Soft Comput (2008) 12:1105–1113
DOI 10.1007/s00500-008-0278-z

ORIGINAL PAPER

Forecasting analysis by using fuzzy grey regression model for solving limited time series data

Ruey-Chyn Tsaur

Published online: 5 February 2008

Abstract The grey model GM(1,1) is a popular forecasting method when using limited time series data and is successfully applied to management and engineering applications. On the other hand, the reliability and validity of the grey model GM(1,1) have never been discussed. First, without considering other causes when using limited time series data, the forecasting of the grey model GM(1,1) is unreliable, and provide insufficient information to a decision maker. Therefore, for the sake of reliability, the fuzzy set theory was hybridized into the grey model GM(1,1). This resulted in the fuzzy grey regression model, which granulates a concept into a set with membership function, thereby obtaining a possible interval extrapolation. Second, for a newly developed product or a newly developed system, the data collected are limited and rather vague with the result that the grey model GM(1,1) is useless for solving its problem with vague or fuzzy-input values. In this paper the fuzzy grey regression model is verified to show its validity in solving crisp-input data and fuzzy-input data with limited time series data. Finally, two examples for the LCD TV demand are illustrated using the proposed models.

Keywords Grey model GM(1,1) · Forecasting · Fuzzy regression model · Fuzzy grey regression model

1 Introduction

The time series model is a popular method for the forecasting of economic, marketing, as well as social problems that require at least 50 and preferably 100 observations. However, it is sometimes impossible to collect 50 data or more for forecasting a new product demand or a new system developing in today's rapidly changing socio-economic situations. Therefore, other forecasting models have been developed to cope with the problem when collected data is limited and violates the basic assumption of normal distribution of standard statistical models. The grey model with first-order differential equation and one dependent variable model is referred to as the grey model GM(1,1) (Deng 1986, 1989) and was introduced in management and engineering applications for solving limited time series data. For example, Hsu and Wen (2000) hybridized the grey model GM(1,1) and applied it to a multi-objective programming model to forecast airline city-pair passenger traffic; Lin and Yang (2003) applied the grey model GM(1,1) to forecast the output value of Taiwan's optoelectronics industry; Hsu and Chen (2003) improved the grey model GM(1,1) to forecast the power demand of Taiwan; Chiao and Wang (2002) provided a practical method to improve the lifetime of fluorescent lamps; and Hsu (2003) applied three types of residual modification models to revise the grey model GM(1,1) and forecast the short-term demand in the integrated circuit industry. It is obvious that the grey model GM(1,1) is a good method to use in the forecasting with limited time series data. However, without considering other causes, the reliability of the forecasting point-estimations obtained from the grey model GM(1,1) are not stable when we take another sampling, and it does not provide sufficient information to a decision maker. In addition, when it comes to validity, a newly developed product or a newly developed system, the data collected are usually limited and rather vague, with the result that the validity of the grey model GM(1,1) becomes useless for solving a problem using vague or fuzzy-input values.

R.-C. Tsaur (✉)
Department of Finance, Hsuan Chuang University,
48, Hsuan Chuang Road, Hsinchu 300, Taiwan
e-mail: rctsaur@hcu.edu.tw

Zadeh (1965) introduced the concept of the fuzzy set theory, rather then use the point estimation in the conventional probability theory. Also, he used the fuzzy set theory to granulate a concept into a set with membership function thereby decreasing the amount of data required. Tanaka et al. (1982) extended this idea, and introduced the fuzzy regression in his proposal of a non-parameter approach for evaluating the relation between independent variables and dependent variables. Kim et al. (1996) showed that its forecasting error was better than that of statistical regression if the collected data were smaller by simulation and in comparison with those two models. In addition, fuzzy regression models have been applied to various problems such as forecasting and engineering. In forecasting, Watada (1992) used fuzzy regression for time series analysis; Chang (1997) showed that the fuzzy regression model could be better explained in seasonal analysis; Tseng et al. (2001) stated that Watada's method had a large forecasting error, and proposed the fuzzy ARIMA (Auto-Regressive Integrated Moving Average) method to obtain a reliable forecasting interval. The lags of the differenced series appearing in the forecasting equation are called "auto-regressive" terms, lags of the forecast errors are called "moving average" terms, and a time series which needs to be differenced in order to be made stationary is said to be an "integrated" version of a stationary series. Tsaur et al. (2002) used two independent variables of preceding periodical data and index of time to show the pattern of the seasonal variation; Tsaur (2003) showed that the fuzzy regression model could be used to forecast the demand of internet users under different product life cycles. In engineering, Chang et al. (1996) used the fuzzy linear regression model to analyze a typical ergonomics problem, and showed that these types of analyses could accurately represent the experimental data; Funga et al. (2006) used an asymmetric fuzzy linear regression to estimate the functional relationships for product planning based on quality function deployment by integrating the least-squares regression into fuzzy linear regression, and proposed a hybrid linear programming model in order to cope with the typical vagueness or imprecision of functional relationships in a product under a fuzzy environment; Abdalla and Buckley (2007) apply fuzzy Monte Carlo method to a certain fuzzy linear regression problem to estimate the best solution, and find that the best solution is a vector of triangular fuzzy numbers for the fuzzy coefficients in the model. In addition, they use a quasi-random number generator to produce random sequences of these fuzzy vectors which uniformly fill the search space.

However, for the fuzzy ARIMA model it is still necessary to collect a lot of data to derive the parameters of ARIMA (p,d,q), otherwise the model cannot work. In this study, for reliability under limited time series, the fuzzy set theory is hybridized into the grey model GM(1,1) to obtain the fuzzy grey regression model, and to granulate a concept into a set with membership function, thereby obtaining the possibly interval extrapolation. Next, the fuzzy grey regression model is verified to show its validity in solving crisp-input data or fuzzy-input data with limited time series data. Finally, two examples for LCD TV demand are illustrated by our proposed models. The rest of this paper is organized as follows. Section 2 reviews the grey model GM(1,1) and the fuzzy regression model. In Sect. 3, the fuzzy grey regression model for solving limited data is constructed. Two examples are illustrated in Sect. 4, and finally, conclusions are drawn in Sect. 5.

