PUMS99:1

UNIVERSITI MAŁAYSIA SABAH

DOL.	M	60 PM		LIN	AND	AMA	MOD	=1	Vanier	
A 7 A11.	PACHE	LOR	WITH	SCI	ENCE	DEGR	FE I	LITH	HONO	URS
AZAR.					22					
YA_	GUAH	21M	PENG (HURUF	BES	AR)		SESI	PENG	AJIAN:	2007/2002
ngaku Ilaysia	membenar Sabah den	kan tesi gan syai	s (LPSM. at-syarat	Sarja kegui	na/Dokto naan sep	or Falsafah) erti berikut	ini disin -	npan di	Perpustak	aan Universiti
1. 2.	Tesis adalı Perpustaka sahaja.	ah hakm aan Univ	ilik Univ versiti Ma	ersiti Ilaysia	Malaysi Sabah (a Sabah. dibenarkan	membuat	salina	n untuk tu	juan pengajian
3. 4.	Perpustaka pengajian Sila tandal	an dibe tinggi. can (/)	narkan m	embu	at salina	n tesis ini s	ebagai ba	han pe	rtukaran a	ntara institutsi
		LIT			(M Ke Ak	engandungi pentingan N CTA RAHS	maklum Malaysia IA RASN	at yang seperti AI 1972	berdarjah yang term 2)	keselamatan ata aktub di dalam
					(Me ole	engandungi h organisas	makluma i/badan d	at TERI i mana	HAD yang penyelidi	g telah ditentukan kan dijalankan)
-	Simler		RHAD					Disah	kan Oleh	
TANE	ATANGA	N PEN	ULIS)				(TAND	ATAN	GAN PUS	TAKAWAN)
mat Te	tap: BLO	- D-7	-7. SRI				PROF	. DR.	ZAINODI	N BIN HAJI JU
TERI	, 57000	K.L.						-	Nama	Penyelia
ikh: 2	4 APRIL	2008					Tarikh:_		-	
TATA	N:- *Poton **Jika t /orga	g yang esis ini nisasi bo	idak berl SULIT at	au TE	n. RHAD, n menya	sila lampir atakan sekal	kan surat i sebab d	daripad an tem	la pihak b poh tesis i	erkuasa ni perlu



FORECASTING MAYBANK'S SHARE PRICES USING MARKOV CHAIN AND ARMA MODEL

QUAH SIM PENG

THIS DISSERTATION IS PRESENTED TO FULFIL PART OF THE REQUIREMENT TO OBTAIN A BACHELOR OF SCIENCE DEGREE WITH HONOURS

MATHEMATICS WITH ECONOMIC PROGRAMME SCHOOL OF SCIENCE AND TECHNOLOGY UNIVERSITI MALAYSIA SABAH

APRIL 2008



DECLARATION

I, hereby declare that this is my original work except for quotation that I have clearly noticeable for sources.

24 APRIL 2008

QUAH SIM PENG HS2005-4789



1. SUPERVISOR

(PROF. DR. ZAINODIN HJ. JUBOK)

2. EXAMINER 1

(ASSOC. PROF. DR. HO CHONG MUN)

3. EXAMINER 2

(MADAM DARMESAH GABDA)

4. DEAN

(ASSOC. PROF. DR. SHARIFF A. K. OMANG)

SHen Mar Min 2



Signature

tu

AKNOWLEDGEMENT

I would like to express my gratitude to all those who gave me the possibility to complete this project. I am deeply indebted to my supervisor Prof. Dr. Zainodin Hj. Jubok from School of Science and Technology, University of Malaysia Sabah whose help, giving suggestions and encouragement helped me in all the time of writing this project.

My coursemate, Mr. Ng Yin Kee who wrote the C programme in estimation of missing values using cubic spline interpolation. I want to thank him for helping me in this project. I want to thank to my other course mates and friends who gave me support and encouragement.

Lastly, I would like to give my special thanks to my family whose patient love and support enabled me to complete this work.



