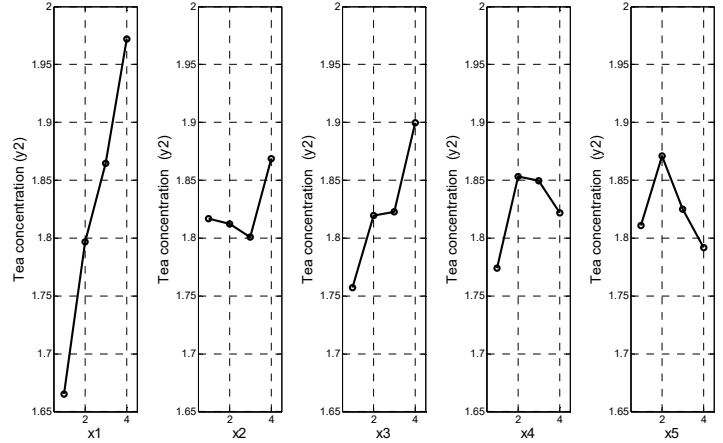


Figure 3: Main effects with respect to Tea concentration, y_2

The linear regressors, x_1 and x_3 , in (23) can be explained by the main effects of the engineering characteristics with respect to the tea concentration, y_2 . As the experiments were conducted based on the orthogonal array shown in Table 2, the combinations of engineering characteristics of each experimental trial are also orthogonal. Hence, the main effect of each engineering characteristic can be separated out [36]. The main effects of the five engineering characteristics at each of the four levels were computed and the main effect plot with respect to y_2 , are shown graphically in Figure 3. The main effect for an engineering characteristic is computed by averaging the experimental results obtained for the engineering characteristic at a given level [36, 55]. As an example, Table 2 shows that x_1 is at level two in the 2nd, 6th, 10th and 14th experiments. Hence, the main effects of x_1 at level two for y_2 are 1.81, which are obtained by averaging the 2nd, 6th, 10th and 14th experimental results for y_2 . This procedure is repeated until all main effects for all engineering characteristics with different levels are computed.

The main effect graphs can provide insight at a glance, and also the main effects can be used to determine the sensitivities of the engineering characteristics with respect to the customer preferences, where the sensitivity of each engineering characteristic is determined by the difference between the largest and smallest main effects for a given engineering characteristic [36, 55]. Table 5 shows the sensitivities of the engineering characteristics with respect to y_2 . Table 5 illustrates that the engineering characteristics, x_1 and x_3 , show the greatest sensitivity to y_2 when they are compared with those of x_2 , x_4 and x_5 . Hence, it explains why x_1 and x_3 exist in the customer preference model for y_2 in (23). These analysed sensitivities validate the sustainability of the customer preference model developed by the proposed F-SR.

Table 5: Sensitivities to tea concentration, y_2

Engineering characteristics	x_1	x_2	x_3	x_4	x_5
Sensitivities to Tea concentration, y_2	0.31	0.07	0.14	0.08	0.08

Apart from the results obtained by the F-SR, Table 6 illustrates the customer preference models developed by the other tested methods, TS-FR, Peters-FR, H-FLSR and GP-FR. The training errors obtained by the customer preference models and the number of engineering characteristics included in the models are also illustrated. For all the customer preference models, the training errors obtained by the five methods are smaller than 5%. Five engineering characteristics are included in the models developed by TS-FR, Peters-FR and H-FLSR, as all five are required in developing the models. The number of engineering characteristics used by GP-FR is smaller than those used by TS-FR, Peters-FR and H-FLSR. The engineering characteristics used by S-FR are the smallest in number, where only 3 engineering characteristics are used for catechin content; and only 2 engineering characteristics are used for tea concentration. Therefore, the number of engineering characteristics used by S-FR is less than those used by the other methods.

When a smaller number of engineering characteristics is included in the models, a simpler product design can be produced. For example, when we control the catechin content of the tea maker, we only need to adjust the parameters, x_1 , x_3 , and x_5 for the model developed by the proposed F-SR. As more parameters are involved in the models developed by the other four tested methods, a more complex design is required in order to control the catechin content and it is necessary to adjust the additional parameters to determine the desired catechin content. Models with a simpler structure can be interpreted more easily by users. When the models are used to determine the optimal setting of engineering characteristics, the complexity of the optimization problem can be reduced. Also, based on the mechanism of the proposed F-SR, the regressors included in the models are significantly related to the customer preferences.

To further validate the generalization capability of the customer preference models developed by the five methods, leave-one-out cross-validation (namely LOOCV) was conducted for the 16 experimental observations. It uses a single observation from the whole data set as the validation data, and the remaining observations as the training data. This is repeated so that all observations are used once as the validation data. Also, the generalization capability of the F-SR is evaluated by using different significance levels. Here, we use $\alpha = 0.02$, $\alpha = 0.05$, $\alpha = 0.10$, and $\alpha = 0.15$, as these significance levels can be considered to be a low value. For modelling the catechin content, the LOOCV results are shown in Table 7 which

summarizes the generalization errors of the five methods, and the mean generalization error of each method. Also, the rank of each method is shown in the table. The results show that the generalization errors obtained by the four F-SR with different significance levels are generally smaller than those of the other four tested methods, TS-FR, Peters-FR, H-FLSR and GP-FR.

Similar LOOCV results are demonstrated in Table 8 which shows the generalization errors obtained by the tested methods for tea concentration. It shows that the ranks of the F-SR for tea concentration are higher than the other four tested methods, TS-FR, Peters-FR, H-FLSR and GP-FR. The generalization errors obtained by the F-SR are generally better than the other four methods; hence, better generalization capability can be obtained by the F-SR in modelling the customer preferences. Better generalization can be explained by the mechanisms of

the F-SR whereby only significant regressors are included in the models, while the other four tested methods cannot guarantee that insignificant regressors are excluded and that the significant ones have not been excluded.

The LOOCV results show that F-SR obtained the best generalization results for the untrained data, although Table 6 shows that GP-FR obtained smaller training error in tea concentration than the F-SR. The better generalization can be explained by the mechanisms of the two methods, in that higher order and interaction regressors are included in the models developed by the GP-FR and they may overtrain with the experimental observations used for developing models. Hence, the generalization capabilities obtained by GP-FR are poorer than those obtained by the F-SR which can generate models which include only the significant regressors.

