Journal of Applied Sciences

ISSN 1812-5654

Journal of Applied Sciences 14 (7): 631-640, 2014
ISSN 1812-5654 / DOI: 10.3923/jas.2014.631.640

Parameter Estimation of Fuzzy Linear Regression Model: The Extension of Chen and Hsueh Method

Atchanut Rattanalertnusorn, Ampai Thongteeraparp and Winai Bodhisuwan
Department of Statistics, Faculty of Science,
Kasetsart University, Bangkhen, Bangkok, 10900, Thailand

Abstract: Chen and Hsueh (2009) proposed the new method to construct a fuzzy regression model which is based on distance criterion. In this work, Chen and Hsueh model was extended, where the model parameters are estimated by minimized the total estimated error, i.e., the sum of the average squared distance between the observed and estimated responses by using some α-cuts. Also three cases of the estimated coefficients were derived by using the model that the explanatory variables and response variable are Trapezoidal Fuzzy Numbers (TrFN): All of positives coefficients, all of negatives coefficients and positives and negatives coefficients. Two numerical examples are demonstrated the extension of Chen and Hsueh model. Finally, the useful conclusions are provided.

Key words: Fuzzy number, triangular fuzzy number, trapezoidal fuzzy number, fuzzy linear regression model, the α-cuts method

INTRODUCTION

The Fuzzy Linear Regression Model (FLRM) was proposed by Tanaka *et al.* (1982). They developed a linear regression model with fuzzy response, crisp explanatory and fuzzy coefficients. The model can be written in the general form as:

$$\tilde{y}_i = \tilde{A}_0 + \tilde{A}_1 x_{i1} + \tilde{A}_2 x_{i2} + \cdots + \tilde{A}_k x_{ik}, \quad i = 1, \ldots, n \qquad (1)$$

where, $\tilde{y}_i$ is the *i*th fuzzy response, $\tilde{A}_0$ is a fuzzy intercept, $\tilde{A}_j$ is the jth fuzzy coefficient which is corresponding to x_{ij} and x_{ij} is the jth explanatory variable for the ith sample. However, the model has some drawbacks. For example, when using more observations to establish the model leads to fuzzier estimation of parameters, thus making the spread of the estimated fuzzy response wider as stated by Tanaka (1987), Tanaka and Watada (1988), Tanaka *et al.* (1989) and Tanaka and Lee (1998). In addition, Redden and Woodall (1994) pointed out that Tanaka's method is very sensitive to outliers (Chen and Hsueh, 2009).

To prevent the above problems, Diamond (1988) and Kao and Chyu (2002, 2003) adopted numeric coefficients to describe the fuzzy relationship between the fuzzy response variable and the fuzzy explanatory variables. Diamond (1988) also proposed a fuzzy least squares approach and defined a new distance metric to measure error of two fuzzy numbers. Kao and Chyu (2003) also proposed the two-stage method to formulate the FLRM. In the first stage, the fuzzy data, fuzzy response or fuzzy explanatory or both, are defuzzified into the crisp data, then they used the classical least squares method to estimate a regression line for showing the general trend of data. In the second stage, they calculated the estimated error in the model which is based on the criterion of Kim and Bishu (1998). This criterion minimizes the difference of the membership values between the observed and estimated fuzzy responses.

In addition, Buckley (2005) proposed a new method to estimate parameters in the FLRM by using a set of confidence intervals. This set can produced the approximated triangular shaped fuzzy numbers. The approximated triangular shaped fuzzy numbers can be used as the estimated coefficients. He also employed fuzzy prediction and fuzzy hypothesis testing. Kim *et al.* (2005) proposed the fuzzy least absolute deviation method and constructed the FLRM with fuzzy input. They also provided some numerical examples and evaluated an effectiveness of the fuzzy absolute deviation method compared with the fuzzy least squares method.

Recently, Chen and Hsueh (2007) proposed a mathematical programming to construct the FLRM based on the concept of distance. The total estimation error of the model is the sum of distances between the observed and estimated responses. Later Chen and Hsueh (2009)

Corresponding Author: Winai Bodhisuwan, Department of Statistics, Faculty of Science, Kasetsart University, Bangkhen, Bangkok, 10900, Thailand

also proposed the FLRM based on the concept of distance and adapted the least squares method to estimate model parameters. They also derived the estimated model parameters and the estimated adjustment term. The explanatory variables and response variable in the model are Triangular Fuzzy Numbers (TFN). However, those variables can be the other types of fuzzy number such as trapezoidal, normal, sigmoid and etc. In particular, TFN are the special case of Trapezoidal Fuzzy Numbers (TrFN).

The purpose of this study is to estimate model parameters, when the response and fuzzy explanatory are TrFN. Also the estimated model parameters and the estimated adjustment term are derived.

FUZZY LINEAR REGRESSION MODEL

In this study, the FLRM model can be written in the general form as:

$$\tilde{y}_i = \beta_0 + \beta_1\tilde{x}_{i1} + \beta_2\tilde{x}_{i2} + \cdots + \beta_p\tilde{x}_{ip}, \quad i = 1, 2, 3, ..., n \tag{2}$$

where, $\tilde{y}_i$ is the observed fuzzy response for the ith sample, $\tilde{x}_{ij}$ is the jth fuzzy explanatory variable for the ith sample, j = 1, 2,..., p, β_0 is an intercept and β_j is the jth regression coefficient. This section can be divided into the two subsection. First, Chen and Hsueh method is briefly introduced. Second, the extension of Chen and Hsueh method is presented.

Chen and Hsueh method: In the classical linear regression model, the model can be written in the general form as:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + \varepsilon_i, \quad i = 1, 2, 3, ..., n \tag{3}$$

where, y_i is the response variable for the ith sample, x_{ij} is the jth explanatory variable for the ith sample, β_j is the jth regression coefficient for j = 0, 1, 2..., p, where, β_0 is the intercept and ε_i is a random error for the ith sample. Normally, ε_i is assumed to be independent and identical distribution (iid). The ordinary least squares method (OLS) is widely used to estimate the regression coefficients. The estimated response can be obtained as:

$$\hat{y}_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \cdots + b_p x_{ip}, \quad i = 1, ..., n \tag{4}$$

where, $\hat{y}_i$ is the estimated response for the ith sample, b_0, b_1,... b_p are the estimated intercept and coefficients by OLS method.

By substituting $\tilde{x}_{ij}$ into Eq. 4, the FLRM can be constructed as:

$$\tilde{y}_i = b_0 + b_1\tilde{x}_{i1} + b_2\tilde{x}_{i2} + \cdots + b_p\tilde{x}_{ip}, \quad i = 1, ..., n \tag{5}$$

where, $\tilde{y}_i$ is the fuzzy response for the ith sample, $\tilde{x}_{ij}$ is the jth fuzzy explanatory for the ith sample.

Since TFN are convenient to compute and implement, let $\tilde{x}_{ij}$ is a TFN, denoted by $\tilde{x}_{ij} = (x_{ijl}, x_{ijm}, x_{iju})$ with the membership function:

$$\mu_{\tilde{x}_{ij}}(x) = \begin{cases} \dfrac{x - x_{il}}{x_{ijm} - x_{ijl}}, & x_{ijl} \le x \le x_{ijm} \\ \dfrac{x_{iju} - x}{x_{iju} - x_{ijm}}, & x_{ijm} \le x \le x_{iju} \end{cases}$$

where, x_{ijl}, x_{ijm}, x_{iju} are the left point, the middle point and the right point of $\tilde{x}_{ij}$, respectively, $\mu_{\tilde{x}_{ij}}(x_{ijm}) = 1$ and $\mu_{\tilde{x}_{ij}}(x_{ijl}) = \mu_{\tilde{x}_{ij}}(x_{iju}) = 0$. The support of $\tilde{x}_{ij}$ is x_{iju}-x_{ijl}.

