STOCHASTIC DIFFERENTIAL EQUATION APPROACH FOR DAILY GOLD PRICES IN SRI LANKA

Weerasinghe Mohottige Hasitha Nilakshi Weerasinghe

(148914G)

Dissertation submitted in partial fulfillment of the requirements for the Degree Master of Science in Financial Mathematics

Department of Mathematics

University of Moratuwa

Sri Lanka

May 2018

Declaration of the Candidate

"I declare that this is my own work and this thesis/dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any University or other institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text"

Signature:

……………………　　　　　　　　　　　　　　　……………………

W.M.H.N. Weerasinghe　　　　　　　　　　　　　　　Date

148914G

Declaration of the Supervisor

I have supervised and accepted the thesis titled "Stochastic Differential Equation Approach for Daily Gold Prices in Sri Lanka" for the submission of the degree.

Signature of the supervisor:

…………………… ……………………

Mr. A.R. Dissanayake Date
Senior Lecturer,
Department of Mathematics,
Faculty of Engineering,
University of Moratuwa

ACKNOWLEDGEMENT

First and foremost, I would like to thank my supervisor, Mr. A.R. Dissanayaka, Senior Lecturer, Department of Mathematics, University of Moratuwa, for the patient guidance, encouragement and advice he has provided throughout my time as his student. I have been extremely lucky to have a supervisor who cared so much about my work, and who responded to my questions and queries so promptly.

I would like to express my deepest appreciation to the course coordinator, T.M.J.A. Cooray, Senior Lecturer, Department of Mathematics, University of Moratuwa and all the other members of staff at the department of Mathematics, University of Moratuwa for helping me keep things in perspective.

And also I would be thankful to all the friends. Without them, completing this work would have been more difficult.

I would like to extend my thanks to my colleagues in the Department of Mathematics, University of Kelaniya for their support and encouragement.

A special mention for my family: without your kindness, love, dedication and support nothing would be possible.

W.M.H.N. Weerasinghe
148914G

ABSTRACT

In our day to day life, predictability of gold prices is significant in many domains such as economic, financial and political environment. The objectives of this research are to study the behavior of the gold price in Sri Lanka, to forecast the daily gold prices making use of four Stochastic Differential Equation (SDE) models, Brownian motion, Geometric Brownian motion, Cox-Ingersoll-Ross (CIR) model and Vasicek model and compare the results with an ARIMA (2,1,2) model which is used to forecast the Sri Lankan gold prices in a previous research. The daily gold prices per troy ounce in Sri Lanka are obtained from 01st of October 2015 to 14th of October 2016 from the website http://www.cbsl.gov.lk/htm/english/_cei/er/g_1.asp on 1st of November, 2016. The gold prices from 01st of October 2015 to 07th of October 2016 are used to estimate the parameters of the four models and the parameter estimation is done using maximum likelihood estimation method. The gold prices from 10th of October 2016 to 14th of October 2016 are used to forecast the gold price. By taking the gold price on 10th of October 2016 as the initial value, daily gold prices from 11th of October 2016 to 14th of October 2016 are forecasted. Numerical approximations are carried out using Euler-Maruyama approximation method and the Monte Carlo simulation technique is used to simulate the daily gold prices. After evaluating forecasting accuracy of estimated models and existing ARIMA (2,1,2) model by root mean square error (RMSE) and mean absolute percentage error (MAPE), it turns out that the Vasicek model has the minimum RMSE and MAPE values for the given data set. The price of the gold may change rapidly because of some economic factors such as inflation, currency exchange rates etc. In these situations the best SDE model to forecast the daily gold price in Sri Lanka may be changed to another model. Hence this method is suitable for short runs only.

Keywords: Gold Price, Stochastic differential equations, Maximum likelihood estimation, Monte Carlo method, Euler-Maruyama method

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Description
AIC	Akaike Information Criterion
ARFIMA	Auto Regressive Fractionalized Integrated Moving Average
ARIMA	Auto Regressive Integrated Moving Average
ARMA	Auto Regressive Moving Average
BIC	Bayesian Information Criterion
BMA	Bayesian Model Averaging
CIR	Cox Ingersoll Ross Model
CRB	Commodity Research Bureau
DMA	Dynamic Model Averaging
DMS	Dynamic Model Selection
ERC	Earnings Response Coefficients
ETF	Exponential Smoothing
INF	Inflation
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MLR	Multiple Linear Regression
RMSE	Root Mean Square Error
RW	Random Walk
SDE	Stochastic Differential Equations
TBATS	Trend and Seasonal components
VAR	Vector Auto Regressive

LIST OF APPENDICES

CHAPTER 01

INTRODUCTION

In this chapter, a brief description about the research is provided. This chapter describes the background of the study in section 1.1, objectives in section 1.2 and the significance of the study in section 1.3. Organization of the dissertation is given in the section 1.4.

1.1 Background of the Study

The Gold has a long history as a valuable metal and its history is far from over. From the ancient Egyptians to the modern Treasury, there are few metals that have such an influential role in human history as gold.

Human fascination with gold is as old as recorded history. However, flakes of gold have been found in Paleolithic caves dating back as far as 40,000 B.C. Most archaeological evidence shows that human who came into contact with gold were impressed by the metal. Since gold is found all over the world, it has been mentioned numerous times throughout ancient historical texts.

The Egyptians produced the first known currency exchange ratio which mandated the correct ratio of gold to silver: one piece of gold is equal to two and a half parts of silver. This is also the first recorded measurement of the lower value of silver in comparison to gold. They also produced gold maps, some of which survive to this day. These gold maps described where to find gold mines and various gold deposits around the Egyptian kingdom. As much as the Egyptians loved gold, they never used it as a bartering tool. Instead, most Egyptians used agricultural products like barley as a de-facto form of money. The first known civilization to use gold as a form of currency was the Kingdom of Lydia, an ancient civilization centered in western Turkey.

In 1792, the United States Congress made a decision that would change the modern history of gold. Congress passed the Mint and Coinage Act. This Act established a fixed price of gold in terms of U.S. dollars. At the time, gold was worth approximately 15 times more than silver. That ratio would change after the Civil War. In 1862, paper money was declared to be legal tender, marking the first time a fiat currency (not convertible on demand at a fixed rate) was used as an official currency.

After decades of war and conflict, world leaders came together under the Bretton Wood's Agreements. This system created a gold exchange standard where the price of gold was fixed to the U.S. dollar. The day the price of gold was pegged to the U.S. dollar is one of the most important points of U.S. history because it helped make the United States the global superpower it is today. In 1944, gold was fixed at $35 per ounce for the foreseeable future. In the early 1970s, the Vietnam War caused the gold exchange standard to collapse. America's budget was in ruin and in 1971, President Nixon suddenly decided to end the Bretton Woods system with a moment known in history as the Nixon Shock.

Today, no countries in the world use a gold standard. In other words, no currency in the world is backed by gold. The last major currency to use a gold standard was the Swiss Franc, which used a 40% gold reserve until the year 2000.

In present, gold has been seen as a smart investment. However, the use of gold as an investment became hugely popular after the end of the Bretton Woods system in 1971. Since the 1970s, the price of gold has steadily increased. In 1970, gold was pegged at $35 per ounce. However, the years in between were not a smooth upward slope and gold – like any other investment – has gone through a number of ups and downs over the past few decades.

Karat is the term used to measure the gold content or purity. The higher the karatage, the purer the gold. 24k gold is also called pure gold or 100 per cent gold. This means that all

24 parts in the gold are all pure gold without traces of any other metals. It is known to be 99.9 per cent pure and takes on a distinct bright yellow color. There is no higher form of gold than 24K. Since this is the purest form of gold, it is naturally more expensive than 22K or 18K gold. 22K gold implies that 22 parts of the jewelry amounts to gold and the balance 2 parts are some other metals. This kind of gold is commonly used in jewelry making. In 22K gold, of the 100 per cent, only 91.67 per cent is pure gold. The other 8.33 per cent comprises metals like silver, zinc, nickel and other alloys. 18K gold is 75 per cent gold mixed with 25 per cent of other metals like copper /6or silver etc. This kind of gold is less expensive compared to 24K and 22K. This one has a slightly dull gold color. Troy ounce is another measure for gold. It is a unit of measure for weight that dates backs to the middle age. One troy ounce is equal to 31.1034768 grams.

By carefully weighing all of this information and current trends, we can build an accurate view of the present value and future value of gold. The high value of gold is generally accepted to be the result of a combination of factors such as scarcity, physical characteristics**,** aesthetic attributes and wealth storage.

The gold market is deep and liquid and there are many ways for investors to buy physical gold or gain an exposure to movements in the gold. Some of them are gold coins, gold bars, gold exchange traded funds, gold mining equities, gold accounts, gold futures and options and the over-the-counter market.

Because of the gold has a higher demand, lots of people are interested in dealing with gold. Some of them are government, banks and jewelry makers. In present, each bank has the pawning facility. To attract people to the bank, the bank try to give the maximum credit for the gold. Because of that, there is a competition between banks. Hence all the banks try to find the future price of the gold.

By observing historical daily gold prices in Sri Lanka, it can be assumed that there will be a higher demand for the gold in near future. Because of that many investors will try to invest their money in gold market and they will be very interested to get some idea about the future gold prices.

Just like any commodity, it's impossible to accurately predict the price of gold. There are many factors affecting for the price of gold. Demand for consumer goods, investment, inflation prospects, value of the dollar, gold reserves, lack of the safe havens, stock market and speculation are some factors among them.

Every day, thousands of investors around the world study all of the metrics involved in the price of gold. Some of these experts will take all of this information and accurately predict the future price of gold, while other experts will see the same information and guess wrong.

By reviewing the literature of the gold price forecasting, it can be observed that researchers used time series models very frequently and few researchers used methods such as wavelet schemes, dynamic models and Bayesian models. In literature it could not be found any research which used stochastic differential equation (SDE) models to forecast the gold price. In this research, the daily gold price of Sri Lanka is forecasted using four SDE models named Brownian motion, Geometric Brownian motion, Cox-Ingersoll-Ross model and Vasicek model. Model parameters are estimated using maximum likelihood estimation method. MATLAB software is used for computation and graphical plotting of data. To approximate the solution of an SDE, Euler- Maruyama method or Milstein method can be used. In this study Euler-Maruyama approximation method is used to approximate the SDE. To simulate the predicted gold prices, Monte-Carlo technique is used. There are several measures for evaluating forecasts. For this study, the root mean square error (RMSE) and the mean absolute percentage error (MAPE) are used. When comparing performance of models, smaller values of RMSE and MAPE indicate the better model. In literature, a model to predict daily gold prices in Sri Lanka cannot be found. Because of that, forecasting accuracy measures of four SDE models were compared with the forecasting accuracy measures of an ARIMA (2, 1, 2) model which is used to predict the monthly gold price in Sri Lanka [14].

1.2 Data Collection

To test the accuracy of the model, daily gold prices per troy ounce from 01/10/2015 to 14/10/2016 were obtained from http://www.cbsl.gov.lk/htm/english/_cei/er/g_1.asp on 01/11/2016. Among these data, daily gold prices from 01^{st} of October 2015 to 07^{th} of October 2016 were used to estimate the model parameters and by taking the initial value as the daily gold price on 10^{th} of October, daily gold prices from 11^{th} of October, 2016 to 14^{th} of October, 2016 are predicted.

1.3 Objectives of the Study

The objectives of this research are:

1. Study the behavior of the gold price in Sri Lanka.
2. Find the suitable SDE model among four SDE models, Brownian motion, Geometric Brownian motion, Cox-Ingersoll- Ross model and Vasicek model for Sri Lankan gold price.
3. Forecast the daily gold price using the most suitable model among four SDE models.
4. Compare the results with previously used ARIMA model to forecast the gold price in Sri Lanka.

1.4 Significance of the Study

If we refer previous research studies on forecasting Gold prices, many researchers used time series models to forecast the gold price. It should be noted that any researcher did not use SDE models to forecast the gold price. But SDE models are used in finance for various purposes such as stock pricing. Therefore SDEs may be suitable to forecast daily gold prices.