2 Reviewing the grey model GM(1,1) and the fuzzy regression model

2.1 Grey model GM(1,1)

If an original time series set f^0 is defined as

$$f^0 = \{ f_t^0 \,|\, t \in 1, 2, \ldots, n\} \tag{1}$$

where t denotes the number of data observed in period t, then the AGO value f_t^1 of the original time series f_t^0 is obtained as

$$f_t^1 = \left(\sum_{k=1}^{t} f_k^0 \right) \quad t = 1, 2, \ldots, n. \tag{2}$$

The grey model GM(k, N) (Deng 1986, 1989) is defined as Eq. (3) where k stands for the kth-order derivative of the dependent variables F_t^1, and N stands for N variables (i.e. one dependent variable F_t^1 and $N-1$ independent variables $X_1^1(t), X_2^1(t), \ldots, X_{N-1}^1(t)$).

$$\begin{aligned}\frac{d^k F_t^1}{dt^k} + a_1 \frac{d^{k-1} F_t^1}{dt^{k-1}} + \cdots + a_{k-1} \frac{d F_t^1}{dt} + a_k F_t^1 \\ = b_1 X_1^1(t) + b_2 X_2^1(t) + \cdots + b_{N-1} X_{N-1}^1(t)\end{aligned} \tag{3}$$

where $a_1, a_2, \ldots, a_k$ and $b_1, b_2, \ldots, b_{N-1}$ are unknown parameters. If $k = 1$ and $N = 1$, then the grey model GM(1,1) with first-order differential equation and one dependent variable model can be constructed as

$$\frac{dF_t^1}{dt} + aF_t^1 = b, \quad t = 1, 2, \ldots, n \tag{4}$$

where a represents the unknown developed parameter, b represents the unknown grey controlled parameter, and F_t^1 is the dependent variable with AGO input value f_t^1. For solving model (4), the derivative $\frac{dF_t^1}{dt}$ for the dependent variable is represented as

$$\frac{dF_t^1}{dt} = \lim_{h \to 0} \frac{F_{t+h}^1 - F_t^1}{h}, \quad \forall t \geq 1. \tag{5}$$

Because the collected data is a set of time-series, we assume the sampling time interval between period t and $t+1$ to be one unit. Then, the derivative $\frac{dF_t^1}{dt}$ can be approximated to be the difference between two successive periods of F_t^1 and F_{t+1}^1, defined as an inverse accumulated generating operation (IAGO) variable F_{t+1}^0 as

$$\frac{dF_t^1}{dt} \approx \frac{F_{t+1}^1 - F_t^1}{1} = F_{t+1}^1 - F_t^1 = F_{t+1}^0, \quad \forall t \geq 1 \tag{6}$$

for the original $(t+1)$-th time series data $f_{t+1}^0, \forall t \geq 1$. In order to have a more steady value for the dependent variable $F_t^1, \forall t \geq 1$, the second part of model (4) is suggested as the average of two successive periods of F_t^1 and $F_{t+1}^1, \forall t \geq 1$ (Deng 1986, 1989). Then, we can rewrite model (4) as

$$F_{t+1}^0 = a\left[-\frac{1}{2}(F_{t+1}^1 + F_t^1)\right] + b, \quad \forall t \geq 1. \tag{7}$$

If $t = 1, 2, \ldots, n-1$, then (7) can be rewritten into matrix form as

$$\begin{bmatrix} F_2^0 \\ F_3^0 \\ \cdots \\ F_n^0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}(F_2^1 + F_1^1) & 1 \\ -\frac{1}{2}(F_3^1 + F_2^1) & 1 \\ \cdots & . \\ -\frac{1}{2}(F_n^1 + F_{n-1}^1) & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}. \tag{8}$$

By applying the least square method with input data sets f^1 and f^0, the parameters of a and b in matrix $\hat{\mathbf{a}}$ can be solved as

$$\hat{\mathbf{a}} = \begin{bmatrix} a \\ b \end{bmatrix} = (\mathbf{B}^{\mathbf{T}}\mathbf{B})^{-\mathbf{1}}\mathbf{B}^{\mathbf{T}}\mathbf{F}^0 \tag{9}$$

where matrices $\mathbf{F^0} = \begin{bmatrix} f_2^0 \\ f_3^0 \\ \cdots \\ f_n^0 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} -\frac{1}{2}(f_2^1 + f_1^1) & 1 \\ -\frac{1}{2}(f_3^1 + f_2^1) & 1 \\ \cdots & . \\ -\frac{1}{2}(f_n^1 + f_{n-1}^1) & 1 \end{bmatrix}$ and $\mathbf{B^T}$ are the transpose of matrix $\mathbf{B}$. Then, the differential equation of model (4) can be solved to obtain the estimated value $\hat{f}_{t+1}^1$ for the dependent variable $F_t^1, \forall t \geq 1$ as

$$\hat{f}_{t+1}^1 = \left(f_1^0 - (b/a)\right) e^{-at} + (b/a), \quad \forall t \geq 1. \tag{10}$$

Finally, the estimated value $\hat{f}_{t+1}^0$ for the IAGO variable F_{t+1}^0 is obtained as

$$\hat{f}_{t+1}^0 = \hat{f}_{t+1}^1 - \hat{f}_t^1, \quad \forall t = 1, 2, \ldots, n. \tag{11}$$

Therefore, by inputting the time series $f_1^0, f_2^0, \ldots, f_n^0$ into the grey model GM(1,1), it can obtain the extrapolative value of $\hat{f}_2^0, \hat{f}_3^0, \ldots, \hat{f}_n^0$, and $\hat{f}_{n+1}^0$.