ABSTRACT

Data that is used to forecast in this study is closing prices of Malayan Banking Berhad (MAYBANK). The trend and changes of share prices are studied and observed. The closing prices from 3 January 2000 to 29 December 2006 are used for data analysis. The closing prices from 3 January 2007 to 31 January 2007 are used for accuracy checking. Bursa Malaysia does not operate on weekend and public holidays. Thus, missing values exist and are estimated using Cubic spline interpolation. Five-state Markov chain, seven-state Markov chain, nine-state Markov chain and eleven-state Markov chain are being analysed. The forecast performances of different states of Markov chain are being compared and found that seven-state Markov chain obtained the best results. The better forecasting tool in forecasting share prices is investigated. ARMA model is better forecasting tool in forecasting share prices compared to Markov chain. The significance difference between model with and without missing values is being investigated. It is concluded that there is significance difference between models with and without missing values. Mean absolute percentage error (MAPE) is used in comparisons among the models.



ABSTRAK

Dalam kajian ini, data yang digunakan untuk ramalan adalah harga saham Malayan Banking Berhad (MAYBANK). Aliran dan perubahan harga saham telah dikaji.Harga saham bermula daripada 3 Januari 2000 hingga 29 Disember 2006 digunakan untuk menjalani data analisis. Harga saham daripada 3 Januari 2007 hingga 31 Januari 2007 digunakan untuk memeriksa ketepatan keputusan. Bursa Malaysia tidak beroperasi pada cuti-cuti am, Sabtu dan Ahad. Maka, nilai tidak diniagakan wujud dan dianggarkan dengan menggunakan cubic spline interpolasi. 5x5 rangkaian Markov, 7x7 rangkaian Markov, 9x9 rangkaian Markov dan 11x11 rangkaian Markov telah dikaji. Didapati 7x7 rangkaian Markov mendapat ramalan yang paling tepat. Kesesuaian rangkaian Markov dan model ARMA dalam meramalkan harga saham telah dikaji.Didapati model ARMA lebih sesuai dalam meramalkan harga saham berbanding dengan rangkaian Markov. Perbezaan antara keputusan model dengan nilai tidak diniagakan dan model tanpa nilai tidak diniagakan telah dikaji. Didapati keputusan model tanpa nilai tidak diniagakan mendapat ramalan yang lebih tepat. MAPE digunakan untuk membuat perbandingan antara model.



LIST OF CONTENT

		Page
DEC	LARATION	ü
VER	IFICATION	iii
AKN	IOWLEGMENT	iv
ABS	TRACT	v
ABS	TRAK	vi
LIST	OF CONTENT	vii
LIST	OF TABLE	xi
LIST	OF FIGURE	xiii
LIST	OF SYMBOL	xv
СНА	APTER 1 INTRODUCTION	
1.1	Introduction	1
1.2	Problems	1
1.3	Importance	3
1.4	Bursa Malaysia Berhad	4
1.5	Malayan Banking Berhad	5
1.6	Stock Indices	6
1.7	Objectives of study	7
1.8	Scope of study	8
1.9	Terms and definitions	8
	1.9.1 Missing values	8



	1.9.2 Share	9
CHA	PTER 2 LITERATURE REVIEW	
2.1	Cubic spline interpolation	11
2.2	Forecasting share prices	12
2.3	Time series	14
	2.3.1 Stationarity	14
	2.3.2 ARIMA model	15
2.4	Markov chain	18
CHA	PTER 3 METHODOLOGY	
3.1	Cubic spline interpolation	19
3.2	ARIMA model	24
	3.2.1 Stationary	25
	3.2.2 Stationarity test	28
	3.2.3 Possible Model	34
	3.2.4 Global Test	36
	3.2.5 Elimination Procedure	36
	3.2.6 Coefficient Test	37
	3.2.7 Wald Test	38
	3.2.8 Obtain Best Model Using Eight Selection Criteria	39
	3.2.9 Residuals	41
	3.2.10 Randomness Test	42
	3.2.11 Forecasting	43
3.3	Markov chain	44



