Journal of Marine Science and Technology

Volume 20 | Issue 4 Article 9

FUZZY STATISTICAL REFINEMENT FOR THE FORECASTING OF TENDERS FOR ROADWAY CONSTRUCTION

Jeng-Wen Lin
Department of Civil Engineering, Feng Chia University, Taichung, Taiwan, R.O.C.

Cheng-Wu Chen
Institute of Maritime Information and Technology, National Kaohsiung Marine University, Kaohsiung, Taiwan, R.O.C.; Global Earth Observation and Data Analysis Center, National Cheng Kung University, Tainan, Taiwan, R.O.C., chengwu@mail.nkmu.edu.tw

Ting-Chang Hsu
Department of Civil Engineering, Feng Chia University, Taichung, Taiwan, R.O.C.

Recommended Citation

Lin, Jeng-Wen; Chen, Cheng-Wu; and Hsu, Ting-Chang (2012) "FUZZY STATISTICAL REFINEMENT FOR THE FORECASTING OF TENDERS FOR ROADWAY CONSTRUCTION," *Journal of Marine Science and Technology*: Vol. 20: Iss. 4, Article 9.
DOI: 10.6119/JMST-011-0318-1
Available at: https://jmstt.ntou.edu.tw/journal/vol20/iss4/9

FUZZY STATISTICAL REFINEMENT FOR THE FORECASTING OF TENDERS FOR ROADWAY CONSTRUCTION

Acknowledgements

The authors would like to thank the National Science Council, R.O.C., for the financial support of this research under Contract Nos. NSC 97-2221-E-035-045, NSC 100- 2221-E-022-013-MY2, and NSC 100-2628-E-022-002-MY2.

Journal of Marine Science and Technology, Vol. 20, No. 4, pp. 410-417 (2012)
DOI: 10.6119/JMST-011-0318-1

FUZZY STATISTICAL REFINEMENT FOR THE FORECASTING OF TENDERS FOR ROADWAY CONSTRUCTION

Jeng-Wen Lin[1], Cheng-Wu Chen[2], and Ting-Chang Hsu[1]

Key words: data classification, fuzzy theory, model refinement, power series model, tendering price prediction.

ABSTRACT

Due to market competition, construction companies often place low bids when tenders are invited for domestic public construction projects. Over-competition can lead to vicious price wars to win a tender, which can in turn seriously affect the quality of construction. This study aims to establish an accurate Taiwan based model for the forecasting of the tendered price for roadway construction. This model is designed to assist the public sector to determine what would be a reasonable reserve price or award price. In order to ensure accurate predictions, a data classification system is established using fuzzy set theory. For each category of classified data, multiple regression analysis is applied to the linear model, the power series model, and the refined power series model. Multiple factors in the regression for the tender price prediction include the contract schedule, the budget price, and the tender bond. It is shown that the average relative error of the final reserve price model is about 3%, while that for the price of award model is 9%. In comparison, the developed reserve price model is more feasible than the price of award model.

I. INTRODUCTION

The construction industry in Taiwan has been facing long-term problems of price competition, forcing low price bidding to win tenders, which has often squeezed their profit margin. Bid winners have been forced to adopt the business pattern of sub-contracting in order to split their risks, and transfer costs and responsibilities. The sub-contracting pattern may have several negative consequences. One is a degradation on the quality of construction, and the other is difficulty in managing multiple tasks at the same time. For the contracting owner, it is a very important matter to select a contractor in excellent financial condition. Sound finances are a reflection of the contractor's reliability for planning, organizing, control and human resource management. Under the circumstances, therefore, prior to accepting a tender, one needs to examine the engineering experience and financial condition of possible contractors [1, 6, 23, 26]. The contractor's financial condition can be assessed through the observation of its working capital management.

In order to solve the problem of over-competitiveness in the construction industry, Ng *et al*. suggested that clients needed to be informed in advance of their likely future financial commitments and cost implications with the design evolution [22]. This requires the estimation of building costs which is done based on historical cost data updated by the forecast tender price index (TPI). Reliable short- to medium-term prediction of the TPI is crucial to construction company stakeholders [36]. The model for TPI forecasting can also assist the public sector in planning the construction workload to improve the stability of the construction market. Yu and Ive carried out a critical review of the methods for compilation of building price indexes in Britain [39]. They argued that the importance of accurate measurement and pertinent modeling of the general level of construction prices cannot be overemphasized. Uses range from macroeconomic statistics such as the real value of the investment to micro-level budgeting such as the forecast price of the construction project.

