Research Article

Utilizing an Adaptive Grey Model for Short-Term Time Series Forecasting: A Case Study of Wafer-Level Packaging

Che-Jung Chang,[1] Der-Chiang Li,[2] Wen-Li Dai,[3] and Chien-Chih Chen[2]

[1] *Department of Business Administration, Chung Yuan Christian University, No. 200, Chung-Pei Road, Chung-Li City, Taoyuan County 32023, Taiwan*

[2] *Department of Industrial and Information Management, National Chen Kung University, No. 1, University Road, Tainan City 70101, Taiwan*

[3] *Department of Information Management, Tainan University of Technology, No. 529, Zhongzheng Road, Tainan City 71002, Taiwan*

Correspondence should be addressed to Der-Chiang Li; lidc@mail.ncku.edu.tw

The wafer-level packaging process is an important technology used in semiconductor manufacturing, and how to effectively control this manufacturing system is thus an important issue for packaging firms. One way to aid in this process is to use a forecasting tool. However, the number of observations collected in the early stages of this process is usually too few to use with traditional forecasting techniques, and thus inaccurate results are obtained. One potential solution to this problem is the use of grey system theory, with its feature of small dataset modeling. This study thus uses the AGM(1,1) grey model to solve the problem of forecasting in the pilot run stage of the packaging process. The experimental results show that the grey approach is an appropriate and effective forecasting tool for use with small datasets and that it can be applied to improve the wafer-level packaging process.

1. Introduction

Firms can gain competitive advantages by more effectively controlling their manufacturing systems [1]. In this context, the early stages of a manufacturing system are especially important for managers, because any slight change in the parameters used at this time will significantly influence final product quality and manufacturing performance [2]. Timely and precise information is thus needed for effective operations management [3]. However, few observations are available at the early stages of a manufacturing system, meaning that statistical theory and data mining techniques cannot obtain accurate results [4, 5].

The introductory stage of the wafer-level packaging (WLP) process is an example of a small dataset problem. The WLP process is a new, advanced technology for the packaging industry, and thus manufacturers do not have much experience to draw its form when seeking to improve it. However, despite the lack of data, engineers still need to understand the features of this production system to maintain a high manufacturing yield [2]. Consequently, an appropriate forecasting tool that can deal with incomplete information based on small datasets is required for effective management of the WLP process.

Grey system theory was proposed by Deng (1982) [6] and is used to examine uncertain systems based on incomplete information [7, 8]. The main principle of this approach is to process the data indirectly through data mapping to transform the state space, so as to extract hidden information in the data [9, 10]. Because of its ease of use, grey system theory has been successfully applied in various domains [11–17].

Many studies have demonstrated that the grey model is an effective approach for data analyses with small samples, and it is applied in this research to the packaging process of semiconductor manufacturers, in order to solve forecasting problems in the pilot run stage. The grey model developed in this work is preexamined by four different measurements and compared with two other popular forecasting methods

to examine its flexibility and value in this context. The results show that the grey approach is an effective tool that can enable manufacturers to make more accurate forecasts based on limited observations collected in the early stages of the packaging process.

The rest of this paper is organized as follows. Section 2 introduces the development of the adaptive grey model used in this work. In Section 3, we briefly describe the problem of interest in the pilot run of the wafer-level packaging process, where the applicability of the grey approach is demonstrated in this study. Finally, the conclusions are presented in Section 4.

2. Methodology

While the conventional grey forecasting model, GM(1,1), has been widely applied, it is still possible to improve its forecasting performance. Li et al. [18] proposed a revised grey model, called AGM(1,1), which is one of the best and most accurate of the existing models. It uses the trend and potency tracking method [19] to form a function to determine the background values of the grey model. This model can better reflect the data growth trends at different stages of the process and overcomes some of the weaknesses of the ordinary modeling procedure in the GM(1,1), thus producing more accurate forecasts. Therefore, this study uses this method to solve the forecasting problem that arises in the pilot run of the wafer-level packaging process. The AGM(1,1) modeling procedure has two main parts, which are (1) determining the trend and potency value and (2) building the grey forecasting model. Both of these will be described in the following subsections.