Table 6: Customer preference models developed by F-FR, TS-FR, Peters-FR, H-FLSR and GP-FR

Customer preference	Modeling methods	Developed models	Training errors	No. engineering characteristics
Catechin content	F-SR ($\alpha = 0.05$)	$\tilde{y}_1 = (1.627, 0.009, 0.009) + (0.180, 0.003, 0.003) \cdot x_1 \cdot x_5 - (0.131, 0.002, 0.002) \cdot x_4 \cdot x_5$	2.729%	3
	TS-FR	$\tilde{y}_1 = (1.621, 0.180) + (0.2410, 2.274 \times 10^{-13}) \cdot x_1 + (0.045, -8.526 \times 10^{-14}) \cdot x_2 + (-0.006, 4.548 \times 10^{-13}) \cdot x_3 + (-0.201, 0) \cdot x_4 + (0.009, 3.979 \times 10^{-13}) \cdot x_5$	4.384%	5
	Peters-FR	$\tilde{y}_1 = (1.596, 0.045) + (0.176, 0.092) \cdot x_1 + (0.059, 0.092) \cdot x_2 + (-0.036, 0.092) \cdot x_3 + (-0.099, 0.092) \cdot x_4 + (0.015, 0.092) \cdot x_5$	3.359%	5
	H-FLSR	$\tilde{y}_1 = (1.609, 0.009, 0.009) + (0.171, 2.859 \times 10^{-13}, 2.842 \times 10^{-13}) \cdot x_1 + (0.054, 0.026, 0.026) \cdot x_2 + (-0.042, 0.001, 0.001) \cdot x_3 + (-0.105, 0.005, 0.005) \cdot x_4 + (0.010, 0.000, 0.000) \cdot x_5$	3.225%	5
	GP-FR	$\tilde{y}_1 = (-0.1783, 0.913 \times 10^{-12}) \cdot x_3 \cdot x_1^3 + (-0.125, 0.127 \times 10^{-13}) \cdot x_4 + (0.258, 0.906 \times 10^{-12}) \cdot x_1^2 + (1.646, 0.008)$	3.119%	3
Tea concentration	F-SR ($\alpha = 0.05$)	$\tilde{y}_2 = (1.638, 0.051, 0.051) + (0.320, 0.029, 0.029) \cdot x_1 + (0.088, 0.005, 0.005) \cdot x_3 - (0.099, 0.002, 0.002) \cdot x_1 \cdot x_3$	2.051%	2
	TS-FR	$\tilde{y}_2 = (1.618, 0.070) + (0.282, 0.022) \cdot x_1 + (0.035, 0.000) \cdot x_2 + (0.088, 0) \cdot x_3 + (-0.0220, 0) \cdot x_4 + (0.042, 0) \cdot x_5$	2.206%	5
	Peters-FR	$\tilde{y}_2 = (1.618, 0.184) + (0.347, 0.085) \cdot x_1 + (0.022, 0.007) \cdot x_2 + (0.053, 0.007) \cdot x_3 + (-0.052, 0.007) \cdot x_4 + (0.019, 0.007) \cdot x_5$	2.592%	5
	H-FLSR	$\tilde{y}_1 = (1.612, 0.0676, 0.0670) + (0.1296, 0.039, 0.039) \cdot x_1 + (0.043, 3.454 \times 10^{10}, 3.437 \times 10^{10}) \cdot x_2 + (0.075, 0.005, 0.005) \cdot x_3 + (-0.031, 0.003, 0.003) \cdot x_4 + (0.041, 0.009, 0.009) \cdot x_5$	2.135%	5
	GP-FR	$\tilde{y}_2 = (0.860, 0) \cdot x_1 \cdot x_3 \cdot x_4 \cdot x_5 + (0.129, 0.279 \times 10^{-12}) \cdot x_2 + (0.248, 0) \cdot x_1 + (1.612, 0.063)$	1.8885%	5

Table 7: Generalization errors for catechin content obtained by the tested methods

Cross validation number	F-SR (%)	F-SR (%)	F-SR (%)	F-SR (%)	TS-FR (%)	Peters-FR (%)	H-FLSR (%)	GP-FR (%)
	$\alpha = 0.15$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.02$				
1	6.762	4.869	6.652	6.652	15.823	7.096	0.3839	4.921
2	1.431	1.431	1.588	1.588	5.691	3.556	6.902	5.704
3	3.166	0.687	0.688	2.901	13.607	9.801	7.479	3.126
4	9.187	9.003	10.873	10.873	4.531	8.175	6.522	26.679
5	2.420	2.900	2.900	4.191	25.794	18.932	12.775	11.624
6	0.698	0.414	3.437	3.437	18.113	10.941	2.784	10.658
7	2.030	3.791	2.446	2.446	0.855	2.217	1.880	1.704
8	2.941	4.681	4.681	4.799	10.187	13.152	7.144	2.889
9	3.014	2.579	2.579	2.579	4.558	7.393	8.181	14.057
10	0.5214	2.126	3.035	3.035	3.034	9.971	1.290	1.870
11	3.087	0.195	2.830	2.830	8.746	2.762	0.287	18.913
12	3.445	3.445	2.843	2.843	5.464	3.027	3.775	3.879
13	0.285	1.686	1.686	8.461	1.137	7.222	2.463	1.202
14	7.470	9.282	9.282	9.282	2.812	2.189	5.794	3.465
15	1.990	1.108	1.108	5.841	10.039	10.039	12.954	19.434
16	12.546	11.236	11.236	11.236	4.524	1.368	10.613	9.721
Mean	3.812	3.715	4.241	5.187	8.432	7.983	5.702	8.740
Rank	2	1	3	4	7	6	5	8

Table 8: Generalization errors for tea concentration obtained by the tested methods

