A New Prediction Method for Short-Term Forecast

Bao Rong Chang
Dept. of Electrical Engineering, Cheng Shiu Institute of Technology, 840 Cheng-Ching Rd., Neau-Song Hsiang, Kaoshiung, Taiwan
e-mail: brchang@cc.csit.edu.tw

Shiou Fen Tsai
Dept. of International Trades, Kao Yuan Institute of Technology, 1821 Chung-Shan Rd., Lu-Chu Hsiang, Kaoshiung, Taiwan
e-mail: sftsai@cc.kyit.edu.tw

ABSTRACT

This study introduced a new method for the short-term forecast that actually is a hybrid model, combing the grey model and the cumulative least squared linear model, with the special feature of automatically adjusting the overestimated or underestimated predicted value around the points having extreme value. The verification of this study is also tested successfully in two experiments. This demonstrated that the proposed method has the best accuracy of predicted value among four short-term forecasting models discussed in this study.

Keywords: Grey model, Cumulative least squared linear model

1. Introduction

The methods of regression, moving average, and exponential smoothing [1] in statistics are traditionally employed for the long-term data fitting or the future trend forecasting. However, in order to fit the sampled data, some of restrictions are required to get the quite many data and to assume sampled data distributed normally for establishing the models just mentioned above. In contrast for the short-term forecasting applications, the GM(1,1| α) model ,Simple exponential, Holt's exponential, Winter's exponential, Causal regression, Time series depression, and Box Jenkins. [2] are of popular models for being widely applicable to the variety of forecasting topics, but they still encountered some shortcomings about the predictive generalization [8]. However, the GM(1,1|α) model [3] just need a few data to construct a forecasting model, and the output to the model clearly is a simple exponential function. This implies that this kind of grey prediction model meets the conditions — (i) simple and (ii) more accurate in the predicted results. Thus, the GM(1,1|α) model is often employed to the short-term forecast for many applications in recent years. Although the GM(1,1|α) model equipped the advantage of simple and fast to predict the future output, the precision limitation is also still arguable in many papers [4][5] since it get a drawback about the predicted output with overshooting at turning points. Another 3 points cumulative least squared linear model introduced in this study can achieved the pretty good results on the predicted values. Unfortunately, this model does also run into an undershooting situation around the predicted output with the extreme value. Therefore, a new method for the short-term prediction proposed in this study introduces a compromise algorithm to offset the problem of overshooting and undershooting predicted results to improve the prediction accuracy. A data preprocess, called accumulated generating operation (AGO) [3], works into the grey prediction model and the cumulative least squared linear model. The aim of this AGO is try to smooth the original given data to analogue an exponential data distribution for easily matching exponential or linear form under the few input data provided. In this study, the most recent four actual values is considered as a set of input data used to predict the next desired value. As the next desired value is observed, the first value of the current input data set is discarded and joins this latest desired value into the input data set at the last place in the order of a update data sequence for keeping four input value in a data set to ready for next prediction procedure.

2. Prediction Models

2.1 Grey Prediction Model

A prototype of grey prediction model GM(1,1| α) is introduced in the grey system theory 1982 [3].

Step 1: accumulated generating operation once (1-AGO)

$$x^{(1)}(k)=\sum_{j=1}^{k}x^{(0)}(j),\quad k=1,2,\ldots,n \tag{1}$$

$x^{(0)}(k)$: the original sampled data that is a nonnegative sequence

Step 2: finding developing coefficient and control coefficient by using grey difference equation

$$x^{(0)}(k)+az^{(1)}(k)=b,\ \ k=2,3,\ldots,n \tag{2}$$

$$z^{(1)}(k)=\alpha x^{(1)}(k)+(1-\alpha)x^{(1)}(k-1),\ \ 0\le\alpha\le 1 \tag{3}$$

$z^{(1)}(k)$: the background value

$$a=\frac{\sum_{k=2}^{n}x^{(0)}(k)\sum_{k=2}^{n}z^{(1)}(k)-(n-1)\sum_{k=2}^{n}x^{(0)}(k)z^{(1)}(k)}{(n-1)\sum_{k=2}^{n}z^{(1)}(k)^2-(\sum_{k=2}^{n}z^{(1)}(k))^2}=\frac{\Delta_a}{\Delta} \tag{4}$$

$$b=\frac{\sum_{k=2}^{n}x^{(0)}(k)\sum_{k=2}^{n}z^{(1)}(k)^2-\sum_{k=2}^{n}z^{(1)}(k)\sum_{k=2}^{n}x^{(0)}(k)z^{(1)}(k)}{(n-1)\sum_{k=2}^{n}z^{(1)}(k)^2-(\sum_{k=2}^{n}z^{(1)}(k))^2}=\frac{\Delta_b}{\Delta} \tag{5}$$