2.1. Trend and Potency Tracking Method. Li and Yeh [19] proposed the trend and potency tracking method (TPTM), which is an analysis method that uses the characteristics of data to explore possible changes in data behavior in different stages of a process, and this is the key concept used in AGM(1,1) to improve the accuracy of the conventional grey forecasting model.

The detailed procedure of the TPTM is as follows.

Step 1. Given the original data series, $X = \{x_1, x_2, \ldots, x_n\}$, let $x_{\min}$ be the minimum value in X and $x_{\max}$ the maximum one.

Step 2. Figure the variations σ_i of paired data (x_{i-1}, x_i), $i = 2, 3, \ldots, n$, and determine the increasing or decreasing potencies according to the data sequence. If σ_i concerning the paired data (x_{i-1}, x_i) is positive, it indicates that the trend of data at phase i is moving upward. Conversely, the movement of the data at phase i is descending if $\sigma_i < 0$.

Step 3. Set the biggest weight to the latest datum as the intensity of different phases, since it dominates the characteristics of the upcoming datum. The importance w_i of the datum at phase i equals $i - 1$, $i = 2, 3, \ldots, n$.

Step 4. Let $A_i = \sigma_i \times w_i$, $i = 2, 3, \ldots, n$, as an operator to strengthen the data trend and potency at different phases by multiplying weights and variations together. Note that $A_i > 0$
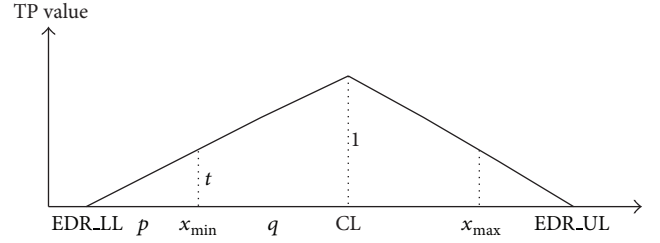

FIGURE 1: An example of the triangular TP function.

is an increasing potency (IP) and $A_i < 0$ is a decreasing potency (DP).

Step 5. Determine the central location (CL) of existing data using the equation CL $= (x_{\min} + x_{\max})/2$. The CL is then utilized as the main point to conduct the asymmetric domain range expansion.

Step 6. Compute the average of increasing potencies (AIP) and the average of the decreasing potencies (ADP), and then use them to asymmetrically extend the domain range. The upper limit of the extended domain range is EDR_UL $= x_{\max} + $ AIP, and the lower limit is EDR_LL $= x_{\min} + $ ADP. We thus obtain an extended domain range within which to explore extrainformation about the data trend and potency.

Step 7. Form a triangular TP function using the CL, EDR_UL, and EDR_LL. Here, we set the TP value of the CL to be 1, and then we can obtain the TP values of existing data through the ratio rule of a triangle. Figure 1 illustrates an example and the TP value of $x_{\min}$ is $t = p/(p + q)$, where p is the distance between EDR_LL and $x_{\min}$, and q is the distance between $x_{\min}$ and CL. The range of the TP value is between 0 and 1, and the TP value reveals that the current datum's intensity is close to the CL.

2.2. Adaptive Grey Forecasting Model. The background value of GM(1,1) is the most important factor that affects model construction and the final forecasting results. Li et al. [18] studied the impact of differences in the background value on forecasting performance and proposed an improved grey forecasting model, AGM(1,1), by integrating the concept of TPTM into the formula for calculating the background value. This study uses this approach to deal with forecasting problems in a small dataset.

The AGM(1,1) model is described as follows.

Step 1. Suppose the original data series is $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)\}$, $n \geq 4$.

