Composed Fuzzy Measure of Maximized L-Measure and Delta-Measure

Hsiang-Chuan Liu, Hsien-Chang Tsai, Yu-Du Jheng, and Tung-Sheng Liu

Abstract—The well known fuzzy measures, λ-measure and P-measure, have only one formulaic solution. Two multivalent fuzzy measures with infinitely many solutions, L-measure and δ-measure, were proposed by our previous works, but the former do not include the additive measure as the latter and the latter has not so many measure solutions as the former, therefore, a composed fuzzy measure of above two measures, called L_δ -measure was proposed by our additional previous work. However, all of abovementioned fuzzy measures do not contain the largest measure, B-measure, which all not completed measures. In this paper, an improved completed fuzzy measure composed of maximized L-measure and δ-measure, denoted $L_{m\delta}$ -measure, is proposed. For evaluating the Choquet integral regression models with our proposed fuzzy measure and other different ones, two real data experiments by using a 5-fold cross-validation mean square error (MSE) were conducted. The performances of Choquet integral regression models with fuzzy measure based $L_{m\delta}$ -measure, $L_{m\delta}$ -measure, L_δ -measure, L-measure, δ-measure, λ-measure, and P-measure, respectively, a ridge regression model, and a multiple linear regression model are compared. Both of two experimental results show that the Choquet integral regression models with respect to our new measure based on γ-support outperforms others forecasting models.

***Keywords*—**Lambda-measure, P-measure, Delta-measure, L-measure, composed fuzzy measure.

I. Introduction

When there are interactions among independent variables, traditional multiple linear regression models do not perform well enough. The traditional improved methods exploited ridge regression models [1]. Recently, some Choquet integral regression models with respect to different fuzzy measures were used [2-5, 7-15] to improve this situation. The well-known fuzzy measures, λ-measure [2-4] and P-measure [5] have only one formulaic solution of fuzzy measure, the former is not a closed form, and the latter is not sensitive enough. Two multivalent fuzzy measures with infinitely many solutions were proposed by our previous works, called L-measure [8-11] and δ-measure [10,11], but L-measure do not include the additive measure and δ-measure has not so many measure solutions as L-measure. Due to the above drawbacks, an improved fuzzy measure composed of above two multivalent fuzzy measures, denoted L_δ -measure, was proposed by our additional previous work. This improved multivalent fuzzy measure is not only including the additive measure, but also having the same infinitely many measure solutions as L-measure. However, all of above mentioned fuzzy measures do not contain the largest measure, B-measure, which all not completed measures. In this paper, an improved completed fuzzy measure composed of maximized L-measure and δ-measure, denoted $L_{m\delta}$ -measure, is proposed. This new measure not only contains the additive measure and B-measure but also have more fuzzy measure solutions than L_δ -measure. For evaluating the Choquet integral regression models with our proposed fuzzy measure and other different ones, two real data experiments by using a 5-fold cross-validation mean square error (MSE) are conducted. The performances of Choquet integral regression models with fuzzy measure based on $L_{m\delta}$ -measure, L_m -measure, L_δ -measure, L-measure, δ-measure, λ-measure, and P-measure, respectively, a ridge regression model, and a multiple linear regression model are compared.

This paper is organized as follows: The multiple linear regression and ridge regression [1] are introduced in section II; two well known fuzzy measure, λ-measure [2] and P-measure [5], L -measure, δ -measure and L_δ -meassure are introduced in section III; B-measure, completed fuzzy measure and our new measure, $L_{m\delta}$ -measure, are introduced in section IV; the fuzzy support, γ-support [7] is described in section V; the Choquet integral regression model [6-8] based on fuzzy measures are described in section VI; two experiments and results are described in section VII; and final section is for conclusions and future works.

Manuscript received Dec. 11, 2009: Revised version received Jen. 26, 2010. This work was supported in part by the National Science Council grant of Taiwan Government (NSC 98-2410-H-468-014).

H.-C. Liu, is with the Department of Bioinformatics, Asia University, Taichung, 41345, Taiwan (corresponding author: phone: 886-4-23384072, Fax: 886-4-23383640, e-mail: lhc@asia.edu.tw).
H.-C. Tsai, is with Department of Biology, National Changhua University of Education Changhua, 50001, Taiwan.(bihft@cc.ncue.edu.tw).
Y.-D. Jheng, is with Department of Education, National Taichung University, Taichung, 40306 Taiwan. (e-mail: doora0622@yahoo.com.tw).
T.-S. Liu, is with Forensic Science Center, Military Police Command, 104, Taiwan, R.O.C. (e-mail: tommy0366@yahoo.com.tw).

II. THE MULTIPLE LINEAR REGRESSION, RIDGE REGRESSION

Let $\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$, $\underline{\varepsilon} \sim \mathrm{N}\left(\underline{0}, \sigma^2 I_n\right)$ be a multiple linear model, $\hat{\underline{\beta}} = \left(X'X\right)^{-1} X'Y$ be the estimated regression coefficient vector, and $\hat{\underline{\beta}}_k = \left(X'X + kI_n\right)^{-1} X'Y$ be the estimated ridge regression coefficient vector, Hoerl, Kenard and Baldwin [1] suggested

$$\hat{k} = \frac{n\hat{\sigma}^2}{\hat{\underline{\beta}}'\hat{\underline{\beta}}} \tag{1}$$

III. FUZZY MEASURES

The two well known fuzzy measures, the λ-measure proposed by Sugeno in 1974, and P-measure proposed by Zadah in 1978, are concisely introduced as follows.

