

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

Grey Differential Equation Modeling on Stock Prices

by

Qifeng Xue



Submitted to the Department of Statistical Sciences
in partial fulfillment of the requirements for the degree of

Master of Science in Financial Mathematics

at the

UNIVERSITY OF CAPE TOWN

August 2005

©

Supervisor: Professor Renkuan Guo

Declaration

1. Plagiarism is to use another's work and pretend that it is one's own. I know that plagiarism is wrong and therefore my work does not contain any plagiarism.
2. Each significant contribution to, and quotation in this mini-dissertation from the work(s) of other people has been attributed, and has been cited and referenced.
3. I have not allowed, and will not allow any individual to copy my work with the intention of passing it off as his or her own work.
4. The author grants to University of Cape Town a permission to reproduce and to distribute this mini-dissertation in whole or in part for scientific research or teaching purposes.

Signature.....

Abstract

The economical fundamentals of stock market and the moods of market traders are both changing with time. Particularly the market mood may change extremely fast and dramatic within a very short time. Therefore, the traditional time-series analysis or econometric models which require large sample of data may not be suitable to reflect market changes in short period because of the sparse data availability. For example, a market trader may only have current week five daily close indices for a particular stock, he/she wants to know what is the following Monday's close price. It is obvious that standard large-sample based statistical methodology is powerless here. In order to predict market movement with 4 or 5 stock prices, in this thesis, the first order one variable grey differential equation model (abbreviated as GM(1, 1)) from grey theory is reviewed and applied to short stock price sequence for prediction. The efficiency and predictive power of GM(1, 1) are examined based on a few stock prices of Johannesburg Stock Exchange. Also, the second order one variable grey differential equation model (abbreviated as GM(2, 1)) is used for the stock price prediction for the first time and it demonstrates having a better performance than that of GM(1, 1) in some circumstances. Furthermore, we investigate the extended versions of GM(1, 1) model and their predictive power in stock price. The CGM(1, 1) (Cavity GM) is investigated and the result backs up the using of trading days rather than natural days in modeling on daily stock prices. The MGM(1, 1) (Modified GM(1, 1)) is implemented and shows to be with the capacity of general improvement. A type of transformation on the original data series is proposed and the EGM(1, 1) (Exponential Grey Model) is constructed for the transformed data. Finally an MEGM(1, 1), which is the combination of MGM and EGM, is developed and exhibited being superior to EGM(1, 1).

Acknowledgments

I would like to express my great appreciation and gratitude to Professor R. Guo for his enthusiastic guidance. He gives me lots of valuable advices on the theory and methodology. His insight leads my research to the right way. Moreover, his positive attitude toward work and life also inspires me a lot.

I would also like to thank Professor C. Troskie for all he has done for my coursework and dissertation. I learn methodologies from his teaching, which is definitely very helpful for my present and future work. He has also generously provided us a lot of materials and data.

My gratitude goes to Professor T. Dunne as well, for his firm helps and supports. Every time I have any problems, I know I can go to him.

Finally I would like to thank the support of my family. Through the whole MSc course, I owe them a lot.

University of Cape Town

Contents

Declaration	ii
Abstract	iii
Acknowledgments	iv
List of Tables	vii
List of Figures	viii
1 Introduction.....	1
2 General Characteristics of Stock Market.....	7
2.1 Intrinsic Value.....	7
2.2 EMH and Random Walk.....	8
2.3 Human Behavior.....	10
3 A Review on Grey Theory and Grey Differential Equation Models.....	12
3.1 Basic Concept of Grey System Theory.....	12
3.1.1 The Grey System.....	12
3.1.2 Grey Number and Grey Whitening.....	13
3.1.3 Grey Series.....	13
3.2 Data Processing in Grey System Theory.....	14
3.2.1 Holistic Generating Operation.....	14
3.2.2 Partial Generating Operation.....	18
3.2.3 Smooth Discrete Series.....	20
3.3 Grey Differential Equation and Grey Model.....	25
3.3.1 $GM(n, h)$	26
3.3.2 $GM(1, 1)$	29
3.3.3 $GM(2, 1)$	31
3.4 The Error of GM.....	37
3.4.1 The Average Relative Error.....	38
3.4.2 The Posterior Error.....	38
3.4.3 The Error of Prediction.....	40
3.5 The Extension of GM to Data Series Containing Negative Values.....	44
4 Empirical Modeling.....	47
4.1 $GM(1, 1)$ Modeling on Monthly Data.....	47
4.2 $GM(2, 1)$ Modeling on Monthly Data.....	52
4.3 $GM(1, 1)$ Modeling on Daily Data.....	54
4.4 $CGM(1, 1)$ Modeling on Daily Data.....	59
5 Extensions to $GM(1, 1)$ Model.....	63
5.1 $MGM(1, 1)$ Model.....	63
5.2 $EGM(1, 1)$ Model.....	65
5.3 Implementation of the Extended Models.....	68
5.3.1 $MGM(1, 1)$ Model.....	68
5.3.2 $EGM(1, 1)$ Model.....	71
5.3.3 $MEGM(1, 1)$ Model.....	75
6 Conclusion and Discussion.....	83

6.1	Errors of the Models on Different Time Span	83
6.2	Prediction Error Estimation	84
6.2.1	Prediction Error with Relation to the Model Fittings	84
6.2.2	Prediction Error Estimated by Variance Propagation	90
6.2.3	Improvement on Variance Propagation	94
6.3	Limitations of Grey Modeling.....	97
Appendix A	Data	99
1.	Monthly data	99
2.	Daily data	99
Appendix B	VBA Codes	100
1.	Flow Chart of the Program.....	100
2.	Codes for GM(1, 1).....	101
3.	Codes for GM(2, 1).....	105
References	110

University of Cape Town

List of Tables

Table 3-1 GM Model Precision Grades.....	40
Table 3-2 Classification for Prediction Error	43
Table 4-1 Statistics of Fitting by GM(1, 1)	49
Table 4-2 A GM(1, 1) on TONGAT	50
Table 4-3 Model Fitting Comparison of GM(1, 1) and GM(2, 1).....	53
Table 4-4 Error Comparison for Each Value of GM(1, 1) and GM(2, 1).....	53
Table 4-5 Statistics of Fitting by GM(1, 1) on Daily Prices.....	56
Table 4-6 Statistics of Fitting by CGM(1, 1) on Daily Prices	60
Table 4-7 Fitting Comparison of GM(1, 1) and MGM(1, 1).....	60
Table 5-1 Statistics of Fitting by MGM(1, 1) on Daily Prices	68
Table 5-2 Fitting Comparison of GM(1, 1) and MGM(1, 1).....	69
Table 5-3 Statistics of Fitting by EGM(1, 1) on Daily Prices.....	72
Table 5-4 Fitting Comparison of GM(1, 1) and EGM(1, 1).....	72
Table 5-5 One Model Comparison Between GM(1, 1) and EGM(1, 1).....	74
Table 5-6 Statistics of Fitting by MEGM(1, 1) on Daily Prices.....	75
Table 5-7 Fitting Comparison of EGM(1, 1) and MEGM(1, 1).....	76
Table 5-8 Fitting Comparison of GM(1, 1) and MEGM(1, 1).....	77
Table 5-9 Fitting Comparison of MGM(1, 1) and MEGM(1, 1).....	79
Table 6-1 Comparison of GM(1, 1) on Daily and Monthly Data	84
Table 6-2 Result from GM(1, 1) on 10 stocks (part I).....	85
Table 6-3 Result from GM(1, 1) on 10 stocks (part II)	86
Table 6-4 Average ARE at Each MP Grade.....	87
Table 6-5 Average RPE at Each MP Grade	87
Table 6-6 Prediction Error Estimation by Propagation (I).....	92
Table 6-7 Prediction Error Estimation by Propagation (II)	93
Table 6-8 Relation of Confidence Level and Prediction Range	95
Table 6-9 Prediction Range of GM(1, 1) at Different Confidence Levels.....	97

List of Figures

Figure 3-1 The Original Series	16
Figure 3-2 The1-AGO Series	16
Figure 3-3 The 2-AGO Series	16
Figure 4-1 Movement of the Monthly Stock Prices	48
Figure 4-2 Average ARE and RPE by Stocks	49
Figure 4-3 A GM(1, 1) Fitting on TONGAT	51
Figure 4-4 The Fitting Comparison of GM(2, 1) and GM(1, 1).....	54
Figure 4-5 Movement of the Daily Stock Prices	55
Figure 4-6 Average ARE and RPE by Stocks for Daily Data.....	57
Figure 4-7 Fittings and Predictions of GM(1, 1) on Daily Stock Prices	58
Figure 4-8 Average ARE Comparison by Stocks Between GM(1, 1) and CGM(1, 1).....	61
Figure 4-9 Average RPE Comparison by Stocks Between GM(1, 1) and CGM(1, 1).....	61
Figure 5-1 Average ARE Comparison by Stocks Between GM(1, 1) and MGM(1, 1).....	69
Figure 5-2 Maximum ARE Comparison by Stocks Between GM(1, 1) and MGM(1, 1).....	70
Figure 5-3 Average RPE Comparison by Stocks Between GM(1, 1) and MGM(1, 1).....	71
Figure 5-4 Maximum RPE Comparison by Stocks Between GM(1, 1) and MGM(1, 1).....	71
Figure 5-5 ARE Comparison by Stocks Between GM(1, 1) and EGM(1, 1).....	73
Figure 5-6 RPE Comparison by Stocks Between GM(1, 1) and EGM(1, 1).....	73
Figure 5-7 One Model Comparison Between GM(1, 1) and EGM(1, 1)	74
Figure 5-8 ARE Comparison by Stocks Between EGM(1, 1) and MEGM(1, 1).....	76
Figure 5-9 RPE Comparison by Stocks Between EGM(1, 1) and MEGM(1, 1).....	77
Figure 5-10 ARE Comparison by Stocks Between GM(1, 1) and MEGM(1, 1).....	78
Figure 5-11 RPE Comparison by Stocks Between GM(1, 1) and MEGM(1, 1)	78
Figure 5-12 ARE Comparison by Stocks Between MGM(1, 1) and MEGM(1, 1).....	80
Figure 5-13 RPE Comparison by Stocks Between MGM(1, 1) and MEGM(1, 1).....	80
Figure 5-14 Fittings and Predictions of MGM(1, 1) on Daily Stock Prices	81
Figure 5-15 Fittings and Predictions of MEGM(1, 1) on Daily Stock Prices	82
Figure 6-1 ARE Relates to MP Grade	87
Figure 6-2 RPE Relates to MP Grade.....	88
Figure 6-3 RPE Scatter over ARE.....	88
Figure 6-4 RPE Scatter over $ a $	89
Figure 6-5 RPE Scatter over $ u/a $	90
Figure 6-6 One Sigma Prediction Range.....	94
Figure 6-7 Confidence Level Changing with Prediction Range.....	95
Figure 6-8 Five and Eight Sigma Prediction Ranges	96

1 Introduction

Risk and return are two fundamental concepts in investment. Most analyses and predictions are involved in dealing with them. The risk is usually expressed as the uncertainty of the future. The prediction is to extract the substantial properties from what is known and reduce the future uncertainty.

Over the past decades, the stock market has been playing a major role among the investment utilities. Many theories and methodologies have been applied in the prediction of the stock prices. Among them, the popular quantitative analysis, which is based on conventional statistical methods, usually depends on a large sample of data information and adopts an assumption on the sample distribution. For this purpose, the time range for the whole data set is wide, so that some portion of data from the past may reflect totally different stage of the macrostructure behind the stocks. Hence the prediction based on data with a different macrostructure is questionable. Sometimes we may face sparse data availability, for example, for a newly listed stock or a newly merged company. It is also noticed that market mood changes very fast and dramatically so that traders may totally temporarily ignore the economic fundamentals behind the stock and only concentrate on the performance in a short period. For example, a market trader may only have current week five daily close indices for a particular stock, but he/she wants to know what is the following Monday's close price. This sparsity will make the analysis impossible if one engages only traditional large-sample based statistical methodology. In view of the above phenomena, we intend to apply the small-sample based grey theory to stock market data analysis.

The main aim of this mini-dissertation is explore the applicability and the predictability of grey differential equation models, particularly, GM(1, 1) model, to the short-term stock prices of Johannesburg Stock Exchange (abbreviated as JSE).

The scope of the study is limited to the investigation of GM(1, 1) model and its extensions. GM(2, 1) model is also introduced for the first time in stock market analysis. The mathematical foundation of grey theory is beyond the discussions of the mini-dissertation. Because of the

short-term nature (4-5 data points) of the grey data analysis, major features of the traditional time-series market data analysis will not be revealable. This research will not engage comparisons between grey prediction and time-series prediction since they are built on totally different data information assumptions.

We start the literature review on applications of grey methodology in general first and application to the Chinese stock market analysis, in order to state the motivation why we are interested to have a grey stock market research on JSE market data.

The *Grey System Theory* was created by Deng in 1982. It incorporates ideas and methodologies from system, cybernetics and information science and is able to deal with analysis, prediction and decision making. With the features of being easy to use and reasonable relative error in prediction, its applications have been rapidly extended to almost all the fields in science, business and society. Its basic idea is the grey model (GM), and among which, GM(1, 1) is the most commonly used. The modeling requires as little as only four data values. So the data used in constructing the GM can be highly time effective. The prices in the past week definitely have much more useful information than those from two years ago. The GM does not make any assumptions about the data distribution. Comparing to the conventional quantitative methods, GM(1, 1) is quite simple and easy to apply.

After first introducing the grey concept in 1982, Deng [1985] published a book to explain his grey theory. Fu [1992] and Liu [1999] each wrote a book with the same name as each others, aiming to let grey model be accepted and applied widely. Among the two books, Liu gave more detailed description and thus his book has more frequently been referenced. Although the grey theory has been accepted internationally, there is almost no book of English version. What we can find is only *Grey Systems, Modeling and Prediction* by Wen [2004], which merely covered parts of the extensive range.

There are many published papers on the variety practical applications of the grey theory. In environment research, Yeh and Chen [2004] used the grey relation analysis for analyzing the optimal artificial lake site in the Pingtung Plain in Taiwan and found it a feasible method to

judge the preferable site. Su and Zhao [2004] quantitatively determined a suitable scheme for wetland protection in Chagan Lake by means of the grey relation grade. The grey clustering method was presented in Li's [2002] conference proceedings paper which was on analysis of city atmosphere pollution. Chen [2003] exploited the grey system theory in his Master's dissertation for the prediction of the traffic noise. Tsai [2003] used grey theory as one of the three methods he chose to predict the quality of effluent water and regarded GM(1, 1) as giving the best estimation.

In the area of agriculture, Shi, Dong and Wang [2004] supported a new agricultural technique on grape planting by using the indicators derived from the grey system theory. Similar applications include Zhang's [2004] research on soybean's characters and Li's [2002] optimization on Chinese prickly ash tree planting.

In the area of industry and engineering, Hou [2004] introduced grey prediction for external random perturbations into industrial design. By incorporating the method with linear quadratic Gaussian method, he claimed the improvement of the performance. W-C. Yan *et al* [2005] tested different GM(1,1) models with dimension 4, 6, 8, and 16. Thus they concluded that the 16-dim model fit best for their feed-back noise reducing filter. While the grey modeling requires a minimum of four data in a series and usually keeps the number of data below 10, Yan's conclusion provides a good reference on the model dimension. Wu [2005] proposed a grey system based motion prediction method to effectively reduce the latency in building walkthrough, based on head-mounted display. His work underlies the generalization of grey data processing methods. L-J. Yan *et al* [2001] used a grey decision making method in finding out the optimal process parameters of EDM. They then verified the parameters by grey relation analysis. They claimed that their method was simpler, more efficient and accurate than other conventional methods. Bauer *et al* [2005] designed and implemented Grey, a set of software extensions that convert an off-the-shelf smartphone-class device into a tool by which its owner exercises and delegates her authority to both physical and virtual resources. Fan [2003] used the improved GM(1, 1) in the modeling of a dynamically tuned gyroscope and exhibited its superiority over a neural network method. Yue [2003] introduced the grey idea into the neural network model,

optimized the input structure, reduced calculation and hence raised the efficiency. Lin and Tsai [2005] created a hierarchical clustering analysis based on grey relation grade and distinguished it from other methods in terms of simplicity, effectiveness and flexibility.

In the area of construction, Z. Feng *et al* [2004] also applied GM(1, 1) in the settlement forecasting of ground. They made a revision after solving the grey differential equation to reduce the error. While most grey models were applied with equal time span between successive data values, Han *et al* [2000] built an unequal time span GM(1, 1) to forecast the ground settlement and highly praised its efficiency.

Other areas include: Lin [2004] found an effective multiple criteria decision making method on selecting the best company in information industry in Taiwan, by applying the grey relational analysis. Hsieh used the GM(1, 1) in the prediction of the cosmetic industry in Taiwan and had a good overall profile of the market. Mu and Kondou *et al* [2005] applied the GM(1, h) in forecasting the long term energy consumption in China and depicted rough changing ranges over the future 50 years. Guo *et al* [2005] proposed an automatic document classification technique based on grey relation analysis. They picked up key words for documents and trained the system to recognize them and proved it a viable way. For convenient use, Ren [2002] encoded GM(1, 1) in Delphi. Also, partly on the basis of grey system theory, L. Feng [2003] created Fan-Shu, a comprehensive theory.

In contrast to the vast pages on the above areas, there are just a few reports of the application of grey theory on stock market. Gao [1993] published the first book on grey stock analysis. He investigated all the stocks of Shanghai and Shenzhen Stock Exchange and claimed that GM(1, 1) is a very powerful approach to analyze stock price movements. Zhang *et al* [1995] implemented the GM(1, 1) on forecasting of Shenzhen Stock Indices, and got the prediction error below 5%. Li [1997] proposed an improved method on GM(1, 1) and applied it to a stock of the Shanghai Stock Exchange. Cong and Ji [2000] made a prediction on Shanghai Stock Index. They applied the GM(1, 1) model on daily, weekly and monthly data, and compared the forecast values from different data sets. Then they suggested that a revision should be made depending on the errors from different groups. Du *et al* [2002] designed a forecasting alarm system based on grey system

theory. They did not enter in the stock market, but their application on the electricity price which was also a time series made good reference to us. They claimed a model forecasting accuracy of 84% - 95%, and a price mutation forecasting accuracy of 96%. They also expressed a condition for judging whether the GM(1, 1) is applicable, i.e., the grey development coefficient should fall in interval $[-2,2]$. Otherwise, the response function would become chaotic. Chen and Li [2003] built a single model to predict Shanghai Stock Index and gave a prediction of very small relative errors. Nie and Li [2003] chose monthly data of Shanghai Stock Index over the period of January 2000 to January 2002 and predicted the mutation date. They found that usually only predicted values within three forward steps acceptable. So they tested putting the predicted values into the model and discarding the oldest one, constructing a new model with equal dimension and rolling forward. They achieved a good trace over the real trend. But they pointed out that there seemed a lag between the predicted and the real values. They also discovered that the prediction accuracy would be below 50%, if the model's grey development coefficient satisfies $a > 1.5$. Hong *et al* [2000] used GM(1, 1) in the volatility analysis of the Nikkei 225 index. By comparing the forecasting relative errors between the consecutive sliding windows, they discovered the constructional shift and hence revealed the profit opportunities.

From the literature review, it is obvious that, except for one paper, all the grey investigations were carried out by Chinese scholars and focused on the Chinese stock market and also most of them tested on index rather than stock prices. Therefore the grey modeling method on stock market is essentially not known to the western stock market researchers and traders that grey model may be an excellent methodology for short-term market behavior analysis.

Johannesburg Stock Exchange (abbreviated as JSE) is the major stock market in Africa because of the strongest economy in the African continent being that of South Africa. On the other hand, JSE responds sensitively to worldwide major stock market sensitively. Therefore it possesses some of the strongest characteristics of the developed countries, say, USA, Japan and UK. Exploration of the applicability of grey models to the JES should be a meaningful task in this dissertation.

The structure of this dissertation is as follows: Chapter two will be given to the discussions on

the general characteristics of the stock market. In Chapter three, we will review the grey theory and the grey differential equation models because most of the grey theory books and papers are written in Chinese and there are only two English books on it. Furthermore the readability of the English versions of research papers available is questionable. We will perform empirical modeling in Chapter four, using monthly data which are supposed to be rougher than daily data and hence we are expecting the result to be more convincing. Besides GM(1, 1), GM(2, 1) which has never been investigated by others on the stock market will also be tested. We then compare the GM(1, 1) on a different time span by implementing it on daily stock prices. Also by using CGM(1, 1) to compare with GM(1, 1), we exhibit the suitability of taking trading days rather than natural days when applying GM on daily stock prices. In Chapter five, we make some attempts on improvements beyond GM(1, 1). A new transformation will be introduced, while the methods that have already been in use will also be further exploited. We check the efficacy of these extensions on daily data based on the Week-Monday fitting-prediction structure. The last Chapter is for conclusions where we will focus on error discussion. The Appendices lists the data source and VBA codes.

2 General Characteristics of Stock Market

Stock market has expanded rapidly over the past century, especially in the emerging markets. It is an aggregation of social, economic, political, cultural and psychological phenomena, and full of contradiction, inconsistency, illogicality and controversy. Its nature has been in dispute and without conclusion since it came into the world. Our concern in this paper will be about its characteristics in the aspect of pricing.

It is very important to understand the mechanism of stock pricing as almost all the investment operations are involved in dealing with the asset prices. The profits from stock market usually come in two ways. One is price difference and the other is dividends. The latter makes sense only in the form of dividend yield which is the dividend per unit price. Everyday, lots of stories of joy or sadness are originated from the surge or dive of the stock prices all over the world. Ceaseless efforts have been put in finding out the regularities of stock pricing but until now, no unanimous theory has been reached.

2.1 Intrinsic Value

Graham and Dodd [1934] pointed out that the stock prices should be based on their intrinsic values which are the present values of their future cash flow. As the stock prices fluctuate up and down, they seldom just stay at the intrinsic values. Therefore the profiting opportunities exist. They recommend investors to buy the stocks with the prices below their intrinsic values and expect to take profits. They emphasized the analysis on the operation of the companies and estimated their future ability of growth. They aimed at buying a stock at the price equivalent or even below its net asset price and often purchased stocks which were neglected by other investors.

Sometimes earning's multiples, such as the P/E ratio, are used to determine value, where cash flows are relatively stable and predictable. The obvious caveat is that the P/E ratio is ultimately not an objective measure, because it must be interpreted. A high P/E ratio might be an overvalued stock, or it might be a company with high potential for growth. Other techniques

include book value and dividend yield analysis.

In practice, however, the operation under the guidance of the intrinsic value theory has not beaten the market. Some statistics showed that between 1985 and 1994 in Wall Street, the conventional mutual fund that believes in value investment only had about 26% growth, underperformed comparing the other fund and the S&P 500. The problem is how to have a correct valuation of the company's future. Even Graham and Dodd themselves admitted the difficulties inherent in measuring the intrinsic values. They also acknowledged that the buying behavior would push up the stock prices and as a result, cause them to deviate from the intrinsic values. And under their theory, rational investors will prevail and make a persistent price fluctuation around their true values.

2.2 EMH and Random Walk

The Efficient Market Hypothesis (EMH) was created by Fama in the 1960's. It stated that at any given time, the prices of all stocks fully reflect all available information about those stocks. There are three forms of market efficiency:

- i. Weak form. In this case, current prices fully reflect the information contained in the record of past prices. Investors cannot outperform the market through studying the history of stocks.
- ii. Semi-strong form. In this case, current prices fully reflect all public information. So it is of no use to analyze either the public information or the price chart.
- iii. Strong form. In this case, not only public information but also inside news as well as rumors have been put into the prices.

Market efficiency usually holds for the following conditions:

- The market is large and has a good liquidity.
- Information has to be widely available, in terms of accessibility and cost, and released to sufficient number of investors at more or less the same time.
- Transaction costs have to be reasonable, comparing to the expected profits of an

investment strategy.

- Investors should also have enough funds to take advantage of the once in a while inefficiency until, according to the EMH, it disappears again.
- There is not an agreement in the market about the implication of the current information and the expectation regarding the future price movements.

While that does not sound so radical, most people who buy and sell stocks do so with the assumption that the stocks they are buying are undervalued and therefore worth more than the purchase price. When they buy a stock, they hope that other investors have overlooked that stock for some reason, in effect giving them the opportunity to buy at a lower price. But under the Efficient Market Hypothesis, they are engaging in a game of chance, not skill. If markets are efficient and current prices always reflect all information, there's no way they will ever be able to buy a stock at a bargain price.

EMH supposes the stock prices will not follow any patterns or trends and move randomly. It does not require prices to be equal to fair values all of the time. Prices may be over or undervalued, but only in random occurrences, so they eventually resort back to their mean value. Past price movements cannot be used to predict the future's. Fama called this the Random Walk Theory. Because the deviations from a stock's fair price are in themselves random, there's no way to ever profit from "inefficiencies" in the price of a stock. The investment strategies that result in beating the market cannot be consistent phenomena. Empirical research has revealed that if the new information comes to the stock market, it will soon cause the changes in demand or supply, and is immediately reflected in the shift of the stock prices.

There is a controversy between the Efficient Market Hypothesis and Random Walk Theory. Suppose the stock prices are pure random walk and not predictable, people's effort in choosing stocks is vain. Even the investment in the stock market is ridiculous by itself. Then how do we explain the existing and development of all the stock exchanges?

2.3 Human Behavior

Contrast to the above two theories, behavioral finance theory argues that stock prices are affected heavily by the operation features of the investors and there are much common and large price deviations in the market. Supporters of behavioral finance think the stock market is irrational, rather than Fama's rational. Shefrin [2002] described the three themes of behavioral finance:

1. Heuristic-driven bias: People hold biased beliefs that predispose them to commit errors, while traditional finance assumes that when processing data, people use statistical tools appropriately and correctly.
2. Frame dependence: Behavioral finance postulates that in addition to objective considerations, practitioners' perception of risk and return are highly influenced by how decision problems are framed. In contrast, traditional finance assumes that people view all decisions through the transparent, objective lens of risk and return.
3. Inefficient markets: the heuristic-driven bias and framing effects cause market price to deviate from fundamental values. This is a contrast to traditional finance's rational expectations and market efficiency.

There is a speculative market view which underpins Shefrin's opinion. Evans [2003] concluded that the "crux of the speculative market view is that assets are purchased based on the belief of future price appreciation, implying that price movements are based primarily on the balance of public opinion rather than objective fundamentals."

Behavioral finance emphasizes the interaction of investors and the market and considers the psychological factors as a non-neglectable factor affecting the stock prices. When most people are optimistic, the stock market will go up even the economy is not so good. Like in the commercial market, the stock price is determined by the purchasing power and stop at a balance where supply meets demand. But the market will not be pushed by people's confidence all the time. Over optimism after a remarkable rally will lead the market turn to the opposite way and result in a sharp decline.

Some phenomena in the emerging stock market can be well explained by the behavioral finance. In the early stage of Chinese and Vietnamese stock market development, people used to experience a sharp rally or sudden dive which was happened in a short term that it could hardly be interpreted by the economic growth and the company's prosperity.

Besides the intrinsic value, random walk and human behavior in the stock pricing, there are definitely countless other proposals, opinions, theories and methodologies. Stocks have already developed well beyond one of the investment utilities. It has far-reaching influences on the economy, policy, culture and ideology. Efforts on research or predicting of the nature of stock pricing will never stop.

Today, we are all evolved in the phase of globalization. South Africa features a small and open economy which is reflected in its stock market characteristics. Any of the emerging markets like South Africa, with stable policy and open economy, will supply more opportunities and draw great attention of international far-sighted investors. South Africa is also opening itself wider to the overseas investors. So understanding the JSE behavior is of significant meaningful for overseas capital investors in exploring the investment opportunities in South Africa.

3 A Review on Grey Theory and Grey Differential Equation Models

In this chapter, we will introduce the major concepts of the grey system theory. As we can only apply a fraction of the grey methods, we will focus in those that could be related to our investigation on stock prices.

The grey system theory is to study on grey analysis, grey modeling, grey prediction, grey decision, and grey control. It focuses on:

- 1) Grey Model (GM)
- 2) Grey Prediction
- 3) Grey Relation Analysis (GRA)
- 4) Grey Statistics and Analogy
- 5) Grey Decision Making
- 6) Grey Control

We will apply the grey model to make grey prediction in this paper and will not attempt the other four appliances.

3.1 Basic Concept of Grey System Theory

3.1.1 The Grey System

The grey system is a relative concept. To understand what it is, we have to know about the concept of white system and black system.

A *White System* (W System) is the system in which, according to the extent of knowledge, all the information is known.

A *Black System* (B System) is the system in which, according to the extent of knowledge, all the

information is unknown.

A *Grey System* (G System) is the system in which, according to the extent of knowledge, only part of the information is known.

3.1.2 Grey Number and Grey Whitening

A *grey number* is the one with incomplete information, marked as \otimes . It is not a number in fact. It is a set of numbers and thus a range into which a number could fall. The concept of grey number is the characteristics of the grey system theory.

A *whitened grey number* is the possible value at which the grey number could be fixed. For a general grey number \otimes , its whitened number is expressed as $\tilde{\otimes}$. The symbol $\otimes(\alpha_i)$ stands for a grey number whose possible whitened value is α_i . It must be noted that $\tilde{\otimes}(\alpha_i)$, the whitened value of $\otimes(\alpha_i)$, is not necessarily to be α_i . It could be other values.

3.1.3 Grey Series

There are three major types of grey series.

Let X_i be the i^{th} series, and

$$\begin{aligned} X_i &= \{X_i(k) | k = 1, 2, \dots, N\} \\ &= \{X_i(1), X_i(2), \dots, X_i(N)\} \end{aligned}$$

- 1 If k is a time serial number, then X_i is the i^{th} time series;
- 2 If k is a space serial number, then X_i is the i^{th} space series;
- 3 If k is an index serial number, then X_i is the i^{th} index series.

For financial analysis, we usually deal with the first type, i.e., the time series.

For any of the above cases, if there is a cavity in the series, it will be called *cavity series*.

For example, in the series

$$X = \{X(1), X(2), \dots, X(k-1), \Phi(k), X(k+1), \dots, X(N)\}$$

$\Phi(k)$ is the cavity at the k^{th} point.

3.2 Data Processing in Grey System Theory

The data processing in grey system theory aims to add more information into the system. The frequently used method is generating operation, which is divided into holistic and partial operations.

3.2.1 Holistic Generating Operation

Holistic generating operation includes accumulated generating operation, inverse accumulated generating operation, initialization, maximization, minimization, averaging, localization, etc.

3.2.1.1 Accumulated Generating Operation (AGO)

The *Accumulated Generating Operation* (AGO) is the most important in the grey system theory. It means to put new information for analysis and weaken the randomness of the original series.