Step 2. Calculate the TP values by TPTM; that is, $\{TP_i\} = \{TP_1, TP_2, \ldots, TP_n\}$, $i = 1, 2, \ldots, n$.

Step 3. Compute the coefficient α_k

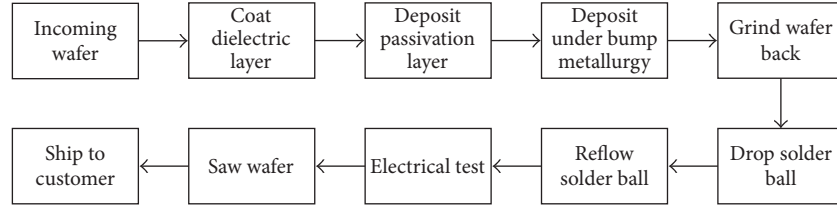$$\alpha_k = \frac{\sum_{i=1}^k 2^{i-1} TP_i}{\sum_{i=1}^k 2^{i-1}}, \quad k \geq 2. \tag{1}$$

FIGURE 2: Flowchart of the wafer-level packaging process.

Step 4. Form a new data series through the accumulating generation operator (AGO)

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

$$
x^{(1)}(1) = x^{(0)}(1);
$$

$$
x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), \quad k = 2, 3, \ldots, n.
$$

(2)

Step 5. Determine the background values $z^{(1)}(k)$

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

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

(3)

Step 6. Estimate the developing coefficient a and the grey input b from (4) by the ordinary least-square method and establish the grey differential equation from (5) to replace the source model

$$
x^{(0)}(k) + az^{(1)}(k) = b,
$$

(4)

$$
\frac{dx^{(1)}}{dt} + ax^{(1)} = b.
$$

(5)

To estimate a and b, we expand (4) as

$$
\begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix} = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} \times \begin{bmatrix} a \\ b \end{bmatrix}.
$$

(6)

Let

$$
Y = \left[x^{(0)}(2), x^{(0)}(3), \ldots, x^{(0)}(n) \right]^{T},
$$

$$
\widehat{a} = [a, b]^{T},
$$

$$
B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix},
$$

(7)

$$
\widehat{a} = \left(B^{T} B \right)^{-1} B^{T} Y.
$$

Step 7. Solve (5) together with initial condition $x^{(0)}(1) = x^{(1)}(1)$, and the desired forecasting output at step $k + 1$ can be obtained by (8)

$$
\widehat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{b}{a} \right) e^{-ak} + \frac{b}{a},
$$

$$
\widehat{x}^{(0)}(k+1) = \widehat{x}^{(1)}(k+1) - \widehat{x}^{(1)}(k).
$$

(8)

3. Experimental Studies: Forecasting the Wafer-Level Packaging Process

The applicability of the adaptive grey forecasting model is examined here by an experiment, and its details are described in the following subsection.

3.1. Background of Wafer-Level Packaging. Semiconductor manufacturing is an important link in the supply chain of consumer electronics and has significant effects on the quality and characteristics of the final products. The packaging method used in this industry is increasingly the wafer-level packaging (WLP) rather than the chip-scale packaging, as this can enable smaller package sizes. Moreover, WLP can allow manufacturers to integrate the wafer fabrication, packaging, testing, and burn-in processes. In practice, WLP is a technology that packages an integrated circuit at the wafer level, instead of the traditional process of assembling individual units in packages after dicing them from a wafer. Many different materials are combined on a substrate in WLP to form a chip, as shown in Figure 2, so the physical properties of this combination depend on the rigidity of materials, the thermal expansion coefficient, and the temperature variation among them. There are five quality inspections in the WLP process to ensure the process standards, and these examine the passivation layer defects, diameter size, coplanarity, metal bump defects, and the height of metal bumps.

