



Forecasting value of agricultural imports using a novel two-stage hybrid model



Yi-Shian Lee^{a,b,*}, Wan-Yu Liu^c

^a Department of Industrial Engineering and Management, National Chiao Tung University, Hsinchu 300, Taiwan, ROC

^b Research Center for Psychological and Educational Testing, National Taiwan Normal University, Taipei 106, Taiwan, ROC

^c Department of Tourism Information, Aletheia University, New Taipei City 251, Taiwan, ROC

ARTICLE INFO

Article history:

Received 23 July 2013

Received in revised form 24 February 2014

Accepted 27 March 2014

Keywords:

Value of agricultural imports

GM(1,1)

Genetic programming

Residual signs

Residual series

ABSTRACT

Agricultural imports are becoming increasingly important in terms of their impact on economic development. An accurate model must be developed for forecasting the value of agricultural imports since rapid changes in industry and economic policy affect the value of agricultural imports. Conventionally, the ARIMA model has been utilized to forecast the value of agricultural imports, but it generally requires a large sample size and several statistical assumptions. Some studies have applied nonlinear methods such as the GM(1,1) and improved GM(1,1) models, yet neglected the importance of enhancing the accuracy of residual signs and residual series. Therefore, this study develops a novel two-stage forecasting model that combines the GM(1,1) model with genetic programming to accurately forecast the value of agricultural imports. Moreover, accuracy of the proposed model is demonstrated based on two agricultural imports data sets from the Taiwan and USA.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Since agricultural development is critical to the economic development of every country, agricultural issues are of global concern. Governments must devise viable economic policies to avoid unnecessary costs that are incurred with increasing agricultural imports. For example, after joining the World Trade Organization (WTO) in 2002, Taiwan signed the Economic Cooperation Framework Agreement (ECFA) in 2012 for reducing commercial barriers with China, drastically changing the value of agricultural imports. Since economic forecasting in the agricultural sector is critical to agricultural business planning and economic policy making, a high-precision forecasting approach must be designed to evaluate agricultural imports to enable policy makers to implement effective policies concerning agricultural imports and enhance economic development.

Relevant literature includes using various forecasting approaches to forecast agricultural demand (Lambert and Cho, 2008). Multiple linear regression and Box–Jenkins models (Agrawal, 2003; Lambert and Cho, 2008) are two conventional statistical methods. However, those approaches may be inaccurate

when data sets are small and nonlinear, as well as fail to meet certain statistical assumptions (Lee and Tong, 2011b; Pao, 2009). Hence, the forecasting accuracy of traditional statistical methods often varies under real-life condition (Yang et al., 2009). With the development of advances in machine-learning methods, some algorithms such as artificial neural network (ANN) and genetic algorithms (GAs), have been utilized in agricultural forecasting. For example, Jutras et al. (2009) adopted the ANN to predict the morphological parameters of street trees and found that the ANN can yield robust and precise results. Yang et al. (2009) combined principal component analysis and ANN to predict the population of the paddy stem borer (*Scirpophaga incertulas*), indicating that their proposed model outperformed other models. Ou (2012) proposed an improved forecasting model that combined improved GM(1,1) (IGM(1,1)) applied in modeling original time series and GAs applied in estimating the parameters of IGM(1,1), and demonstrated that the proposed model outperformed other models. Despite yielding satisfactory results for real-world data sets, the above methods have certain limitations. For instance, the hidden layers in ANN are difficult to explain, and the relationship between the independent and dependent variables cannot be expressed as a clear mathematical equation (Lee and Tong, 2011b). Moreover, the high precision of the above approaches depends on the sample sizes and the parameter settings that are determined by a trial and error approach. Using neural network-based models to construct an optimal network model is often criticized, owing to the

* Corresponding author at: Research Center for Psychological and Educational Testing, National Taiwan Normal University, Taipei 106, Taiwan, ROC.

E-mail address: bill.net.tw@yahoo.com.tw (Y.-S. Lee).

lack of openness and shift of emphasis towards training the network model (Srinivasan, 2008). Since data on agricultural imports are generally few and nonlinear, they may not yield accurate forecasting results when conventional statistical methods are applied.

Nonlinear or small-size time-series data sets are handled using approaches such as fuzzy theory, grey model (GM), and genetic programming (GP). The observations (real numbers) of fuzzy time series in a certain period are converted as discrete fuzzy sets (Egrioglu et al., 2011a). The procedure of fuzzy time series consists of three stages: fuzzification, determination of fuzzy relations and defuzzification (Song and Chissom, 1993). Some studies have attempted to increase forecasting accuracy by developing fuzzy-based approaches. For instances, Egrioglu et al. (2011a) determined an appropriate number of fuzzy clusters by using the Gustafson–Kessel fuzzy clustering algorithm and, later, determined the length of intervals of fuzzy time by using an optimization technique (2011b). Despite the applicability of the fuzzy-based approach to small data sets, determining an appropriate length of intervals based on different algorithms may expend a considerable amount of time. As useful in forecasting problems (Ou, 2012; Lee and Tong, 2011a; Yin and Tang, 2013; Pao et al., 2012; Chang et al., 2013), GM is often used in forecasting when data sets contain more than four samples (Wu et al., 2013). GM can generally be represented as $GM(g, h)$, where g and h denote the order and number of variables in constructing the GM, respectively. For example, $GM(1, 1)$ represents the first-order single-variable GM, and has been used to forecast agricultural output (Ou, 2012). To enhance the accuracy of $GM(1, 1)$ in the construction of agricultural demand values (including the value of agricultural imports/exports), some studies have modified $GM(1, 1)$ models (Ou, 2012). Although capable of yielding accurate forecasting results, the modified $GM(1, 1)$ belong to the GM system in order to obtain values of necessary parameters. However, few studies have improved the residual time-series data of the $GM(1, 1)$ with a machine-learning approach. Recently, some hybrid forecasting models have been proposed to improve the performance, which can be achieved using only a single forecasting method (Zhou and Hu, 2008; Pai and Lin, 2005; Aladag et al., 2009; Wang et al., 2012; Yolcu et al., 2013; Khashei and Bijari, 2012). For instances, Khashei and Bijari (2012) forecasted time series data by using probabilistic neural networks with feed-forward neural networks. Yolcu et al. (2013) performed time series forecasting by using linear and nonlinear ANN model. A criticism of ANN is the difficulty to explain the layers and neurons in its hidden-layer. Moreover, those studies have ignored the importance of residual-sign estimator. According to some studies (Hsu and Chen, 2003; Hsu, 2003; Lee and Tong, 2011a), the accuracy of the estimator of residual signs can influence the performance of a forecasting model. Moreover, using a complex residual equation to obtain the forecast residual values makes it difficult to use the hybrid model.

GP is an approach for evolving the functions that performs well in the defined problems (Koza, 1992) and constructs a forecasting model by using the symbolic regression method. The intelligence scheme can automatically extract knowledge from data sets and construct the model without defining related problems. The approach used in this paper is based on GP, owing to that GP often performs better than conventional statistical methods, in terms of forecasting accuracy. Although the performances of all forecasting models depend on the quality of the data set, these models differ in the ability to mine the inherent relationships in the data set. Most real-world data sets are nonlinear and time-dependent. GP is a relatively easy means of constructing mathematical models since no specialized knowledge. In some modeling time series applications, GP performs well in small data sets. For instance, based on a multi-level genetic programming (MLGP) approach, Forouzanfar et al. (2012) developed a transport energy demand forecasting model

(training set: 35 samples from year 1968 to 2002; testing set: 3 sample size which from year 2003 to 2005), which is more accurate than other models. By using a GP approach, Lee et al. (1997) designed an electric power demand forecasting model (training set: 20 samples from year 1961 to 1980; testing set: 10 samples from year 1981 to 1990), which is more accurate than the conventional regression model. Moreover, while developing the classification model, Lee and Tong (2012) predict the transfer efficiency of photovoltaic systems by using a GP-based model; the classification model outperforms other models on small photovoltaic data sets. Some studies (Huang et al., 2006; Muttill and Lee, 2005) demonstrated that GP can perform well even in small data sets.

This study develops a novel two-stage forecasting model that first utilizes $GM(1, 1)$ to forecast original data based on the advantage of being applied to small data sets, and then uses GP to forecast the residual signs and residual series of $GM(1, 1)$ based on the advantage of adopting symbolic regression to model complex data sets, to increase its accuracy in forecasting the value of agricultural imports. Analysis results demonstrate that the proposed model is easily applied in practice and performs well in modeling time-series data sets. The rest of this paper is organized as follows. Section 2 examines the feasibility of improving the grey forecasting model, which includes $GM(1, 1)$, to forecast the original data sets. The ability to use GP in order to forecast the residual signs and residual series of $GM(1, 1)$ is examined as well. Section 3 then presents two data sets to demonstrate the application of the proposed model, which is compared with other models. Conclusions are finally drawn in Section 4, along with recommendations for future research.

2. Methodology

2.1. $GM(1, 1)$ forecasting model

The $GM(1, 1)$ has been utilized in agriculture (Ou, 2012) and high-tech industry (Hsu, 2003; Hsu and Wang, 2007; Wang et al., 2011). $GM(1, 1)$ usually requires only four or more data points (Hsu, 2009) to construct a forecasting model. $GM(1, 1)$ is constructed as follows.

The general procedure for constructing a $GM(1, 1)$ is given as follows.

Collect an original non-negative time-series data sequence,

$$w^{(0)} = [w^{(0)}(1), w^{(0)}(2), \dots, w^{(0)}(n)], \quad n \geq 4 \quad (1)$$

where n is the total number of periods, and $w^{(0)}(n)$ is the observation that is associated with the n th time period.