Suppose the original series is

$$\begin{aligned} X^{(0)} &= \{X^{(0)}(k) \mid k=1, 2, \dots, N\} \\ &= \{X^{(0)}(1), X^{(0)}(2), \dots, X^{(0)}(N)\} \end{aligned}$$

Make the first order AGO (1-AGO),

$$X^{(1)}(k) = \sum_{i=1}^k X^{(0)}(i)$$

i.e.

$$\begin{aligned}
X^{(1)} &= \{X^{(1)}(k) | k=1,2,\dots,N\} \\
&= \{X^{(1)}(1), X^{(1)}(2), \dots, X^{(1)}(N)\} \\
&= \left\{ X^{(0)}(1), \sum_{i=1}^2 X^{(0)}(i), \dots, \sum_{i=1}^N X^{(0)}(i) \right\}
\end{aligned}$$

Similarly, the second order AGO (2-AGO) is

$$\begin{aligned}
X^{(2)} &= \{X^{(2)}(k) | k=1,2,\dots,N\} \\
&= \left\{ X^{(1)}(1), \sum_{i=1}^2 X^{(1)}(i), \dots, \sum_{i=1}^N X^{(1)}(i) \right\} \\
&= \left\{ X^{(0)}(1), \sum_{i=1}^2 \sum_{j=1}^i X^{(0)}(j), \dots, \sum_{i=1}^N \sum_{j=1}^i X^{(0)}(j) \right\}
\end{aligned}$$

Let us get some feelings on the effect of AGO. We just generate a series with random numbers between 0 and 1 and treat it as the original series

$$X^{(0)} = \{0.204, 0.716, 0.870, 0.589, 0.383\}$$

By making 1-AGO and 2-AGO, we have

$$\begin{aligned}
X^{(1)} &= \{0.204, 0.919, 1.790, 2.379, 2.762\} \\
X^{(2)} &= \{0.204, 1.123, 2.913, 5.291, 8.053\}
\end{aligned}$$

Looking at Figure 3-1, Figure 3-2 and Figure 3-3, we know the regularity has been changed. The 1-AGO and 2-AGO have better linear or exponential characteristics. We just check the linearity of $X(k)$ with relation to k by using R -square, and find that $R^2(X^{(0)})=0.0194$, $R^2(X^{(1)})=0.982$ and $R^2(X^{(2)})=0.967$. So the linearity is considerably improved. But we also noted that the linearity of 2-AGO series is worse than 1-AGO. Liu [1999] has an explanation for this. He suggests that the AGO should stop at a right place where a regularity has been reached. Otherwise, over high order AGO will lead to the opposite and reduce the information we have got.

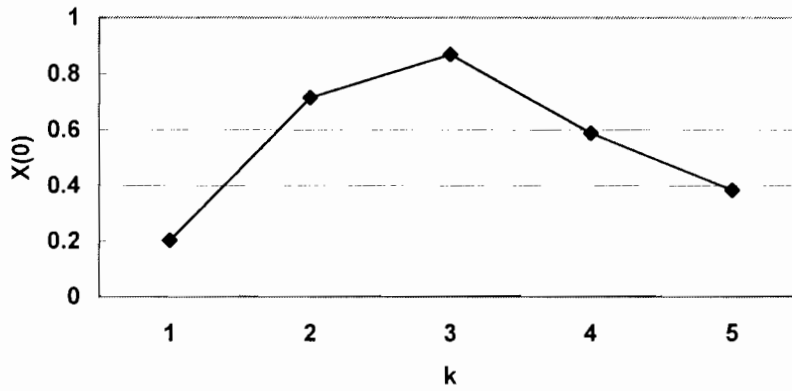


Figure 3-1 The Original Series

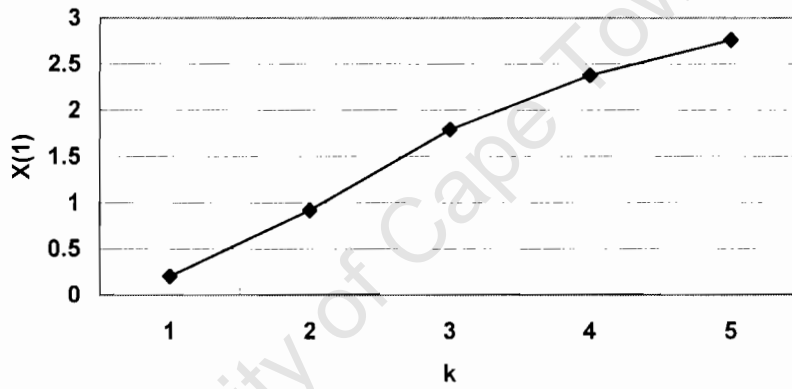


Figure 3-2 The 1-AGO Series

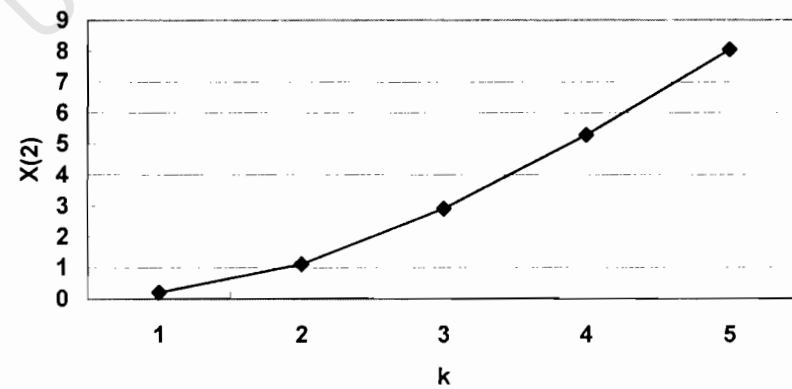


Figure 3-3 The 2-AGO Series

If we make the m^{th} order AGO (m -AGO) on the original series, then

$$X^{(m)}(k) = \sum_{i=1}^k X^{(m-1)}(i)$$

3.2.1.2 Inverse Accumulated Generating Operation (IAGO)

The *Inverse Accumulated Generating Operation* (IAGO) is the counter-operation of AGO. Here we use $I^{(i)}(X)$ to represent the value after i^{th} -IAGO, and

$$I^{(0)}(X(k)) = X^{(m)}(k)$$

So the 1-IAGO is

$$\begin{aligned} I^{(1)}(X(k)) &= I^{(0)}(X(k)) - I^{(0)}(X(k-1)) \\ &= X^{(m)}(k) - X^{(m)}(k-1) \\ &= X^{(m-1)}(k) \end{aligned}$$

and the 2-IAGO is

$$\begin{aligned} I^{(2)}(X(k)) &= I^{(1)}(X(k)) - I^{(1)}(X(k-1)) \\ &= X^{(m-1)}(k) - X^{(m-1)}(k-1) \\ &= X^{(m-2)}(k) \end{aligned}$$

and the m -IAGO is,

$$I^{(m)}(X(k)) = I^{(m-1)}(X(k)) - I^{(m-1)}(X(k-1))$$

So generally,

$$X^{(0)}(k) \xrightleftharpoons[m\text{-IAGO}]{m\text{-AGO}} X^{(m)}(k)$$

3.2.2 Partial Generating Operation

The partial generating operation is applied in the following three cases:

- 1) Non-equal step time series or space series;
- 2) Equal step time series or space series, which has abnormal values;
- 3) Cavity series.

The partial generating operation includes the methods of averaging generation, class ratio generation, interpolating generation and grey relational generation. We introduce the first two.

3.2.2.1 Averaging Generation

For a series

$$X = \{X(k) | k = 1, 2, \dots, N\}$$

If

$$X^*(k) = \alpha X(k-1) + (1-\alpha)X(k)$$

where $\alpha \in [0, 1]$, then we call $X^*(k)$ generated by averaging generation. When the weight $\alpha > 0.5$, the generation is old-info preferring. When $\alpha < 0.5$, the generation is new-info preferring. When $\alpha = 0.5$, the generation is an equal weight operation.

3.2.2.2 Class Ratio Generation

For a series

$$X = \{X(k) | k = 1, 2, \dots, N\}$$

its class ratio is defined as

$$\sigma(k) = \frac{X(k-1)}{X(k)} \quad k \geq 2$$

And further, $\sigma^{(0)}(k)$ is defined as the class ratio of the original series, $X^{(0)}$, $\sigma^{(1)}(k)$, the class ratio of 1-AGO series, $X^{(1)}$, and $\sigma^{(m)}(k)$, the class ratio of m -AGO series, $X^{(m)}$.

The four lemmas below show the properties of the class ratio.

- 1 **Lemma 1** For the original series $X^{(0)} = \{X^{(0)} > 0\}$, $\sigma^{(1)}(k) \in (0, 1]$ $k \geq 2$.
- 2 **Lemma 2** For the original series $X^{(0)} = \{X^{(0)} > 0\}$, $\sigma^{(0)}(k+1) \geq \sigma^{(1)}(k)$ $k \geq 2$.
- 3 **Lemma 3** For the original series $X^{(0)} = \{X^{(0)} > 0\}$, $\sigma^{(1)}(k+1) \geq \sigma^{(1)}(k)$ $k \geq 2$.
- 4 **Lemma 4** For the original series $X^{(0)} = \{X^{(0)} > 0\}$, $\sigma^{(1)}(k) > \sigma^{(2)}(k)$ $k \geq 2$.

Now we use a cavity series to explain the class ratio generation. For

$$X = \{X(1), X(2), \dots, X(k-1), \Phi(k), X(k+1), \dots, X(N)\}$$

we use $X^*(k)$ to substitute the cavity $\Phi(k)$ and make a new series

$$X^* = \{X(1), X(2), \dots, X(k-1), X^*(k), X(k+1), \dots, X(N)\}$$

The $X^*(k)$ is generated by the class ratio method, where the class ratios at point k for series X^* , $X^{*(1)}$ and $X^{*(2)}$ must meet the requirements by the above four lemmas.

The four lemmas build up four inequalities on the variable $X^*(k)$. The solution is usually in the form of $a \leq X^*(k) \leq b$, i.e., $X^*(k)$ is a set of values.

3.2.3 Smooth Discrete Series

The grey system theory requires the series to be a smooth discrete series, and then the theory can analyze it by modeling.

Definition *Smooth Discrete Series*

Suppose $X^{(0)}$ is a non-negative discrete series,

$$X^{(0)} = \{X^{(0)}(k) | k = 1, 2, \dots, N\}$$

let ε_k be the smoothness ratio, and

$$\begin{aligned}\varepsilon_k &= \frac{X^{(0)}(k)}{\sum_{i=1}^{k-1} X^{(0)}(i)} \\ &= \frac{X^{(0)}(k)}{X^{(0)}(k-1)}\end{aligned}$$

If $0 \leq \varepsilon_k < 1$, for $k \geq 3$, and $\varepsilon_k = \{\varepsilon_3, \varepsilon_4, \dots, \varepsilon_n\}$ is a decreasing series, then $X^{(0)}$ is called a *smooth discrete series*.

By the definitions of class ratio and smoothness ratio, we have

$$\sigma(k) = \frac{\varepsilon_{k-1}}{\varepsilon_k(1 + \varepsilon_{k-1})}$$

In the example of the class ratio generation,

$$\sigma^*(k) = \frac{\varepsilon_{k-1}^*}{\varepsilon_k^*(1 + \varepsilon_{k-1}^*)}$$

Since

$$\varepsilon_{k-1}^* = \frac{X^{(0)}(k-1)}{X^{(0)}(k-2)} = \varepsilon_{k-1}$$

hence

$$\sigma^*(k) = \frac{\varepsilon_{k-1}}{\varepsilon_k^*(1 + \varepsilon_{k-1})}$$

As

$$\sigma^*(k) = \frac{X(k-1)}{X^*(k)}$$

so

$$\frac{X(k-1)}{X^*(k)} = \frac{\varepsilon_{k-1}}{\varepsilon_k^*(1 + \varepsilon_{k-1})}$$

and

$$\frac{X^*(k)}{\varepsilon_k^*} = X(k-1) \frac{(1 + \varepsilon_{k-1})}{\varepsilon_{k-1}}$$

One prerequisite for the GM is the series $X^{(0)}$ must be smooth, then we can apply the differential equation on the 1-AGO series $X^{(1)}$. If the $X^{(0)}$ is not smooth, we will check whether or not $X^{(1)}$ is smooth, and if so, will model on the 2-AGO series $X^{(2)}$. But we cannot allow the AGO to go further, because the 2-AGO series will never be a smooth series under the condition of $0 \leq \varepsilon_k < 1$ for $k \geq 3$, and $\varepsilon_k = \{\varepsilon_3, \varepsilon_4, \dots, \varepsilon_N\}$ decreasing.

Lemma For any non-negative series $X^{(0)}$, its 2-AGO and any higher order AGO will never generate a smooth series.

Proof The result can be proved by the method of induction.

Consider

$$X^{(0)} = \{X^{(0)}(k) | k = 1, 2, \dots, N\}$$

Using $\varepsilon_k^{(i)}$ to represent the smoothness factors of $X^{(i)}$, for $k = 3, 4, \dots, N$, we have

$$\varepsilon_k^{(i)} = \frac{X^{(i)}(k)}{X^{(i+1)}(k-1)} \quad (3.2.1)$$

STEP 1: $i = 2$.

Let $k = 3$, then

$$\varepsilon_3^{(2)} = \frac{X^{(2)}(3)}{X^{(3)}(2)} \quad (3.2.2)$$

and

$$\begin{aligned} X^{(2)}(3) &= X^{(1)}(1) + X^{(1)}(2) + X^{(1)}(3) \\ &= X^{(0)}(1) + [X^{(0)}(1) + X^{(0)}(2)] + [X^{(0)}(1) + X^{(0)}(2) + X^{(0)}(3)] \\ &= 3X^{(0)}(1) + 2X^{(0)}(2) + X^{(0)}(3) \end{aligned}$$

$$\begin{aligned} X^{(3)}(2) &= X^{(2)}(1) + X^{(2)}(2) \\ &= X^{(1)}(1) + X^{(1)}(1) + X^{(1)}(2) \\ &= X^{(0)}(1) + X^{(0)}(1) + [X^{(0)}(1) + X^{(0)}(2)] \\ &= 3X^{(0)}(1) + X^{(0)}(2) \end{aligned}$$

Substitute $X^{(2)}(3)$ and $X^{(3)}(2)$ into (3.2.2),

$$\begin{aligned} \varepsilon_3^{(2)} &= \frac{3X^{(0)}(1) + 2X^{(0)}(2) + X^{(0)}(3)}{3X^{(0)}(1) + X^{(0)}(2)} \\ &= 1 + \frac{X^{(0)}(2) + X^{(0)}(3)}{3X^{(0)}(1) + X^{(0)}(2)} \end{aligned}$$

Since $X^{(0)}$ is a non-negative series, as long as $X^{(0)}(1) = 0$ and $X^{(0)}(2) = 0$ do not happen at the same time,

$$\frac{X^{(0)}(2) + X^{(0)}(3)}{3X^{(0)}(1) + X^{(0)}(2)} \geq 0$$

So

$$\varepsilon_3^{(2)} \geq 1$$

which means $X^{(2)}$ is not smooth. We do not need to test if $\varepsilon_k^{(2)}$ are decreasing.

STEP 2: $i = 3$.

Similarly, let $k = 3$, then

$$\varepsilon_3^{(3)} = \frac{X^{(3)}(3)}{X^{(4)}(2)} \quad (3.2.3)$$

and

$$\begin{aligned} X^{(3)}(3) &= X^{(2)}(1) + X^{(2)}(2) + X^{(2)}(3) \\ &= X^{(1)}(1) + [X^{(1)}(1) + X^{(1)}(2)] + [X^{(1)}(1) + X^{(1)}(2) + X^{(1)}(3)] \\ &= 3X^{(1)}(1) + 2X^{(1)}(2) + X^{(1)}(3) \\ &= 3X^{(0)}(1) + 2[X^{(0)}(1) + X^{(0)}(2)] + [X^{(0)}(1) + X^{(0)}(2) + X^{(0)}(3)] \\ &= 6X^{(0)}(1) + 3X^{(0)}(2) + X^{(0)}(3) \end{aligned}$$

$$\begin{aligned} X^{(4)}(2) &= X^{(3)}(1) + X^{(3)}(2) \\ &= 4X^{(0)}(1) + X^{(0)}(2) \end{aligned}$$

Substitute $X^{(3)}(3)$ and $X^{(4)}(2)$ into (3.2.3),

$$\begin{aligned} \varepsilon_3^{(3)} &= \frac{6X^{(0)}(1) + 3X^{(0)}(2) + X^{(0)}(3)}{4X^{(0)}(1) + X^{(0)}(2)} \\ &= 1 + \frac{2X^{(0)}(1) + 2X^{(0)}(2) + X^{(0)}(3)}{4X^{(0)}(1) + X^{(0)}(2)} \\ &\geq 1 \end{aligned}$$

Thus $X^{(3)}$ is not smooth either.

STEP 3: Suppose for $i = N - 1$,

$$\varepsilon_3^{(N-1)} = \frac{X^{(N-1)}(3)}{X^{(N)}(2)} \geq 1$$

then

$$X^{(N-1)}(3) \geq X^{(N)}(2) \quad (3.2.4)$$

For $i = N$

$$\varepsilon_3^{(N)} = \frac{X^{(N)}(3)}{X^{(N+1)}(2)} \quad (3.2.5)$$

and

$$X^{(N)}(3) = X^{(N-1)}(1) + X^{(N-1)}(2) + X^{(N-1)}(3)$$

$$\begin{aligned} X^{(N+1)}(2) &= X^{(N)}(1) + X^{(N)}(2) \\ &= X^{(N-1)}(1) + X^{(N)}(2) \end{aligned}$$

Substitute $X^{(N)}(3)$ and $X^{(N+1)}(2)$ into (3.2.5), then

$$\varepsilon_3^{(N)} = \frac{X^{(N-1)}(1) + X^{(N-1)}(2) + X^{(N-1)}(3)}{X^{(N-1)}(1) + X^{(N)}(2)} \quad (3.2.6)$$

Now substitute (3.2.4) into the right hand of (3.2.6),

$$\begin{aligned} \frac{X^{(N-1)}(1) + X^{(N-1)}(2) + X^{(N-1)}(3)}{X^{(N-1)}(1) + X^{(N-1)}(2)} &\geq \frac{X^{(N-1)}(1) + X^{(N-1)}(2) + X^{(N-1)}(3)}{X^{(N-1)}(1) + X^{(N-1)}(3)} \\ &= 1 + \frac{X^{(N-1)}(2)}{X^{(N-1)}(1) + X^{(N-1)}(3)} \\ &\geq 1 \end{aligned}$$

i.e.

$$\varepsilon_3^{(N)} \geq 1$$

which means $X^{(N)}$ is not smooth and the lemma is true.

Therefore, if we cannot get a smooth series for $X^{(0)}$ and $X^{(1)}$, we should stop trying any higher order AGO for the purpose of modeling. We have to turn to other generating methods, such as partial grey operations.

3.3 Grey Differential Equation and Grey Model

The grey system theory treats a random process as a grey process which keeps changing in a specific time and space profile. The random discrete series is considered implying a potential internal relation and hence the randomness can be effectively weakened through transformation. The grey system theory features in modeling on the generated data rather than the original ones.

Another characteristic of grey modeling is that it may use a small set of data, with the minimum of four. The conventional statistical methods often require large amount of data which is not always available. Even if we acquire the whole data, the information is in fact not with the same contribution. Old information should have less effect on the moving trend than the new one. Grey modeling just overcomes the shortcoming of sample size requirement. With a limited number of time effective data, it gives quite reasonable analysis and accurate prediction.

The Grey Differential Equation is the core idea of the grey system theory, based on which the Grey Model (GM) is constructed. A GM is usually built on a series by 1-AGO or 2-AGO, which requires the original series to be nonnegative. For different series, and depending on our various

objectives, it is possible to build different GM models. A GM is usually expressed in the form of GM(n, h), where n is the order of the differentiation and h , the number of series (not the number of data in a series) being used in constructing the model.

3.3.1 GM(n, h)

Using N to represent the number of data in a series on which GM is built and calling it N -dimension (short as N -dim), let us consider the time series

$$X_i^{(0)} = \{ X_i^{(0)}(k) \mid k = 1, 2, \dots, N; \quad i = 1, 2, \dots, h \}$$

Its GM(n, h) can be written in the form of

$$\sum_{i=0}^n a_i \frac{d^{n-i} X_1^{(1)}}{dt^{n-i}} = \sum_{j=1}^{h-1} b_j X_{j+1}^{(1)}$$

where

$$a_0 = 1$$

$$X_i^{(1)}(k) = \sum_s^k X_i^{(0)}(s) \quad \text{for } i = 1, 2, \dots, h; \quad k = 1, 2, \dots, N$$

For equal time step,

$$\Delta t = t_{i+1} - t_i = \text{Const} = 1$$

then the differentiation can be replaced with

$$\frac{d^{n-i} X_1^{(1)}}{dt^{n-i}} = \frac{\Delta^{n-i} X_1^{(1)}}{\Delta t^{n-i}}$$

and

$$\frac{\Delta^{n-i} X_1^{(1)}}{\Delta t^{n-i}} = \Delta^{n-i} X_1^{(1)}$$

Now the GM(n, h) becomes

$$\sum_{i=0}^n a_i \Delta^{n-i} X_1^{(1)} = \sum_{j=1}^{h-1} b_j X_{j+1}^{(1)}$$

Since

$$\Delta^{n-i} X_1^{(1)} = I^{(n-i)} \left(X_1^{(1)} \right)$$

the GM(n, h) further becomes

$$\sum_{i=0}^n a_i I^{(n-i)} \left(X_1^{(1)} \right) = \sum_{j=1}^{h-1} b_j X_{j+1}^{(1)}$$

i.e.,

$$\sum_{i=0}^n a_i I^{(n-i)} \left(X_1^{(1)}(k) \right) = \sum_{j=1}^{h-1} b_j X_{j+1}^{(1)}(k)$$

Taking $k = 2, 3, \dots, N$, we have the following equations

$$\begin{cases} \sum_{i=0}^n a_i I^{(n-i)} \left(X_1^{(1)}(2) \right) = \sum_{j=1}^{h-1} b_j X_{j+1}^{(1)}(2) \\ \sum_{i=0}^n a_i I^{(n-i)} \left(X_1^{(1)}(3) \right) = \sum_{j=1}^{h-1} b_j X_{j+1}^{(1)}(3) \\ \vdots \\ \sum_{i=0}^n a_i I^{(n-i)} \left(X_1^{(1)}(N) \right) = \sum_{j=1}^{h-1} b_j X_{j+1}^{(1)}(N) \end{cases}$$

Extend the left side of the equations,

$$\begin{cases} I^{(n)} \left(X_1^{(1)}(2) \right) + a_1 I^{(n-1)} \left(X_1^{(1)}(2) \right) + \dots + a_n I^{(0)} \left(X_1^{(1)}(2) \right) = \sum_{j=1}^{h-1} b_j X_{j+1}^{(1)}(2) \\ I^{(n)} \left(X_1^{(1)}(3) \right) + a_1 I^{(n-1)} \left(X_1^{(1)}(3) \right) + \dots + a_n I^{(0)} \left(X_1^{(1)}(3) \right) = \sum_{j=1}^{h-1} b_j X_{j+1}^{(1)}(3) \\ \vdots \\ I^{(n)} \left(X_1^{(1)}(N) \right) + a_1 I^{(n-1)} \left(X_1^{(1)}(N) \right) + \dots + a_n I^{(0)} \left(X_1^{(1)}(N) \right) = \sum_{j=1}^{h-1} b_j X_{j+1}^{(1)}(N) \end{cases}$$

By transforming,

$$\begin{cases} I^{(n)}(X_1^{(1)}(2)) = -\left[a_1 I^{(n-1)}(X_1^{(1)}(2)) + \dots + a_n I^{(0)}(X_1^{(1)}(2)) \right] + \sum_{j=1}^{h-1} b_j X_{j+1}^{(1)}(2) \\ I^{(n)}(X_1^{(1)}(3)) = -\left[a_1 I^{(n-1)}(X_1^{(1)}(3)) + \dots + a_n I^{(0)}(X_1^{(1)}(3)) \right] + \sum_{j=1}^{h-1} b_j X_{j+1}^{(1)}(3) \\ \vdots \\ I^{(n)}(X_1^{(1)}(N)) = -\left[a_1 I^{(n-1)}(X_1^{(1)}(N)) + \dots + a_n I^{(0)}(X_1^{(1)}(N)) \right] + \sum_{j=1}^{h-1} b_j X_{j+1}^{(1)}(N) \end{cases}$$

Now let

$$Y_N = \left(I^{(n)}(X_1^{(1)}(2)), I^{(n)}(X_1^{(1)}(3)), \dots, I^{(n)}(X_1^{(1)}(N)) \right)^T$$

$$A = \begin{pmatrix} -I^{(n-1)}(X_1^{(1)}(2)) & -I^{(n-2)}(X_1^{(1)}(2)) & \dots & -I^{(1)}(X_1^{(1)}(2)) \\ -I^{(n-1)}(X_1^{(1)}(3)) & -I^{(n-2)}(X_1^{(1)}(3)) & \dots & -I^{(1)}(X_1^{(1)}(3)) \\ \vdots & \vdots & \ddots & \vdots \\ -I^{(n-1)}(X_1^{(1)}(N)) & -I^{(n-2)}(X_1^{(1)}(N)) & \dots & -I^{(1)}(X_1^{(1)}(N)) \end{pmatrix}$$

$$B = \begin{pmatrix} -I^{(0)}(X_1^{(1)}(2)) & X_2^{(1)}(2) & \dots & X_h^{(1)}(2) \\ -I^{(0)}(X_1^{(1)}(3)) & X_2^{(1)}(3) & \dots & X_h^{(1)}(3) \\ \vdots & \vdots & \ddots & \vdots \\ -I^{(0)}(X_1^{(1)}(N)) & X_2^{(1)}(N) & \dots & X_h^{(1)}(N) \end{pmatrix}$$

and write the equations in matrix form

$$Y_N = A \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n-1} \end{pmatrix} + B \begin{pmatrix} a_n \\ b_1 \\ \vdots \\ b_{h-1} \end{pmatrix}$$

i.e.,

$$Y_N = (A \vdots B)(a_1, a_2, \dots, a_{n-1} \vdots a_n, b_1, b_2, \dots, b_{h-1})^T$$

Let

$$\hat{a} = (a_1, a_2, \dots, a_{n-1} \mid a_n, b_1, b_2, \dots, b_{h-1})^T$$

then the residual is

$$e = Y_N - (A \mid B)\hat{a}$$

By using the *least square estimation*, we can find

$$\hat{a} = \left((A \mid B)^T (A \mid B) \right)^{-1} (A \mid B)^T Y_N$$

Generally, $I^{(0)}(X_1^{(0)}(k))$ is replaced by

$$I^{(0)}(X_1^{(0)}(k)) = \frac{1}{2}(X_1^{(0)}(k) + X_1^{(0)}(k-1))$$

Substitute it into matrix B ,

$$B = \begin{pmatrix} -\frac{1}{2}(X_1^{(0)}(2) + X_1^{(0)}(1)) & X_2^{(0)}(2) & \dots & X_h^{(0)}(2) \\ -\frac{1}{2}(X_1^{(0)}(3) + X_1^{(0)}(2)) & X_2^{(0)}(3) & \dots & X_h^{(0)}(3) \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{2}(X_1^{(0)}(N) + X_1^{(0)}(N-1)) & X_2^{(0)}(N) & \dots & X_h^{(0)}(N) \end{pmatrix}$$

After solving for \hat{a} , we get the differential functions with known coefficients. We can further solve the differential functions and get the *time response function*.

3.3.2 GM(1, 1)

Among GM(n, h), if $h \geq 2$, there are more than one series in the grey differential equation. That means more factors are considered to affect the system properties. But the grey methods are hard to attribute the main reason among the factors and give a viable model for prediction. So the grey system theory usually suggest the GM(n, h) with $h \geq 2$ be used for analysis rather than prediction. For prediction purpose we use GM($n, 1$). The most important and popular one is

GM(1, 1).

GM(1, 1) has the form of

$$\frac{dX^{(1)}}{dt} + aX^{(1)} = u \quad (3.3.1)$$

Similar to the solving process of GM(n , h), we have the following

$$Y_N = \left(X^{(0)}(2), X^{(0)}(3), \dots, X^{(0)}(N) \right)^T$$
$$B = \begin{pmatrix} -\frac{1}{2}(X^{(1)}(2) + X^{(1)}(1)) & 1 \\ -\frac{1}{2}(X^{(1)}(3) + X^{(1)}(2)) & 1 \\ \vdots & \vdots \\ -\frac{1}{2}(X^{(1)}(N) + X^{(1)}(N-1)) & 1 \end{pmatrix}$$

The estimation for the coefficients a , u is

$$\hat{a} = (a, u)^T = (B^T B)^{-1} B^T Y_N$$

Now consider the first order differential equation of the form

$$\frac{dy}{dx} + py = q$$

where p and q are constant. Its solution is

$$y = e^{-\int_0^x p dx} \left[q e^{\int_0^x p dx} dx + C \right]$$
$$= e^{-px} \left[\frac{q}{p} (e^{px} - 1) + C \right]$$

where C is a constant.

Similarly, for the discrete case, the solution to (3.3.1) is

$$\hat{X}^{(1)}(i+1) = e^{-ai} \left[\frac{u}{a}(e^{ai} - 1) + C \right] \quad \text{for } i=0,1,2,\dots$$

The constant C is determined by the initial value. Letting $i=0$, we get

$$C = X^{(1)}(1) = X^{(0)}(1)$$

Thus the *time response function* is

$$\begin{aligned} \hat{X}^{(1)}(i+1) &= e^{-ai} \left[\frac{u}{a}(e^{ai} - 1) + X^{(0)}(1) \right] \\ &= \left(X^{(0)}(1) - \frac{u}{a} \right) e^{-ai} + \frac{u}{a} \quad \text{for } i=0,1,2,\dots \end{aligned}$$

Hence by 1-IAGO

$$\begin{cases} \hat{X}^{(0)}(1) = \hat{X}^{(1)}(1) \\ \hat{X}^{(0)}(i+1) = \hat{X}^{(1)}(i+1) - \hat{X}^{(1)}(i), \quad \text{for } i=1,2,\dots \end{cases}$$

where $\hat{X}^{(0)}$ is the estimation on the original series by the GM(1, 1). For $i=N$, $\hat{X}^{(0)}(N+1)$ is one step forward prediction. If we take $i>N$, we can get more predictive values. But the farther we step forward, the larger the error will be.