Further, because the metal bump height affects the overall quality in the packaging process, the detection of this is very important and needs to be effectively controlled to maintain a good yield. In addition, while solder balls are generally used for surface mounting, because they have a relatively low melting point, they can also suffer from creep and plastic behavior during mounting under common operating conditions. This means that the height of solder balls is unstable in the packaging process, which can adversely affect the quality of the final products. In short, if engineers can better understand the moving trend of the solder ball height and

adjust the process parameters in a timely fashion, then this can help improve overall manufacturing efficiency.

3.2. Experimental Data and Procedure. This research employs data collected from a leading packaging manufacturer in Taiwan. The case company is a wafer original equipment manufacturer, with relatively little experience of WLP production at the time when the data was collected.

The experimental time series data were collected at the end of 2011, the period when the case company was implementing a new WLP process and subjecting it to postinstallation testing. This pilot run yielded a total of thirteen observations, as shown in Table 1. The unit of measurement for the height in this table is micrometers (μm). The aim of this study is to build an effective forecasting model for the height of solder balls over time. Four data are used to construct the forecasting model, with the rest of the data used one at a time for testing, so that a total of nine sets of data are obtained.

3.3. Computation of the AGM(1,1) Model. To illustrate the modeling of the AGM(1,1) in more detail, this study employs the preceding four data {192.16, 192.78, 192.95, 193.11} as an example to build the model.

The computation process is shown below, and the results are listed in Table 2.

  (1) The original dataset is $X^{(0)}$ = {192.16, 192.78, 192.95, 193.11}.

  (2) Apply TPTM to calculate the TP values and get TP = {0, 0.8482, 0.6702, 0.5026}.

  (3) Compute the coefficient α_k = {0.5654, 0.6253, 0.5599}.

  (4) Use AGO to form a new series $X^{(1)}$ = {192.16, 384.94, 577.89, 771.00}.

  (5) Determine the background values $Z^{(1)}$ = {301.1665, 505.5879, 686.0046}.

  (6) Calculate $\hat{a} = (B^T B)^{-1} B^T Y$, where

$$Y = [192.78, 192.95, 193.11]^T, \qquad B = \begin{bmatrix} -301.1665 & 1 \\ -505.5879 & 1 \\ -686.0046 & 1 \end{bmatrix}. \tag{9}$$

We can get $\hat{a} = [-0.000857, 192.52027]^T$.

  (7) Use (8) to create the model $x^{(1)}(k + 1) = 224854.4073e^{0.000857k} - 224662.2473$ and the corresponding output is predicted as $\hat{x}_5 = 193.264$. See Table 2.

3.4. Preevaluation Indexes for the Performance of AGM(1,1) Model. To assess the reasonableness and applicability of the adaptive grey forecasting model in the early stage of the WLP process, we first use four different measurements that are often used in pretesting modeling fitness in the grey system theory [20] to evaluate the forecasting performance. These are the mean relative error (MRE), absolute degree of grey incidence (ADGI), ratio of standard deviation (RSD), and

Table 1: Wafer-level packaging process data (unit: micrometers).

Order	Height (μm)	Order	Height (μm)
01	192.16	08	194.48
02	192.78	09	194.42
03	192.95	10	195.50
04	193.11	11	195.67
05	193.50	12	195.98
06	194.13	13	196.23
07	194.30		

probability of small error (PSE). We let m be the sum of testing samples, while $\hat{x}_i$, x_i and e_i are the predicted outputs, actual values and errors of the ith testing sample, respectively. The average and standard deviations of x_i and e_i, written as $\bar{x}$, $\bar{e}$, s_x and s_e, are then computed by basic formulas in the statistics. The definitions of MRE, RSD, and PSE are presented below. Since ADGI is relatively complex, it is not easy to define briefly here, and so please refer to an earlier study [21] for further information about this

$$\begin{aligned} \text{MRE} &= \frac{1}{m}\sum_{i=1}^{m}\left|\frac{\hat{x}_i - x_i}{x_i}\right|, \\ \text{RSD} &= \frac{s_e}{s_x}, \\ \text{PSE} &= P\left(\left|e_i - \bar{e}\right| < 0.6745 \times s_x\right). \end{aligned} \tag{10}$$