A. *Axioms of Fuzzy Measures*

Definition 1 fuzzy measure [2-4]
A fuzzy measure μ on a finite set X is a set function $\mu : 2^X \rightarrow [0,1]$ satisfying the following axioms:

1) $\mu(\phi) = 0$, $\mu(X) = 1$ (boundary conditions) (2)
2) $A \subseteq B \Rightarrow \mu(A) \leq \mu(B)$ (monotonicity) (3)

B. *Fuzzy density function [2-7]*

A fuzzy density function of a fuzzy measure μ on a finite set X is a function $s : X \rightarrow [0,1]$ satisfying:

$$s(x) = \mu(\{x\}), x \in X \tag{4}$$

$s(x)$ is called the fuzzy density of singleton x.

C. *λ-measure*

Definition 3 λ-measure [3]
For a given fuzzy density function s, λ-measure, g_λ, is a fuzzy measure on a finite set X, satisfying:

$$\begin{aligned} &A, B \in 2^X, A \cap B = \phi, A \cup B \neq X \\ &\Rightarrow g_\lambda(A \cup B) \\ &= g_\lambda(A) + g_\lambda(B) + \lambda g_\lambda(A) g_\lambda(B) \end{aligned} \tag{5}$$

$$\prod_{i=1}^{n}\left[1 + \lambda s(x_i)\right] = \lambda + 1 > 0, s(x_i) = g_\lambda(\{x_i\}) \tag{6}$$

Where the real number, λ, is also called the determine coefficient of λ-measure.

Note that once the fuzzy density function is known, we can obtain the values of λ uniquely by using the previous polynomial equation. In other words, λ-measure has a unique solution without closed form. Moreover, for given singleton measures s,

If $\sum_{x \in X} s(x) = 1$ *then* $g_\lambda(A) = \sum_{x \in A} s(x)$, in other word, if $\sum_{x \in X} s(x) = 1$ then λ-measure is just the additive measure

D. *P-measure*

Definition 4. P-measure [5]
For given a fuzzy density function s, P-measure, g_P, is a fuzzy measure on a finite set X, satisfying:

$$\begin{aligned} &\forall A \in 2^X \\ &\Rightarrow g_P(A) = \max_{x \in A}\{s(x)\} = \max_{x \in A}\{g_P(\{x\})\} \end{aligned} \tag{7}$$

Note that for any subset of X, A, P-measure considers only the maximum value and will lead to insensitivity.

E. *L-measure*

Definition 5. L-measure [8,11]
For given a fuzzy density function $s(x)$, L-measure, g_L, is a fuzzy measure on a finite set X, $|X| = n$, satisfying:

1) $L \in [0, \infty)$ (8)

2) $\forall A \subset X, n - |A| + (|A| - 1)L > 0 \Rightarrow$

$$g_L(A) = \max_{x \in A}\left[s(x)\right] + \frac{(|A|-1)L\sum_{x \in A} s(x)\left[1 - \max_{x \in A}\left[s(x)\right]\right]}{\left[n - |A| + (|A|-1)L\right]\sum_{x \in X} s(x)} \tag{9}$$

Where the real number, L, is also called the determine coefficient of L-measure.

Note that L-measure satisfies the following properties
(i) For each $L \in [0, \infty)$, L-measure is a fuzzy measure, in other words, L-measure has infinitely many solutions of fuzzy measures, for each $L \in [0, \infty)$.
(ii) $L \in [0, \infty)$, L-measure is an increasing function on real number L.
(iii) if $L = 0$ then L-measure is just the P-measure

F. *δ-measure*

Definition 6 δ-measure [9-10]
For given fuzzy density function $s(x)$, a δ-measure, g_δ, is a fuzzy measure on a finite set X, $|X| = n$, satisfying:

1) $\delta \in [-1, 1]$, $\sum_{x \in X} s(x) = 1$ (10)

2) $g_\delta(\phi) = 0, g_\delta(X) = 1$ (11)

3) $\forall A \subset X, A \neq X \Rightarrow$

$$g_\delta(A)=\left[1+\delta\max_{x\in A}s(x)\right]\frac{(1+\delta)\sum_{x\in A}s(x)}{1+\delta\sum_{x\in A}s(x)}-\delta\max_{x\in A}s(x) \quad (12)$$

Note that δ -measure satisfies the following properties

(i) For each $\delta\in[-1,1]$, δ -measure is a fuzzy measure, i.e., δ -measure has infinite many fuzzy measures with determine coefficient δ , $\delta\in[-1,1]$.

