# A Modified Grey Prediction Method to Early Manufacturing Data Sets

Chun-Wu Yeh, Che-Jung Chang, and Der-Chiang Li

*Abstract*—**In the age of global competition, product life cycles are shortened and the volatility increases. Industrial demands no longer follows linear trend which could be forecasted relatively easily. As a consequent, enterprise to survival depends on its capability to accommodating changes in manufacturing environment. Due to the rareness of data in the early stages of manufacturing management, traditional prediction techniques hardly obtain ideal results. The urgency of learning and acquirement of the management knowledge is unavoidable for decision makers. Therefore, small-data-set prediction cannot be ignored and should be taken seriously.**

**Nowadays the grey prediction model is one of the important methods for small-data-set prediction. But its application is limited because the way of building model is fixed and could not be modified by the characteristics of data. This research explores the extra information with data analysis through trend and potency tracking method, and uses the generation of trend and potency value of each datum to develop the adaptive grey prediction model as a tool of small-data-set prediction based on the theory of grey system. The example verification displayed that the presented method establishes the suitable model according to the characteristics of data, and also improves the prediction accuracy of small data sets.**

*Index Terms*—**Grey theory, Small data sets, Trend and potency tracking method**

Manuscript received December 9, 2008. This research was supported by the National Science Council under Grant 97-2410-H-168-017.

Chun-Wu Yeh is an assistant professor in the Department of Information Management, Kun Shan University, No. 949, Da-Wan Road, Yung-Kang City, Tainan Hsien, 71003, Taiwan, ROC. (phone: +886(6)205-1053; fax: +886(6)205-0545; e-mail: davidyeh929@gmail.com).

Che-Jung Chang is a Ph.D. student at the Department of Industrial and Information Management, National Cheng Kung University, No. 1, University Road, Tainan City, 70101, Taiwan, ROC. (e-mail: kentaaa@ms51.hinet.net).

Der-Chiang Li is a professor in the Department of Industrial and Information Management, National Cheng Kung University, No. 1, University Road, Tainan City, 70101, Taiwan, ROC. (e-mail: lidc@mail.ncku.edu.tw).

## I. Introduction

The era of global competition with economy as main shaft is coming, and it makes international funds move faster than ever. In addition, the number of products increases dramatically and the standard of services becomes higher. Organizations have to change their strategies so as to adapt to the changing environment, and make adequate responses to customers' demand. Information such as demand and sales are often limited in the initial stage of the process, and the problem that we face is quite uncertain and incomplete. How to increase the core ability of competition, speed, and knowledge is crucial for companies to survive in this ambiguous world. Therefore, the importance of current management is far more than it in the past. The suitable management model can be used to forecast future demands by looking into the limited data obtained and to collect valuable information. Moreover, globalization and innovation make manufacturing environment change expeditiously, and the product life cycle compresses rapidly. It is important for companies to control manufacturing systems more efficiently. The initial stage of manufacturing systems is the key point for management. Managers should take right actions and set up management knowledge by detecting manufacturing problems beforehand.

Due to unsatisfactory data in the initial stage of manufacturing systems, traditional statistical theory and data mining techniques may not be suitable for prediction. The grey system theory developed by Deng [1] is to deal with information uncertainty and insufficiency of systems. Unlike statistical methods, this theory mainly deals indirectly with original data by accumulated generating operation (AGO) and tries to find out its intrinsic regularity. Besides, assumptions concerning statistical distribution of data are not necessary when applying grey theory, and the potency of the original series in the grey model, called GM(1,1), has been proven to be as few as four [2]. Grey theory has been widely applied in various fields including systems analysis, data processing, forecasting, decision-making and so on. In addition, grey forecasting models have been used in many applications, such as prediction in agriculture, traffic, transportation, marketing, engineering, and so on [3]-[17].

Hence, the interesting issue when predicting under the condition of fewer samples is objective of this study. Nonetheless, the drawback of using GM(1,1) model is that it can not reflect growth trends of data in real systems. Therefore, this study applies the trend and potency tracking method (TPTM) proposed by Li and Yeh [18] for analyzing data behavior, and utilizes the trend and potency value to construct an adaptive grey prediction model, AGM(1,1). The Synthetic Control Chart Time Series dataset (SCCTS) will be the experimental study to verify the proposed method, and we also compare the performance with conventional grey model GM(1,1) and nonlinear grey Bernoulli model, NGBM. The experimental results illustrate that AGM(1,1) indeed improves the model precision. The following of this study is that section 2 briefly introduces the concept of trend and potency tracking method (TPTM), and the proposed model AGM(1,1). Section 3 uses AGM(1,1) to forecast the Synthetic Control Chart Time Series dataset comparing with other forecasting methods, and conclusions are presented in section 4.