The MRE is a measurement that is used to estimate the forecasting accuracy. The ADGI judges the closeness of the relationship between the actual and forecast values, based on the similarity of the geometric patterns of the sequence curves. It is often used to check whether or not a model is able to appropriately reflect the data profile. The RSD is an indicator of the degree of discreteness of the errors, that is, the whole amplitude of the error fluctuation. The PSE shows what ratio of errors lie in the acceptable range. With these four measurement indexes, smaller values of MRE and RSD represent better simulated results, while bigger ones are needed with ADGI and PSE [20, 21].

The levels used to examine the accuracy of the model and the critical values of the previous four indexes are shown in Table 3. These four measurements can help determine whether a model is accurate or not. If they fall within the range of levels 1 or 2, this suggests that the model has high forecasting accuracy.

3.5. Comparison with Other Conventional Forecasting Approaches. To verify the effectiveness of the AGM(1,1) model, we compare its results with those from two popular forecasting techniques, the backpropagation neural network (BPNN) and support vector regression (SVR) approaches.

BPNN is a neural network that has shown good forecasting ability in many contexts [22]. The learning tool used in this work is the Pythia software. We use four training samples to train a BPNN, where each sample includes one input attribute and one output attribute, such as $(1, x_1), (2, x_2), (3, x_3),$

TABLE 2: Computation of the AGM(1,1).

Order	x_k	TP value	$x_1^{(1)}$	α_k	$z_k^{(1)}$	$\widehat{x}_k$
1	192.16	0.0000	192.16	—	—	—
2	192.78	0.8482	384.94	0.5654	301.1665	192.768
3	192.95	0.6702	577.89	0.6253	505.5879	192.933
4	193.11	0.5026	771.00	0.5599	686.0046	193.098
5	193.50	—	—	—	—	193.264

$x^{(1)}(k+1) = 224854.4073e^{0.000857k} - 224662.2473.$

TABLE 3: Level of accuracy.

Level	MRE	ADGI	RSD	PSE
1	0.01	0.90	0.35	0.95
2	0.05	0.80	0.50	0.80
3	0.10	0.70	0.65	0.70
4	0.20	0.60	0.80	0.60

TABLE 4: Forecasting results of AGM(1,1).

Order	Actual values	Forecasting values
01	192.16	—
02	192.78	—
03	192.95	—
04	193.11	—
05	193.50	193.264
06	194.13	193.640
07	194.30	194.414
08	194.48	194.731
09	194.42	194.637
10	195.50	194.505
11	195.67	195.798
12	195.98	196.284
13	196.23	196.169

and $(4, x_4)$. The topology of the BPNN is a 1-2-1 structure, and after it is trained it returns the corresponding value for $\widehat{x}_5$.

SVR is a nonparametric estimation learning algorithm based on statistical learning theory, which is used to solve training problems with limited samples [23, 24]. Due to its good performance, it is one of the primary methods used in machine learning. We use Weka 3.6.9 as the learning tool for the SVR, and the training set is the same as that used with the BPNN.

Since accuracy is a critical index when measuring the ability of a forecasting method [25], we employ three error indexes in this work, namely, the mean square error (MSE), mean absolute error (MAE), and mean absolute percentage error (MAPE), and these are calculated as follows:

$$MSE = \frac{1}{m}\sum_{i=1}^{m}(\widehat{x}_i - x_i)^2,$$

$$MAE = \frac{1}{m}\sum_{i=1}^{m}|\widehat{x}_i - x_i|, \tag{11}$$

$$MAPE = \frac{1}{m}\sum_{i=1}^{m}\left|\frac{\widehat{x}_i - x_i}{x_i}\right| \times 100\%,$$

where m is the sum of testing samples and $\widehat{x}_i$ and x_i are the predicted and actual values of the ith testing sample, respectively.