If the explanatory variables are crisp numbers, the estimated responses in Eq. 5 are also crisp numbers which leads to a large fuzzy error for a fuzzy response. To deal with this problem, Chen and Hsueh (2007) added a triangular fuzzy adjustment term $\tilde{\delta}$, defined as $\tilde{\delta} = (\delta_l, \delta_m, \delta_u)$ with the membership function:

$$\mu_{\tilde{\delta}}(\delta) = \begin{cases} \dfrac{\delta - \delta_l}{\delta_m - \delta_l}, & \delta_l \le \delta \le \delta_m \\ \dfrac{\delta_u - \delta}{\delta_u - \delta_m}, & \delta_m \le \delta \le \delta_u. \end{cases}$$

where, δ_l, δ_m and δ_u are regarded as the parameters which are determined. When the triangular fuzzy adjustment term is added into Eq. 5, the FLRM becomes:

$$\tilde{y}_i^* = b_0 + b_1\tilde{x}_{i1} + b_2\tilde{x}_{i2} + \cdots + b_p\tilde{x}_{ip} + \tilde{\delta}^*, \quad i = 1, \ldots, n \tag{6}$$

where, $\tilde{y}_i^*$ is the fuzzy response for the *i*th sample and $\tilde{\delta}^*$ is the triangular fuzzy adjustment term including the intercept $(\tilde{\delta} + b_0)$. This model is used to build up the fuzzy linear regression models. The number of parameters is p+4.

For convenient to implement, Chen and Hsueh method can be summarized as follows:

First, the observed fuzzy response and fuzzy explanatory are defuzzified into crisp number by using the centroid formulae of TFN. Then the Pearson correlation coefficient of the crisp response and explanatory can be calculated.

Second, using the least squares method to estimate the model parameters, the estimated parameters are b_0, b_1..., b_p.

Third, three parameters of the triangular fuzzy adjustment term including the intercept are determined. That is:

$$\hat{\hat{\delta}}^* = (\hat{\delta}_l + b_0, \hat{\delta}_m + b_0, \hat{\delta}_u + b_0)$$

Fourth, the estimated fuzzy response for the ith sample can be calculated by:

$$\hat{\tilde{y}}_i^* = b_1\tilde{x}_{i1} + b_2\tilde{x}_{i2} + \cdots + b_p\tilde{x}_{ip} + \hat{\hat{\delta}}^*, \quad i = 1,\ldots,n \tag{7}$$

where, b_0, b_1..., b_p are the estimated coefficients by the least squares method and $\hat{\hat{\delta}}^*$ is the estimated triangular fuzzy adjustment term including the intercept.

Fifth, the estimated error based on using m α-cuts for the ith sample is defined as:

$$E_i = \frac{1}{2m}\sum_{s=0}^{m-1}\left[\left(\left(\hat{\tilde{y}}_i^*\right)_{\alpha_s}^L - (y_i)_{\alpha_s}^L\right)^2 + \left(\left(\hat{\tilde{y}}_i^*\right)_{\alpha_s}^U - (y_i)_{\alpha_s}^U\right)^2\right] \tag{8}$$

and the total estimated error for n samples can be formulated as:

$$\sum_{i=1}^{n} E_i = \frac{1}{2m}\sum_{i=1}^{n}\sum_{s=0}^{m-1}\left[\left(\left(\hat{\tilde{y}}_i^*\right)_{\alpha_s}^L - (y_i)_{\alpha_s}^L\right)^2 + \left(\left(\hat{\tilde{y}}_i^*\right)_{\alpha_s}^U - (y_i)_{\alpha_s}^U\right)^2\right] \tag{9}$$

Where:

$$(y_i)_{\alpha_s}^L = y_{il} + \alpha_s(y_{im} - y_{il}),\ (y_i)_{\alpha_s}^U = y_{iu} - \alpha_s(y_{iu} - y_{im})$$

$$(x_{ij})_{\alpha_s}^L = x_{ijl} + \alpha_s(x_{ijm} - x_{ijl}),\ (x_{ij})_{\alpha_s}^U = x_{iju} - \alpha_s(x_{iju} - x_{ijm})$$

$$(\delta)_{\alpha_s}^L = \delta_l + \alpha_s(\delta_m - \delta_l),\ (\delta)_{\alpha_s}^U = \delta_u - \alpha_s(\delta_u - \delta_m)$$

$$(\hat{\tilde{y}}_i^*)_{\alpha_s}^L = b_0 + \sum_{j=1}^{p} b_j(x_{ij})_{\alpha_s}^L + (\delta)_{\alpha_s}^L,\ (\hat{\tilde{y}}_i^*)_{\alpha_s}^U = b_0 + \sum_{j=1}^{p} b_j(x_{ij})_{\alpha_s}^U + (\delta)_{\alpha_s}^U$$

and:

$$\alpha_s = \frac{s}{m-1}, \quad s = 0,1,\ldots,m-1$$

Therefore, an approach should be developed to determine the optimal parameters for achieving the minimum total error, $\min\sum_i E_i$, the further details (Chen and Hsueh, 2009).

Extension of Chen and Hsueh method: The model in Eq. 2 is used to study when fuzzy observations $(\tilde{y}_i, x_i^T)$ are TrFN. Then the trapezoidal fuzzy adjustment term, $\tilde{\delta}$ is added into the model which becomes:

$$\tilde{y}_i = \beta_0 + \beta_1\tilde{x}_{i1} + \cdots + \beta_p\tilde{x}_{ip} + \tilde{\delta}, \quad i = 1,\ldots,n \tag{10}$$

where, $\tilde{y}_i$ is the trapezoidal fuzzy response for the ith sample, $\tilde{x}_{ij}$ is the jth trapezoidal fuzzy explanatory for the ith sample, β_0 is an intercept, β_j is the jth parameter and $\tilde{\delta}$ is the trapezoidal fuzzy adjustment term which is denoted by $\tilde{\delta} = (\delta_l, \delta_{m_1}, \delta_{m_2}, \delta_u)$.

This study the estimated fuzzy response for the ith sample which is based on a particular α_S level can be defined as:

$$(\hat{\tilde{y}}_i^*)_{\alpha_s} = \hat{\beta}_0 + \hat{\beta}_1\tilde{x}_{i1} + \cdots + \hat{\beta}_p\tilde{x}_{ip} + (\hat{\tilde{\delta}})_{\alpha_s}, \quad i = 1,\ldots,n \tag{11}$$

or can be rewritten in the interval form:

$$(\hat{\tilde{y}}_i^*)_{\alpha_s} = \left[(\hat{y}_i^*)_{\alpha_s}^L,\ (\hat{y}_i^*)_{\alpha_s}^U\right] \tag{12}$$

Where:

$$(\hat{y}_i^*)_{\alpha_s}^L = \hat{\beta}_0 + \sum_{j=1}^{p}\hat{\beta}_1(\tilde{x}_{i1})_{\alpha_s}^L + (\hat{\tilde{\delta}})_{\alpha_s}^L$$

$$(\hat{y}_i^*)_{\alpha_s}^U = \hat{\beta}_0 + \sum_{j=1}^{p}\hat{\beta}_1(\tilde{x}_{i1})_{\alpha_s}^U + (\hat{\tilde{\delta}})_{\alpha_s}^U$$

and $[\hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_p]^T$ is a vector of the estimated parameters.

For estimating model parameters, we present the two estimation methods as follows:

- **Ordinary least squares method (OLS):** Generally the OLS method is used to estimate regression parameters in the linear regression model. Similarly, this study the OLS method is applied to fit the model in Eq. 10. Further details of OLS the method can see in Myers and Milton (1991). By using the OLS method, we obtained a vector of estimated parameters as $\hat{\beta}_{LSE} = [\hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_p]^T$ where $\hat{\beta}_j$ is the jth estimated parameter for j = 1,..., p except $\hat{\beta}_0$ is called the estimated intercept
- **α-cuts method:** This method can be estimated model parameters in Eq. 10 by minimizing the total estimated error in Eq. 8 which is based on the α-cuts method. By using m α-cuts, we obtained a vector of the estimated parameters as $\hat{\beta}_\alpha = [\hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_p]^T$, where $\hat{\beta}_0 = 0$ and $\hat{\beta}_j$ is the jth estimated parameter.