Step 3: solving the predicted value through the grey differential equation

$$\frac{dx^{(1)}(k)}{dk}+ax^{(1)}(k)=b \tag{6}$$

$$\hat{x}^{(0)}(k)=(x^{(0)}(1)-\frac{b}{a})(e^{-a(k-1)}-e^{-a(k-2)}),\ \ k=2,3,\ldots \tag{7}$$

2.2 Least Squared Polynomial Model

The least squared polynomial [6] is applied to study the statistic relation between a set of independent variables and a dependent variable. This model can be utilized for estimating or predicting the future output. Some of phenomena in the real world can be realized to be a multivariate model so that the accuracy of estimated (or predicted) would be improved. **Step 1:** Building least squared polynomial model generally in the following way:

$$\hat{y}^{(0)}(i)=b_0+b_1x^{(0)}(i)+b_2{x^{(0)}}^2(i)+\cdots\cdots+b_k{x^{(0)}}^k(i) \tag{8}$$

In order to solve the coefficients $b_0,b_1,\cdots\cdots,b_k$ in Eq. (8), the least square method [6] is employed so as to minimize the sum of square of the residual error in Eq. (9) expressed below:

$$\min.\quad q=\sum_{i=1}^{n}\varepsilon_i^2=\sum_{i=1}^{n}(y^{(0)}(i)-\hat{y}^{(0)}(i))^2 \tag{9}$$

$$s.t.\quad \hat{y}^{(0)}(i)=b_0+b_1x^{(0)}(i)+b_2{x^{(0)}}^2(i)+\cdots\cdots+b_k{x^{(0)}}^k(i)$$

The least square method [6] is again used to solve the best approximation solution for x to the equation of

$$X_a B = Y \ . \tag{10}$$

According to the definition of the following vectors:

$$X_a = \begin{bmatrix} X_a(1)^T \\ X_a(2)^T \\ \cdot \\ \cdot \\ \cdot \\ X_a(n)^T \end{bmatrix}$$: Matrix X_a

$X_a(i) = [1, x^{(0)}(i), x^{(0)^2}(i), ..., x^{(0)^k}(i)]^T$: Augment vector $X_a(i)$

$Y = [y^{(0)}(1), y^{(0)}(2), ..., y^{(0)}(k)]^T$: Vector Y

$B = [b_0, b_1, b_2, ..., b_k]^T$: Vector B

Step 2: Derived the normal equation to find pseudo inverse matrix

Solving for Eq. (10) typically turns out to be a normal equation [7],

$$X_a^T X_a B = X_a^T Y \tag{11}$$

, in which matrix B is a coefficient vector for $b_0, b_1, \cdots, b_k$ in Eq. (8) and Y is observed values given by in Eq. (8).

Step 3: Solving the appropriate coefficients and Predicting the next output

The solution to B in the normal equation is equal to $X_a^+ Y$ where X_a^+ is a pseudo inverse [7] of matrix X_a defined as $(X_a^T X_a)^{-1} X_a^T$.

$$B = (X_a^T X_a)^{-1} X_a^T Y = X_a^+ Y \tag{12}$$

$$\hat{y}^{(0)}(n+1) = X_a(n+1) B \ . \tag{13}$$

2.3 Cumulative Least Squared Polynomial Model

The cumulative least squared polynomial model is constructed as follows:

Step 1: accumulated generating operation once (1-AGO)

$$y^{(1)}(k) = \sum_{j=1}^{k} y^{(0)}(j), \quad k = 1,2,...,n \tag{14}$$

$y^{(0)}(k)$: the original sampled data that is a nonnegative sequence.