Cross validation number	F-SR (%)	F-SR (%)	F-SR (%)	F-SR (%)	TS-FR (%)	Peters-FR (%)	H-FLSR (%)	GP-FR (%)
	$\alpha = 0.15$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.02$				
1	0.029	0.029	0.029	1.202	26.842	18.124	9.690	11.381
2	0.773	0.773	0.77325	0.136	12.437	11.128	5.087	5.926
3	0.130	0.130	2.2888	2.289	18.476	16.092	2.607	5.449
4	1.515	0.604	0.60385	0.604	17.157	17.113	3.517	3.228
5	0.208	0.208	2.0623	2.062	8.518	17.625	1.356	1.358
6	1.087	1.087	2.9859	2.984	14.071	13.361	1.457	1.115
7	2.692	3.481	3.4806	3.481	6.318	7.303	0.635	3.523
8	0.934	0.934	1.0998	1.100	8.899	8.799	2.561	16.469
9	3.746	3.746	1.568	1.568	5.431	9.492	0.630	8.472
10	0.711	0.7108	0.71077	1.611	5.855	11.843	2.533	10.875
11	2.842	2.842	1.9961	1.996	7.891	7.839	1.847	25.644
12	3.541	3.541	5.388	5.388	4.591	5.812	0.698	0.566
13	0.176	0.694	0.69428	0.486	6.646	1.937	6.646	4.282
14	2.243	2.243	2.2231	2.223	0.924	2.947	6.628	7.023
15	0.841	0.8413	0.5081	0.508	8.330	18.922	6.697	17.759
16	6.577	5.975	4.278	4.278	1.258	7.976	5.412	1.442
Mean	1.753	1.740	1.918	1.995	9.603	11.019	3.625	7.782
Rank	2	1	3	4	7	8	5	6

Also, the average numbers of engineering characteristics involved in the models developed by the tested methods are shown in Figure 4 for the two customer preferences. Figure 4 illustrates that the models developed by TS-FR, Peters-FR, and H-FLSR all have five engineering characteristics, as the methods require all the five engineering characteristics in order to generate the models. For the GP-FR, the number of engineering characteristics is smaller than those used by TS-FR, Peters-FR and H-FLSR, where on average 3.267, and 4.533 engineering characteristics are used by the GP-FR when modelling catechin content and tea concentration respectively. The numbers of engineering characteristics used by the F-SR for different significant levels are also shown. Generally, the numbers of engineering characteristics used by the F-SR are less than those used by the other methods. When fewer engineering characteristics are involved in the models, customer preferences can be controlled using fewer engineering

characteristics. Therefore, the simpler tea maker design can be produced.

Also, the results indicate that the F-SR generates the large models with more engineering characteristics when a low significance level is used, since the F-SR rejects fewer hypotheses with a lower significant level. The F-SR generates the smaller models with fewer engineering characteristic when a higher significant level is used, since more hypotheses are rejected. When we look back to [Table 4](#) [Tables 7 and 8](#), the F-SR involved with the significant level of $\alpha = 0.10$ can obtain the smaller errors compared with the other tested F-SR involved with $\alpha = 0.15$, $\alpha = 0.05$, and $\alpha = 0.02$. Hence, the significant level of $\alpha = 0.10$ could be a better choice for this tea maker design. However, [Table 4](#) [Figure 7](#) shows that more engineering characteristics are required on the models which are developed based on the significant level of $\alpha = 0.10$ compared with those of $\alpha = 0.05$, and $\alpha = 0.02$. Hence, the simpleness of the models is degraded when larger significant levels are used. We

determine the trade-off between the errors and the simpleness of the models.

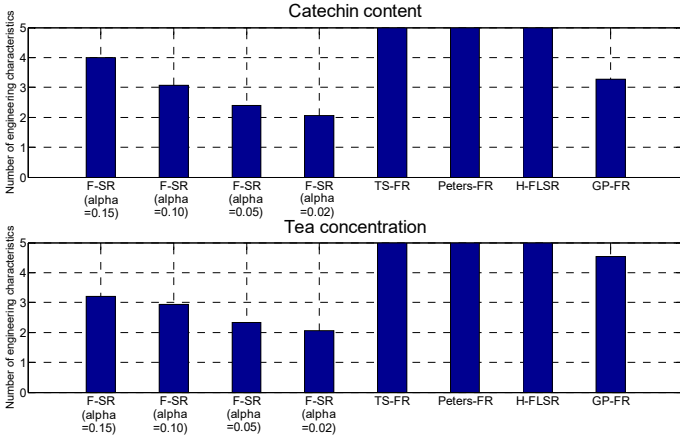


Figure 4: Numbers of engineering characteristics involved in the models obtained by the tested algorithms

D. Analysis and optimization of customer preferences

After the customer preference models are developed, new tea makers can be developed with respect to the specifications of the customer preference. The relationship between the two customer preferences, catechin contents and tea concentration, can be illustrated, based on the customer preference models. Figure 5 illustrates both the catechin contents and tea concentration increase when the engineering characteristics (x_1, x_3, x_4 and x_5), which are engaged with the two customer preference models in (22) and (23), are increased gradually. We can adjust the engineering characteristics with respect to the two consumer preferences. For example, when catechin contents with 1.66% and tea concentration with 1.90% are required by the users, the engineering characteristics with $x_1=97.8$ degree cesius, $x_3=9.7$ minutes, $x_4=4.4$ seconds and $x_5=34$ seconds can be set. This graphic information provides the insights for the production management to specify the customer preferences for the development of new tea makers.

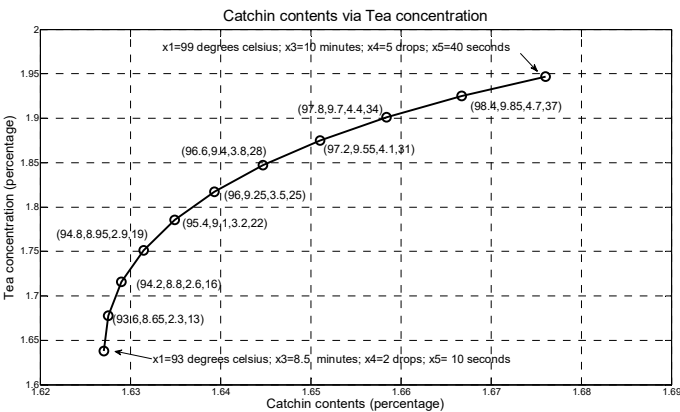


Figure 5: Relationship between catechin content and tea concentration

Based on the customer preference models, the overall customer satisfaction, namely OCS of the tea maker, can be formulated in order to indicate the degree of satisfaction with which a new design of tea maker can fulfil all customer preferences [3]. The OCS can be maximized by optimizing the customer preferences with respect to the engineering characteristics of the tea maker. Here, the linear weighted sum

method was used to derive OCS [3], which can be expressed as equation (24).