ESTIMATION OF MODEL PARAMETERS

Based on the α-cuts method, the estimated model parameters were derived as follows.

Let $\tilde{y}_i, \tilde{x}_{ij}$ and $\tilde{\delta}$ are trapezoidal fuzzy numbers defined as $\tilde{y}_i=(y_{il},y_{im_1},y_{im_2},y_{iu})$, $\tilde{x}_{ij}=(x_{ijl},x_{ijm_1},x_{ijm_2},x_{iju})$ and $\tilde{\delta}=(\delta_l,\delta_{m_1},\delta_{m_2},\delta_u)$, respectively. Let the α-cuts of $\tilde{y}_i, \tilde{x}_{ij}$ and $\tilde{\delta}$ are the closed intervals which are denoted as $(\tilde{y}_i)_\alpha=\left[(y_i)_\alpha^L,\ (y_i)_\alpha^U\right]$, $(\tilde{x}_{ij})_\alpha=\left[(x_{ij})_\alpha^L,\ (x_{ij})_\alpha^U\right]$ and $(\tilde{\delta})_\alpha=\left[(\delta)_\alpha^L,\ (\delta)_\alpha^U\right]$, respectively, where $\alpha\in[0,1]$.

According to Chen and Hsueh (2007, 2009), the estimate error for the ith sample is the average squared distance between $\tilde{y}_i$ and $\hat{y}_i^*$ on the X-axis at the particular α_S level, defined by:

$$e_i=\frac{1}{2}\left[\left((\hat{y}_i^*)_{\alpha_s}^L-(y_i)_{\alpha_s}^L\right)^2+\left((\hat{y}_i^*)_{\alpha_s}^U-(y_i)_{\alpha_s}^U\right)^2\right] \qquad (13)$$

where, $(y_i)_{\alpha_s}^L,(y_i)_{\alpha_s}^U$ are the lower bound and the upper bound of the observed fuzzy response $\tilde{y}_i$ and $\left(\hat{y}_i^*\right)_{\alpha_s}^L,\left(\hat{y}_i^*\right)_{\alpha_s}^U$ are the lower bound and the upper bound of the estimated fuzzy response $\hat{y}_i^*$ at the particular α_S level.

Since a group of α-cuts can be represented as an approximated fuzzy number, thus the estimated error for the ith sample which is based on using m α-cuts can be formulated as:

$$e_i^m=\frac{1}{m}\sum_{s=0}^{m-1}e_i \qquad (14)$$

But, the model has n samples, thus the total estimated error which is based on using m α-cuts can be defined by:

$$T_e=\sum_{i=1}^{n}e_i=\sum_{i=1}^{n}\frac{1}{2m}\left[\left((\hat{y}_i^*)_{\alpha_s}^L-(y_i)_{\alpha_s}^L\right)^2+\left((\hat{y}_i^*)_{\alpha_s}^U-(y_i)_{\alpha_s}^U\right)^2\right] \qquad (15)$$

where, α_S is a particular α-cut level, s is an index number and m is the number of α-cuts which are given by:

$$\alpha_s=\frac{s}{m-1},\quad s=0,1,\ldots,m-1 \qquad (16)$$

Based on the particular α_S level, the lower bound and the upper bound of trapezoidal fuzzy adjustment term $(\tilde{\delta})$ can be expressed as:

$$(\tilde{\delta})_{\alpha_s}^L=\delta_l+\frac{s}{m-1}(\delta_{m_1}-\delta_l)\text{ and }(\tilde{\delta})_{\alpha_s}^U=\delta_u-\frac{s}{m-1}(\delta_u-\delta_{m_2}) \qquad (17)$$

Since the model parameters can be all of positives or all of negatives or mixed (some of positives and negatives), thus the estimated model parameters can be derived in the three cases as follows:

Let:

$$1=\begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}_{n\times1},\quad ß=\begin{bmatrix}\beta_1\\ \vdots\\ \beta_p\end{bmatrix}_{p\times1}$$

$$X_{\alpha_S}^L=\begin{bmatrix}(x_{11})_{\alpha_S}^L & \cdots & (x_{1p})_{\alpha_S}^L\\ \vdots & \cdots & \vdots\\ (x_{n1})_{\alpha_S}^L & \cdots & (x_{np})_{\alpha_S}^L\end{bmatrix}_{n\times p},\quad X_{\alpha_S}^U=\begin{bmatrix}(x_{11})_{\alpha_S}^U & \cdots & (x_{1p})_{\alpha_S}^U\\ \vdots & \cdots & \vdots\\ (x_{n1})_{\alpha_S}^U & \cdots & (x_{np})_{\alpha_S}^U\end{bmatrix}_{n\times p}$$

$$\overline{x}_{\alpha_S}^L=\left[(\overline{x}_1)_{\alpha_S}^L,\ \cdots\ ,(\overline{x}_p)_{\alpha_S}^L\right]_{1\times p},\quad \overline{x}_{\alpha_S}^U=\left[(\overline{x}_1)_{\alpha_S}^U,\ \cdots\ ,(\overline{x}_p)_{\alpha_S}^U\right]_{1\times p}$$

$$X_{\alpha_S}^*=\begin{bmatrix}(x_{11})_{\alpha_S}^L & \cdots & (x_{1k})_{\alpha_S}^L & (x_{1(k+1)})_{\alpha_S}^U & \cdots & (x_{1p})_{\alpha_S}^U\\ \vdots & & \vdots & \vdots & & \\ (x_{n1})_{\alpha_S}^L & \cdots & (x_{nk})_{\alpha_S}^L & (x_{n(k+1)})_{\alpha_S}^U & \cdots & (x_{np})_{\alpha_S}^U\end{bmatrix}_{n\times p}$$

$$\overline{x}_{\alpha_S}^*=\left[(\overline{x}_1)_{\alpha_S}^L,\ldots,(\overline{x}_k)_{\alpha_S}^L,(\overline{x}_{k+1})_{\alpha_S}^U,\ldots,(\overline{x}_p)_{\alpha_S}^U\right]_{1\times p}$$

$$X_{\alpha_S}^{**}=\begin{bmatrix}(x_{11})_{\alpha_S}^U & \cdots & (x_{1k})_{\alpha_S}^U & (x_{1(k+1)})_{\alpha_S}^L & \cdots & (x_{1p})_{\alpha_S}^L\\ \vdots & & \vdots & \vdots & & \\ (x_{n1})_{\alpha_S}^U & \cdots & (x_{nk})_{\alpha_S}^U & (x_{n(k+1)})_{\alpha_S}^L & \cdots & (x_{np})_{\alpha_S}^L\end{bmatrix}_{n\times p}$$

$$\overline{x}_{\alpha_S}^{**}=\left[(\overline{x}_1)_{\alpha_S}^U,\ldots,(\overline{x}_k)_{\alpha_S}^U,(\overline{x}_{k+1})_{\alpha_S}^L,\ldots,(\overline{x}_p)_{\alpha_S}^L\right]_{1\times p}$$

$$(\overline{x}_j)_{\alpha_S}^L=\frac{\sum_{i=1}^{n}(x_{ij})_{\alpha_S}^L}{n},\quad (\overline{x}_j)_{\alpha_S}^U=\frac{\sum_{i=1}^{n}(x_{ij})_{\alpha_S}^U}{n},\quad j=1,2,\ldots,p$$

$$y_{\alpha_S}^L=\begin{bmatrix}(y_1)_{\alpha_S}^L\\ \vdots\\ (y_n)_{\alpha_S}^L\end{bmatrix}_{n\times1},\quad y_{\alpha_S}^U=\begin{bmatrix}(y_1)_{\alpha_S}^U\\ \vdots\\ (y_n)_{\alpha_S}^U\end{bmatrix}_{n\times1}$$

$$\overline{y}_{\alpha_S}^L=\frac{\sum_{i=1}^{n}(y_i)_{\alpha_S}^L}{n},\quad \overline{y}_{\alpha_S}^U=\frac{\sum_{i=1}^{n}(y_i)_{\alpha_S}^U}{n}$$