Step 2: Building least squared polynomial model generally in the following way:

$$\hat{y}^{(1)}(i) = b_0 + b_1 x^{(0)}(i) + b_2 x^{(0)^2}(i) + \cdots + b_k x^{(0)^k}(i) \tag{15}$$

In order to solve the coefficients $b_0, b_1, \cdots, b_k$ in Eq. (15), the least square method [6] is employed so as to minimize the sum of square of the residual error in Eq. (16) expressed below:

$$\text{min.} \quad q = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y^{(1)}(i) - \hat{y}^{(1)}(i))^2 \tag{16}$$

$$s.t. \quad \hat{y}^{(1)}(i) = b_0 + b_1 x^{(0)}(i) + b_2 x^{(0)^2}(i) + \cdots + b_k x^{(0)^k}(i)$$

The least square method [6] is again used to solve the best approximation solution for x to the equation of

$$X_a B = Y \ . \tag{17}$$

According to the definition of the following vectors:

$$X_a = \begin{bmatrix} X_a(1)^T \\ X_a(2)^T \\ \cdot \\ \cdot \\ \cdot \\ X_a(n)^T \end{bmatrix}$$: Matrix X_a

$X_a(i) = [1, x^{(0)}(i), x^{(0)^2}(i), ..., x^{(0)^k}(i)]^T$: Augment vector $X_a(i)$

$Y = [y^{(1)}(1), y^{(1)}(2), ..., y^{(1)}(k)]^T$: Vector Y

$B = [b_0, b_1, b_2, ..., b_k]^T$: Vector B

Step 3: Derived the normal equation to find pseudo inverse matrix

Solving for Eq. (17) typically turns out to be a normal equation [7],

$$X_a^T X_a B = X_a^T Y \tag{18}$$

, in which matrix B is a coefficient vector for $b_0, b_1, \cdots, b_k$ in Eq. (15) and Y is observed values given by in Eq. (15).

Step 4: Solving the appropriate coefficients and Predicting the next output

The solution to B in the normal equation is equal to $X_a^+ Y$ where X_a^+ is a pseudo inverse [7] of matrix X_a defined as $(X_a^T X_a)^{-1} X_a^T$.

$$B = (X_a^T X_a)^{-1} X_a^T Y = X_a^+ Y \tag{19}$$

$$\hat{y}^{(1)}(n+1) = X_a(n+1)B . \tag{20}$$

$$\hat{y}^{(0)}(n+1) = \hat{y}^{(1)}(n+1) - \hat{y}^{(1)}(n) . \tag{21}$$

3. A New Prediction Method

According to the analysis mentioned in [8][9], decreasing the number of sampling points as possible as we can, and lessening the effect of the magnitude of original data can lower the residual error of GM(1,1|α) model. Thus, using a few sampled points for GM(1,1|α) prediction would achieve the better prediction accuracy. This imply that this kind of GM(1,1|α) model is therefore applicable for short-term forecasting application. Next, how to alleviate the effect of the magnitude of the original given data so as to reduce the residual error of GM(1,1|α) model is another crucial issue [8][9]. Based on the phenomena discovered in [8][9], the prediction of GM(1,1|α) model is always to reveal an overshooting around the turning points since the extreme magnitude (too high or too low) happens there as shown in Figure 2. Accordingly, the predicted value from the grey prediction model will turn out to be an overestimated (or underestimated) result at the position of turning points. However, a cumulative least squared linear model using the most recent 3 sampled points is introduced herein to compromise the problem in GM(1,1|α) so as for lessening the effect of the magnitude of the original given data. This 3 points cumulative least squared linear model can be set up in the following steps.

Step 1: accumulated generating operation once (1-AGO)

$$x^{(1)}(i) = \sum_{j=k-3}^{k-4+i} x^{(0)}(j), \quad i = 1,2,3 \tag{22}$$

$x^{(0)}(j)$: three successive sampled data $x^{(0)}(k-3)$, $x^{(0)}(k-2)$, and $x^{(0)}(k-1)$ before the next predicted point $\hat{x}^{(0)}(k)$.