2.2 The fuzzy regression model

The fuzzy regression was first introduced by Tanaka et al. (1982). It is an alternative approach to evaluating the relationship between independent variables and the dependent variable. The basic model assumes a fuzzy regression equation as follows

$$\tilde{Y}_i = \tilde{A}_0 + \tilde{A}_1 X_{i1} + \cdots + \tilde{A}_N X_{iN} = \tilde{A} X_i \tag{12}$$

where $X_i = [1, X_{i1}, \ldots, X_{iN}]^T$ is a vector of independent variables for the i-th data; $\left[\tilde{A}_0, \tilde{A}_1, \ldots, \tilde{A}_N\right]$ is a vector of the fuzzy parameters presented in the form of symmetric triangular fuzzy numbers denoted by $\tilde{A}_j = (\alpha_j, c_j)$, $j = 0, 1, \ldots, N$, with its membership function described as (13) below where α_j is its central value and c_j is its spread value.

$$\mu_{\tilde{A}_j}(a_j) = \begin{cases} 1 - \frac{|\alpha_j - a_j|}{c_j}, & \alpha_j - c_j \leq a_j \leq \alpha_j + c_j, \\ & \forall j = 0, 1, \ldots, N \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

By applying the Extension Principle (Zadeh 1965), the derived membership function of fuzzy number $\tilde{Y}_i$ is shown as (14)

$$u(Y_i) = \begin{cases} 1 - \frac{|Y_i - X_i\alpha|}{c^T|X_i|}, & X_i \neq 0, \\ 1, & X_i = 0, Y_i \neq 0 \\ 0, & X_i = 0, Y_i = 0 \end{cases} \quad \forall i = 1, 2, \ldots, M \tag{14}$$

The aim of a fuzzy regression model is to find the narrowest fuzzy regression interval as described in (15) by requiring that the membership degree of each observation Y_i is at least equal to the value h as shown in (16) below.

$$\text{MIN} \sum_{j=0}^{N} \left[c_j \sum_{i=1}^{M} |X_{ij}|\right] \tag{15}$$

$$1 - \frac{|Y_i - X_i^T\alpha|}{c^T|X_i|} \geq h, \quad \forall i = 1, 2, \ldots, M \tag{16}$$

Therefore, the following linear programming model can be obtained for solving the fuzzy regression equation:

$$\begin{aligned} &\text{MIN} \sum_{j=0}^{N} \left[c_j \sum_{i=1}^{M} |X_{ij}|\right] \\ &s.t. \sum_{j=0}^{N} \alpha_j X_{ij} + (1-h) \sum_{j=0}^{N} c_j |X_{ij}| \geq Y_i, \quad \forall i = 1, 2, \ldots, M \\ &\sum_{j=0}^{N} \alpha_j X_{ij} - (1-h) \sum_{j=0}^{N} c_j |X_{ij}| \leq Y_i, \quad \forall i = 1, 2, \ldots, M \\ &c \geq 0, a \in \Re, X_{i0} = 1, 0 \leq h < 1; \quad \forall i = 1, 2, \ldots, M \end{aligned} \tag{17}$$

Then, formula (12) can be rewritten as

$$\tilde{Y}_i = (\alpha_0, c_0) + (\alpha_1, c_1) X_{i1} + \cdots + (\alpha_N, c_N) X_{iN} \tag{18}$$

Each value of the dependent variable can be estimated as a fuzzy number $\tilde{Y}_i = (Y_i^L, Y_i^{h=1}, Y_i^U)$, $i = 1, 2, \ldots, M$ where the lower bound of $\tilde{Y}_i$ is $Y_i^L = (\alpha - c)^T X_i$; the center value of $\tilde{Y}_i$ is $Y_i^{h=1} = \alpha^T X_i$; the upper of $\tilde{Y}_i$ is $Y_i^U = (\alpha + c)^T X_i$ and $c^T = (c_0, c_1, \ldots, c_N), \alpha^T = (\alpha_0, \alpha_1, \ldots, \alpha_N)$.

The degree of fitness of the estimated fuzzy regression equation $\tilde{Y}_i = \tilde{A}_0 + \tilde{A}_i X_i$ to the given data Y_i is measured by index h with $Y_i^h = \{Y_i \mid u_{\tilde{Y}_i}(Y_i) \geq h\}$. The value of h is a membership degree which requires that the collected data are included in the derived fuzzy regression interval at least to the degree h. Moskowitz and Kim (1993) proposed that if one is confident with the collected data, then a smaller value h is assigned, otherwise, a larger value h should be given. Besides, Moskowitze and Kim also suggested that "If the solution for a fuzzy regression model is obtained as $\tilde{A}_{j,h_1} = (\alpha_j^*, c_j^*)$, then the solution is changed into $\tilde{A}_{j,h_2} = (\alpha_j^*, \frac{1-h_1}{1-h_2} c_j^*)$ when the confidence value h is adjusted from h_1 to h_2".

3 Fuzzy grey regression model

In this section, for the purpose of reliability, we first hybridize the fuzzy set theory into the grey model GM(1,1) in Sect. 3.1 to evaluate the fuzzy relationship between dependent and independent variables, and granulate a concept into a set with membership function, thereby obtaining the possible interval extrapolation. Second, for a newly developed product or a newly developed system the data collected are limited and rather vague such that the grey model GM(1,1) is useless for solving a problem with vague or fuzzy-input values.

Therefore, the fuzzy grey regression model is verified in Sect. 3.2 to show its validity in solving fuzzy-input data under limited time series data.

3.1 Fuzzy grey regression model with crisp-input and fuzzy-output model

For a limited time series set $\mathrm{f}^0 = \{f_t^0 \mid t \in 1, 2, \ldots, n\}$, an AGO time series set $\mathrm{f}^1 = \{f_t^1 \mid t \in 1, 2, \ldots, n\}$ is obtained as $f_t^1 = \sum_{k=1}^{t} f_k^0$. In this study, in order to forecast with limited time series data, we first hybridize the fuzzy set theory into the grey model GM(1,1) to obtain the fuzzy grey regression model using limited data. Then, the fuzzy grey regression model with crisp-input and fuzzy-output value is constructed as per Definition 1 by fuzzyfying the grey model GM(1,1) of Eq. (7).

Definition 1 The fuzzy grey regression model with crisp-input and fuzzy-output value is defined as $\tilde{F}_{t+1}^0 = \tilde{A}_0 + \tilde{A}_1 F_t^1, \forall t \geq 1$, where the fuzzy parameters $\tilde{A}_0$ and $\tilde{A}_1$ are fuzzy sets on the product space of the parameter for mapping the input AGO time series f_t^1 of the independent variable F_t^1 to the fuzzy-output value of the fuzzy dependent variable $\tilde{F}_{t+1}^0$.