By examining the past data of the price of gold, it can be concluded that the demand for the gold in Sri Lanka will be increased in future. Because of this reason, many investors will be invested their money in gold market than other financial markets. Hence the investors will be very interested to get most accurate predicted daily gold prices in Sri Lanka. Because they can invest their money with law risk if they have an idea about the future gold prices in Sri Lanka. If the models discussed in this research will predict the daily gold prices accurately, investors can be used these models to get some idea about the future daily gold prices in Sri Lanka.

1.5 Outline of the Thesis

The rest of the chapters were organized as follows.

In chapter 2, a brief synoptic review of the empirical literature will be provided and that chapter included various time series models, dynamic model averaging and dynamic model selection methods which are used to predict the gold price.

Next, chapter 3 discussed the empirical methodology employed in the study. It included mathematical preliminaries, stochastic processes, stochastic differential equation, estimation methods, approximation methods and forecasting accuracy measures which are used in this study.

The research was followed by the result interpretations based on the estimation outputs in chapter 4. In this chapter, parameter estimations, forecasted daily gold prices, forecasting accuracy measures of four SDEs and ARIMA (2, 1, 2) model are included.

Lastly, chapter 5 had concluded our research by summarize the major findings, contributions of study, limitations of study and some of the recommendations for future research.

CHAPTER 02

LITERATURE REVIEW

The comprehensive review of research from the existing researchers related to forecasting gold price had been documented in this chapter.

2.1 Review of the Literature

Alessio Azzutti (2016) evaluated the use of 6 different parametric and nonparametric time series analysis and forecasting techniques using monthly gold price data. The six models are Auto Regressive Integrated Moving Average (ARIMA), Random walk (RW), Auto Regressive Fractionalized Integrated Moving Average (ARFIMA), Exponential Smoothing (ETS), Exponential smoothing state space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components (TBATS) and Multiple Linear Regression (MLR). This research concluded that, among these six models ARIMA model is better.

Banhi Guha and Gautam Bandyopadhyay (2016) forecasted Indian gold price using ARIMA models. The research suggested that ARIMA (1, 1, 1) is the best model among six different models.

Asad Ali, Muhammad Iqbal Ch., Sadia Qamar, Noureen Akhtar, Tahir Mahmoods, Mehvish Hyder (2016) proposed a time series model for forecasting the daily Gold price and used the data set of United State Dollars per ounce from Jan 02, 2014 to Jul 03, 2015 for the said purpose. By using the Box-Jenkins methodology, Autoregressive Integrated Moving Average (ARIMA) model is selected and the model selection criterion (AIC and BIC) shows that ARIMA (1,1,0) and (0,1,1) are close to each other for forecasting the daily Gold price. The forecasted values reveal that ARIMA (0,1,1) is more efficient than ARIMA (1,1,0) on the base of model selection criteria's, Mean

Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) .

Pitigalaarachchi P. A. A. C., Jayasundara D. D. M., Chandrasekara N. V. (2016) developed two models Auto Regressive Integrated Moving Average (ARIMA) model and Vector Auto Regressive (VAR) model for forecasting monthly gold prices per troy ounce in Sri Lanka. The research concluded that ARIMA(2,1,2) is the best model to forecast the gold price in Sri Lanka.

M. Khalid, Mariam Sultana, Faheem Zaidi (2014) forecasted the price of Gold in Pakistan market using ARIMA and two distinct versions of wavelet scheme. After evaluating the accuracy of those models by mean absolute error and mean square error, it turned out that wavelet neural transformation has better prediction accuracy than rest of the models.

Goodness Aye, Rangan Gupta, Shawkat Hammoudeh, Won Joong Kim (2014) developed models for examining possible predictors of the return on gold that embrace six global factors (business cycle, nominal, interest rate, commodity, exchange rate and stock price factors) and two uncertainty indices (the Kansas City Fed's financial stress index and the U.S. Economic uncertainty index). Specifically, by comparing with other alternative models, the research showed that the dynamic model averaging (DMA) and dynamic model selection (DMS) models outperform not only a linear model (such as random walk) but also the Bayesian model averaging (BMA) model for examining possible predictors of the return of gold. The DMS is the best overall across all forecast horizons.

Dirk G. Baur, Joscha Beckmann, Robert Czudaj (2014) showed that Dynamic Model Averaging (DMA) improves forecasts compared to other frameworks and provided a clear evidence for time variation of gold price predictors.

Rebecca Davis, Vincent Kofi Dedu, Freda Bonye (2014) forecasted the price of gold using an ARMA model.

Abdullah Lazim (2012) in his paper has addressed the forecasting of gold bullion coin prices through ARIMA model and had concluded by suggesting that the gold bullion coin selling prices are in upward trends and could be considered as a worthy investment.

Deepika M G, Gautam Nambiar & Rajkumar M (2012) has tried to study the forecasting of gold price through ARIMA model & Regression but their finding suggests that suitable model was not identified to forecast Gold price through ARIMA Model hence Regression analysis was carried out in the later part of their study.

Shahriar Shafie and Erkan Topal (2010) have forecasted the gold price by applying a modified econometric version of the long term trend reverting jump and dip diffusion model.

Z. Ismail, A. Yahya and A. Shabri (2009) developed a Multiple Linear Regression (MLR) model for predicting gold prices based on economic factors such as inflation, currency price movements and others. Two models were considered. The first model considered all possible independent variables. The second model considered only four independent variables the Commodity Research Bureau future index (CRB lagged one), USD/Euro Foreign Exchange Rate (EUROUSD lagged one), Inflation rate (INF lagged two) and Money Supply (M1 lagged two) to be significant. In terms of prediction, the second model achieved high level of predictive accuracy.

M.M. Ali Khan (2008) used Box-Jenkins, Auto Regressive Integrated Moving Average (ARIMA) methodology for building forecasting model. Results suggested that ARIMA(0,1,1) is the most suitable model to be used for predicting the gold price.

Pravit Khaemasunun (2006) applied two forecasting models, Multiple-Regression, and Auto-Regressive Integrated Moving Average (ARIMA), are applied to forecast the Thai gold price. The research result suggested that ARIMA (1, 1, 1) is the most suitable model to be used for forecasting gold price in the short term. The second method, multiple-regression, showed that Australian Dollars, Japanese Yen, US dollars, Canadian Dollars, EU Ponds, Oil prices and Gold Future prices have effect on the change of Thai gold price.

Selvanathan (1991) has analyzed the accuracy of the gold price forecasts gathered from a panel of gold experts and concluded that forecasts from a simple random walk model are superior to the ERC panel forecasts and simple random walk model forecasts are cheap as compared to the efforts of the panel of experts.

2.2 Chapter Summary

According to the literature review, it can be observed that the most researchers used time series models such as ARMA, ARIMA, ARFIMA and MLR to forecast the price of gold. Some researchers used dynamic model averaging (DMA) and dynamic model selection (DMS) models for gold price forecasting. And also there exists one random walk model to forecast the gold price. But, the mathematical models such as stochastic differential equations were not used by any researcher to forecast the gold price. Hence this research is a new approach to forecast the gold price.

To forecast the gold prices in Sri Lanka, time series models were used and that model is used to forecast the monthly gold prices in Sri Lanka. Since the price of gold can be changed more rapidly, forecasting daily gold prices is most preferable than forecasting monthly gold prices.

CHAPTER 03

METHODOLOGY

This chapter introduces the research methodologies used for this research and how it has guided data collection, analysis and development of theory. Section 3.1 describes the mathematical preliminaries used in this study. Stochastic processes, stochastic integrals and stochastic differential equations are explained in sections 3.2, 3.3 and 3.4 respectively. Maximum error of the estimates and forecasting accuracy measures are described in sections 3.5 and 3.6.

3.1 Mathematical Preliminaries

The Law of Large Numbers is an important limit theorem that is used in a variety of fields including statistics, probability theory and areas of economics and finance. It can be used to optimize sample sizes as well as approximate calculations that could otherwise be troublesome.

Theorem 3.1: Law of large numbers

Let $X_1, X_2, \ldots$ be independent and identically distributed random variables. Let $\mu = E(X_n)$ and $\sigma^2 = Var(X_n)$. Define $S_n = \sum_{i=1}^{n} X_i$. Then,

$$\lim_{n\to\infty} E\left(\left|{S_n}/{n} - \mu\right|^2\right) = 0 \text{ and } \lim_{n\to\infty} {S_n}/{n} = \mu \text{ with probability 1.}$$

Another important theorem is the central limit theorem. This theorem gives the ability to measure how much the means of various samples will vary, without having to take any other sample means to compare it with.

Theorem 3.2: Central Limit Theorem

Define $S_n = \sum_{i=1}^{n} X_i$ .Let $Z_n = {(S_n - n\mu)}/{\sigma\sqrt{n}}$. Then Z_n converges in distribution to $Z \sim N(0,1)$.

3.1.1 Monte Carlo method

Monte Carlo method means using random numbers in scientific computing. More precisely, it means using random numbers as a tool to compute something that is not random.

In principle, Monte Carlo methods can be used to solve any problem having a probabilistic interpretation. By the law of large numbers. Integrals described by the expected value of some random variable can be approximated by taking the empirical mean of independent samples of the variable.

As an example, let X be a random variable and write its expected value as $A = E(X)$. If we can generate $X_1, X_2, \dots, X_n$, n independent random variables with the same distribution, then we can make the approximation,

$$\widehat{A_n} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

By law of large numbers, $\widehat{A_n} \to A$ as $n \to \infty$. The X_i and $\widehat{A_n}$ are random and could be different each time we run the program. Still, the target number A is not random.

Hence a Monte Carlo method is a technique that can be used to solve a mathematical or statistical problem, and a Monte Carlo simulation uses repeated sampling to determine the properties of some phenomenon.

3.2 Stochastic Processes

A **stochastic process** is a family of random variables $\{X(t): t\epsilon\tau\}$ defined on a probability space and indexed by a parameter t where t varies over a set τ. The mapping $t \longmapsto X(t,\omega)$ for each $\omega \in \Omega$ is known as a sample path. One of the main characteristics of stochastic processes is that if multiple experiments were run, different paths would be observed. If the set τ is discrete, then the stochastic process is discrete. If the set τ is continuous, stochastic process is continuous.

3.2.1 Discrete stochastic processes

If the set $\tau = \{t_0, t_1, ...\}$ is a set of discrete times, stochastic process is discrete. Let the sequence of random variables $X(t_0), X(t_1), ...$ be defined on the sample space Ω.

If only the present value $X(t_n) = X_n$ is needed to determine the future value of X_{n+1}, then the sequence $\{X_n\}$ is said to be a **Markov Process**. A discrete valued Markov Process is called a **Markov Chain**.

Let $P(X_{n+1} = x_{n+1} | X_n = x_n)$ be the **one-step transition probability** of a Markov chain. That is, $P(X_{n+1} = x_{n+1}, X_n = x_n) = P(X_{n+1} = x_{n+1} | X_n = x_n) P(X_n = x_n)$.

If the transition probabilities are independent of time t_n, Markov chain is **homogeneous**.

Homogeneous markov chains

Let $\{X_n : n \geq 0\}$ be a homogeneous markov chain defined at discrete times $\tau = \{t_0, t_1, ...\}$. Let X_n be non negative and integer valued for each $t_n, n = 0,1,2, ...$

Let $p_{ij} = P(X_{n+1} = j | X_n = i) \quad i \geq 0, j \geq 0$ be the transition probabilities.

Transition probability matrix is defined as $P = [p_{ij}]$ and $\sum_{j=0}^{\infty} p_{ij} = 1$ for $i = 0,1,2, ...$

Define $P^k = [p_{ij}{}^{(k)}]$.

Since$P^{l+n} = P^l P^n$,

$$\boldsymbol{p_{ij}{}^{(l+n)} = \sum_{m=0}^{\infty} p_{im}{}^{(l)} p_{mj}{}^{(n)} \quad for\ l, n \geq 0} \text{ ,}$$

where $P^0 = \mathrm{I}$.

This relation is called the **Chapman Kolmogorov formula** for a homogeneous Markov chain.