The technique applies the accumulated generating operator (AGO) to $w^{(0)}$ to obtain an accumulated data sequence, as follows.

$$\begin{aligned} w^{(1)} &= \left(w^{(0)}(1), \sum_{m=1}^2 w^{(0)}(m), \dots, \sum_{m=1}^n w^{(0)}(m) \right) \\ &= (w^{(1)}(1), w^{(1)}(2), \dots, w^{(1)}(n)), \end{aligned} \quad (2)$$

where $w^{(0)}(1)$ equals $w^{(1)}(1)$.

$GM(1, 1)$ is constructed using the following grey differential equation.

$$w^{(0)}(k) + a \times s^{(1)}(k) = u, \quad k = 2, 3, \dots, n \quad (3)$$

where a , u , and $s^{(1)}(k)$ represent the development coefficient, grey input, and background value, respectively. Notably, $s^{(1)}(k)$ is obtained by applying the mean operator to $w^{(1)}$, as follows.

$$s^{(1)}(k) = \frac{w^{(1)}(k) + w^{(1)}(k-1)}{2}, \quad k = 2, 3, \dots, n. \quad (4)$$

The solution for $w^{(1)}(k)$ in Eq. (4) can be estimated using the ordinary least squares (OLS) method, as follows.

$$w^{(1)}(k) = \left(w^{(0)}(1) - \frac{\hat{u}}{\hat{a}} \right) \times e^{-\hat{a}(k-1)} + \frac{\hat{u}}{\hat{a}}, \quad k = 2, 3, \dots, n \quad (5)$$

where

$$\begin{bmatrix} \hat{a} \\ \hat{u} \end{bmatrix} = (B^T B)^{-1} B^T W, \quad (6)$$

and

$$B = \begin{bmatrix} -w^{(1)}(2) & -w^{(1)}(3) & \dots & -w^{(1)}(n) \\ 1 & 1 & \dots & 1 \end{bmatrix}^T, \quad (7)$$

$$W = [w^{(0)}(2), w^{(0)}(3), \dots, w^{(0)}(n)]^T. \quad (8)$$

Finally, the GM(1,1) forecasting equation can be obtained using the inverse AGO technique, as follows.

$$\begin{aligned} \hat{w}^{(0)}(t) &= \hat{w}^{(1)}(t) - \hat{w}^{(1)}(t-1) \\ &= \left(w^{(0)}(1) - \frac{\hat{u}}{\hat{a}} \right) \times (1 - e^{\hat{a}}) \times e^{-\hat{a} \times (t-1)}, \quad t = 2, 3, \dots \end{aligned} \quad (9)$$

2.2. Forecasting residual signs and residual series of GM(1,1) based on GP

This study develops a novel improved GM(1,1) model to enhance the accuracy of GM(1,1). The difference between the target values $w^{(0)}$ and the predicted values $\hat{w}^{(0)}$ is called the residual series. The modified forecasted values are obtained by combining the original GM(1,1) and the residual component: residual signs and the residual series of GM(1,1), which increase the accuracy based on GP. Some studies have demonstrated that the effectiveness of the residual series of GM(1,1) depends on the number of observations with the same sign (Hsu and Chen, 2003; Hsu, 2003; Lee and Tong, 2011a). Notably, the residual GM(1,1) model cannot be constructed if the number of observations with the same sign does not exceed four (Hsu and Chen, 2003; Lee and Tong, 2011a). Despite their use of different machine-learning approaches to forecast the residual signs (ANN and GP, respectively), Hsu and Chen (2003) and Lee and Tong (2011a) failed to consider the weak performance of GM(1,1) model when its residual series are complex or do not fit the exponential curve. Some studies have attempted to enhance the accuracy of residual series of GM(1,1) by using various approaches. For example, although Zhou and Hu (2008) modeled the residual series of GM(1,1) by using ARIMA, its performance may be weak when the sample was small or did not meet statistical assumptions.

Koza (1992) proposed the GP as a new algorithm for computer programs that exploits the concept of evolution to identify problems (as in modeling time-series data sets). GP can automatically create computer programs to solve problems according to the following two principles (Robinson, 2001): (a) it can make developing difficult algorithms easier (b) it might perform better than other fitness-driven automatic programming techniques such as hill climbing and simulated annealing. Like GAs, the GP uses mutation, crossover, and reproduction (Sette and Boullart, 2001) in modeling identification problem. Fig. 1 presents a GP parse tree that is used to express a simple example of $\exp[(6-x) \times (8+y)]$. To obtain an appropriate forecasting model, the user has to adopt a trial-and-error method to identify a number of GP parameters when applying the approach in the data-set modeling. GP has become more popular than conventional linear forecasting methods (e.g., ARIMA) because it can be applied to find complex nonlinear solutions. As compared to other popular machine-learning methods such as

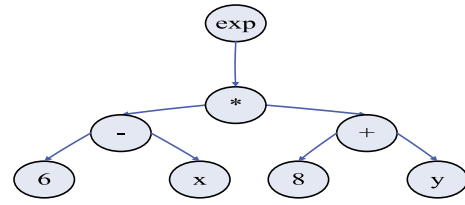


Fig. 1. Demonstration of GP parse tree.

ANN, GP is more accurate in modeling and does not need a large sample size to train the data sets (Forouzanfar et al., 2012).

This study attempts to increase the forecasting accuracy of GM(1,1) by combining GM(1,1) applied in the original time-series data and GP applied to forecast the residual signs and residual series of GM(1,1) based on the strength of the parse-tree function and their satisfactory performance with small sample sizes. Fig. 2 depicts the construction of the proposed model.

Firstly, in forecasting the residual signs, a dummy variable $d(t)$ is adopted to reveal the sign of the residual in the t th year. If the residual sign for the t th year is positive, then the value of $d(t)$ is one; otherwise, $d(t)$ is zero. For selecting the number of lagged residual variables, some hybrid forecasting models (Aladag et al., 2009; Zhang, 2003; Forouzanfar et al., 2012; Hsu and Chen, 2003; Lee and Tong, 2011a) adopted different approaches such as a trial and error method (Aladag et al., 2009; Lee and Tong, 2011b) or given selected criterion (Zhang, 2003; Forouzanfar et al., 2012; Hsu and Chen, 2003; Lee and Tong, 2011a) to determine the lagged residual variables. This study attempts to determine the lagged residual variables (including residual signs and series) by using a given selected criterion which is referred to Hsu and Chen (2003) and Lee and Tong (2011a). Moreover, in order to compare with different selected lagged residual variables, this study also adopted one lagged residual variable as the input element of GP model. The GP model for residual signs (giving two lagged residual signs) is represented as follows.

$$\hat{d}(t) = f(d(t-1), d(t-2)), \quad (10)$$

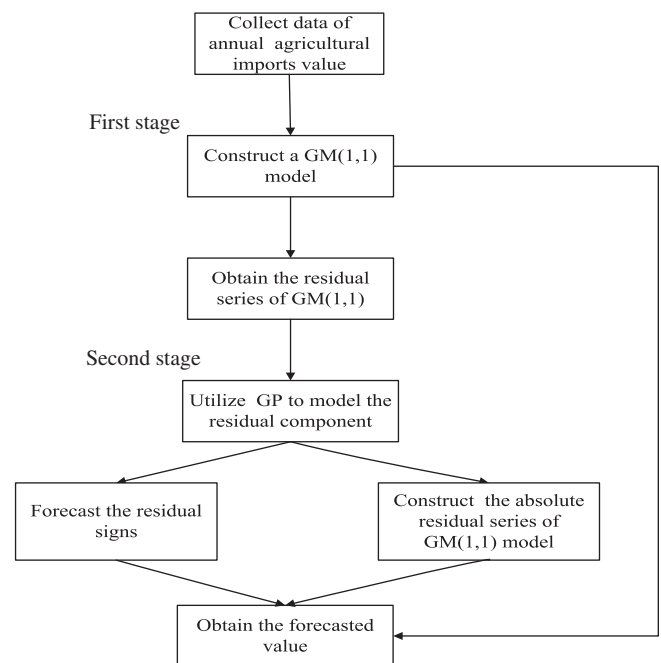


Fig. 2. Procedure of the proposed two-stage approach.

Table 1
Parameter settings of GP for the one and two lagged residual signs.

Parameter	Value
Population size	100
Maximum number of generation	1000
Fitness function	Minimize: $\sum_{t=1}^n \hat{d}(t) - d(t) $, $\sum_{t=1}^n \hat{d}'(t) - d'(t) $
Function set	+, -, ×, ÷, sin, cos, exp, log, constant
Crossover rate	0.9
Mutation rate	0.1
Simulation time	10

where $\hat{d}(t)$ represents the forecasted residual sign in the t th year. The mathematical function $f(d(t-1), d(t-2))$ represents the nonlinear function that is constructed using GP, with independent variables $d(t-1)$ and $d(t-2)$. Table 1 lists the parameters of GP when applied in one and two lagged residual signs. Based on the GP settings, the sign of the t th year residual, $s(t)$ (based on two lagged residual signs as input variables of GP model), can be expressed as

$$s(t) = \begin{cases} 1, & \text{if } \hat{d}(t) = 1 \\ -1 & \text{if } \hat{d}(t) = 0 \end{cases}, \quad t = 1, 2, 3, \dots \quad (11)$$

Secondly, the residual series of GM(1,1) is constructed using GP. The algorithm is as follows.