Step 2: finding a linear approximate polynomial for fitting three successive sampled data $x^{(1)}(1)$, $x^{(1)}(2)$, and $x^{(1)}(3)$

$$x^{(1)}(k) = c_1 k + c_0 , \quad k = 1,2,3 \tag{23}$$

That is,

$$X = KC , \quad C = (K^T K)^{-1} K^T X \tag{24}$$

where $X = [x^{(1)}(1), x^{(1)}(2), x^{(1)}(3)]^T$,

$K = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix}$, and $C = [c_1, c_0]^T$

Step 3: obtaining a predicted value from the linear approximate polynomial

$$\tilde{x}^{(1)}(k+1) = c_1(k+1) + c_0 , \quad k = 3 \tag{25}$$

Step 4: inverse accumulated generating operation once (1-IAGO)

$$\tilde{x}^{(0)}(k+1) = \tilde{x}^{(1)}(k+1) - x^{(1)}(k) , \quad k = 3 \tag{26}$$

This cumulative least squared linear model have the problem about the undershooting around turning points, as shown in Figure 2, that is conversely to the situation happened to GM(1,1|α) model. Therefore, we can apply this characteristic to offset the magnitude overshooting such that alleviating the effect of the magnitude of the original given data for GM(1,1|α) prediction can be achieved. A 3 points cumulative least squared linear model combining with GM(1,1|α) model thus is exploited for the prediction as follows.

$$\breve{x}^{(0)}(k) = w_1 \hat{x}^{(0)}(k) + w_2 \tilde{x}^{(0)}(k) , \tag{27}$$

$$w_1 + w_2 = 1$$

In Eq. (27), $\hat{x}^{(0)}(k)$ and $\tilde{x}^{(0)}(k)$ stand for the predicted value of a grey model and the predicted value of a 3 points cumulative least squared linear model, respectively; moreover, the w_1 and w_2 represent the weight of $\hat{x}^{(0)}(k)$ and $\tilde{x}^{(0)}(k)$, respectively.

The value of w_1 or w_2 can be evaluated by a weighting algorithm shown below where $\bar{x}^{(0)}(k)$ represents the predicted output from a least squared linear model using 4 sampled data without cumulative data preprocess.

Input Parameters:

e(1)= $x^{(0)}(k-3)-x^{(0)}(k-4)$

e(2)= $x^{(0)}(k-2)-x^{(0)}(k-3)$

e(3)= $x^{(0)}(k-1)-x^{(0)}(k-2)$

m2= $(x^{(0)}(k-1)+x^{(0)}(k-2))/2$

m3= $(x^{(0)}(k-1)+x^{(0)}(k-2)+x^{(0)}(k-3))/3$

m4= $(x^{(0)}(k-1)+x^{(0)}(k-2)+x^{(0)}(k-3)+x^{(0)}(k-4))/4$

prdvalue1= $\hat{x}^{(0)}(k)$

prdvalue2= $\tilde{x}^{(0)}(k)$

refvalue= $\bar{x}^{(0)}(k)$

Output Data: w_1 and w_2

Algorithm:

```
if refvalue is very close to prdvalue2
        t=(prdvalue1+prdvalue2)/2;
elseif refvalue is not located in validation region
    if e(3)* e(2)>0 & e(2)*e(1)>0
        t=m4;
    else
        t=m2;
    end
else
    if e(p-1)* e(p-2)<0 & e(p-2)*e(p-3)<0
        t=(refvalue+m4)/2;
    elseif e(p-1)* e(p-2)<0 & e(p-2)*e(p-3)>0
        t=(refvalue+m4)/2;
    elseif e(p-1)* e(p-2)>0 & e(p-2)*e(p-3)<0
        t=(refvalue+m3f)/2;
    else
        t=(prdvalue1+prdvalue2)/2;
    end
end
q1=abs(t-prdvalue1);
q2=abs(t-prdvalue2);
w1=q2/(q1+q2);
w2=q1/(q1+q2);
#
```

4. Experimental Results

As shown in Figure 1 to Figure 8, the predicted sequence 1 indicates the predicted results of the new prediction method, the predicted sequence 2 represents the predicted results of the GM(1,1|α) model, the predicted sequence 3 stands for the predicted results of a 3 points least squared linear model, and the predicted sequence 4 is denoted by the predicted results of a 4 points least squared linear model.

4.1 Indexes of stock price

The stock price index prediction for four countries (U.S.A. New York Dow Jones, Taiwan TAIEX, Japan Nikkei Index, and South Korea-Stock Index) [10] have been experimented as shown in Figure 1 to Figure 4. Their accuracy of four prediction methods, which are a grey model, a 3 points cumulative least squared linear model, a 4 points least squared linear model without cumulative data preprocess, the proposed method, is also compared and the summary of this experiment is listed in Table 1.