Case 1: The estimated parameters are all of positives ($b_j\geq0$, j = 1,..., p), then total error T_e in Eq. 15 can be rewritten in the matrix form as:

$$T_e=\sum_{i=1}^{n}\frac{1}{2m}\{[y_{\alpha_s}^L-(X_{\alpha_s}^Lß+(1\overline{y}_{\alpha_s}^L-1\overline{x}_{\alpha_s}^Lß))]^T[y_{\alpha_s}^L-(X_{\alpha_s}^Lß+(1\overline{y}_{\alpha_s}^L-1\overline{x}_{\alpha_s}^Lß))]+$$
$$[y_{\alpha_s}^U-(X_{\alpha_s}^Uß+(1\overline{y}_{\alpha_s}^U-1\overline{x}_{\alpha_s}^Uß))]^T[y_{\alpha_s}^U-(X_{\alpha_s}^Uß+(1\overline{y}_{\alpha_s}^U-1\overline{x}_{\alpha_s}^Uß))]\} \qquad (18)$$

By minimizing T_e, the first partial derivative of T_e with respect to β and set to zero, the estimated parameter vector $\hat{\beta}_1$ is obtained as:

$$\hat{\beta}_1 = \left[\sum_{s=0}^{m-1}\{[X_{\alpha_s}^L - 1\bar{x}_{\alpha_s}^L]^T[X_{\alpha_s}^L - 1\bar{x}_{\alpha_s}^L] + [X_{\alpha_s}^U - 1\bar{x}_{\alpha_s}^U]^T[X_{\alpha_s}^U - 1\bar{x}_{\alpha_s}^U]\}\right]^{-1} \times\sum_{s=0}^{m-1}\{[(X_{\alpha_s}^L)^T y_{\alpha_s}^L - n(\bar{x}_{\alpha_s}^L)^T\bar{y}_{\alpha_s}^L] + [(X_{\alpha_s}^U)^T y_{\alpha_s}^U - n(\bar{x}_{\alpha_s}^U)^T\bar{y}_{\alpha_s}^U]\}. \tag{19}$$

Case 2: The estimated parameters are all of negatives ($b_j \geq 0$, j = 1,…,p), then total error T_e in Eq. 15 can be rewritten in the matrix form as:

$$T_e = \sum_{i=1}^{n}\frac{1}{2m}\Big\{[y_{\alpha_s}^L - (X_{\alpha_s}^U\beta + (1\bar{y}_{\alpha_s}^L - 1\bar{x}_{\alpha_s}^U\beta))]^T[y_{\alpha_s}^L - (X_{\alpha_s}^U\beta + (1\bar{y}_{\alpha_s}^L - 1\bar{x}_{\alpha_s}^U\beta))] + [y_{\alpha_s}^U - (X_{\alpha_s}^L\beta + (1\bar{y}_{\alpha_s}^U - 1\bar{x}_{\alpha_s}^L\beta))]^T[y_{\alpha_s}^U - (X_{\alpha_s}^L\beta + (1\bar{y}_{\alpha_s}^U - 1\bar{x}_{\alpha_s}^L\beta))]\Big\} \tag{20}$$

For minimizing T_e, the first partial derivative of T_e with respect to β and set $\partial T_e/\partial\beta = 0$, the estimated parameter vector $\hat{\beta}_2$ is obtained as:

$$\hat{\beta}_2 = \left[\sum_{s=0}^{m-1}\{[X_{\alpha_s}^L - 1\bar{x}_{\alpha_s}^L]^T[X_{\alpha_s}^L - 1\bar{x}_{\alpha_s}^L] + [X_{\alpha_s}^U - 1\bar{x}_{\alpha_s}^U]^T[X_{\alpha_s}^U - 1\bar{x}_{\alpha_s}^U]\}\right]^{-1} \times\sum_{s=0}^{m-1}\{[(X_{\alpha_s}^U)^T y_{\alpha_s}^L - n(\bar{x}_{\alpha_s}^U)^T\bar{y}_{\alpha_s}^L] + [(X_{\alpha_s}^L)^T y_{\alpha_s}^U - n(\bar{x}_{\alpha_s}^L)^T\bar{y}_{\alpha_s}^U]\} \tag{21}$$

Case 3: The estimated parameters are mixed (some of positives and negatives) ($b_1,\ldots, b_k \geq 0$ and $b_{k+1},\ldots, b_p < 0$ where $1 \leq k < p$), then total error T_e in Eq. 15 can be rewritten in the matrix form as:

$$T_e = \sum_{i=1}^{n}\frac{1}{2m}\Big\{[y_{\alpha_s}^L - (X_{\alpha_s}^*\beta + (1\bar{y}_{\alpha_s}^L - 1\bar{x}_{\alpha_s}^*\beta))]^T[y_{\alpha_s}^L - (X_{\alpha_s}^*\beta + (1\bar{y}_{\alpha_s}^L - 1\bar{x}_{\alpha_s}^*\beta))] + [y_{\alpha_s}^U - (X_{\alpha_s}^{**}\beta + (1\bar{y}_{\alpha_s}^U - 1\bar{x}_{\alpha_s}^{**}\beta))]^T[y_{\alpha_s}^U - (X_{\alpha_s}^{**}\beta + (1\bar{y}_{\alpha_s}^U - 1\bar{x}_{\alpha_s}^{**}\beta))]\Big\} \tag{22}$$

To minimize T_e, the first partial derivative of T_e with respect to β and set $\partial T_e/\partial\beta = 0$, the estimated parameter vector $\hat{\beta}_3$ is obtained as:

$$\hat{\beta}_3 = \left[\sum_{s=0}^{m-1}\{[X_{\alpha_s}^* - 1\bar{x}_{\alpha_s}^*]^T[X_{\alpha_s}^* - 1\bar{x}_{\alpha_s}^*] + [X_{\alpha_s}^{**} - 1\bar{x}_{\alpha_s}^{**}]^T[X_{\alpha_s}^{**} - 1\bar{x}_{\alpha_s}^{**}]\}\right]^{-1} \times\sum_{s=0}^{m-1}\{[(X_{\alpha_s}^*)^T y_{\alpha_s}^L - n(\bar{x}_{\alpha_s}^*)^T\bar{y}_{\alpha_s}^L] + [(X_{\alpha_s}^{**})^T y_{\alpha_s}^U - n(\bar{x}_{\alpha_s}^{**})^T\bar{y}_{\alpha_s}^U]\}. \tag{23}$$