$$\begin{aligned}
 \text{OCS} &= \sum_{i=1}^2 w_i \tilde{y}_i \\
 &= 0.5 \left[(1.627, 0.009, 0.009) + (0.180, 0.003, 0.003) \cdot x_1 \cdot x_5 \right. \\
 &\quad \left. - (0.131, 0.002, 0.002) \cdot x_4 \cdot x_5 \right] + 0.5 \left[(1.638, 0.051, 0.051) \right. \\
 &\quad \left. + (0.320, 0.029, 0.029) \cdot x_1 + (0.088, 0.005, 0.005) \cdot x_3 \right. \\
 &\quad \left. - (0.099, 0.002, 0.002) \cdot x_1 \cdot x_3 \right]
 \end{aligned} \tag{24}$$

where \tilde{y}_1 and \tilde{y}_2 can be referred to (22) and (23) respectively. Here, we assume that both the customer preferences are equally important. Hence, we set $w_1 = w_2 = 0.5$. The commonly-used optimization method namely genetic algorithm is used to maximize the OCS. As the development of the genetic algorithm is not main contribution of this research, the detailed mechanism of the genetic algorithm is not discussed. The operations and the parameters used in the genetic algorithm can be referred to in [3]. The optimal normalized engineering characteristics are found as follows, and OCS with 1.822 can be achieved based on these optimal engineering characteristics, $x_1 = 0.999, x_3 = 0.891, x_4 = 0.000$, and $x_5 = 0.996$, where the values of the engineering characteristics are normalized. The real values of the engineering characteristics are $x_1 = 99.9$ degrees Celsius, $x_3 = 9.84$ minutes, $x_4 = 2$ drops, and $x_5 = 69.88$ seconds. Hence, this result illustrates that better OCS can be obtained when the engineering characteristics, namely reheating temperature, the dipping time, and the immersion time, are higher. As higher catechin content and tea concentration can be obtained with higher values of these engineering characteristics, better OCS can be obtained. These results provide the insights for the product management to develop the new tea makers which can achieve the optimal OCS.

V. FURTHER EVALUATION

To further evaluate the modelling effectiveness of the F-SR approach, the F-SR was used to model a solder paste dispenser for electronic packaging automation, which is used to deposit small amounts of solder paste on printed circuit boards. The size of the solder paste must be controlled correctly in order to precisely align the IC leads with each solder spot. For this design problem, the significance level of $\alpha = 0.05$ is used in the F-SR.

The detailed description of the solder paste dispenser can be referred to [51], where the dispenser has three axes of motion (i.e. x, y and z). The (x, y) axes 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 motion of the solder paste dispenser is controlled by a microcontroller whereby the coordinates of the solder paste to be deposited are entered into the computer program. In the solder paste dispenser, the key consumer preference is the diameter of the circular solder pads namely, y , which are controlled by four engineering characteristics: (1) the injection pressure of the screw pump, x_1 , which is used to drive the solder paste; (2) the inside diameter, x_2 , of the interchangeable needle which is used to inject the solder paste; (3) the shot size, x_3 , of the clutch which is engaged

to deposit the amount of solder paste; and (4) the dwell time, x_4 , which is the amount of time the dispenser remains over the location after the pump has been disengaged. Experiments based on a full factorial design with two levels of each engineering characteristics were conducted, where levels 1 and 2 for the pump pressure, x_1 , were set at 4 psi and 12 psi respectively; levels 1 and 2 for needle inner diameter, x_2 , were set at 0.0016 inch and 0.023 inch respectively; levels 1 and 2 for the shot size, x_3 , were set at 150 ms and 500 ms respectively; and levels 1 and 2 for the dwell time, x_4 , were set at 0 second and 0.5 second respectively.

To measure the diameter of the circular solder paste, y , the images of the solder pastes were first captured by an image

scanner and the diameters of the solder pastes were estimated by a vision builder for automated inspection. Table 9 shows the experimental results for the diameter of the circular solder paste Y of which five replications were involved. To address the randomness of the customer preferences, the real observation, $(\tilde{y}^C(k), \tilde{y}^L(k), \tilde{y}^R(k))$ with $k=1,2,\dots,16$, are generated based on (18), (19) and (20), where $(\tilde{y}^C(k), \tilde{y}^L(k), \tilde{y}^R(k))$ represents the k -th real observation for the solder spot diameters. Five experimental replications (i.e. $N_r = 5$) were collected for this design problem.

Table 9: The 2^4 factorial experiment with five replications [54]

The k -th experiment	Levels				Solder spot diameters (Replications)					fuzzy number		
	x_1	x_2	x_3	x_4	$y_1(k)$	$y_2(k)$	$y_3(k)$	$y_4(k)$	$y_5(k)$	$y^c(k)$	$y^L(k)$	$y^R(k)$
1	1	1	1	1	0.040	0.045	0.045	0.040	0.050	0.044	0.004	0.006
2	2	1	1	1	0.035	0.045	0.045	0.050	0.045	0.044	0.004	0.006
3	1	2	1	1	0.045	0.050	0.055	0.055	0.050	0.051	0.006	0.004
4	2	2	1	1	0.040	0.040	0.060	0.050	0.050	0.048	0.008	0.012
5	1	1	2	1	0.080	0.085	0.080	0.085	0.085	0.083	0.003	0.002
6	2	1	2	1	0.085	0.085	0.085	0.080	0.085	0.084	0.004	0.001
7	1	2	2	1	0.085	0.085	0.085	0.090	0.090	0.087	0.002	0.003
8	2	2	2	1	0.080	0.080	0.080	0.075	0.070	0.077	0.007	0.003
9	1	1	1	2	0.055	0.055	0.055	0.055	0.055	0.055	0.000	0.000
10	2	1	1	2	0.055	0.055	0.055	0.050	0.055	0.054	0.004	0.001
11	1	2	1	2	0.050	0.050	0.050	0.050	0.050	0.050	0.000	0.000
12	2	2	1	2	0.060	0.045	0.040	0.040	0.040	0.045	0.005	0.015
13	1	1	2	2	0.085	0.085	0.085	0.085	0.085	0.085	0.000	0.000
14	2	2	1	1	0.080	0.085	0.085	0.085	0.085	0.084	0.004	0.001
15	1	2	2	2	0.085	0.085	0.085	0.085	0.085	0.085	0.000	0.000
16	2	2	2	2	0.075	0.075	0.075	0.075	0.080	0.076	0.001	0.004