4.2 National Economy Growth Rate

The national economy growth rate prediction for four countries (U.S.A., Taiwan, Japan, and South Korea) [11] are also tested and their results are demonstrated in Figure 5 to Figure 8. The comparison of the accuracy for four prediction methods, which are a grey model, a 3 points cumulative least squared linear model, a 4 points least squared linear model without cumulative data preprocess, the proposed method, is also made and the brief of this test is listed in Table 2.

5. Conclusions

This study introduced a new method for the short-term forecast that actually is a hybrid model, combing the grey model and cumulative least squared linear model, with the special feature of automatically adjusting the overestimated or underestimated predicted value around the points having extreme value. We summarized the following statements for this study.

As shown in Figure 1 to Figure 8, The proposed method in fact has the advantage of compromising the crucial problem of overshooting occurred in the grey model and the severe issue of undershooting happened to the 3 points cumulative least squared linear model.

According to Table 1 and Table 2, we can conclude that the mean square error of the predicted results in the proposed method is less than that of the other three models. This implies that the proposed method is the best choice for the short-term forecasting application.

6. References

[1] J. Neter, M. H. Kutner, C. J. Nachtsheim, and W. Wasserman, Applied Linear Regression Models, 3rd Ed., Chicago: Irwin, 1996.

[2] S. Haykin, “Adaptive Filter Theory,” *3rd Edition*, New Jersey: Prentice Hall, 1996.

[3] J. L. Deng, “Introduction to Grey System Theory,” *Journal of Grey System*, Vol. 1, No. 1, pp. 1-24, 1989.

[4] K. L. Wen, T.C. Chang, H. T. Chang, and M. L. You, “The Adaptive α in GM(1,1) model,” *IEEE SMC International Conference*, 1999, pp. 304-308.

[5] C. K. Chen, T. l and Tien, “A New Method of System Parameter Identification for Grey Model GM(1,1),” *Journal of Grey System*, Vol. 4, No. 1, pp. 321-330, 1996.

[6] O. Bretscher, Linear Algebra with Applications, New Jersey: Prentice Hall, 1997.

[7] R. A. Johnson, Miller & Freund’s Probability & Statistics for Engineers, 6th edition, Prentice Hall, New Jersey, 2000.

[8] B. R. Chang, “A Profile of Residual Error Fluctuation for GM(1,1| α) Grey Prediction Model”, *IJFS*, Aug. 2001. (submitted)

[9] B. R. Chang, “A Mathematical Analysis of Residual Error Reduction for GM(1,1,α) Grey Prediction Model”, *2001 Symposium of Grey System Theory and Applications*, National Yulin University of Technology and Science , Taiwan, Oct. 25-26, 2001.

[10] *The capitalization weighted stock index* (TAIEX) from January 1998 to June 2001, Taiwan Stock Exchange Cooperation (TSEC), Taipei, Taiwan, 2001.

[11] National Economy Growth Rate Report from 1984 to 2000, Ministry of Economic Affairs, Taiwan, R.O.C, 2001.

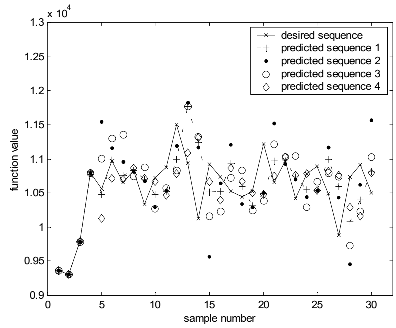

Figure 1. Forecasting results of New York Dow Jones Indus. Index sequence.

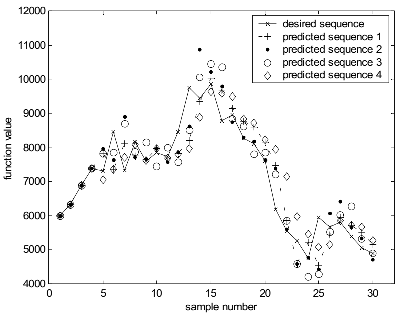

Figure 2. Forecasting results of Taiwan TAIEX sequence.