viii

	3.3.1	Classifications of States	45
	3.3.2	Initial Probability	48
	3.3.3	Transition matrix	48
	3.3.4	Comparisons	50
	3.4.1	Comparisons among Markov chain	51
	3.4.2	Comparisons between Markov chain and ARMA model	53
	3.4.3	Comparisons between series with and without missing values	53
СНА	PTER 4	DATA DESCRIPTION	
4.1	Introd	uction	54
4.2	Estim	ation of missing values	55
4.3	Descr	iption of data	61
4.4	Statio	narity test	63
	4.4.1	Series with missing values	64
	4.4.2	Series without missing values	69
СНА	PTER	5 DATA ANALYSIS AND RESULTS	
5.1	Introd	uction	74
5.2	ARM	A model	75
	5.2.1	Possible models	75
	5.2.2	Global Test	76
	5.5.3	Elimination procedure	78
	5.2.4	Coefficient Test	86
	5.2.5	Wald Test	90
	5.2.6	Obtain Best Model Using Eight Selection Criteria	93



	5.2.7	Residuals	97
	5.2.8	Randomness test	103
	5.2.9	Forecasting	106
5.3	Marko	v chain	109
	5.3.1	Example of computation Markov chain	109
	5.3.2	Classifications of States	116
	5.3.3	Initial Probability	116
	5.3.4	Transition Matrix	118
5.4	Compa	arison	121
CHAP	TER 6	CONSLUSION	
6.1	Conslu	usion	122
6.2	Recon	umendations	126
REFF	EREN	CES	128
APPE	NDIX		131



LIST OF TABLE

No. of	fTable	Page No.
3.1	Conditions of cubic spline interpolation	20
3.2	Example of Possible Models	35
3.3	Eight Selection Criteria	40
3.4	Example of Seven Intervals	47
3.5	Number Assigned to Each State	51
3.6	Computation of MAPE	52
4.1	Closing prices of Maybank from 3 January 2000 to 31 January 2000	55
4.2	Interior Points and Endpoints	56
4.3	Cubic Spline Model for Data of Table 4.1	58
4.4	Missing Values Estimated Using Cubic Spline Models in Table 4.3	59
5.1	Output For M53	76
5.2	ANOVA Table For M53	77
5.3	Output For M53	79
5.4	Output For M53.1	80
5.5	Output For M53.2	81
5.6	Output For M53.3	82
5.7	Output For M53.4	83
5.8	Output For M53.5	84
5.9	Output For M53.6	85
5.10	Output For M53.3	86



xi

xi

5.11	Output For M53	90
5.12	ANOVA table For M53	91
5.13	Output For M53.6	92
5.14	ANOVA Table For M53.6	92
5.15	ANOVA Table For Wald test	93
5.16	Output For M53.6	94
5.17	Output of M'64.5	95
5.18	Forecast results of ARIMA Model (Data with missing values)	107
5.19	Forecast results of ARIMA Model (Data without missing values)	108
5.20	Sample of closing prices	110
5.21	Intervals of example	112
5.22	Event table	112
5.23	Transition matrix for example	113
5.24	Transition table	115
5.25	Summary table of MAPE	121



xii

LIST OF FIGURE

No. o	of Figure	Page
3.1	Cubic spline functions	21
3.2	Nonstationary time series	27
3.3	Stationary time series	27
3.4	Correlogram of stationary time series	29
3.5	Correlogram of nonstationary time series	30
3.6	Normal distribution curve	46
4.1	Closing price of Maybank against day (data without missing values)	61
4.2	Closing price of Maybank against day (data with missing values)	62
4.3	Correlogm of closing prices of Maybank with missing values	64
4.4	Correlogram of logarithmic of closing prices with missing values	66
4.5	Correlogram of first-order values of differencing for closing prices	68
	with missing values	
4.6	Correlogram of closing prices of Maybank without missing values	69
4.7	Correlogram of logarithmic of closing prices with missing values	71
4.8	Correlogram of first-order values of differencing for closing prices	
	with missing values	72
5.1	Residuals plot against days (series with missing values)	97
5.2	Residuals plot against days (series without missing values)	98
5.3	Correlogram For Residuals of M53.6	99
5.4	PACF For Residuals of M53.6	99



xiii

5.5	Correlogram of residuals of M'64.5	100
5.6	PACF of residuals of M'64.5	100
5.7	Histogram of Residuals (Series With Missing Values)	101
5.8	Histogram of Residuals (Series Without Missing Values)	102
5.9	Procedure of Transformation and Retransformation of Data	106
5.10	Standard Normal Distribution of five-state matrix	112



xiv

LIST OF SYMBOL

5	less than or equal to
<	less than
α	alpha
β	beta
∇	difference
=	equal to
+	plus
-	minus
¥	not equal to
Σ	summation
H ₀	null hypotheses
H ₁	alternative hypotheses
ρ	rho
μ	miu
E	subset