(ii) $\delta\in[-1,1]$, δ -measure is an increasing function on δ ,

(iii) if $\delta=-1$ then δ -measure is just P-measure

(iv) if $\delta=0$ then δ -measure is just additive measure

(v) if $-1\le\delta<0$ then δ -measure is a sub- additive measure,

(vi) if $0<\delta\le 1$ then δ -measure is a supper- additive measure,

G. L_δ -measure

Definition 7 L_δ -measure [11]

For given fuzzy density function $s(x)$, the composed measure of L-measure and δ-measure , denoted L_δ -measure, g_{L_δ} , is a fuzzy measure on a finite set X, $|X|=n$, satisfying:

1) $L\in[-1,\infty),\ \sum_{x\in X}s(x)=1$ (13)

2) $g_{L_\delta}(\phi)=0, g_{L_\delta}(X)=1$ (14)

3) $\forall A\subset X\Rightarrow$

$$g_L(A)=\begin{cases}\max_{x\in A}s(x) & \text{if } L=-1\\ \dfrac{(1+L)\sum_{x\in A}s(x)\left[1+L\max_{x\in A}s(x)\right]}{1+L\sum_{x\in A}s(x)}-L\max_{x\in A}s(x) & \text{if } L\in(-1,0]\\ \dfrac{L(|A|-1)\sum_{x\in A}s(x)\left[1-\sum_{x\in A}s(x)\right]}{\left[n-|A|+L(|A|-1)\right]\sum_{x\in X}s(x)}+\sum_{x\in A}s(x) & \text{if } L\in(0,\infty)\end{cases} \quad (15)$$

Note that L_δ -measure satisfies the following properties

(i) $L\in[-1,\infty)$, L_δ -measure is a fuzzy measure family

(ii) $L\in[-1,\infty)$, L_δ -measure is an increasing function on L

(iii) if $L=-1$ then L_δ -measure is just the P-measure

(iv) if $L=0$ then L_δ -measure is just the additive measure

(v) if $-1<L<0$ then L_δ -measure is a sub-additive measure

(vi) if $0<L<\infty$ then L_δ -measure is a supper-additive measure

(vii) If $\sum_{x\in X}s(x)=1$ and $L=0$ then L_δ -measure is just the λ -measure

IV. Composed completed fuzzy measure of maximized L-measure and Delta-measure

A. B-measure

Definition 8 B- measure [8]

For any given fuzzy density function, $s(x)$, on a finite set, X, a B-measure is a set function, $g_B:2^X\to[0,1]$, satisfying:

$$g_B(A)=\begin{cases}0 & A=\phi\\ s(x) & A=\{x\},x\in X\\ 1 & |A|>1, A\subset X\end{cases} \quad (16)$$

Note that for any given fuzzy density function, $s(x)$, on a finite set, X, B-measure is not smaller than any fuzzy measure, that is to say, B-measure is the largest one.

B. Completed measure

Definition 9 Completed measure [8]

For any given fuzzy density function, $s(x)$, on a finite set, X, a multivalent normalized monotone measure, μ -measure, with determined coefficient, μ , is called a completed measure , if it satisfies following conditions

1) μ -measure is a monotone increasing function of its determined coefficient, μ
2) if μ =0 then μ -measure is just the P-measure
3) If the upper limit fuzzy measure of μ - measure is just the B-measure.

Note that all of abovementioned fuzzy measures are not completed fuzzy measures

C. Maximized L- measure, L_m -measure

Definition 9 L_m -measure

For given fuzzy density function $s(x)$, the composed measure of maximized L-measure and δ-measure , denoted $L_{M\delta}$ -measure, $g_{L_{M\delta}}$, is a fuzzy measure on a finite set X, $|X|=n$, satisfying:

1) $L\in[-1,\infty),\ \sum_{x\in X}s(x)=1$ (16)

2) $g_{L_{M\delta}}(\phi)=0, g_{L_{M\delta}}(X)=1$ (17)

3) $\forall A\subset X, 0<|A|<n\Rightarrow$

$$g_{L_m}(A)=\frac{L(|A|-1)\sum_{x\in A}s(x)\left[1-\sum_{x\in A}s(x)\right]}{(n-|A|)\sum_{x\in X-A}s(x)+L(|A|-1)\sum_{x\in A}s(x)}+\sum_{x\in A}s(x) \quad (18)$$

Note that L_m -measure satisfies the following properties

(i) $L\in[0,\infty)$, L_δ -measure is a fuzzy measure family

(ii) $L\in[0,\infty)$, L_δ -measure is an increasing function of L

(iii) if $L=0$ then L_δ -measure is just the P-measure

(iv) if $L=0$ then L_δ -measure is just the additive measure

(v) L_δ -measure is a completed measure

(vi) L_δ -measure does not contain the additive measure.

D. Composed measure of Maximized L- measure and δ -measure, $L_{M\delta}$ -measure

Definition 9 $L_{M\delta}$ -measure

For given fuzzy density function $s(x)$, the composed measure of maximized L-measure and δ-measure , denoted $L_{M\delta}$ -measure, $g_{L_{M\delta}}$, is a fuzzy measure on a finite set X, $|X| = n$, satisfying:

1) $L \in [-1,\infty), \sum_{x\in X} s(x) = 1$ (19)

2) $g_{L_{M\delta}}(\phi) = 0, g_{L_{M\delta}}(X) = 1$ (20)

3) $\forall A \subset X \Rightarrow$

$$g_{L_{M\delta}}(A)=\begin{cases}\max_{x\in A} s(x) & \text{if } L=-1\\ \dfrac{(1+L)\sum_{x\in A}s(x)\left[1+L\max_{x\in A}s(x)\right]}{1+L\sum_{x\in A}s(x)}-L\max_{x\in A}s(x) & \text{if } L\in(-1,0]\\ \dfrac{L(|A|-1)\sum_{x\in A}s(x)\left[1-\sum_{x\in A}s(x)\right]}{(n-|A|)\sum_{x\in X-A}s(x)+L(|A|-1)\sum_{x\in A}s(x)}+\sum_{x\in A}s(x) & \text{if } L\in(0,\infty)\end{cases} \quad (21)$$