Let $p_i(t_k) = P(X(t_k) = i)$ for $i = 0,1,2, \ldots$be the **probability distribution of $\boldsymbol{X_k}$**. Let

$$\mathrm{p}(t_k) = [p_0(t_k), p_1(t_k), p_2(t_k), \ldots]^T,$$

where$(\mathrm{p}(t_0))_i = P(X(t_0) = i)$is the **initial probability distribution of $\boldsymbol{X(t_0)}$**.

Then,

$$(\mathrm{p}(t_n))^T = \left(\mathrm{p}(t_{n-1})\right)^T P = \left(\mathrm{p}(t_0)\right)^T P^n.$$

Thus, $p_i(t_n) = \sum_{m=0}^{\infty} p_m(t_{n-1})\, p_{mi} = \sum_{m=0}^{\infty} p_m(t_0) {p_{mi}}^{(n)}$.

3.2.2 Continuous stochastic processes

If $\{X(t): t\epsilon\tau\}$ is a stochastic process such that $\tau = [0, T]$ is an interval in time and the process is defined at all instant τ, then the process is a continuous stochastic process.

A continuous time stochastic process is a function such that

$$X: \tau \times \Omega \to \mathbb{R}.$$

$X(t) = X(t,\cdot)$is a random variable for each value of $t \in \tau$. $X(\cdot) = X(\cdot, \omega)$ maps the interval τ into $\mathbb{R}$ and is called a sample path or trajectory.

The stochastic process X is a **Markov Process** if the state of the process at any time $t_n \in \tau$ determines the future state of the process.

Transition probability density function from x at time s to y at time t for a continuous Markov process is given by

$$P[X(s) = x | X(t) = y] = p(y, t, x, s) = \int p(y, t, z, u) p(z, u, x, s) dz.$$

If $p(y, t + u, x, s + u) = p(y, t, x, s)$, $X(t)$ is **homogeneous**.

3.2.3 Wiener process

Robert Brown was a 18th century Botanist and was the first scientist who would observe and document the seemingly random motion of certain particles moving of the surface of water. Brown was initially observing pollen particles under a microscope, and his first thought was that the motion was caused by the particles being alive. He abandoned this theory after observing with dust particles. After this scientist, the mathematician Norbert Wiener defined it in mathematical terms. It is the **Wiener process** or **Brownian motion**.

A Wiener process $\{W(t): t\epsilon[0,T]\}$ is a continuous stochastic process which satisfies following conditions.

a) $W(0) = 0$.
b) For $0 \leq s \leq t \leq T$, the increment $W(t) - W(s)$ is normally distributed with mean 0 and variance $|t - s|$.
c) For $0 \leq s < t < u < v \leq T$, $W(t) - W(s)$ and $W(v) - W(u)$ are independent increments.

Also Wiener process is a homogeneous Markov process.

Generating a Sample Path of a Wiener Process

Suppose that a wiener process trajectory is desired on the interval $[t_0, t_N]$ at the points

$\{t_i: i = 0, 1, \dots, N\}$where $t_0 = 0$. Then $W(t_0) = 0$ and the values of a Wiener process trajectory at the points $t_0, t_1, t_2, \dots, t_N$ is given by

$$\boldsymbol{W(t_i) = W(t_{i-1}) + \eta_{i-1}\sqrt{t_i - t_{i-1}}} ,$$

where $\eta_{i-1} \sim N(0,1)$ for $i = 1, 2, \dots, N$.

Using these $N + 1$ values, the wiener process sample path can be approximated everywhere on$[t_0, t_N]$.

Probability density of normally distributed random variables with mean m and variance $|t|$ is given by

$$p(x,m,t) = \frac{1}{(2\pi|t|)^{1/2}} \exp\left(\frac{-(x-m)^2}{2|t|}\right).$$

Let $W(t)$ be a Wiener process on $[0,T]$.

For $t_1 \in [0,T]$ and $: \mathbb{R} \to \mathbb{R}$,

$$E\big[G\big(W(t_1)\big)\big] = \int_{-\infty}^{\infty} G(x_1) p(x_1, 0, t_1) dx_1.$$

$$P(W(t_1) \le z_1) = \int_{-\infty}^{z_1} p(x_1, 0, t_1) d\, x_1.$$

Let $p(y,t,x,s)$ is the **transition probability** density for the Wiener process from x at time s to y at time t. Then,

$$p(y,t,x,s) = \frac{1}{(2\pi|t-s|)^{1/2}} \mathrm{e}^{\left(\frac{-(x-y)^2}{2|t-s|}\right)}.$$

Since Wiener process is a continuous homogeneous Markov process,

$$p(y,t,x,s) = p(y,x,|t-s|).$$

Chapman Kolmogorov equation for this transition probability is,

$$p(y,t,x,s) = \int_{-\infty}^{\infty} p(y,t,z,u)\, p(z,u,x,s) dz \qquad for\ s < u < t\ .$$

An approximation to the Wiener Process

Consider the interval $0 \le t \le T$. Let $t_i = ih$ where $h = \frac{T}{N} \quad for\ i = 0,1,2, \ldots N$.

Let $W(t) \sim N(0,t)$ be a Wiener process.

Define the continuous linear stochastic process $X_N(t)$ on this partition of [0,T] by ,

$$X_N(t) = W(t_i)\frac{(t_{i+1}-t)}{h} + W(t_{i+1})\frac{(t-t_i)}{h} \text{ for } t_i \le t \le t_{i+1} \text{ and } i = 0,1,2,\ldots,N-1.$$

Then, $X_N(t_i) = W(t_i) \;\; for\ i = 0,1,2, \ldots N$ and $X_N(t)$ is continuous on [0, T].

Also $E(X_N - W)^2 = \sum_{i=0}^{N-1}\int_{t_i}^{t_{i+1}} E[W(t_i)\frac{t_{i+1}-t}{h} + W(t_{i+1})\frac{t-t_i}{h} - W(t)]^2 dt$

$$=\sum_{i=0}^{N-1}\int_{t_i}^{t_{i+1}} E[\big(W(t_i) - W(t)\big)\frac{t_{i+1}-t}{h} + (W(t_{i+1}) - W(t))\frac{t-t_i}{h}]^2 dt$$

$$=\sum_{i=0}^{N-1}\int_{t_i}^{t_{i+1}} \frac{2(t-t_i)(t_{i+1}-t)}{h} dt$$

$$=\sum_{i=0}^{N-1}\frac{h^2}{3} = \frac{T^2}{3N}$$

Thus, $E(X_N - W)^2 \to 0$ as $N \to \infty$.

That is $X_N \to W \;\; as\ N \to \infty$.

Therefore the graph of a Wiener process trajectory is represented by plotting $X_N(t)$ for a large value of N.

Consider the Wiener process $W(t)$ on [0,T].

Let $X(t_{k+1}) = X(t_k) + \eta_k \quad where\ \ t_k = kh\ and\ h = \frac{T}{N}\ for\ k = 0,1,2, \ldots N$

$$X(t_0) = 0\ and\ \ \eta_k \sim N(0,h)$$

Since $W(t_k) = X(t_k) for\ k = 0,1,2, \ldots, N,$ each sample path of $W(t)$ is computed at the discrete times $t_0, t_1, \ldots, t_N$ .To estimate $W(t)$ at $t \neq t_k$ for any k, a linear interpolation is used. In particular,

$$W(t) \approx X(t_k)\frac{t_{k+1}-t}{h} + X(t_{k+1})\frac{t-t_k}{h} \text{ for } t_k \le t \le t_{k+1} \ .$$

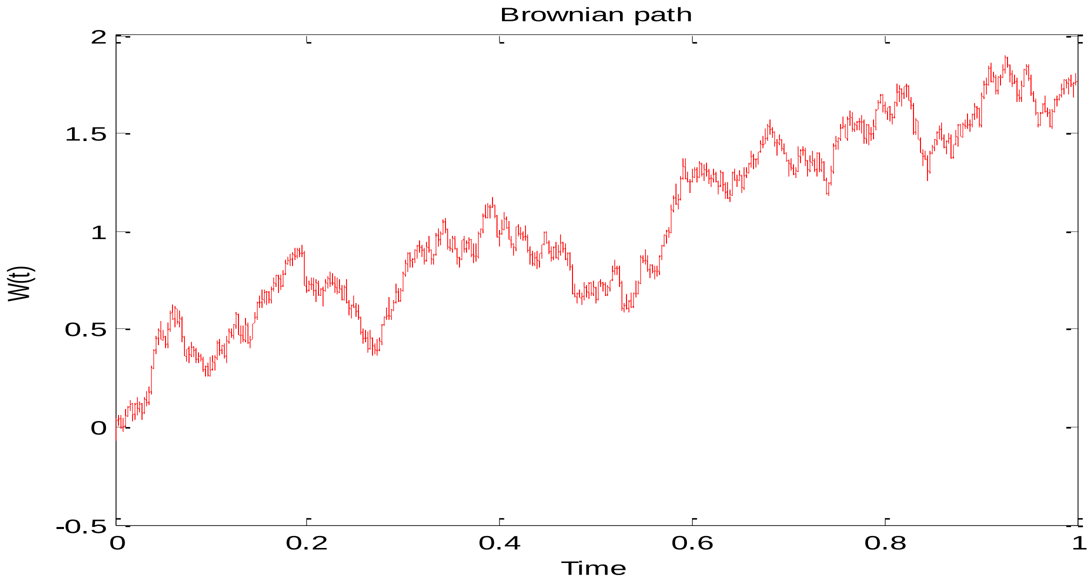


Figure 3.1: A Sample of Brownian path generated by MATLAB

3.3 Stochastic Integral

Consider integral of the form $\int_0^t g(s)dW(s)$, where g is a stochastic process on Ω . Then it can be approximated by

$$\sum_{j=0}^{N-1} g(t_j)[W(t_{j+1}) - W(t_j)]$$

where $t_j = j\left(\frac{T}{N}\right)$, N is the number of steps.

This is the **Ito integral**.

If we approximate $\int_0^t g(s,\omega)dW(s,\omega)$ by

$$\sum_{j=0}^{N-1} g\left(\frac{t_j + t_{j+1}}{2}\right)[W(t_{j+1}) - W(t_j)]$$

It is called the **Stratonovich Integral.**

An Ito integral has following properties.

a) $\int_S^T f dW(t) = \int_S^U f dW(t) + \int_U^T f dW(t)$ for $0 \leq S \leq U \leq T$.

b) $\int_S^T (cf + g) dW(t) = c \int_S^T f dW(t) + \int_S^T g dW(t)$

c) $E\left[\int_S^T f dW(t)\right] = 0$

d) $E\left[\left|\int_S^T f dW(t)\right|^2\right] = \int_S^T E|f(t)|^2 dt$

An **Ito Process** is a stochastic process X_t on $(\Omega, \mathcal{A}, P)$of the form

$$\boldsymbol{X_t = X_0 + \int_0^t u(s,\omega) ds + \int_0^t v(s,\omega) dW(s)}.$$

Sometimes it can be written as differential form

$$\boldsymbol{dX_t = udt + vdW(t)}.$$

Let X_t be an Ito process given by,

$$dX_t = udt + vdW(t).$$

Let $g(t, x)$ be a twice continuously differentiable function.

Then,

$$Y_t = g(t, X_t)$$

is again an Ito process and

$$dY_t = \frac{\partial g}{\partial t}(t, X_t)dt + \frac{\partial g}{\partial x}(t, X_t)dt + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t, X_t).$$

where$(dX_t)^2 = dX_t . dX_t$ is computed according to the rules

$$dt. dt = dt. dW(t) = dW(t). dt = 0 \ , \ dW(t). dW(t) = dt.$$

3.4 Stochastic Differential Equations

An Ito stochastic differential equation on [0, T] has the form

$$\boldsymbol{dX(t) = f\big(t, X(t)\big)dt + g\big(t, X(t)\big)dW(t)} \text{ for } 0 \le t \le T$$

Function f is called the drift coefficient and function g is called the diffusion coefficient of the stochastic differential equation. $W(t)$ is an independent Wiener process. It is assumed that the functions f and g are none anticipating and satisfy following conditions for some constant $k \geq 0$.

a) $|f(t,x) - f(s,y)|^2 \le k(|t-s| + |x-y|^2)$ for $0 \le s, t \le T$ and $x, y \in \mathbb{R}$.
b) $|f(t,x)|^2 \le k(1 + |x|^2)$ for $0 \le t \le T$ and $x \in \mathbb{R}$.