## II. The proposed methodology: AGM(1,1)

To improve the accuracy of forecasting, this study applies the trend and potency tracking method (TPTM) to construct an adaptive grey prediction model, AGM(1,1), based on grey theory. The trend and potency tracking method systematically estimates the variation of sequential data obtained to get hidden message, and builds the trend and potency function with an asymmetrical domain range. TPTM explores data behavior and estimates possible change in different time stages. Therefore, this research utilizes the concept of TPTM onto grey prediction models to be the adaptive grey model named as AGM(1,1).

TP values in TPTM examine data behavior, and display the approaching degree between data and the center of original data (CL). If TP values are bigger, that means the fluctuation of data is less concerning the whole data. Moreover, the crucial factor that influences GM(1,1) adoption and forecasting results is the background value. Using background values is to smooth data and reducing randomness. Different background values will affect the value of developing coefficient and the model precision. Researchers generally regard as each data with equal importance and set the adjusting factor $\alpha$ equal 0.5 for easily solving the problem. However, this will ignore the difference of data characteristics and produce more prediction errors. In other words, one needs to choose larger $\alpha$ for emphasizing the importance of the newest data or one should choose smaller $\alpha$ for reducing randomness to reduce prediction errors. In this condition, the relation between $\alpha$ and TP value is positive and both of the values lie between 0 and 1. Hence, it's suitable to replace $\alpha$ with TP values. This research proposes the concept for deciding the value of $\alpha$ by considering data behavior and the whole data characteristics to revise the grey model into the adaptive grey model. Its procedure is as follows:

Step 1: Suppose we have n data

$$X^{(0)} = \left\{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right\}$$

Step 2: Calculate the TP values by TPTM, that is

$$\left\{TP_i\right\} = \left\{TP_1, TP_2, ..., TP_n\right\},$$

$$i = 1, 2, ..., n$$

Step 3: $\alpha_k$ is computed by $\alpha_k = \dfrac{\sum_{i=1}^{k} 2^{i-1} TP_i}{\sum_{i=1}^{k} 2^{i-1}}$,

$$k \ge 2$$

Step 4: Use AGO $x^{(1)}(1) = x^{(0)}(1)$;

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(k), \qquad k = 2, 3, \cdots, n$$

to form a new data series $X^{(1)} = \left\{x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\right\}$

Step 5: Calculate background values

$$z^{(1)}(k) = x^{(1)}(k-1) + \alpha_k x^{(0)}(k),$$

$$\alpha \in (0,1), \quad k = 2, 3, \cdots, n$$

Step 6: Establish the grey differential equation and estimate the developing coefficient $a$ and the grey input $b$ by the ordinary least square (OLS) method

Step 7: Construct the adaptive grey model, and substitute $k = 2, 3, \cdots, n+1$ into model to obtain forecasts of each period.

## III. Experimental Studies: Analysis of the Synthetic Control Chart Time Series (SCCTS) Data

This research chooses a Synthetic Control Chart Time Series data set collected by Knowledge Discovery Database (http://kdd.ics.uci.edu/databases/synthetic_control/synthetic_control.html). There are 600 observations of control charts and there are six different classes (normal, cyclic, increasing trend, decreasing trend, upward shift, downward shift) in SCCTS. We choose the normal (1-100), cyclic (101-200), increasing trend (201-300), decreasing trend (301-400) examples as listed in Table I for the experimental analysis. Suppose each example has only 4-stage data at present, given $\{x_1, x_2, x_3, x_4\}$, and the work is to construct a prediction model based on the current observations. Hence, this learned model is to predict the values of the incoming data such as the value of $x_5$. Namely, the 4-stage data will be adopted to induce certain knowledge for predicting $x_5$.

Table I: The Synthetic Control Chart Time Series (SCCTS) Data (1-400)

| | Time stage | | | | |
|---|---|---|---|---|---|
| No. | 1 | 2 | 3 | … | 60 |
| 1 | 28.7812 | 34.4632 | 31.3381 | | 25.8717 |
| 2 | 24.8923 | 25.741 | 27.5532 | | 26.691 |
| 3 | 31.3987 | 30.6316 | 26.3983 | | 29.343 |
| 4 | 25.774 | 30.5262 | 35.4209 | | 25.3069 |
| 5 | 27.1798 | 29.2498 | 33.6928 | … | 31.0179 |
| 6 | 25.5067 | 29.7929 | 28.0765 | | 35.4907 |
| 7 | 28.6989 | 29.2101 | 30.9291 | | 26.4637 |
| 8 | 30.9493 | 34.317 | 35.5674 | | 34.523 |
| 9 | 35.2538 | 34.6402 | 35.7584 | | 32.3833 |
| ⋮ | | ⋮ | | ⋱ | ⋮ |
| 398 | 26.2727 | 31.2289 | 29.7408 | | 7.90108 |
| 399 | 26.1482 | 30.8284 | 27.1216 | … | 8.98081 |
| 400 | 30.8314 | 31.9645 | 31.0937 | | 14.6646 |