After acquiring accurate price indexes, tendering becomes an important task for construction companies. The tendering results have a great influence on the operating performance and profits of the construction company. Money, time and manpower must be invested to submit tenders. If the company fails to win the tender, those resources have been wasted. Therefore, it is very important for construction companies to offer suitable prices for tenders for construction projects they are about to bid on based on the price of awards from previous tenders. Preparation of such a prediction is the motivation for this study. The aim is to increase the chances of clients to win

Paper submitted 09/02/10; revised 01/22/11; accepted 03/18/11. Author for correspondence: Cheng-Wu Chen (e-mail: chengwu@mail.nkmu.edu.tw).
[1] *Department of Civil Engineering, Feng Chia University, Taichung, Taiwan, R.O.C.*
[2] *Institute of Maritime Information and Technology, National Kaohsiung Marine University, Kaohsiung, Taiwan, R.O.C.; Global Earth Observation and Data Analysis Center, National Cheng Kung University, Tainan, Taiwan, R.O.C.*

tenders and help the public sector to establish reasonable prices of awards and related regulations.

McCaffer *et al*. [21] predicted the tender price of buildings during the early design stage. They provided estimates using a library of data containing rate, quantity and date for the constituent elements of previously construction buildings, inflation indexes and statistical models. Thirty-two different models were included, together with the criterion for selecting the most appropriate, which ensured the most precise prediction possible. An analysis of pre-tender building price forecasts was made by a Hong Kong consulting organization for a series of 89 building projects from 1995 to 1997 [25]. Identification of the factors influencing the accuracy of the forecasts was made for possible improvement in performance. The purpose is to identify and explain the underlying systematic causes of errors, in addition to assist in improving the predictive ability of the forecasts using statistical methods such as analysis of variance. In Taiwan, Chiu compared the tendering systems and the tender award systems before and after the implementation of the Government Purchase Act for highway construction works [5]. They collected data from 2251 construction purchase cases. Stepwise regression analysis and cluster analysis were applied to find possible explanatory variables and factors of influence on differences between the reserve prices and award prices. The results obtained could be a reference for the reserve price review committee to set up reserve prices and for purchasing departments to determine reasonable award prices. An introduction to the Government Purchase Act is available in the work of Lo [20]. Analyses of the regulations in the Act related to prices, reserve prices, and award prices were provided. They also discussed an example, a public tender for a construction project opened by the Laiyi Township Office in Pingtung County. Simple linear regression analysis was performed to obtain the relationships between budget prices and prices of awards, and between prices of award and reserve prices. Ranges for reserve prices were determined according to prices of awards and showed a normal distribution. The implementation results show that the practicality of the model.

This study aims to establish an accurate model based on Taiwan practices for the prediction of price tenders on roadway construction. The goal is to prepare winning tenders, and to assist the public sector to determine reasonable reserve prices and award prices. This is done by (1) assessing the established models using classified data so as to promote the accuracy of prediction and (2) comparing the efficiency of the models established by power series with those refined using a proposed statistical approach. First, fuzzy set theory is introduced to classify the obtained data for multiple regression analysis. Multiple factors in the regression for tender price prediction include the contract schedule (calendar days), budget price, and tender bond after referencing the documents. These three variables must be correlated with purchase information and are related to the reserve price and price of award. Second, the three variables are applied in power series modeling. The verification and prediction errors from the power series models are compared and data are classified into different categories. Finally, a statistical model refinement approach, based on the confidence interval, is introduced to refine the established power series models, both for the reserve price and price of award predictions, and to reduce possible modelling uncertainties.

II. FUZZY SET THEORY

A data classification system was established using fuzzy set theory in order to determine the optimal number of categories and ensure the accuracy of the forecast tendering price [2, 9].

Definition 1: Let R be a real number set. A fuzzy set $\tilde{A}$ on R is said to be a fuzzy number if the following conditions are satisfied:

(1) $\exists x_0 \in R$, such that $\mu_{\tilde{A}}(x_0)=1$, and membership function $\mu_{\tilde{A}}(x)$ is piecewise continuous; and

(2) $\forall \alpha \in (0,1]$, $A_\alpha \equiv \left\{x \,\middle|\, \mu_{\tilde{A}}(x) \geq \alpha, x \in R\right\}$ is a convex set on R,

where x_0 is the mean value of $\tilde{A}$ and A_α is a crisp set. The convex set means that $\forall x \in [x_1, x_2]$,

$$f(x) \geq \min(f(x_1),\ f(x_2)). \tag{1}$$

Evidently for any $\forall \alpha \in (0, \alpha]$ the α-level set $\tilde{A}_\alpha$ will be expressed as a closed interval $[p, q]$. Based on the fuzzy extension principle (Zadeh, 1965), linear operations about closed intervals are obtained as follows:

Lemma 1: Let $[a, \mathrm{b}]$, $[d, \mathrm{e}]$ be closed intervals of real numbers. Then

$$\begin{aligned}&[a,\ \mathrm{b}]+[d,\ \mathrm{e}]=[a+d,\ b+\mathrm{e}];\\&[a,\ \mathrm{b}]-[d,\ \mathrm{e}]=[a-e,\ b-\mathrm{d}],\end{aligned} \tag{2}$$

$$[a,\ \mathrm{b}]\cdot[d,\ \mathrm{e}]=[\min(ad,\ ae,\ b\ d,\ be),\ \max(ad,\ ae,\ b\ d,\ be)]; \tag{3}$$

$$\begin{aligned}&[a,\ \mathrm{b}]/[d,\ \mathrm{e}]=[a,\ \mathrm{b}]\cdot[1/e,\ 1/d]\\&=[\min(a/d,\ a/e,\ b\ /d,\ b/e),\ \max(a/d,\ a/e,\ b\ /d,\ b/e)].\end{aligned} \tag{4}$$

Remark 1: Given any operations which have commutative and associative characteristics, the operations of extension still have these characteristics.