Assume that the original absolute values of a residual sequence are given by $r^{(0)}$,

$$r^{(0)} = (e^{(0)}(2), e^{(0)}(3), e^{(0)}(4), \dots, e^{(0)}(t)) \quad (12)$$

where

$$r^{(0)}(t) = e^{(0)}(t) = |w^{(0)}(t) - \hat{w}^{(0)}(t)|, \quad t = 2, 3, \dots \quad (13)$$

First, based on the lagged absolute residual time sequence of $e^{(0)}(t-1)$ and $e^{(0)}(t-2)$, $e^{(0)}(t)$ is forecasted using the parse-tree function in GP. Some studies (Hsu and Chen, 2003; Lee and Tong, 2011a) focus mainly on the selected lagged residual series ($e^{(0)}(t-1)$ and $e^{(0)}(t-2)$) when selecting the number of residual variables. Hence, the GP model, used in the absolute residual series of GM(1,1) can be represented as follows.

$$\hat{e}^{(0)}(t) = f(e^{(0)}(t-1), e^{(0)}(t-2)), \quad (14)$$

where $\hat{e}^{(0)}(t)$ represents the forecasted value of residual series in the t th year. Also, the mathematical function $f(e^{(0)}(t-1), e^{(0)}(t-2))$ is the nonlinear function that is constructed using GP, with independent variables $e^{(0)}(t-1)$ and $e^{(0)}(t-2)$. Table 2 lists the GP parameters in the absolute one and two lagged residual series of GM(1,1).

Hence, the two-stage forecasting model when adopting one and two lagged residual variables can be obtained, respectively

$$\hat{w}_{one-lagged}^{(0)}(t) = \left(w^{(0)}(1) - \frac{\hat{u}}{\hat{a}} \right) \times (1 - e^{\hat{a}}) \times e^{-\hat{a} \times (t-1)} + \{s'(t) \times f(e^{(0)}(t-1))\}, \quad (15)$$

Table 2
Parameter settings of GP for the absolute one and two lagged residual series of GM(1,1).

Parameter	Value
Population size	100
Maximum number of generation	1000
Fitness function	Minimize: $\sum_{t=1}^n \hat{e}^{(0)}(t) - e^{(0)}(t) $, $\sum_{t=1}^n (\hat{e}^{(0)}(t) - e^{(0)}(t)) $
Function set	+, -, ×, ÷, sin, cos, exp, log, constant
Crossover rate	0.9
Mutation rate	0.1
Simulation time	10

$s'(t)$ can be expressed as:

$$s'(t) = \begin{cases} 1, & \text{if } \hat{d}'(t) = 1 \\ -1, & \text{if } \hat{d}'(t) = 0 \end{cases}, \quad t = 1, 2, 3, \dots \quad (16)$$

and $\hat{d}'(t) = f(d(t-1))$, $t = 1, 2, \dots, n, n+1, \dots$

$$\hat{w}_{two-lagged}^{(0)}(t) = \left(w^{(0)}(1) - \frac{\hat{u}}{\hat{a}} \right) \times (1 - e^{\hat{a}}) \times e^{-\hat{a} \times (t-1)} + \{s(t) \times f(e^{(0)}(t-1), e^{(0)}(t-2))\}, \quad t = 1, 2, \dots, n, n+1, \dots \quad (17)$$

3. Computational results

3.1. Data sources

The performance of the proposed model is evaluated using two agricultural imports data sets. First, effectiveness of the proposed two-stage model is demonstrated based on agricultural import data in Taiwan from 2002 to 2011. Above data are obtained from the Annual Report of the Council of Agriculture, Executive Yuan (Taiwan). The historical values of agricultural imports in Taiwan from 2002 to 2009 are utilized as the training data and the data for 2010–2011 are utilized for testing. The second data sets are annual agricultural imports of USA during 2002–2011, based on data are obtained from the United States Department of Agriculture (USDA). The historical values of agricultural imports in USA from 2002 to 2009 are utilized as the training data and the data for 2010–2011 are utilized for testing. Moreover, the proposed model (including one and two lagged residual variables as input variables of GP model) is also compared with other models,

Table 3
MAPE criteria for model assessment. Source: Delurgio (1998).

MAPE (%)	Forecasting level
<10	High forecasting
10–20	Good forecasting
20–50	Reasonable forecasting
>50	Weak forecasting

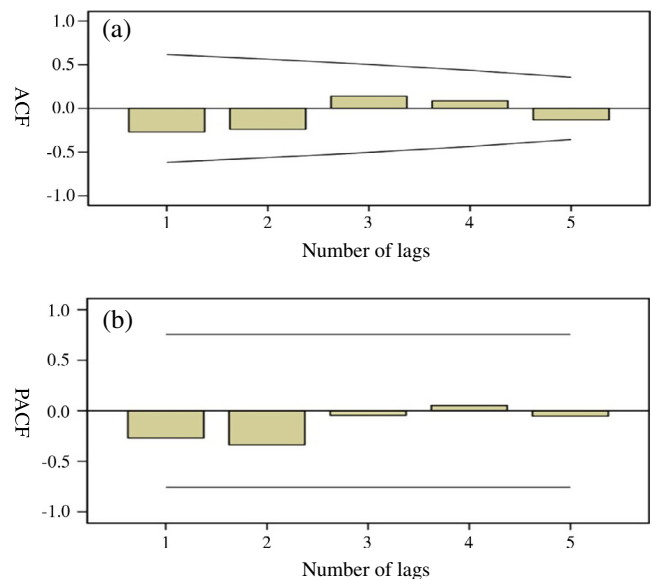


Fig. 3. ACF and PACF results for the agricultural imports data sets in Taiwan.

including GM(1,1) model, the traditional ARIMA model, the improved GM(1,1) model by Lee and Tong (2011a).

3.2. Performance evaluation

The above forecasting models are compared in terms of prediction accuracy based on two evaluation indices. Predictive accuracies of the above four forecasting models are compared using two indices. The first index is the percentage error (PE), which is defined as follows.

$$PE = \frac{\hat{w}^{(0)}(t) - w^{(0)}(t)}{w^{(0)}(t)} \times 100\%, \tag{18}$$

where $\hat{w}^{(0)}(t)$ denotes the forecasted value, and $w^{(0)}(t)$ denotes the actual value. Similar to the first index, the accuracy of the conven-

tional forecasting model popularly is evaluated using the mean absolute percentage error (MAPE). MAPE is defined as follows.

$$MAPE = \frac{\sum_{t=1}^n |\hat{w}^{(0)}(t) - w^{(0)}(t)/w^{(0)}(t)|}{N} \times 100\%, \tag{19}$$

DeLurgio (1998) indicated that using MAPE facilitates the evaluation of the forecasting model accuracy. Table 3 summarizes the criteria for evaluating the models.

3.3. The results of forecasting the agricultural imports data sets in Taiwan

Results of the four forecasting models (i.e. the GM(1,1) model, the improved GM(1,1) model by Lee and Tong (2011a), the ARIMA model, and the proposed model) for Taiwan's agricultural imports data sets are expressed as follows:

Autocorrelations: Residual series of ARIMA model

Lag	Auto-Corr.	Stand. Err.	-1	-.75	-.5	-.25	0	.25	.5	.75	1	Box-Ljung Prob.
1	-.270	.309					*****					.763 .382
2	-.240	.282					*****					1.487 .475
3	.140	.252						***				1.796 .616
4	.088	.218						**				1.957 .744
5	-.132	.178						***				2.502 .776

Plot Symbols: Autocorrelations * Two Standard Error Limits .

Fig. 4. Box–Ljung evaluation results for the agricultural imports data sets in Taiwan.

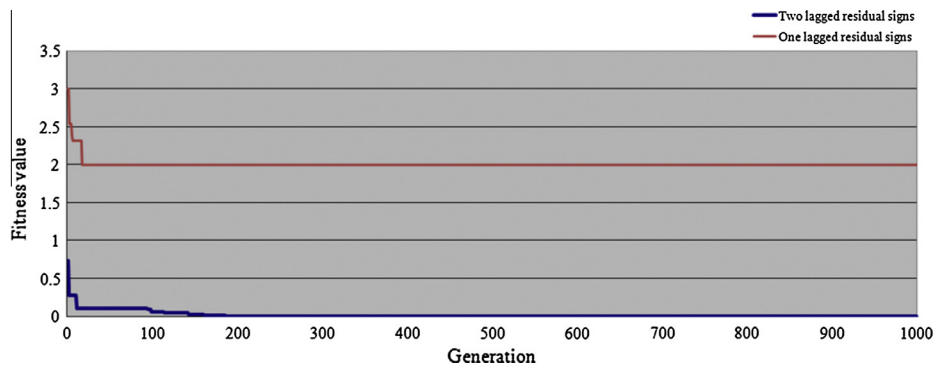


Fig. 5. Fitness value of the best simulated one and two lagged residual signs over 1000 generations (Taiwan).