3.3.3 GM(2, 1)

GM(2, 1) has the form

$$\frac{d^2 X^{(1)}}{dt^2} + a_1 \frac{dX^{(1)}}{dt} + a_2 X^{(1)} = u$$

Use $n=2$ and $h=1$ to substitute into the solution for GM(n, h), we have

$$Y_N = \left(I^{(2)}(X^{(1)}(2)), I^{(2)}(X^{(1)}(3)), \dots, I^{(2)}(X^{(1)}(N)) \right)^T$$

$$A = \begin{pmatrix} -I^{(1)}(X(2)) \\ -I^{(1)}(X(3)) \\ \vdots \\ -I^{(1)}(X(N)) \end{pmatrix}$$

$$B = \begin{pmatrix} -\frac{1}{2}(X^{(1)}(2) + X^{(1)}(1)) & 1 \\ -\frac{1}{2}(X^{(1)}(3) + X^{(1)}(2)) & 1 \\ \vdots & 1 \\ -\frac{1}{2}(X^{(1)}(N) + X^{(1)}(N-1)) & 1 \end{pmatrix}$$

Since

$$\begin{aligned} I^{(1)}(X^{(1)}(k)) &= I^{(0)}(X^{(1)}(k)) - I^{(0)}(X^{(1)}(k-1)) \\ &= X^{(0)}(k) \end{aligned}$$

$$\begin{aligned} I^{(2)}(X^{(1)}(k)) &= I^{(1)}(X^{(1)}(k)) - I^{(1)}(X^{(1)}(k-1)) \\ &= X^{(0)}(k) - X^{(0)}(k-1) \end{aligned}$$

we finally get

$$Y = \begin{pmatrix} X^{(0)}(2) - X^{(0)}(1) \\ X^{(0)}(3) - X^{(0)}(2) \\ \vdots \\ X^{(0)}(N) - X^{(0)}(N-1) \end{pmatrix}$$

$$A = \begin{pmatrix} -X^{(0)}(2) \\ -X^{(0)}(3) \\ \vdots \\ -X^{(0)}(N) \end{pmatrix}$$

Thus the fitting coefficients

$$\begin{aligned} \hat{a} &= (a_1, a_2, u)^T \\ &= \left[(A:B)^T (A:B) \right]^{-1} (A:B)^T Y_N \end{aligned}$$

The solution is changed into solving the second order normal differential equation with known factors.

3.3.3.1 The solution for the second order differential equation

Let us consider the *homogeneous* equation

$$\frac{d^2y}{dx^2} + a\frac{dy}{dx} + by = 0$$

Its eigenfunction

$$\lambda^2 + a\lambda + b = 0$$

and eigenvalues

$$\lambda_1 = \frac{-a + \sqrt{a^2 - 4b}}{2}$$
$$\lambda_2 = \frac{-a - \sqrt{a^2 - 4b}}{2}$$

Hence if

1 $\Delta = a^2 - 4b > 0$, then

$$y = C_1 e^{\lambda_1 x} + C_2 e^{\lambda_2 x}$$

2 $\Delta = a^2 - 4b < 0$, then

$$\lambda_1 = \alpha + i\beta$$
$$\lambda_2 = \alpha - i\beta$$

where

$$\alpha = -\frac{1}{2}a$$

$$\beta = \frac{1}{2}\sqrt{4b - a^2} = \frac{1}{2}\sqrt{-\Delta}$$

Thus

$$y_1^* = e^{\alpha x} e^{i\beta x}$$

$$= e^{\alpha x} (\cos \beta x + i \sin \beta x)$$

$$y_2^* = e^{\alpha x} e^{-i\beta x}$$

$$= e^{\alpha x} (\cos \beta x - i \sin \beta x)$$

Because we need the real solutions, so let

$$y_1 = \frac{1}{2}(y_1^* + y_2^*) = e^{\alpha x} \cos \beta x$$

$$y_2 = \frac{1}{2i}(y_1^* - y_2^*) = e^{\alpha x} \sin \beta x$$

and the general solution becomes

$$y = C_1 e^{\alpha x} \cos \beta x + C_2 e^{\alpha x} \sin \beta x$$

3 $\Delta = a^2 - 4b = 0$, then

$$\lambda_1 = \lambda_2 = -\frac{1}{2}a$$

and

$$y_1 = y_2 = e^{\lambda x}$$

The general solution for the *homogeneous* case is

$$y = C_1 e^{\lambda x} + C_2 x e^{\lambda x}$$

For the *non-homogeneous* equation of the form

$$\frac{d^2y}{dx^2} + a\frac{dy}{dx} + by = m$$

Since $y = \frac{m}{b}$ is a specific solution of it, the general solution can thus be expressed as

$$y = y_1 + y_2 + \frac{m}{b}$$

Now by substituting the coefficients a_1 into a , a_2 into b , and u into m , we can get the response functions for each situation.

If $\Delta > 0$, then

$$\lambda_{1,2} = \frac{-a_1 \pm \sqrt{a_1^2 - 4a_2}}{2}$$

and

$$\hat{X}^{(1)}(k+1) = C_1 e^{\lambda_1 k} + C_2 e^{\lambda_2 k} + \frac{u}{a_2}, \text{ for } k = 0, 1, 2, \dots$$

Else if $\Delta < 0$, then

$$\alpha = -\frac{1}{2}a_1$$

$$\beta = \frac{1}{2}\sqrt{4a_2 - a_1^2}$$

and

$$\hat{X}^{(1)}(k+1) = C_1 e^{\alpha k} \cos \beta k + C_2 e^{\alpha k} \sin \beta k + \frac{u}{a_2}, \text{ for } k = 0, 1, 2, \dots$$

Else if $\Delta = 0$, then

$$\lambda_1 = \lambda_2 = -\frac{1}{2}a_1$$

and

$$\hat{X}^{(1)}(k+1) = C_1 e^{\lambda_1 k} + C_2 k e^{\lambda_2 k} + \frac{u}{a_2}, \text{ for } k = 0, 1, 2, \dots$$

Hence as in the case of GM(1, 1), by 1-IAGO

$$\begin{cases} \hat{X}^{(0)}(1) = \hat{X}^{(1)}(1) \\ \hat{X}^{(0)}(i+1) = \hat{X}^{(1)}(i+1) - \hat{X}^{(1)}(i), \text{ for } i = 1, 2, \dots \end{cases}$$

The constants C_1 and C_2 can be solved by using the initial values.

3.3.3.2 The Solving for the Constants in Response Function

Without loss of generality, we just take an example for the case of $\Delta > 0$. The response function is

$$\hat{X}^{(1)}(k+1) = C_1 e^{\lambda_1 k} + C_2 e^{\lambda_2 k} + \frac{u}{a_2}, \text{ for } k = 0, 1, 2, \dots \quad (3.3.2)$$

where

$$\lambda_{1,2} = \frac{-a_1 \pm \sqrt{a_1^2 - 4a_2}}{2}$$

Now we need to solve for C_1 and C_2 through the initial values.

For $k = 0$, $X^{(1)}(1) = X^{(0)}(1)$ and

$$C_1 + C_2 + \frac{u}{a_2} = X^{(0)}(1) \quad (3.3.3)$$

By differentiation on both sides of (3.3.2), we have

$$\frac{dX^{(1)}}{dk} = \lambda_1 C_1 e^{\lambda_1 k} + \lambda_2 C_2 e^{\lambda_2 k} \quad (3.3.4)$$

The left side of (3.3.4) is estimated by

$$\frac{dX^{(1)}}{dk} = X^{(1)}(k+1) - X^{(1)}(k) \quad (3.3.5)$$

For $k=1$,

$$\frac{dX^{(1)}}{dk} = X^{(1)}(2) - X^{(1)}(1) = X^{(0)}(2) \quad (3.3.6)$$

Substitute (3.3.6) into (3.3.4),

$$\lambda_1 e^{\lambda_1} C_1 + \lambda_2 e^{\lambda_2} C_2 = X^{(0)}(2) \quad (3.3.7)$$

Combine (3.3.3) and (3.3.7) to obtain

$$\begin{cases} C_1 + C_2 + \frac{u}{a_2} = X^{(0)}(1) \\ \lambda_1 e^{\lambda_1} C_1 + \lambda_2 e^{\lambda_2} C_2 = X^{(0)}(2) \end{cases}$$

and then solve for C_1 and C_2 .

3.4 The Error of GM

Although some preparatory estimation is taken before constructing the GM, we still need to check its accuracy after we fit it. For the application in time series, there are usually two methods checking the precision of the fitting GM. They are:

- 1) The Average Relative Error (ARE);
- 2) The Posterior Error;

The two methods only count in the observed data that have been used in constructing the model and do not include any prediction values.

3.4.1 The Average Relative Error

The Average Relative Error (ARE) is quite simple and easy to understand. We use the GM(1, 1) as an example to explain how to apply it.

For the original series

$$X^{(0)} = \{X^{(0)}(k) | k = 1, 2, \dots, N\}$$

we get the solution series

$$\hat{X}^{(0)} = \{\hat{X}^{(0)}(k) | k = 1, 2, \dots, N\}$$

Calculate the residuals

$$e(k) = X^{(0)}(k) - \hat{X}^{(0)}(k)$$

then

$$\text{ARE} = \frac{1}{N} \sum_{k=1}^N \frac{|e(k)|}{X^{(0)}(k)}$$

3.4.2 The Posterior Error

The standard error of the original series is

$$S_1 = \sqrt{\frac{1}{N} \sum_{k=1}^N (X^{(0)}(k) - \bar{X}^{(0)})^2}$$

where

$$\bar{X}^{(0)} = \frac{1}{N} \sum_{k=1}^N X^{(0)}(k)$$

Similarly, for the residual series

$$e = \{e(k) | k = 1, 2, \dots, N\}$$

the standard error is

$$S_2 = \sqrt{\frac{1}{N} \sum_{k=1}^N (e(k) - \bar{e})^2}$$

where

$$\bar{e} = \frac{1}{N} \sum_{k=1}^N e(k)$$

Now the *Posterior Error Ratio* is

$$C = \frac{S_2}{S_1}$$

and the *Small Error Probability*

$$P = P\{|e(k) - \bar{e}| < 0.6745S_1\}$$

where the constant 0.6745 is derived from the probability theory. A standard normal distribution has the probability of 50% in the range of $(-0.6745, 0.6745)$.

The precision of a GM is decided by C and P . Using M.P. to represent the Model Precision Grade, then

$$\text{M.P.} = \text{Max}\{\text{Grade of } C, \text{Grade of } P\}$$

Table 3-1 shows the precision grades, where the smaller the grade number, the better the GM is. Theoretically, Grade 4 means the model is not acceptable in terms of either C or P or both, and must be reconstructed.

Table 3-1 GM Model Precision Grades

M.P.	C	P
1	$C \leq 0.35$	$0.95 \leq P$
2	$0.35 < C \leq 0.50$	$0.80 \leq P < 0.95$
3	$0.50 < C \leq 0.65$	$0.70 \leq P < 0.80$
4	$0.65 < C$	$P < 0.70$

3.4.3 The Error of Prediction

The ARE and the posterior error ratio are the indicators of the fitting and not suitable for the judgment of the prediction. A method for the estimation of the prediction error is now presented.

The fitting solution for the GM(1, 1) is expressed as

$$\hat{X}^{(1)}(i+1) = \left(X^{(0)}(1) - \frac{u}{a} \right) e^{-ai} + \frac{u}{a}$$

Making differentiation on both sides, we get

$$\frac{d\hat{X}^{(1)}}{di} = \left(u - aX^{(0)}(1) \right) e^{-ai}$$

Taking approximation on the left side by

$$\begin{aligned} \frac{d\hat{X}^{(1)}}{di} &= \hat{X}^{(1)}(i+1) - \hat{X}^{(1)}(i) \\ &= \hat{X}^{(0)}(i+1) \end{aligned}$$

we have

$$\hat{X}^{(0)}(i+1) = \left(u - aX^{(0)}(1) \right) e^{-ai}$$

According to the propagation of the variance, the variance of the predicted value is

$$\begin{aligned} \sigma_{\hat{X}^{(0)}(i+1)}^2 &= \left(\frac{\partial \hat{X}^{(0)}(i+1)}{\partial a} \right)^2 \sigma_a^2 + \left(\frac{\partial \hat{X}^{(0)}(i+1)}{\partial u} \right)^2 \sigma_u^2 \\ &+ 2 \left(\frac{\partial \hat{X}^{(0)}(i+1)}{\partial a} \frac{\partial \hat{X}^{(0)}(i+1)}{\partial u} \right) \sigma_{au} + \left(\frac{\partial \hat{X}^{(0)}(i+1)}{\partial X^{(0)}(1)} \right)^2 \sigma_0^2 \end{aligned} \quad (3.4.1)$$

where σ_a^2 and σ_u^2 are the variances of a and u respectively. σ_{au} is the covariance of a and u . σ_0^2 is the variance of the original series.

We assume the last part of the right hand is very small and can be neglected for the calculation.

Let

$$Q = (B^T B)^{-1}$$

$$I_1 = (1, 0)^T$$

$$I_2 = (0, 1)^T$$

since

$$\hat{a} = (a, u)^T = (B^T, B)^{-1} B^T Y_N$$

then

$$a = I_1^T \hat{a} = I_1^T Q B^T Y_N$$

and

$$u = I_2^T \hat{a} = I_2^T Q B^T Y_N$$

So

$$\sigma_a^2 = (I_1^T Q B^T) (I_1^T Q B^T)^T \sigma_0^2 = I_1^T Q I_1 \sigma_0^2$$

$$\sigma_u^2 = (I_2^T Q B^T) (I_2^T Q B^T)^T \sigma_0^2 = I_2^T Q I_2 \sigma_0^2$$

$$\sigma_{au} = (I_1^T Q B^T) (I_2^T Q B^T)^T \sigma_0^2 = I_1^T Q I_2 \sigma_0^2$$

Suppose further that

$$Q = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}$$

where $Q_{12} = Q_{21}$, then

$$\sigma_a^2 = Q_{11} \sigma_0^2$$

$$\sigma_u^2 = Q_{22} \sigma_0^2$$

$$\sigma_{au} = Q_{12} \sigma_0^2 = Q_{21} \sigma_0^2$$

Substituting these expressions into equation (3.4.1) and omitting the last item,

$$\begin{aligned} \sigma_{\hat{X}^{(0)}(i+1)}^2 &= \left(aiX^{(0)}(1) - X^{(0)}(1) - ui \right)^2 e^{-2ai} Q_{11} \sigma_0^2 + e^{-2ai} Q_{22} \sigma_0^2 \\ &\quad + 2 \left(aiX^{(0)}(1) - X^{(0)}(1) - ui \right) e^{-2ai} Q_{12} \sigma_0^2 \end{aligned} \quad (3.4.2)$$

where the variance of the original series $X^{(0)}$ can be estimated by the residual series

$$\sigma_0^2 = \frac{1}{N-1} e^T e$$

and

$$\sigma_0 = \sqrt{\frac{1}{N-1} e^T e}$$

Thus taking square root of (3.4.2)

$$\sigma_{\hat{X}^{(0)}(i+1)} = \left[\left(aiX^{(0)}(1) - X^{(0)}(1) - ui \right)^2 Q_{11} + Q_{22} + 2 \left(aiX^{(0)}(1) - X^{(0)}(1) - ui \right) Q_{12} \right]^{\frac{1}{2}} e^{-ai} \sigma_0$$

As the data at time points $1, 2, \dots, N$ are used in modeling, from time point $N+1$, the values are predicted. The predicted values of $X^{(0)}(i+1)$ for $i = N+1, N+2, \dots$ will be in the range of

$$\hat{X}^{(0)}(i+1) \pm \sigma_{\hat{X}^{(0)}(i+1)} \quad (3.4.3)$$

When we investigate the historical data, we can use (3.4.3) to estimate the prediction range and examine if the actual value falls into this range. We can also compare the predicted value with the actual value and calculate the relative error. As a matter of fact, (3.4.3) is applied when the prediction point is in the real future and we do not know the actual value for which we predict. Otherwise, we prefer using the relative error to evaluate the prediction, which is definitely easier.

Fu [1992] stated that if a prediction reaches the precision of 85%, it is successful. See for the classification he suggested, where RPE stands for the absolute Relative Prediction Error.

Table 3-2 Classification for Prediction Error

RPE	Class
RPE < 10%	Excellent
10% ≤ RPE < 20%	Good
20% ≤ PRE ≤ 50%	Acceptable
50% < PRE	Unacceptable

3.5 The Extension of GM to Data Series Containing Negative Values

The grey modeling is usually limited to the nonnegative series. For series with negative numbers, we cannot directly build a model. We have to convert the original series into a nonnegative one before AGO.

Suppose the negative series

$$X_1^{(0)} = \{X_1^{(0)}(i) | i = 1, 2, \dots, N\}$$

Let

$$X_{1,\min}^{(0)} = \min \{X_1^{(0)}(i) | i = 1, 2, \dots, N\}$$

then set

$$X_2^{(0)}(i) = X_1^{(0)}(i) + 2|X_{1,\min}^{(0)}|$$

So we have a positive series

$$X_2^{(0)} = \{X_2^{(0)}(i) | i = 1, 2, \dots, N\}$$

and can thus treat it in the standard way. This method is often applied in the modeling on the residual series, within which positive and negative values occur.

It should be noted that the modeling on residuals is seldom performed for the sole purpose of fitting an independent residual series. It is, in practice, a complementary activity to the normal GM model.

The above data processing method is just a general principle. In practice, the residual model has some requirements for the residuals, which are stricter than the conditions above.

The most commonly used residual model is for the revision of $X^{(1)}$, i.e., using the residual of

$X^{(1)}$ rather than $X^{(0)}$, to build a residual model. The fitting values of the residual model will be applied to the fitting values of the original model. The original model's accuracy will thus be enhanced.

Consider

$$\varepsilon^{(0)} = \{\varepsilon^{(0)}(1), \varepsilon^{(0)}(2), \dots, \varepsilon^{(0)}(N)\}$$

where

$$\varepsilon^{(0)}(k) = X^{(1)}(k) - \hat{X}^{(1)}(k)$$

If there exists a k_0 fulfilling the two conditions

1. for any $k > k_0$, $\varepsilon^{(0)}(k)$ has the same sign (plus or minus);
2. $N - k_0 \geq 4$.

then we call $\{|\varepsilon^{(0)}(k_0)|, |\varepsilon^{(0)}(k_0+1)|, \dots, |\varepsilon^{(0)}(N)|\}$ a residual tail for modeling.

By modeling on the residual tail, we can get the response function

$$\hat{\varepsilon}^{(1)}(k+1) = \left(\varepsilon^{(0)}(k_0) - \frac{u_\varepsilon}{a_\varepsilon} \right) e^{-a_\varepsilon(k-k_0)} + \frac{u_\varepsilon}{a_\varepsilon}, \quad k \geq k_0$$

and by 1-IAGO,

$$\hat{\varepsilon}^{(0)}(k+1) = (-a_\varepsilon) \left(\varepsilon^{(0)}(k_0) - \frac{u_\varepsilon}{a_\varepsilon} \right) e^{-a_\varepsilon(k-k_0)}, \quad k \geq k_0$$

Thus the revised response function for $X^{(1)}$ is

$$\hat{X}^{(1)}(k+1) = \begin{cases} \left(X^{(0)}(1) - \frac{u}{a} \right) e^{-ak} + \frac{u}{a}, & k < k_0 \\ \left(X^{(0)}(1) - \frac{u}{a} \right) e^{-ak} + \frac{u}{a} \pm a_\varepsilon \left(\varepsilon^{(0)}(1) - \frac{u_\varepsilon}{a_\varepsilon} \right) e^{-a_\varepsilon(k-k_0)}, & k \geq k_0 \end{cases}$$

where the plus/minus sign is in accordance with the residual tail's sign. Thus $\hat{X}^{(0)}(k+1)$ is calculated by 1-AGO. This kind of revision is also called residual GM.

It should be noted that to obtain at least four residual tail data with the same sign, usually the model's dimension is large (relatively), say 10-dim. For the model with dimension of only four or five, it is not easy to obtain the residual tail with the same sign. In this paper we plan to construct a 5-dim model on different time span data. So we will not attempt the improvement by using the residual model.

University of Cape Town

4 Empirical Modeling

In this chapter we implement the grey models on stock prices from the Johannesburg Stock Exchange in South Africa. We will also use different time span data, monthly and daily, to examine the model fitting and prediction accuracy.

The minimum requirement of grey modeling is four data. We use five, i.e., our models are 5-dim. There is no special reason for us to choose the 5-dim. We just want to test the efficacy of grey modeling on small data set. We also take it into consideration that there are five trading days in a week, but do not think it will make the 5-dim model perform better than the models with other dimensions.

For each stock, the price data are divided into groups. There are six data in a group. Five are used to build the model and the rest one, to check the prediction. We only make one step prediction.

We compute ARE and Relative Prediction Error (RPE) for each model, i.e., for each group of data set. For each stock, we take the average of all the ARE's and RPE's of the models built on it. We also record the maximum and minimum ARE and RPE for each stock.

We test the smoothness for each group of data. If the original data set $X^{(0)}$ is smooth, we will build the model on $X^{(1)}$. If $X^{(0)}$ is not smooth, we check if $X^{(1)}$ is and if so, build model on $X^{(2)}$. If $X^{(1)}$ is not smooth either, we will stop trying higher order AGO and to seek other data processing methods.

We write the program in VBA and run it in MS Excel.

4.1 GM(1, 1) Modeling on Monthly Data

We randomly take nine JSE stocks and examine the monthly prices by using GM(1, 1). The data period is from July 1988 to December 2004, 198 months. Thus we build 33 models for each

stock. There are totally 297 models for the nine stocks.

Figure 4-1 shows the stock prices. It should be noted that we have made an adjustment so the prices in the chart are not their real values. In order to put the nine stocks in one chart, for each stock we set the price in the first month to 100 and each of the rest prices to its percentage to the price in the first month. So the plot reflects only the *relative* price movement, not the actual prices.

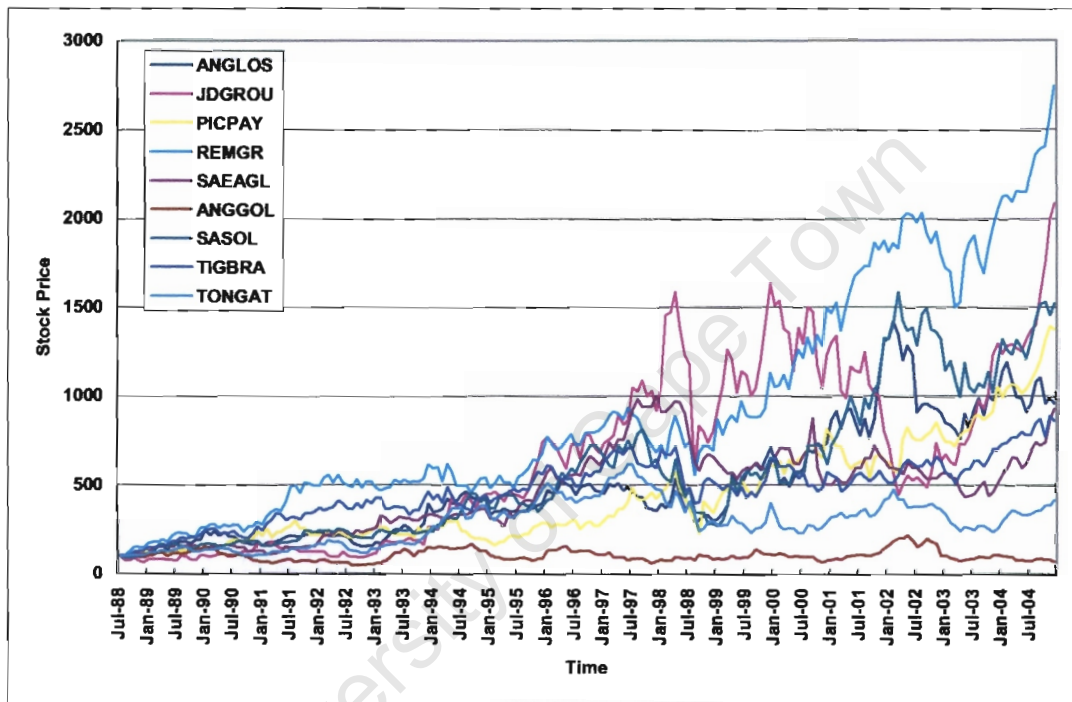


Figure 4-1 Movement of the Monthly Stock Prices

Table 4-1 lists the statistics of the modeling results. For each stock, the “Ave”, “Max” and “Min” are the average, maximum and minimum of the 33 models respectively.

ANGGOL has the highest average ARE, 3.469%, while TIGBRA has the lowest, 2.202%. The average ARE of the nine stocks is 2.697%, which is statistically the fitting error (excluding predictions) of GM(1, 1) on monthly stock prices.

JDGROU has the highest average RPE, 14.089%, while SAEAGL has the lowest, 6.557%. The average RPE of the nine stocks is 9.558%. According to Fu’s [1992] claim, a prediction with

precision of 85% is successful. Referring to Table 3-2, the prediction with the RPE of less than 10% is excellent. However, the classification of prediction is applied in all kinds of data series. For financial time series, we would rather take a higher standard and think the RPE of 9.558% acceptable, but not excellent.

Table 4-1 Statistics of Fitting by GM(1, 1)

Stock	Model Fitting Error (ARE)			Relative Prediction Error		
	Ave	Max	Min	Ave	Max	Min
ANGLOS	2.588%	7.349%	0.192%	8.535%	46.693%	0.189%
JDGROU	3.414%	9.254%	0.428%	14.089%	40.426%	0.280%
PICPAY	2.841%	8.933%	0.621%	10.158%	74.264%	0.072%
REMGR	2.361%	5.783%	0.289%	8.258%	31.478%	0.352%
SAEAGL	2.266%	9.559%	0.393%	6.557%	23.860%	0.169%
ANGGOL	3.469%	7.237%	0.722%	13.323%	73.375%	0.155%
SASOL	2.833%	8.739%	0.310%	7.341%	32.049%	0.064%
TIGBRA	2.202%	4.907%	0.243%	8.706%	35.015%	0.042%
TONGAT	2.302%	9.574%	0.466%	9.053%	30.769%	0.290%
Average	2.697%	7.926%	0.407%	9.558%	43.103%	0.179%

We make a plot in Figure 4-2 for the average ARE and RPE by stocks.

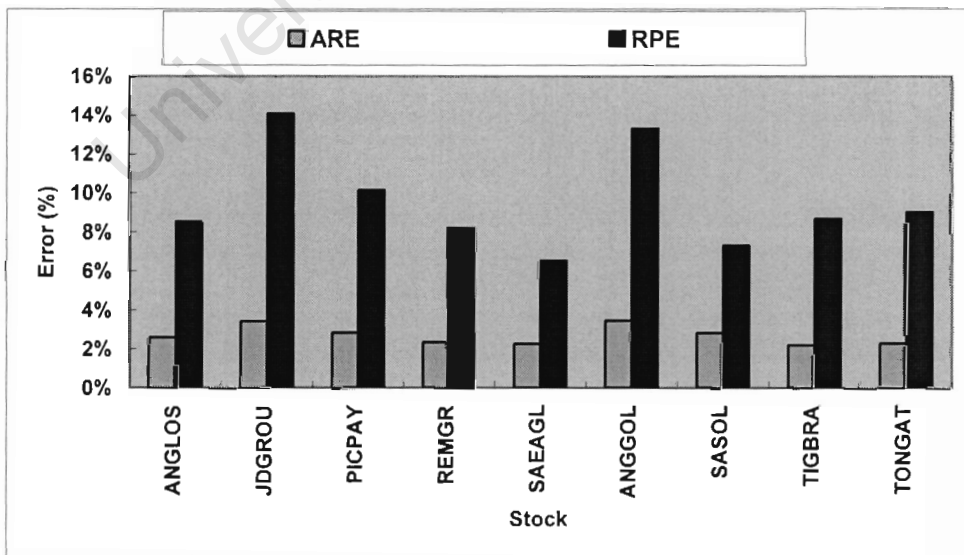


Figure 4-2 Average ARE and RPE by Stocks

Among the 297 models, the maximum and minimum ARE are 9.574% and 0.192% respectively.

The maximum RPE is 74.264% which is absolutely too high, while the minimum RPE is merely 0.042%. The unusually large ARE and RPE are probably caused by the volatile change in stock prices on which the model is built. For a specific data set that has large error, some sophisticated data processing methods before modeling or their combination may reduce the error. But as we are investigating the general performance of GM(1, 1), there is no need to further develop a method only applying in a specific data set.

From the total 297 models we pick one with the nearest absolute ARE and RPE to the averages and show its fitted results in Table 4-2 and Figure 4-3. This is a model on TONGAT for the data period of July 1993 to November 1993. The value in December 1993 is used to check the prediction. Its ARE is 2.688%, RPE, 7.749%, and MP grade, 1. The fitting coefficients are $a = -0.0746$ and $u = 1813.518$. Its response function is

$$\hat{X}^{(1)}(i+1) = \left(X^{(0)}(1) + 24309.893 \right) e^{0.0746i} - 24309.893 \quad \text{for } i = 0, 1, 2, \dots$$

Table 4-2 A GM(1, 1) on TONGAT

Date	Price	Fitted / Pred	Relative Error
Jul-93	2200	2200.00	0.000%
Aug-93	2150	2053.34	-4.496%
Sep-93	2100	2212.43	5.354%
Oct-93	2350	2383.84	1.440%
Nov-93	2625	2568.53	-2.151%
Dec-93	3000	2767.53	-7.749%

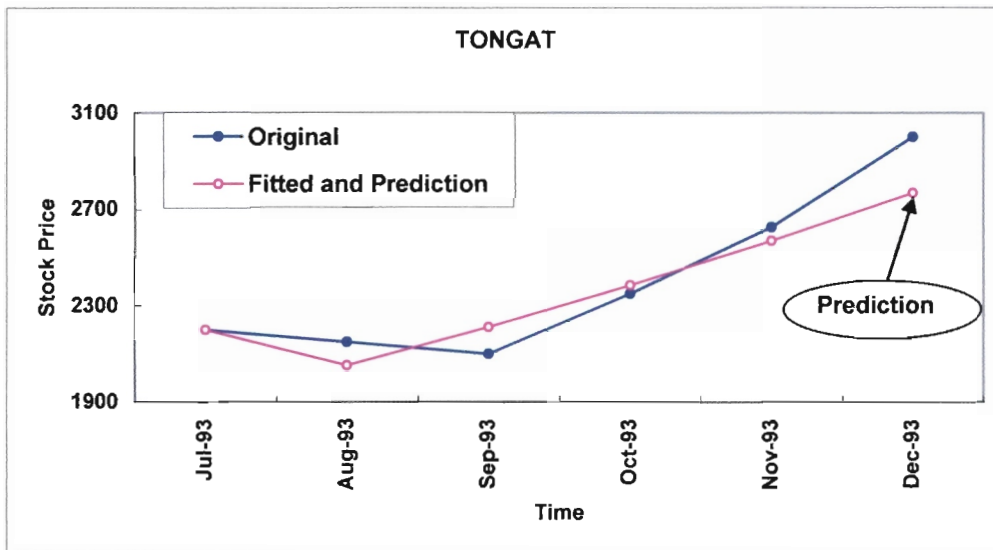


Figure 4-3 A GM(1, 1) Fitting on TONGAT

In Figure 4-3, the fitted values from the second time point is like a straight line. The reason is the exponential function can be approximated with a linear function, when the absolute value of development coefficient a is small.