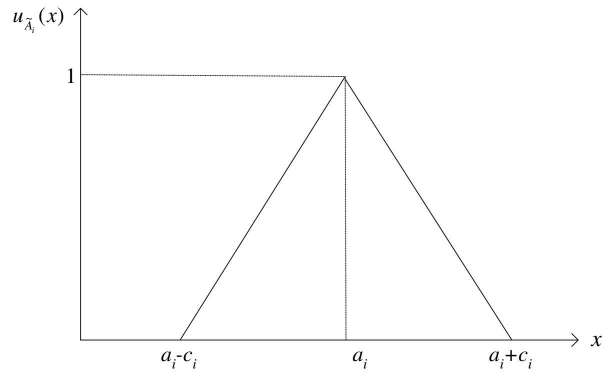

Fig. 1 The membership function of fuzzy parameter $\tilde{A}_i$

For the sake of simplicity, we define the fuzzy parameters $\tilde{A}_0 = (a_0, c_0)$ and $\tilde{A}_1 = (a_1, c_1)$ in the fuzzy grey regression model as symmetrical triangular fuzzy numbers with central values a_0, a_1, and spread values c_0, c_1, respectively. Besides, their figure and membership functions are formulated in Fig. 1 and Eq. (19)

$$u_{\tilde{A}_i}(x) = \begin{cases} 1 - \frac{(a_i - x)}{c_i}, & a_i - c_i \leq x \leq a_i \\ 1 - \frac{(x - a_i)}{c_i}, & a_i \leq x \leq a_i + c_i \\ 0, & \text{otherwise.} \end{cases}, \quad \forall i = 0, 1 \tag{19}$$

By applying the Extension Principle, the derived membership function of the fuzzy-output value of $\tilde{F}_{t+1}^0$ is shown as (20)

$$u(f_{t+1}^0) = \begin{cases} 1 - \frac{\left|f_{t+1}^0 - (a_0 + a_1 f_t^1)\right|}{c_0 + c_1 f_t^1}, & \left(a_0 + a_1 f_t^1\right) - \left(c_0 + c_1 f_t^1\right) \\ & \leq f_{t+1}^0 \leq \left(a_0 + a_1 f_t^1\right) \\ & + \left(c_0 + c_1 f_t^1\right) \\ 0, & \text{o.w.} \end{cases} \tag{20}$$

The aim of a fuzzy grey regression model is to find the narrowest regression interval by requiring that the membership degree of each observation f_{t+1}^0 is at least equal to the value h as shown in (21) below with $0 \leq h < 1$.

$$1 - \frac{\left|f_{t+1}^0 - \left(a_0 + a_1 f_t^1\right)\right|}{c_0 + c_1 f_t^1} \geq h, \quad \forall t = 1, 2, 3, \ldots, n-1. \tag{21}$$

Finally, in order to obtain the minimum fuzzy relation in the fuzzy grey regression model, it is necessary to require the spread values of the fuzzy-output value $\tilde{F}_{t+1}^0, \forall t = 1, 2, 3, \ldots, n-1$ to be as small as possible. Based on this objective, we can obtain the following linear programming model

$$\begin{aligned}
&Min \sum_{t=1}^{n-1} (c_0 + c_1 f_t^1) \\
&f_{t+1}^0 \geq a_0 - (1-h)c_0 + (a_1 - (1-h)c_1) * (f_t^1) \\
&f_{t+1}^0 \leq a_0 + (1-h)c_0 + (a_1 + (1-h)c_1) * (f_t^1) \\
&a_0, a_1 \in R, c_0, c_1 \geq 0, 0 \leq h < 1, \forall t = 1, 2, \ldots, n-1
\end{aligned} \tag{22}$$

By solving the linear programming model (22), the fuzzy parameters $\tilde{A}_0 = (a_0, c_0)$ and $\tilde{A}_1 = (a_1, c_1)$ can be solved. The equation of the fuzzy grey regression model is obtained as $\tilde{F}_{t+1}^0 = (a_0, c_0) + (a_1, c_1)F_t^1$. This equation provides a forecasting interval constructed by each estimated fuzzy output $\tilde{F}_{t+1}^0 = \left(F_{t+1}^{0,L}, F_{t+1}^{0,h_t=1}, F_{t+1}^{0,U}\right) \quad t = 1, 2, \ldots, n$ where $F_{t+1}^{0,L} = (a_0 - c_0) + (a_1 - c_1) \cdot f_t^1$ is the lower bound of $\tilde{F}_{t+1}^0$; $F_{t+1}^{0,h_t=1} = a_0 + a_1 \cdot f_t^1$ is the center value of $\tilde{F}_{t+1}^0$, and $F_{t+1}^{0,U} = (a_0 + c_0) + (a_1 + c_1) \cdot f_t^1$ is the upper bound of $\tilde{F}_{t+1}^0$.

Therefore, by the crisp-input values $f_1^0, f_2^0, \ldots, f_n^0$, the extrapolative fuzzy-output values $\hat{\tilde{F}}_2^0, \hat{\tilde{F}}_3^0, \ldots, \hat{\tilde{F}}_n^0$, and $\hat{\tilde{F}}_{n+1}^0$ can be obtained to granulate a concept into a set with membership function.

3.2 Fuzzy grey regression model with fuzzy-input and fuzzy-output model

If an original fuzzy time series set $\tilde{\mathrm{f}}^0 = \{ \tilde{f}_t^0 \mid t \in 1, 2, \ldots, n\}$ with central value f_t^0, and spread value e_t^0, $\forall t \in 1, 2, \ldots, n$ is collected, then an AGO fuzzy time series set $\tilde{\mathrm{f}}^1 = \{ \tilde{f}_t^1 \mid t \in 1, 2, \ldots, n\}$ is obtained with central value $f_t^1 = \sum_{k=1}^t f_k^0$, and spread value $e_t^1 = \sum_{k=1}^t e_k^0 \quad \forall t \in 1, 2, \ldots, n$, by the concept of fuzzy addition. Then, the fuzzy grey regression model with fuzzy-input and fuzzy-output value is constructed as per Definition 2 by fuzzyfying the grey model GM(1,1) of Eq. (7).

Definition 2 The fuzzy grey regression model with fuzzy-input and fuzzy-output value is defined as $\tilde{F}_{t+1}^0 = \tilde{A}_0 + \tilde{A}_1 \tilde{F}_t^1$,$\forall t \geq 1$, where the fuzzy parameters $\tilde{A}_0$ and $\tilde{A}_1$ are fuzzy sets on the product space of the parameter for mapping the AGO fuzzy time series $\tilde{f}_t^1$ of the fuzzy independent variable $\tilde{F}_t^1$ to the fuzzy-output value of fuzzy dependent variable $\tilde{F}_{t+1}^0$.