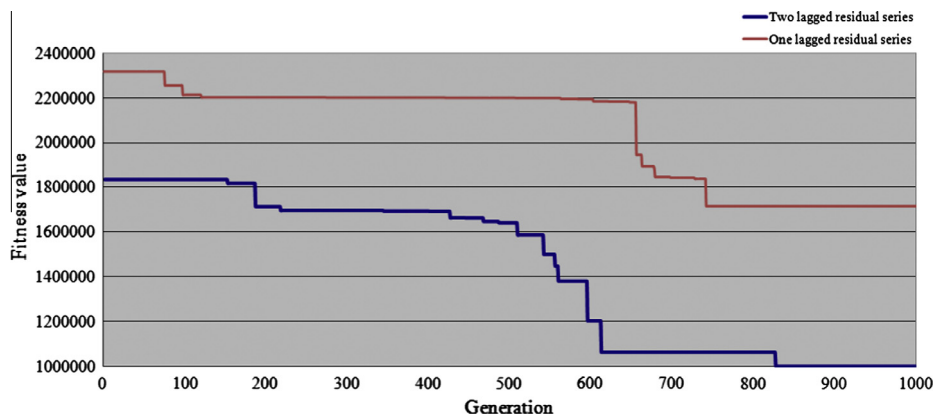


Fig. 6. Fitness value of the best simulation absolute one and two lagged residual series of GM(1,1) over 1000 generations (Taiwan).

Table 4
Forecasted values and errors among models for the agricultural imports value in Taiwan (unit: US\$1000).

Year	Original value	GM(1,1)		Lee and Tong (2011a)		ARIMA		Proposed (one lagged)		Proposed (two lagged)	
		Model value	PE	Model value	PE	Model value	PE	Model value	PE	Model value	PE
2002	7105407	7105407.00	0.00	7105407.00	0.00	7105407.00	0.00	7105407.00	0.00	7105407.00	0.00
2003	7829519	8312247.17	6.17	7829519.00	0.00	7525528.43	-3.88	7829512.64	0.00	7829519.00	0.00
2004	8862024	8745198.19	-1.32	8899769.77	0.43	8249640.43	-6.91	9228093.16	4.13	8958911.39	1.09
2005	9355094	9200699.87	-1.65	9450835.99	1.02	9282145.43	-0.78	9083707.20	2.90	9366267.02	0.12
2006	9428136	9679926.77	2.67	9275142.91	-1.62	9775215.43	3.68	9525365.80	1.03	9442101.34	0.15
2007	10456064	10184114.64	-2.60	10839157.86	3.66	9848257.43	-5.81	10436072.22	0.19	10450568.81	-0.05
2008	12121293	10714563.60	-11.61	11774590.10	-2.86	10876185.43	-10.27	10442447.42	13.85	11463397.25	-5.43
2009	10046257	11272641.48	12.21	9557249.04	-4.87	12541414.43	24.84	9865745.29	1.80	10260280.11	2.13
MAPE (%) (2002–2009)			4.78		1.81		7.02		2.99		1.12
2010	12759852	11859787.36	-7.05	14635728.55	14.70	10466378.43	-17.97	12949018.19	1.48	13673908.76	7.16
2011	14842035	12477515.27	-15.93	16969694.03	14.34	10886499.86	-26.65	11510081.71	22.45	14243014.18	-4.04
MAPE (%) (2010–2011)			11.49		14.52		22.31		11.97		5.60

$$PE = \frac{\hat{w}^{(0)}(t) - w^{(0)}(t)}{w^{(0)}(t)} \times 100\%$$

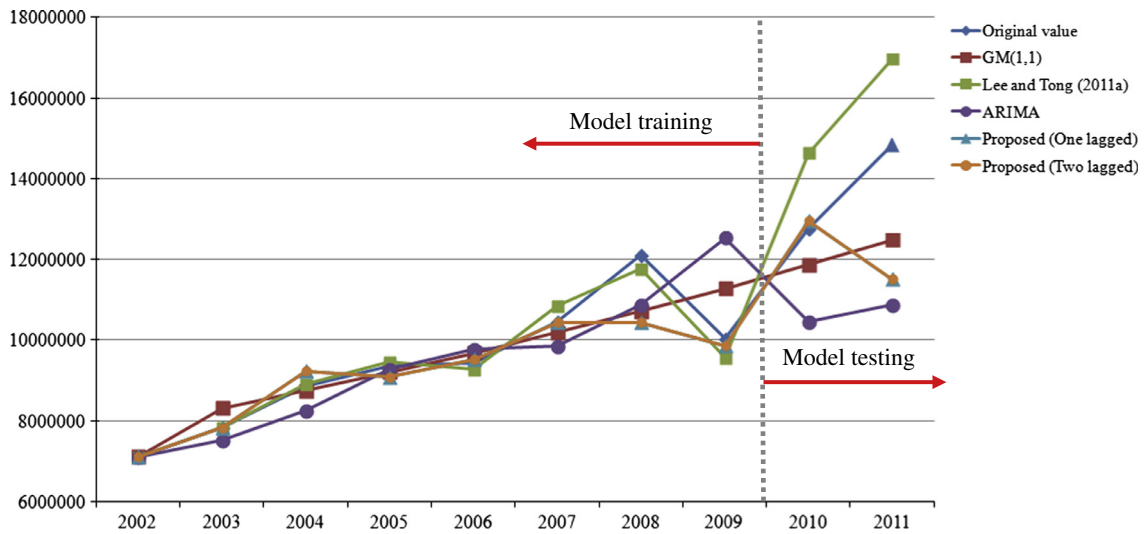


Fig. 7. Distributions of forecasted values and real values from 2002 to 2011 for the agricultural imports data sets in Taiwan.

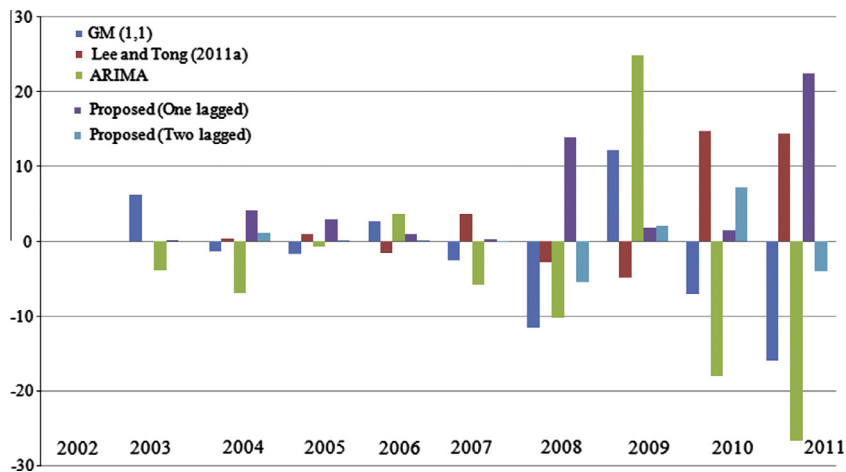


Fig. 8. Trend of percentage error (%) among the four models from 2002 to 2011 for the agricultural imports data sets in Taiwan.

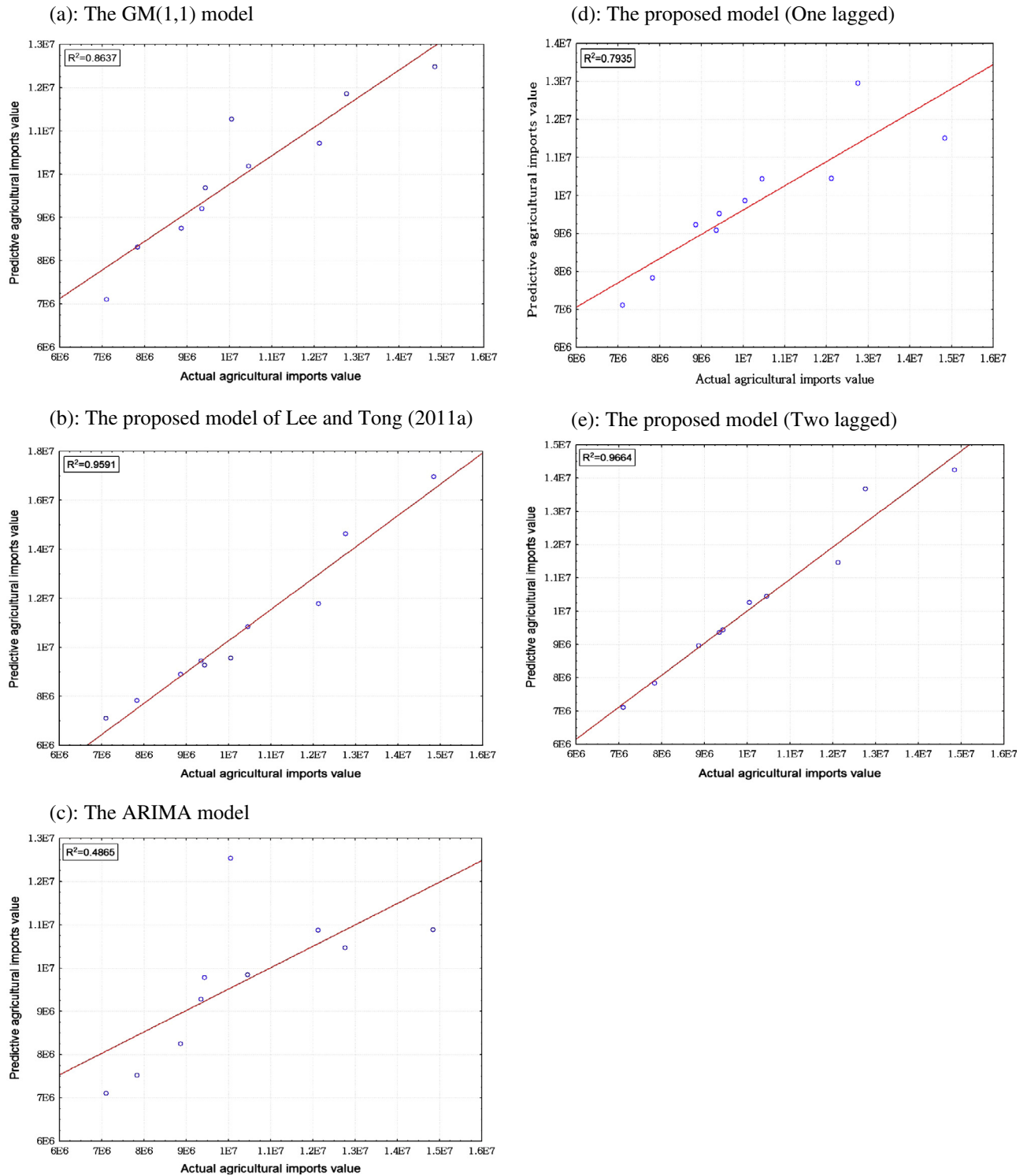


Fig. 9. Scatter plots among different forecasting models for the agricultural imports data sets in Taiwan.