For the response function

$$\hat{X}^{(1)}(i+1) = \left(X^{(0)}(1) - \frac{u}{a} \right) e^{-ai} + \frac{u}{a} \quad \text{for } i = 0, 1, 2, \dots$$

When $i = 0$, $\hat{X}^{(0)}(1) = \hat{X}^{(1)}(1) = X^{(0)}(1)$. The first fitted value is equal to the first original value.

The development coefficient a only affect the rest fitted values.

For $i = 1, 2, \dots$,

$$\begin{aligned}
\hat{X}^{(0)}(i+1) &= \hat{X}^{(1)}(i+1) - \hat{X}^{(1)}(i) \\
&= \left(X^{(0)}(1) - \frac{u}{a} \right) e^{-ai} + \frac{u}{a} - \left[\left(X^{(0)}(1) - \frac{u}{a} \right) e^{-a(i-1)} + \frac{u}{a} \right] \\
&= \left(X^{(0)}(1) - \frac{u}{a} \right) (e^{-ai} - e^{-a(i-1)}) \\
&= \left(X^{(0)}(1) - \frac{u}{a} \right) (e^{-ai} - e^{-ai} e^a) \\
&= \left(X^{(0)}(1) - \frac{u}{a} \right) (1 - e^a) e^{-ai}
\end{aligned}$$

If $|a|$ is small enough so that we can make the approximations of $e^a \approx 1+a$ and $e^{-ai} \approx 1-ai$, then

$$\begin{aligned}
\hat{X}^{(0)}(i+1) &= \left(X^{(0)}(1) - \frac{u}{a} \right) (-a)(1-ai) \\
&= (u - aX^{(0)}(1))(1-ai) \\
&= (u - aX^{(0)}(1)) - (u - aX^{(0)}(1))ai
\end{aligned}$$

That means $\hat{X}^{(0)}(i+1)$ is a linear function of the variable i , for $i=1,2,\dots$.

However, as i becomes larger, the difference of e^{-ai} and $1-ai$ gets larger, the approximation tends to lose its efficiency and the linearity becomes weaker.

4.2 GM(2, 1) Modeling on Monthly Data

The GM(1, 1) is widely exploited because it is easy to apply. Comparatively, GM(2, 1) is much more complicated. No attempt has been reported of forecasting the financial market by GM(2, 1). As sometimes GM(1, 1) does not fit well, we implement GM(2, 1) by using the same monthly data as GM(1, 1) and examine whether it is able to enhance the fitting. There are also 33 models built for each stock.

We find that GM(2, 1) generally fits much worse than GM(1, 1). For example of ANGLOS, the average ARE of the 33 models is 242.383%. Although the minimum ARE is only 0.970%, the maximum is as large as 2492.071%. The average RPE is 15439.440%. Other stocks have the

similar situation.

But in some occasions, GM(2, 1) fits better than GM(1, 1). The model on ANGLOS for the period of July 1993 to November 1993 does so. Table 4-3 and Table 4-4 list the comparative statistics. Both models are grade 1 in MP. The model ARE changes from 4.419% in GM(1, 1) to 3.178% in GM(2, 1). The prediction is also enhanced from -28.858% to -17.553%. We see from Figure 4-4 that the fitting line of GM(2, 1) bends up, better reflects the trend of the stock prices.

Table 4-3 Model Fitting Comparison of GM(1, 1) and GM(2, 1)

••	ARE	a / a1	a2	u	RPE	M.P.
GM(1, 1)	4.419%	-0.0407	••	3095.332	-28.858%	1
GM(2, 1)	3.178%	-0.4525	-0.0602	-2224.568	-17.553%	1

Table 4-4 Error Comparison for Each Value of GM(1, 1) and GM(2, 1)

		GM(1, 1)		GM(2, 1)		
DATE	ANGLOS	Fitted	Error	Fitted	Error	Diff. of Abs. Error
Jul-93	3825	3825.000	0.000%	3825	0.000%	0.000%
Aug-93	3550	3317.966	-6.536%	3654.879	2.954%	-3.582%
Sep-93	3175	3455.726	8.842%	3491.639	9.973%	1.131%
Oct-93	3500	3599.206	2.834%	3502.262	0.065%	-2.770%
Nov-93	3900	3748.643	-3.881%	3787.055	-2.896%	-0.985%
Dec-93	5488	3904.284	-28.858%	4524.714	-17.553%	-11.305%

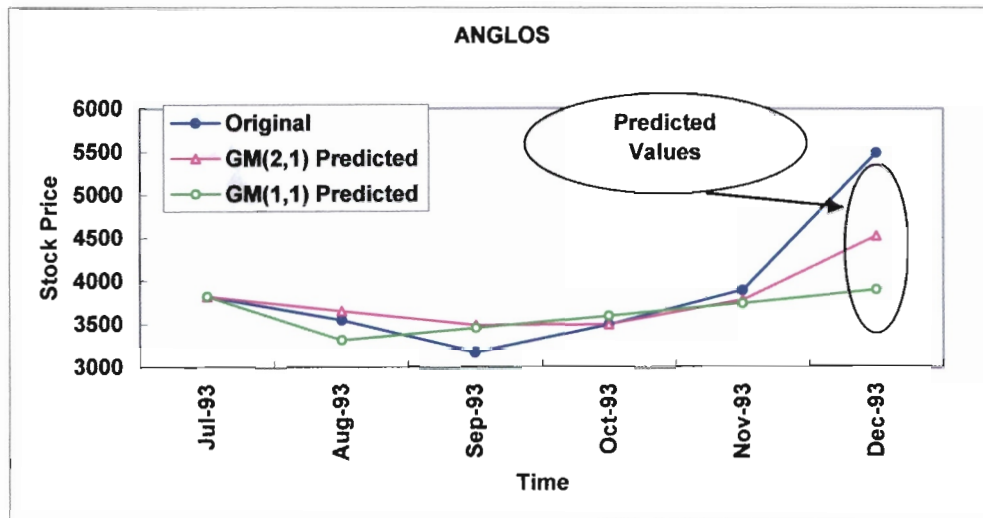


Figure 4-4 The Fitting Comparison of GM(2, 1) and GM(1, 1)

So if GM(1, 1) does not satisfy us, we can try GM(2, 1) on the same data set and see if the latter improves the fitting. If the ARE is reduced, we can use GM(2, 1) instead of GM(1, 1). Otherwise we keep using GM(1, 1) or turn to other methods.

From the solution of GM(2, 1), we can see there are two exponential items in any of the three cases. It should be able to reflect richer variations of the price movement. We think it need more experiences in implementing GM(2, 1). In some steps of the fitting, such as the determination of the matrix B and the solution for the constants in the response function, C_1 and C_2 , it should be possible to attempt other ways. Until now, no such report has been found in the literature. It is sophisticated and beyond the intention of this paper.

4.3 GM(1, 1) Modeling on Daily Data

As the GM(1, 1) is quite simple and provides good fitting, we will focus on it. In this section, we will test GM(1, 1) on daily stock prices. We randomly take 10 stocks from JSE. The time period is from September 6, 2004 to October 29, 2004, eight weeks. We model on the five week days and predict the price for the following Monday. There are eight models for each stock and 80 models totally. We also take the data in November 1 for checking the predictions.

For a stock, there might be a trading suspension because of dividend announcement, important

press release, share split, etc. The weekday without trading is a gap and the cavity idea of the grey system theory applies. If there is no price for a stock on a certain trading day, we can make a cavity value by a grey generation and then build a model. We checked all the data and found no cavity. So we can build the models without any data processing other than AGO's.

We make the same adjustment on the daily stock prices as we did to the monthly's and show the relative movement in Figure 4-5, where number 1 in time axis representing for September 6, 2004. The daily prices are much smoother than the monthly's. In the period we investigate, all the prices change in the range of $\pm 25\%$.

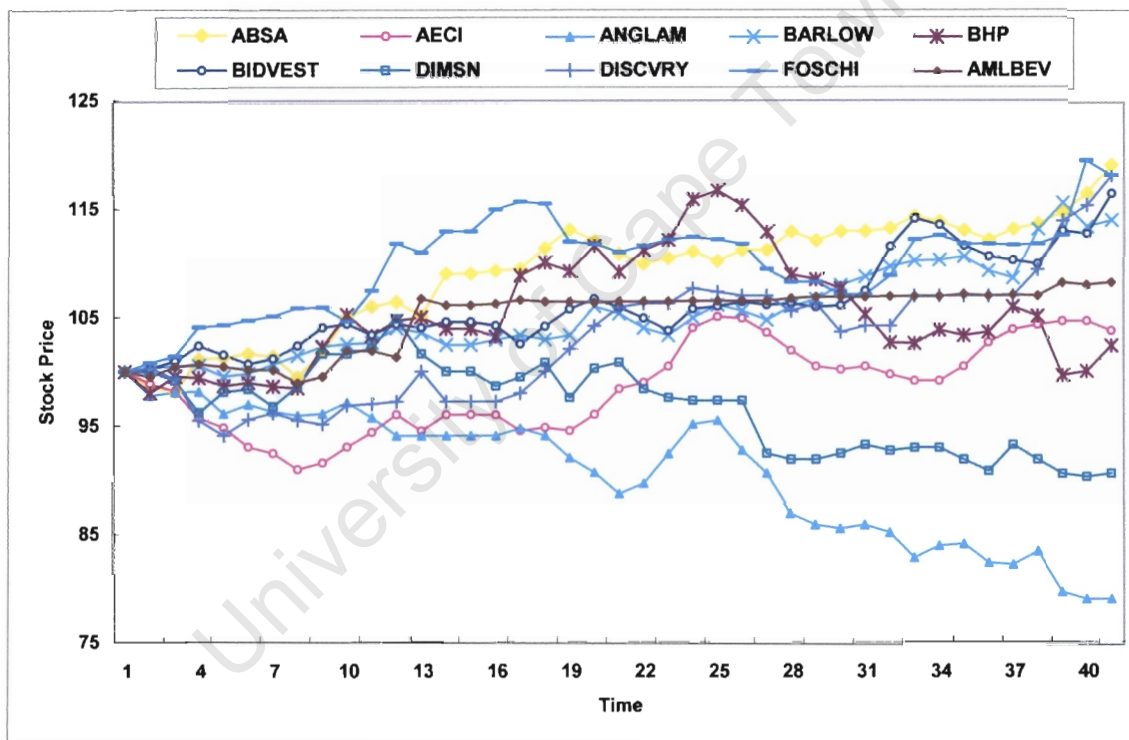


Figure 4-5 Movement of the Daily Stock Prices

The fitting statistics is listed in Table 4-5.

Table 4-5 Statistics of Fitting by GM(1, 1) on Daily Prices

Stock	Model Fitting Error (ARE)			Relative Prediction Error		
	Ave	Max	Min	Ave	Max	Min
ABSA	0.478%	0.999%	0.212%	1.104%	2.765%	0.386%
AECI	0.340%	0.641%	0.086%	1.582%	2.638%	0.037%
ANGLAM	0.456%	0.811%	0.000%	1.835%	5.978%	0.000%
BARLOW	0.352%	1.206%	0.058%	0.976%	2.537%	0.131%
BHP	0.455%	0.821%	0.184%	2.001%	5.283%	0.232%
BIDVEST	0.316%	0.804%	0.069%	1.519%	2.388%	0.009%
DIMSN	0.447%	0.852%	0.176%	1.438%	2.187%	0.550%
DISCVRY	0.412%	0.789%	0.007%	1.111%	3.938%	0.009%
FOSCHI	0.476%	1.196%	0.146%	1.289%	2.262%	0.542%
AMLBEV	0.286%	1.036%	0.011%	0.516%	2.153%	0.045%
Average	0.402%	0.916%	0.095%	1.337%	3.213%	0.194%

Comparing to monthly prices, GM(1, 1) fits much better on daily prices. The average ARE of the 80 models is reduced to 0.402% from 2.697% for monthly prices, while the maximum ARE is 1.206% and the minimum, 0.000%. Taking average for each stock, ABSA has the highest ARE, 0.478%, while AMLBEV has the lowest, 0.286%.

For the prediction, the average RPE of the 80 models is 1.337%, considerably reduced from 9.558% for monthly data. It is definitely successful, according to Fu's [1992] suggestion. Taking average for each stock, BHP has the highest RPE, 2.001%, while AMLBEV has the lowest, 0.516%.

Figure 4-6 shows the average ARE and RPE by stocks for daily data.

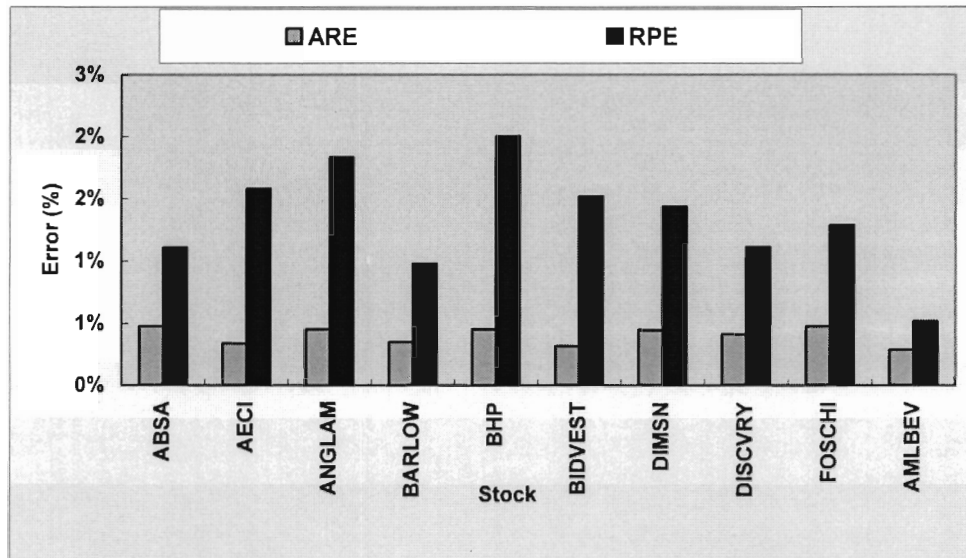


Figure 4-6 Average ARE and RPE by Stocks for Daily Data

The maximum ARE of all the 80 models is 1.206% and the maximum RPE is 5.978%. There is no unusually large error occurred. So the application of GM(1, 1) on daily stock prices is viable. Its prediction is accurate and reliable.

In Figure 4-7, the fitted pattern for each stock is consisted of the fitting results of eight models. Each model is built on five data and gives one prediction. The modeling data do not overlap. For each stock, a model's prediction is for the following Monday and compared with the actual price that is also used as the first datum in the following model, except for the last model.

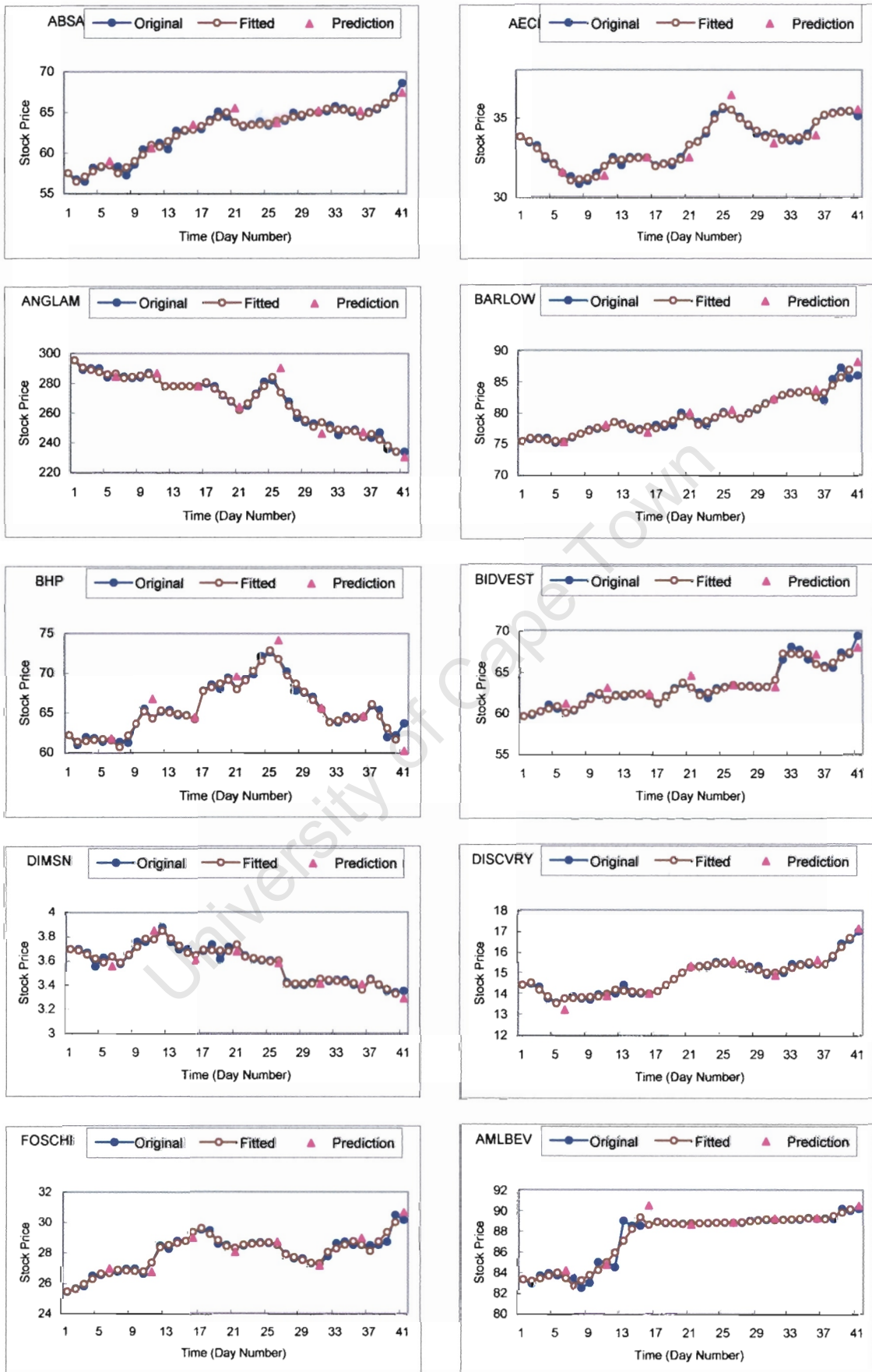


Figure 4-7 Fittings and Predictions of GM(1, 1) on Daily Stock Prices

4.4 CGM(1, 1) Modeling on Daily Data

If we consider all the natural days, the daily stock prices are not continuous because there is no trade on weekends. The two days on a weekend can be treated as cavity points by the grey system theory. Since we construct the model on the data from Monday to Friday, we can also insert two cavity values in the data series and then use all the seven data to construct a model and predict for the following Monday.

For a series with five data

$$X^{(0)} = \{X^{(0)}(1), X^{(0)}(2), X^{(0)}(3), X^{(0)}(4), X^{(0)}(5)\}$$

by using a 5-dim model, we can get the predicted value $\hat{X}^{(0)}(6)$. Now we treat it as the cavity $\Phi^{(0)}(6)$, which is put into the series

$$X^{(0)} = \{X^{(0)}(1), X^{(0)}(2), X^{(0)}(3), X^{(0)}(4), X^{(0)}(5), \Phi^{(0)}(6)\}$$

We use this series to build a 6-dim model and get the predicted value $\hat{X}^{(0)}(7)$, which is again treated as the cavity value $\Phi^{(0)}(7)$ and put into the series

$$X^{(0)} = \{X^{(0)}(1), X^{(0)}(2), X^{(0)}(3), X^{(0)}(4), X^{(0)}(5), \Phi^{(0)}(6), \Phi^{(0)}(7)\}$$

Now we have fill up both the two cavities and can build a 7-dim model on the generated series, and then will get the predicted value $\hat{X}^{(0)}(8)$, which corresponds the estimation of price for the following Monday.

The above cavities are at the end point and generated by predictions based on the previous data. So they will abide by the internal *law* reflected by the grey models. Thus the randomness of the new series which has included the cavity values is weakened. Modeling on the cavity series is supposed to fit better. If we compare the ARE of the derived 7-dim model with that of the 5-dim

model, the former should be smaller. It makes more sense to use RPE, rather than ARE, to judge the performance of cavity models.

We write the Cavity Grey Model as CGM and implement CGM(1, 1) in the same data set as GM(1, 1). The fitting statistics of the 80 models is listed in Table 4-6.

Table 4-6 Statistics of Fitting by CGM(1, 1) on Daily Prices

Stock	Model Fitting Error (ARE)			Relative Prediction Error		
	Ave	Max	Min	Ave	Max	Min
ABSA	0.341%	0.714%	0.152%	1.923%	4.551%	0.179%
AECI	0.243%	0.459%	0.061%	2.606%	7.084%	0.343%
ANGLAM	0.326%	0.580%	0.000%	3.581%	10.615%	0.000%
BARLOW	0.252%	0.862%	0.041%	2.415%	5.555%	0.564%
BHP	0.325%	0.588%	0.132%	4.213%	9.516%	0.263%
BIDVEST	0.225%	0.574%	0.050%	2.205%	5.032%	0.188%
DIMSN	0.317%	0.608%	0.118%	2.810%	5.547%	0.816%
DISCVRY	0.295%	0.564%	0.005%	3.403%	8.291%	0.315%
FOSCHI	0.340%	0.856%	0.105%	2.843%	6.268%	0.320%
AMLBEV	0.204%	0.741%	0.008%	1.168%	4.838%	0.099%
Ave	0.287%	0.655%	0.067%	2.717%	6.730%	0.309%

Table 4-7 Fitting Comparison of GM(1, 1) and MGM(1, 1)

Stock	ARE		RPE	
	GM(1,1)	CGM(1,1)	GM(1,1)	CGM(1,1)
ABSA	0.478%	0.341%	1.104%	1.923%
AECI	0.340%	0.243%	1.582%	2.606%
ANGLAM	0.456%	0.326%	1.835%	3.581%
BARLOW	0.352%	0.252%	0.976%	2.415%
BHP	0.455%	0.325%	2.001%	4.213%
BIDVEST	0.316%	0.225%	1.519%	2.205%
DIMSN	0.447%	0.317%	1.438%	2.810%
DISCVRY	0.412%	0.295%	1.111%	3.403%
FOSCHI	0.476%	0.340%	1.289%	2.843%
AMLBEV	0.286%	0.204%	0.516%	1.168%
Average	0.402%	0.287%	1.337%	2.717%

We list the comparison of GM(1, 1) and CGM(1, 1) in Table 4-7. Of the 80 models, the average

ARE is reduced to 0.287% for CGM(1, 1) from 0.402% for GM(1, 1), but the average RPE is, on the opposite, increased to 2.717% from 1.337%. Adding cavities to the prices does not lead to more accurate prediction. This conclusion is further supported by the comparison by stocks.

For each stock, the ARE and RPE are the averages of the eight models. For each one, CGM(1, 1) obviously reduced the ARE, as showed in Figure 4-8, but considerably increased the RPE, as showed in Figure 4-9.

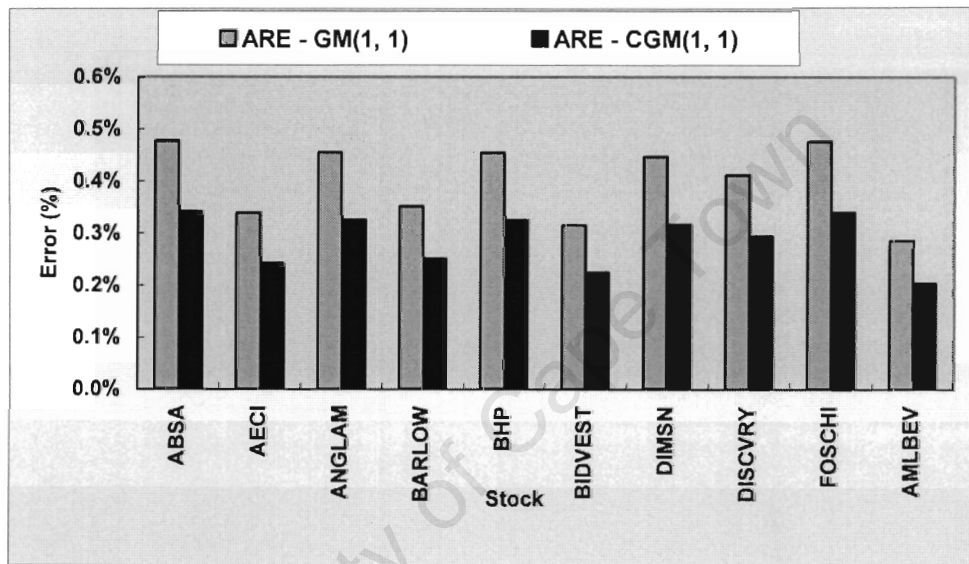


Figure 4-8 Average ARE Comparison by Stocks Between GM(1, 1) and CGM(1, 1)

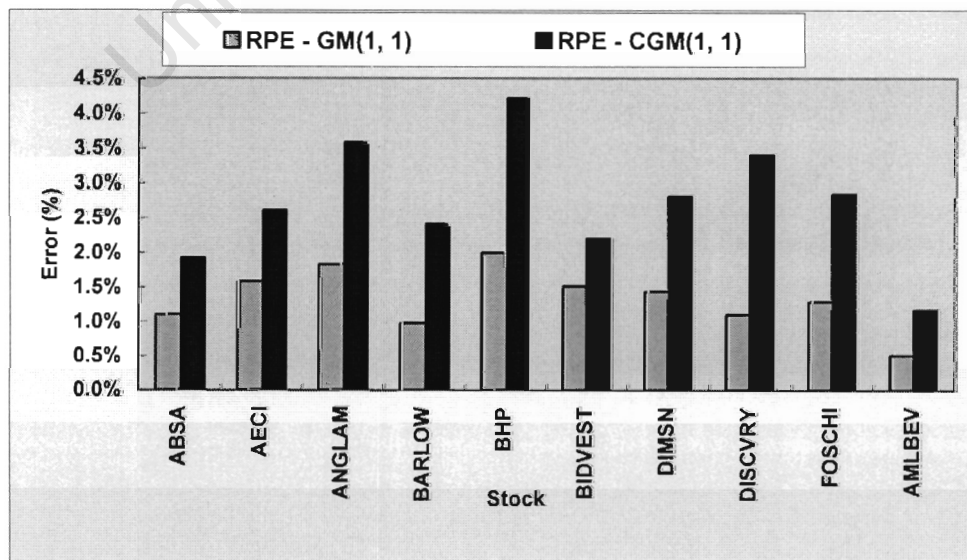


Figure 4-9 Average RPE Comparison by Stocks Between GM(1, 1) and CGM(1, 1)

The above investigation has exhibited that by the indicator of RPE, CGM(1, 1) is inferior to GM(1, 1). So there is no need to consider weekends as gaps and build cavity models. It is well accepted using trading days to analyze financial time series. For the application of grey models on stock prices, we should do the same.

University of Cape Town

5 Extensions to GM(1, 1) Model

As the GM(1, 1) is quite simple and provides good fitting, we will focus on it and try to make some improvements. Fu [1992] described using residual tail to revise the model. But as we have mentioned before, the residual model is not suitable for our 5-dim GM. There are other ways of improvement of the GM(1, 1) fitting being proposed. Li [1997] used a method of changing the approximation of the background value and improve the GM(1, 1) prediction on a stock price. We will extend his method and make a MGM(1, 1)

5.1 MGM(1, 1) Model

We examine the equation

$$\frac{dX^{(1)}}{dt} + aX^{(1)} = u \quad (6.2.1)$$

When applying to a discrete series, $dX^{(1)}(k)/dt$ is approximated by

$$\frac{dX^{(1)}(k)}{dt} = X^{(1)}(k) - X^{(1)}(k-1) = X^{(0)}(k)$$

where $k = 2, 3, \dots, N$. We can rewrite (6.2.1) as

$$X^{(0)}(k) + aX^{(1)}(k) = u, \quad \text{for } k = 2, 3, \dots, N \quad (6.2.2)$$

In which the data 2 to N are used in the estimation of the coefficients. Datum 1 is left as the initial value.

Because $dX^{(1)}(k)/dt$ is substituted by values at the two time points, $k-1$ and k , the background value $X^{(1)}(k)$ is approximated in the same way. Here we represent the background value by $Z^{(1)}$ and

$$Z^{(1)} = \{Z^{(1)}(k) | k = 2, 3, \dots, N\}$$

i.e., the dimension of $Z^{(1)}$ is one less than that of $X^{(1)}$.

Equation (6.2.2) can be rewritten as

$$X^{(0)}(k) + aZ^{(1)}(k) = u \quad \text{for } k = 2, 3, \dots, N \quad (6.2.3)$$

Now

$$B = \begin{pmatrix} -Z^{(1)}(2) & 1 \\ -Z^{(1)}(3) & 1 \\ \vdots & 1 \\ -Z^{(1)}(N) & 1 \end{pmatrix}$$

Generally, for simplification, $Z^{(1)}(k)$ just takes the equal-weight average of the values at the two points, $k-1$ and k , and the process is called neighbor values average generation as in the form

$$Z^{(1)}(k) = \frac{1}{2}(X^{(1)}(k) + X^{(1)}(k-1))$$

Now we consider changing its form to

$$Z^{(1)}(k) = \alpha X^{(1)}(k-1) + (1-\alpha)X^{(1)}(k), \quad 0 \leq \alpha \leq 1$$

If $\alpha > 0.5$, $Z^{(1)}$ is called old-info preferring; if $\alpha < 0.5$, $Z^{(1)}$ is called new-info preferring. When implementing the model, we change α from 0 to 1 at tiny steps, say 0.01. For each step, we fit the coefficients and write the response function, then compute the ARE. After the iteration, we can find the value of α corresponding to the least ARE. Since $\alpha = 0.5$ has been included in the loop, this method is obviously an optimization to the simple GM(1, 1). We call a GM(1, 1) using this method MGM(1, 1) (Modified Grey Model). Li [1997] utilized this model in the analysis of a stock in the Shanghai Stock Exchange, and confirmed its efficiency. But he only

investigated the situations for $\alpha=0$, $\alpha=0.5$ and $\alpha=1$. Besides Li's work, no other applications of the similar model for stock market data have been found in the literature.