For the sake of simplicity, we define the fuzzy parameters $\tilde{A}_0 = (a_0, c_0)$ and $\tilde{A}_1 = (a_1, c_1)$ in the fuzzy grey regression model as symmetrical triangular fuzzy numbers with central values a_0, a_1, and spread values c_0,c_1, respectively. By interval arithmetic, as in Fig. 2, a requirement of the maximum degree of fit h between the fuzzy-output value $\tilde{F}_{t+1}^0$ and the fuzzy-input value $\tilde{f}_{t+1}^0 = (f_t^0, e_t^0)$ is presented as $[\tilde{f}_{t+1}^0]^h \subset [\tilde{F}_{t+1}^0]^h$ (Tanaka et al. 1982) with

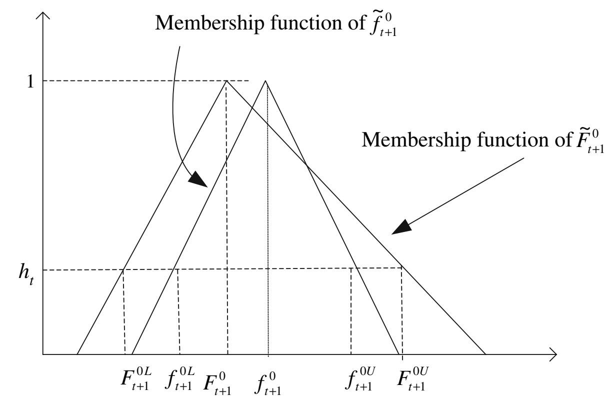

Fig. 2 The relation between membership functions of fuzzy -input and fuzzy-output values

$0 \leq h < 1, \forall t = 1, 2, 3, \ldots, n-1$. Then, we can derive

$$f_{t+1}^0 - (1-h)e_t^0 \geq a_0 - (1-h)c_0 + (a_1 - (1-h)c_1) * (f_t^1 - (1-h)e_t^1) \tag{23}$$

$$f_{t+1}^0 + (1-h)e_t^0 \leq a_0 + (1-h)c_0 + (a_1 + (1-h)c_1) * (f_t^1 + (1-h)e_t^1). \tag{24}$$

Finally, in order to obtain the minimum fuzzy relation in the fuzzy grey model GM(1,1), it is necessary to require the spread values of the fuzzy-output value $\tilde{F}_{t+1}^0$, $\forall t = 1, 2, 3, \ldots, n-1$ to be as small as possible. Based on this objective, we can obtain the following linear programming model:

$$\begin{aligned}
&\mathrm{Min} \sum_{\mathrm{t}=1}^{\mathrm{n}-1} (c_0 + a_1 e_t^1 + c_1 f_t^1) \\
&f_{t+1}^0 - (1-h)e_t^0 \geq a_0 - (1-h)c_0 \\
&\quad + (a_1 - (1-h)c_1) * (f_t^1 - (1-h)e_t^1) \\
&f_{t+1}^0 + (1-h)e_t^0 \leq a_0 + (1-h)c_0 + (a_1 + (1-h)c_1) \\
&\quad * (f_t^1 + (1-h)e_t^1) \\
&a_0, a_1 \in R, c_0, c_1 \geq 0, 0 \leq h < 1, \forall t = 1, 2, \ldots, n-1
\end{aligned} \tag{25}$$

By solving the linear programming model (25), the fuzzy parameters $\tilde{A}_0 = (a_0, c_0)$ and $\tilde{A}_1 = (a_1, c_1)$ can be solved. The equation of the fuzzy grey regression model is obtained as $\tilde{F}_{t+1}^0 = (a_0, c_0) + (a_1, c_1)\tilde{F}_t^1$. This equation provides a forecasting interval constructed by each estimated fuzzy-output value $\tilde{F}_{t+1}^0 = \left(F_{t+1}^{0,L}, F_{t+1}^{0,h_t=1}, F_{t+1}^{0,U}\right)$, $\forall t = 1, 2, \ldots, n$ where $F_{t+1}^{0,L} = (a_0 - c_0) + (a_1 - c_1) \cdot (f_t^1 - e_t^1)$ is the lower bound of $\tilde{F}_{t+1}^0$; $F_{t+1}^{0,h_t=1} = a_0 + a_1 \cdot f_t^1$ is the center value of $\tilde{F}_{t+1}^0$, and $F_{t+1}^{0,U} = (a_0 + c_0) + (a_1 + c_1) \cdot (f_t^1 + e_t^1)$ is the upper bound of $\tilde{F}_{t+1}^0$. Finally, by the fuzzy-input values $\tilde{f}_1^0, \tilde{f}_2^0, \ldots, \tilde{f}_n^0$, the extrapolative fuzzy-output values $\hat{\tilde{F}}_2^0, \hat{\tilde{F}}_3^0, \ldots, \hat{\tilde{F}}_n^0$, and $\hat{\tilde{F}}_{n+1}^0$ can be obtained. Therefore, the fuzzy grey regression model verified its validity for solving fuzzy-input data under limited time series data.

Furthermore, the value h is referred to as the degree of fit between the fuzzy-input values and the extrapolative fuzzy-output values of the fuzzy grey regression model, which thus determines the range of the possible distributions of the proposed model. That is, the value h is subjectively selected by a decision maker as a membership degree which requires that the collected data are included in the derived fuzzy interval to at least the degree h. Moskowitz and Kim (1993) proposed that if one is confident with the collected data, then a smaller value h can be assigned, otherwise a larger value h should be given. When h increases, the spread of the fuzzy-output value becomes wider. Tanaka and Watada (Tsaur et al. 2002) suggested that $h = 0$ when the data set is sufficiently large, and to use a comparatively higher h as the size of the data set becomes smaller. Bardossy et al. (1990) selected the value for h according to the decision maker's belief in the model, generally recommending an h value between 0.5 and 0.7. With the property of h, the resultant forecasting interval provides more flexibility for a decision maker.