(1) The GM(1,1) forecasting equation is:

$$\hat{w}^{(0)}(t) = 7900730.716 \times \exp(0.0507747821 \times (t - 1)), \quad t = 2, 3, \dots \quad (20)$$

(2) The improved GM(1,1) forecasting equation by Lee and Tong (2011a) is:

$$\hat{w}^{(0)}(t) = 7900730.716 \times \exp(0.0507747821 \times (t - 1)) + s(t) \times \{95517.486 \times \exp(0.4813479730 \times (t - 1))\}, \quad t = 1, 2, \dots, \quad (21)$$

where $s(t)$ is the binary variable (1 or -1). The forecasted residual sign is obtained using GP (independent variables: $d(t - 1)$ and $d(t - 2)$, dependent variable: $d(t)$). The parameter settings of GP

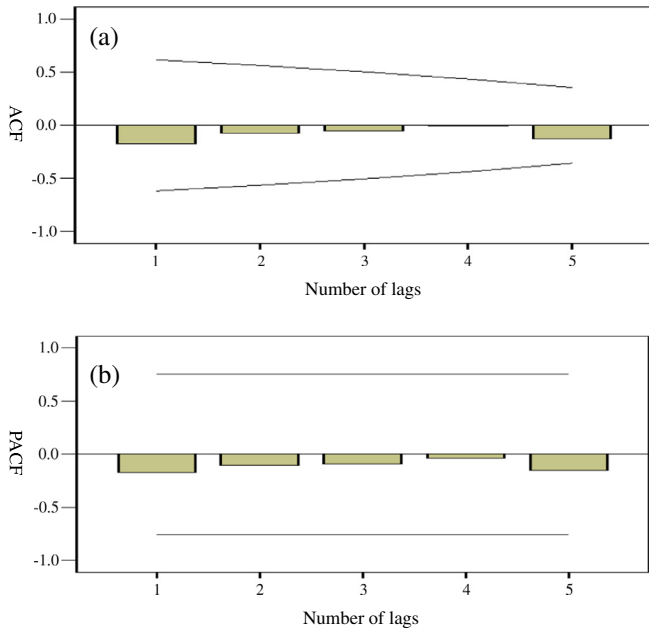


Fig. 10. ACF and PACF results for the agricultural imports data sets in USA.

for the population size, maximum number of generation, crossover rate, and mutation rate are set to 100, 1000, 0.9, and 0.1, respectively. If $\hat{d}(t)$ is 1, then $s(t)$ represents 1; otherwise, $s(t)$ represents -1 ($\hat{d}(t)$ is 0).

(3) The ARIMA forecasting equation is:

Based on SPSS software, the order of the ARIMA model is modified using the autocorrelation function (ACF) and partial autocorrelation function (PACF) criteria in the training data. Fig. 3

summarizes the results of ACF and PACF. Therefore, it can be obviously seen that the differencing order: 1 to be the stationary state. The forecasting model can be represented as follows:

$$\hat{w}^{(0)}(t) = 420121.4300 + w^{(0)}(t - 1), \quad t = 2, 3, \dots \quad (22)$$

Moreover, based on the Akaike Information Criterion (AIC) rule (Harvey, 1981) and Box–Ljung test, an appropriate ARIMA model is determined and the residual series of ARIMA is evaluated. In this study, the AIC value of Eq. (22) is 216.9143. Fig. 4 shows the Box–Ljung test. This figure demonstrates that the residual series of ARIMA does not exhibit significant lack of fit.

(4) The proposed two-stage model can be obtained as follows.

Based on the GM(1,1) model, the residual signs can then be evaluated by the GP model. Table 1 lists the parameter settings based on GP for residual signs. Among the 10 simulation times, this study demonstrates the best simulation results (based on the fitness function) in Fig. 5. Hence, the forecasting equations of one and two lagged residual signs are represented, respectively, as follows:

$$\hat{d}'(t) = f(d(t - 1)) = (\exp(0) - d(t - 1)). \quad (23)$$

$$\begin{aligned} \hat{d}(t) &= f(d(t - 1), d(t - 2)) \\ &= (\exp(\exp(\exp(\exp(d(t - 2)) - 1.852) - 1.852) - 1.852) - 1.852) \\ &\quad \times ((\exp(1.852)) \times (1.852^2)) \times (((d(t - 2)) \times ((d(t - 1)) \\ &\quad \times 1.852)) \times (1.852^2))). \end{aligned} \quad (24)$$

This study also demonstrates the best one among 10 times for the absolute one and two lagged residual series of GM(1,1). Table 2 lists the parameter settings of GP for the absolute residual series of GM(1,1). Fig. 6 shows the fitness value of the best simulation for absolute one and two lagged residual series of GM(1,1). Hence, the forecasting equation for absolute one and two lagged residual series of GM(1,1) can be represented, respectively.

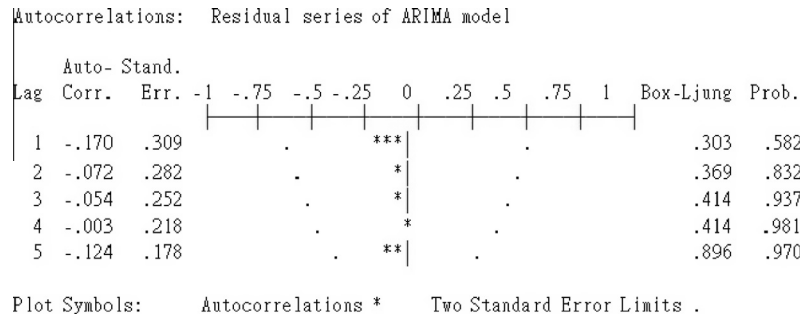


Fig. 11. Box–Ljung evaluation results for the agricultural imports data sets in USA.

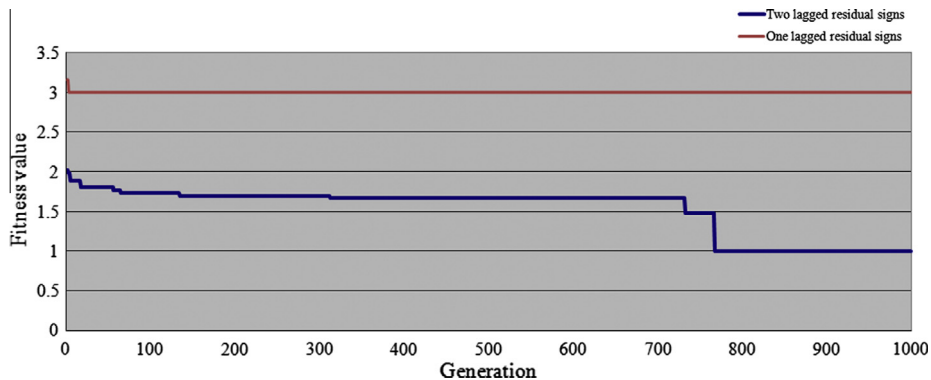


Fig. 12. Fitness value of the best simulated one and two lagged residual signs over 1000 generations (USA).

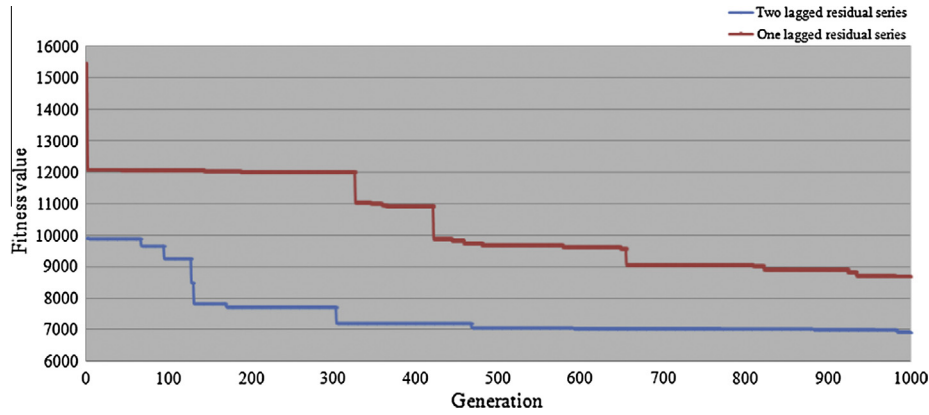


Fig. 13. Fitness value of the best simulation absolute one and two lagged residual series of GM(1,1) over 1000 generations (USA).

Table 5

Forecasted values and errors among the models for the agricultural imports value in USA (unit: \$ Million).