To investigate the generalization capabilities of the models developed by the F-SR, cross-validation was used, where 4 pieces of data were randomly selected from 16 pieces of data as a test set, and the remaining 12 pieces of data were used to develop the fuzzy linear regression models. The prediction errors were calculated. Ten cross-validations with different test sets and training sets were conducted. Table 10 summarizes the prediction errors for the 10 cross-validations, and the mean prediction errors of the five tested methods namely F-SR, TS-FR, Peters-FR, H-FLSR and GP-FR. Table 10 shows that the prediction accuracies obtained by the F-SR are generally better than those obtained by the tested linear fuzzy regression approaches, TS-FR, Peters-FR and H-FLSR. GP-FR is only slightly better than F-SR. Table 11 shows the number of engineering characteristics used on the models generated by the five tested methods. It shows that the number of engineering characteristics used on the models generated by the F-SR is smaller than those generated by the other four tested methods. Hence, the simpler but more effective models can be generated by the F-SR.

Table 10: Cross-validation of the five algorithms

Validation data sets	F-SR	TS-FR	Peters-FR	H-FLSR	GP-FR
	$\alpha = 0.05$ (%)	(%)	(%)	(%)	(%)
3,6,10,12	6.226	9.352	6.782	5.488	5.147
2,5,9,10	6.7354	8.791	8.264	7.554	5.370
1,5,8,15	7.678	8.310	9.541	6.204	5.079
10,11,15,16	4.826	7.307	9.353	7.303	7.134
3,4,10,11	5.315	13.720	10.321	9.263	7.324
5,6,10,16	3.320	5.595	4.128	4.757	3.109
2,7,13,15	7.814	5.633	7.718	8.237	6.278
5,6,10,15	1.962	7.974	6.938	5.503	6.469
6,7,10,11	3.736	7.743	8.283	6.954	6.649
8,9,13,16	5.644	12.680	12.339	5.676	6.576
Mean	5.326	8.711	8.367	6.694	5.914

Table 11: Numbers of engineering characteristics involved on the models

Algorithms	F-SR	TS-FR	Peters-FR	H-FLSR	GP-FR
	$\alpha = 0.05$				
No. eng. char.	3.3	4	4	4	3.3

VI. CONCLUSION AND FUTURE RESEARCH

In this paper, a fuzzy stepwise regression (F-SR) approach is proposed that incorporates the mechanisms of fuzzy least square regression and statistical stepwise regression, in order to

develop a customer preference model which is structured with an appropriate polynomial that includes significant regressors only. It overcomes the limitations of the commonly-used fuzzy regression which do not develop appropriate polynomial structures that include only significant regressors, prior to developing the customer preference model. The F-SR uses the statistical stepwise regression to develop the appropriate polynomial structures. It uses the fuzzy least square regression to determine the fuzzy coefficients of the determined polynomial, as the fuzzy least square regression is effective in addressing randomness of consumer preferences. A case study of a tea maker design showed that the F-SR was able to generate consumer preference models with better generalization capabilities than the usual fuzzy regression methods. Also, the F-SR was able to generate consumer preference models with fewer engineering characteristics than those generated by the other tested methods. This case study demonstrated that the F-SR was able to develop a simpler and more effective consumer preference model for new product development. Also another case study of solder paste dispenser design was employed to further validate the effectiveness of the F-SR.

In the future, we will further improve the effectiveness of the proposed F-SR by modifying the strategy of regressor selection. Here the F-SR generates the initial regressor polynomial based on the one with the largest F -value to the system response. In the future, we will further improve the F-SR by modifying the strategy of generating the initial regressor polynomial. The heuristic optimization methods, such as evolutionary algorithms which have commonly been used on determining polynomial structures on models, will be used to generate the initial regressor polynomial. It would increase the chance to generate a better model. Also, we will apply the proposed method for selecting significant regressors to other fuzzy linear models, in order to determine more appropriate fuzzy coefficients. The resulting methods will be applied to model the other real-world systems involving feature selection.

Acknowledgement: The authors would like to thank the four anonymous referees and the associate editor for their helpful comments. Also the first authors would like to thank C.K. Kwong for many useful discussions and providing the data sets for this research.