CHAPTER 1

INTRODUCTION

1.1 Introduction

In recent years, more people invest in stock market with the introduction of online trading. Everyone are expecting for high returns in their investments. If we can predict the unseen future in stock market, then those investors know which, where and when to invest to make more profit. In this study, Markov chain and ARIMA model are used to forecast the Maybank's closing prices. Before the closing prices are being forecasted, the missing values are estimated by using Cubic spline interpolation.

1.2 Problems

Everyone wants to win in stock market, but picking a profitable share is difficult. Everyone has their own technique to spot a share that will rise in shares. Generally, there



are two basic techniques to spot profitable shares namely fundamental analysis and technical analysis.

According to Bowerman *et al.* (2005), fundamental analysis include look at the business of company, the financial health of company, the state of current economy, the performance of company and the share prices of company. The most common method to rate the share is measure the profit of company and payment of dividends. Besides that, the annual report and company accounts also important. By comparing a few companies, it is easier to spot a profitable share.

Technical analysis is forecast the future share prices using historical data. Historical data is a good indicator of the future prices. It is because technical analysts believe that there is trend or season exists in the price movement and the pattern might repeat. The pattern in the chart might indicate the next move of the price of a company.

This study is using the technical analysis to forecast the share prices. The methods used are ARIMA model and Markov chain to forecast. A model may produce a good fit to historical data, but fail to predict the unseen future. This research is designed to find the best model in Markov chain and ARIMA model. By using the best model, predict the most accurate results. There are many models to make predictions such as Markov chain, Time series, artificial neural networks, multiple regressions and many more.



In this research, only Markov chain and ARIMA model are used to make forecast. Compare the results of Markov chain and ARIMA model and determine which method is a better tool to forecast share prices. The forecast errors are the difference between the actual value and the forecasted values. Mean Absolute Percentage Error is used to determine the reliability of ARIMA models.

1.3 Importance

Numerous studies on forecasting have been done by using different methods. This is because forecasting is very important for us to predict our unseen future. Referring to Frances (1998), for example in the fields of economics and business, variables such as inflation rates, stock market indices, unemployment rates and market shares are forecasted because the out-of-sample forecasts for this variables are often needed to set policy targets. For example, the forecast for next year's inflations rate can lead to a change in the monetary policy of a central bank. A forecast of a company's share in the next few months may lead to changes in the allocation of the advertising budget.

In this research, Maybank's share prices from 3 January 200 to 31 December 2006 are used to forecast using markov chain and ARIMA model. The share of Maybank was chosen because Maybank is the largest bank in Malaysia. It is listed in Main Board of Bursa Malaysia Berhad It plays a very important role in economy Malaysia.



1.4 Bursa Malaysia Berhad

According to Chang *et al.* (1994), the Kuala Lumpur Stock Exchange (KLSE) provides trading facilities for common equities, preferred stocks, Malaysian government securities, corporate debt securities, transferable subscription rights and warrants. There is only one stock exchange market in Malaysia known as Bursa Malaysia nowadays. It consists of main board, second board and MESDAQ.

According to Ding *et al.* (1994), in May 1960 shares were publicly traded for the first time with the formation of the Malaysia Stock Exchange. At that time, Singapore was part of Malaysia. The board system had trading rooms in Singapore and Kuala Lumpur, linked by direct telephone lines. Since the reunion of Malaysia on 9 August 1963, Stock Exchange Malayan and Singapore changed to Stock Exchange of Malaysia. In 1965, with the secessions of Singapore from Malaysia, the stock exchange continued functioning as a single entity as the Stock Exchange of Malaysia and Singapore. This was because the interchangeability of the two countries' currencies. In 1973, the currency interchangeability was terminated, leading to the formation of the Kuala Lumpur Stock Exchange and the Stock Exchange of Singapore as two separate entities.