Integral form of the Ito stochastic differential equation is,

$$\boldsymbol{X(t) = X(0) + \int_0^t f\big(X(s)\big)ds + \int_0^t g\big(X(s)\big)dW(s)}.$$

The exact solution of a stochastic differential equation is generally difficult to obtain. But can be approximated using numerical methods.

a) Euler-Maruyama Method

Consider the stochastic differential equation

$$\boldsymbol{dX(t) = f\big(t, X(t)\big)dt + g\big(t, X(t)\big)dW(t)} \text{ for } 0 \le t \le T$$

Let $\Delta t = \frac{T}{N}$ for some integer N and $t_j = j\Delta t$ for $j = 1,2, \dots, N$.

Then,

$$X\big(t_j\big) = X(0) + \int_0^{t_j} f\big(X(S)\big)ds + \int_0^{t_j} g\big(X(s)\big)dW(s)$$

$$X\big(t_{j-1}\big) = X(0) + \int_0^{t_{j-1}} f\big(X(S)\big)ds + \int_0^{t_{j-1}} g\big(X(s)\big)dW(s)$$

Subtracting above two equations, we obtain

$$X(t_j) = X(t_{j-1}) + \int_{t_{j-1}}^{t_j} f(X(S))ds + \int_{t_{j-1}}^{t_j} g(X(s))dW(s). \quad \textbf{3.4.1}$$

For the first integral

$$\int_{t_{j-1}}^{t_j} f(X(S))ds \approx f(X(S))(t_j - t_{j-1}) = f(X(S))\Delta t. \quad \textbf{3.4.2}$$

For the second integral

$$\int_{t_{j-1}}^{t_j} g(X(s))dW(s) \approx g(X(s))\left(W(t_j) - W(t_{j-1})\right). \quad \textbf{3.4.3}$$

By substituting 3.4.2 and 3.4.3 to 3.4.1, we can obtain

$$\boldsymbol{X_j = X_{j-1} + f(X_{j-1})\Delta t + g(X_{j-1})(W(t_j) - W(t_{j-1}))} \text{ for } j = 1, 2, \dots, N.$$

where, $X_j = X(t_j)$,$\Delta t = \frac{T}{N}$ and $(\mathrm{W}(\mathrm{t_j}) - \mathrm{W}(\mathrm{t_{j-1}})) \sim \mathrm{N}(0, \Delta t)$.

b) Milstein's Method

Milstein method is a popular second order method. It has mean square error proportional to $(\Delta t)^2$ rather than Δt. Milstein method has the form

$$\boldsymbol{X_j = X_{j-1} + f(X_{j-1})\Delta t + g(X_{j-1})\Delta W_j + \frac{1}{2} g(X_{j-1}) \frac{\partial g(X_{j-1})}{\partial x}\left[(\Delta W_j)^2 - \Delta t\right]}$$

3.4.1 SDE models in finance

We now explore the solution of four SDE models used in this research. Let $X(t)$ be the gold price of Sri Lanka at time t.

a) Brownian Motion

A Brownian motion $X(t)$ is the solution of an SDE with constant drift and diffusion coefficients

$$\boldsymbol{dX(t) = \alpha dt + \sigma dW(t)}$$

where α and σ are parameters to be determined with initial value $X(0) = x_0$.

b) Geometric Brownian Motion

A geometric Brownian motion $X(t)$ is the solution of an SDE with linear drift and diffusion coefficients.

$$\boldsymbol{dX(t) = \beta X(t) dt + \sigma X(t) dW(t)}$$

where β and σ are parameters to be determined with initial value $X(0) = x_0$.

c) Cox- Ingersol- Ross Model

In the Cox–Ingersoll–Ross model, briefly CIR model, $X(t)$ is assumed to satisfy the stochastic differential equation

$$\boldsymbol{dX(t) = (\alpha + \beta X(t)) dt + \sigma \sqrt{X(t)} dW(t)}$$

where α, β and σ are parameters to be determined with initial value $X(0) = x_0$.

d) Vasicek Model

In the Vasicek model $X(t)$ is assumed to satisfy the stochastic differential equation

$$\boldsymbol{dX(t) = (\alpha + \beta X(t)) dt + \sigma dW(t)}$$

where α, β and σ are parameters to be determined with initial value $X(0) = x_0$.

3.4.2 Parameter estimation of SDEs

In this section, a stochastic differential equation of the form

$$dX(t) = f(t, X(t); \theta)\, dt + g(t, X(t); \theta)\, dW(t)$$

is considered where $\theta \in \mathbb{R}^m$ is a vector of parameters that are unknown. It is assumed that

$$x_0, x_1, x_2, \dots, x_N$$

are observed values of $X(t)$ at the respective uniformly distributed times $t_i = i\Delta t$ for $i = 0, 1, \dots, N$ where $\Delta t = T/N$. The problem is to find an estimate of the vector θ given these $N + 1$ data points. In this research**, maximum likelihood estimation method** is considered.

Let $p(t_k, x_k | t_{k-1}, x_{k-1};\ \theta)$ be the transition probability density of (t_k, x_k) starting from (t_{k-1}, x_{k-1}) given the vector θ. Suppose that the density of the initial state is $p_0(x0|\theta)$.

In maximum likelihood estimation of θ, the joint density

$$\boldsymbol{D(\theta) = p_0(x_0|\theta) \prod_{k=1}^{N} p\left(t_k, x_k \middle| t_{k-1}, x_{k-1};\ \theta\right)}$$

is maximized over $\theta \in \mathbb{R}^m$. The value of θ that maximizes $D(\theta)$ will be denoted as θ^*. It is more convenient to minimize the function

$$\boldsymbol{L(\theta) = -ln\,(D(\theta))}$$

which has the form

$$\boldsymbol{L(\theta) = -\ln\left(p_0(x_0|\theta)\right) - \sum_{k=1}^{N} \ln\left(p\left(t_k, x_k \middle| t_{k-1}, x_{k-1};\ \theta\right)\right)}$$

One difficulty in finding the optimal value θ^* is that the transition densities are not generally known. However, by considering the Euler approximation and letting $X(t_{k-1}) = x_{k-1}$ at

$t = t_{k-1}$.

$$\boldsymbol{X(t_k) \approx x_{k-1} + f(t_{k-1}, x_{k-1};\ \theta)\Delta t + g(t_{k-1}, x_{k-1};\ \theta)\sqrt{\Delta t}\,\eta_k}$$

where $\eta_k \sim N(0, \Delta t)$.

This implies that

$$\boldsymbol{p\left(t_k, x_k \middle| t_{k-1}, x_{k-1};\ \theta\right) = \frac{1}{\sqrt{2\pi\sigma_k{}^2}} exp\left(\frac{-(x_k-\mu_k)^2}{2\sigma_k{}^2}\right)}$$

where $\mu_k = x_{k-1} + f(t_{k-1}, x_{k-1};\ \theta)\Delta t$ and $\sigma_k = g(t_{k-1}, x_{k-1};\ \theta)\sqrt{\Delta t}$.

This transition density can be substituted into the expression for $L(\theta)$ which can subsequently be minimized over $\mathbb{R}^m$.

In this research study, we considered four SDE models Brownian motion, Geometric Brownian motion, CIR model and Vasicek model. We need to estimate the parameters of four SDE models using Maximum likelihood estimation method.

First consider the Brownian motion.

Let α_{BM} and σ_{BM} be the parameters of Brownian motion model. By considering the Euler- Maruyama approximation, the gold price at time t can be approximated using the equation,

$$X(t) = X(t-1) + \alpha_{BM}\Delta t + \sigma_{BM}\sqrt{\Delta t}\eta_t$$

where $X(t)$ is the gold price at time t, $\Delta t = {}^{1}/_{252}$ and $\eta_t \sim N(0, \Delta t)$.

This implies that

$$p(t_k, x_k | t_{k-1}, x_{k-1};\ \theta) = \frac{1}{\sqrt{2\pi{\sigma_k}^2}} exp\left(\frac{-(x_k - \mu_k)^2}{2{\sigma_k}^2}\right)$$

where $\mu_k = \ x_{k-1}\ + \alpha_{BM}\Delta t$ and $\sigma_k = \sigma_{BM}\sqrt{\Delta t}$.

By substituting this transition density to the equation $L(\theta)$,

$$L(\theta) = -\sum_{k=1}^{N} \ln\left(\frac{1}{\sqrt{2\pi\left(\sigma_{BM}\sqrt{\Delta t}\right)^2}} exp\left(\frac{-(x_k - x_{k-1} - \alpha_{BM}\Delta t)^2}{2\left(\sigma_{BM}\sqrt{\Delta t}\right)^2}\right)\right)$$

$$= -\sum_{k=1}^{N} \ln\left(\frac{1}{\sqrt{2\pi\left(\sigma_{BM}\sqrt{\Delta t}\right)^2}}\right) - \sum_{k=1}^{N}\left(\frac{-(x_k - x_{k-1} - \alpha_{BM}\Delta t)^2}{2\left(\sigma_{BM}\sqrt{\Delta t}\right)^2}\right)$$

$$= -N \ln\left(\frac{1}{\sqrt{2\pi\left(\sigma_{BM}\sqrt{\Delta t}\right)^2}}\right) + \frac{1}{2\left(\sigma_{BM}\sqrt{\Delta t}\right)^2}\sum_{k=1}^{N}(x_k - x_{k-1} - \alpha_{BM}\Delta t)^2.$$

To minimize $L(\theta)$, differentiate with respect to α_{BM} and σ_{BM} .

By differentiating with respect to α_{BM},

$$\frac{\partial L}{\partial \alpha_{BM}} = \frac{1}{2\left(\sigma_{BM}\sqrt{\Delta t}\right)^2} \times \sum_{k=1}^{N} 2(x_k - x_{k-1} - \alpha_{BM}\Delta t) \times -\Delta t$$

$$= \frac{1}{(\sigma_{BM})^2} \sum_{k=1}^{N} (x_k - x_{k-1} - \alpha_{BM}\Delta t)$$

At the optimal value of α_{BM}, $\frac{\partial L}{\partial \alpha_{BM}} = 0$.

$$\left.\frac{\partial L}{\partial \alpha_{BM}}\right|_{\alpha_{BM}=\hat{\alpha}_{BM}} = 0$$

$$\frac{1}{(\sigma_{BM})^2} \sum_{k=1}^{N} (x_k - x_{k-1} - \hat{\alpha}_{BM}\Delta t) = 0$$

$$\sum_{k=1}^{N} (x_k - x_{k-1} - \hat{\alpha}_{BM}\Delta t) = 0$$

$$\hat{\alpha}_{BM} = \sum_{k=1}^{N} \frac{(x_k - x_{k-1})}{N\Delta t} = \frac{(x_N - x_0)}{N\Delta t}$$

By differentiating $L(\theta)$ with respect to σ_{BM},

$$\frac{\partial L}{\partial \sigma_{BM}} = -N \ln\left(\frac{1}{\sqrt{2\pi\left(\sigma_{BM}\sqrt{\Delta t}\right)^2}}\right) + \frac{1}{2\left(\sigma_{BM}\sqrt{\Delta t}\right)^2} \sum_{k=1}^{N} (x_k - x_{k-1} - \alpha_{BM}\Delta t)^2$$

$$= -N\sqrt{2\pi\left(\sigma_{BM}\sqrt{\Delta t}\right)^2} \times \frac{1}{\sqrt{2\pi\Delta t}} \times -\frac{1}{(\sigma_{BM})^2} - \frac{2}{2\Delta t(\sigma_{BM})^3} \sum_{k=1}^{N} (x_k - x_{k-1} - \alpha_{BM}\Delta t)^2$$

$$= \frac{N}{\sigma_{BM}} - \frac{1}{\Delta t(\sigma_{BM})^3}\sum_{k=1}^{N}(x_k - x_{k-1} - \alpha_{BM}\Delta t)^2$$

$$= \frac{N\Delta t(\sigma_{BM})^2 - \sum_{k=1}^{N}(x_k - x_{k-1} - \alpha_{BM}\Delta t)^2}{\Delta t(\sigma_{BM})^3}$$