REFERENCES

- [1] A. Bissierier, R. Boukezzoula, and S. Galichet, A revisited approach to linear fuzzy regression using trapezoidal fuzzy Intervals, *Information Sciences*, vol. 180, pp. 3653-3673, 2010.
- [2] K.Y. Chan, C.K. Kwong, T.S. Dillon, and K.Y. Fung, An intelligent fuzzy regression approach for affective product design that captures nonlinearity and fuzziness, *Journal of Engineering Design*, vol. 22, no. 8, pp. 523-542, 2011.
- [3] K.Y. Chan, C.K. Kwong, and T.S. Dillon, *Computational Intelligence Techniques for New Product Design*, Springer: Studies in Computational Intelligence 403, 2012.
- [4] Y.H.O. Chang, Hybrid fuzzy least squares regression analysis and its reliability measures, *Fuzzy Sets and Systems*, vol. 119, pp. 225-246, 2001.
- [5] L.H. Chen, C.C. Hsueh, and C.J. Chang, A two-stage approach for formulating fuzzy regression models, *Knowledge-Based Systems*, vol. 52, pp. 302-310, 2013.
- [6] Y. Chen, J. Tang, R.Y.K. Fung, and Z. Ren, Apply fuzzy regression to model functional relationship in product planning, *Proceedings of the IEEE Conference on Decision and Control*, pp. 4747-4752, 2003.
- [7] Y. Chen, J. Tang, R.Y.K. Fung, and Z. Ren, Fuzzy regression-based mathematical programming model for quality function deployment, *International Journal of Production Research*, vol. 42, pp. 1009-10027, 2004.
- [8] Y. Chen and L. Chen, A non-linear possibilistic regression approach to model functional relationships in product planning, *International Journal of Advanced Manufacturing Technology*, vol. 28, no. 11-12, 1175-1181, 2005.
- [9] S. Cheng, Y. Zhou, J. Li, J. Lang, and H. Wang, A new statistical modeling and optimization framework for establishing high-resolution PM₁₀ emission inventory - I. Stepwise regression model development and application, *Atmospheric Environment*, vol. 60, pp. 613-622, 2012.
- [10] P. D'Urso, R. Massar and A. Santoro, Robust fuzzy regression analysis, *Information Sciences*, vol. 181, pp. 4154-4174, 2011.
- [11] D. Dubois and H. Prade, The three semantics of fuzzy sets. *Fuzzy Sets and Systems*, vol. 90, no. 2, pp. 141 – 150, 1997.
- [12] D. Dubois, and H. Prade, Gradualness, uncertainty and bipolarity: Making sense of fuzzy sets, *Fuzzy Sets and Systems*, vol. 192, no. 1, pp. 3 – 24, 2012.
- [13] R.Y.K. Fung, Y. Chen, J. Tang, and Y. Tu, Estimating functional relationships for product planning under uncertainties, *Fuzzy Sets and Systems*, vol. 157, pp. 98-120, 2006.
- [14] R.A. Ghunem, K. Assaleh and A.H. El-Hag, Artificial neural networks with stepwise regression for predicting transformer oil furan content, *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 19, no. 2, pp. 414-420, 2012.
- [15] A. Griffin and J.R. Hauser, The voice of the customer, *Marketing Sciences*, vol. 12, no. 1, pp. 1-27, 1993.
- [16] E. Grigoroudis and Y. Siskos, Preference disaggregation for measuring and analyzing customer satisfaction -The MUSA method, *European Journal of Operational Research*, vol. 143, pp. 148-170, 2002.
- [17] E. Grigoroudis, C. Litos, V.A. Moustakis, Y. Politis, and L. Tsironis, The assessment of user-perceived web quality: Application of a satisfaction benchmarking approach, *European Journal of Operational Research*, vol. 187, pp. 1346-1357, 2008.
- [18] S.H. Han, M.H. Yun, K.J. Kim, J. Kwahk, Evaluation of product usability: development and validation of usability dimensions and design elements based on empirical models, *International Journal of Industrial Ergonomics*, vol. 26, pp. 477-488, 2000.
- [19] J.R. Hauser and D. Clausing, The house of quality, *Harvard Business Review*, 66, 63-73, 1988.
- [20] Y.Q. He, L.K. Chan and M.L. Wu, Balancing productivity and consumer satisfaction for profitability: Statistical and fuzzy regression analysis, *European Journal of Operational Research*, vol. 176, pp. 252-263, 2007.
- [21] A.G. Ivakhnenko, Polynomial theory of complex systems, *IEEE Transactions on Systems, Man and Cybernetics*, vol. 1, no. 4, pp. 364-378, 1971.
- [22] H.M. Jiang, C.K. Kwong, W.H. Ip and Z.Q. Chen, Chaos-based fuzzy regression approach to modelling customer satisfaction for product design, *IEEE Transactions on Fuzzy Systems*, vol. 21, no. 5, pp. 926-936, 2013.
- [23] T. Johnsson, A procedure for stepwise regression analysis, *Statistical Papers*, vol. 33, pp. 21-29, 1992.
- [24] E.E. Karsak, Robot selection using an integrated approach based on quality function deployment and fuzzy regression, *International Journal of Product Research*, vol. 46, no. 3, pp. 723-738, 2008.
- [25] M. Kelkinna and S.M. Taheri, Fuzzy least-absolutes regression using shape-preserving operations, *Information Sciences*, vol. 214, pp. 105-120, 2012.
- [26] K.J. Kim, H. Moskowit, A. Dhingra and G. Evans, Fuzzy multi-criteria models for quality function deployment, *European Journal of Operational Research*, vol. 121, pp. 504-518, 2000.