From the theory of the α-level described above and the decomposition theorem discussed by Klir and Yuan [12] we obtain

$$(A*B)_\alpha \equiv A_\alpha * B_\alpha, \tag{5}$$

$$A*B \equiv \bigcup_{\alpha\in(0,1]} (A*B)_\alpha, \tag{6}$$

where $*$ denotes any arithmetic operation; A and B are fuzzy numbers; and $A*B$ will be a fuzzy number.

Remark 2: (Wang and Chiu [34]): the resultant fuzzy numbers described above are the same type as the original fuzzy numbers after the operation of addition or subtraction.

Definition 2: Extended Operations for LR-Representation of Fuzzy Sets.

A triangular fuzzy number $\tilde{A}$ denoted by (m, β, γ) is defined as

$$\mu_{\tilde{A}}(x) = \begin{cases} 1\text{-}\dfrac{|m-x|}{\beta} & \text{if } m-\beta \le x \le m \\ R\left(\dfrac{x-m}{\gamma}\right) & \text{if } m+\gamma \ge x \ge m \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

where $m \in R$ is the center; $\beta > 0$ is the left spread; and $\gamma > 0$ is the right spread of $\tilde{A}$.

If $\beta = \gamma$, then the triangular fuzzy number is called a symmetric triangular fuzzy number and is denoted by (m, β).

An LR-type fuzzy number $\tilde{A} = (m, \beta, \gamma)_{LR}$ is a function from the real into the interval [0,1] satisfying

$$\mu_{\tilde{A}}(x) = \begin{cases} L\left(\dfrac{m-x}{\beta}\right) & \text{for } m-\beta \le x \le m \\ R\left(\dfrac{x-m}{\gamma}\right) & \text{for } m+\gamma \ge x \ge m \\ 0 & \text{else,} \end{cases} \tag{8}$$

where L and R are nondecreasing and continuous functions from [0, 1] to [0, 1] satisfying $L(0) = R(0) = 1$ and $L(1) = R(1) = 0$. If $L = R$ and $\beta = \gamma$, then the symmetric L-L fuzzy number is denoted as $(m, \beta)_L$.

Lemma 2: Given two LR-type fuzzy numbers $\tilde{A}$ and $\tilde{B}$, we have

$$(m, \beta, \gamma)_{LR} + (n, \delta, \eta)_{LR} = (m+n, \beta+\delta, \gamma+\eta)_{LR} \tag{9}$$

$$(m, \beta, \gamma)_{LR} - (n, \delta, \eta)_{LR} = (m-n, \beta+\eta, \gamma+\delta)_{LR}. \tag{10}$$

Table 1. Fuzzy inference rules.

Number	Contract schedule	Budget price	Tender bond	Total	Category
1	W = 0	W = 0	W = 0	W = 0	Category 1
2	W = 0	W = 0	W = 1	W = 1	
3	W = 0	W = 1	W = 0	W = 1	
4	W = 1	W = 0	W = 0	W = 1	
5	W = 1	W = 1	W = 0	W = 2	
6	W = 1	W = 0	W = 1	W = 2	
7	W = 0	W = 1	W = 1	W = 2	
8	W = 2	W = 0	W = 0	W = 2	
9	W = 0	W = 2	W = 0	W = 2	
10	W = 0	W = 0	W = 2	W = 2	
11	W = 1	W = 1	W = 1	W = 3	Category 2
12	W = 0	W = 1	W = 2	W = 3	
13	W = 0	W = 2	W = 1	W = 3	
14	W = 1	W = 0	W = 2	W = 3	
15	W = 2	W = 0	W = 1	W = 3	
16	W = 1	W = 2	W = 0	W = 3	
17	W = 2	W = 1	W = 0	W = 3	
18	W = 0	W = 2	W = 2	W = 4	
19	W = 2	W = 0	W = 2	W = 4	
20	W = 2	W = 2	W = 0	W = 4	
21	W = 2	W = 1	W = 1	W = 4	
22	W = 1	W = 2	W = 1	W = 4	
23	W = 1	W = 1	W = 2	W = 4	
24	W = 2	W = 2	W = 1	W = 5	Category 3
25	W = 2	W = 1	W = 2	W = 5	
26	W = 1	W = 2	W = 2	W = 5	
27	W = 2	W = 2	W = 2	W = 6	

In the next section, fuzzy set theory is applied to classify the obtained data, in order to promote the accuracy of construction tender price forecasting.