At the optimal value of σ_{BM}, $\frac{\partial L}{\partial \sigma_{BM}} = 0$.

$$\left.\frac{\partial L}{\partial \sigma_{BM}}\right|_{\sigma_{BM}=\hat{\sigma}_{BM}} = 0$$

$$\frac{N\Delta t(\hat{\sigma}_{BM})^2 - \sum_{k=1}^{N}(x_k - x_{k-1} - \hat{\alpha}_{BM}\Delta t)^2}{\Delta t(\hat{\sigma}_{BM})^3} = 0$$

$$N\Delta t(\hat{\sigma}_{BM})^2 - \sum_{k=1}^{N}(x_k - x_{k-1} - \hat{\alpha}_{BM}\Delta t)^2 = 0$$

$$\hat{\sigma}_{BM}{}^2 = \frac{\sum_{k=1}^{N}(x_k - x_{k-1} - \hat{\alpha}_{BM}\Delta t)^2}{N\Delta t}$$

Hence, the parameters for Brownian motion can be obtained from the equations:

$$\hat{\alpha}_{BM} = \frac{(x_N - x_0)}{N\Delta t}$$

and

$$\hat{\sigma}_{BM}{}^2 = \frac{\sum_{k=1}^{N}(x_k - x_{k-1} - \hat{\alpha}_{BM}\Delta t)^2}{N\Delta t}$$

Let $\hat{\beta}_{GBM}$ and σ_{GBM} be the parameters of Geometric Brownian motion model. By considering the Euler- Maruyama approximation, the gold price at time t can be approximated using the equation,

$$X(t) = \beta_{GBM}X(t-1)\Delta t + \sigma_{GBM}X(t-1)\sqrt{\Delta t}\eta_t$$

where $X(t)$ is the gold price at time t, $\Delta t = {}^{1}/_{252}$ and $\eta_t \sim N(0, \Delta t)$.

This implies that

$$p(t_k, x_k | t_{k-1}, x_{k-1};\ \theta) = \frac{1}{\sqrt{2\pi{\sigma_k}^2}} exp\left(\frac{-(x_k - \mu_k)^2}{2{\sigma_k}^2}\right)$$

where $\mu_k = \ x_{k-1} \ + \beta_{GBM} x_{k-1} \Delta t$ and $\sigma_k = \sigma_{GBM} x_{k-1} \sqrt{\Delta t}$.

By substituting this transition density to the equation $L(\theta)$,

$$L(\theta) = -\sum_{k=1}^{N} \ln\left(\frac{1}{\sqrt{2\pi\left(\sigma_{GBM} x_{k-1}\sqrt{\Delta t}\right)^2}} exp\left(\frac{-(x_k - x_{k-1} - \beta_{GBM} x_{k-1}\Delta t)^2}{2\left(\sigma_{GBM} x_{k-1}\sqrt{\Delta t}\right)^2}\right)\right)$$

$$= -\sum_{k=1}^{N} \ln\left(\frac{1}{\sqrt{2\pi\left(\sigma_{GBM} x_{k-1}\sqrt{\Delta t}\right)^2}}\right) - \sum_{k=1}^{N}\left(\frac{-(x_k - x_{k-1} - \beta_{GBM} x_{k-1}\Delta t)^2}{2\left(\sigma_{GBM} x_{k-1}\sqrt{\Delta t}\right)^2}\right)$$

$$= -\sum_{k=1}^{N} \ln\left(\frac{1}{\sqrt{2\pi\left(\sigma_{GBM} x_{k-1}\sqrt{\Delta t}\right)^2}}\right) + \frac{1}{2\left(\sigma_{GBM}\sqrt{\Delta t}\right)^2}\sum_{k=1}^{N}\frac{(x_k - x_{k-1} - \beta_{GBM} x_{k-1}\Delta t)^2}{{x_{k-1}}^2}$$

To minimize $L(\theta)$, differentiate with respect to β_{GBM} and σ_{GBM} .

By differentiating with respect to β_{GBM},

$$\frac{\partial L}{\partial \beta_{GBM}} = \frac{1}{2\left(\sigma_{GBM}\sqrt{\Delta t}\right)^2} \times \sum_{k=1}^{N} \frac{2(x_k - x_{k-1} - \beta_{GBM} x_{k-1}\Delta t) \times -x_{k-1}\Delta t}{{x_{k-1}}^2}$$

$$= \frac{-1}{(\sigma_{GBM})^2}\sum_{k=1}^{N}\frac{(x_k - x_{k-1} - \beta_{GBM} x_{k-1}\Delta t) x_{k-1}}{{x_{k-1}}^2}.$$

At the optimal value of β_{GBM}, $\frac{\partial L}{\partial \beta_{GBM}} = 0$.

$$\left.\frac{\partial L}{\partial \beta_{GBM}}\right|_{\beta_{GBM}=\hat{\beta}_{GBM}} = 0$$

$$\frac{-1}{(\sigma_{GBM})^2}\sum_{k=1}^{N}\frac{(x_k - x_{k-1} - \beta_{GBM}x_{k-1}\Delta t)x_{k-1}}{{x_{k-1}}^2} = 0$$

$$\sum_{k=1}^{N}\frac{(x_k - x_{k-1} - \beta_{GBM}x_{k-1}\Delta t)x_{k-1}}{{x_{k-1}}^2} = 0$$

$$\sum_{k=1}^{N}\frac{(x_k - x_{k-1})}{x_{k-1}} - \hat{\beta}_{GBM}N\,\Delta t = 0$$

$$\hat{\beta}_{GBM} = \frac{1}{N\Delta t}\sum_{k=1}^{N}\frac{(x_k - x_{k-1})}{x_{k-1}}$$

By differentiating $L(\theta)$ with respect to σ_{GBM},

$$\frac{\partial L}{\partial \sigma_{GBM}} = -\sum_{k=1}^{N}\sqrt{2\pi\left(\sigma_{GBM}x_{k-1}\sqrt{\Delta t}\right)^2}\frac{1}{\sqrt{2\pi\Delta t}x_{k-1}} \times -\frac{1}{{\sigma_{GBM}}^2} + \frac{-2}{2\Delta t(\sigma_{GBM})^3}\sum_{k=1}^{N}\frac{(x_k - x_{k-1} - \beta_{GBM}x_{k-1}\Delta t)^2}{{x_{k-1}}^2}$$

$$= -\sum_{k=1}^{N}\sigma_{GBM} \times -\frac{1}{{\sigma_{GBM}}^2} - \frac{1}{\Delta t(\sigma_{GBM})^3}\sum_{k=1}^{N}\frac{(x_k - x_{k-1} - \beta_{GBM}x_{k-1}\Delta t)^2}{{x_{k-1}}^2}$$

$$= \frac{N}{\sigma_{GBM}} - \frac{1}{\Delta t(\sigma_{GBM})^3}\sum_{k=1}^{N}\frac{(x_k - x_{k-1} - \beta_{GBM}x_{k-1}\Delta t)^2}{{x_{k-1}}^2}$$

$$= \frac{N\Delta t(\sigma_{GBM})^2 - \sum_{k=1}^{N}\frac{(x_k - x_{k-1} - \beta_{GBM}x_{k-1}\Delta t)^2}{{x_{k-1}}^2}}{\Delta t(\sigma_{GBM})^3}.$$

At the optimal value of σ_{GBM}, $\frac{\partial L}{\partial \sigma_{GBM}} = 0$.

$$\left.\frac{\partial L}{\partial \sigma_{GBM}}\right|_{\sigma_{BM}=\hat{\sigma}_{GBM}} = 0$$

$$\frac{N\Delta t(\hat{\sigma}_{GBM})^2 - \sum_{k=1}^{N} \frac{\left(x_k - x_{k-1} - \hat{\beta}_{GBM} x_{k-1}\Delta t\right)^2}{{x_{k-1}}^2}}{\Delta t(\sigma_{GBM})^3} = 0$$

$$N\Delta t{\hat{\sigma}_{GBM}}^2 - \sum_{k=1}^{N} \frac{\left(x_k - x_{k-1} - \hat{\beta}_{GBM} x_{k-1}\Delta t\right)^2}{{x_{k-1}}^2} = 0$$

$${\hat{\sigma}_{GBM}}^2 = \frac{\sum_{k=1}^{N} \frac{\left(x_k - x_{k-1} - \hat{\beta}_{GBM} x_{k-1}\Delta t\right)^2}{{x_{k-1}}^2}}{N\Delta t}$$

$${\hat{\sigma}_{GBM}}^2 = \frac{1}{N\Delta t} \sum_{k=1}^{N} \left(\frac{x_k - \left(1 + \hat{\beta}_{GBM}\Delta t\right) x_{k-1}}{x_{k-1}}\right)^2$$

Hence, the parameters for Geometric Brownian motion can be obtained from the equations:

$$\hat{\beta}_{GBM} = \frac{1}{N\Delta t} \sum_{k=1}^{N} \frac{(x_k - x_{k-1})}{x_{k-1}}$$

and

$${\hat{\sigma}_{GBM}}^2 = \frac{1}{N\Delta t} \sum_{k=1}^{N} \left(\frac{x_k - \left(1 + \hat{\beta}_{GBM}\Delta t\right) x_{k-1}}{x_{k-1}}\right)^2$$

Then consider the CIR model.

Let $\alpha_{CIR}, \beta_{CIR}$ and σ_{CIR} be the parameters to be determined.

By considering the Euler- Maruyama approximation, the gold price at time t can be approximated using the equation,

$$X(t) = \big(\alpha_{CIR} + \beta_{CIR} X(t-1)\big)\Delta t + \sigma_{CIR}\sqrt{X(t-1)}\sqrt{\Delta t}\eta_t$$

where $X(t)$ is the gold price at time t, $\Delta t = {}^{1}/_{252}$ and $\eta_t \sim N(0, \Delta t)$.

This implies that

$$p(t_k, x_k | t_{k-1}, x_{k-1};\ \theta) = \frac{1}{\sqrt{2\pi{\sigma_k}^2}} exp\left(\frac{-(x_k - \mu_k)^2}{2{\sigma_k}^2}\right)$$

where $\mu_k = \ x_{k-1}\ + (\alpha_{CIR} + \beta_{CIR} x_{k-1})\mathit{\Delta}t$ and $\sigma_k = \sigma_{CIR}\sqrt{x_{k-1}\Delta t}$.

By substituting this transition density to the equation $L(\theta)$,

$$L(\theta)$$

$$= -\sum_{k=1}^{N} \ln\left(\frac{1}{\sqrt{2\pi\left(\sigma_{CIR}\sqrt{x_{k-1}\Delta t}\right)^2}} exp\left(\frac{-(x_k - x_{k-1} - (\alpha_{CIR} + \beta_{CIR} x_{k-1})\mathit{\Delta}t\ )^2}{2\left(\sigma_{CIR}\sqrt{x_{k-1}\Delta t}\right)^2}\right)\right)$$

$$= -\sum_{k=1}^{N} \ln\left(\frac{1}{\sqrt{2\pi\left(\sigma_{CIR}\sqrt{x_{k-1}\Delta t}\right)^2}}\right) - \sum_{k=1}^{N}\left(\frac{-(x_k - x_{k-1} - (\alpha_{CIR} + \beta_{CIR} x_{k-1})\mathit{\Delta}t\ )^2}{2\left(\sigma_{CIR}\sqrt{x_{k-1}\Delta t}\right)^2}\right)$$

$$= -\sum_{k=1}^{N} \ln\left(\frac{1}{\sqrt{2\pi\left(\sigma_{CIR}\sqrt{x_{k-1}\Delta t}\right)^2}}\right)$$

$$+ \frac{1}{2\left(\sigma_{CIR}\sqrt{\mathit{\Delta}t}\right)^2}\sum_{k=1}^{N}\frac{(x_k - x_{k-1} - (\alpha_{CIR} + \beta_{CIR} x_{k-1})\mathit{\Delta}t\ )^2}{x_{k-1}}$$

To minimize $L(\theta)$, differentiate with respect to $\alpha_{CIR}, \beta_{CIR}$ and σ_{CIR} .