4 Illustrated examples

Since 1998, the growth of the TFT-LCD industry in Taiwan has been more rapid than the growth in the semiconductor industry. The initial application of TFT-LCD was in the notebook, but it is now also applied to PC and TV monitors. Display Search (Optotech 2005) reported that LCD TV demand has surged worldwide, with all regions enjoying at least 41% sequential growth. Topology research institute (2005) predicted that LCD TV will have a 50% market share in the TV industry by 2008 if the LCD TV price can be reduced through lower panel costs and new capacity optimization. In order to get a larger market share, TFT-LCD companies have been forecasting future demand in the LCD TV market and planning a new generation of production line for large size panels. However, it is difficult to collect the world-wide sales of different sizes of LCD TVs so we roughly collected the secondary data from some research institutes. Because the collected time series data is limited and the secondary data collected from different research institutes is vague, the fuzzy grey regression model as proposed in Sect. 3 is a good method to forecast the future demand of LCD TV for the time period 2002 to 2005.

4.1 Example 1 for crisp-input LCD TV demand

The secondary data of LCD TV demand with crisp-input value is shown in the second row of Table 1. By inputting the values of LCD TV sales into the model (22) and setting the value h as 0, 0.3, 0.5, and 0.7, respectively, as per the suggestions of Tanaka et al, and Bardossy et al, the solutions of the fuzzy parameters are obtained in Table 2. Next, the extrapolative fuzzy-output LCD TV demand is estimated and shown in Table 3, with a lower and upper value for each year.

Table 1 The sales of LCD TVs per year (unit: ten thousand)

Year	2001	2002	2003	2004
LCD TV Sales	81	150	393	870

Table 2 The solutions for the variables

Variable	a_0	a_1	c_0	c_1
Solution with value $h = 0$	64.65	1.33	22.05	0
Solution with value $h = 0.3$	64.65	1.33	31.5	0
Solution with value $h = 0.5$	64.65	1.33	44.11	0
Solution with value $h = 0.7$	64.65	1.33	73.52	0

A better forecasting model means that it has a smaller estimated error and can use the training data to obtain a more accurate extrapolative value (compared to the testing value). The central value of the fuzzy-output LCD TV demand is used for defuzzying the fuzzy-output values of column 1 and column 3 in Table 3. Then we find that the fuzzy grey regression model has the smallest forecasting error among the grey model GM(1,1), Watada's model, and the linear regression model, with a estimated error MAPE (Mean Absolute Percentage Error, MAPE) of 29.44% from year 2002 to 2004. In addition, the testing data of the year 2005 for the LCD TV demand shows that the proposed model extrapolates the LCD TV demand in 2005 to be between 20,296,200 units (lower value under $h = 0$) and 20,737,200 units (upper value under $h = 0$), which matches the possible demand forecasting of some research institutes. For example, it is believed that the LCD TV demand in 2005 was more than 20,000,000 units. The comparison results are plotted in Figs. 3 and 4.

In Fig. 3, the linear regression model forecasting value for the 2005 LCD TV demand is 10,260,000, and the grey model GM(1,1) forecasting value for the 2005 LCD TV demand is 16,698,000 which does not fit the rapid increasing LCD TV demand in 2005. In Fig. 4, based on the empirical results of this example, we found that the predictive capability of the fuzzy grey regression model is rather encouraging and that the possible interval of the fuzzy grey regression model is narrower than Watada's interval. For example, for the forecasting demand in 2002, the lower bound and upper bound of the proposed model are 1,503,300 and 1,944,300 whereas the lower bound and upper bound in Watada's model are 1,261,824 and 3,399,960, a range that is too wide for a decision maker to make a decision on under such an uncertain environment. We also found that Watada's model is a fuzzy linear regression model that cannot fit the rapid increasing LCD TV demand. However, our proposed model is a piecewise model which can more reliably forecast a rapidly

Table 3 The extrapolative value $\hat{\hat{F}}_t^0$ with different models (unit: ten thousand)

Forecasting value	Fuzzy grey regression model with value h=0	Grey model GM(1,1)	Watada's Fuzzy regression model with value h=0	Linear regression model
$\hat{\hat{F}}_{2002}^0$	(150.33,194.43)	156.13	(126.1824, 339.996)	243
$\hat{\hat{F}}_{2003}^0$	(349.83,393,93)	343.99	(389.0736, 602.994)	504
$\hat{\hat{F}}_{2004}^0$	(872.52,916.62)	757.89	(651.9648, 865.992)	765
$\hat{\hat{F}}_{2005}^0$	(2029.62, 2073.72)	1669.80	(914.856, 1128.99)	1026
MAPE	23.12%	29.44%	94.37%	101.16%

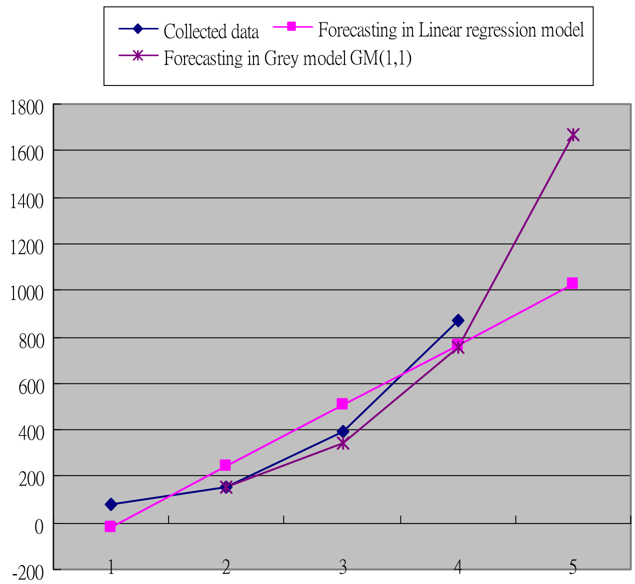

Fig. 3 Comparison results between linear regression model and grey model GM(1,1)