E. Important Properties $L_{M\delta}$ -measure

Theorem 1. $L_{M\delta}$ -measure is a fuzzy measure

Proof: if $L \in [-1,1]$, from the property of δ -measure, we know that $L_{M\delta}$ -measure is a fuzzy measure, hereafter, we need only to prove that if $L \in (0,\infty)$, $L_{M\delta}$ -measure is also a fuzzy measure,

The boundary conditions are trivial, now to prove that the monotonicity condition is also satisfied,

(i) Let $\forall A, B \in 2^X, A \subset B$ to prove $g_{L_{M\delta}}(A) \le g_{L_{M\delta}}(B)$

Let $g_{L_{M\delta}}(A) = \dfrac{N(A,L)}{D(A,L)}, g_{L_{M\delta}}(B) = \dfrac{N(B,L)}{D(B,L)}$ (22)

Where

$$N(A,L) = L(|A|-1)\sum_{x\in A}s(x) + (n-|A|)\sum_{x\in A}s(x)\sum_{x\in X-A}s(x)$$
$$D(A,L) = (n-|A|)\sum_{x\in X-A}s(x) + L(|A|-1)\sum_{x\in A}s(x) \quad (23)$$

and

$$N(B,L) = L(|B|-1)\sum_{x\in B}s(x) + (n-|B|)\sum_{x\in B}s(x)\sum_{x\in X-B}s(x)$$
$$D(B,L) = (n-|B|)\sum_{x\in X-B}s(x) + L(|B|-1)\sum_{x\in B}s(x) \quad (24)$$

Then

$$g_{L_M}(B) - g_{L_M}(A) = \frac{N(B,L)}{D(B,L)} - \frac{N(A,L)}{D(A,L)} = \frac{N(B,L)D(A,L) - N(A,L)D(B,L)}{D(B,L)D(A,L)} \quad (25)$$

to prove $g_{L_{M\delta}}(B) - g_{L_{M\delta}}(A) \ge 0$

(ii) Since $|B| \ge |A|, \sum_{x\in B} s(x) \ge \sum_{x\in A} s(x)$ (26)

and $\sum_{x\in X-A} s(x) \ge \sum_{x\in X-B} s(x)$

We can obtain

$$\begin{aligned} &N(B,L)D(A,L) - N(A,L)D(B,L)\\ &= L(|B|-1)(n-|A|)\sum_{x\in B}s(x)\sum_{x\in X-A}s(x)\\ &+L(|A|-1)(n-|B|)\sum_{x\in B}s(x)\sum_{x\in X-B}s(x)\sum_{x\in A}s(x)\\ &-L(|B|-1)(n-|A|)\sum_{x\in A}s(x)\sum_{x\in X-A}s(x)\sum_{x\in B}s(x)\\ &-L(|A|-1)(n-|B|)\sum_{x\in A}s(x)\sum_{x\in X-B}s(x)\\ &= L(|B|-1)(n-|A|)\sum_{x\in B}s(x)\left[\sum_{x\in X-A}s(x)\right]^2\\ &-L(|A|-1)(n-|B|)\sum_{x\in A}s(x)\left[\sum_{x\in X-B}s(x)\right]^2 \ge 0 \end{aligned} \quad (27)$$

(iii) from (i) and (ii), the proof is completed.

Theorem 2 *Important Properties of $L_{M\delta}$ -measure*

(i) $L \in [-1,\infty)$, L_δ -measure is an increasing function on L

(ii) if $L = -1$ then L_δ -measure is just the P-measure

(iii) if $L = 0$ then L_δ -measure is just the additive measure

(iv) if $-1 < L < 0$ then L_δ -measure is a sub-additive measure

(v) if $0 < L < \infty$ then L_δ -measure is a supper-additive measure

(vi) If $\sum_{x\in X} s(x) = 1$ and $L = 0$ then L_δ -measure is just the λ -measure

(vii) $L_{M\delta}$ -measure is a completed fuzzy measure.

Proof:

(i) if $L \in [-1,1]$, from the property of δ -measure, we know that $L_{M\delta}$ -measure is an increasing function of L, need only to prove that if $L \in (0,\infty)$, $L_{M\delta}$ -measure is also an increasing function of L, Let

$$f(L) = g_{L_{M\delta}}(A) = \frac{L(|A|-1)\sum_{x\in A}s(x)\left[1-\sum_{x\in A}s(x)\right]}{(n-|A|)\sum_{x\in X-A}s(x)+L(|A|-1)\sum_{x\in A}s(x)} + \sum_{x\in A}s(x) \quad (28)$$

Then we can obtain

$$f'(L) = \frac{(n-|A|)(|A|-1)\sum_{x\in X-A}s(x)\sum_{x\in A}s(x)\left[1-\sum_{x\in A}s(x)\right]}{\left[(n-|A|)\sum_{x\in X-A}s(x)+L(|A|-1)\sum_{x\in A}s(x)\right]^2} \ge 0 \quad (29)$$

(ii)-(vi) trivial!

(vii) Since $\lim_{L\to\infty} g_{L_{M\delta}}(A) = g_B(A), \forall A \in X$ (30)

Hence $L_{M\delta}$ -measure is a completed fuzzy measure.