By differentiating with respect to α_{CIR},

$$\frac{\partial L}{\partial \alpha_{CIR}} = \frac{1}{2\left(\sigma_{CIR}\sqrt{\Delta t}\right)^2} \times \sum_{k=1}^{N} \frac{2(x_k - x_{k-1} - (\alpha_{CIR} + \beta_{CIR} x_{k-1})\Delta t) \times -\Delta t}{x_{k-1}}$$

$$= \frac{-1}{(\sigma_{CIR})^2} \sum_{k=1}^{N} \frac{(x_k - x_{k-1} - (\alpha_{CIR} + \beta_{CIR} x_{k-1})\Delta t)}{x_{k-1}}$$

At the optimal value of α_{CIR}, $\frac{\partial L}{\partial \alpha_{CIR}} = 0$.

$$\left.\frac{\partial L}{\partial \alpha_{CIR}}\right|_{\alpha_{CIR} = \hat{\alpha}_{CIR}} = 0$$

$$\frac{-1}{(\sigma_{CIR})^2} \sum_{k=1}^{N} \frac{\left(x_k - x_{k-1} - \left(\hat{\alpha}_{CIR} + \hat{\beta}_{CIR} x_{k-1}\right)\Delta t\right)}{x_{k-1}} = 0$$

$$\sum_{k=1}^{N} \frac{\left(x_k - x_{k-1} - \left(\hat{\alpha}_{CIR} + \hat{\beta}_{CIR} x_{k-1}\right)\Delta t\right)}{x_{k-1}} = 0$$

$$\sum_{k=1}^{N} \frac{(x_k - x_{k-1})}{x_{k-1}} - \hat{\alpha}_{CIR}\Delta t \sum_{k=1}^{N} \frac{1}{x_{k-1}} - \hat{\beta}_{CIR} N\, \Delta t = 0$$

$$\hat{\alpha}_{CIR}\Delta t \sum_{k=1}^{N} \frac{1}{x_{k-1}} + \hat{\beta}_{CIR} N\, \Delta t = \sum_{k=1}^{N} \frac{(x_k - x_{k-1})}{x_{k-1}} \qquad \textbf{3.4.4}$$

By differentiating with respect to β_{CIR},

$$\frac{\partial L}{\partial \beta_{CIR}} = \frac{1}{2\left(\sigma_{CIR}\sqrt{\Delta t}\right)^2} \times \sum_{k=1}^{N} \frac{2(x_k - x_{k-1} - (\alpha_{CIR} + \beta_{CIR} x_{k-1})\Delta t) \times -x_{k-1}\Delta t}{x_{k-1}}$$

$$= \frac{-1}{(\sigma_{CIR})^2} \sum_{k=1}^{N} (x_k - x_{k-1} - (\alpha_{CIR} + \beta_{CIR} x_{k-1}) \Delta t)$$

At the optimal value of β_{CIR}, $\frac{\partial L}{\partial \beta_{CIR}} = 0$.

$$\left. \frac{\partial L}{\partial \beta_{CIR}} \right|_{\beta_{CIR} = \hat{\beta}_{CIR}} = 0$$

$$\frac{-1}{(\sigma_{CIR})^2} \sum_{k=1}^{N} (x_k - x_{k-1} - (\alpha_{CIR} + \beta_{CIR} x_{k-1}) \Delta t) = 0$$

$$\sum_{k=1}^{N} (x_k - x_{k-1} - (\alpha_{CIR} + \beta_{CIR} x_{k-1}) \Delta t) = 0$$

$$x_N - x_0 - \hat{\alpha}_{CIR} N \Delta t - \hat{\beta}_{CIR} \Delta t \sum_{k=1}^{N} x_{k-1} = 0$$

$$\hat{\alpha}_{CIR} N \Delta t + \hat{\beta}_{CIR} \Delta t \textstyle\sum_{k=1}^{N} x_{k-1} = x_N - x_0 \qquad \textbf{3.4.5}$$

Considering equations 3.4.4 and 3.4.5,

$$\frac{x_N - x_0 - \hat{\alpha}_{CIR} N \Delta t}{\Delta t \sum_{k=1}^{N} x_{k-1}} = \frac{\sum_{k=1}^{N} \frac{(x_k - x_{k-1})}{x_{k-1}} - \hat{\alpha}_{CIR} \Delta t \sum_{k=1}^{N} \frac{1}{x_{k-1}}}{N \Delta t}$$

$$\hat{\alpha}_{CIR} \left(\frac{N}{\sum_{k=1}^{N} x_{k-1}} - \frac{\sum_{k=1}^{N} \frac{1}{x_{k-1}}}{N} \right) = \frac{x_N - x_0}{\Delta t \sum_{k=1}^{N} x_{k-1}} - \frac{\sum_{k=1}^{N} \frac{(x_k - x_{k-1})}{x_{k-1}}}{N \Delta t}$$

$$\hat{\alpha}_{CIR} \frac{N^2 - \sum_{k=1}^{N} x_{k-1} \sum_{k=1}^{N} \frac{1}{x_{k-1}}}{N \sum_{k=1}^{N} x_{k-1}} = \frac{N(x_N - x_0) - \sum_{k=1}^{N} x_{k-1} \sum_{k=1}^{N} \frac{(x_k - x_{k-1})}{x_{k-1}}}{N \Delta t \sum_{k=1}^{N} x_{k-1}}$$

$$\hat{\alpha}_{CIR} = \frac{N(x_N - x_0) - \sum_{k=1}^{N} x_{k-1} \sum_{k=1}^{N} \frac{(x_k - x_{k-1})}{x_{k-1}}}{\Delta t\left(N^2 - \sum_{k=1}^{N} x_{k-1} \sum_{k=1}^{N} \frac{1}{x_{k-1}}\right)}$$

Substituting $\hat{\alpha}_{CIR}$ to the equation 3.4.4,

$$\hat{\beta}_{CIR} = \frac{(x_N - x_0) - N\hat{\alpha}_{CIR}\Delta t}{\Delta t \sum_{k=1}^{N} x_{k-1}}$$

By differentiating $L(\theta)$ with respect to σ_{CIR},

$$\frac{\partial L}{\partial \sigma_{CIR}} = -\sum_{k=1}^{N} \sqrt{2\pi\left(\sigma_{CIR}\sqrt{x_{k-1}\Delta t}\right)^2} \frac{1}{\sqrt{2\pi\Delta t x_{k-1}}} \times -\frac{1}{{\sigma_{CIR}}^2} + \frac{-2}{2\Delta t(\sigma_{CIR})^3} \sum_{k=1}^{N} \frac{(x_k - x_{k-1} - (\alpha_{CIR} + \beta_{CIR} x_{k-1})\Delta t\,)^2}{x_{k-1}}$$

$$= -\sum_{k=1}^{N} \sigma_{CIR} \times -\frac{1}{{\sigma_{CIR}}^2} - \frac{1}{\Delta t(\sigma_{CIR})^3} \sum_{k=1}^{N} \frac{(x_k - x_{k-1} - (\alpha_{CIR} + \beta_{CIR} x_{k-1})\Delta t\,)^2}{x_{k-1}}$$

$$= \frac{N}{\sigma_{CIR}} - \frac{1}{\Delta t(\sigma_{CIR})^3} \sum_{k=1}^{N} \frac{(x_k - x_{k-1} - (\alpha_{CIR} + \beta_{CIR} x_{k-1})\Delta t\,)^2}{x_{k-1}}$$

$$= \frac{N\Delta t(\sigma_{CIR})^2 - \sum_{k=1}^{N} \frac{(x_k - x_{k-1} - (\alpha_{CIR} + \beta_{CIR} x_{k-1})\Delta t\,)^2}{x_{k-1}}}{\Delta t(\sigma_{CIR})^3}$$

At the optimal value of σ_{CIR}, $\frac{\partial L}{\partial \sigma_{CIR}} = 0$.

$$\left.\frac{\partial L}{\partial \sigma_{CIR}}\right|_{\sigma_{CIR} = \hat{\sigma}_{CIR}} = 0$$

$$\frac{N\Delta t(\sigma_{CIR})^2 - \sum_{k=1}^{N} \frac{(x_k - x_{k-1} - (\alpha_{CIR} + \beta_{CIR} x_{k-1})\Delta t\,)^2}{x_{k-1}}}{\Delta t(\sigma_{CIR})^3} = 0$$

$$N\Delta t(\sigma_{CIR})^2 - \sum_{k=1}^{N} \frac{(x_k - x_{k-1} - (\alpha_{CIR} + \beta_{CIR} x_{k-1})\Delta t\,)^2}{x_{k-1}} = 0$$

$$\hat{\sigma}_{CIR}{}^2 = \frac{\sum_{k=1}^{N} \frac{(x_k - x_{k-1} - (\alpha_{CIR} + \beta_{CIR} x_{k-1})\Delta t\,)^2}{x_{k-1}}}{N\Delta t}$$

$$\hat{\sigma}_{CIR}{}^2 = \frac{1}{N\Delta t} \sum_{k=1}^{N} \frac{(x_k - x_{k-1} - (\alpha_{CIR} + \beta_{CIR} x_{k-1})\Delta t\,)^2}{x_{k-1}}$$

Hence, the parameters for the CIR model can be obtained from the equations:

$$\hat{\alpha}_{CIR} = \frac{N(x_N - x_0) - \sum_{k=1}^{N} x_{k-1} \sum_{k=1}^{N} \frac{(x_k - x_{k-1})}{x_{k-1}}}{\Delta t\left(N^2 - \sum_{k=1}^{N} x_{k-1} \sum_{k=1}^{N} \frac{1}{x_{k-1}}\right)}$$

$$\hat{\beta}_{CIR} = \frac{(x_N - x_0) - N\hat{\alpha}_{CIR}\Delta t}{\Delta t \sum_{k=1}^{N} x_{k-1}}$$

and

$$\hat{\sigma}_{CIR}{}^2 = \frac{1}{N\Delta t} \sum_{k=1}^{N} \frac{(x_k - x_{k-1} - (\alpha_{CIR} + \beta_{CIR} x_{k-1})\Delta t\,)^2}{x_{k-1}}.$$

Finally consider the Vasicek model.

Let $\alpha_{Vas}, \beta_{Vas}$ and σ_{Vas} be the parameters to be determined.

By considering the Euler- Maruyama approximation, the gold price at time t can be approximated using the equation,

$$X(t) = \big(\alpha_{Vas} + \beta_{Vas} X(t-1)\big)\Delta t + \sigma_{Vas}\sqrt{\Delta t}\,\eta_t$$

where $X(t)$ is the gold price at time t, $\Delta t = {}^{1}\!/_{252}$ and $\eta_t \sim N(0, \Delta t)$.

This implies that

$$p(t_k, x_k | t_{k-1}, x_{k-1};\ \theta) = \frac{1}{\sqrt{2\pi{\sigma_k}^2}} exp\left(\frac{-(x_k - \mu_k)^2}{2{\sigma_k}^2}\right)$$

where $\mu_k = \ x_{k-1}\ + (\alpha_{Vas} + \beta_{Vas} x_{k-1})\Delta t$ and $\sigma_k = \sigma_{Vas}\sqrt{\Delta t}$.