5.2 EGM(1, 1) Model

From the structure of the response function we can see that the idea underlying the modeling of GM(1, 1) is to use an exponential curve to fit the data sequence $X^{(1)}$, generated by the 1-AGO on the original data sequence $X^{(0)}$. As a matter of fact, the data processing of a nonnegative series in terms of 1-AGO can make the accumulated sequence have monotone-increasing trend. However, 1-AGO does not guarantee that the generated sequence is monotone-increasing at an exponential changing rate. This fact is the main reason why GM(1, 1) modeling in actual applications may lose forecasting efficiency. Targeting an improvement in model accuracy, we propose a multi-transformation on the original data and attempt an Exponential Grey Model EGM(1, 1).

Lemma Given a uniform random variable U on $(0, 1]$, $E = -\frac{1}{\lambda} \ln U$ is exponentially distributed with rate parameter $\lambda > 0$.

The EGM(1, 1) method requires some transformation on the original series.

STEP 1. A grey localization on

$$X^{(0)} = \{X^{(0)}(k) | k = 1, 2, \dots, N\}$$

will help us to get a quasi-uniform distributed $U^{(0)}$ where

$$U^{(0)} = \{U^{(0)}(k) | k = 1, 2, \dots, N\}$$

and

$$U^{(0)}(k) = \frac{X^{(0)}(k) - \min(X^{(0)})}{\max(X^{(0)}) - \min(X^{(0)})}$$

We call it *quasi* because $U^{(0)}(k)$ is zero when $X^{(0)}(k) = \min(X^{(0)})$. At that point, we will use a tiny number, say 0.0001, to replace $U^{(0)}(k)$ and it will not affect the further operation.

STEP 2. Set

$$E^{(0)}(k) = -\ln U^{(0)}(k)$$

then

$$E^{(0)} = \{E^{(0)}(k) | k = 1, 2, \dots, N\}$$

should be exponentially distributed with unit rate ($\lambda = 1$).

STEP 3. However, the rate change of $E^{(0)}$ is not necessarily exponential. We further transform it to

$$D^{(0)}(k) = \exp[-(E^{(0)}(k) - E^{(0)}(1))]$$

It is actually the discrete form of standard exponential density function. So the transformed series

$$D^{(0)} = \{D^{(0)}(k) | k = 1, 2, \dots, N\}$$

is what we are seeking. Perform 1-AGO on it

$$\begin{aligned} D^{(1)}(k) &= \sum_{i=1}^k D^{(0)}(i) \\ &\approx \int_0^{E^{(0)}(k) - E^{(0)}(1)} \exp(-z) dz \\ &= 1 - \exp[-(E^{(0)}(k) - E^{(0)}(1))] \end{aligned}$$

So

$$D^{(1)} = \{D^{(1)}(k) | k = 1, 2, \dots, N\}$$

possesses not only monotone increasing trend but will also have an exponential changing rate.

STEP 4. Treating $D^{(0)}$ as the original series and modeling on $D^{(1)}$

$$\frac{dD^{(1)}}{dt} + aD^{(1)} = u$$

Now apply routine GM(1, 1). After finding the response function

$$\hat{D}^{(1)}(k+1) = \left(D^{(0)}(1) - \frac{u}{a} \right) e^{-ak} + \frac{u}{a}, \quad k = 0, 1, 2, \dots$$

Making several inverse transformations will solve for $\hat{X}^{(0)}$. In fact, there is just one step leading us back.

Notice that

$$\begin{aligned} D^{(0)}(k) &= \exp \left[- \left(E^{(0)}(k) - E^{(0)}(1) \right) \right] \\ &= \exp \left[- \left((-\ln U^{(0)}(k)) - (-\ln U^{(0)}(1)) \right) \right] \\ &= \exp \left[\ln \left(\frac{U^{(0)}(k)}{U^{(0)}(1)} \right) \right] \\ &= \frac{U^{(0)}(k)}{U^{(0)}(1)} \\ &= \frac{X^{(0)}(k) - \min X^{(0)}}{X^{(0)}(1) - \min X^{(0)}} \end{aligned}$$

We get

$$\hat{X}^{(0)}(k) = \left(X^{(0)}(1) - \min X^{(0)} \right) \hat{D}^{(0)}(k) + \min X^{(0)}$$

Again we do not need to worry about the case of $X^{(0)}(1) - \min X^{(0)} = 0$. When zero occurs, we make an adjustment by setting $\min X^{(0)}$ to be little bit smaller (say one per mill) than the old one.

5.3 Implementation of the Extended Models

In this section we implement the improved models by focusing on the daily prices.

5.3.1 MGM(1, 1) Model

We apply MGM(1, 1) on the same data set we used for GM(1, 1). We also use the 5-dim model, modeling on the prices from Monday to Friday and predict for the following Monday. There is the same number of models for each stock as the implementation of GM(1, 1).

The fitting statistics is listed in Table 5-1. Of the 80 models, the average ARE is 0.393% and the average RPE, 1.291%. Both are reduced from what GM(1, 1) gave out. The average maximum ARE and average maximum RPE are reduced too, from 0.916% and 3.213% of GM(1, 1) to 0.901% and 2.964% respectively.

Table 5-1 Statistics of Fitting by MGM(1, 1) on Daily Prices

Stock	Model Fitting Error (ARE)			Relative Prediction Error		
	Ave	Max	Min	Ave	Max	Min
ABSA	0.476%	0.993%	0.211%	1.105%	2.311%	0.190%
AECI	0.332%	0.641%	0.085%	1.550%	2.551%	0.037%
ANGLAM	0.441%	0.759%	0.000%	1.920%	5.337%	0.000%
BARLOW	0.350%	1.197%	0.053%	0.827%	1.767%	0.239%
BHP	0.453%	0.818%	0.184%	1.921%	5.477%	0.078%
BIDVEST	0.303%	0.804%	0.070%	1.503%	2.283%	0.110%
DIMSN	0.421%	0.848%	0.164%	1.455%	1.878%	0.715%
DISCVRY	0.399%	0.730%	0.005%	1.095%	4.232%	0.150%
FOSCHI	0.474%	1.188%	0.147%	1.118%	2.327%	0.438%
AMLBEV	0.284%	1.029%	0.011%	0.419%	1.471%	0.039%
Average	0.393%	0.901%	0.093%	1.291%	2.964%	0.200%

We list the comparison of GM(1, 1) and MGM(1, 1) in Table 5-2. The ARE and RPE for each stock are the averages of eight models.

Table 5-2 Fitting Comparison of GM(1, 1) and MGM(1, 1)

Stock	ARE		RPE	
	GM(1,1)	MGM(1,1)	GM(1,1)	MGM(1,1)
ABSA	0.478%	0.476%	1.104%	1.105%
AECI	0.340%	0.332%	1.582%	1.550%
ANGLAM	0.456%	0.441%	1.835%	1.920%
BARLOW	0.352%	0.350%	0.976%	0.827%
BHP	0.455%	0.453%	2.001%	1.921%
BIDVEST	0.316%	0.303%	1.519%	1.503%
DIMSN	0.447%	0.421%	1.438%	1.455%
DISCVRY	0.412%	0.399%	1.111%	1.095%
FOSCHI	0.476%	0.474%	1.289%	1.118%
AMLBEV	0.286%	0.284%	0.516%	0.419%
Average	0.402%	0.393%	1.337%	1.291%

For each stock, the average and maximum ARE's are also reduced, although the reduction is not so considerable. We show that in Figure 5-1 and Figure 5-2. That means for fitting the *historical* data, MGM(1, 1) is an efficient and reliable improvement to GM(1, 1).

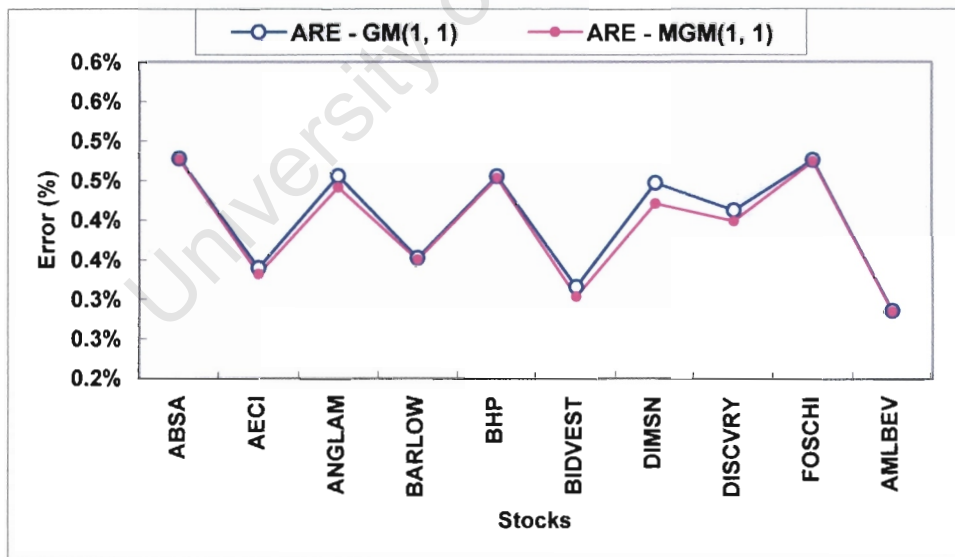


Figure 5-1 Average ARE Comparison by Stocks Between GM(1, 1) and MGM(1, 1)

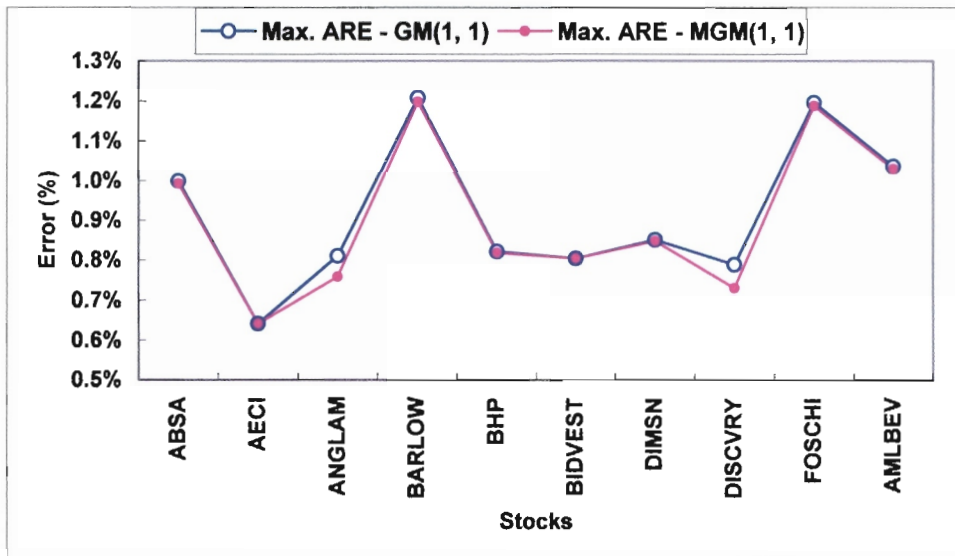


Figure 5-2 Maximum ARE Comparison by Stocks Between GM(1, 1) and MGM(1, 1)

But for prediction, MGM(1, 1) does not lower the RPE for every stock. It reduced the RPE's for most of the 10 stocks. As being showed in Figure 5-3 and Figure 5-4, MGM(1, 1) increased the RPE of ANGLAM and DIMSN, the maximum RPE of BHP, DISCVRY and FOSCHI. Since MGM(1, 1) works on seeking the minimum ARE for each model, the increase of RPE reveals that a smaller ARE does not necessarily lead to a smaller RPE.

Modeling method has a presumption that the value for which the model predicts has an internal relation to the modeling data, so the prediction based on the observed data makes sense. If the relation exists, a reasonable model gives accurate prediction. When the model fits better on observed data, the prediction is supposed to be more accurate. If it is not so, that means the relation between the modeling data and the predicted value is weak and the data series is more random. In our case, the prices of ANGLAM and DIMSN are more random than the others, because their RPE's are increased with lower ARE's. The maximum RPE for each stock could be an accidental result and not as convincing as the average.

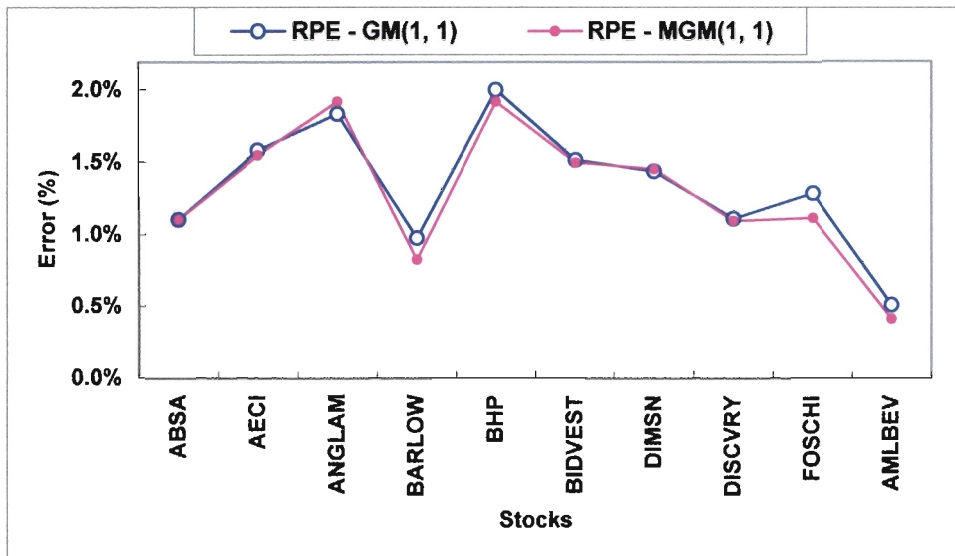


Figure 5-3 Average RPE Comparison by Stocks Between GM(1, 1) and MGM(1, 1)

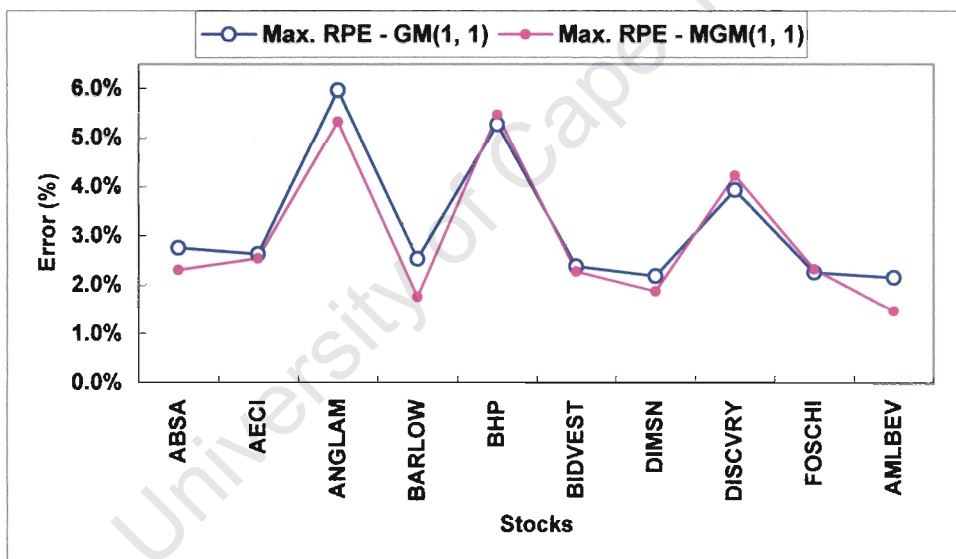


Figure 5-4 Maximum RPE Comparison by Stocks Between GM(1, 1) and MGM(1, 1)

Generally speaking, MGM(1, 1) is superior to GM(1, 1). It can be widely utilized wherever the GM(1, 1) is applicable.

5.3.2 EGM(1, 1) Model

We implement the EGM(1, 1) in the same way as GM(1, 1) and MGM(1, 1). The fitting statistics is listed in Table 5-3.

Table 5-3 Statistics of Fitting by EGM(1, 1) on Daily Prices

Stock	Model Fitting Error (ARE)			Relative Prediction Error		
	Ave	Max	Min	Ave	Max	Min
ABSA	0.606%	1.964%	0.090%	2.390%	6.558%	0.496%
AECI	1.361%	4.698%	0.093%	14.892%	72.629%	0.115%
ANGLAM	0.852%	3.219%	0.000%	5.581%	31.554%	0.000%
BARLOW	0.418%	1.330%	0.061%	1.315%	2.825%	0.116%
BHP	0.594%	1.861%	0.180%	4.820%	25.164%	0.272%
BIDVEST	0.392%	0.804%	0.069%	1.859%	5.214%	0.369%
DIMSN	0.419%	0.887%	0.035%	1.557%	4.920%	0.005%
DISCVRY	0.556%	0.857%	0.230%	1.425%	4.509%	0.095%
FOSCHI	0.491%	1.055%	0.153%	1.569%	3.496%	0.219%
AMLBEV	0.589%	2.999%	0.011%	2.809%	18.050%	0.047%
Average	0.628%	1.967%	0.092%	3.822%	17.492%	0.173%

Of the 80 models on the 10 stocks, the average ARE is 0.628% and the average RPE, 3.822%. It is quite good. But the modeling of EGM(1, 1) varies greatly from different data sets. The ARE's range from 0.000% to 4.698% and the RPE's, from 0.000% to 72.629%. That means in some cases, EGM(1, 1) might lose its efficiency.

Table 5-4 Fitting Comparison of GM(1, 1) and EGM(1, 1)

Stock	ARE		RPE	
	GM(1,1)	EGM(1,1)	GM(1,1)	EGM(1,1)
ABSA	0.478%	0.606%	1.104%	2.390%
AECI	0.340%	1.361%	1.582%	14.892%
ANGLAM	0.456%	0.852%	1.835%	5.581%
BARLOW	0.352%	0.418%	0.976%	1.315%
BHP	0.455%	0.594%	2.001%	4.820%
BIDVEST	0.316%	0.392%	1.519%	1.859%
DIMSN	0.447%	0.419%	1.438%	1.557%
DISCVRY	0.412%	0.556%	1.111%	1.425%
FOSCHI	0.476%	0.491%	1.289%	1.569%
AMLBEV	0.286%	0.589%	0.516%	2.809%
Average	0.402%	0.628%	1.337%	3.822%

As we aimed to improve the performance of GM(1, 1) by EGM(1, 1), we need to compare

between them. As being showed in Table 5-4, EGM(1, 1) does not give what we expected. Comparing to GM(1, 1), the average ARE of EMG(1, 1) increased by over 50%, and the average RPE nearly tripled. The ARE is increased for every stock, except for DIMSN. The RPE is increased for all the stock, considerably for some of them. They are showed in Figure 5-5 and Figure 5-6. So EGM(1, 1) generally failed to enhance the performance of GM(1, 1).

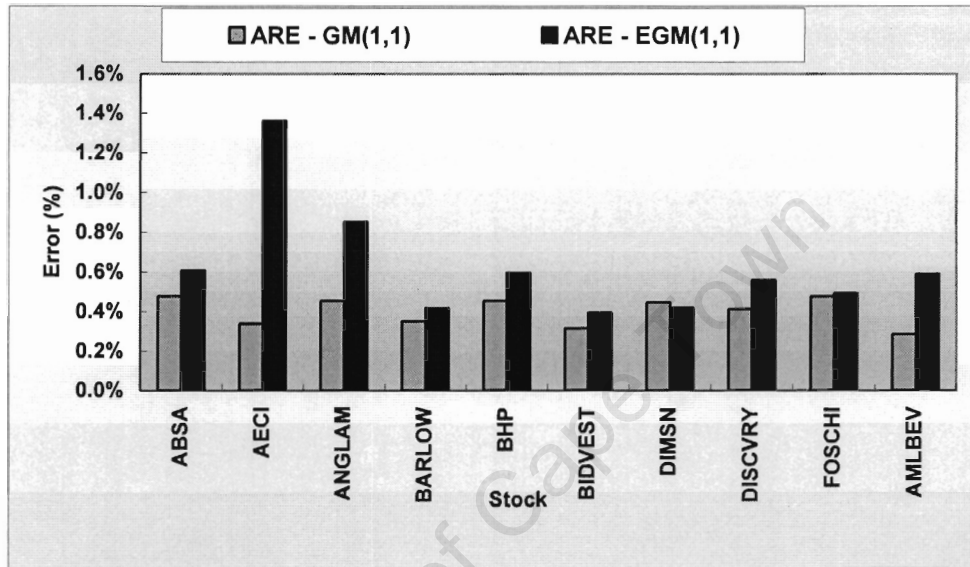


Figure 5-5 ARE Comparison by Stocks Between GM(1, 1) and EGM(1, 1)

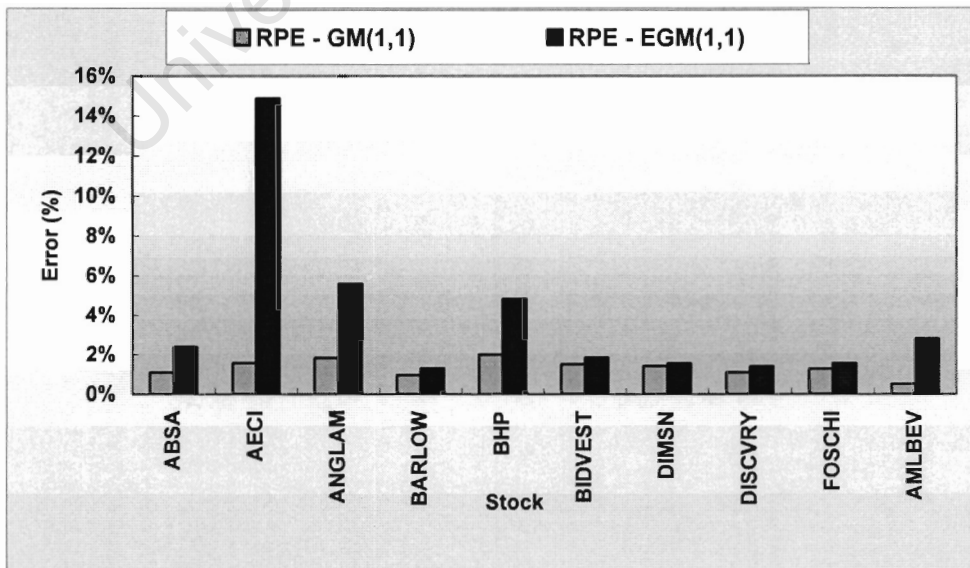


Figure 5-6 RPE Comparison by Stocks Between GM(1, 1) and EGM(1, 1)

But EGM(1, 1) has its advantage on some sort of data set. We exhibit this by picking up a model on ABSA. The modeling data period is from October 25, 2004 to October 29, 2004 and the prediction is for the price on November 1, 2004. The statistics is listed in Table 5-5.

Table 5-5 One Model Comparison Between GM(1, 1) and EGM(1, 1)

DATE	WEEKDAY	ABSA	GM(1,1) Fitted / Pred	EGM(1,1) Fitted / Pred
04-10-25	Mon	64.55	64.5492	64.5500
04-10-26	Tue	65.11	64.9369	64.9908
04-10-27	Wed	65.40	65.5592	65.4227
04-10-28	Thu	66.00	66.1874	66.0261
04-10-29	Fri	67.00	66.8217	66.8694
04-11-01	Mon	68.60	67.4620	68.0479

As showed in Figure 5-7, the EGM(1, 1) fitted values bend up and trace the stock price well, while the GM(1, 1) fitted values lose in stiffness. That typically reflects the advantage of the transformations we proposed. EGM(1, 1) fits all the values better than GM(1, 1). The model ARE is reduced to 0.090% from 0.212% for GM(1, 1). The RPE is reduced to 0.805% from 1.658% for GM(1, 1). Thus although EGM(1, 1) performs worse in general, for a particular case, it still can be used in combination with GM(1, 1) to provide a reference for the future values.

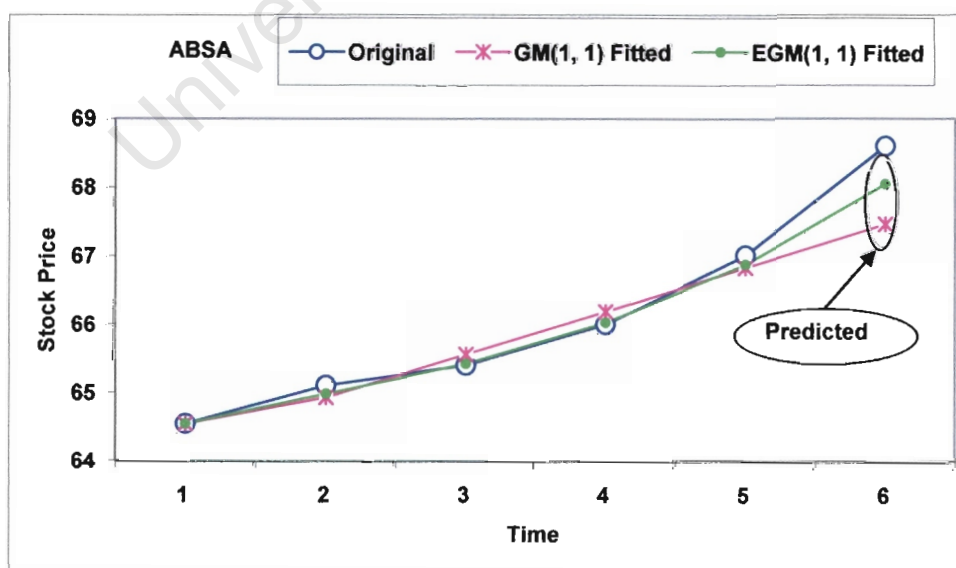


Figure 5-7 One Model Comparison Between GM(1, 1) and EGM(1, 1)

5.3.3 MEGM(1, 1) Model

We find the EGM does well at tracing the exponential trends. Seeing that the MGM(1, 1) reliably improve the GM(1, 1), we would like to try to incorporate the modification method into the EGM(1, 1) and construct an MEGM(1, 1) (Modified EGM). We implement the MEGM(1, 1) exactly as we did for the GM(1, 1), MGM(1, 1) and EGM(1, 1) on daily stock prices.

The fitting statistics is listed in Table 5-6. Of all the 80 models, the average ARE is reduced to 0.381% from 0.628% for EMG(1, 1) and the average RPE is reduced to 2.575% from 3.822% for EGM(1, 1). The ARE's range from 0.000% to 1.292% and the RPE's, from 0.000% to 46.630%. That means in some cases, EGM(1, 1) might lose its efficiency. Both of the ranges narrowed. So the modification to EGM(1, 1) is successful. The comparison listed in Table 5-7.

Table 5-6 Statistics of Fitting by MEGM(1, 1) on Daily Prices

Stock	Model Fitting Error (ARE)			Relative Prediction Error		
	Ave	Max	Min	Ave	Max	Min
ABSA	0.397%	0.644%	0.083%	1.976%	6.353%	0.511%
AECI	0.418%	0.822%	0.093%	7.100%	46.630%	0.037%
ANGLAM	0.429%	0.979%	0.000%	1.874%	6.247%	0.000%
BARLOW	0.341%	1.292%	0.056%	1.616%	3.773%	0.230%
BHP	0.388%	0.863%	0.143%	2.773%	9.861%	0.033%
BIDVEST	0.344%	0.804%	0.059%	1.665%	3.348%	0.463%
DIMSN	0.365%	0.804%	0.021%	1.230%	2.237%	0.002%
DISCVRY	0.511%	0.729%	0.207%	0.967%	2.282%	0.095%
FOSCHI	0.355%	0.777%	0.096%	5.692%	37.968%	0.211%
AMLBEV	0.264%	1.130%	0.011%	0.851%	4.026%	0.039%
Average	0.381%	0.884%	0.077%	2.575%	12.273%	0.162%

Table 5-7 Fitting Comparison of EGM(1, 1) and MEGM(1, 1)

Stock	ARE		RPE	
	EGM(1,1)	MEGM(1,1)	EGM(1,1)	MEGM(1,1)
ABSA	0.606%	0.397%	2.390%	1.976%
AECI	1.361%	0.418%	14.892%	7.100%
ANGLAM	0.852%	0.429%	5.581%	1.874%
BARLOW	0.418%	0.341%	1.315%	1.616%
BHP	0.594%	0.388%	4.820%	2.773%
BIDVEST	0.392%	0.344%	1.859%	1.665%
DIMSN	0.419%	0.365%	1.557%	1.230%
DISCVRY	0.556%	0.511%	1.425%	0.967%
FOSCHI	0.491%	0.355%	1.569%	5.692%
AMLBEV	0.589%	0.264%	2.809%	0.851%
Average	0.628%	0.381%	3.822%	2.575%

Like MGM(1, 1) did to GM(1, 1), MEMG(1, 1) effectively reduced the ARE for each stock, as showed in Figure 5-8. But still, better fitting does not guarantee a better prediction. As showed in Figure 5-9, MEGM(1, 1) reduced the RPE's for eight stocks. But the ARE for BARLOW increased and the ARE for FOSCHI was even enlarged to over three times.