### *A. Case study of GM(1,1), AGM(1,1), and NGBM(1,1)*

This study takes $\{28.7812, 34.4632, 31.3381, 31.2834\}$ for instance to construct the prediction model of the three methods. In GM(1,1), we let $\alpha = 0.5$ to create the model for predicting $\hat{x}_5$. In AGM(1,1), we compute the TP value of each observation and then use them to generate the coefficient $\alpha$ of the background value. In NGBM(1,1), we choose power $r$ in accordance with the research of Chen et al. [19] that used the criteria of minimizing mean absolute percentage errors. Therefore, the GM(1,1), AGM(1,1), and NGBM(1,1) model is shown in Table II, III, and IV.

Table II: Experimental study of GM(1,1)

| Time | Analysis case | | | |
|---|---|---|---|---|
| Stage | $x_k$ | $x_k^{(1)}$ | $z_k^{(1)}$ | $\hat{x}_k$ |
| 1 | 28.7812 | 28.7812 | – | – |
| 2 | 34.4632 | 63.2444 | 46.0128 | 33.9820 |
| 3 | 31.3381 | 94.5825 | 78.9135 | 32.3277 |
| 4 | 31.2834 | 125.8659 | 110.2242 | 30.7540 |
| 5 | 28.9207 | – | – | 29.2569 |
| $\hat{x}^{(1)}(k+1)=(-698.066)e^{-0.0499k}+726.847$ | | | | |

Table III: Experimental study of AGM(1,1)

| Time | Analysis case | | | | | |
|---|---|---|---|---|---|---|
| Stage | $x_k$ | TP value | $x_k^{(1)}$ | $\alpha_k$ | $z_k^{(1)}$ | $\hat{x}_k$ |
| 1 | 28.7812 | 0.53027 | 28.7812 | – | – | – |
| 2 | 34.4632 | 0.66667 | 63.2444 | 0.6212 | 50.1898 | 34.2095 |
| 3 | 31.3381 | 0.95303 | 94.5825 | 0.8108 | 88.6538 | 32.7015 |
| 4 | 31.2834 | 0.94398 | 125.8659 | 0.8818 | 122.1694 | 31.2600 |
| 5 | 28.9207 | – | – | – | – | 29.8820 |
| $\hat{x}^{(1)}(k+1)=(-776.045)e^{-0.0451k}+804.826$ | | | | | | |

Table IV: Experimental study of NGBM(1,1)

| Time | Analysis case | | | |
|---|---|---|---|---|
| Stage | $x_k$ | $x_k^{(1)}$ | $z_k^{(1)}$ | $\hat{x}_k$ |
| 1 | 28.7812 | 28.7812 | – | – |
| 2 | 34.4632 | 63.2444 | 42.1482 | 37.7090 |
| 3 | 31.3381 | 94.5825 | 75.3993 | 31.5451 |
| 4 | 31.2834 | 125.8659 | 106.7162 | 31.2580 |
| 5 | 28.9207 | – | – | 32.0257 |
| $\hat{x}^{(1)}(k+1)=[(1677.899)e^{0.1251k}-1567.555]^{0.7143}$ | | | | |

### *B. Experimental results among three prediction methods*

This study implemented the prediction task of four different cyclic time series data in SCCTS and used four measurements to evaluate accuracy and stability of prediction methods in Table V. Four measurements are mean square error (MSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and standard deviation (SD). Table V shows the results between the predicted value ($\hat{X}_k$) and the actual value ($X_k$) of the fifth stage among the forecasting methods. In the experiment, the proposed approach AGM(1,1) outperforms the other two methods. In addition, the standard deviation of AGM(1,1) is the smallest among contrasting methods. It proves clearly that AGM(1,1) is adequate for both accuracy and stability.

Table V: The performance of forecasting methods of SCCTS data

| Methods | MSE | MAE | MAPE(%) | SD |
|---|---|---|---|---|
| GM(1,1) | 54.39 | 5.81 | 18.18 | 7.11 |
| AGM(1,1) | 46.37 | 5.35 | 16.76 | 6.60 |
| NGBM(1,1) | 112.34 | 8.75 | 27.75 | 10.57 |

## IV. Conclusions

This study proposed a revised prediction model that combines the advantage of grey theory and TPTM to conquer the problem that occurs in the early stages of manufacturing systems. From the experimental results, the proposed method AGM(1,1) systematically analyzes data behavior at hand and can help raise the prediction performance. Besides, it is crucial to decide the proper value of coefficient $\alpha$ for model construction, and it will make prediction results better. Therefore, AGM(1,1) is an adequate forecasting tool for small data sets.