ESTIMATION OF FUZZY ADJUSTMENT TERM PARAMETERS

The total error in Eq. 15 can be minimized by the first partial differential of the total error with respect to each parameter of the trapezoidal fuzzy adjustment term $\tilde{\delta}$, that is:

$$\frac{\partial T_e}{\partial\delta_1}, \frac{\partial T_e}{\partial\delta_{m_1}}, \frac{\partial T_e}{\partial\delta_{m_2}}$$

and $\partial T_e/\partial\delta_u$, respectively. The results are summarized as follows:

$$\begin{aligned}
\frac{\partial T_e}{\partial\delta_1} &= \sum_{i=1}^{n}\frac{1}{2m}\sum_{s=0}^{m-1}2\left[\left((\hat{y}_i^*)_{\alpha_s}^L - (y_i)_{\alpha_s}^L\right)\left(1-\frac{s}{m-1}\right)\right],\\
\frac{\partial T_e}{\partial\delta_{m_1}} &= \sum_{i=1}^{n}\frac{1}{2m}\sum_{s=0}^{m-1}2\left[\left((\hat{y}_i^*)_{\alpha_s}^L - (y_i)_{\alpha_s}^L\right)\left(\frac{s}{m-1}\right)\right],\\
\frac{\partial T_e}{\partial\delta_{m_2}} &= \sum_{i=1}^{n}\frac{1}{2m}\sum_{s=0}^{m-1}2\left[\left((\hat{y}_i^*)_{\alpha_s}^U - (y_i)_{\alpha_s}^U\right)\left(\frac{s}{m-1}\right)\right],\\
\frac{\partial T_e}{\partial\delta_u} &= \sum_{i=1}^{n}\frac{1}{2m}\sum_{s=0}^{m-1}2\left[\left((\hat{y}_i^*)_{\alpha_s}^U - (y_i)_{\alpha_s}^U\right)\left(1-\frac{s}{m-1}\right)\right]
\end{aligned} \tag{24}$$

Similarly, we have three cases to estimate parameter of trapezoidal fuzzy adjustment term as following:

Case 1: The estimated parameters are all of positives ($b_j \geq 0$, j = 1,…,p).

To obtain the estimated parameter of trapezoidal fuzzy adjustment term, the first partial differential of the total error T_e with respect to δ_l, δ_{m_1}, δ_{m_2} and δ_u, respectively should be set to zero and solved. The simplified system equations are summarized as follows:

$$\begin{bmatrix} 2m-1 & m-2 & 0 & 0\\ m-2 & 2m-1 & 0 & 0\\ 0 & 0 & 2m-1 & m-2\\ 0 & 0 & m-2 & 1m-1 \end{bmatrix}\begin{bmatrix}\delta_1\\ \delta_{m1}\\ \delta_{m2}\\ \delta_u\end{bmatrix} = \begin{bmatrix} \sum_{s=0}^{m-1}\frac{6}{m}(m-1-s)\left((\bar{y})_{\alpha_s}^L - \sum_{j=1}^{p}b_j\left(\bar{x}_j\right)_{\alpha_s}^L\right) - 3(m-1)b_0\\ \sum_{s=0}^{m-1}\frac{6S}{m}\left((\bar{y})_{\alpha_s}^L - \sum_{j=1}^{p}b_j\left(\bar{x}_j\right)_{\alpha_s}^L\right) - 3(m-1)b_0\\ \sum_{s=0}^{m-1}\frac{6S}{m}\left((\bar{y})_{\alpha_s}^U - \sum_{j=1}^{p}b_j\left(\bar{x}_j\right)_{\alpha_s}^U\right) - 3(m-1)b_0\\ \sum_{s=0}^{m-1}\frac{6}{m}(m-1-s)\left((\bar{y})_{\alpha_s}^U - \sum_{j=1}^{p}b_j\left(\bar{x}_j\right)_{\alpha_s}^U\right) - 3(m-1)b \end{bmatrix} \tag{25}$$

Case 2: The estimated parameters are all of negatives ($b_j < 0$, j = 1,…,p).

By applying the same manner as Case 1, the simplified system equations are summarized as follows:

$$\begin{bmatrix} 2m-1 & m-2 & 0 & 0 \\ m-2 & 2m-1 & 0 & 0 \\ 0 & 0 & 2m-1 & m-2 \\ 0 & 0 & m-2 & 1m-1 \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_{m_1} \\ \delta_{m2} \\ \delta_u \end{bmatrix} = \begin{bmatrix} \sum_{s=0}^{m-1} \frac{6}{m}(m-1-s)\left((\bar{y})_{\alpha_s}^L - \sum_{j=1}^{p} b_j\left(\bar{x}_j\right)_{\alpha_s}^U\right) - 3(m-1)b_0 \\ \sum_{s=0}^{m-1} \frac{6S}{m}\left((\bar{y})_{\alpha_s}^L - \sum_{j=1}^{p} b_j\left(\bar{x}_j\right)_{\alpha_s}^U\right) - 3(m-1)b_0 \\ \sum_{s=0}^{m-1} \frac{6S}{m}\left((\bar{y})_{\alpha_s}^U - \sum_{j=1}^{p} b_j\left(\bar{x}_j\right)_{\alpha_s}^L\right) - 3(m-1)b_0 \\ \sum_{s=0}^{m-1} \frac{6}{m}(m-1-s)\left((\bar{y})_{\alpha_s}^U - \sum_{j=1}^{p} b_j\left(\bar{x}_j\right)_{\alpha_s}^L\right) - 3(m-1)b \end{bmatrix} \quad (26)$$

Case 3: The estimated parameters are mixed (some of positives and negatives) ($b_1,\ldots, b_k \geq$ and $b_{k+1},\ldots b_p<0$ where, $1 \leq k<p$).

By applying the same manner as Case 1, the simplified system equations are summarized as follows:

$$\begin{bmatrix} 2m-1 & m-2 & 0 & 0 \\ m-2 & 2m-1 & 0 & 0 \\ 0 & 0 & 2m-1 & m-2 \\ 0 & 0 & m-2 & 2m-1 \end{bmatrix} \begin{bmatrix} \delta_\ell \\ \delta_{m_1} \\ \delta_{m_2} \\ \delta_u \end{bmatrix} = \begin{bmatrix} \sum_{s=0}^{m-1} \frac{6}{m}(m-1-s)\left((\bar{y})_{\alpha_s}^L - \left(\sum_{j=1}^{k} b_j\left(\bar{x}_j\right)_{\alpha_s}^L + \sum_{j=k+1}^{p} b_j\left(\bar{x}_j\right)_{\alpha_s}^U\right)\right) - 3(m-1)b_0 \\ \sum_{s=0}^{m-1} \frac{6s}{m}\left((\bar{y})_{\alpha_s}^L - \left(\sum_{j=1}^{k} b_j\left(\bar{x}_j\right)_{\alpha_s}^L + \sum_{j=k+1}^{p} b_j\left(\bar{x}_j\right)_{\alpha_s}^U\right)\right) - 3(m-1)b_0 \\ \sum_{s=0}^{m-1} \frac{6s}{m}\left((\bar{y})_{\alpha_s}^U - \left(\sum_{j=1}^{k} b_j\left(\bar{x}_j\right)_{\alpha_s}^U + \left(\sum_{j=1}^{k} b_j\left(\bar{x}_j\right)_{\alpha_s}^U + \sum_{j=k+1}^{p} b_j\left(\bar{x}_j\right)_{\alpha_s}^L\right)\right)\right) - 3(m-1)b_0 \\ \sum_{s=0}^{m-1} \frac{6}{m}(m-1-s)\left((\bar{y})_{\alpha_s}^U - \left(\sum_{j=1}^{k} b_j\left(\bar{x}_j\right)_{\alpha_s}^U + \sum_{j=k+1}^{p} b_j\left(\bar{x}_j\right)_{\alpha_s}^L\right)\right) - 3(m-1)b_0 \end{bmatrix} \quad (27)$$

In practical, the extension of Chen and Hsueh method can be expressed by five main steps as follows:

Step 1: Trapezoidal fuzzy data are defuzzified into crisp data by using the centroid formulae (Wang *et al.*, 2006) which are denoted by:

$$y_i^c = \frac{1}{3}\left[y_{il} + y_{im_1} + y_{im_2} + y_{iu} - \frac{y_{iu}y_{im_2} - y_{il}y_{im_1}}{\left(y_{iu} + y_{im_2}\right) - \left(y_{il} + y_{im_1}\right)}\right] \quad (28)$$

and:

$$x_{ij}^c = \frac{1}{3}\left[x_{ijl} + x_{ijm_1} + x_{ijm_2} + x_{iju} - \frac{x_{iju}x_{ijm_2} - x_{ijl}x_{ijm_1}}{\left(x_{iju} + x_{ijm_2}\right) - \left(x_{ijl} + x_{ijm_1}\right)}\right] \quad (29)$$