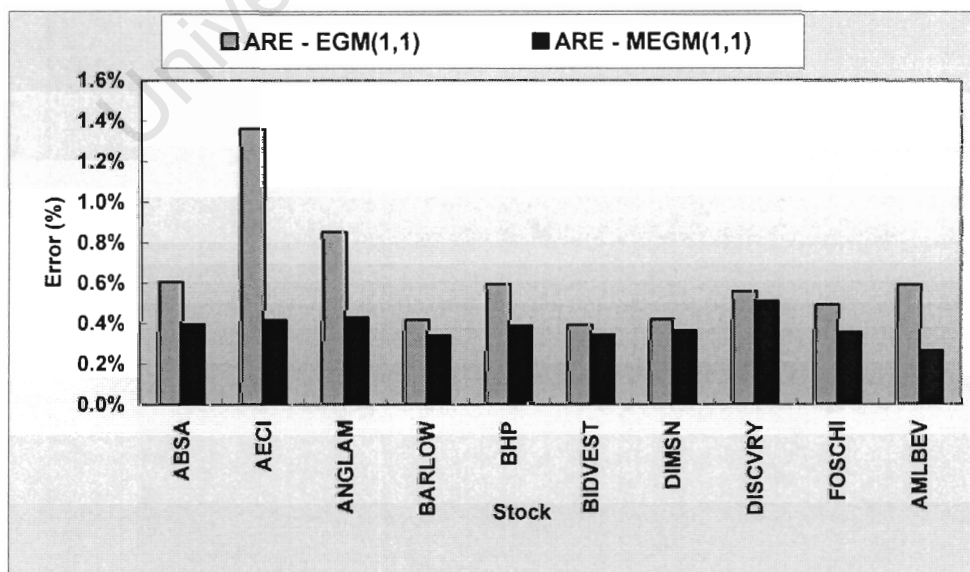


Figure 5-8 ARE Comparison by Stocks Between EGM(1, 1) and MEGM(1, 1)

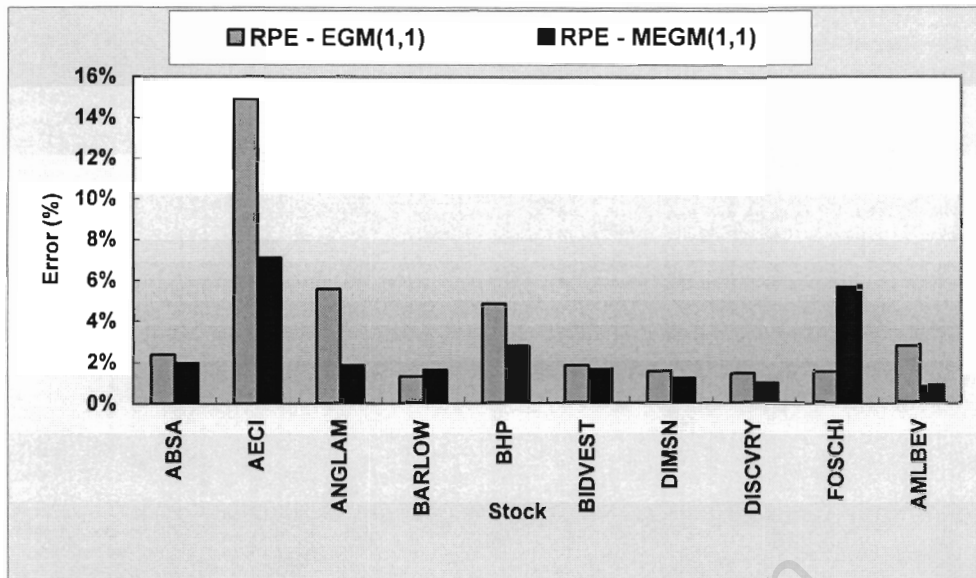


Figure 5-9 RPE Comparison by Stocks Between EGM(1, 1) and MEGM(1, 1)

Now we compare MEGM(1, 1) with GM(1, 1) in Table 5-8. The average ARE for MEGM(1, 1) is reduced to 0.381% from 0.402% for GM(1, 1), while the average RPE is increased to 2.575% from 1.337% for GM(1, 1). So the overall performance is better fitting, but worse prediction.

Table 5-8 Fitting Comparison of GM(1, 1) and MEGM(1, 1)

Stock	ARE		RPE	
	GM(1,1)	MEGM(1,1)	GM(1,1)	MEGM(1,1)
ABSA	0.478%	0.397%	1.104%	1.976%
AECI	0.340%	0.418%	1.582%	7.100%
ANGLAM	0.456%	0.429%	1.835%	1.874%
BARLOW	0.352%	0.341%	0.976%	1.616%
BHP	0.455%	0.388%	2.001%	2.773%
BIDVEST	0.316%	0.344%	1.519%	1.665%
DIMSN	0.447%	0.365%	1.438%	1.230%
DISCVRY	0.412%	0.511%	1.111%	0.967%
FOSCHI	0.476%	0.355%	1.289%	5.692%
AMLBEV	0.286%	0.264%	0.516%	0.851%
Average	0.402%	0.381%	1.337%	2.575%

As showed in Figure 5-10, there are three stocks, AECI, BIDVEST and DISCVRY, having the increased ARE's. The increment is to the maximum extent of around 20%. However, there are

eight stocks having the increased RPE, as showed in Figure 5-11. For AECI and FOSCHI, the RPE's are increased to 4.5 and times that for GM(1, 1) respectively.

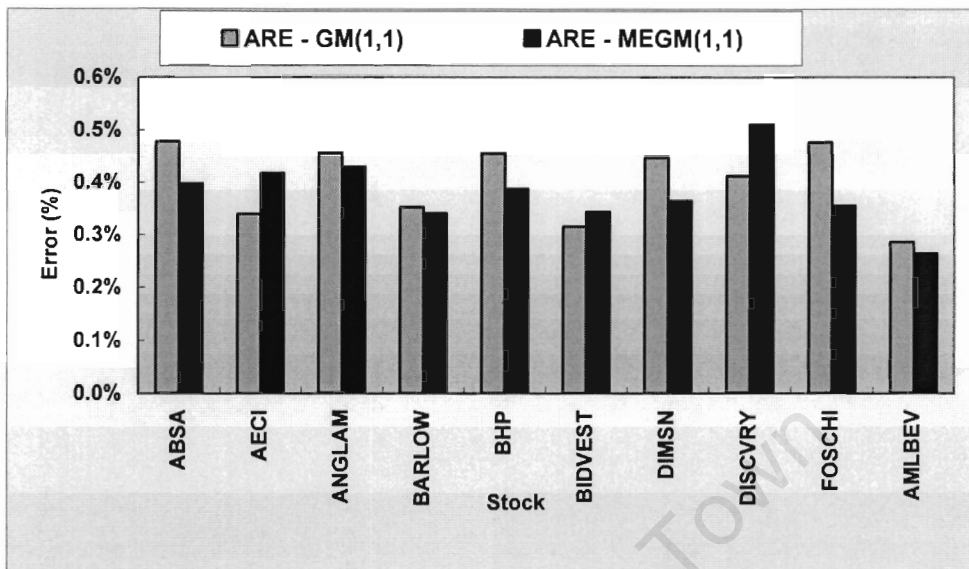


Figure 5-10 ARE Comparison by Stocks Between GM(1, 1) and MEGM(1, 1)

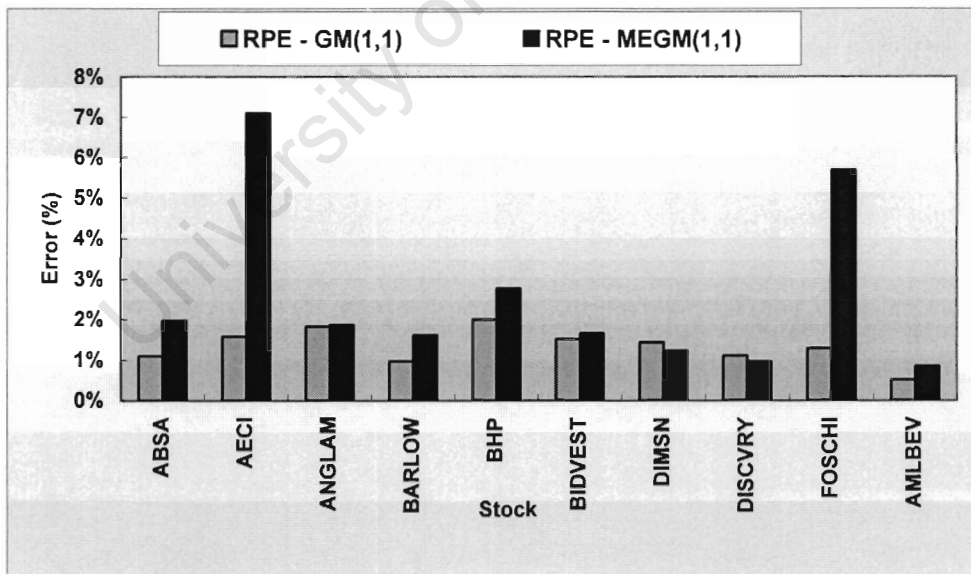


Figure 5-11 RPE Comparison by Stocks Between GM(1, 1) and MEGM(1, 1)

For the explanation of the abnormality that a smaller average ARE corresponds to a larger average RPE, we attribute it to the internal randomness of the data set. For some data, say BHP,

the application of MEGM(1, 1) made the ARE reduced, but the RPE increased. For some other data, says DIMSN, the ARE reduced, and the RPE reduced too.

Finally we compare MEGM(1, 1) with MGM(1, 1), as listed in Table 5-9. Of the 80 models, the ARE for MEGM(1, 1) is even smaller than that for MGM(1, 1). It is reduced from 0.393% to 0.381%. But the REP for MEGM(1, 1) is twice as large as that for MGM(1, 1).

Table 5-9 Fitting Comparison of MGM(1, 1) and MEGM(1, 1)

Stock	ARE		RPE	
	MGM(1,1)	MEGM(1,1)	MGM(1,1)	MEGM(1,1)
ABSA	0.476%	0.397%	1.105%	1.976%
AECI	0.332%	0.418%	1.550%	7.100%
ANGLAM	0.441%	0.429%	1.920%	1.874%
BARLOW	0.350%	0.341%	0.827%	1.616%
BHP	0.453%	0.388%	1.921%	2.773%
BIDVEST	0.303%	0.344%	1.503%	1.665%
DIMSN	0.421%	0.365%	1.455%	1.230%
DISCVRY	0.399%	0.511%	1.095%	0.967%
FOSCHI	0.474%	0.355%	1.118%	5.692%
AMLBEV	0.284%	0.264%	0.419%	0.851%
Average	0.393%	0.381%	1.291%	2.575%

The comparisons of ARE and RPE by stocks are showed in Figure 5-12 and Figure 5-13. They are very similar to the case of GM(1, 1) comparing with MEGM(1, 1).

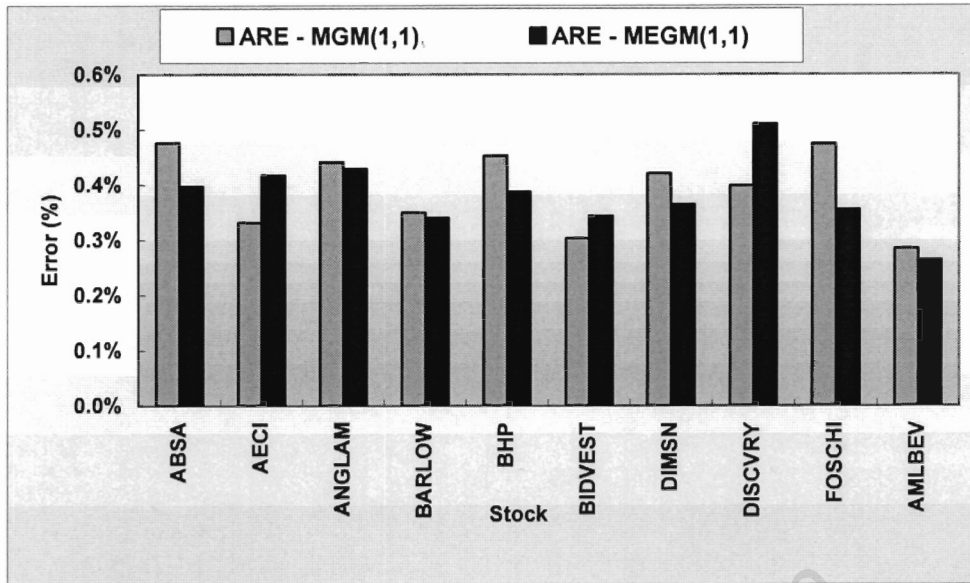


Figure 5-12 ARE Comparison by Stocks Between MGM(1, 1) and MEGM(1, 1)

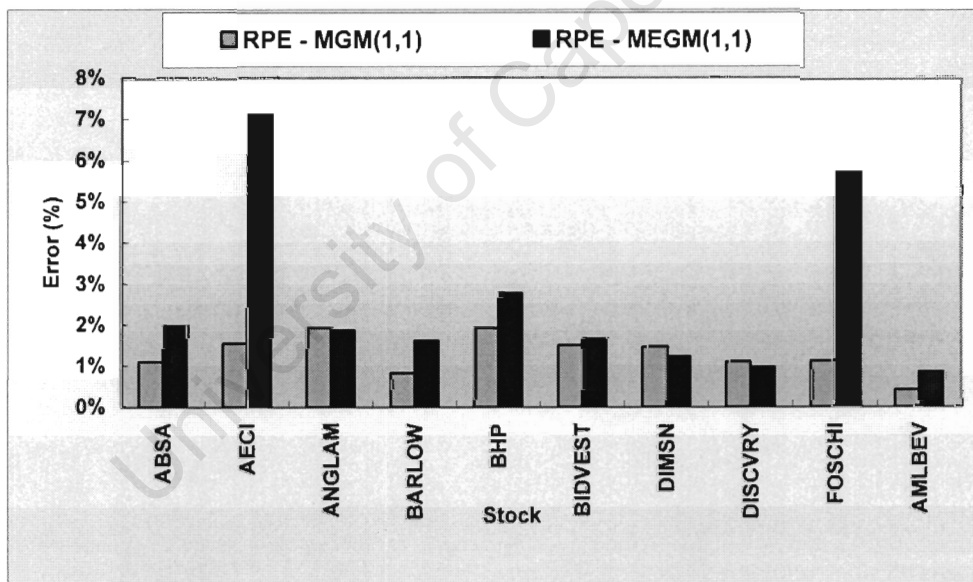


Figure 5-13 RPE Comparison by Stocks Between MGM(1, 1) and MEGM(1, 1)

So when we emphasize the analysis on historical data (without prediction), MEGM(1, 1) is superior. But when we need the prediction, MGM(1, 1) is still much better. In some particular cases, they can be exploited in combination.

To end this chapter, we exhibit the fittings and predictions by MGM(1, 1) and MEGM(1, 1) below.

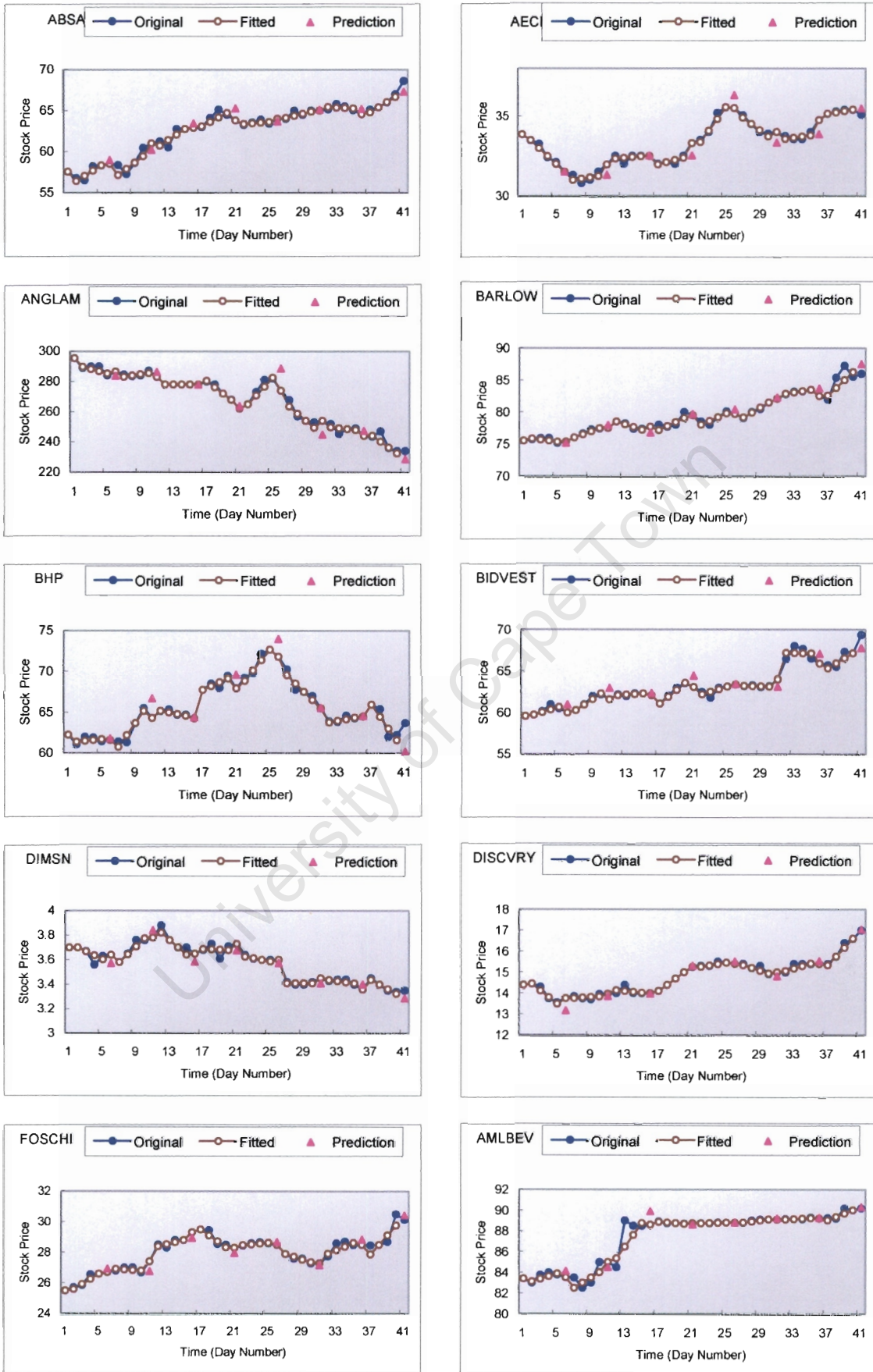


Figure 5-14 Fittings and Predictions of MGM(1, 1) on Daily Stock Prices

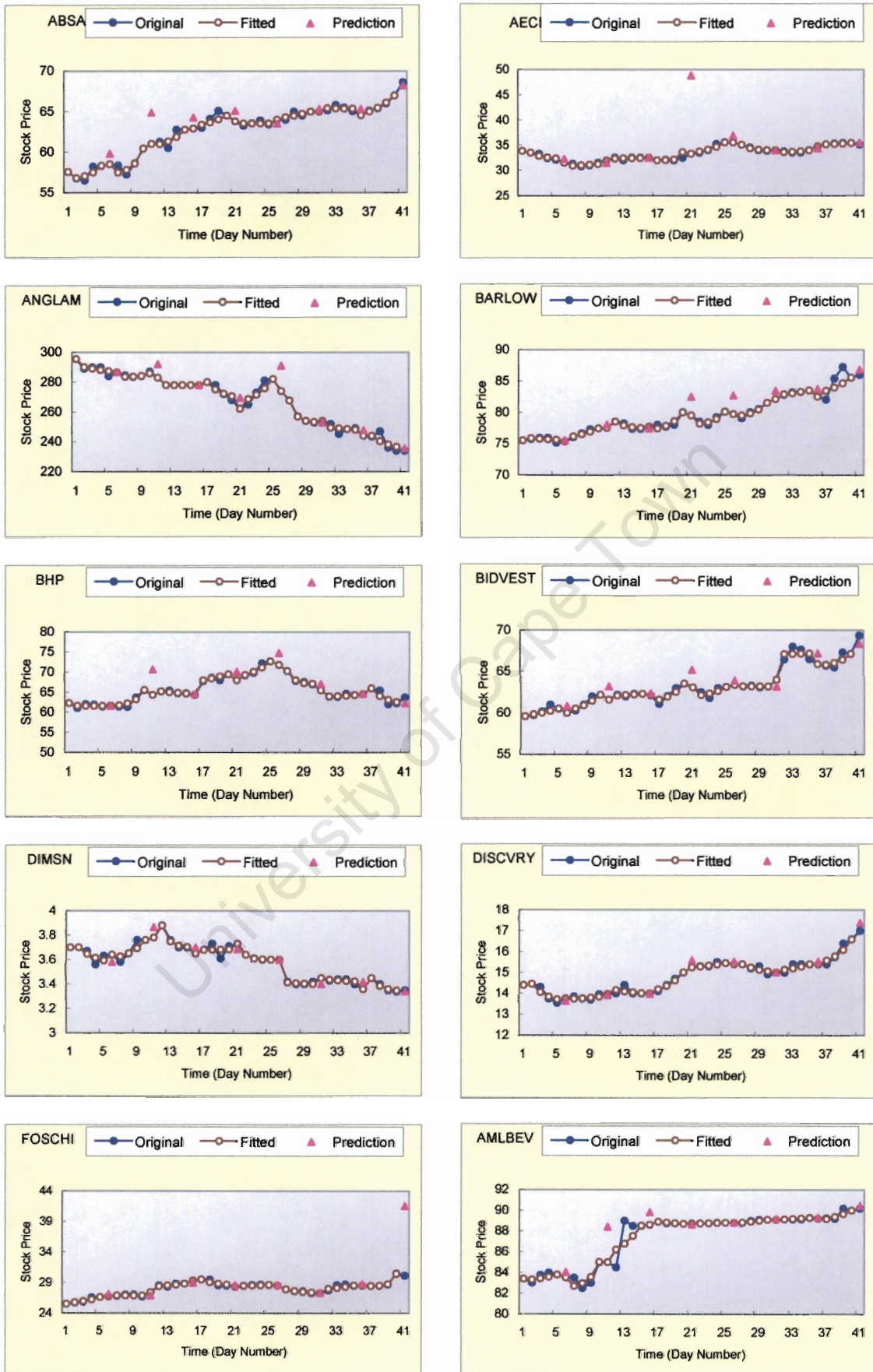


Figure 5-15 Fittings and Predictions of MEGM(1, 1) on Daily Stock Prices

6 Conclusion and Discussion

Our investigations suggest that the grey modeling can fully meet the requirement for accurate short term prediction on JSE stocks. In practice, we recommend that one take the MGM(1, 1) as a major utility, with reference to the fittings of MEGM(1, 1), and GM(2, 1).

Among the models, MGM(1, 1) has the overall superiority, as it is easy to use, precise and quite adaptable to most cases. By average on our test, it has only 0.393% model fitting error and 1.291% prediction error. The MEGM(1, 1) is impressively good at fitting the observed data, with the average ARE of 0.381%, the least of all types of the models we implemented, but on some data sets, its prediction deviates much more than what we expect from the good fitting and so the average RPE is 2.575%, almost twice as high as that given by MGM(1, 1). The MEGM(1, 1) is definitely superior to EGM(1, 1), just like MGM(1, 1) to GM(1, 1). For GM(2, 1), although sometimes it can generate better fitting and prediction, it is still very sophisticated comparatively. Only skilled users are advised to apply it.

Now we would like to discuss about the errors. Without loss of generality, we take the example of GM(1, 1).

6.1 Errors of the Models on Different Time Span

As we have tested GM(1, 1) on monthly and daily data, we can compare the results. Although there are different numbers of stocks, it won't affect the conclusion at large.

We examine Table 6-1. Almost all indicators show that GM(1, 1) fits daily data much better than monthly data, except for the minimum prediction errors which are at the same level. The daily data has an average ARE of 0.402% comparing to 2.697% monthly, and an average prediction error of 1.337% comparing to 9.558% monthly. The reason is just that daily data changes are generally smaller than monthly changes, and GM(1, 1) is able to give more accurate prediction on smooth data series.

Table 6-1 Comparison of GM(1, 1) on Daily and Monthly Data

••	Model Fitting Error (ARE)			Relative Prediction Error		
	Ave	Max	Min	Ave	Max	Min
Daily	0.402%	0.916%	0.095%	1.337%	3.213%	0.194%
Monthly	2.697%	7.926%	0.407%	9.558%	43.103%	0.179%
Difference	2.296%	7.011%	0.312%	8.220%	39.890%	-0.015%

6.2 Prediction Error Estimation

We calculate the relative prediction error (PE) by comparing with the actual values in historical data. For investigation that's the best way to check the predictability of the model. But in practice, we don't have the real future price, which is why we predict. Then how do we estimate the efficiency of the prediction? Can we just trust the model error?

6.2.1 Prediction Error with Relation to the Model Fittings

There are two indicators we used for the model fitting error. One is ARE (Average Relative Error) and the other is Model Precision (MP) Grade which is classified into four grades. We use results given for GM(1, 1) on the 10 stocks' daily data to test the relationship between the prediction error and the other indicators. Each model is built on one week's data and there are eight models for one stock. The prediction is for the price on the following Monday. The statistics is listed in Table 6-2 and Table 6-3.

Table 6-2 Result from GM(1, 1) on 10 stocks (part I)

	Week	a	u	ARE	M. P.	RPE
ABSA	1	-0.01101	55.54962	0.00498	2	0.00920
	2	-0.01313	56.36450	0.00999	3	0.00640
	3	-0.01095	59.80116	0.00685	3	0.00989
	4	-0.00862	62.52836	0.00537	3	0.02765
	5	-0.00126	63.28554	0.00251	4	0.00430
	6	-0.00389	63.86951	0.00303	3	0.00386
	7	0.00091	65.55438	0.00336	4	0.01048
	8	-0.00953	64.01357	0.00212	1	0.01658
AECI	1	0.01521	34.33043	0.00285	1	0.00248
	2	-0.00258	30.90784	0.00641	4	0.01875
	3	-0.00156	32.22414	0.00435	4	0.00037
	4	-0.00437	31.72719	0.00274	1	0.02395
	5	-0.02125	32.40588	0.00355	1	0.02638
	6	0.01152	35.56883	0.00257	1	0.01814
	7	-0.00223	33.48625	0.00385	2	0.02444
	8	-0.00239	35.06029	0.00086	1	0.01209
ANGLAM	1	0.00517	292.76850	0.00486	2	0.00686
	2	-0.00287	282.30937	0.00284	4	0.01350
	3	0.00000	278.00084	0.00000	1	0.00000
	4	0.01528	287.21542	0.00207	1	0.00767
	5	-0.02144	258.04149	0.00527	1	0.05978
	6	0.01840	272.56538	0.00737	1	0.03043
	7	0.00229	250.40619	0.00592	4	0.01340
	8	0.01639	252.07508	0.00811	3	0.01515
BARLOW	1	0.00217	76.15592	0.00251	4	0.00131
	2	-0.00631	75.37077	0.00118	1	0.00686
	3	0.00553	79.14304	0.00147	2	0.01185
	4	-0.00808	76.53679	0.00583	4	0.00656
	5	-0.00765	77.13418	0.00359	3	0.00961
	6	-0.00974	77.93467	0.00093	1	0.00167
	7	-0.00258	82.54681	0.00058	1	0.01486
	8	-0.01452	81.40530	0.01206	4	0.02537
BHP	1	-0.00163	61.28013	0.00452	4	0.00380
	2	-0.02345	58.60502	0.00603	1	0.03788
	3	0.00307	65.57306	0.00184	2	0.00354
	4	-0.00674	67.07349	0.00382	1	0.02406
	5	-0.01794	67.21723	0.00404	1	0.03328
	6	0.01485	71.25338	0.00559	1	0.00232
	7	-0.00305	63.57506	0.00235	2	0.00241
	8	0.02285	68.32944	0.00821	2	0.05283

Table 6-3 Result from GM(1, 1) on 10 stocks (part II)

	Week	a	u	ARE	M. P.	RPE
BIDVEST	1	-0.00513	59.42736	0.00328	3	0.01879
	2	-0.01095	59.38013	0.00191	1	0.02388
	3	-0.00107	62.00318	0.00109	2	0.00434
	4	-0.01333	59.94907	0.00132	1	0.02273
	5	-0.00501	61.67617	0.00411	4	0.00009
	6	0.00040	63.28263	0.00069	1	0.01336
	7	0.00030	67.20838	0.00804	2	0.01820
	8	-0.00923	64.58052	0.00481	1	0.02012
DIMSN	1	0.00879	3.73660	0.00711	4	0.02187
	2	-0.01758	3.49529	0.00444	1	0.01912
	3	0.01613	3.94335	0.00632	2	0.01081
	4	0.00078	3.69116	0.00852	4	0.01462
	5	0.00363	3.65260	0.00176	1	0.00550
	6	0.00004	3.40976	0.00232	1	0.01186
	7	0.00267	3.45516	0.00289	2	0.01333
	8	0.01133	3.50019	0.00240	1	0.01792
DSICVRY	1	0.02319	15.00577	0.00423	1	0.03938
	2	-0.00179	13.73850	0.00506	4	0.00683
	3	0.00282	14.21937	0.00789	4	0.00009
	4	-0.02041	13.68032	0.00007	1	0.00393
	5	-0.00422	15.19346	0.00234	3	0.00976
	6	0.00879	15.60856	0.00431	2	0.00843
	7	-0.00781	14.94552	0.00525	3	0.01308
	8	-0.02644	14.79643	0.00382	1	0.00737
FOSCHI	1	-0.01283	25.17619	0.00358	1	0.01198
	2	0.00130	26.97188	0.00381	2	0.02262
	3	-0.00480	28.19715	0.00341	1	0.01273
	4	0.01347	30.18262	0.00470	1	0.00932
	5	-0.00180	28.42381	0.00146	2	0.00723
	6	0.00653	28.14814	0.00173	1	0.00542
	7	-0.00821	27.70029	0.00744	2	0.01666
	8	-0.02191	27.16243	0.01196	2	0.01716
AMLBEV	1	-0.00306	82.86009	0.00227	4	0.00898
	2	-0.00598	82.00478	0.00712	4	0.00275
	3	-0.01300	84.26177	0.01036	4	0.02153
	4	0.00068	88.95528	0.00025	1	0.00140
	5	-0.00027	88.70839	0.00011	2	0.00045
	6	-0.00116	88.69427	0.00043	1	0.00180
	7	-0.00050	89.05273	0.00033	4	0.00112
	8	-0.00365	88.69039	0.00200	2	0.00324

Firstly, we examine whether ARE and MP are compatible, i.e., if small ARE corresponds to optimal MP (small in number). We take the averages of ARE's for each MP Grade and list them in Table 6-4. We see that smaller ARE relates to higher MP Grade and MP Grade one has the smallest ARE, 0.275%. As the MP Grade is lowered (with bigger figure), the average ARE increases.

Table 6-4 Average ARE at Each MP Grade

MP Grade	1	2	3	4
ARE	0.275%	0.428%	0.531%	0.543%

As showed in Figure 6-1. The ARE for MP Grade one is considerably smaller. The ARE's for MP Grade three and four are near in value. The compatibility is acceptable.

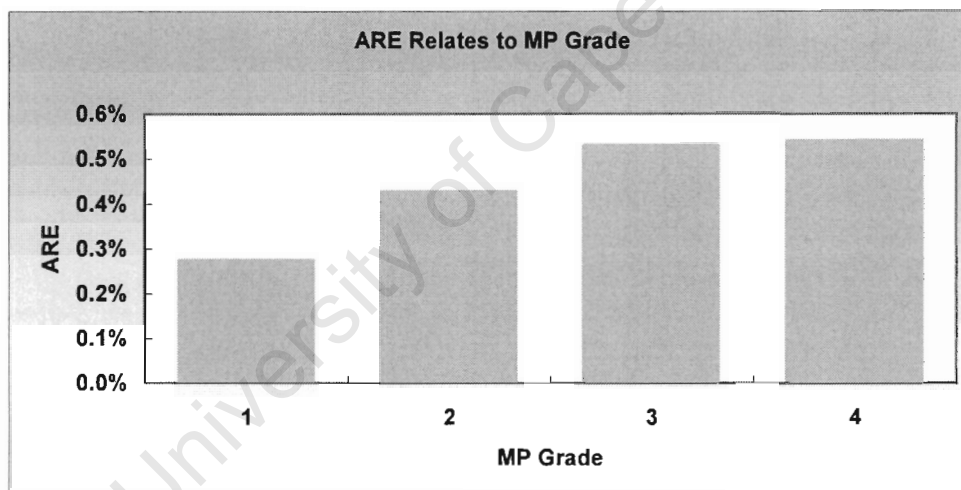


Figure 6-1 ARE Relates to MP Grade

Secondly, we examine RPE over MP Grade in the same way. The result is listed in Table 6-5. To the opposite of ARE, RPE increases as MP Grade is lowered. MP Grade four has the smallest RPE, 0.925%, while MP Grade one which is supposed to be the best fitted class of model, has the largest, 1.607%.