V. Γ- SUPPORT

Definition 10: γ -support [7]
For a given fuzzy density function s of a fuzzy measure μ on a finite set X, if $\sum_{x\in X} s(x)=1$, then s is called a fuzzy support measure of μ, or a fuzzy support of μ, or a support of μ. One of fuzzy supports is introduced as below.
Let μ be a fuzzy measure on a finite set $X=\{x_1,x_2,...,x_n\}$, y_i be global response of subject i and $f_i(x_j)$ be the evaluation of subject i for singleton x_j, satisfying:

$$0<f_i(x_j)<1, i=1,2,...,N,\ j=1,2,...,n \quad (31)$$

$$\gamma(x_j)=\frac{1+r(f(x_j))}{\sum_{k=1}^{n}\left[1+r(f(x_k))\right]},\ \ j=1,2,...,n \quad (32)$$

where $r(f(x_j))=\dfrac{S_{y,x_j}}{S_y S_{x_j}}$ (33)

$$S_y^2=\frac{1}{N}\sum_{i=1}^{n}\left(y_i-\frac{1}{N}\sum_{i=1}^{N}y_i\right)^2 \quad (34)$$

$$S_{x_j}^2=\frac{1}{N}\sum_{i=1}^{n}\left[f_i(x_j)-\frac{1}{N}\sum_{i=1}^{N}f_i(x_j)\right]^2 \quad (35)$$

$$S_{y,x_j}=\frac{1}{N}\sum_{i=1}^{n}\left(y_i-\frac{1}{N}\sum_{i=1}^{N}y_i\right)\left[f_i(x_j)-\frac{1}{N}\sum_{i=1}^{N}f_i(x_j)\right] \quad (36)$$

satisfying $0\le\gamma(x_j)\le 1$ and $\sum_{j=1}^{n}\gamma(x_j)=1$ (37)

then the function $\gamma: X\to[0,1]$ satisfying $\mu(\{x\})=\gamma(x),\ \forall x\in X$ is a fuzzy support of μ, called γ-support of μ.

VI. CHOQUET INTEGRAL REGRESSION MODELS

A. Choquet Integral

Definition 11 Choquet Integral [2-6]
Let μ be a fuzzy measure on a finite set X. The Choquet integral of $f_i: X\to R_+$ with respect to μ for individual i is denoted by

$$\int_C f_i d\mu=\sum_{j=1}^{n}\left[f_i(x_{(j)})-f_i(x_{(j-1)})\right]\mu(A_{(j)}^i)\ ,i=1,2,...,N \quad (38)$$

where $f_i(x_{(0)})=0$, $f_i(x_{(j)})$ indicates that the indices have been permuted so that

$$0\le f_i(x_{(1)})\le f_i(x_{(2)})\le...\le f_i(x_{(n)}) \quad (39)$$

$$A_{(j)}=\{x_{(j)},x_{(j+1)},...,x_{(n)}\} \quad (40)$$

B. Choquet Integral Regression Models

Definition 12 Choquet Integral Regression Models [7-15]
Let $y_1,y_2,...,y_N$ be global evaluations of N objects and $f_1(x_j), f_2(x_j),..., f_N(x_j),\ j=1,2,...,n$, be their evaluations of x_j, where $f_i: X\to R_+$, $i=1,2,...,N$.

Let μ be a fuzzy measure, $\alpha,\beta\in R$,

$$y_i=\alpha+\beta\int_C f_i dg_\mu+e_i\ ,\ e_i\sim N(0,\sigma^2)\ ,\ i=1,2,...,N \quad (41)$$

$$(\hat{\alpha},\hat{\beta})=\arg\min_{\alpha,\beta}\left[\sum_{i=1}^{N}\left(y_i-\alpha-\beta\int_C f_i dg_\mu\right)^2\right] \quad (42)$$

then $\hat{y}_i=\hat{\alpha}+\hat{\beta}\int f_i dg_\mu,\ i=1,2,...,N$ is called the Choquet integral regression equation of μ, where

$$\hat{\beta}=S_{yf}/S_{ff} \quad (43)$$

$$\hat{\alpha}=\frac{1}{N}\sum_{i=1}^{N}y_i-\hat{\beta}\frac{1}{N}\sum_{i=1}^{N}\int f_i dg_\mu \quad (44)$$

$$S_{yf}=\frac{\sum_{i=1}^{N}\left[y_i-\frac{1}{N}\sum_{i=1}^{N}y_i\right]\left[\int f_i dg_{\mu^*}-\frac{1}{N}\sum_{k=1}^{N}\int f_k dg_{\mu^*}\right]}{N-1} \quad (45)$$

$$S_{ff}=\frac{\sum_{i=1}^{N}\left[\int f_i dg_{\mu^*}-\frac{1}{N}\sum_{k=1}^{N}\int f_k dg_{\mu^*}\right]^2}{N-1} \quad (46)$$

VII. EXPERIMENTS AND RESULTS

A. Education Data

The total scores of 60 students from a junior high school in Taiwan are used for this research [9-13]. The examinations of four courses, physics and chemistry, biology, geoscience and mathematics, are used as independent variables, the score of the Basic Competence Test of junior high school is used as a dependent variable.