## References

[1] J. L. Deng, Grey System Fundamental Method, Huazhong University of Science and Technology, Wuhan, China, 1982.

[2] J. L. Deng, Grey Prediction and Decision, Huazhong University of Science and Technology, Wuhan, China, 1986.

[3] S. C. Chang, H. C. Lai, and H. C. Yu, "A variable P value rolling Grey forecasting model for Taiwan semiconductor industry production," Technological Forecasting and Social Change, Vol. 72, No. 5, 2005, pp. 623-640.

[4] Y. H. Hao, and X. M. Wang, "The Study of Grey System Models of Niangziguan Spring," Journal of systems Engineering, Vol. 16, No. 3, 2000, pp. 39-44.

[5] C. C. Hsu, and C. Y. Chen, "Application of Grey Theory to Regional Electricity Demand Forecasting," Energy Quarterly, Vol. 24, No. 4, 1999, pp. 96-108.

[6] C. I. Hsu, and Y. H. Wen, "Improved Grey Prediction Models for the Trans-Pacific Air Passenger Market," Transportation Planning and Technology, Vol. 22, No. 2, 1998, pp. 87-107.

[7] Y. P. Huang, and S. F. Wang, "The Identification of Fuzzy Grey Prediction System by Genetic Algorithm," International Journal of Systems Sciences, Vol. 28, No. 3, 1997, pp. 15-26.

[8] L. C. Hsu, "Applying the Grey Prediction Model to the Global Integrated Circuit Industry," Technological Forecasting and Social Change, Vol. 70, No. 6, 2003, pp. 563-574.

[9] L. C. Hsu, and C. H. Wang, "Forecasting the Output of Integrated Circuit Industry Using a Grey Model Improved by the Bayesian Analysis," Technological Forecasting and Social Change, Vol. 74, No. 6, 2007, pp. 843-853.

[10] Y. Jiang, Y. Y. Yao, S. Deng, and Z. Ma, "Applying Grey Forecasting to Predicting the Operating Energy Performance of Air Cooled Water Chillers." International Journal of Refrigeration, Vol. 27, No. 4, 2004, pp.385-392.

[11] C. T. Lin, and S. Y. Yang, "Forecast of the Output Value of Taiwan's Opto-electronics Industry Using the Grey Forecasting Model," Technological Forecasting and Social Change, Vol. 70, No. 2, 2003, pp. 177-186.

[12] M. Mao, and E. C. Chirwa, "Application of Grey Model GM(1, 1)

ISBN: 978-988-17012-2-0 

to Vehicle Fatality Risk Estimation," Technological Forecasting and Social Change, Vol. 73, No. 5, 2006, pp. 588–605.

[13] H. Morita, N. F. Hubele, and G. G. Karady, "Interval Prediction of Annual Maximum Demand Using Grey Dynamic Model," Electrical Power and Energy Systems, Vol. 18, No. 3, 1996, pp. 409-413.

[14] G. Sun, "Prediction of Vegetable Yields by Grey Model GM(1,1)," Journal of Grey System, Vol. 2, No. 2, 1991, pp. 187-197.

[15] M. Xing, "Research on Combined Grey Neural Retwork Model of Seasonal Forecast," System Engineering Theory and Application, Vol. 1, No. 1, 2001, pp. 31-35. (in Chinese).

[16] H. Yong, "A New Forecasting Model for Agricultural Commodities," Journal of Agricultural Engineering Research, Vol. 60, No. 4, 1995, pp. 227-235.

[17] C. L. Yue, and L. Wang, "Grey-Markov Forecast of the Stock Price," System Engineering, Vol. 16, No. 3, 2000, pp. 54-59. (in Chinese).

[18] D.C. Li, and C.W. Yeh, "A Non-parametric Learning Algorithm for Small Manufacturing Data Sets," Expert Systems with Applications, Vol. 34, 2008, pp. 391-398.

[19] C. I. Chen, H. L. Chen, and S. P. Chen, "Forecasting of Foreign Exchange Rates of Taiwan's Major Trading Partners by Novel Nonlinear Grey Bernoulli Model NGBM (1,1)," Communications in Nonlinear Science and Numerical Simulation, Vol. 13, No. 6, 2008, pp.1194-1204.