Year	Original value	GM(1,1)		Lee and Tong (2011a)		ARIMA		Proposed (one lagged)		Proposed (two lagged)	
		Model value	PE	Model value	PE	Model value	PE	Model value	PE	Model value	PE
2002	41915.26	41915.26	0.00	41915.26	0.00	41915.26	0.00	41915.26	0.00	41915.26	0.00
2003	47383.70	50963.85	7.56	47383.70	0.00	46167.51	-2.57	51384.68	8.44	47383.70	0.00
2004	53989.18	54880.07	1.65	53923.47	-0.12	51635.94	-4.36	58604.00	8.55	53722.47	-0.49
2005	59291.06	59097.23	-0.33	60633.82	2.26	58241.43	-1.77	60489.62	2.02	60152.38	1.45
2006	65325.76	63638.44	-2.58	66106.68	1.20	63543.31	-2.73	64187.68	1.74	63866.32	-2.23
2007	71913.00	68528.62	-4.71	72493.35	0.81	69578.01	-3.25	71659.27	0.35	70410.01	-2.09
2008	80487.69	73794.57	-8.32	80163.12	-0.40	76165.24	-5.37	80477.61	0.01	77782.76	-3.36
2009	71680.99	79465.18	10.86	89694.99	25.13	84739.94	18.22	87366.06	21.88	87383.12	21.91
MAPE (%) (2002–2009)			4.50		3.74		4.78		5.38		3.94
2010	81857.55	85571.53	4.54	69139.38	-15.54	75933.24	-7.24	93654.64	14.41	76331.03	-6.75
2011	98948.79	92147.12	-6.87	118542.11	19.80	80185.49	-18.96	104217.70	5.32	103117.90	4.21
MAPE (%) (2010–2011)			5.71		17.67		13.10		9.87		5.48

$$PE = \frac{w^{(0)}(t) - w^{(0)}(t)}{w^{(0)}(t)} \times 100\%$$

$$\begin{aligned} \hat{e}^{(0)}(t) &= f(e^{(0)}(t-1)) \\ &= ((((((e^{(0)}(t-1)) - (-27.488)) - ((-27.488) \\ &+ (-37.114))) - (((e^{(0)}(t-1)) - (-8.931) \\ &- (-8.931)))/(((e^{(0)}(t-1)) \times 0) - (-8.931)))) \\ &\times (\exp(((-18.557 \times -3.17) - (-17.863)))/((e^{(0)}(t-1)) \\ &- (-8.931))))). \end{aligned} \tag{25}$$

$$\begin{aligned} \hat{e}^{(0)}(t) &= f(e^{(0)}(t-1), e^{(0)}(t-2)) \\ &= ((((((e^{(0)}(t-1)) \times (\text{Sin}((e^{(0)}(t-2)))) \times (\text{Sin}((e^{(0)}(t-2)))) \\ &- ((e^{(0)}(t-1)) - (e^{(0)}(t-2)))) - (((e^{(0)}(t-2)) \\ &\times (\text{Sin}((e^{(0)}(t-2)))) \times (\text{Cos}((e^{(0)}(t-1)))) \\ &- ((e^{(0)}(t-1)) - (((e^{(0)}(t-1)) - (e^{(0)}(t-2)))) \\ &\times (\text{Sin}(-13.025)))))). \end{aligned} \tag{26}$$

Based on the use of GM(1,1) in the original data sets and the use of GP in residual signs as well as absolute residual series of GM(1,1), the proposed two-stage forecasting model which includes one and two lagged residual variables can be represented, respectively.

$$\begin{aligned} \hat{w}_{one-lagged}^{(0)}(t) &= \left(7105407 - \frac{(7742231.1299162430)}{(-0.0507747821)} \right) \\ &\times (1 - e^{(-0.0507747821)}) \times e^{0.0507747821(t-1)} \\ &+ \{s'(t) \times f(e^{(0)}(t-1))\}. \end{aligned} \tag{27}$$

where $s'(t)$ can be obtained from Eqs. (16) and (23), and $f(e^{(0)}(t-1))$ can be estimated using Eq. (25).

$$\begin{aligned} \hat{w}_{two-lagged}^{(0)}(t) &= \left(7105407 - \frac{(7742231.1299162430)}{(-0.0507747821)} \right) \\ &\times (1 - e^{(-0.0507747821)}) \times e^{0.0507747821(t-1)} \\ &+ \{s(t) \times f(e^{(0)}(t-1), e^{(0)}(t-2))\}. \end{aligned} \tag{28}$$

where $s(t)$ can be obtained from Eqs. (11) and (24), and $f(e^{(0)}(t-1), e^{(0)}(t-2))$ can be estimated using Eq. (26).

Table 4 and Fig. 7 summarize the results obtained using the four forecasting models with the data on the annual agricultural imports values in Taiwan from 2002 to 2011. Fig. 8 displays the percentage error (%) of the forecasting models. In Table 4, MAPE of the GM(1,1) model, the forecasting model of Lee and Tong (2011a), the ARIMA model, and the proposed two-stage model (adopting one and two lagged residual variables as input variables of GP model) applied to the training data (2002–2009) are 4.78%, 1.81%, 7.02%, 2.99%, and 1.12%, respectively. For the testing data, the MAPE are 11.49%, 14.52%, 22.31%, 11.97%, and 5.60% from 2010 to 2011, respectively. Above results indicate that the proposed forecasting model which adopts two lagged residual variables as input variables of the GP model, has a higher forecasting precision than that of the other models when applied to both training and testing data sets. The proposed model, which adopts two lagged residual variables as input variables, outperforms the proposed model, which adopts one lagged residual variable. This difference may explain why adopting more input variables can

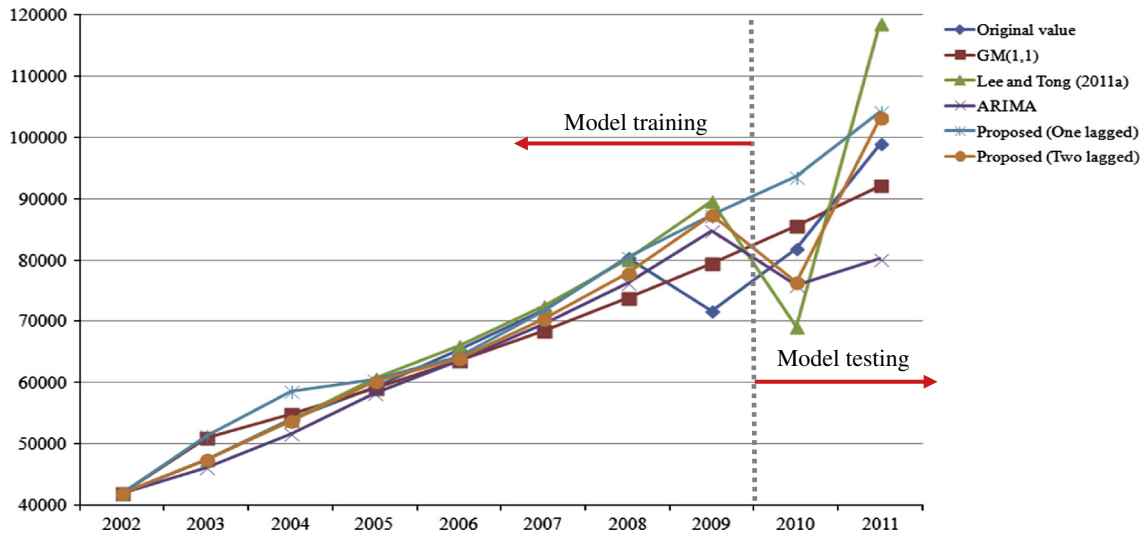


Fig. 14. Distributions of forecast values and real values from 2002 to 2011 for the agricultural imports data sets in USA.

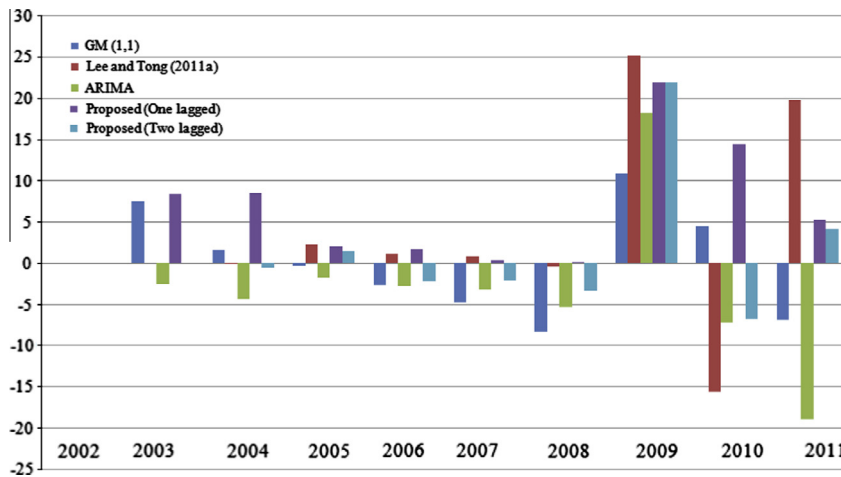


Fig. 15. Trend of percentage error (%) among the four models from 2002 to 2011 for the agricultural imports data sets in USA.

obtain more information and forecasting accuracy than only adopting one input variable when constructing a forecasting model. As for forecasting the testing data, the proposed model which adopts two lagged residual variables achieves good forecasting based on the MAPE criterion. The ARIMA model yields less satisfactory results than those of the other models, perhaps owing to the statistical assumptions. It may be the reasons which come from lacking of large-size samples and linear model. Fig. 9 presents the scatter plots for the GM(1,1) model, the proposed model of Lee and Tong (2011a), the ARIMA model, and the proposed model (adopting one and two lagged residual variables), respectively. This figure reveals that when adopting two lagged residual variables as input variables of GP model, the proposed forecasting model has a higher R^2 value ($R^2 = 0.9664$) than that of the other models.