By substituting this transition density to the equation $L(\theta)$,

$$L(\theta) = -\sum_{k=1}^{N} \ln\left(\frac{1}{\sqrt{2\pi\left(\sigma_{Vas}\sqrt{\Delta t}\right)^2}} exp\left(\frac{-(x_k - x_{k-1} - (\alpha_{Vas} + \beta_{Vas} x_{k-1})\Delta t\ )^2}{2\left(\sigma_{Vas}\sqrt{\Delta t}\right)^2}\right)\right)$$

$$= -\sum_{k=1}^{N} \ln\left(\frac{1}{\sqrt{2\pi\left(\sigma_{Vas}\sqrt{\Delta t}\right)^2}}\right) - \sum_{k=1}^{N}\left(\frac{-(x_k - x_{k-1} - (\alpha_{Vas} + \beta_{Vas} x_{k-1})\Delta t\ )^2}{2\left(\sigma_{Vas}\sqrt{\Delta t}\right)^2}\right)$$

$$= -\sum_{k=1}^{N} \ln\left(\frac{1}{\sqrt{2\pi\left(\sigma_{Vas}\sqrt{\Delta t}\right)^2}}\right)$$

$$+ \frac{1}{2\left(\sigma_{Vas}\sqrt{\Delta t}\right)^2}\sum_{k=1}^{N}(x_k - x_{k-1} - (\alpha_{Vas} + \beta_{Vas} x_{k-1})\Delta t)^2$$

To minimize $L(\theta)$, differentiate with respect to $\alpha_{Vas}, \beta_{Vas}$ and σ_{Vas} .

By differentiating with respect to α_{Vas},

$$\frac{\partial L}{\partial \alpha_{Vas}} = \frac{1}{2\left(\sigma_{Vas}\sqrt{\Delta t}\right)^2} \times \sum_{k=1}^{N} 2(x_k - x_{k-1} - (\alpha_{Vas} + \beta_{Vas} x_{k-1})\Delta t) \times -\Delta t$$

$$= \frac{-1}{(\sigma_{Vas})^2}\sum_{k=1}^{N}(x_k - x_{k-1} - (\alpha_{Vas} + \beta_{Vas} x_{k-1})\Delta t).$$

At the optimal value of α_{Vas}, $\frac{\partial L}{\partial \alpha_{Vas}} = 0$.

$$\left.\frac{\partial L}{\partial \alpha_{Vas}}\right|_{\alpha_{Vas}=\hat{\alpha}_{Vas}} = 0$$

$$\frac{-1}{(\sigma_{Vas})^2}\sum_{k=1}^{N}\left(x_k - x_{k-1} - \left(\hat{\alpha}_{Vas} + \hat{\beta}_{Vas}x_{k-1}\right)\Delta t\right) = 0$$

$$\sum_{k=1}^{N}\left(x_k - x_{k-1} - \left(\hat{\alpha}_{Vas} + \hat{\beta}_{Vas}x_{k-1}\right)\Delta t\right) = 0$$

$$\sum_{k=1}^{N}(x_k - x_{k-1}) - \hat{\alpha}_{Vas}N\Delta t - \hat{\beta}_{Vas}\Delta t\sum_{k=1}^{N}x_{k-1} = 0$$

$$\hat{\alpha}_{Vas}N\Delta t + \hat{\beta}_{Vas}\Delta t\sum_{k=1}^{N}x_{k-1} = x_N - x_0 \qquad \textbf{3.4.6}$$

By differentiating with respect to β_{Vas},

$$\frac{\partial L}{\partial \beta_{Vas}} = \frac{1}{2\left(\sigma_{Vas}\sqrt{\Delta t}\right)^2} \times \sum_{k=1}^{N} 2(x_k - x_{k-1} - (\alpha_{Vas} + \beta_{Vas}x_{k-1})\Delta t) \times -x_{k-1}\Delta t$$

$$= \frac{-1}{(\sigma_{Vas})^2}\sum_{k=1}^{N}(x_k - x_{k-1} - (\alpha_{Vas} + \beta_{Vas}x_{k-1})\Delta t)x_{k-1}$$

At the optimal value of β_{Vas}, $\frac{\partial L}{\partial \beta_{Vas}} = 0$.

$$\left.\frac{\partial L}{\partial \beta_{Vas}}\right|_{\beta_{Vas}=\hat{\beta}_{Vas}} = 0$$

$$\frac{-1}{(\sigma_{Vas})^2}\sum_{k=1}^{N}\left(x_k - x_{k-1} - \left(\hat{\alpha}_{Vas} + \hat{\beta}_{Vas}x_{k-1}\right)\Delta t\right)x_{k-1} = 0$$

$$\sum_{k=1}^{N}\left(x_k - x_{k-1} - \left(\hat{\alpha}_{Vas} + \hat{\beta}_{Vas}x_{k-1}\right)\Delta t\right)x_{k-1} = 0$$

$$\sum_{k=1}^{N}(x_k - x_{k-1})x_{k-1} - \hat{\alpha}_{Vas}\Delta t\sum_{k=1}^{N}x_{k-1} - \hat{\beta}_{Vas}\Delta t\sum_{k=1}^{N}{x_{k-1}}^2 = 0$$

$$\hat{\alpha}_{Vas}\Delta t\sum_{k=1}^{N}x_{k-1} + \hat{\beta}_{Vas}\Delta t\sum_{k=1}^{N}{x_{k-1}}^2 = \sum_{k=1}^{N}(x_k - x_{k-1})x_{k-1} \qquad \textbf{3.4.7}$$

Considering equations 3.4.6 and 3.4.7,

$$\frac{x_N - x_0 - \hat{\beta}_{Vas}\Delta t\sum_{k=1}^{N}x_{k-1}}{N\Delta t} = \frac{\sum_{k=1}^{N}(x_k - x_{k-1})x_{k-1} - \hat{\beta}_{Vas}\Delta t\sum_{k=1}^{N}{x_{k-1}}^2}{\Delta t\sum_{k=1}^{N}x_{k-1}}$$

$$(x_N - x_0)\sum_{k=1}^{N}x_{k-1} - \hat{\beta}_{Vas}\Delta t\left(\sum_{k=1}^{N}x_{k-1}\right)^2 = N\sum_{k=1}^{N}(x_k - x_{k-1})x_{k-1} - N\hat{\beta}_{Vas}\Delta t\sum_{k=1}^{N}{x_{k-1}}^2$$

$$\hat{\beta}_{Vas}\Delta t\left(\left(\sum_{k=1}^{N}x_{k-1}\right)^2 - N\sum_{k=1}^{N}{x_{k-1}}^2\right) = (x_N - x_0)\sum_{k=1}^{N}x_{k-1} - N\sum_{k=1}^{N}(x_k - x_{k-1})x_{k-1}$$

$$\hat{\beta}_{Vas} = \frac{(x_N - x_0)\sum_{k=1}^{N}x_{k-1} - N\sum_{k=1}^{N}(x_k - x_{k-1})x_{k-1}}{\Delta t\left(\left(\sum_{k=1}^{N}x_{k-1}\right)^2 - N\sum_{k=1}^{N}{x_{k-1}}^2\right)}$$

By substituting $\hat{\beta}_{Vas}$to the equation 3.4.6,

$$\hat{\alpha}_{Vas} = \frac{\left(\frac{(x_N - x_0)}{\Delta t} - \hat{\beta}_{Vas}\sum_{k=1}^{N}x_{k-1}\right)}{N}.$$

By differentiating $L(\theta)$ with respect to σ_{Vas},

$$-\sum_{k=1}^{N} \ln\left(\frac{1}{\sqrt{2\pi\left(\sigma_{Vas}\sqrt{\Delta t}\right)^2}}\right) + \frac{1}{2\left(\sigma_{Vas}\sqrt{\Delta t}\right)^2}\sum_{k=1}^{N}(x_k - x_{k-1} - (\alpha_{Vas} + \beta_{Vas}x_{k-1})\Delta t)^2$$

$$\frac{\partial L}{\partial \sigma_{Vas}} = -\sum_{k=1}^{N}\sqrt{2\pi\left(\sigma_{Vas}\sqrt{\Delta t}\right)^2}\frac{1}{\sqrt{2\pi\Delta t}} \times -\frac{1}{\sigma_{Vas}{}^2}$$

$$+\frac{-2}{2\Delta t(\sigma_{Vas})^3}\sum_{k=1}^{N}(x_k - x_{k-1} - (\alpha_{Vas} + \beta_{Vas}x_{k-1})\Delta t)^2$$

$$= -\sum_{k=1}^{N}\sigma_{Vas} \times -\frac{1}{\sigma_{Vas}{}^2} - \frac{1}{\Delta t(\sigma_{Vas})^3}\sum_{k=1}^{N}(x_k - x_{k-1} - (\alpha_{Vas} + \beta_{Vas}x_{k-1})\Delta t)^2$$

$$= \frac{N}{\sigma_{Vas}} - \frac{1}{\Delta t(\sigma_{Vas})^3}\sum_{k=1}^{N}(x_k - x_{k-1} - (\alpha_{Vas} + \beta_{Vas}x_{k-1})\Delta t)^2$$

$$= \frac{N\Delta t(\sigma_{Vas})^2 - \sum_{k=1}^{N}(x_k - x_{k-1} - (\alpha_{Vas} + \beta_{Vas}x_{k-1})\Delta t)^2}{\Delta t(\sigma_{Vas})^3}$$

At the optimal value of σ_{Vas}, $\frac{\partial L}{\partial \sigma_{Vas}} = 0$.

$$\left.\frac{\partial L}{\partial \sigma_{Vas}}\right|_{\sigma_{Vas}=\hat{\sigma}_{Vas}} = 0$$

$$\frac{N\Delta t(\sigma_{Vas})^2 - \sum_{k=1}^{N}(x_k - x_{k-1} - (\alpha_{Vas} + \beta_{Vas}x_{k-1})\Delta t)^2}{\Delta t(\sigma_{Vas})^3} = 0$$

$$N\Delta t(\sigma_{Vas})^2 - \sum_{k=1}^{N}(x_k - x_{k-1} - (\alpha_{Vas} + \beta_{Vas}x_{k-1})\Delta t)^2 = 0$$

$$\hat{\sigma}_{Vas}{}^{2} = \frac{\sum_{k=1}^{N}(x_k - x_{k-1} - (\alpha_{Vas} + \beta_{Vas}x_{k-1})\Delta t)^2}{N\Delta t}$$

$$\hat{\sigma}_{Vas}{}^{2} = \frac{1}{N\Delta t}\sum_{k=1}^{N}(x_k - x_{k-1} - (\alpha_{Vas} + \beta_{Vas}x_{k-1})\Delta t)^2$$

Hence, the parameters for the Vasicek model can be obtained from the equations:

$$\hat{\beta}_{Vas} = \frac{(x_N - x_0)\sum_{k=1}^{N} x_{k-1} - N\sum_{k=1}^{N}(x_k - x_{k-1})x_{k-1}}{\Delta t((\sum_{k=1}^{N} x_{k-1})^2 - N\sum_{k=1}^{N} x_{k-1}{}^2)}$$

$$\hat{\alpha}_{Vas} = \frac{\left(\frac{(x_N - x_0)}{\Delta t} - \hat{\beta}_{Vas}\sum_{k=1}^{N} x_{k-1}\right)}{N}$$

and

$$\hat{\sigma}_{Vas}{}^{2} = \frac{1}{N\Delta t}\sum_{k=1}^{N}(x_k - x_{k-1} - (\alpha_{Vas} + \beta_{Vas}x_{k-1})\Delta t)^2.$$

Table 3.1 contains the maximum likelihood estimators of the four SDE models described in this research. Using a sample of past gold prices in Sri Lanka, it can be determined the above parameters. In this study, gold prices per troy ounce from 01st of October, 2015 to 07th of October, 2016 were used to estimate the parameters given in table 3.1.