The data of all variables listed in Table IV is applied to evaluate the performances of seven Choquet integral regression models with P-measure, λ-measure and δ-measure, L-measure measure, L_δ -measure, and $L_{m\delta}$ -measure based on γ-support respectively, a ridge regression model, and a multiple linear regression model by using 5-fold cross validation method to compute the mean square error (MSE) of the dependent variable. The formula of MSE is

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \qquad (47)$$

The fuzzy density function, γ-support of the P-measure, λ-measure, δ-measure, L-measure, L_δ -measure, L_m -measure, and $L_{m\delta}$ -measure are listed as follows which can be obtained by using the formula (32).

$$\{0.2488, 0.2525, 0.2439, 0.2547\} \qquad (48)$$

Due to above same fuzzy support, all event measures of any fuzzy measure can be found, and then, the Choquet integral based on abovementioned fuzzy measures and the Choquet integral regression equation based on those fuzzy measures can also be found by using above corresponding formulae.

The experimental results of nine forecasting models are listed in Table I. We find that the Choquet integral regression model with $L_{m\delta}$ -measure based on γ-support outperforms other forecasting regression models.

TABLE I MSE OF REGRESSION MODELS

Regression model	measure	5-fold CV MSE
Choquet Integral Regression model	$L_{m\delta}$	47.5698
	L_m	47.6187
	L_δ	47.9742
	L	48.4610
	δ	48.7672
	λ	49.1832
	p	53.9582
Ridge regression		59.1329
Multiple regression		65.0664

B. Thermostable Proteins Data

A thermostable proteins data set was downloaded from the Protein Data Bank (PDB), http://www.rcsb.org, [16]. By substituting four physicochemical quantities of each residue of amino acid in sequence of the thermostable proteins using the four feature scaling estimators, we can obtain four non-symbolic sequences of the thermostable proteins. Then we can estimate the Hurst exponents of each non-symbolic sequence of the thermostable proteins by using R/S method [17]. It is based on empirical observations by Hurst in 1965 and estimates Hurst exponents are based on the R/S statistic. It indicates (asymptotically) second-order self-similarity. Hurst exponent is roughly estimated through the slope of the linear line in a log-log plot, depicting the R/S statistics over the number of points of the aggregated series. That is, given a time sequence of observations, w_t define the Series

$$W(t,\tau) = \sum_{u=1}^{t}\left(w_u - \overline{w}_\tau\right), 1 \le t \le \tau \qquad (49)$$

where $$\overline{w}_\tau = \frac{1}{\tau}\sum_{t=1}^{\tau} w_t \qquad (50)$$

Define $$R(\tau) = \max_{t=1}^{\tau} W(t,\tau) - \min_{t=1}^{\tau} W(t,\tau) \qquad (51)$$

and $$S(\tau) = \sqrt{\left(\frac{1}{\tau}\sum_{t=1}^{\tau}\left(w_t - \overline{w}_\tau\right)^2\right)} \qquad (55)$$

In plotting $\log\frac{R(\tau)}{S(\tau)}$ against $\log\tau$, we expect to get a line whose slope determines the Hurst exponent.

We can obtain four features of Hurst exponents in each sequences of the thermostable protein. The data of all features are listed in Table IV. Using these extracted features, we can predict the temperature of the 40 thermostable proteins.

For evaluating the performances of seven Choquet integral regression models with P-measure, λ-measure and δ-measure, L-measure measure, L_δ -measure, and $L_{m\delta}$ -measure based on γ-support respectively, a ridge regression model, and a multiple linear regression model we can use 5-fold cross validation method to compute the mean square error (MSE) of the dependent variable by equation (47).

The fuzzy density function, γ-support of the P-measure, λ-measure, δ-measure, L-measure, L_δ -measure, L_m -measure, and $L_{m\delta}$ -measure are listed as follows which can be obtained by using the formula (32).

$$\{0.1697, 0.2633, 0.31799, 0.2490\} \qquad (56)$$

TABLE II MSE OF REGRESSION MODELS

Regression model	measure	5-fold CV MSE
HE-Choque t Integral Regression model with fuzzy measure	$L_{m\delta}$	21.4563
	L_m	21.5634
	L_δ	21.6173
	L	21.7794
	δ	22.3164
	λ	22.0117
	P	22.5051
HE-Ridge regression		23.9718
HE-Multiple regression regression		25.3937

For any fuzzy measure, μ-measures, once the fuzzy support of the μ-measure is given, all event measures of μ can be found, and then, the Choquet integral based on μ and the Choquet integral regression equation based on μ can also be found by using above corresponding formulae.

The experimental results of nine forecasting models are listed in Table II. We find that the Choquet integral regression model with $L_{m\delta}$-measure based on γ-support outperforms other forecasting regression models.

VIII. Conclusion

In this paper, a improved multivalent composed fuzzy measure of maximized L-measure and δ-measure, called $L_{m\delta}$-measure, is proposed. This new measure is proved that it is of closed form with infinitely many solutions, and it can be considered as an extension of the two well known fuzzy measures, λ-measure, P-measure and B-measure. Furthermore, this improved multivalent fuzzy measure is not only including the additive measure, but also a completed fuzzy measure having the more infinitely many measure solutions than other abovementioned multivalent fuzzy measures. By using 5-fold cross-validation MSE, two experiments are conducted for comparing the performances of a multiple linear regression model, a ridge regression model, and the Choquet integral regression model with respect to P-measure, λ-measure, δ-measure, L-measure, L_{δ}-measure and the new measure, $L_{m\delta}$-measure, based on γ-support respectively. Both of two results show that the Choquet integral regression models with respect to the proposed $L_{m\delta}$-measure based on γ-support outperforms other forecasting models.