According to Ding *et al.* (1994), on 13 November 1989, the KLSE replaced the traditional open outcry system of trading in favor of a semi-automated trading system known as the System on Computerized Order Routing and Execution (SCORE). The KLSE employed two different processes of order-matching price discovery known as call



market system and continuous market system. The call market represents a periodic single-price auction which is a computerized form of the clearing of house auction system where bids and offers are submitted continuously over time and transactions occur when the orders cross. The KLSE used the call market to determine the both the stock's opening price and it's closing price for a particular trading session, while its continuous market system was used to determine matching prices throughout the trading session

On 14 December 1976, KLSE was incorporated as a company limited by guarantee, took over the operations of the Kuala Lumpur Stock Exchange Berhad. On 14 April 2004, it had been changed to Bursa Malaysia Berhad. On 18 March 2005, Bursa Malaysia was listed on the Main Board of Bursa Securities Berhad.

1.5 Malayan Banking Berhad (Maybank)

According to Maybank Annual Report 2006, Maybank was founded by Malaysian business tycoon Khoo Tech Puat, who died in 2004. However, the company has been led by President and CEO Armisham A. Aziz for some two decades.

Malayan Banking Berhad (Maybank) is the largest bank and financial group in Malaysia. Maybank is playing a very important role in economy of Malaysia. Maybank is listed in Main Board of Bursa Malaysia Berhad. At the end of June 2007, Maybank is the second top listed company in Main Board with the market value RM 47.04 millions. Maybank have more than 450 branch offices and more than 2500 ATM machines in



Malaysia. Maybank also operate branches in some Southeast Asian countries, New York, London and China. Maybank was the first Malaysian bank establishes a branch office in China.

Besides commercial banking network, Maybank operates a number of specialized subsidiaries in the insurance, investment banking and assets management, finance sectors. The group's subsidiaries include Mayban general Assurance, Mayban Life Assurance and Mayban Takaful.

1.6 Stock indices

According to Chia and Soh (1994), a stock index is a basket of shares that indicates the overall situation and trend of the stock market. It is more like a barometer. In other words, it is a system that shows the price changes of overall market. So we can compare the level of prices and wages with those of a previous date.

Most of the indices are designed to fulfill the requirement for representative value without including every single share. Since only some samples of index were involve, it may give poor indication about the movement of market. But, the indexes still remain 'flat' as price changes are buried.



There are two types of index namely weighted indices and unweighted indices. The unweighted indices are those indices that the importance of each company share is equal. In contrast, the weighted indices give priority to those companies which are more widely traded, or whose activity contributes more to country's economy.

1.7 Objectives of study

Objective is very important for a study. The objective is act as a guideline in the process of completing the study. Every objective must be carried out before this study end. The objectives of this study is

- a. To observe the trend of Maybank's share prices.
- b. To estimate missing values using cubic spline interpolation.
- c. To find the best ARIMA model.
- d. To forecast Maybank's share prices using ARIMA model.
- e. To forecast Maybank's share prices using Markov chain.
- f. To investigate which dimensions of matrix provide more reliable forecast result.
- g. To investigate whether ARMA model or Markov chain is more suitable in forecasting share prices.
- h. To investigate the significance of missing values in the obtained best model.



1.8 Scope of study

This research is focus on the closing prices of Maybank which is now listed in the main board of Bursa Malaysia Berhad. The data used in model fitting covered from 3 January 2000 to 29 December 2006. The data from 3 January 2007 to 31 January 2007 are used for validation. The methods used to forecast are Markov Chain and ARIMA model. Besides, cubic spline interpolation is used to estimate the missing values.

1.9 Terms and definitions

Some terms and definitions are explained in this section. There are two terms are explained in this study namely missing values and share.

1.9.1 Missing values

Bursa Malaysia only operates on weekdays which are from Monday to Friday. Saturday, Sunday and public holidays are holidays which yield no stock price. Out of 365 days a year, there are more than 100 such cases and thus we called "missing value". Missing value can easily affect the result of forecasting. Therefore, missing values must be estimated before data analysis start.