III. CONSTRUCTION OF THE FUZZY RULE MODEL

The data used in the study include the contract schedule (calendar days), the budget price, and the tender bond. These data are retrieved from public tendering information and used as parameters for system classification. The data related to public tendering information used in this study were input into the system. Every tendering case could be classified using this system. The system output included only one value: the result of classification. The output value was used to determine which category the input case belonged to. Since there was interaction between the three factors used in this study, the “and” fuzzy inference rules were adopted. For completeness of system rules, the total degree of belonging of parameter to the categories was considered. For example, three degrees were considered for each parameter, including 0, 1, and 2. Since there were three parameters, 27 fuzzy inference rules covered all combinations, including the rules for categories 1 to 3. The fuzzy inference rules for this classification system are listed in Table 1.

Table 2. Case profiles (first 20).

Case \ Variable	Contract schedule (days)	Budget price (10,000 dollars)	Tender bond (dollars)	System output value
Case 1	320	261.4	130000	1.05
Case 2	334	350000	9990000	0.979
Case 3	300	500000	9946000	1.210
Case 4	300	430000	8800000	1.170
Case 5	300	300000	8628000	1.160
Case 6	40	850000	18953000	0.947
Case 7	300	350000	7770000	1.160
Case 8	280	120000	4106000	0.948
Case 9	50	760000	19143962	0.940
Case 10	290	195000	3915000	1.050
Case 11	300	500000	9946000	1.210
Case 12	300	430000	8800000	1.170
Case 13	400	800000	19257000	2.870
Case 14	100	1450000	29670000	2.940
Case 15	180	1200000	38581472	3.140
Case 16	365	950000	19179000	2.980
Case 17	400	800000	19257000	2.870
Case 18	100	1450000	29670000	2.940
Case 19	180	1200000	38581472	3.140
Case 20	350	400000	11031240	0.903

Table 1 shows the fuzzy inference rules constructed for this study. The output values of w are the sums of the w values from the input variables. Using the output w values as a gauge, it is possible to determine which real data case belongs to which category. Table 2 shows the classification result (only the first 20 cases are listed). Once each case is classified to find numbers of categories, multiple linear regression is performed in each category using the three factors considered. In order to enhance the regression analysis, these three factors are applied in power series modeling. A comparison of prediction errors from the power series models is made so to determine the optimal number of classifications.

IV. POWER SERIES MODELING

The power series modeling procedure adopted in this study expands a linear model into a nonlinear model [13, 17]. A linear model can be written as

$$y = a_1x_1 + a_2x_2 + a_3x_3, \tag{11}$$

where x_1, x_2, and x_3 are the explanatory variables (predictors); y is the response variable. The power series model expands the linear model into

$$y = \sum_{i=1}^{N}(a_1x_1 + a_2x_2 + a_3x_3)^i, \tag{12}$$

where N represents the highest power required for the expansion. It is user-defined. For example,

$$\begin{aligned} N = 3, \quad y &= (a_1x_1 + a_2x_2 + a_3x_3) \\ &+ (a_4x_4 + a_5x_5 + a_6x_6 + a_7x_7 + a_8x_8 + a_9x_9) \\ &+ \begin{pmatrix} a_{10}x_{10} + a_{11}x_{11} + a_{12}x_{12} + a_{13}x_{13} + a_{14}x_{14} + a_{15}x_{15} \\ + a_{16}x_{16} + a_{17}x_{17} + a_{18}x_{18} + a_{19}x_{19} \end{pmatrix}, \end{aligned} \tag{13}$$

where the second-order terms include

$$x_4 = x_1^2 \quad x_5 = x_2^2 \quad x_6 = x_3^2 \quad x_7 = x_1 * x_2$$
$$x_8 = x_1 * x_3 \quad x_9 = x_2 * x_3,$$

and the third-order terms include

$$x_{10} = x_1^3 \quad x_{11} = x_2^3 \quad x_{12} = x_3^3 \quad x_{13} = x_1^2 * x_2$$
$$x_{14} = x_1^2 * x_3 \quad x_{15} = x_2^2 * x_1 \quad x_{16} = x_2^2 * x_3$$
$$x_{17} = x_3^2 * x_1 \quad x_{18} = x_3^2 * x_2 \quad x_{19} = x_1 * x_2 * x_3.$$

In real-life applications in civil and mechanical engineering, the response variable y is usually represented in a series of powers up to the third-order terms, as shown in Eq. (13). In the following section, a numerical example is given to prove the efficiency of the proposed approach at the reserve price and price of award predictions in Taiwan. Furthermore, a statistical model refinement procedure is introduced to refine the established power series models.