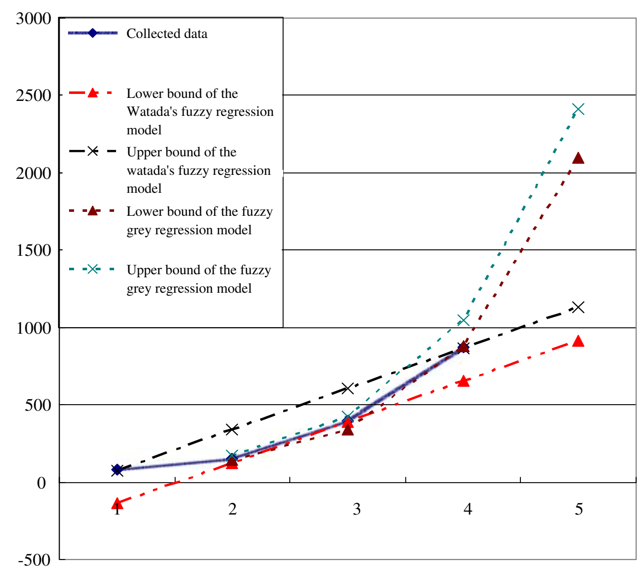

Fig. 4 Comparison results between fuzzy regression model and fuzzy grey regression model

Table 4 The sales of LCD TVs per year (unit: ten thousand)

Year	2001	2002	2003	2004
LCD TV Sales	(81, 8)	(150, 12)	(393, 16)	(870, 20)

increasing demand. These evidences show that the performance of the fuzzy grey regression model is better than that of the grey model GM(1,1), the linear regression model and Watada's fuzzy regression model under limited information.

4.2 Example 2 for the fuzzy-input LCD TV demand

In this example, we revised the collected secondary data in Table 1 into fuzzy data by adding the spread value, as shown in Table 4. For example, the LCD TV demand in 2001 is listed as (81, 8) which means that the central value is 81 and the spread value is 8 to express fuzzy number $\tilde{81}$. We know that the grey model GM(1,1) and the statistical regression model are not used for analyzing fuzzy-input data. In order to analyze the fuzzy-input data and transfer to useable information, we use the fuzzy grey regression model for forecasting the LCD TV demand, and compare to Watada's model. By inputting the fuzzy-input values of the LCD TV sales into the model (25) and by setting the value h as 0, 0.3, 0.5, and 0.7, respectively, as per the suggestions of Tanaka et al, and Bardossy et al, then the solutions of the fuzzy parameters can be obtained in Table 5. Next, the LCD TV demand can be estimated, as shown in Table 6, as a fuzzy-output value with a lower and upper value for each year. The results show that the proposed model extrapolates the LCD TV demand in 2005 to be between 20,251,110 units (lower value at $h = 0$) and 22,048,610 units (upper value at $h = 0$), which matches the possible demand forecasting of some research institutes. Figure 5 shows that Watada's fuzzy regression model has a larger forecasting interval than the fuzzy grey regression interval, and Watada's model remains a linear trend capturing the future demand, while the fuzzy grey regression model is a piecewise trend that fits the rapidly increasing LCD TV

Table 5 The solutions for the variables

Variable	a_0	a_1	c_0	c_1
Solution with value $h = 0$	49.5311	1.3825	12.4553	0
Solution with value $h = 0.3$	40.2435	1.4239	0	0.0616
Solution with value $h = 0.5$	39.7807	1.4262	0	0.0727
Solution with value $h = 0.7$	40.8943	1.4241	0	0.0818

Table 6 The extrapolative values $\hat{\tilde{F}}_t^0$ with different models (unit: ten thousand)

Forecasting value	Fuzzy grey regression model with value h=0	Tanaka's Fuzzy regression model with value h=0
$\hat{\tilde{F}}_{2002}^0$	(137.9983, 185.0289)	(110.1544, 356)
$\hat{\tilde{F}}_{2003}^0$	(328.7833, 408.9939)	(377.0316, 623)
$\hat{\tilde{F}}_{2004}^0$	(849.9858, 974.4364)	(643.9088, 890)
$\hat{\tilde{F}}_{2005}^0$	(2025.111, 2204.861)	(910.786, 1157)

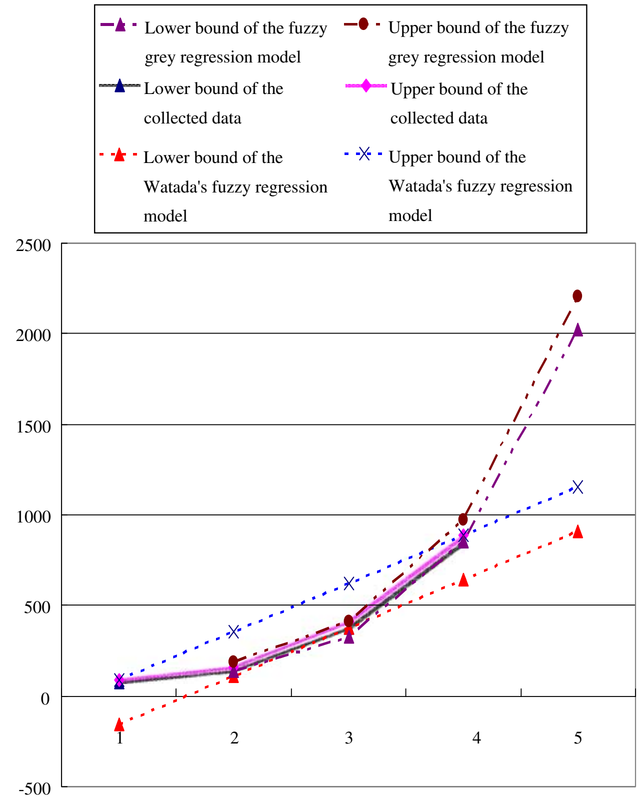

Fig. 5 Comparison results between Watada's fuzzy regression model and fuzzy grey regression model

demand. For example, in the forecasting demand of 2002, the lower bound and upper bound of the proposed model were 1,379,983 and 1,850,289, while the lower bound and upper bound in Watada's model was 1,101,544 and 3,560,000 which is too wide for a decision maker to make a decision on under an uncertain environment. In addition, as per the forecasting demand in some research institutes, the forecasting interval (20,251,110, 22,048,610) in the fuzzy grey regression model is more precise forecasting for the LCD TV demand in 2005 than Watada's fuzzy regression model.