According to Yaffee (2000), if some social or economic indicators have too much missing data, then the series may not be amenable to time series analysis. When series



REFFERENCES

- Ahmad, Z. & Kothyari, U. C. 2001. Time-line cubic spline interpolation scheme for solution of advection equation. *Computers & Fluids*. 30 (6): 737-752.
- Bowerman, B. L., O'Connell R. & T., Kochler, A. B. 2005. Forecasting Time Series & Regression. 4th Ed. Thomson Brooks/Cole, California.
- Chang, R. P., Kang, J. K. & Rhee, S. G. 1994. The Behavior of Malaysian Stock Price. Capital Market Review. 2 (1): 15-41.
- Chatfield, C. 1996. The Analysis of Time Series. 5th Ed. Chapman & Hall, New York.
- Chia, R. & Soh, D. 1994. How to invest in stocks & shares. 4th Ed. Times Books International, Singapore.
- Frances, P. H. 1998. *Time series models for business and economic forecasting*. Cambridge University Press, New York.
- Feldman, R. M. 1996. Applied probability and stochastic process. PWS Publishing Company, Boston.



- Givon, M. & Grosfeld-Nir, A. 2008. Usinf partially observed Markov processes to select optimal termination time of TV shows. *Omega.* **36** (3): 477-485.
- Gooijer, J. G. D. & Hyndman, R. J. 2006. 25 years of time series forecasting. International journal of Forecasting. 22 (3): 443-473.
- Hiller, F. S. & Lieberman, G. J. 1995. Introduction to Operations Research, 6th Ed. McGrraw-Hill, Inc., Singapore: 630.
- Ismail Mohd., Siska Chandra Ningsih & YoszaDasril. 2007. Unimodality Tests For Global Optimization of Single Variable Functions Using Statistical Method. *Malaysian Journal of Mathematical Sciences*, 1 (2): 43-53.
- Jensen, P.A. & Bard, J. F. 2003. Operations Research Model and Method. John Wiley & Sons, Inc, New Jersey.
- Klapper, D. & Herwartz, H. 2000. Forecasting market share using predicted values of competitive behavior: further empirical results. *International Journal of Forecast.* 16 (3): 399-421.
- Kohzadi, N., Boyd, M. S., Kermanshahi, B. & Kaastra, I. 1996. A comparison of artificial neural network and time series models for forecasting commodity prices. *Neurocomputing.* 10 (2): 169-181.



- Kumar, V., Nagpal, A. & Venkatesan, R. 2002. Forecasting category sales and market share for wireless telephone subscribers: a combined approach. *International Journal of Forecasting.* 18 (4): 583-603.
- Lim, T. S. & Loh, W. Y. 1996. A comparison of tests of equality of variances. Computataional Statistics & Data Analysis. 22 (3): 287-301.
- Lim, C. & McAleer, M. 2002. Time series forecasts of international travel demand for Australia. *Tourism Management.* 23 (4): 389-396.
- Lind, D.A., Marchal, W.G. & Wathen, S.A. 2005. Statistical Technique in Business & Economics. 12th Ed. McGraw-Hill/Irwin, New York.
- Pei, C. C. & Liu, C. H. 2008. A TSK type fuzzy rule based system for stock price prediction. Expert Systems with Applications. 34 (1): 135-144.
- Pitfield, D. E. 2007. The impact on traffic, market shares and concentration of airline alliances on selected European-US routes. *Journal of Air Transport Management*. 13 (4): 192-202.
- Ramanathan, R. 2002. Introductory Econometrics with Applications. 5th Ed. South Western, Ohio.



- Shiavi, R. 2006. Introduction to Applied Statistical Signal Analysis: Guide to Biomedical and Electrical Engineering Applications. Elsevier, Inc., Tennessee.
- Wei, W. W. S. 1990. Time Series Analysis Univariate and Multivariate Mehtods. Addisson – Wesley Publishing Company, Boston.
- Vogelvang, B. 2003. Econometrics Theory and Applications with EViews. Prentice Hall, London.
- Yaffee, Robert A. 2000. Introduction to Time Series Analysis and forecasting with applications of SAS and SPSS. Academic Press, California.