For triangular fuzzy data, $y_{im_1} = y_{im_2} = y_m$ and $x_{ijm_1} = x_{ijm_2} = x_{ijm}$, the above formulae become:

$$y_i^c = \frac{1}{3}\left[y_{il} + y_{im} + y_{iu}\right] \quad (30)$$

and:

$$x_{ij}^c = \frac{1}{3}\left[x_{ijl} + x_{ijm} + x_{iju}\right] \quad (31)$$

Step 2: To calculate Pearson correlation coefficients of the crisp response and explanatory data by using the equation:

$$\rho_{x_{ij}^c y_i^c} = \frac{\sum_{i=1}^{n}(x_{ij}^c - \bar{x}_j^c)(y_i^c - \bar{y}^c)}{\sqrt{\sum_{i=1}^{n}(x_{ij}^c - \bar{x}_j^c)^2 \sum_{i=1}^{n}(y_i^c - \bar{y}^c)^2}}, \quad j = 1,\ldots,p \quad (32)$$

where, y^c_i is the crisp response data, x_{ij}^c the jth crisp explanatory data:

$$\bar{y}^c = \frac{1}{n}\sum_{i=1}^{n} y_i^c$$

and:

$$\bar{x}_j^c = \frac{1}{n}\sum_{i=1}^{n} x_{ij}^c$$

Step 3: The model parameters are estimated by using the two estimation methods, one is the OLS method and the other is the α-cuts method. Based on the OLS method, we obtained a vector of estimated parameters, $\hat{\beta}_{LSE} = [\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_p]^T$. Similarly, we obtained a vector of estimated parameters when using m α-cuts $\hat{\beta}_\alpha = [\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_p]^T$

Step 4: To calculate the estimated fuzzy adjustment term and the estimated fuzzy response

Step 5: To calculate the error for the ith sample which is based on m α-cuts (e_i^m) and the total error for n sample (T_e) in the model. For a comparative of the models, whose the value of T_e is smallest implied to the most effectiveness

NUMERICAL EXAMPLES

In this section, two numerical examples are presented to demonstrate the extension of Chen and Hsueh method.

Example 1 (crisp input and fuzzy output model): This example is designed by Tanaka *et al.* (1989). There are five sets of triangular fuzzy response and crisp explanatory

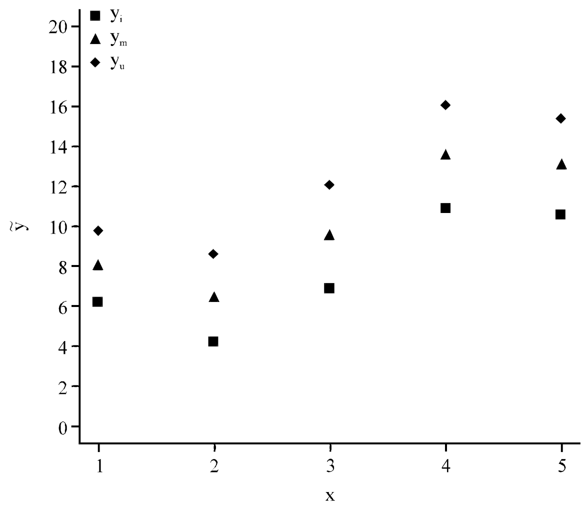

Fig. 1: Scatter plot of triangular fuzzy response data

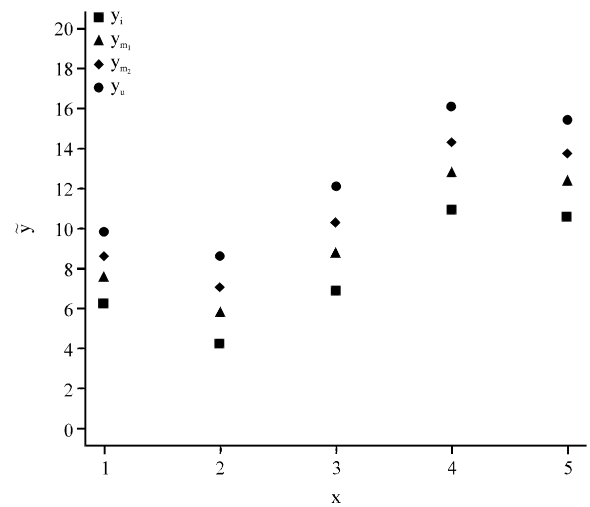

Fig. 2: Scatter plot of trapezoidal fuzzy response data

Table 1: Tanaka *et al.* (1989) data and trapezoidal fuzzy response data with $\alpha_S = 0.7$

Explanatory x_{i1}	Triangular response $\tilde{y}_i = (y_{il}, y_{im}, y_{iu})$	Trapezoidal response $\tilde{y}_i = (y_{il}, y_{im_1}, y_{im_2}, y_{iu})$
1	(6.2,8.0,9.8)	(6.20, 7.46, 8.54, 9.80)
2	(4.2,6.4,8.6)	(4.20, 5.74, 7.06, 8.60)
3	(6.9,9.5,12.1)	(6.90, 8.72, 10.28, 12.10)
4	(10.9,13.5,16.1)	(10.90, 12.72, 14.28, 16.10)
5	(10.6,13.0,15.4)	(10.60, 12.28, 13.72, 15.40)

observations $(\tilde{y}_i, x_{i1})$ as shown in the first two columns of Table 1. The scatter plot of triangular fuzzy response as shown in Fig. 1.

Before we apply the extension of Chen and Hsueh method, we have to transform triangular fuzzy data into trapezoidal fuzzy data with the same support. The α-cut technique is presented as follows:

- Let $(\tilde{y}_i)_{\alpha_s}$ is an α-cut of triangular fuzzy response $\tilde{y}_i$ for a particular α_S level which is denoted by $(\tilde{y}_i)_{\alpha_s} = [y_{il} + \alpha_s(y_{im} - y_{il}), y_{iu} - \alpha_s(y_{iu} - y_{im})]$
- Choosing an α_S level from a set of alpha levels, s = {0.5, 0.6, 0.7, 0.8} Then $(\tilde{y}_i)_{\alpha_s}$ can be calculated by using above formulae
- Assigning the lower bound and the upper bound of $(\tilde{y}_i)_{\alpha_s}$ to the left mode point (y_{im_1}) and the right mode point (y_{im_2}) of trapezoidal fuzzy number, respectively. Thus the trapezoidal fuzzy number can be obtained as $\tilde{y}_i = (y_{il}, y_{im_1}, y_{im_2}, y_{iu})$, where y_{il} and y_{iu} come form the left point and the right point of TFN but y_{im_1} and y_{im_2} come from the α-cuts of $\tilde{y}_i$ with a specified α_S level

This example is used the α-cuts technique with a specified α_S = 0.7, thus the trapezoidal fuzzy responses data are obtained as shown in the last column of Table 1. Also the scatter plot of trapezoidal fuzzy response as shown in Fig. 2.