3.4. Forecasting results of the agricultural imports data sets in USA

Results of the four forecasting models (i.e. GM(1,1) model, the improved GM(1,1) model by Lee and Tong (2011a), the ARIMA model, and the proposed model) for the agricultural imports data sets in the USA are expressed as follows:

(1) The GM(1,1) forecasting equation is:

$$\hat{w}^{(0)}(t) = 47327.0901 \times \exp(0.0740337257 \times (t - 1)), \quad t = 2, 3, \dots \quad (29)$$

(2) The improved GM(1,1) forecasting equation by Lee and Tong (2011a) is:

$$\hat{w}^{(0)}(t) = 47327.0901 \times \exp(0.0740337257 \times (t - 1)) + s(t) \times \{595.5328 \times \exp(0.4739341266 \times (t - 1))\}, \quad t = 1, 2, \dots, \quad (30)$$

where $s(t)$ is the binary variable (1 or -1). The forecasted residual sign are obtained using GP (independent variables: $d(t - 1)$ and $d(t - 2)$, dependent variable: $d(t)$). The parameter settings of GP for the population size, maximum number of generation, crossover rate, and mutation rate are 100, 1000, 0.9, and 0.1, respectively. If $\hat{d}(t)$ is 1, then $s(t)$ represents 1; otherwise, $s(t)$ represents -1 ($\hat{d}(t)$ is 0).

(3) The ARIMA forecasting equation is:

Based on the ACF and PACF criteria in the training data to identify the order of the ARIMA model, Fig. 10 summarizes the results of ACF and PACF. This figure reveals that the differencing order: 1

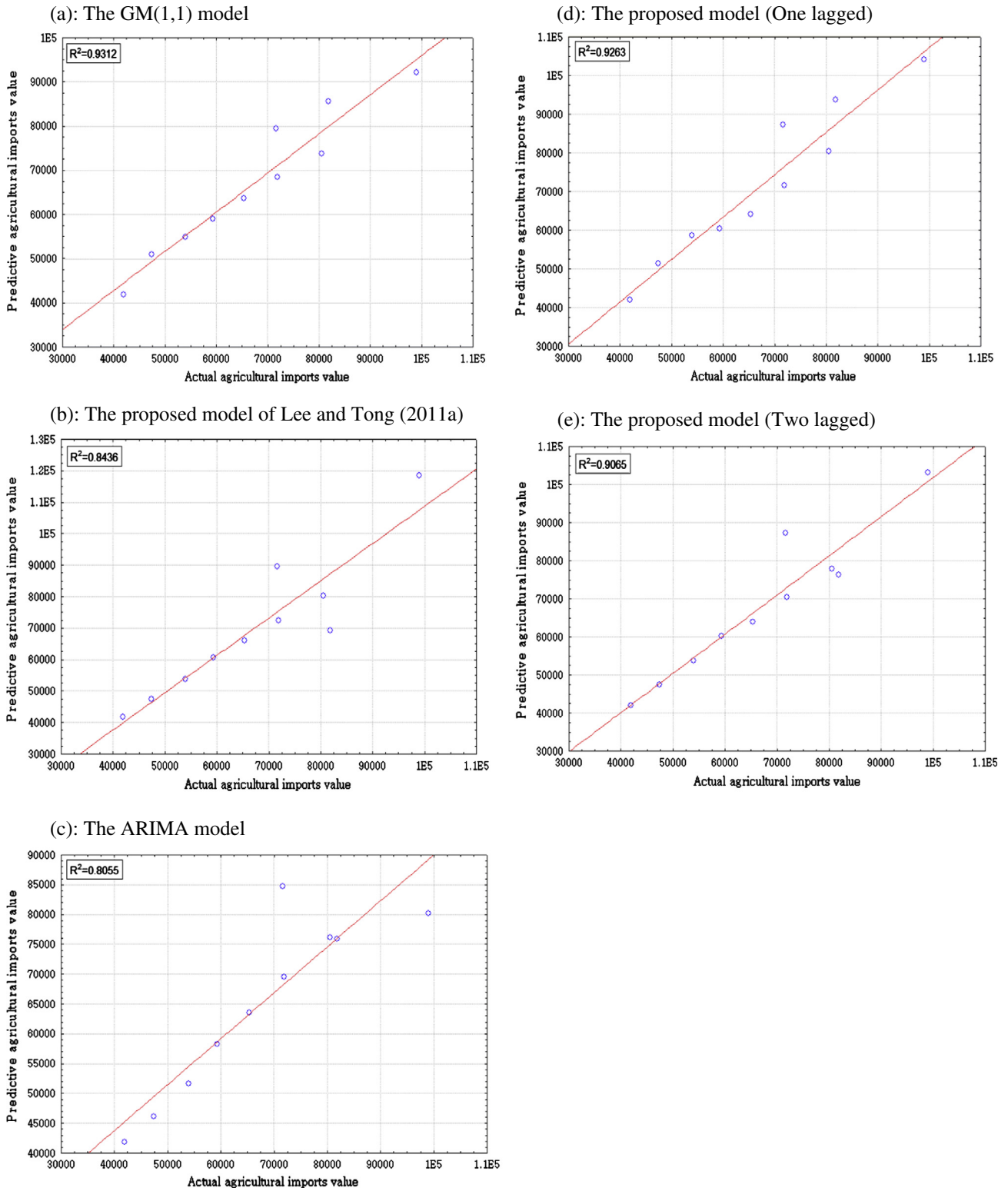


Fig. 16. Scatter plots among different forecasting models for the agricultural imports data sets in USA.

to be the stationary state. The forecasting model can be represented as follows:

$$\hat{w}^{(0)}(t) = 4252.2480 + w^{(0)}(t - 1), \quad t = 2, 3, \dots \quad (31)$$

Moreover, the AIC value of Eq. (31) is 142.32542, and Fig. 11 shows the Box–Ljung test. According to this figure, the residual series of ARIMA does not exhibit a significant lack of fit.

(4) The proposed two-stage model can be obtained as follows.

Based on the GM(1,1) model, then the residual signs can be measured by the GP model. The parameter settings based on GP for residual signs are the same as those in Table 1. Among the 10 simulation times, this study demonstrates the best simulation results (based on the fitness function) in Fig. 12. Hence, the forecasting equations of one and two lagged residual signs are represented, respectively.

$$\begin{aligned}\hat{d}'(t) &= f(d(t-1)) \\ &= (\exp((\log(\cos(d(t-1))))/(\cos(d(t-1)))) \\ &\quad \times (\cos((d(t-1) \times 0))))). \quad (32)\end{aligned}$$

$$\begin{aligned}\hat{d}(t) &= f(d(t-1), d(t-2)) \\ &= ((\cos(d(t-2))) + (((d(t-1)) - (\cos(d(t-2))) \\ &\quad \times (d(t-2)))). \quad (33)\end{aligned}$$

This study also demonstrates the best one among 10 times for the absolute one and two lagged residual series of GM(1,1). Table 2 lists the parameter settings of GP for the absolute residual series of GM(1,1). Fig. 13 demonstrates the fitness value of the best simulation for the absolute one and two lagged residual series of GM(1,1). Hence, the forecasting equation for absolute one and two lagged residual series of GM(1,1) can be represented, respectively.

$$\begin{aligned}\hat{e}'^{(0)}(t) &= f(e^{(0)}(t-1)) \\ &= ((-288.682) + (((e^{(0)}(t-1)) - (-94.667)) \\ &\quad - (-288.682))/(\cos(\cos((e^{(0)}(t-1)) \\ &\quad \times -11.851))))). \quad (34)\end{aligned}$$

$$\begin{aligned}\hat{e}^{(0)}(t) &= f(e^{(0)}(t-1), e^{(0)}(t-2)) \\ &= (((e^{(0)}(t-1)) - ((e^{(0)}(t-1)) + 18.954/(e^{(0)}(t-2)) \\ &\quad + 9.762))/(\exp((\log((e^{(0)}(t-1))))/(e^{(0)}(t-2)) \\ &\quad + 8.016)))/(\sin(\exp((\log((e^{(0)}(t-1))))/(e^{(0)}(t-2)) \\ &\quad + 18.954))))). \quad (35)\end{aligned}$$

Based on use of the GM(1,1) in the original data sets and use of the GP in residual signs as well as absolute residual series of GM(1,1), the proposed two-stage forecasting model (which include one and two lagged residual variables) can be represented, respectively, as follows:

$$\begin{aligned}\hat{w}_{one-lagged}^{(0)}(t) &= \left(41915.217 - \frac{(45997.4625505823)}{(-0.0740337257)}\right) \\ &\quad \times (1 - e^{(-0.0740337257)}) \times e^{0.0740337257(t-1)} \\ &\quad + \{s'(t) \times f(e^{(0)}(t-1))\}. \quad (36)\end{aligned}$$

where $s'(t)$ can be obtained from Eqs. (16) and (32), and $f(e^{(0)}(t-1))$ can be estimated using Eq. (34).

$$\begin{aligned}\hat{w}_{two-lagged}^{(0)}(t) &= \left(41915.217 - \frac{(45997.4625505823)}{(-0.0740337257)}\right) \\ &\quad \times (1 - e^{(-0.0740337257)}) \times e^{0.0740337257(t-1)} \\ &\quad + \{s(t) \times f(e^{(0)}(t-1), e^{(0)}(t-2))\}. \quad (37)\end{aligned}$$

where $s(t)$ can be obtained from Eqs. (11) and (33), and $f(e^{(0)}(t-1), e^{(0)}(t-2))$ can be estimated using Eq. (35).