V. EXAMPLE

The reserve price model was built using the reserve price as the dependent variable (y) and the budget price (x_1), the contract schedule (x_2), and the tender bond (x_3) as the independent variables. The price of the award model was built using the price of award as the dependent variable (y) and the same multiple factors x_1, x_2, and x_3 as the independent variables. These three variables must be provided with open purchase information and are related to the reserve price and price of award [19]. In this study, the data were collected from public sector roadway constructions in 2005. The amount of the reserve price and price of award was under NT$50 millions. There were a total of 400 construction project cases. The first 300 cases were used to build the "verification models," while the remaining 100 cases were used to build the "prediction models".

First, a statistical software package, Stata 9, was used to test and anaylyze the multiple regression models. For example, the 300 cases used to build the verification models were classified into 4 categories through the classification system. Multiple regression analysis was then performed in each of the 4 categories to obtain 4 linear regression equations for the reserve price verification models.

Table 3. Comparison of verification and prediction errors for the linear models, with data classified into 5 different numbers of categories, for reserve price predictions.

Reserve price model		Error comparison for linear models					
	category type	1	2	3	4	5	Average error
Verification model	No classification	6.03%					6.03%
	2 categories	5.03%	6.04%				5.54%
	3 categories	5.73%	4.84%	2.93%			4.5%
	4 categories	4.37%	5.88%	3.99%	0.08%		3.58%
	5 categories	3.93%	5.92%	4.18%	6.68%	Fail	
Prediction model	No classification	5.9%					5.9%
	2 categories	5.13%	4.55%				4.84%
	3 categories	5.94%	6.3%	11.49%			7.91%
	4 categories	5.28%	3.91%	5.2%	15.74%		7.53%
	5 categories	4.83%	5.61%	4.33%	7.38%	Fail	

$$y = 206775.7 + (0.79)x_1 - (4901.64)x_2 + (3.25)x_3 \tag{14}$$

$$y = -124576.3 + (0.78)x_1 + (279.82)x_2 + (2.58)x_3 \tag{15}$$

$$y = 5321930 + (0.74)x_1 - (5846.5)x_2 - (0.72)x_3 \tag{16}$$

$$y = -3751049 + (0.8)x_1 - (9669.8)x_2 + (5.3)x_3 \tag{17}$$

The x variables in Eqs. (14)-(17) were then replaced with real data values from the 4 categories to predict the y value (reserve price). In this way, the relative prediction errors were obtained from the verification models for the 4 categories, the verification errors. Similarly, the 100 cases used to build prediction models were classified into 4 categories using the same classification system. The x variables in Eqs. (14)-(17) were then replaced with the real data values from the 100 cases using the verification models, to obtain the relative prediction errors for each category.

The same procedure was performed with data being classified into 2, 3, and 5 categories. Table 3 lists the comparison results showing verification errors and prediction errors for the linear models for 5 different classifications, including one category with all cases, and 2, 3, 4, and 5 categories. Obviously, the best results were obtained from the linear models when the data were classified into 4 categories. The average relative error was 3.58%. When the data were classified into 5 categories cases in the 5th category were too few. Therefore it was not possible to explore the performance of models where the data were classified into 5 different categories.

Second, the expanded form of the linear terms (x_1, x_2, and x_3), with data classified into 4 categories, represented the power series up to the 3rd-order terms with 19 independent variables as shown in Eq. (13). Multiple regression analysis was performed to obtain 4 power series Eqs. (18)-(21).

$$y = -1.02\times10^6 + (7.68\times10^{-1})x_1 + (4.55\times10^4)x_2 + \cdots\cdots + (4.66\times10^{-8})x_{19} \tag{18}$$

$$y = 2.04\times10^6 + (1.74\times10^{-1})x_1 - (1.97\times10^4)x_2 \cdots\cdots - (1.47\times10^{-8})x_{19} \tag{19}$$

$$y = 1.46\times10^8 - (2.78)x_1 - (7.38\times10^5)x_2 + \cdots\cdots(3.09\times10^{-9})x_{19} \tag{20}$$

$$y = -3751049 + (0.7961728)x_1 - (9669.808)x_2 + (5.33037)x_3 \tag{21}$$

For comparison purposes, the respective verification errors and prediction errors of the power series models were calculated and are listed in Table 4. The best result was obtained when using the power series model with data classified into 4 categories for reserve price predictions. The average relative error was reduced from 3.58% to 2.65%.

Third, the power series model for each category mentioned above was refined through a statistical model refinement approach. The 95% confidence intervals (CIs) of the parameters (a_i, i = 1, 2, 3, $\cdots$, 19 in Eq. (13)) corresponding to the 19 independent variables in the power series were tested using an exclusion criterion to determine their statistical significances. If the zero (null) value fell within the CI for a parameter, the corresponding term was excluded, so that the CIs of all parameters would not cover the zero value and thus their statistical significance could be sustained. The procedure for building the regression model and selecting CIs was repeated to refine the model until there were no CIs within the zero value. The model refinement approach was applied to the power series models. Each model uses data classified into different numbers of categories. The verification errors and prediction errors were calculated and are listed in Table 5. It can be seen that better results are obtained for the refined model without data classification when compared to those of the power series models. Tables 3, 4, and 5, show the prediction errors for the linear model, power series model, and refined power series model, respectively. It can be seen that for the category without data classification, the reserve price

Table 4. Comparison of the verification and prediction errors obtained from power series models, with 5 different categories, for reserve price predictions.