3.6. Experimental Results. This research uses the four pretesting measurements in Section 3.4 to assess the forecasting performance from various perspectives. The forecasting results of the proposed approach are shown in Table 4, and the four measurements are as follows: MRE is 0.0016, ADGI is 0.9515, RSD is 0.4496, and PSE is 0.8889. This shows that the AGM(1,1) falls within level 2 and thus has a high forecasting accuracy.

Further, the MRE value, which is smaller than 0.01, suggests that the risk when using this grey model is controllable. The ADGI value is about 0.95, indicating a high similarity

between the actual and forecast results with regard to the geometric shape, and this also explains that the proposed approach can properly reflect the data trend, with no serious inconsistencies between the forecast results and the real situation. Lastly, both the measurements of RSD and PSE are at an acceptable level, so there are neither extreme values nor a high level of error discreteness, which both indirectly confirm the stability of the AGM(1,1) model.

The results shown in Table 5 indicate that the AGM(1,1) model has better forecasting performance and outperforms the other two methods, with its average improvements being over 20% for all three error indexes. In addition, the MAPE of AGM(1,1) is less than 0.2%, which falls into an acceptable range in the WLP industry. These results clearly show that the AGM(1,1) model is a feasible tool to deal with forecasting task in the WLP process.

4. Conclusions and Discussion

In recent years, the packaging technology has focused on the WLP process. However, due to the limited data available at the early stages of the introduction of this technology, it is difficult for companies to effectively control it. To improve production yields, it is thus necessary for engineers to have an effective approach to forecast in this context, as if adverse phenomena are discovered in time; they can be addressed in a timely manner.

This paper presented a new adaptive model based on grey system theory to solve the small dataset forecasting problem

TABLE 5: Comparison of the forecasting methods.

Methods	MSE	MAE	MAPE (%)
AGM(1,1)	0.1692	0.3109	0.1595
BPNN	0.4283	0.5608	0.2876
SVR	0.2354	0.3784	0.1941

in this context. The experimental results show that the AGM(1,1) model can obtain satisfactory outcomes in the four pretesting measurements, outperforming the two popular forecasting techniques and thus that the grey approach has practical value for use in the packaging industries. Furthermore, one direction for future research would be to examine how to apply this method to other real-world industrial cases, such as the panel and biotechnology industries.

References

[1] D.-C. Li, C. Wu, and F. M. Chang, "Using data-fuzzification technology in small data set learning to improve FMS scheduling accuracy," *International Journal of Advanced Manufacturing Technology*, vol. 27, no. 3-4, pp. 321–328, 2005.

[2] D.-C. Li, C.-J. Chang, C.-C. Chen, and W.-C. Chen, "A grey-based fitting coefficient to build a hybrid forecasting model for small data sets," *Applied Mathematical Modelling*, vol. 36, pp. 5101–5108, 2012.

[3] V. C. Ivănescu, J. W. M. Bertrand, J. C. Fransoo, and J. P. C. Kleijnen, "Bootstrapping to solve the limited data problem in production control: an application in batch process industries," *Journal of the Operational Research Society*, vol. 57, no. 1, pp. 2–9, 2006.

[4] Y.-S. Lin and D.-C. Li, "The Generalized-Trend-Diffusion modeling algorithm for small data sets in the early stages of manufacturing systems," *European Journal of Operational Research*, vol. 207, no. 1, pp. 121–130, 2010.

[5] I. A. Gheyas and L. S. Smith, "A neural network-based framework for the reconstruction of incomplete data sets," *Neurocomputing*, vol. 73, no. 16-18, pp. 3039–3065, 2010.

[6] J. L. Deng, "Control problems of grey systems," *Systems and Control Letters*, vol. 1, no. 5, pp. 288–294, 1982.

[7] N. M. Xie, S. F. Liu, Y. J. Yang, and C. Q. Yuan, "On novel grey forecasting model based on non-homogeneous index sequence," *Applied Mathematical Modelling*, vol. 37, pp. 5059–5068, 2013.