Table 5 and Fig. 14 summarize the results obtained using the four forecasting models with the data on the annual agricultural imports values in USA from 2002 to 2011. Fig. 15 displays the percentage error (%) of the forecasting models. In Table 5, MAPE of the GM(1,1) model, the forecasting model of Lee and Tong (2011a), the ARIMA model, and the proposed two-stage model (adopting one and two lagged residual variables as input variables of GP model) applied to the training data (2002–2009) are 4.50%, 3.74%, 4.78%, 5.38%, and 3.94%, respectively. For the testing data, the MAPE are 5.71%, 17.67%, 13.10%, 9.87%, and 5.48% from 2010 to 2011, respectively. Above results indicate that the proposed forecasting model, which adopts two lagged residual variables as input variables of the GP model, has a higher forecasting precision than that of the other models overall. Additionally, the proposed model, which adopts two lagged residual variables as input variables, outper-

forms the proposed model, which adopts one lagged residual variable. This difference may explain why adopting more input variables can yield more information and a higher forecasting accuracy than when only adopting one input variable in forecasting. Fig. 16 presents the scatter plots for the GM(1,1) model, the proposed model of Lee and Tong (2011a), the ARIMA model, and the proposed model (adopting one and two lagged residual variables), respectively. This figure reveals that the proposed forecasting models, when adopting one and two lagged residual variables as input variables of GP model, have a high R^2 value as well as that of GM(1,1).

4. Conclusions

Developing a high-precision model for forecasting the value of agricultural imports is quite challenging since many factors affect the value, including the economy, changes in industry, and governmental policies. Decision makers thus heavily depend on the prediction accuracy of such models. This study adopts a novel two-stage forecasting model that combines GM(1,1) model applied to the original time series and the GP model to improve the residual component: residual signs and absolute residual series of GM(1,1). Computational results indicate that the proposed model outperforms the other previous forecasting models. In the future work, this study will combine other feature selection techniques such as GAs to select appropriate lagged residual variables and compare with the proposed model. Moreover, the study should apply the fuzzy time series method to forecast agricultural imports value and use more predicted variables such as annual gross domestic product value as input variables of GM in order to construct an accurate forecasting model.

Acknowledgements

The authors thank reviewers for providing useful comments in order to enhance the quality of this manuscript. Finally, the first author thanks Jesus to make this study become good results.

References

- Agrawal, R., 2003. *Forecasting Techniques in Agriculture*. Indian Agricultural Statistics Research Institute.
- Aladag, C.H., Egrioglu, E., Kadilar, C., 2009. Forecasting nonlinear time series with a hybrid methodology. *Appl. Math. Lett.* 22 (9), 1467–1470.
- Chang, T.S., Ku, C.Y., Fu, H.P., 2013. Grey theory analysis of online population and online game industry revenue in Taiwan. *Technol. Forecast. Soc. Change* 80 (1), 175–185.
- DeLurgio, S.A., 1998. *Forecasting Principles and Applications*. Irwin/McGraw-Hill, New York.
- Egrioglu, E., Aladag, C.H., Basaran, M.A., Yolcu, U., Uslu, V.R., Erilli, N.A., 2011a. Fuzzy time series forecasting method based on Gustafson–Kessel fuzzy clustering. *Exp. Syst. Appl.* 38 (8), 10355–10357.
- Egrioglu, E., Aladag, C.H., Basaran, M.A., Yolcu, U., Uslu, V.R., 2011b. A new approach based on the optimization of the length of intervals in fuzzy time series. *J. Intell. Fuzzy Syst.* 22, 15–19.
- Forouzanfar, M., Doustmohammadi, A., Hasanzadeh, S., Shakouri, G.H., 2012. Transport energy demand forecast using multi-level genetic programming. *Appl. Energy* 91 (1), 496–503.
- Harvey, A.C., 1981. *Time Series Models*. John Wiley & Sons Inc., New York.
- Hsu, L.C., 2003. Applying the Grey prediction model to the global integrated circuit industry. *Technol. Forecast. Soc. Change* 70 (6), 563–574.
- Hsu, L.C., 2009. Forecasting the output of integrated circuit industry using genetic algorithm based multivariable grey optimization models. *Exp. Syst. Appl.* 36 (4), 7898–7903.
- Hsu, C.C., Chen, C.Y., 2003. Applications of improved grey prediction model for power demand forecasting. *Energy Convers. Manage.* 44 (14), 2241–2249.
- Hsu, L.C., Wang, C.H., 2007. Forecasting the output of integrated circuit industry using a grey model improved by the Bayesian analysis. *Technol. Forecast. Soc. Change* 74 (6), 843–853.
- Huang, J.J., Tzeng, G.H., Ong, C.S., 2006. Two-stage genetic programming (2SGP) for the credit scoring model. *Appl. Math. Comput.* 174 (2), 1039–1053.
- Jutras, P., Prasher, S.O., Mehuys, G.R., 2009. Prediction of street tree morphological parameters using artificial neural networks. *Comput. Electron. Agric.* 67 (1–2), 9–17.

- Khashei, M., Bijari, M., 2012. Hybridization of the probabilistic neural networks with feed-forward neural networks for forecasting. *Eng. Appl. Artif. Intell.* 25 (6), 1277–1288.
- Koza, J., 1992. *Genetic Programming: On the Programming of Computers by Natural Selection*. MIT Press, Cambridge, MA.
- Lambert, D.M., Cho, S., 2008. Forecasting input demand shocks on China's gross value of agricultural output. *Rev. Reg. Stud.* 38, 343–359.
- Lee, Y.S., Tong, L.I., 2011a. Forecasting energy consumption using a grey model improved by incorporating genetic programming. *Energy Convers. Manage.* 52 (1), 147–152.
- Lee, Y.S., Tong, L.I., 2011b. Forecasting time series using a methodology based on autoregressive integrated moving average and genetic programming. *Knowl.-Based Syst.* 24 (1), 66–72.
- Lee, Y.S., Tong, L.I., 2012. Predicting high or low transfer efficiency of photovoltaic systems using a novel hybrid methodology combining rough set theory, data envelopment analysis and genetic programming. *Energies* 5, 545–560.
- Lee, D.G., Lee, B.W., Chang, S.H., 1997. Genetic programming model for long-term forecasting of electric power demand. *Electr. Pow. Syst. Res.* 40 (1), 17–22.
- Muttill, N., Lee, J.H.W., 2005. Genetic programming for analysis and real-time prediction of coastal algal blooms. *Ecol. Modell.* 189 (3–4), 363–376.
- Ou, S.L., 2012. Forecasting agricultural output with an improved grey forecasting model based on the genetic algorithm. *Comput. Electron. Agric.* 85, 33–39.
- Pai, P.F., Lin, C.S., 2005. A hybrid ARIMA and support vector machines model in stock price forecasting. *Omega* 33 (6), 497–505.
- Pao, H.T., 2009. Forecasting energy consumption in Taiwan using hybrid nonlinear models. *Energy* 34 (10), 1438–1446.
- Pao, H.T., Fu, H.C., Tseng, C.L., 2012. Forecasting of CO₂ emissions, energy consumption and economic growth in China using an improved grey model. *Energy* 40 (1), 400–409.
- Robinson, A., 2001. *Genetic Programming: Theory, Implementation and the Evolution of Unconstrained Solutions*. Hampshire College Division III.
- Sette, S., Boullart, L., 2001. Genetic programming: principles and applications. *Eng. Appl. Artif. Intell.* 14 (6), 727–736.
- Song, Q., Chissom, B.S., 1993. Fuzzy time series and its models. *Fuzzy Set Syst.* 54 (3), 269–277.
- Srinivasan, D., 2008. Energy demand prediction using GMDH networks. *Neurocomputing* 72 (1–3), 625–629.
- Wang, S.J., Wang, W.L., Huang, C.T., Chen, S.C., 2011. Improving inventory effectiveness in RFID-enabled global supply chain with Grey forecasting model. *J. Strategic Inf. Syst.* 20 (3), 307–322.
- Wang, J.J., Wang, J.Z., Zhang, Z.G., Guo, S.P., 2012. Stock index forecasting based on a hybrid model. *Omega* 40 (6), 758–766.
- Wu, L., Liu, S., Yao, L., Yan, S., 2013. The effect of sample size on the grey system model. *Appl. Math. Comput.* 37 (9), 6577–6583.
- Yang, L., Peng, L., Zhang, L., Yang, S., 2009. A prediction model for population occurrence of paddy stem borer (*Scirpophaga incertulas*), based on back propagation artificial neural network and principal component analysis. *Comput. Electron. Agric.* 68 (2), 200–206.
- Yin, M.S., Tang, H.W.V., 2013. On the fit and forecasting performance of grey prediction models for China's labor formation. *Math. Comput. Model.* 57 (3–4), 357–365.
- Yolcu, U., Egrioglu, E., Aladag, C.H., 2013. A new linear & nonlinear artificial neural network model for time series forecasting. *Decis. Support Syst.* 54 (3), 1340–1347.
- Zhang, G.P., 2003. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* 50, 159–175.
- Zhou, Z.J., Hu, C.H., 2008. An effective hybrid approach based on grey and ARMA for forecasting gyro drift. *Chaos, Solitons Fractals* 35 (3), 525–529.