Reserve price model		Error comparison for power series models					
	Category	1	2	3	4	5	Average error
Verification model	No classification	5.05%					5.05%
	2 categories	4.8%	30.8%				17.8%
	3 categories	4.63%	2.15%	7.81%			4.86%
	4 categories	3.61%	5.31%	1.59%	0.08%		2.65%
	5 categories	3.21%	5.67%	3.48%	33.83%	Fail	
Prediction model	No classification	5.48%					5.48%
	2 categories	5.03%	115.3%				60.17%
	3 categories	4.54%	12.24%	18.77%			11.85%
	4 categories	4.89%	4.47%	7.64%	15.74%		8.19%
	5 categories	6.04%	4.66%	3.95%	104.25%	Fail	

Table 5. Comparison of the verification and prediction errors of the power series models refined through a statistical refinement approach for reserve price predictions.

Reserve price model		Error comparison for refined power series models					
	Category	1	2	3	4	5	Average error
Verification model	No classification	5.48%					5.48%
	2 categories	5.41%	Fail				
	3 categories	4.95%	Fail	Fail			
	4 categories	4.68%	63.88%	Fail	Fail		
	5 categories	4.30%	50.97%	51.54%	Fail	Fail	
Prediction model	No classification	4.64%					4.64%
	2 categories	5.76%	Fail				
	3 categories	4.7%	Fail	Fail			
	4 categories	3.5%	79.28%	Fail	Fail		
	5 categories	3.29%	64.33%	58.62%	Fail	Fail	

predictions are reduced from 5.9% to 5.48% to 4.64%, respectively. Such a reduction in the prediction error illustrates the capacity of the proposed model refinement approach.

Finally, in order to explore what number of data classification categories will lead to the best price, price of award models were built for different numbers of categories. The verification and prediction errors of the linear models are shown in Table 6. For each model data are classified into different categories. Obviously, the best result is obtained with the linear models built when the data are classified into 3 categories, with the average relative verification error being 8.77% and the corresponding prediction error being 15.25%.

The linear models previously built were expanded into power series models, with data classified into different numbers of categories. This is done to explore which power series model leads to the best performance. The verification and prediction errors for the power series models for the price of award predictions are listed in Table 7. The best results were obtained using the power series model with data classified into 4 categories for the price of award predictions, with the average relative error being 8.59%.

According to the comparisons shown in Tables 6 and 7, the difference between the average errors of the linear model built with data classified into 3 categories and that of the power series model built with data classified into 4 categories were the smallest. The prediction error for the former was 15.25%, while that of the latter was 41.09%. Because of the relatively lower prediction error, the linear model built with data classified into 3 categories was chosen in this study for the price of award modeling.

VI. CONCLUSIONS

There are increasing AI and computational methodologies that are proposed to overcome the hazard or practical problems (see [3, 4, 7, 8, 10, 11, 14-16, 18, 24, 27-33, 35, 37, 38,

Table 6. Comparison of verification and prediction errors for linear models, with data classified into 5 different categories, for the price of award predictions.

Price of award model		Error comparison for linear models					
	Category	1	2	3	4	5	Average error
Verification model	No classification	13.79%					13.79%
	2 categories	18.15%	7.22%				12.69%
	3 categories	19.46%	5.71%	1.14%			8.77%
	4 categories	15.81%	20.15%	8.51%	0.33%		11.20%
	5 categories	15.65%	18.66%	18.01%	6.68%	Fail	
Prediction model	No classification	15.09%					15.09%
	2 categories	17.61%	16.29%				16.95%
	3 categories	18.8%	8.85%	18.09%			15.25%
	4 categories	16.71%	16.59%	4.38%	119.71%		39.35%
	5 categories	13.49%	18.08%	13.76%	7.37%	Fail	

Table 7. Comparison of the verification and prediction errors for the power series models, with 5 different categories, for the price of award predictions.

Price of award models		Error comparison for power series models					
	Category	1	2	3	4	5	Average error
Verification model	No classification	18.26%					18.26%
	2 categories	16.07%	28.54%				22.31%
	3 categories	16.42%	10.01%	13.85%			13.43%
	4 categories	12.16%	19.21%	2.67%	0.33%		8.59%
	5 categories	12.3%	18.74%	11.59%	38.83%	Fail	
Prediction model	No classification	17.96%					17.96%
	2 categories	15.84%	52.57%				34.21%
	3 categories	16.75%	29.81%	117.51%			54.69%
	4 categories	17.86%	16.7%	10.09%	119.71%		41.09%
	5 categories	19.01%	24.47%	13.87%	104.25%	Fail	

40, 41] and the references therein). In this paper, the proposed regression model can be used to accurately interpret the successful tendering of price bids for roadway construction. The approach from the data preprocessing using fuzzy inference rules, to the power series modeling, to the model refinement approach, and to the comparison between the reserve price model and the price of award model. By comparing the errors obtained from models built with data that have not been classified and those built with data which have been classified it is found that the classification system is able to successfully reduce errors. Among the reserve price models, the error was smallest for power series model built with data classified into 4 categories. The average verification error was 2.65% and the average prediction error was 8.19%. Among the price of award models, the error was smallest for the linear model built with data classified into 3 categories. The average verification error was 8.77% and the average prediction error was 15.25%. This study provides useful suggestions for future studies which should help researchers to come up with a more appropriate forecasting model, and to assist construction firms to save costs.