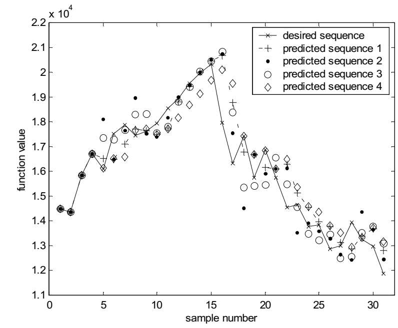

Figure 3. Forecasting results of Japan Nikkei Index sequence.

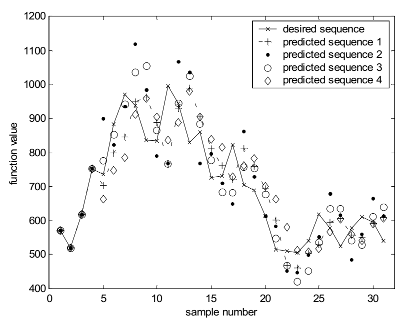

Figure 4. Forecasting results of South Korea-Stock Index sequence.

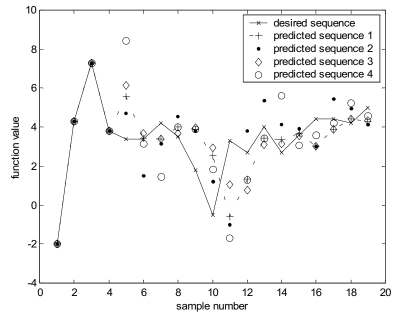

Figure 5. Forecasting results of U.S.A economy growth rate sequence.

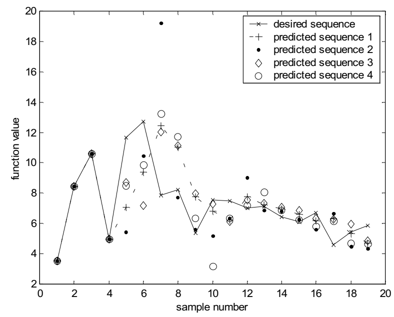

Figure 6. Forecasting results of Taiwan economy growth rate sequence.

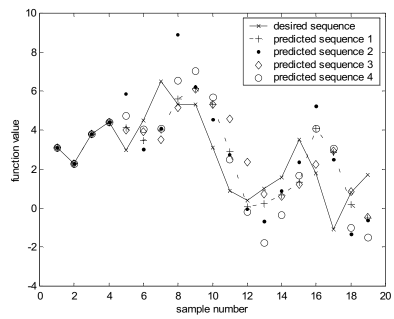

Figure 7. Forecasting results of Japan economy growth rate sequence.

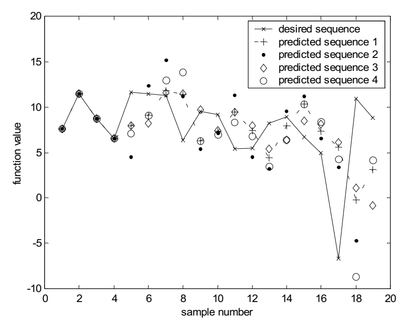

Figure 8. Forecasting results of South Korea economy growth rate sequence.

Table 1. The MSE based on the relative errors of predicted results on stock price index for four countries.

SI	GM	3CLSLM	LSLM	NM
NYDJ	0.0038	0,0019	0.0032	0.0019
TAIEX	0.0102	0.0136	0.0114	0.0090
NIKKEI	0.0051	0.0047	0.0042	0.0037
SKSI	0.0209	0.0166	0.0167	0.0129
Average	0.0100	0.0092	0.0089	0.0069

Abbreviation:

SI-Stock index, GM-Grey Model, 3CLSLM- 3 Points Cumulative Least Squared Linear model, LSLM-Least Squared Linear Model, and NM-Proposed New Method

Table 2. The MSE based on the relative errors of predicted results on economy growth rate for four countries.

EGR	GM	3CLSLM	LSLM	NM
USA	1.0609	3.3859	1.9984	1.0582
TWAN	0.1986	0.0760	0.0936	0.0751
JAPAN	2.2282	3.8441	2.7228	1.6381
KOREA	0.5324	0.5278	0.6271	0.4750
Average	1.0050	1.9585	1.3605	0.8116

Abbreviation:

EGR-Economy Growth Rate, GM-Grey Model, 3CLSLM- 3 Points Cumulative Least Squared Linear model, LSLM-Least Squared Linear Model, and NM- Proposed New Method