Table 6-5 Average RPE at Each MP Grade

MP Grade	1	2	3	4
Ave RPE	1.607%	1.298%	1.269%	0.925%

Referring to Figure 6-2, it does not make any sense to seek a high MP Grade as higher MP Grade models generate worse prediction. The MP Grade totally loses its indicative efficiency for prediction and hence is not applicable here.

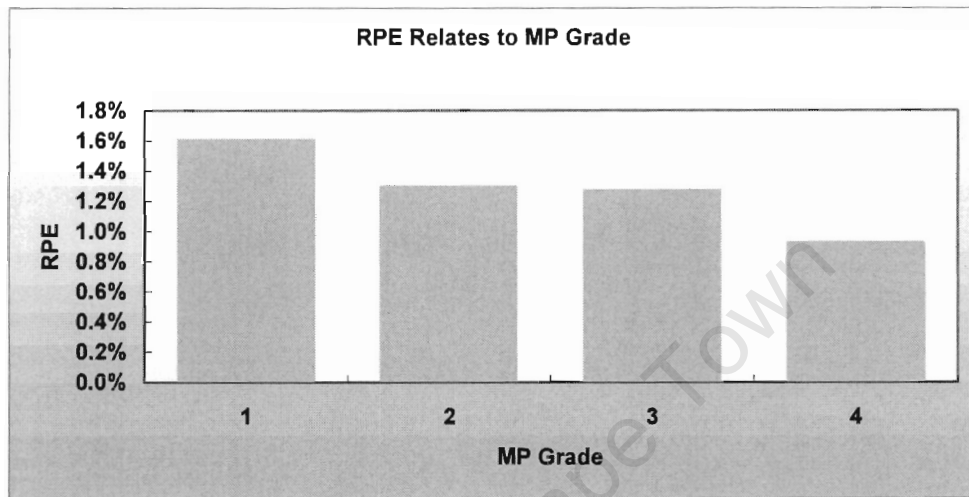


Figure 6-2 RPE Relates to MP Grade

Thirdly, in Figure 6-3 we plot RPE scatter over ARE and examine the linearity by R^2 . There is a poor linear relationship as $R^2 = 0.1229$. Smaller ARE hardly implies smaller RPE.

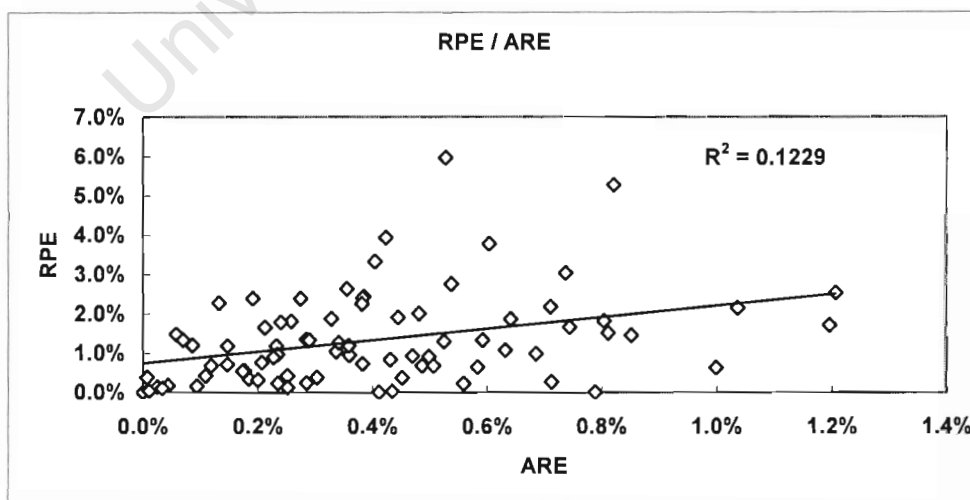


Figure 6-3 RPE Scatter over ARE

Fourthly, we examine the relation between RPE and a , the development coefficient. We use the absolute value of a in the scatter chart, Figure 6-4. The linear indicator increased to $R^2 = 0.2894$. So we can say that the prediction error is better associated with the value of a rather than the model error, ARE. We think the reason is that, being the development coefficient, a reflects the changing speed of the series. Smaller $|a|$ reflects a smooth or gentle series, so the prediction error diminishes. But the linkage is not so close, so it is not wise to merely depend on $|a|$ to guess the prediction range.

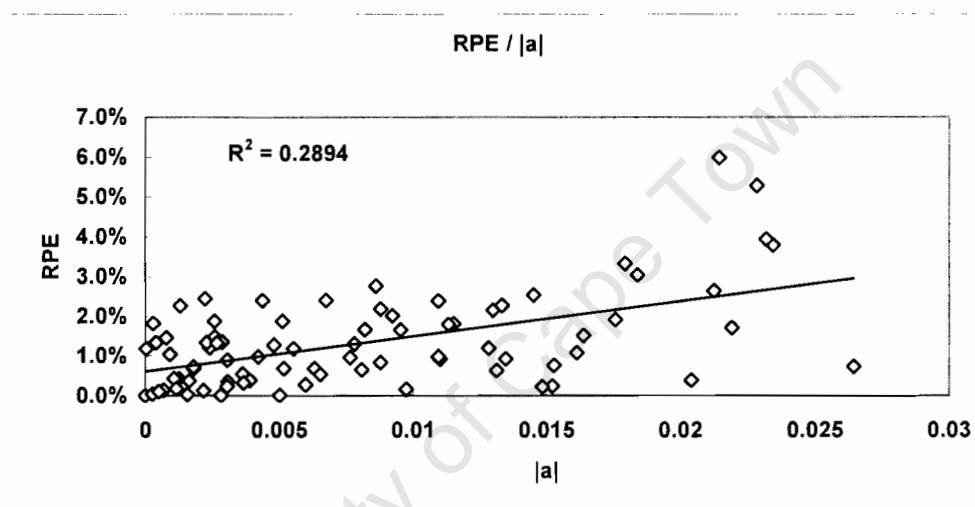


Figure 6-4 RPE Scatter over $|a|$

Finally, we wonder if the control variable u can reveal something. Because we have only the form u/a in the solution, we make the scatter chart, Figure 6-5, using u/a rather than a . According to $R^2 = 0.0472$ and what the pattern has showed to us, we know the value of u/a contributes nothing to the estimation of the prediction error.

So considering all the factors, there appears to be no quantitative relationship between the prediction error and other factors including ARE. The efficacy of ARE and MP grade is only limited to the analysis of model fitting on the known data, but can not be extended to the prediction. Therefore, it does not seem reasonable to claim the predictability of a model, solely depending on its ARE or MP Grade.

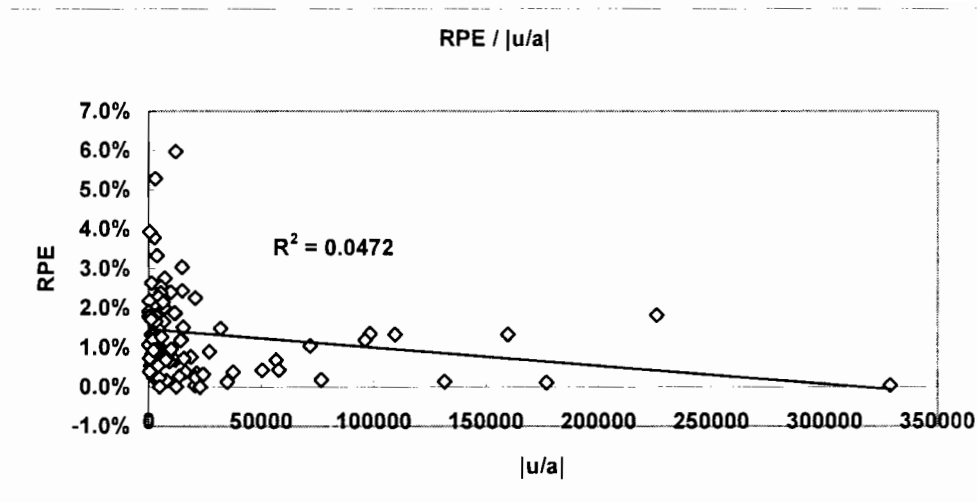


Figure 6-5 RPE Scatter over $|u/a|$

6.2.2 Prediction Error Estimated by Variance Propagation

So we have to seek other ways to estimate the prediction error. Let us recall what we described in Chapter 3, by variance propagation, the prediction error can be calculated by:

$$\sigma_{\hat{X}^{(i+1)}} = \left[\left(aiX^{(0)}(1) - X^{(0)}(1) - ui \right)^2 Q_{11} + Q_{22} + 2 \left(aiX^{(0)}(1) - X^{(0)}(1) - ui \right) Q_{12} \right]^{\frac{1}{2}} e^{-ai} \sigma_0$$

For an N -dim model, the prediction value $\hat{X}^{(0)}(N+1)$ will be in the range of

$$\hat{X}^{(0)}(N+1) \pm \sigma_{\hat{X}^{(0)}(N+1)}$$

Now for each model we tested, we can check if the observed price, for which we predicted, falls into the above range. Then we compute the percentage frequency of observed values in predicted range. This percentage is the confidence range of the prediction method.

As a matter of fact, for

$$e^{(0)}(N+1) = X^{(0)}(N+1) - \hat{X}^{(0)}(N+1)$$

if

$$\left| e^{(0)}(N+1) \right| \leq \left| \sigma_{\hat{X}^{(0)}(N+1)} \right|$$

the observed value falls into the estimation range and the prediction method is reliable.

Because the stock prices are not at the same level, in order to render all the models comparable when making chart, we divide the error $e^{(0)}(N+1)$ by $\hat{X}^{(0)}(N+1)$, and the relative error is reported. Table 6-6 and Table 6-7 list the statistical results.

University of Cape Town

Table 6-6 Prediction Error Estimation by Propagation (I)

	Week	$X^{(0)}(N+1)$	$\hat{X}^{(0)}(N+1)$	$ e^{(0)}(N+1) $	$\frac{ e^{(0)}(N+1) }{\hat{X}^{(0)}(N+1)}$	$ \sigma_{\hat{X}^{(0)}(N+1)} $	$\frac{ \sigma_{\hat{X}^{(0)}(N+1)} }{\hat{X}^{(0)}(N+1)}$
ABSA	1	58.50	59.039	0.539	0.913%	0.599	1.015%
	2	61.00	60.608	0.392	0.646%	1.131	1.867%
	3	62.90	63.522	0.622	0.979%	0.905	1.425%
	4	63.80	65.564	1.764	2.691%	0.675	1.030%
	5	64.00	63.725	0.275	0.432%	0.321	0.504%
	6	65.00	65.251	0.251	0.385%	0.431	0.660%
	7	64.55	65.226	0.676	1.036%	0.413	0.632%
	8	68.60	67.463	1.137	1.686%	0.256	0.380%
AECI	1	31.50	31.579	0.079	0.250%	0.183	0.579%
	2	31.95	31.352	0.598	1.908%	0.367	1.170%
	3	32.49	32.501	0.011	0.032%	0.301	0.926%
	4	33.30	32.502	0.798	2.454%	0.194	0.597%
	5	35.50	36.436	0.936	2.569%	0.270	0.741%
	6	34.00	33.383	0.617	1.847%	0.163	0.488%
	7	34.75	33.901	0.849	2.504%	0.238	0.701%
	8	35.10	35.525	0.425	1.197%	0.055	0.154%
ANGLAM	1	286.51	284.547	1.963	0.690%	2.593	0.911%
	2	282.99	286.811	3.821	1.332%	1.470	0.512%
	3	278.00	278.000	0.000	0.000%	0.000	0.000%
	4	262.16	264.170	2.010	0.761%	1.180	0.447%
	5	274.00	290.381	16.381	5.641%	2.993	1.031%
	6	254.00	246.271	7.729	3.138%	3.385	1.375%
	7	244.00	247.270	3.270	1.322%	3.243	1.312%
	8	233.99	230.446	3.544	1.538%	4.088	1.774%
BARLOW	1	75.35	75.252	0.098	0.130%	0.349	0.464%
	2	77.50	78.031	0.531	0.680%	0.188	0.241%
	3	77.70	76.780	0.920	1.198%	0.245	0.318%
	4	79.50	80.021	0.521	0.652%	0.857	1.071%
	5	79.70	80.467	0.767	0.953%	0.615	0.765%
	6	82.10	82.239	0.139	0.169%	0.155	0.188%
	7	82.50	83.726	1.226	1.464%	0.104	0.124%
	8	86.00	88.181	2.181	2.474%	1.923	2.181%
BHP	1	61.60	61.834	0.234	0.379%	0.510	0.825%
	2	64.30	66.736	2.436	3.650%	0.869	1.302%
	3	64.25	64.477	0.227	0.352%	0.249	0.386%
	4	67.95	69.585	1.635	2.349%	0.579	0.832%
	5	71.80	74.189	2.389	3.220%	0.602	0.812%
	6	65.50	65.652	0.152	0.231%	0.760	1.157%
	7	64.50	64.656	0.156	0.241%	0.317	0.491%
	8	63.69	60.327	3.363	5.575%	1.011	1.676%

Table 6-7 Prediction Error Estimation by Propagation (II)

	Week	$X^{(0)}(N+1)$	$\hat{X}^{(0)}(N+1)$	$ e^{(0)}(N+1) $	$\frac{ e^{(0)}(N+1) }{\hat{X}^{(0)}(N+1)}$	$ \sigma_{\hat{X}^{(0)}(N+1)} $	$\frac{ \sigma_{\hat{X}^{(0)}(N+1)} }{\hat{X}^{(0)}(N+1)}$
BIDVEST	1	60.00	61.127	1.127	1.843%	0.438	0.716%
	2	61.60	63.070	1.470	2.331%	0.263	0.417%
	3	62.10	62.370	0.270	0.433%	0.147	0.235%
	4	63.10	64.534	1.434	2.222%	0.165	0.256%
	5	63.40	63.406	0.006	0.009%	0.557	0.878%
	6	64.00	63.145	0.855	1.354%	0.088	0.139%
	7	65.90	67.100	1.200	1.789%	0.978	1.458%
	8	69.35	67.954	1.396	2.055%	0.653	0.961%
DIMSN	1	3.64	3.560	0.080	2.242%	0.054	1.507%
	2	3.78	3.852	0.072	1.874%	0.037	0.955%
	3	3.65	3.611	0.039	1.068%	0.041	1.127%
	4	3.73	3.675	0.055	1.500%	0.065	1.766%
	5	3.60	3.580	0.020	0.556%	0.011	0.301%
	6	3.45	3.410	0.040	1.174%	0.014	0.422%
	7	3.36	3.405	0.045	1.323%	0.018	0.538%
	8	3.35	3.291	0.059	1.800%	0.015	0.456%
DSICVRY	1	13.76	13.219	0.541	4.095%	0.115	0.874%
	2	13.97	13.875	0.095	0.681%	0.132	0.954%
	3	14.00	14.001	0.001	0.007%	0.237	1.696%
	4	15.25	15.310	0.060	0.391%	0.002	0.013%
	5	15.40	15.551	0.151	0.968%	0.075	0.482%
	6	15.00	14.874	0.126	0.850%	0.138	0.926%
	7	15.40	15.601	0.201	1.286%	0.160	1.027%
	8	17.00	17.125	0.125	0.731%	0.144	0.840%
FOSCHI	1	26.70	27.020	0.320	1.183%	0.199	0.736%
	2	27.40	26.780	0.620	2.313%	0.183	0.684%
	3	29.32	28.947	0.373	1.288%	0.210	0.726%
	4	28.30	28.038	0.262	0.935%	0.263	0.939%
	5	28.50	28.705	0.205	0.714%	0.076	0.264%
	6	27.30	27.153	0.147	0.542%	0.095	0.348%
	7	28.50	28.975	0.475	1.641%	0.391	1.348%
	8	30.15	30.668	0.518	1.688%	0.704	2.296%
AMLBEV	1	83.52	84.271	0.751	0.891%	0.343	0.407%
	2	84.99	84.756	0.234	0.276%	1.084	1.278%
	3	88.60	90.507	1.907	2.107%	1.896	2.095%
	4	88.75	88.625	0.125	0.141%	0.048	0.054%
	5	88.80	88.840	0.040	0.045%	0.019	0.022%
	6	89.10	89.260	0.160	0.180%	0.068	0.077%
	7	89.20	89.300	0.100	0.112%	0.059	0.066%
	8	90.20	90.493	0.293	0.324%	0.364	0.403%

Our calculation indicates the confidence level is 0.35, i.e. only 35% of the observed prices fall into the prediction range of $\hat{X}^{(0)}(N+1) \pm \sigma_{\hat{X}^{(0)}(N+1)}$, as showed in Figure 6-6. It is quite low and can not be used in practice to provide a reliable prediction range. In Figure 6-6, the relative error refers to the predicted $\hat{X}^{(0)}(N+1)$, but not the observed $X^{(0)}(N+1)$. We mirrored the observed errors to positive values and examine if they fall into the right half range of the prediction. We call it one sigma range in single side, where sigma means the error $\sigma_{\hat{X}^{(0)}(N+1)}$.

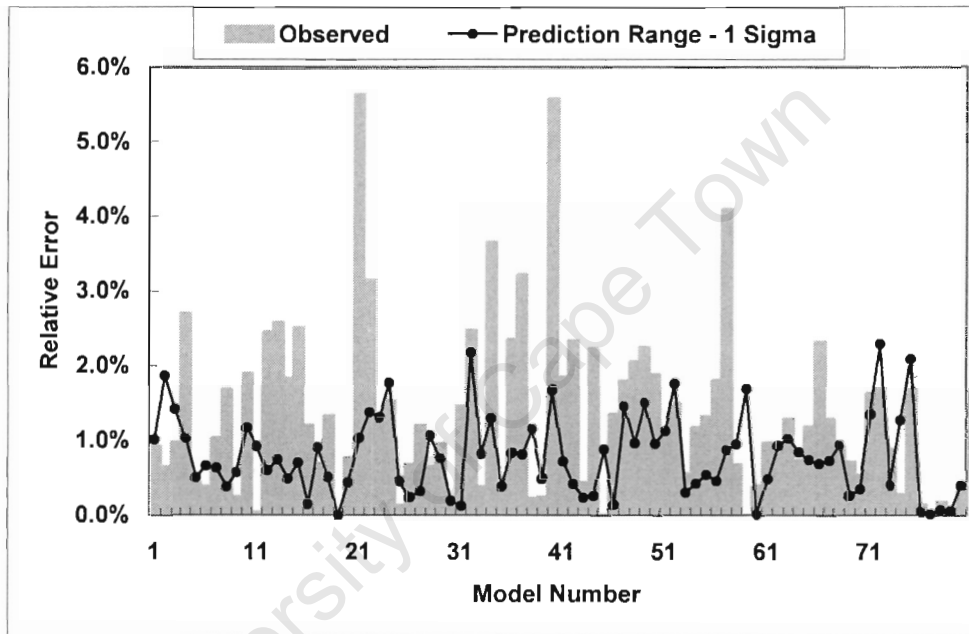


Figure 6-6 One Sigma Prediction Range

6.2.3 Improvement on Variance Propagation

We aware that when we enlarge the range centering at the predicted value $\hat{X}^{(0)}(N+1)$, there will be more observed values falling into it. We make the increment by using a factor m to multiply $\sigma_{\hat{X}^{(0)}(N+1)}$ and compute the percentage of the observed values falling into the range of $\hat{X}^{(0)}(N+1) \pm m\sigma_{\hat{X}^{(0)}(N+1)}$, where m is an integer. The statistics is listed in Table 6-8. In the range of $\left(\hat{X}^{(0)}(N+1) - 8\sigma_{\hat{X}^{(0)}(N+1)}, \hat{X}^{(0)}(N+1) + 8\sigma_{\hat{X}^{(0)}(N+1)}\right)$, the confidence level is 95%, and in the range of $\left(\hat{X}^{(0)}(N+1) - 10\sigma_{\hat{X}^{(0)}(N+1)}, \hat{X}^{(0)}(N+1) + 10\sigma_{\hat{X}^{(0)}(N+1)}\right)$, 97.5%.

Table 6-8 Relation of Confidence Level and Prediction Range

Confidence	35.00%	57.50%	77.50%	87.50%	91.25%
Factor m	1	2	3	4	5
Confidence	93.75%	93.75%	95.00%	96.25%	97.50%
Factor m	6	7	8	9	10

As showed in Figure 6-7, the confidence level changes with the prediction range rapidly when $m < 4$, and slowly from there on. At $m = 4$ and $m = 5$, the confidence level is 87.5% and 91.25% respectively. If we take the 90% confidence, then $m = 5$ is enough. Or we can take the multiple of 8 or 10, to get the confidence level of 95% or 97.5%.

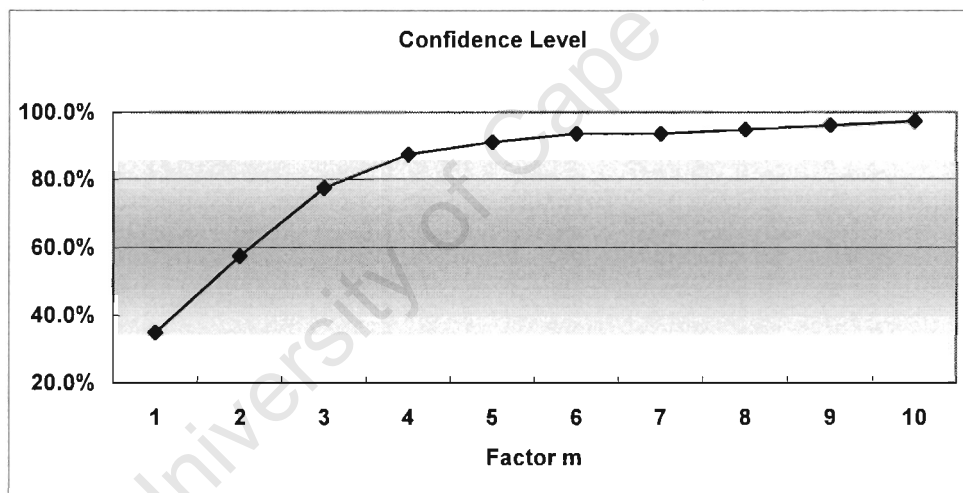


Figure 6-7 Confidence Level Changing with Prediction Range

In Figure 6-8, we show the prediction ranges corresponding to the confidence level of 90% and 95%.

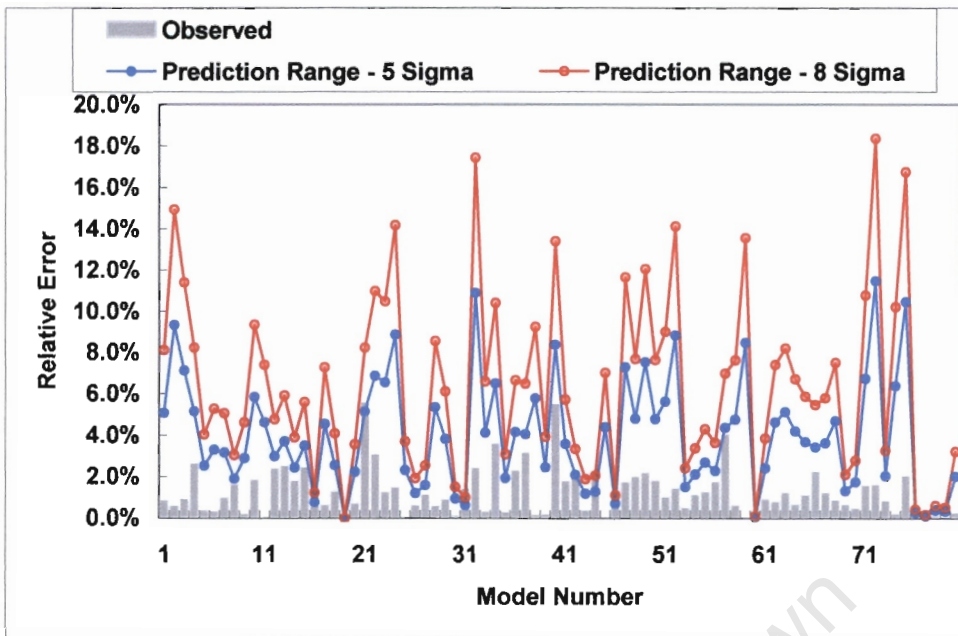


Figure 6-8 Five and Eight Sigma Prediction Ranges

Now we change $\hat{X}^{(0)}(N+1) \pm m\sigma_{\hat{X}^{(0)}(N+1)}$ to the form of

$$\hat{X}^{(0)}(N+1) \pm m\hat{X}^{(0)}(N+1) \frac{\sigma_{\hat{X}^{(0)}(N+1)}}{\hat{X}^{(0)}(N+1)}$$

From Table 6-6 and Table 6-7 and by computation, the average of $\frac{\sigma_{\hat{X}^{(0)}(N+1)}}{\hat{X}^{(0)}(N+1)}$ for the 80 models is 0.797%. Substitute this value, we get

$$\hat{X}^{(0)}(N+1) \pm 0.797\%m\hat{X}^{(0)}(N+1)$$

That is the prediction range for the confidence level. For GM(1, 1), the prediction range is $\hat{X}^{(0)}(N+1) \pm 3.985\%\hat{X}^{(0)}(N+1)$ at 90% confidence, $\hat{X}^{(0)}(N+1) \pm 6.376\%\hat{X}^{(0)}(N+1)$ at 95%, and $\hat{X}^{(0)}(N+1) \pm 7.970\%\hat{X}^{(0)}(N+1)$ at 97.5%, as listed in Table 6-9.

Table 6-9 Prediction Range of GM(1, 1) at Different Confidence Levels

•• Confidence Level	•• Factor m	Prediction Range
90%	5	•• $\hat{X}^{(0)}(N+1) \pm 3.985\% \hat{X}^{(0)}(N+1)$
95%	8	•• $\hat{X}^{(0)}(N+1) \pm 6.376\% \hat{X}^{(0)}(N+1)$
97.5%	10	•• $\hat{X}^{(0)}(N+1) \pm 7.970\% \hat{X}^{(0)}(N+1)$

6.3 Limitations of Grey Modeling

It seems the feature of the grey system theory that it emphasizes multiple possibilities in constructing and fitting the grey models. The *grey* represents unclearness and uncertainty. The *whitened grey*, the solution to a problem, is fully based on the information acquired. The model structure could change observably as the new information is added into the system. So the grey model is unstable to new information. That is the reason the predictions of more than one step have larger errors. For the second step prediction value, the observed datum at the first step point has not been associated in the modeling. Without the new information, the real relation between the actual value at the second prediction point and its past data are weakened, and hence the prediction might lose its efficiency. As we are seeking the accurate prediction for stock prices, we exploit one step prediction.

Deng, the creator of the grey system theory and his successors have incorporated much ancient Chinese philosophic thoughts into the theory. They use dialectic idea to explain the things and attribute the problems to the four states of embryo, growth, maturity and exact proof. But there is less strong mathematical logic in the theory. It emphasizes the practicality and means to give out the solutions. So it seems more like a methodology than a theory. It is far from perfect in the aspect of theory.

Until now, the grey modeling is mainly limited to positive series and lower order differential equation, mostly one order as in the popular GM(1, 1). The extension to negative series is

primarily a complementary to revise the fitting on the original by residuals, but seldom implemented separately. Comparing to GM(1, 1), the GM(2, 1) is much more complicated in the fitting. No reasonable application of GM(2, 1), especially in financial market, has been found in the literature. But the second order differential equation has a richer form solution with two exponential terms. So the response function for GM(2, 1) is much more flexible than that of GM(1, 1). The ability of GM(2, 1) tracing to the trend has been showed by our investigation. The problem is how to increase the reliability. That is beyond the coverage of this paper.

Further studies are needed on error evaluation. The fitting of GM is evaluated by ARE, the average relative error of each modeling datum, and the Model Precision Grade. The prediction is checked by RPE, the relative prediction error. The MP Grade does not make any sense and totally of no use. The ARE has a poor relation to the RPE and loses the indicative efficiency for prediction. The MEGM(1, 1) we proposed effectively improved the fitting and lowered the ARE, but to the opposite, the RPE was not reduced, but increased. It is the error indicators that have put the improvement model in a dilemma. There must be the reliable error indicators at first, then the reasonable improvement may follow. The improvement on variance propagation is quite rough, and hence a better methodology is expected from the people working in the field related to grey system theory.