4.3 Discussions and analysis

It is evident that the reliability and validity of the fuzzy grey regression model were examined successfully. By using the collected limited time series data, we obtained the possible forecasting interval for the LCD TV demand. In addition, we found that when the collected LCD TV demand is fuzzy-input time series data, then the proposed fuzzy grey regression model can still be used for forecasting, and the forecasting intervals are narrower allowing the decision maker to easily determine the trend of future LCD TV demand. In short, our proposed method can be used to assist managers in the LCD TV industry to understand the possible interval in the macro-economic environment.

Although the basic concept of the grey model GM(1,1) and the fuzzy set theory is used to formulate the fuzzy grey regression model, the output of the fuzzy grey regression model requires fewer observations than the linear regression model. There are several situations for which the fuzzy grey regression model appears to be the most appropriate tool, such as:

(i) Fuzzy grey regression model can provide the decision makers the best- and worst-possible situations.
(ii) The required observations are as little as four.

A comparison of four kinds of time-series methods is shown in Table 7.

5 Conclusion

In this paper, based on the basic concepts of the grey model GM(1,1) and Tanaka's fuzzy regression model, we proposed a new method (the fuzzy grey regression model) and applied it to forecasting the LCD TV demand for showing the reliability and validity of the proposed method. From the examples it is evident that the proposed method not only makes good forecasts but also provides the decision makers with the best and worst-possible scenarios. The performance of the fuzzy grey regression model is better than the linear regression model, the grey model GM(1,1) and Watada's fuzzy regression model. Although Display Search, iSuppli, and PIPA research institutes collected sufficient information from the LCD TV market, in general, they could have made more precise forecasts. Using limited data with uncertain information, and by using the fuzzy grey regression model we were able to

Table 7 Comparison of four kinds of methods

Grey Model GM(1,1)	Linear Regression Model	Watada's Fuzzy Regression Model	Fuzzy Grey Regression Model
Apply a one-order AGO to assemble the collected time series data and obtain internal regularity in order to manage the disorganized original data	A functional relationship between independent variables and the dependent variable.	The relationship between the independent variables and the dependent variable is a fuzzy function	The relationship between the input AGO variable and the output variable is a fuzzy function
Point estimation.	Provides confidence interval.	Provides possibility interval.	Provides possibility interval.
Limited time series data	Under the assumption of Normal distribution, at least 20 or more observations are preferable	This model is seldom used to solve limited time series data	Limited time series data

forecast the LCD TV demand with crisp-input or fuzzy-input data, such that the TFT-LCD producers can use it to plan their new generation plant expansion and strengthen their competitive edge, meet new and ongoing challenges, and maximize their profit in the LCD TV market.

Acknowledgments The authors acknowledge the financial support from the National Science Council of Taiwan, R.O.C., under project number NSC 93-2213-E-364-002.

References

Abdalla A, Buckley JJ (2007) Monte Carlo methods in fuzzy linear regression. Soft Computing 11(10):991–996

Bardossy A, Bogardi I, Duckstein L (1990) Fuzzy Regression in Hydrology. Water Resources Research 26(7):1497–1508

Chang PT (1997) Fuzzy seasonality forecasting. Fuzzy Sets and Systems 90(1):1–10

Chang PT, Lee ES, Konz SA (1996) Applying fuzzy linear regression to VDT legibility. Fuzzy Sets and Systems 80(2):197–204

Chang SC (2005) The TFT–LCD industry in Taiwan: competitive advantages and future developments. Technology in Society 27(2): 199–215

Cheng CB (2005) Fuzzy process control: construction of control charts with fuzzy numbers. Fuzzy Sets and Systems 154(2):287–303

Chiao CH, Wang WY (2002) Reliability improvement of fluorescent lamp using grey forecasting model. Microelectronics Reliability 42(1):127–134

Deng LJ (1986) Grey forecasting and decision. Huazhong of Society and Technology Press, Wuhan

Deng LJ (1989) Introduction to grey system theory. The Journal of Grey System 1(1):1–24

Funga RYK, Chen Y, Tang J (2006) Estimating the functional relationships for quality function deployment under uncertainties. Fuzzy Sets and Systems 157(1):98–120

Hsu CI, Wen YH (2000) Application of grey theory and mutiobjective programming towards airline network design. European Journal of Operational Research 127(1):44–68

Hsu CC, Chen CY (2003) Applications of improved grey prediction model for power demand forecasting. Energy Conversion and Management 44(14):2241–2249

Hsu LC (2003) Applying the grey prediction model to the global integrated circuit industry. Technological Forecasting & Social Change 70(6):563–574

Kim KJ, Moskowitz H, Koksalan M (1996) Fuzzy versus statistical linear regression. European Journal of Operational Research 92(2):417–434

Lin CT, Yang SY (2003) Forecast of the output value if Taiwan's optoelectronics industry using the grey forecasting model. Technological Forecasting & Social Change 70(2):177–186

Menozzi M, Napflin U, Krueger H (1999) CRT versus LCD: A pilot study on visual performance and suitability of two display technologies for use in office work. Displays 20(1):3–10

Moskowitz H, Kim K (1993) On assessing the h value in Fuzzy Linear Regression. Fuzzy Sets and Systems 58(3):303–327

Tanaka H, Uejima S, Asai K (1982) Fuzzy linear regression model. IEEE Transaction on SMC 12(6):903–907

Tsaur RC, Wang HF, O Yang JC (2002) Fuzzy Regression for Seasonal Time Series Analysis. International Journal of Information Technology & Decision Making 1(1):165–175

Tsaur RC (2003) Extrapolating internet users in Taiwan by risk assessment. Computers and Mathematics with Applications 46(10–11): 1725–1734

Tseng FM, Tzeng GH, Yu HC, Yuan Benjamin JC (2001) Fuzzy ARIMA model for forecasting the foreign exchange market. Fuzzy Sets and Systems 118(1):9–19

Watada J (1992) Fuzzy time series analysis and forecasting of sales volume. In: Kacprzyk J, Fedrizzi M (eds.), Fuzzy Regression Analysis, Omnitech Press, Warsaw and Physica-Verlag, Heidelberg 211–227

Zadeh LA (1965) Fuzzy sets, Inform. Control l.8(3):338–353

Optotech, Chinese Information Co. LTD, Taipei, *n*.74 (2005) (In Chinese)

Topology research institute, Open a new world of LCD : Trend Analysis in LCD TV market, Topology, Taipei (2005) (In Chinese)