In the future, we will apply the proposed Choquet integral regression model with the new fuzzy measure based on γ-support to develop multiple classifier systems and multi-criteria decision making systems. .

Acknowledgment

This paper is partially supported by the grant of National Science Council of Taiwan Government (NSC 98-2410--H-468-014).

References

[1] A. E. Hoerl, R. W. Kenard, and K. F. Baldwin, Ridge Regression: Some Simulation, *Communications in Statistics*, vol. 4, No. 2, pp. 105-123, 1975.

[2] Z. Wang, and G. J. Klir, *Fuzzy Measure Theory*, Plenum Press, New York, 1992.

[3] Z. Wang, and G. J. Klir, *Generalized Measure Theory*, Springer Press, New York, 2009.

[4] M. Sugeno, *Theory of Fuzzy Integrals and its Applications*, unpublished doctoral dissertation, Tokyo Institute of Technology, Tokyo, Japan, 1974.

[5] L. A. Zadeh, *Fuzzy Sets and Systems*, vol. 1, pp. 3, 1978.

[6] G. Choquet, Theory of Capacities, *Annales de l'Institut Fourier*, vol. 5, pp. 131-295, 1953.

[7] H.-C. Liu, Y.-C Tu,. C.-C. Chen, and W.-S. Weng, "The Choquet Integral with Respect to λ-Measure Based on γ-support", *2008 International Conferences on Machine Learning and Cybernetics*, Kunming, China, July 2008.

[8] H.-C. Liu, "A theoretical approach to the completed L-fuzzy measure", 2009 International Institute of Applied Statistics Studies (2009IIASS) 2nd conference, Qindao, China, July 24-29, 2009.

[9] H.-C. Liu, D.-B. Wu, Y-D Jheng, and T.-W. Sheu, "Delta-Fuzzy Measures and its Application", *Proceedings of the 8th WSEAS International Conference on Applied Computer and Applied Computational Science*, pp. 289-293, May 2009.

[10] H.-C. Liu, D.-B. Wu, Y-D Jheng, and T.-W. Sheu, "Theory of Multivalent Delta-Fuzzy Measures and its Application", *WSEAS Transactions on Information Science and Application*, vol. 6, no. 6, pp. 1061-1070, July 2009.

[11] H.-C. Liu, C.-C. Chen, D.-B. Wu, and T.-W. Sheu, "Theory and Application of the Composed Fuzzy Measure of L-Measure and Delta-Measures", *WSEAS TRANSACTION ON INTERNATIONAL SCIENCE AND CONTRAL* , Issue 8. Vol. 4, pp. 359-368, Augest 2009

[12] J.-I Shieh, H.-H. Wu, H.-C. Liu, Applying Complexity-based Choquet Integral to Evaluate Students' Performance, *Expert Systems with Applications*, 36, pp. 5100-5106, 2009.

[13] S. Srivastava, A. Bansal, D. Chopra, G. Goe, Modeling and Control of a Choquet Fuzzy Integral Based Controller on a Real Time System. *WSEAS Transactions on Systems*, vol. 7, no. 5, pp. 1571-1578, July 2006.

[14] H.-C. Liu, Chen, C.-C., Chen, G.-S., & Jheng, Y.-D. (2007). Power-transformed-measure and its Choquet Integral Regression Model, *Proc. of 7th WSEAS International Conference on Applied Computer Science* (ACS'07), pp.101-104, 2007.

[15] G.S. Chen, Y.-D. Jheng, H.-C. Yao, and H.-C. Liu, Stroke Order Computer-based Assessment with Fuzzy Measure Scoring. *WSEAS Transactions Information Science & Applications*, vol. 2, no. 5, 62-68, 2008.

[16] Protein Data Bank (PDB), http://www.rcsb.org/

[17] H. E. Hurst: Long term storage capacity of reservoirs, *Transactions of the American Society of Civil Engineers*, *Vol.116*, pp. 770-799, 1951.

Hsiang-Chuan Liu received the Ph.D. degree in Statistics from National Taiwan University, Taiwan. He is a professor at the Department of Bioinformatics, Asia University, Taiwan since August 2001 and also a honored professor at the Graduate Institute of Educational Measurement and Statistics, National Taichung University, Taiwan. He was the President of National Taichung University, Taiwan from 1993 to 2000. Dr. Liu is a member of IEEE since 2007. He has funded research and published articles in the areas of Biostatistics, Bioinformatics, Fuzzy Theory, Educational Measurement, and E-Learning.

Hsien-Chang Tsai was born in Chiayi, Taiwan in 1961. He was admitted to candidacy for Ph.D. in Graduate Institute of Science Education, National Changhua University of Education, Taiwan. He is a lecturer at the Department of Biology, National Changhua University of Education, Taiwan. His research and teaching focus on Biostatistics, Computational Biology, Educational and Psychological Measurement and Evaluation, Biology Education, Animal Histology, and Embryology, etc.

Yu-Du Jheng, He was born in Taiwan on July, 1970 and received the Ph.D. in Educational Measurement and Statistics from National Taichung University, Taiwan. He is an Assistant Professor at the Department of Education, National Taichung University. Dr. Jheng is a member of IEEE since 2007. His studies focus in Fuzzy Integral, Fuzzy Measure, Artificial Neural Network, Measurement and Statistics, etc.