ACKNOWLEDGMENTS

The authors would like to thank the National Science Council, R.O.C., for the financial support of this research under Contract Nos. NSC 97-2221-E-035-045, NSC 100-2221-E-022-013-MY2, and NSC 100-2628-E-022-002-MY2.

REFERENCES

1. Bubshalt, A. A. and Al-Gobali, K. H., "Contractor prequalification in Saudi Arabia," *Journal of Management in Engineering*, ASCE, Vol. 12, No. 2, pp. 50-54 (1996).
2. Chen, C. W., Wang, M. H. L., Liu, K. F. R., and Chen, T. H., "Application of project cash management and control for infrastructure," *Journal of Marine Science and Technology*, Vol. 18, No. 5, pp. 644-651 (2010).

3. Chen, C. Y., "Assessment of the major hazard potential of interfacial solitary waves moving over a trapezoidal obstacle on a horizontal plateau," *Natural Hazards*, Vol. 62, No. 3, pp. 841-852 (2012).
4. Chen, C. Y., "Disaster prevention and reduction for exploring teachers' technology acceptance using a virtual reality system and partial least squares techniques," *Natural Hazards*, Vol. 62, No. 3, pp. 1217-1231 (2012).
5. Chiu, C. T., *The Gaps between the Government Estimate and the Tender Awarding Value of Government Procurements for Highway Construction Works*, Master Thesis, the Institute of Transportation Engineering and Management, Feng Chia University, Taiwan (2007). (in Chinese)
6. Edum-Fotwe, F., Price, A., and Thorpe, A., "A review of financial ratio tools for contractor insolvency," *Construction Management and Economics*, Vol. 14, pp. 189-198 (1996).
7. Hsiao, F. H., Chen, C. W., Liang, Y. W., Xu, S. D., and Chiang, W. L., "T-S fuzzy controllers for nonlinear interconnected systems with multiple time delays," *IEEE Transactions on Circuits and Systems I*, Vol. 52, pp. 1883-1893 (2005).
8. Hsiao, F. H., Hwang, J. D., Chen, C. W., and Tsai, Z. R., "Robust stabilization of nonlinear multiple time-delay large-scale systems via decentralized fuzzy control," *IEEE Transactions on Fuzzy Systems*, Vol. 13, pp. 152-163 (2005).
9. Hsieh, T. Y., Wang, M. H. L., and Chen, C. W., "A case study of S-curve regression method to project control of construction management via T-S fuzzy model," *Journal of Marine Science and Technology*, Vol. 12, No. 3, pp. 209-216 (2004).
10. Hsu, W. K., "An integrated flood risk assessment model for property insurance industry in Taiwan," *Natural Hazards*, Vol. 58, No. 3, pp. 1295-1309 (2011).
11. Hsu, W. K., "Risk and uncertainty analysis in the planning stages of a risk decision-making process," *Natural Hazards*, Vol. 61, No. 3, pp. 1355-1365 (2012).
12. Klir, G. J. and Yuan, B., *Fuzzy Sets and Fuzzy Logic Theory and Applications*, Prentic-Hall, Englewood Cliffs, NJ (1995).
13. Lin, J. W., *Adaptive Algorithms for the Identification of Nonlinear Structural Systems*, Ph.D. Dissertation, Columbia University, New York, USA (2001).
14. Lin, J. W., "Kalman filter decision systems for debris flow hazard assessment," *Natural Hazards*, Vol. 60, No. 3, pp. 1255-1266 (2012).
15. Lin, J. W. "Modeling and assessment of bridge structure for seismic hazard prevention," *Natural Hazards*, Vol. 61, No. 3, pp. 1115-1126 (2012).
16. Lin, J. W., "Potential hazard analysis and risk assessment of debris flow by fuzzy modeling," *Natural Hazards*, DOI 10.1007/s11069-012-0236-z (2012).
17. Lin, J. W. and Betti, R., "On-line identification and damage detection in non-linear structural systems using a variable forgetting factor approach," *Earthquake Engineering and Structural Dynamics*, Vol. 33, No. 4, pp. 419-444 (2004).
18. Lin, M. L., "Using GIS-based spatial geocomputation from remotely sensed data for drought risk-sensitive assessment," *International Journal of Innovative Computing, Information and Control*, Vol. 7, No. 2, pp. 657-668 (2011).
19. Lin, P. H., Lin, J. W., and Hsu, P. C., "Kalman filter, multiple regression and artificial neural network for the prediction of tendering price on roadway construction," *Quarterly Journal of Construction Management*, Vol. 72, pp. 44-49 (2007). (in Chinese)