The extension of Chen and Hsueh method can be summarized as follows:

Step 1: The trapezoidal fuzzy response data is defuzzified to crisp data by using the centroid equation in Eq. 28. Thus, we obtained y^c_1 = 8.0, y^c_2 = 6.4, y^c_3 = 9.5, y^c_4 = 13.5 and y^c_5 = 13.0

Step 2: Pearson correlation coefficient of x_{i1} and y^c_i can be calculated by the formulae in Eq. 32, where:

$$\bar{x}_1 = \frac{1}{5}\sum_{i=1}^{5} x_{i1} = 3.0$$

and:

$$\bar{y}^c_i = \frac{1}{5}\sum_{i=1}^{5} y^c_i = 10.08$$

Thus, we obtained $\rho_{x_{i1}y^c_i} = 0.8723$ means a high relationship between the response and explanatory variables. Moreover, the sign of $\rho_{x_{i1}y^c_i}$ can be implied to the sign of corresponding estimated parameter $\hat{\beta}_1$

Step 3: The two estimation methods are used to estimate model parameters as follows:

- By using the OLS method, then obtained $\hat{\beta}_0 = 4.95$ and $\hat{\beta}_1 = 1.71$

- By using the α-cuts method with m = 2, then obtained $\hat{\beta}_{0\alpha} = 0.00$ and $\hat{\beta}_{1\alpha} = 1.71$

Step 4: Since $\hat{\beta}_{1\alpha} = 1.71$ and positive number, thus the system equation in Eq. 25 is used to calculate the estimated trapezoidal adjustment term, $\hat{\hat{\delta}} = (\hat{\delta}_\ell + \hat{\beta}_0, \hat{\delta}_{m_1} + \hat{\beta}_0, \hat{\delta}_{m_2} + \hat{\beta}_0, \hat{\delta}_u + \hat{\beta}_0)$. Substituting m = 2 into Eq. 25, then the system equation can be reduced to these equation:

$$\hat{\delta}_\ell + \hat{\beta}_0 = (\bar{y})^L_{\alpha_0} - b_1(\bar{x}_j)^L_{\alpha_0},\ \hat{\delta}_{m_1} + \hat{\beta}_0 = (\bar{y})^L_{\alpha_1} - b_1(\bar{x}_j)^L_{\alpha_1},$$
$$\hat{\delta}_{m_2} + \hat{\beta}_0 = (\bar{y})^U_{\alpha_1} - b_1(\bar{x}_j)^U_{\alpha_1},\ \hat{\delta}_u + \hat{\beta}_0 = (\bar{y})^U_{\alpha_0} - b_1(\bar{x}_j)^U_{\alpha_0}$$

Where:

$$\alpha_0 = 0, \alpha_1 = 1, (\bar{y})^L_{\alpha_0} = 38.80, (\bar{y})^L_{\alpha_1} = 46.92, (\bar{y})^U_{\alpha_1} = 53.88$$
$$(\bar{y})^U_{\alpha_0} = 62.00, (\bar{x})^L_{\alpha_0} = (\bar{x})^L_{\alpha_1} = (\bar{x})^U_{\alpha_1} = (\bar{x})^U_{\alpha_0} = 3.00$$

and b_1 is the estimated parameter ($\hat{\beta}_1 = 1.71$ for OLS method or $\hat{\beta}_{1\alpha} = 1.71$ for the α-cuts method).

Substituting those values into above formulae, then the estimated trapezoidal adjustment term can be obtained as $\hat{\hat{\delta}} = (2.630, 4.254, 5.646, 7.270)$ for the both method. Consequently, the estimated fuzzy response for the *i*th sample can be expressed by:

$$\hat{\tilde{y}}_i^* = 1.71x_{i1} + (2.630, 4.254, 5.646, 7.270),\ i = 1, \ldots, n$$

where, the second term on the right hand side as the intercept and fuzzy adjustment term are combined. Substituting x_{ij} into above equation, we obtained:

$$\hat{\tilde{y}}_1^* = (4.340, 5.964, 7.356, 8.980),\ \hat{\tilde{y}}_2^* = (6.050, 7.674, 9.066, 10.690),$$
$$\hat{\tilde{y}}_3^* = (7.76, 9.384, 10.776, 12.40),\ \hat{\tilde{y}}_4^* = (9.470, 11.094, 12.486, 14.110)$$

and:

$$\hat{\tilde{y}}_5^* = (11.180, 12.804, 14.196, 15.820)$$

Step 5: To calculate the estimated error for the *i*th sample which is based on m α-cuts, e_i^m in Eq. 14 and the total estimated error for n samples, T_e in Eq. 15 as the following

By using m = 2 (α_0 = 0 and α_1 = 1), then obtained:

$$e_1^{m=2} = 1.942968,\ e_2^{m=2} = 3.888748,\ e_3^{m=2} = 0.379128$$
$$e_4^{m=2} = 2.966828,\ e_5^{m=2} = 0.253488$$

and T_e = 9.43116.

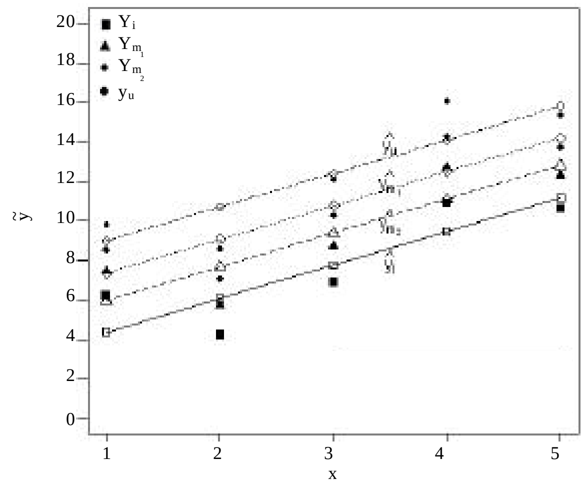

Fig. 3: Predicted values of trapezoidal fuzzy response data

Table 2: Summarized results of the extension of Chen and Hsueh method for Example 1

Sample i	The estimated response $\hat{\tilde{y}}_i^* = (\hat{y}_{il}^*, \hat{y}_{im_1}^*, \hat{y}_{im_2}^*, \hat{y}_{iu}^*)$	The estimated error ($e_i^{m=2}$) OLS	α-cuts
1	(4.340, 5.964, 7.356, 8.980)	1.942968	1.942968
2	(6.050, 7.674, 9.066, 10.690)	3.888748	3.888748
3	(7.760, 9.384, 10.776, 12.400)	0.379128	0.379128
4	(9.470, 11.094, 12.486, 14.110)	2.966828	2.966828
5	(11.180, 12.804, 14.196, 15.820)	0.253488	0.253488
	Total error, T_e	9.431160	9.431160

Table 3: A comparison of the estimated errors of the extension of Chen and Hsueh method

Sample i	Chen and Hsueh method E_i, m = 2	Extended Chen and Hsueh method e^m_i, m = 2
1	1.931	1.9308
2	3.888	3.8881
3	0.376	0.3756
4	2.963	2.9633
5	0.253	0.2532
Total error, T_e	9.411	9.4110

The summarized results of the extension of Chen and Hsueh method is shown in Table 2.

Also the predicted values of trapezoidal fuzzy response data are shown in Fig. 3.

Based on the results, the OLS method and the α-cuts method are given the same total error, T_e = 9.43116. However, the two methods can be given the different of the total error. Further details see in Example 2.

For comparing the estimated error of the extension of Chen and Hsueh method with Chen and Hsueh method, we used the triangular fuzzy response (α_S = 1) and triangular adjustment term to formulate the model. The comparative of the results of the extension of Chen and Hsueh method and Chen and Hsueh method is shown in Table 3.