[8] D.-C. Li, C.-J. Chang, C.-C. Chen, and W.-C. Chen, "Forecasting short-term electricity consumption using the adaptive grey-based approach-An Asian case," *Omega*, vol. 40, no. 6, pp. 767–773, 2012.

[9] J. Deng, "Introduction to grey system theory," *Journal of Grey System*, vol. 1, pp. 1–24, 1989.

[10] J. L. Deng, *The Primary Methods of Grey System Theory*, Huazhong University of Science and Technology Press, Wuhan, China, 2nd edition, 2005.

[11] Y.-Y. Tai, J.-Y. Lin, M.-S. Chen, and M.-C. Lin, "A grey decision and prediction model for investment in the core competitiveness of product development," *Technological Forecasting and Social Change*, vol. 78, no. 7, pp. 1254–1267, 2011.

[12] L.-C. Hsu and C.-H. Wang, "Forecasting integrated circuit output using multivariate grey model and grey relational analysis," *Expert Systems with Applications*, vol. 36, no. 2, pp. 1403–1409, 2009.

[13] R.-C. Tsaur and Y.-C. Liao, "Forecasting LCD TV demand using the fuzzy grey model GM(1,1)," *International Journal of Uncertainty, Fuzziness and Knowlege-Based Systems*, vol. 15, no. 6, pp. 753–767, 2007.

[14] D. Yamaguchi, G.-D. Li, and M. Nagai, "A grey-based rough approximation model for interval data processing," *Information Sciences*, vol. 177, no. 21, pp. 4727–4744, 2007.

[15] W. Köse, I. Temiz, and S. Erol, "Grey system approach for Economic Order Quantity models under uncertainty," *Journal of Grey System*, vol. 23, no. 1, pp. 71–82, 2011.

[16] Z. X. Wang, "An optimized Nash nonlinear grey Bernoulli model for forecasting the main economic indices of high technology enterprises in China," *Computers and Industrial Engineering*, vol. 64, pp. 780–787, 2013.

[17] E. Kayacan, B. Ulutas, and O. Kaynak, "Grey system theory-based models in time series prediction," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1784–1789, 2010.

[18] D.-C. Li, C.-W. Yeh, and C.-J. Chang, "An improved grey-based approach for early manufacturing data forecasting," *Computers and Industrial Engineering*, vol. 57, no. 4, pp. 1161–1167, 2009.

[19] D.-C. Li and C.-W. Yeh, "A non-parametric learning algorithm for small manufacturing data sets," *Expert Systems with Applications*, vol. 34, no. 1, pp. 391–398, 2008.

[20] D.-C. Li, C.-J. Chang, W.-C. Chen, and C.-C. Chen, "An extended grey forecasting model for omnidirectional forecasting considering data gap difference," *Applied Mathematical Modelling*, vol. 35, no. 10, pp. 5051–5058, 2011.

[21] S. F. Liu and Y. Lin, *Grey Information: Theory and Practical Applications*, Springer, London, UK, 1st edition, 2006.

[22] J. W. S. Hu, Y. C. Hu, and R. R. W. Lin, "Applying neural networks to prices prediction of crude oil futures," *Mathematical Problems in Engineering*, vol. 2012, Article ID 959040, 12 pages, 2012.

[23] J. L. Wu and P. C. Chang, "A trend-based segmentation method and the support vector regression for financial time series forecasting," *Mathematical Problems in Engineering*, vol. 2012, Article ID 615152, 20 pages, 2012.

[24] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufmann Publishers, San Francisco, Calif, USA, 2nd edition, 2005.

[25] J. T. Yokuma and J. S. Armstrong, "Beyond accuracy: comparison of criteria used to select forecasting methods," *International Journal of Forecasting*, vol. 11, no. 4, pp. 591–597, 1995.