Tung-Sheng Liu graduated from the Department of Electrical Engineering, Chung-Cheng Institute of Technology and the Department of Foreign Language, Aletheia University, Taiwan. He began to devote his career life in military since 1996. Now he is a major serving in Forensic Science Center, Military Police Command, R.O.C and the leader of CSI section. In military, he spent a lot of time researching Forensic Science, also got the opportunity to be the instructor in Military Police School, R.O.C. Since serving in Forensic Science Center, due to his interest, he has been in charge of polygraph and pattern recognition and also been the member in Taiwan Academy of Forensic Science.

TABLE III THE DATA SET WITH FOUR COURSES AND SCIENCE SCORES OF THE BCT

No.	C1	C2	C3	C4	BCT	No.	C1	C2	C3	C4	BCT
1	72	66	78	72	19	31	66	68	75	74	25
2	86	80	82	81	35	32	68	70	74	76	40
3	56	63	69	75	21	33	57	65	75	70	24
4	78	86	86	86	33	34	74	70	80	75	35
5	66	72	80	76	23	35	49	60	69	64	13
6	68	74	77	80	28	36	51	60	63	64	18
7	74	86	87	88	44	37	58	64	68	66	32
8	54	56	62	68	7	38	73	78	84	81	39
9	71	74	80	77	26	39	56	56	65	61	6
10	68	70	80	75	33	40	61	62	70	70	25
11	53	56	70	63	22	41	57	60	68	64	23
12	67	70	80	75	35	42	57	64	67	70	26
13	70	66	70	74	13	43	50	52	68	60	7
14	60	65	75	70	23	44	84	80	76	72	49
15	68	68	78	76	35	45	62	66	76	71	22
16	58	66	76	71	37	46	70	74	78	82	32
17	61	66	72	78	33	47	69	70	80	75	26
18	68	68	80	74	26	48	63	74	74	74	42
19	56	66	76	71	21	49	66	78	80	82	39
20	59	62	70	78	29	50	67	70	80	75	31
21	62	64	76	70	36	51	56	65	75	70	23
22	71	72	78	75	26	52	50	54	66	60	18
23	74	63	69	75	12	53	71	75	85	80	41
24	59	70	80	76	37	54	74	77	80	85	26
25	75	75	85	80	39	55	71	72	76	80	31
26	73	78	84	81	24	56	60	65	75	70	21
27	62	68	72	74	29	57	59	57	70	68	17
28	77	74	80	76	42	58	50	56	65	68	13
29	63	60	68	69	17	59	72	76	80	78	38
30	56	61	75	68	22	60	81	76	78	80	33

C1 : physics and chemistry
C2 : biology
C3 : geoscience
C4 : mathematics
BCT : Basic Competence Test of nature science

Table IV Hurst exponents of four feature scaling of Thermostable Proteins

Code of Proteins	Temperature	ASA	Electrostatic Interactions	Contact Energy	Solvent Accessibility Percentages
1L	37	0.3832	0.4125	0.3335	0.7335
1H	100	0.4691	0.6572	0.3411	0.5636
2L	35	0.4129	0.3985	0.3315	0.4772
2H	60	0.5119	0.537	0.4524	0.5345
3L	37	0.5079	0.4489	0.2766	0.5512
3H	71	0.4224	0.4183	0.4881	0.5818
4L	27.5	0.4463	0.3964	0.3807	0.6805
4H	47.5	0.4936	0.5039	0.501	0.629
5L	37	0.3577	0.4718	0.4509	0.6078
5H	100	0.4751	0.5812	0.4299	0.5699
6L	30	0.2847	0.5155	0.3618	0.6286
6H	60	0.4314	0.6042	0.3244	0.4178
7L	37	0.5432	0.5496	0.4466	0.6699
7H	90	0.255	1	0.3501	0.5807
8L	28	0.3573	0.5062	0.1892	0.6273
8H	80	0.3461	0.5167	0.2403	0.4284
9L	37	0.5153	0.4902	0.3485	0.6008
9H	72.5	0.3989	0.4223	0.409	0.5215
10L	28	0.4641	0.4432	0.3261	0.5817
10H	80	0.446	0.581	0.2082	0.5506
11L	30	0.3832	0.4125	0.3335	0.7335
11H	80	0.4691	0.6572	0.3411	0.5636
12L	37	0.4129	0.3985	0.3315	0.4772
12H	95	0.5119	0.5371	0.4524	0.5345
13L	35	0.5079	0.4489	0.2766	0.5512
13H	65	0.4224	0.4183	0.4881	0.5818
14L	27.5	0.4463	0.3964	0.3807	0.6805
14H	74.5	0.4936	0.5039	0.501	0.629
15L	47.3	0.3577	0.4718	0.4509	0.6078
15H	94	0.4751	0.5812	0.4299	0.5699
16L	47.5	0.2847	0.5155	0.3618	0.6286
16H	90	0.4314	0.6042	0.3244	0.4178
17L	35	0.5432	0.5496	0.4466	0.6699
17H	86	0.255	0.5078	0.3501	0.5807
18L	27.5	0.3573	0.5062	0.1892	0.6273
18H	85	0.3461	0.5167	0.2403	0.4284
19L	30	0.5153	0.4902	0.3485	0.6008
19H	85	0.3989	0.4223	0.409	0.5215
20L	55	0.4641	0.4432	0.3261	0.5817
20H	113	0.446	0.581	0.2082	0.5506