20. Lo, M. H., *A Study on the Public Construction Price of Contractor Selection of Indigenous Peoples*, Master Thesis, the Institute of Civil Engineering, National Pingtung University of Science and Technology, Taiwan (2007). (in Chinese)
21. McCaffer, R., McCaffrey, M. J., and Thorpe, A., "Predicting the tender price of buildings during early design: Method and validation," *Journal of the Operational Research Society*, Vol. 35, No. 5, pp. 415-424 (1984).
22. Ng, S. T., Cheung, S. O., Skitmore, M., and Wong, T. C. Y., "An integrated regression analysis and time series model for construction tender price index forecasting," *Construction Management and Economics*, Vol. 22, No. 5, pp. 483-493 (2004).
23. Russell, J. S. and Skibniewwski, M. J., "Decision criteria in contractor prequalification," *Journal of Management in Engineering*, ASCE, Vol. 4, No. 2, pp. 148-164 (1988).
24. Shih, B. Y., "Using Lego NXT to explore scientific literacy in disaster prevention and rescue systems," *Natural Hazards*, DOI 10.1007/s11069-012-0233-2 (2012).
25. Skitmore, M. and Drew, D., "The analysis of pre-tender building price forecasting performance: A case study," *Engineering, Construction and Architectural Management*, Vol. 10, No. 1, pp. 36-42 (2003).
26. Tarawneh, S. A., "Evaluation of pre-qualification criteria: client perspective; Jordan case study," *Journal of Applied Sciences*, Vol. 4, No. 3, pp. 354-363 (2004).
27. Tsai, C. H., "An earthquake disaster management mechanism based on risk assessment information for the tourism industry-A case study from the island of Taiwan," *Tourism Management*, Vol. 31, No. 4, pp. 470-481 (2010).
28. Tsai, C. H., "Development of a mechanism for typhoon and flood risk assessment and disaster management in the hotel industry - a case study of the Hualien area," *Scandinavian Journal of Hospitality and Tourism*, Vol. 11, No. 3, pp. 324-341 (2011).
29. Tsai, C. H., "The establishment of a rapid natural disaster risk assessment model for the tourism industry," *Tourism Management*, Vol. 32, No. 1, pp. 158-171 (2011).
30. Tseng, C. P., "A new viewpoint on risk control decision models for natural disasters," *Natural Hazards*, Vol. 59, No. 3, pp. 1715-1733 (2011).
31. Tseng, C. P., "Default risk-based probabilistic decision model for risk management and control," *Natural Hazards*, DOI 10.1007/s11069-012-0183-8 (2012).
32. Tseng, C. P., "Natural disaster management mechanisms for probabilistic earthquake loss," *Natural Hazards*, Vol. 60, No. 3, pp. 1055-1063 (2012).
33. Tseng, C. P., Chen, C. W., and Liu, F. R., "Risk control allocation model for pressure vessels and piping project," *Journal of Vibration and Control*, Vol. 18, No. 3, pp. 385-394 (2012).
34. Wang, W. J. and Chiu, C. H., "Entropy variation on the fuzzy numbers with arithmetic operations," *Fuzzy Sets and Systems*, Vol. 103, No. 3, pp. 443-456 (1999).
35. Wen, Y. K., "Methods of random vibration for inelastic structures," *Applied Mechanics Review*, Vol. 42, No. 2, pp. 39-52 (1989).
36. Wong, J. M. W. and Ng, S. T., "Forecasting construction tender price index in Hong Kong using vector error correction model," *Construction Management and Economics*, Vol. 28, No. 12, pp. 1255-1268 (2010).
37. Yang, H. C., "Potential hazard analysis from the viewpoint of flow measurement in large open-channel junctions," *Natural Hazards*, Vol. 61, No. 2, pp. 803-813 (2012).
38. Yi, C. S., Lee, J. H., and Shim, M. P., "GIS-based distributed technique for assessing economic loss from flood damage: pre-feasibility study for the Anyang Stream Basin in Korea," *Natural Hazards*, Vol. 55, No. 2, pp. 251-272 (2010).
39. Yu, M. K. W. and Ive, G., "The compilation methods of building price indices in Britain: A critical review," *Construction Management and Economics*, Vol. 26, No. 7, pp. 693-705 (2008).
40. Zadeh, L. A., "Fuzzy sets," *Information and Control*, Vol. 8, pp. 338-353 (1965).
41. Zhou, H. J., Wang, J. A., Wan, J. H., and Jia, H. C., "Resilience to natural hazards: a geographic perspective," *Natural Hazards*, Vol. 53, No. 1, pp. 21-41 (2010).