- [27] C.K. Kwong, Y. Chen, K.Y. Chan, and X. Luo, A generalised fuzzy least-squares regression approach to modelling relationships in QFD, *Journal of Engineering Design*, vol. 21, no. 5, pp. 601-613, 2010.
- [28] A. Kusiak and F.A. Salustri, Computational intelligence in product design engineering: review and trends, *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 37, no. 5, pp. 766-778, 2007.
- [29] X. Liao, Q. Li, X. Yang, W. Zhang and W. Li, Multiobjective optimization for crash safety design of vehicles using stepwise regression model, *Structural and Multidisciplinary Optimization*, vol. 35, pp. 561-569, 2008.
- [30] Y.C. Lin, H.H. Lai, and C.H. Yeh, Consumer-oriented product form design based on fuzzy logic: A case study of mobile phones, *International Journal of Industrial Ergonomics*, vol. 37, pp. 531-543, 2007.
- [31] K.P. Lin, P.F. Pai, Y.M. Lu, and P.T. Chang, Revenue forecasting using a least-squares support vector regression model in a fuzzy environment, *Information Sciences*, vol. 220, pp. 196-209, 2013.
- [32] X. Liu, X. Zeng, Y. Xu, and L. Koehl, A fuzzy model for customer satisfaction index in e-commerce, *Mathematics and Computers in Simulation*, vol. 77, pp. 512-521, 2007.
- [33] S. Orsborn, J. Cagan, and P. Boatwright, Quantifying aesthetic form preference in a utility function, *Journal of Mechanical Design*, vol. 131, pp. 1-10, 2009.
- [34] J. Park, and S.H. Han, A fuzzy rule-based approach to modeling affective user satisfaction towards office chair design, *International Journal of Industrial Ergonomics*, vol. 34, pp. 31-47, 2004.
- [35] G. Peters, Fuzzy linear regression with fuzzy intervals, *Fuzzy Sets and Systems*, vol. 63, pp. 45-55, 1994.
- [36] M. S. Phadke, *Quality engineering using robust design*. New York: Prentice-Hall, 1987.
- [37] G. Marsaglia, W.W. Tsang, J. Wang, Evaluating Kolmogorov's Distribution, *Journal of Statistical Software*, vol. 8, no. 18, pp. 1-4, 2003.
- [38] M.H. Mashinchi, M.A. Orgun, M. Mashinchi and W. Pedrycz, A tabu harmony search-based approach to fuzzy linear regression, *IEEE Transactions on Fuzzy Systems*, vol. 19, no. 3, pp. 432-448, 2011.
- [39] G. Mauris, L. Berrah, L. Foulloy and A. Haurat, Fuzzy handling of measurement errors in instrumentation, *IEEE Transactions on Instrumentation and Measurement*, vol. 49, no. 1, pp. 89-93, 2000.
- [40] N. I. Nikolaev and H. Iba, Accelerated genetic programming of polynomials, *Genetic Programming and Evolvable Machines*, vol. 2, no. 3, pp. 231-257, 2001.
- [41] M. Oussalah, On the compatibility between defuzzification and fuzzy arithmetic operations, *Fuzzy Sets and Systems*, vol. 128, pp. 247-260, 2002.
- [42] A.A. Ramli, J. Watada and W. Pedrycz, Real-time fuzzy regression analysis: a convex hull approach, *European Journal of Operational Research*, vol. 201, pp. 606-617, 2011.
- [43] J.O. Rawlings, *Applied Regression Analysis: A Research Tool*, Wadsworth and Brooks, 1988.
- [44] A. Safari, H. Shayeghi, and H. A. Shayanfar, Robust state feedback controller design of STATCOM using chaotic optimization algorithm, *Serbian Journal of Electrical Engineering*, vol. 7, no. 2, pp. 253-268, 2010.
- [45] Z. Sener and E.E. Karsak, A decision model for setting target levels in quality function development using nonlinear-programming-based fuzzy regression and optimization, *International Journal of Advanced Manufacturing Technologies*, vol. 48, pp. 1173-1184, 2010.
- [46] Z. Sener and E.E. Karsak, A combined fuzzy linear regression and fuzzy multiple objective programming approach for setting target levels in quality function deployment, *Expert Systems with Applications*, vol. 38, pp. 3015-3022, 2011.
- [47] S. Sonoda, Y. Takahashi, K. Kawagishi, N. Nishida and S. Wakao, Application of stepwise multiple regression to design optimization of electric machine, *IEEE Transactions on Magnetics*, vol. 43, no. 4, pp. 1609-1612, 2007.
- [48] H. Tanaka and J. Watada, Possibilistic linear systems and their application to the linear regression model, *Fuzzy Sets and Systems*, vol. 272, pp. 275-289, 1988.
- [49] J. Watada, S. Wang and W. Pedrycz, Building confidence-interval-based fuzzy random regression models, *IEEE Transactions on Fuzzy Systems*, vol. 17, no. 6, pp. 1273-1283, 2009.
- [50] Y.M. Wi, S.K. Joo, and K.B. Song, Holiday load forecasting using fuzzy polynomial regression with weather feature selection and adjustment, *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 596-603, 2012.
- [51] R.R. Yager and D.P. Filev, On the issue of defuzzification and selection based on a fuzzy set, *Fuzzy Sets and Systems*, vol. 55, no. 3, pp. 255-272, 1993.
- [52] H. You, T. Ryu, K. Oh, M.H. Yun, and K.J. Kim, Development of customer satisfaction models for automotive interior materials, *International Journal of Industrial Ergonomics*, vol. 36, pp. 323-330, 2006.
- [53] N. Zhou, J.W. Pierre, and D. Trudnowski, A stepwise regression method for estimating dominant electromechanical modes, *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 1051-1059, 2012.
- [54] C. K. Kwong, Y. Chen, K.Y. Chan, and H. Wong, The Hybrid Fuzzy Least-Squares Regression Approach to Modeling Manufacturing Processes, *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 3, pp. 644-651, 2008.
- [55] R. Unal, D.O. Stanley and C.R. Joyner, Propulsion system design optimization using the Taguchi method, *IEEE Transactions on Engineering Management*, vol. 40, no. 3, pp. 315-322, 1993.
- [56] D.G. Kleinbaum and L.L. Kupper, *Applied regression analysis and other multivariable methods*, Duxbury Press, 1978.
- [57] K.Y. Chan, C.K. Kwong and M.C. Law, A fuzzy ordinary regression method for modeling customer preference in tea maker design, *Neurocomputing*, vol. 142, pp. 147-154, 2014.