Table 4: Fuzzy input-output data by Sakawa and Yano (1992)

Observation i	Fuzzy input $\tilde{x}_{i1} = (x_{i1l}, x_{i1m}, x_{i1u})$	Fuzzy output $\tilde{y}_i = (y_{il}, y_{im}, y_{iu})$
1	(1.5, 2.0, 2.5)	(3.5, 4.0, 4.5)
2	(3.0, 3.5, 4.0)	(5.0, 5.5, 6.0)
3	(4.5, 5.5, 6.5)	(6.5, 7.5, 8.5)
4	(6.5, 7.0, 7.5)	(6.0, 6.5, 7.0)
5	(8.0, 8.5, 9.0)	(8.0, 8.5, 9.0)
6	(9.5, 10.5, 11.5)	(7.0, 8.0, 9.0)
7	(10.5, 11.0, 11.5)	(10.0, 10.5, 11.0)
8	(12.0, 12.5, 13.0)	(9.0, 9.5, 10.0)

Table 5: Estimated fuzzy output of the extension of Chen and Hseuh method for Example 2

	Estimated fuzzy output $(\hat{\tilde{y}}_i^*)$	
Observation i	OLS method	α-cuts method
1	4.05, 4.44, 4.78, 5.17	4.05, 4.44, 4.77, 5.17
2	4.83, 5.22, 5.56, 5.95	4.83, 5.22, 5.55, 5.95
3	5.61, 6.18, 6.67, 7.25	5.61, 6.18, 6.67, 7.25
4	6.65, 7.04, 7.38, 7.77	6.65, 7.04, 7.38, 7.77
5	7.43, 7.82, 8.15, 8.55	7.43, 7.82, 8.16, 8.55
6	8.21, 8.78, 9.27, 9.85	8.21, 8.78, 9.27, 9.85
7	8.73, 9.12, 9.45, 9.85	8.73, 9.12, 9.46, 9.85
8	9.50, 9.90, 10.23, 10.62	9.51, 9.90, 10.24, 10.63

Table 6: Estimated errors of the extension of Chen and Hsueh method for Example 2

	Estimated error $(e_i^{m=2})$	
Observation i	OLS method	α-cuts method
1	0.37544584	0.36864843
2	0.01403429	0.01493805
3	1.16507791	1.16944087
4	0.50304298	0.50223622
5	0.26525086	0.26427841
6	1.06952781	1.07550705
7	1.47756684	1.46919271
8	0.32037797	0.32598421
Total error	5.19032500	5.19022600

As the results in Table 3, the estimated errors and total error of the both methods are the same. That is, $E_i = e_i^{m=2}$ for all i and $T_e = 9.411$ It has shown the extension of Chen and Hsueh method has an efficiency equal to Chen and Hsueh method in the case of triangular fuzzy data. This is clearly show that the extension of Chen and Hsueh method more generalized.

Example 2 (fuzzy input and fuzzy output model): This example is designed by Sakawa and Yano (1992). There are eight sets of the fuzzy observations $(\tilde{x}_{i1}, \tilde{y}_i)$ as listed in Table 4.

By using the same manner as in Example 1, the estimated fuzzy output of the extension of Chen and Hseuh method are summarized in Table 5. Also the estimated errors of the two estimation methods are shown in Table 6.

As the results in Table 6, the estimated error for the ith sample $(e_i^{m=2})$ in the both estimation methods are unequal. Similarly, the total error of the OLS method is 5.190325 while the total error of the α-cuts method is 5.190226 .

CONCLUSION

In this study, we extended Chen and Hsueh method to trapezoidal fuzzy numbers case. The OLS method and the α-cuts method are presented for estimating model parameters. Based on the α-cuts method, the estimated parameters $(\hat{\beta}_\alpha)$ and the estimated adjustment term $(\hat{\tilde{\delta}})$ were derived. Two numerical examples are demonstrated the extension of Chen and Hsueh method. Based on the results of this study, some of the useful conclusions are given as follows:

- Chen and Hsueh method can not be used to formulate the fuzzy regression model when the fuzzy components are trapezoidal fuzzy numbers while the extension of Chen and Hsueh method can used in the case of triangular fuzzy numbers. That is, the extended method more generalized
- The α-cuts technique can be applied to transform the mode point of triangular fuzzy numbers into the two mode points of trapezoidal fuzzy numbers as shown in Example 1. The result of this method is a closed interval that you can assigned the lower bound and the upper bound of the closed interval to the left mode point and the right mode point of trapezoidal fuzzy numbers, respectively. For the support of triangular fuzzy numbers, we used it as the support of trapezoidal fuzzy numbers. This approach is an alternative method for the researchers
- For the models that the explanatory variables are crisp numbers, the total estimated errors (T_e) of the two estimation method (the OLS method and the α-cuts method) are equal as shown in Example 1. By contrast, the models that the explanatory variables are fuzzy numbers, T_e of the two estimation method are unequal as shown in Example 2. The advantage of the OLS method is well-known and available in the statistics packages for solving the estimated parameters $(\hat{\beta})$. In the α-cuts method, it is more complex computation and non-existence in the statistical packages. We have to create the function for solving the estimated parameters $(\hat{\beta}_\alpha)$
- For evaluating the efficiency of the models, this study used the estimated error for the ith sample which is based on m α-cuts (e_i^m) and the total error for n sample (T_e). Model which is given the minimum total error (T_e) implied to the best model or in favor model

ACKNOWLEDGMENTS

We are grateful to Rajamangala University of Technology Thanyaburi (RMUTT), Pathum Thani, Thailand, for financial support through the RMUTT scholarship. We also wish to thank the anonymous reviewers for their valuable suggestions.

REFERENCES

Buckley, J.J., 2005. Fuzzy statistics: Regression and prediction. Soft Comput., 9: 769-775.

Chen, L.H. and C.C. Hsueh, 2007. A mathematical programming method for formulating a fuzzy regression model based on distance criterion. IEEE Trans. Syst. Man Cybern. Part B: Cybern., 137: 705-712.

Chen, L.H. and C.C. Hsueh, 2009. Fuzzy regression models using the least-squares method based on the concept of distance. IEEE Trans. Fuzzy Syst., 17: 1259-1272.

Diamond, P., 1988. Fuzzy least squares. Inform. Sci., 46: 141-157.

Kao, C. and C.L. Chyu, 2002. A fuzzy linear regression model with better explanatory power. Fuzzy Sets Syst., 126: 401-409.

Kao, C. and C.L. Chyu, 2003. Least-squares estimates in fuzzy regression analysis. Eur. J. Oper. Res., 148: 426-435.

Kim, B. and R.R. Bishu, 1998. Evaluation of fuzzy linear regression models by comparing membership functions. Fuzzy Sets Syst., 100: 343-352.

Kim, K.J., D.H. Kim and S.H. Choi, 2005. Least absolute deviation estimator in fuzzy regression. J. Appl. Math. Comput., 18: 649-656.

Myers, R.H. and J.S. Milton, 1991. A First Course in the Theory of Linear Statistical Models. PWS-KENT Publishing Company, Boston, USA., ISBN-13: 978-0534916459, Pages: 342.

Redden, D.T. and W.H. Woodall, 1994. Properties of certain fuzzy linear regression methods. Fuzzy Sets Syst., 64: 361-375.

Sakawa, M. and H. Yano, 1992. Multiobjective fuzzy linear regression analysis for fuzzy input-output data. Fuzzy Sets Syst., 47: 173-181.

Tanaka, H. and H. Lee, 1998. Interval regression analysis by quadratic programming. IEEE Trans. Fuzzy Syst., 6: 473-481.

Tanaka, H. and J. Watada, 1988. Possibilistic linear systems and their application to the linear regression model. Fuzzy Sets Syst., 27: 275-289.

Tanaka, H., 1987. Fuzzy data analysis by possibilistic linear models. Fuzzy Sets Syst., 24: 363-375.

Tanaka, H., I. Hayashi and J. Watada, 1989. Possibilistic linear regression analysis for fuzzy Eur. J. Oper. Res., 40: 389-396.

Tanaka, H., S. Uejima and K. Asis, 1982. Linear regression analysis with fuzzy model. IEEE Trans. Syst. Man Cybernet., 12: 903-907.

Wang, Y.M., J.B. Yang, D.L. Xu and K.S. Chin, 2006. On the centroids of fuzzy numbers. Fuzzy Sets Syst., 157: 919-926.