Appendix A Data

1. Monthly data

Data source: Department of Statistical Sciences, UCT;

Data period: Jul 1988 – Dec 2004;

Stock labels:

1. ANGLOS
2. JDGROU
3. PICPAY
4. REMGR
5. SAEAGL
6. SAPPI
7. SASOL
8. TIGBRA
9. TONGAT.

2. Daily data

Data source: Datastream, UCT Library;

Data period: Sep 6, 2004 – Oct 29 2004 (Nov 1 used for checking), 8 weeks;

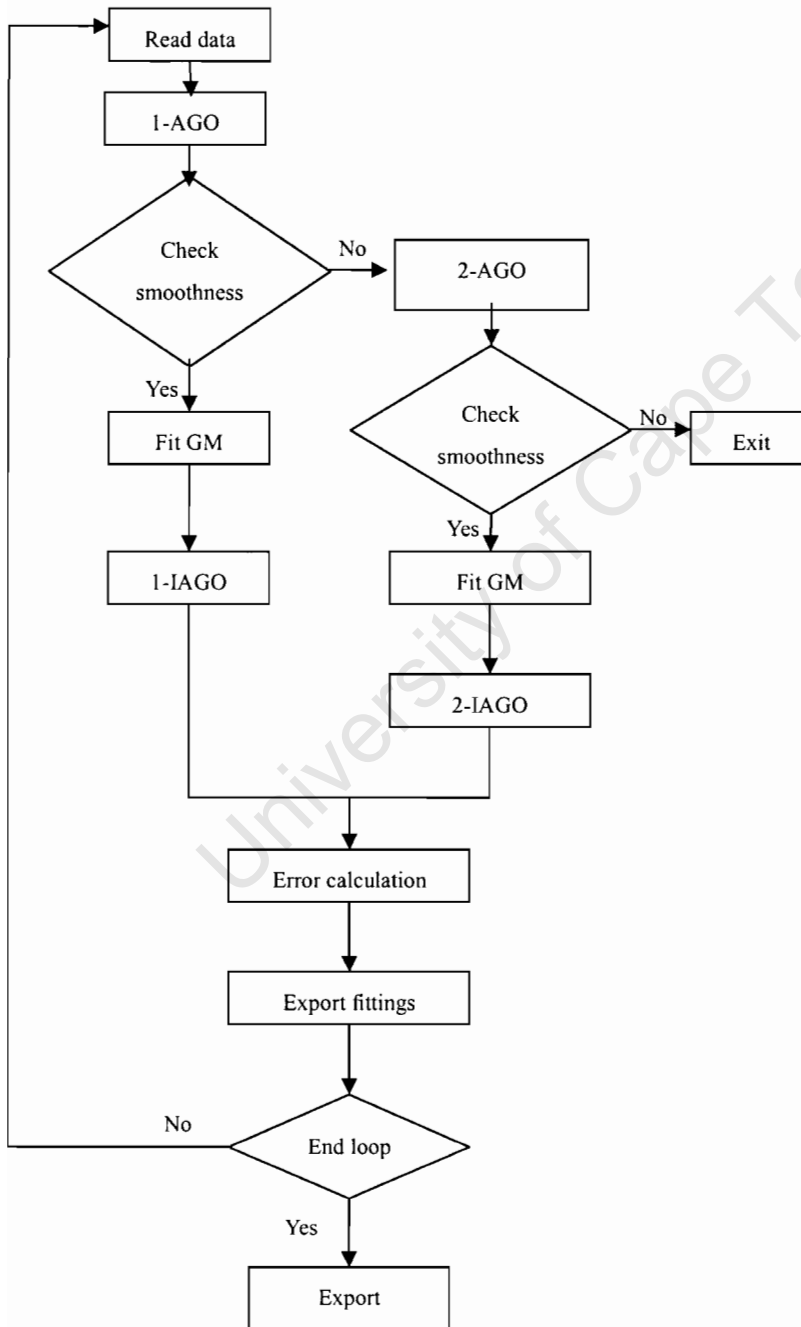
Stock labels:

1. ABSA
2. AECI
3. ANGLAM
4. BARLOW
5. BHP
6. BIDVEST
7. DIMSN
8. DISCVRY
9. FOSCHI
10. AMLBEV

Appendix B VBA Codes

We use MS Excel for the implementation of the models. The program is written with VBA. We list the codes for GM(1, 1) and GM(2, 1).

1. Flow Chart of the Program



2. Codes for GM(1, 1)

Option Base 1

'define sub

Sub GM1_1()

'define N as the model dimension, M as the steps moving forward

Dim N, M As Integer

N = Range("A4")

M = Range("B4")

'X0 is the original series, ie, the stock price

'X1 is for 1-AGO series, and X2, 2-AGO

'X0hat, X1hat and X2hat are the fitting value for X0, X1 and X2 respectively

'B and Y are matrix as in the theory description

'a_hat is the fitting coefficient vector consisting of a and u

'as we only make one step prediction, the X(.)hat variable is defined one dimension higher

Dim X0(), X0hat(), X1(), X1hat(), X2(), X2hat, B(), Y, a_hat(), a, u

ReDim X0(N), X0hat(N + 1), X1(N), X1hat(N + 1), X2(N), X2hat(N + 1), B(N - 1, 2), Y(N - 1, 1)

'j stands the steps moving forward for predicting, totally M steps.

'define AveMdlErr as average, MaxMdlErr as maximum, and MinMdlErr as minimum _

' model fitting error for each stock data

'similarly, AveForeErr, MaxForeErr, MinForeErr are for the prediction

Dim AveMdlErr, MaxMdlErr, MinMdlErr, AveForeErr, MaxForeErr, MinForeErr

AveMdlErr = 0

MaxMdlErr = 0

MinMdlErr = 1

AveForeErr = 0

MaxForeErr = 0

MinForeErr = 1

'this is just to keep the worksheet clean when we try different steps or data

Range("C11:Y2000") = ClearContents

Range("C11:Y2000").Font.Color = Default

'start the iteration, M rounds there

For j = 1 To M * N Step N

'read original data which is arranged in column beginning from "B11"

For i = 1 To N

X0(i) = Range("B" & i + j + 9)

Next i

'1-AGO, calculate X1.

X1(1) = X0(1)

For i = 2 To N

X1(i) = X1(i - 1) + X0(i)

Next i

'start to check the smoothness

'define Smth0(i)=X0(i)/X1(i-1) and Smth1=X1(i)/X2(i-1), for i>=3

Dim Smth0(), Smth1()

ReDim Smth0(N), Smth1(N)

'Smth0 calculation

For i = 3 To N

Smth0(i) = X0(i) / X1(i - 1)

If Smth0(i) > 1 Then GoTo Calc_Smth1 Else

Next i

For i = 3 To N - 1

If Smth0(i) < Smth0(i + 1) Then GoTo Calc_Smth1 Else

Next i

'if X0 is smooth, then continue to fit the GM built on X0

GoTo AGO1

'if X0 is not smooth, then check X1

```

Calc_Smth1:
    X2(1) = X1(1)
    For i = 2 To N
        X2(i) = X2(i - 1) + X1(i)
    Next i
    For i = 3 To N
        Smth1(i) = X1(i) / X2(i - 1)
        If Smth1(i) > 1 Then GoTo NonSmoothInfo Else
    Next i
    For i = 3 To N - 1
        If Smth1(i) < Smth1(i + 1) Then GoTo NonSmoothInfo Else
    Next i

'if X1 is smooth, build GM on it and calculate the fitting
GoTo AGO2

'if neither X0 or X1 is smooth, exit sub and show info
NonSmoothInfo:

'as we are rolling forward, "j" tells us where the codes stop
msg = "Both X0 and X1 are not smooth for step " & j & ", try data processing methods other than AGO."
msgbox msg, vbExclamation
Exit Sub

'calculate the fitting coefficients of the GM on X1
AGO1:

'calculate B and Y, which are 1 less in dimension than X0 and X1
For i = 1 To N - 1
    B(i, 1) = -0.5 * (X1(i) + X1(i + 1))
    B(i, 2) = 1
    Y(i, 1) = X0(i + 1)
Next i

'estimation of the coefficients, by  $\hat{a} = (B'B)^{-1}B'Y$ 
a_hat = Application.MMult(Application.MMult(Application.MInverse_(Application.MMult(Application.Transpose(B), B)),
Application.Transpose(B)), Y)
a = a_hat(1, 1)
u = a_hat(2, 1)

'get X1hat, fitting values of X1, from the response function
For i = 0 To N
    X1hat(i + 1) = (X0(1) - u / a) * Exp(-a * i) + u / a
Next i

'1-AGO, get X0hat
X0hat(1) = X1hat(1)
For i = 1 To N
    X0hat(i + 1) = X1hat(i + 1) - X1hat(i)
Next i

'go to the next step
GoTo StepFwd

'this is estimation of GM on X2, similar to AGO1
AGO2:

For i = 1 To N - 1
    B(i, 1) = -0.5 * (X2(i) + X2(i + 1))
    B(i, 2) = 1
    Y(i, 1) = X1(i + 1)
Next i

a_hat = Application.MMult(Application.MMult(Application.MInverse_(Application.MMult(Application.Transpose(B), B)),
Application.Transpose(B)), Y)
a = a_hat(1, 1)
u = a_hat(2, 1)

For i = 0 To N
    X2hat(i + 1) = (X0(1) - u / a) * Exp(-a * i) + u / a

```

```

Next i

X1hat(1) = X2hat(1)
X0hat(1) = X1hat(1)
For i = 1 To N
    X1hat(i + 1) = X2hat(i + 1) - X2hat(i)
    X0hat(i + 1) = X1hat(i + 1) - X1hat(i)
Next i

'start to calculate the errors
StepFwd:

'define RelErr as relative error of each fitted value
'AveAbsRelErr as the absolute relative error of the model
'ForeErr as the prediction error
Dim RelErr(), AveAbsRelErr, ForeErr
ReDim RelErr(N)
AveAbsRelErr = 0

For i = 1 To N
    RelErr(i) = (X0hat(i) - X0(i)) / X0(i)
    AveAbsRelErr = AveAbsRelErr + Abs(RelErr(i))
Next i

AveAbsRelErr = AveAbsRelErr / N

'read the price we predicted and compute the prediction error
ReDim Preserve X0(N + 1)
X0(N + 1) = Range("B" & N + j + 10)

ForeErr = (X0hat(N + 1) - X0(N + 1)) / X0(N + 1)

'export to worksheet, mark the prediction and its error in red
For i = 1 To N
    Range("C" & i + j + 9) = X0hat(i)
    Range("D" & i + j + 9) = RelErr(i)
Next i

Range("H" & j + 12) = X0hat(N + 1)
Range("H" & j + 12).Font.Color = vbRed
Range("H" & j + 13) = ForeErr
Range("H" & j + 13).Font.Color = vbRed
Range("E" & j + 10) = AveAbsRelErr
Range("F" & j + 10) = a_hat(1, 1)
Range("G" & j + 10) = a_hat(2, 1)

'compute statistics on the errors
AveMdlErr = AveMdlErr + AveAbsRelErr
MaxMdlErr = Application.Max(MaxMdlErr, AveAbsRelErr)
MinMdlErr = Application.Min(MinMdlErr, AveAbsRelErr)
AveForeErr = AveForeErr + Abs(ForeErr)
MaxForeErr = Application.Max(MaxForeErr, Abs(ForeErr))
MinForeErr = Application.Min(MinForeErr, Abs(ForeErr))

'prepare to compute the model precision grade
'define Resid as residual, VarX0 as variance of X0, VarResid as variance of Residuals
'C is posterior ratio, P is small error probability
'PreciC and PreciP is their grade (classified as 1,2,3 or 4)
'Precision is the model precision
Dim Resid(), VarX0, VarResid, C, P, PreciC, PreciP, Precision
ReDim Resid(N)

'compute each residual value
For i = 1 To N
    Resid(i) = X0(i) - X0hat(i)
Next i

'compute variances of prices and residuals
VarX0 = Application.VarP(X0)

```

```

VarResid = Application.VarP(Resid)

'compute C and P, and determine the precision
C = Sqr(VarResid / VarX0)
P = 0
For i = 1 To N
    If Abs(Resid(i) - Application.Average(Resid)) < 0.6745 * Sqr(VarX0) Then
        P = P + 1
    End If
Next i
P = P / N

If C <= 0.35 Then
    PreciC = 1
Elseif C > 0.35 And C <= 0.5 Then
    PreciC = 2
Elseif C > 0.5 And C <= 0.65 Then
    PreciC = 3
Elseif C > 0.65 Then
    PreciC = 4
End If

If P >= 0.95 Then
    PreciP = 1
Elseif P >= 0.8 And P < 0.95 Then
    PreciP = 2
Elseif P >= 0.7 And P < 0.8 Then
    PreciP = 3
Elseif P < 0.7 Then
    PreciP = 4
End If
Precision = Application.Max(PreciC, PreciP)

'export to worksheet
Range("H" & j + 10) = Precision
If Precision = 1 Then
    Range("H" & j + 11) = "Good"
Elseif Precision = 2 Then
    Range("H" & j + 11) = "Satisfactory"
Elseif Precision = 3 Then
    Range("H" & j + 11) = "Acceptable"
Elseif Precision = 4 Then
    Range("H" & j + 11) = "Unacceptable"
    Range("H" & j + 11).Font.Color = vbRed
End If

'move to next step,ie, next model
Next j

'compute teh averages for the stock, consisting of M models
AveMdlErr = AveMdlErr / M
AveForeErr = AveForeErr / M

'export
Range("A7") = AveMdlErr
Range("B7") = MaxMdlErr
Range("C7") = MinMdlErr
Range("D7") = AveForeErr
Range("E7") = MaxForeErr
Range("F7") = MinForeErr
End Sub

```

3. Codes for GM(2, 1)

```

Option Base 1
Sub GM21_1()
    Dim N, M As Integer
    'N is the number of data used for forecasting, M is the number of rounds moving forward
    N = Range("A4")
    M = Range("B4")

    Dim X0(), X0hat(), X1(), X1hat(), X2(), X2hat, B(), Y, a_hat(), a1, a2, u
    ReDim X0(N + 1), X0hat(N + 1), X1(N), X1hat(N + 1), X2(N), X2hat(N + 1), B(N - 1, 3), Y(N - 1, 1)

    'j stands the steps moving forward for predicting, totally M steps. this is the outmost layer loop.
    Dim AveMdlErr, MaxMdlErr, MinMdlErr, AveForeErr, MaxForeErr, MinForeErr
    AveMdlErr = 0
    MaxMdlErr = 0
    MinMdlErr = 1
    AveForeErr = 0
    MaxForeErr = 0
    MinForeErr = 1

    Range("C11:Y2000") = ClearContents
    Range("C11:Y2000").Font.Color = Default
    For j = 1 To M * 6 Step N + 1
        'read original data,beginning from "B11". Add a random number between -0.001 and 0.001 to ensure the solution
        For i = 1 To N + 1
            X0(i) = Range("B" & i + j + 9) + 0.002 * Rnd() - 0.001
        Next i

        '1-AGO, getting X1.
        X1(1) = X0(1)
        For i = 2 To N
            X1(i) = X1(i - 1) + X0(i)
        Next i

        'define Smth0(i)=X0(i)/X1(i-1) and Smth1=X1(i)/X2(i-1), for i>=3
        Dim Smth0(), Smth1()
        ReDim Smth0(N), Smth1(N)

        Dim Delta, Eigen1, Eigen2, C1, C2 As Double
        'Smth0 calculation
        For i = 3 To N
            Smth0(i) = X0(i) / X1(i - 1)
            If Smth0(i) > 1 Then GoTo Calc_Smth1 Else
        Next i
        For i = 3 To N - 1
            If Smth0(i) < Smth0(i + 1) Then GoTo Calc_Smth1 Else
        Next i

        GoTo AGO1

    Calc_Smth1:
        X2(1) = X1(1)
        For i = 2 To N
            X2(i) = X2(i - 1) + X1(i)
        Next i
        For i = 3 To N
            Smth1(i) = X1(i) / X2(i - 1)
            If Smth1(i) > 1 Then GoTo NonSmoothInfo Else
        Next i
        For i = 3 To N - 1
            If Smth1(i) < Smth1(i + 1) Then GoTo NonSmoothInfo Else
        Next i

        GoTo AGO2
    
```

```

NonSmoothInfo:
'j stands the steps moving forward
  msg = "Both X0 and X1 are not smooth for step " & j & ", try data processing methods other than AGO."
  msgbox msg, vbExclamation
  Exit Sub

'start the fitting calculation
AGO1:
'calculate to get B and Y, which are 1 less in dimension than X0 and X1
  For i = 1 To N - 1
    B(i, 1) = -X0(i + 1)
    B(i, 2) = -0.5 * (X1(i) + X1(i + 1))
    B(i, 3) = 1
    Y(i, 1) = X0(i + 1) - X0(i)
  Next i

'Estimation of the coefficients, by  $\hat{a} = (B'B)^{-1}B'Y$ 
  a_hat = Application.MMult(Application.MMult(Application.MInverse(Application.MMult _
(Application.Transpose(B), B)), Application.Transpose(B)), Y)
  a1 = a_hat(1, 1)
  a2 = a_hat(2, 1)
  u = a_hat(3, 1)

'Eigen1 and Eigen2 are the two eigen values
  Delta = a1 ^ 2 - 4 * a2
  If Delta > 0 Then
    Eigen1 = -0.5 * (a1 - Sqr(Delta))
    Eigen2 = -0.5 * (a1 + Sqr(Delta))
    GoTo Solution_1
  ElseIf Delta < 0 Then
    GoTo Solution_2
  Else:
    Eigen1 = -a1 / 2
    GoTo Solution_3
  End If

'there are three cases for the solution of the second order differential equation
'we solve for the constant C1 and C2 for each case
Solution_1:
  C1 = ((X1(1) - u / a2) * Eigen2 * Exp(Eigen2) - X0(2)) / (Eigen2 * Exp(Eigen2) - Eigen1 * Exp(Eigen1))
  C2 = ((X1(1) - u / a2) * Eigen1 * Exp(Eigen1) - X0(2)) / (Eigen1 * Exp(Eigen1) - Eigen2 * Exp(Eigen2))
  For i = 0 To N
    X1hat(i + 1) = C1 * Exp(Eigen1 * i) + C2 * Exp(Eigen2 * i) + u / a2
  Next i
  GoTo IAGO1

Solution_2:
  C1 = X1(1) - u / a2
  C2 = (2 * X0(2) + C1 * Exp(-a1 / 2) * (Sqr(-Delta) * Sin(Sqr(-Delta) / 2) + _
a1 * Cos(Sqr(-Delta) / 2))) / Exp(-a1 / 2) * (Sqr(-Delta) * Cos(Sqr(-Delta) / 2) - a1 * Sin(Sqr(-Delta) / 2))
  For i = 0 To N
    X1hat(i + 1) = C1 * Exp(-a1 * i / 2) * Cos(Sqr(-Delta) * i / 2) + _
    C2 * Exp(-a1 * i / 2) * Sin(Sqr(-Delta) * i / 2) + u / a2
  Next i
  GoTo IAGO1

Solution_3:
  C1 = X1(1) - u / a2
  C2 = (X0(2) - Eigen1 * Exp(Eigen1) * C1) / (Exp(Eigen2) * (Eigen2 + 1))
  For i = 0 To N
    X1hat(i + 1) = C1 * Exp(Eigen1 * i) + C2 * i * Exp(Eigen1 * i) + u / a2
  Next i

IAGO1:
  X0hat(1) = X1hat(1)
  For i = 1 To N
    X0hat(i + 1) = X1hat(i + 1) - X1hat(i)
  Next i

```

GoTo StepFwd

AGO2:

'calculate to get B and Y, which are 1 less in dimension than X1 and X2

For i = 1 To N - 1

B(i, 1) = -X1(i + 1)

B(i, 2) = -0.5 * (X2(i) + X2(i + 1))

B(i, 3) = 1

Y(i, 1) = X0(i + 1)

Next i

a_hat = Application.MMult(Application.MMult(Application.MInverse _
(Application.MMult(Application.Transpose(B), B)), Application.Transpose(B)), Y)

a1 = a_hat(1, 1)

a2 = a_hat(2, 1)

u = a_hat(3, 1)

Delta = a1 ^ 2 - 4 * a2

If Delta > 0 Then

Eigen1 = -0.5 * (a1 - Sqr(Delta))

Eigen2 = -0.5 * (a1 + Sqr(Delta))

GoTo Solution_4

Elsif Delta < 0 Then

GoTo Solution_5

Else:

Eigen1 = -a1 / 2

GoTo Solution_6

End If

Solution_4:

C1 = ((X1(1) - u / a2) * Eigen2 * Exp(Eigen2) - X1(2)) / (Eigen2 * Exp(Eigen2) - Eigen1 * Exp(Eigen1))

C2 = ((X1(1) - u / a2) * Eigen1 * Exp(Eigen1) - X1(2)) / (Eigen1 * Exp(Eigen1) - Eigen2 * Exp(Eigen2))

For i = 0 To N

X2hat(i + 1) = C1 * Exp(Eigen1 * i) + C2 * Exp(Eigen2 * i) + u / a2

Next i

GoTo IAGO2

Solution_5:

C1 = X1(1) - u / a2

C2 = (2 * X1(2) + C1 * Exp(-a1 / 2) * (Sqr(-Delta) * Sin(Sqr(-Delta) / 2) + _
a1 * Cos(Sqr(-Delta) / 2)) / Exp(-a1 / 2) * (Sqr(-Delta) * Cos(Sqr(-Delta) / 2) - a1 * Sin(Sqr(-Delta) / 2))

For i = 0 To N

X2hat(i + 1) = C1 * Exp(-a1 * i / 2) * Cos(Sqr(-Delta) * i / 2) + _

C2 * Exp(-a1 * i / 2) * Sin(Sqr(-Delta) * i / 2) + u / a2

Next i

GoTo IAGO2

Solution_6:

C1 = X1(1) - u / a2

C2 = (X1(2) - Eigen1 * Exp(Eigen1) * C1) / (Exp(Eigen2) * (Eigen2 + 1))

For i = 0 To N

X2hat(i + 1) = C1 * Exp(Eigen1 * i) + C2 * i * Exp(Eigen1 * i) + u / a2

Next i

IAGO2:

X1hat(1) = X2hat(1)

X0hat(1) = X1hat(1)

For i = 1 To N

X1hat(i + 1) = X2hat(i + 1) - X2hat(i)

X0hat(i + 1) = X1hat(i + 1) - X1hat(i)

Next i

StepFwd:

Dim RelErr(), AveAbsRelErr, ForeErr

ReDim RelErr(N)

AveAbsRelErr = 0

```

For i = 1 To N
    RelErr(i) = (X0hat(i) - X0(i)) / X0(i)
    AveAbsRelErr = AveAbsRelErr + Abs(RelErr(i))
Next i
AveAbsRelErr = AveAbsRelErr / N
ForeErr = (X0hat(N + 1) - X0(N + 1)) / X0(N + 1)

For i = 1 To N
    Range("C" & i + j + 9) = X0hat(i)
    Range("D" & i + j + 9) = RelErr(i)
Next i
Range("C" & N + j + 10) = X0hat(N + 1)
Range("C" & N + j + 10).Font.Color = vbRed
Range("D" & N + j + 10) = ForeErr
Range("D" & N + j + 10).Font.Color = vbRed
Range("E" & j + 10) = AveAbsRelErr
Range("F" & j + 10) = a_hat(1, 1)
Range("G" & j + 10) = a_hat(2, 1)
Range("H" & j + 10) = a_hat(3, 1)
Range("K" & j + 10) = AveAbsRelErr - Sheets("GM_(1_1)").Range("E" & j + 10)

AveMdlErr = AveMdlErr + AveAbsRelErr
MaxMdlErr = Application.Max(MaxMdlErr, AveAbsRelErr)
MinMdlErr = Application.Min(MinMdlErr, AveAbsRelErr)
AveForeErr = AveForeErr + Abs(ForeErr)
MaxForeErr = Application.Max(MaxForeErr, Abs(ForeErr))
MinForeErr = Application.Min(MinForeErr, Abs(ForeErr))

Dim Resid(), VarX0, VarResid, C, P, PreciC, PreciP, Precision
ReDim Resid(N)
For i = 1 To N
    Resid(i) = X0(i) - X0hat(i)
Next i

VarX0 = Application.VarP(X0)
VarResid = Application.VarP(Resid)

C = Sqr(VarResid / VarX0)
P = 0
For i = 1 To N
    If Abs(Resid(i) - Application.Average(Resid)) < 0.6745 * Sqr(VarX0) Then
        P = P + 1
    End If
Next i
P = P / N

If C <= 0.35 Then
    PreciC = 1
Elseif C > 0.35 And C <= 0.5 Then
    PreciC = 2
Elseif C > 0.5 And C <= 0.65 Then
    PreciC = 3
Elseif C > 0.65 Then
    PreciC = 4
End If

If P >= 0.95 Then
    PreciP = 1
Elseif P >= 0.8 And P < 0.95 Then
    PreciP = 2
Elseif P >= 0.7 And P < 0.8 Then
    PreciP = 3
Elseif P < 0.7 Then
    PreciP = 4
End If

Precision = Application.Max(PreciC, PreciP)
Range("I" & j + 10) = Precision
If Precision = 1 Then

```

```
    Range("I" & j + 11) = "Good"
Elseif Precision = 2 Then
    Range("I" & j + 11) = "Satisfactory"
Elseif Precision = 3 Then
    Range("I" & j + 11) = "Acceptable"
Elseif Precision = 4 Then
    Range("I" & j + 11) = "Unacceptable"
    Range("I" & j + 11).Font.Color = vbRed
End If
Next j

AveMdlErr = AveMdlErr / M
AveForeErr = AveForeErr / M

Range("A7") = AveMdlErr
Range("B7") = MaxMdlErr
Range("C7") = MinMdlErr
Range("D7") = AveForeErr
Range("E7") = MaxForeErr
Range("F7") = MinForeErr
```

End Sub

University of Cape Town

References

- 1) Bauer, L., L. Garriss, J. M. McCune, M. K. Reiter, J. Rouse and P. Rutenbar. 2005. *Device-Enabled Authorization in the Grey System*. School of Computer Science, Carnegie Mellon University.
- 2) Chen, H. and D. Li. 2003. *The Application of Grey Forecast Model in Stock Price*. Science Research Management. Vol 24, No 2.
- 3) Chen, J-M. 2003. *An application of Grey System Theory to Noise Prediction to Local Roads in Taiwan*. Master's Dissertation, Department of Environment Engineering and Science, Feng Chia University, Taiwan.
- 4) Cong, C. and X. Ji. 2000. *The Application of Grey Prediction in Forecasting the Stock Price Index*. China Statistics, No 5.
- 5) Deng, J. 1985. *The Grey System, Society and Economy*. Press of National Defence Industry.
- 6) Du, S., Z. Hou, C. Jiang and R. Yang. 2002. *Design On Spot Price Forecasting and Mutation Alarm System In Electricity Market*. Power Systems and Communications Infrastructures for the Future, Beijing, China.
- 7) Evans, L. 2003. *Why the Bubble Burst*. Edward Elgar.
- 8) Fan, C., Z. Jin and W. Tian. 2003. *A Novel Hybrid Grey-Based Strategy for Improving the Model Precision of a Dynamically Tuned Gyroscope*. Measurement Science and Technology, Vol 14.
- 9) Feng, L. 2003. *On Modern Chinese Fan-Shu*. <http://www.aideas.com/panshu.pdf>.
- 10) Feng, Z., Z. Xu, L. Wang and W. Yan. 2004. *Grey Prediction of Embankment Settlement and Its Application*. Journal of Northern Jiaotong University, Vol 28, No 4.
- 11) Fu, L. 1992. *Grey System Theory and Its Applications*. The Science and Technology Literature Press, Beijing, China.
- 12) Gao, F. 1993. *Grey Prediction on Stocks*. Shandong People's Press, Jinan, China.
- 13) Graham, B., and D. Dodd. 1934. *Security Analysis*. McGraw-Hill.
- 14) Guo, Y-H, Z-L Lin, X-J Horng and F-L Hsu. 2005. *An Automatic Document Classification Technique Based on Grey Relation Analysis*. Institute of Electrical Engineering, National Taiwan University of Technology.
- 15) Han, R., L. Li and H. Fu. 2000. *Application of Grey Theory in Unequal Time-Span Forecast of Ground Settlement*. Exploration Engineering, No 2.
- 16) Hong, T-P., C-L Wang and C-H Yeh. 2000. *Analysis of Financial Time Series Volatility Using Grey Theory*. Seminar on Grey System Theory and Its Applications, Taiwan.
- 17) Hou G-L. 2004. *Grey Prediction Control for Random Perturbation Servo System*. <http://www.cna.edu.tw/~sas/ccna/article/dec07.pdf>.
- 18) Hsieh, C-S. 2004. *Applying Grey Prediction Theory in the Market Development Trend Forecasting for Cosmetic Industry in Taiwan*. Master's Thesis, Dept. of Business Administration, Chaoyang University of Technology, Taiwan.
- 19) Li, C. 2002. *A Study on Optimum Combination for the Intercrops of Chinese Prickly Ash Tree*. Journal of Southwest China Normal University (Natural Science). Vol 27, No 6.
- 20) Li, P. 1997. *Grey Prediction on Stock Prices*. Eastern China Economic Management, Vol 4.

(Gen Vol 52).

- 21) Li, T., W. Wang, Y. Xu and Y. Chang. 2002. *Grey Clustering Decision-Making Evaluation on City Street Traffic Atmosphere Pollution*. Better Air Quality in Asian and Pacific Rim Cities, Hong Kong.
- 22) Lin, C-T and H-Y Tsai. 2005. *Hierarchical Clustering Analysis Based on Grey Relation Grade*. Information and Management Sciences, Vol 16, No 1.
- 23) Lin, S. 2004. *Evaluation of Business Reputation in Information Service Industry – An analysis of Grey Relational Analysis*. Journal of Information, Technology and Society, No 2.
- 24) Liu, S., T. Guo and Y. Dang. 1999. *Grey System Theory and Its Applications*. The Scientific Press.
- 25) Mu, H., Y. Kondou, W. Zhou, Y. Ning, Y. Tonooka and Y. Sakamoto. 2005. *Development of Model on Energy Consumption and Emission of Air Pollutants and GHGs in China*. Institute of Energy Utilization, AIST, Tsukuba, Japan.
- 26) Nie, R. and B. Li. 2003. *Grey Modeling on Stock Investment*. The Economist.
- 27) Ren, G. 2002. *One Realization of Grey Model Forecasting Based on Delphi Programming*. Information on Microcomputer, Vol 18, No 3.
- 28) Shefrin, H. 2002. *Beyond Greed and Fear*. Oxford University Press.
- 29) Shi, L., Q. Dong and S. Wang. 2004. *Residue of Streptomycin in Berry and the Effect of Expanding Reagents on Berry Quality in 'Muscat Hamburg' Grape*. Chinese Bulletin of Botany, 21.
- 30) Su, Y. and Y. Zhao. 2004. *Method for Quantitative Determination of the Protected Extent of Wetland*. Journal of Water Resources.
- 31) Tsai, C-H. 2003. *Predicting the Effluent Quality from Biological Wastewater Treatment Processes Using ASM2d, ANN and Grey Theory*. Department of Environmental Engineering and Management, Chaoyang University of Technology, Taiwan.
- 32) Wen, K-L. 2004. *Grey Systems, Modeling and Prediction*. Yangsky Press.
- 33) Wu, J-R., Y-W Lei, B-Y Chen and M. Ouhyoung. 2005. *User Interface Issues for a Building Walkthrough System with Motion Prediction*. Communication & Multimedia Lab, Department of Computer Science and Information Engineering, National Taiwan University.
- 34) Yan, L-J., C-S Cheng and M-C Wu. 2001. *The application of Grey System to Optimizing the Process Parameters of EDM*. Journal of Taipei University of Technology, Vol 34, No 2.
- 35) Yan, W-C., T-Y Chen and Y-T Wang. 2005. *The Design and Application of A Grey Model Noise Filter*. Department of Mechanical Engineering, Tan Kiang University, Taiwan.
- 36) Yeh, Y-L and T-C Chen. 2004. *Application of Grey Correlation Analysis for Evaluation the Artificial Lake Site in Pingtung Plain, Taiwan*. Journal of Civil Engineering, Canada, No 31.
- 37) Yue, J. 2003. *The Grey Dynamic Neural Network Model and Its Application*. Journal of Water Resources, 7.
- 38) Zhang, J., P. Wang, W. Yang and K. Zhang. 2004. *Analysis of Grey Correlation Grade among Main Characters of Soybean*. Journal of Shenyang Agricultural University, Vol 35, No 2.
- 39) Zhang, P., Y. Chen, J. Lin and X. Wang. 1995. *The Application of Grey Forecasting in Controlling Stock Investment Risk*. Journal of Hubei Institute of Technology. Vol 10, No 4.