Appendix

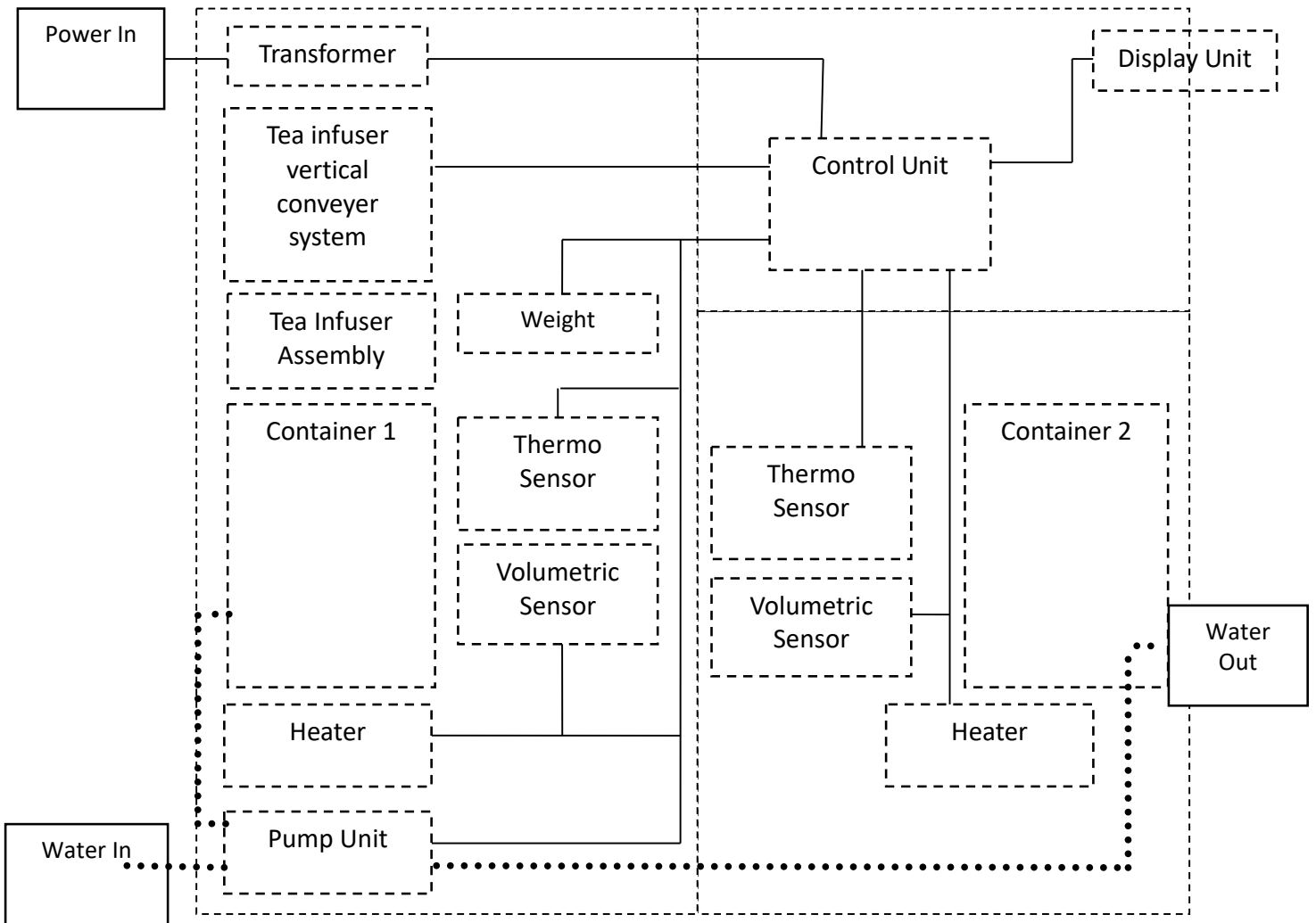
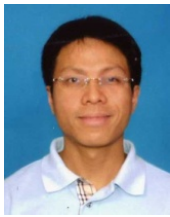


Figure 1 Schematic diagram of conceptual tea maker in the market



Kit Yan Chan (M'11) received the PhD degree in Computing from London South Bank University, United Kingdom in 2006. After his PhD study, he was a post-doctoral researcher at The Hong Kong Polytechnic University, Hong Kong and Curtin University, Australia. Currently, he is a Senior Lecturer in the Department of Electrical and Computer Engineering, Curtin University. He has published 51 journal papers, 41 conference papers, 2 book chapters, and 2 research monographs. His research interests include computational intelligence and its applications to signal processing, new product design, manufacturing process design, and traffic flow control.



Hak Keung Lam (SM'10) received the B.Eng. (Hons.) and Ph.D. degrees from the Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hong Kong, in 1995 and 2000, respectively. During the period of 2000 and 2005, he worked with the Department of Electronic and Information Engineering at The Hong Kong Polytechnic University as Post-Doctoral Fellow and Research Fellow respectively. He joined as a Lecturer at King's College London in 2005 and currently a Reader. His current research interests include intelligent control systems and computational intelligence. He has served as a program committee member and international advisory board member for various international conferences and a reviewer for various books, international journals and international conferences. He is an associate editor for IEEE Transactions on Fuzzy Systems, IET Control Theory and Applications, International Journal of Fuzzy Systems and Neurocomputing; and guest editor for a number of international journals. He is in the editorial board of Journal of Intelligent Learning Systems and Applications, Journal of Applied Mathematics, Mathematical Problems in Engineering, Modelling and Simulation in Engineering, Annual Review of Chaos Theory, Bifurcations and Dynamical Systems and The Open Cybernetics and Systemics Journal?. He is an IEEE senior member. He is the coeditor for two edited volumes: Control of Chaotic Nonlinear Circuits (World Scientific, 2009) and Computational Intelligence and Its Applications (World Scientific, 2012), and the coauthor of the monograph: Stability Analysis of Fuzzy-Model-Based Control Systems (Springer, 2011). His co-authored paper (J.S. Dai, H.K. Lam and S.M. Vahed, "Soil Type Identification for Autonomous Excavation Based on Dissipation Energy," Proceedings of the Institution of Mechanical Engineers, Part I, Journal of Systems and Control Engineering, vol. 225, no. 1, pp. 35-50, 2011) received SAGE Best Paper Award in 2011.



Tharam Dillon (M'74–SM'87–F'98–LF'10) is a Fellow of the IEEE, ACS and IE(Aust). He is the professor at Department of Computer Science and Computer Engineering, La Trobe University, Australia. He is head of the IFIP International Task Force WG2.12/24 on Semantic Web and Web Semantics, the chairman of the IFIP WG12.9 on computational intelligence, the IEEE/IES Technical Committee on Industrial Informatics, and the IFIP Technical Committee 12 on Artificial Intelligence. He has published more than 750 papers published in international conferences and journals and is the author of 5 books and has another 5 edited books. His current research interests include Web semantics, ontologies, Internet computing, e-commerce, hybrid neurosymbolic systems, neural nets, software engineering, database systems, and data mining. He is the Editor-in-Chief of the *International Journal of*

Computer Systems Science and Engineering as well as the *Engineering Intelligent Systems*.



Sai Ho Ling (M'06–SM'12) received the BEng degree from the Department of Electrical Engineering, M.Phil. and Ph.D. degrees from the Department of Electronic and Information Engineering in the Hong Kong Polytechnic University in 1999, 2002 and 2007 respectively. Currently, he works in University of Technology, Sydney, Australia as a senior lecturer. He has authored and coauthored over 140 book chapters, international journal papers and conference papers on computational Intelligence and its industrial applications. His current research interests include evolution computations, fuzzy logics, neural networks, hybrid systems and biomedical applications. Currently, he serves as Co-Editors-in-Chief for Journal of Intelligent Learning Systems and Applications.