Table 3.1: Table of maximum likelihood estimators of the four SDE models

Model	Parameter	Estimated value
Brownian Motion	$\hat{\alpha}_{BM}$	$\frac{(x_N - x_0)}{N\Delta t}$
	$\hat{\sigma}_{BM}{}^2$	$\frac{\sum_{k=1}^{N}(x_k - x_{k-1} - \hat{\alpha}_{BM}\Delta t)^2}{N\Delta t}$
Geometric Brownian Motion	$\hat{\beta}_{GBM}$	$\frac{1}{N\Delta t}\sum_{k=1}^{N}\left(\frac{x_k - x_{k-1}}{x_{k-1}}\right)$
	$\hat{\sigma}_{GBM}{}^2$	$\frac{1}{N\Delta t}\sum_{k=1}^{N}\left(\frac{x_k - (1+\hat{\beta}_{GBM}\Delta t)x_{k-1}}{x_{k-1}}\right)^2$
CIR Model	$\hat{\alpha}_{CIR}$	$\frac{\left(\frac{(x_N - x_0)}{\Delta t\sum_{k=1}^{N} x_{k-1}} - \frac{1}{N\Delta t}\sum_{k=1}^{N}\frac{x_k - x_{k-1}}{x_{k-1}}\right)}{\left(\frac{N}{\sum_{k=1}^{N} x_{k-1}} - \frac{\sum_{k=1}^{N} 1/x_{k-1}}{N}\right)}$
	$\hat{\beta}_{CIR}$	$\frac{\frac{(x_N - x_0)}{\Delta t} - N\hat{\alpha}_{CIR}}{\sum_{k=1}^{N} x_{k-1}}$
	$\hat{\sigma}_{CIR}{}^2$	$\frac{1}{N\Delta t}\sum_{k=1}^{N}\frac{\left(x_k - \hat{\alpha}_{CIR}\Delta t - (1+\hat{\beta}_{CIR}\Delta t)x_{k-1}\right)^2}{x_{k-1}}$
Vasicek Model	$\hat{\beta}_{Vas}$	$\frac{(x_N - x_0)\sum_{k=1}^{N} x_{k-1} - N\sum_{k=1}^{N}(x_k - x_{k-1})x}{\Delta t\left((\sum_{k=1}^{N} x_{k-1})^2 - N\sum_{k=1}^{N} x_{k-1}{}^2\right)}$
	$\hat{\alpha}_{Vas}$	$\frac{\left(\frac{(x_N - x_0)}{\Delta t} - \hat{\beta}_{Vas}\sum_{k=1}^{N} x_{k-1}\right)}{N}$
	$\hat{\sigma}_{Vas}{}^2$	$\frac{1}{N\Delta t}\sum_{k=1}^{N}\left(x_k - \hat{\alpha}_{VAs}\Delta t - (1+\hat{\beta}_{Vas}\Delta t)x_{k-1}\right)^2$

3.5 Maximum Error of the Estimate

The maximum error of the estimate is denoted by E and is one-half the width of the confidence interval.

$$E = Z\alpha/_2 \frac{\sigma}{\sqrt{n}}$$

where $Z\alpha/_2$ is the Z– score obtained from the Z- table, σ is the population standard deviation and n is the sample size.

This formula will work for means and proportions because they will use the Z or T distributions which are symmetric.

3.6 Forecasting Accuracy Measures

Needless to say, forecasting is an important task in this research. With many different methods in

forecasting, understanding their relative performance is critical for more accurate prediction of the daily gold price. Various accuracy measures have been used in the literature. In this research two accuracy measures are considered.

3.6.1 Root Mean Square Error (RMSE)

$$\boldsymbol{RMSE} = \sqrt{\sum_{i=1}^{m} \frac{(x_i - \hat{x}_i)^2}{m}}$$

where x_i is the observed value and $\hat{x}_i$ is the forecasted value at time $t = i$ and m is the number of observations. The minimum value of RMSE indicates the best model.

3.6.2 Mean Absolute Percentage Error (MAPE)

$$\boldsymbol{MAPE} = \left(\left(\sum_{i=1}^{m} \left| \frac{x_i - \hat{x}_i}{x_i} \right| \right) \Big/ m \right) \times 100$$

where x_i is the observed value, $\hat{x}_i$ is the forecasted value at time $t = i$ and m is the number of observations.

Minimum value of MAPE indicates the best model.

CHAPTER 04

DATA ANALYSIS

This chapter describes the analysis of data followed by a discussion of the research findings. Daily gold prices per troy ounce from 1st of October, 2015 to 14th of October, 2016 were obtained from http://www.cbsl.gov.lk/htm/english/_cei/er/g_1.asp on 1st of November 2016. Among these data, daily gold prices from 1st of October, 2015 to 07th of October, 2016 are used for modeling and other 05 observations are used for forecasting.

4.1 Data Analysis

Figure 4.1 represents the graph of daily gold prices per troy ounce in Sri Lanka from 1st of October, 2015 to 14th of October, 2016.

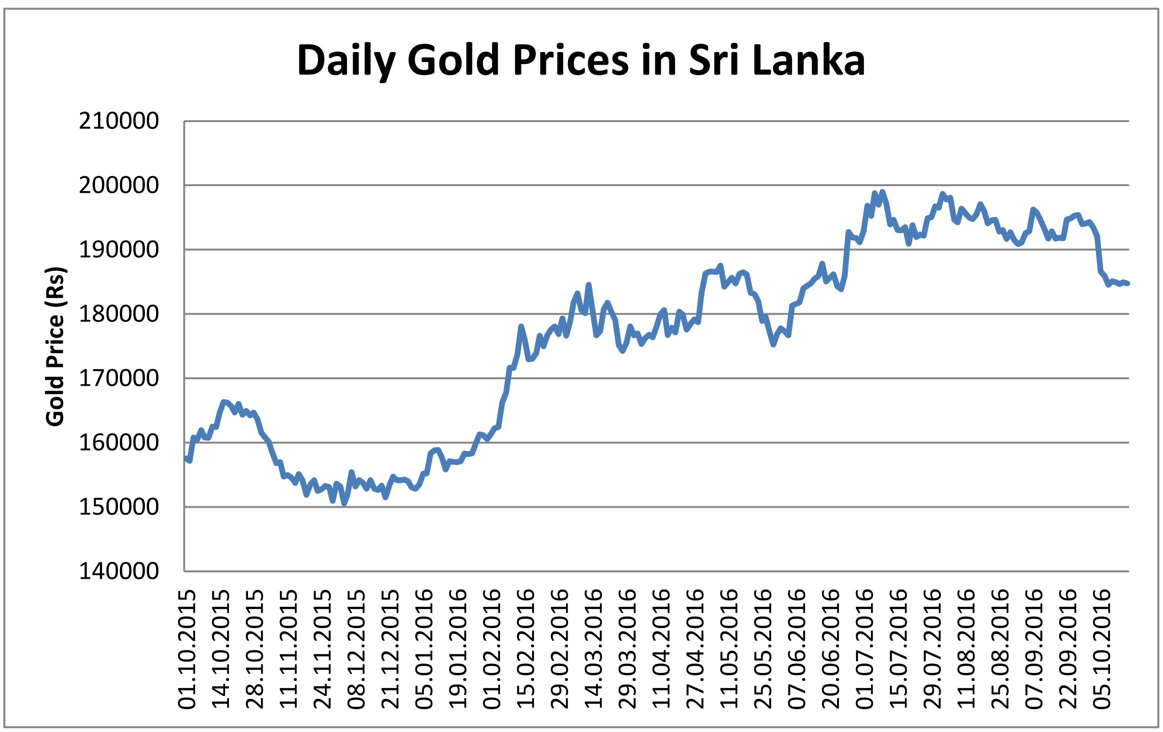


Figure 4.1: The graph of Gold price per troy ounce in Sri Lanka from 01/10/2015 to 14/10/2016

It can be clearly observed that all the gold prices during this period is in between Rs.150000 to Rs.200000. And also there is an upward trend of the daily gold prices of Sri Lanka in this period. Because of several economic factors such as inflation, currency movements, uncertainty etc. the price of gold may be fluctuate rapidly.

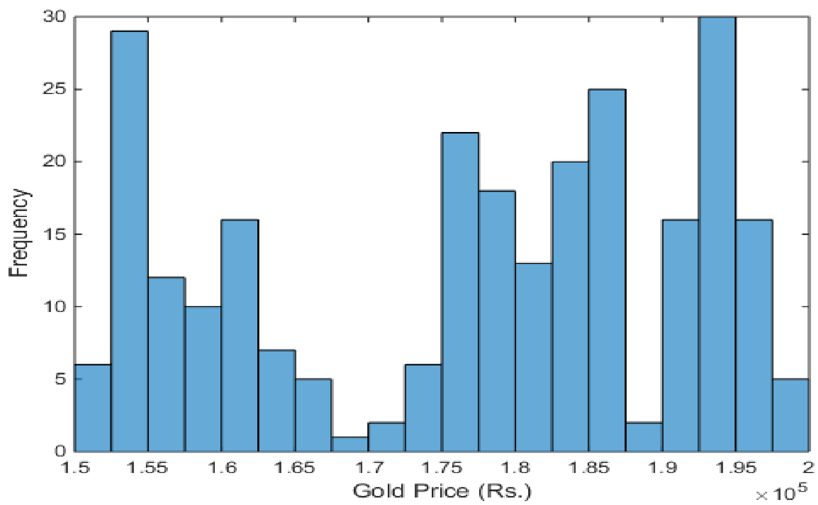


Figure 4.2: The graph of the histogram for the Gold Price in Sri Lanka

To graphically summarize the given data set, the best way is to use a histogram. Figure 4.2 represents the histogram for the daily gold prices in Sri Lanka. It is a non- symmetric graph and it has no apparent pattern. Therefore the price of gold in Sri Lanka is a random distribution. Like the uniform distribution, it may describe a distribution that has several peaks.

Figure 4.3 represents the graph of cumulative distribution function for the daily gold prices in Sri Lanka.

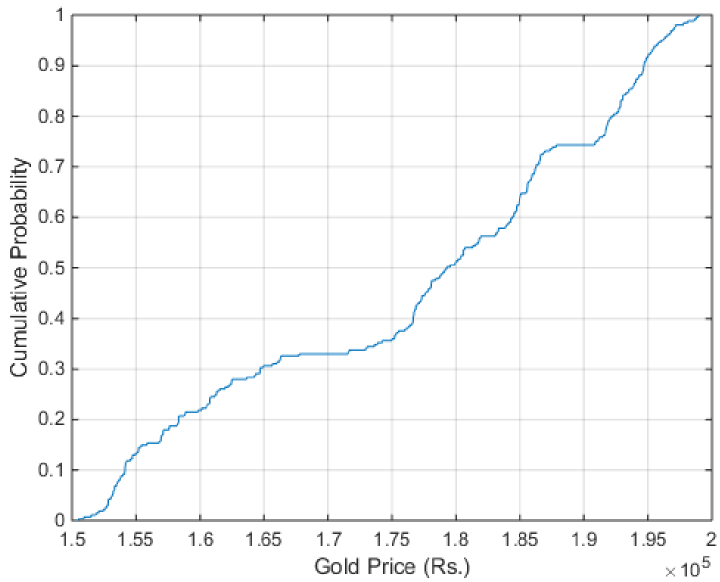


Figure 4.3: The graph of cumulative distribution function for the gold price in Sri Lanka

In this study, four SDE models Brownian motion, Geometric Brownian motion, CIR model and Vasicek model are considered. To forecast the daily gold prices using these SDE models, parameters of the four SDE models should be estimated.

The table 4.1 represents the estimated parameters of the four SDE models. To estimate the parameters, table 3.1 and daily gold prices from 01/10/2015 to 07/10/2016 in Sri Lanka are used.

Using the parameters of the table 4.1, four stochastic differential equations can be written as follows:

- Brownian Motion

$$\boldsymbol{dX(t) = 27722.543dt + 28207.0273dW(t)}$$

- Geometric Brownian Motion

$$\boldsymbol{dX(t) = 0.1751X(t)dt + 0.1593X(t)dW(t)}$$

- CIR Model

$$\boldsymbol{dX(t)} = (\mathbf{406154.478\text{-}2.1487}\boldsymbol{X(t)})\boldsymbol{dt} + \mathbf{66.7345}\sqrt{\boldsymbol{X(t)}}\boldsymbol{dW(t)}$$

- Vasicek Model

$$\boldsymbol{dX(t)} = (\mathbf{439873.2658\text{-}2.3401}\boldsymbol{X(t)})\boldsymbol{dt} + \mathbf{28118.0053}\boldsymbol{dW(t)}$$

where $X(t)$ is the gold price at time t and the $W(t)$ is a linearly independent Wiener process.

Table 4.1: Table of estimated parameters of the four SDE models using maximum likelihood estimation method

Model	Parameter	Estimated value
Brownian Motion	α_{BM}	27722.543
	σ_{BM}	28207.0273
Geometric Brownian Motion	α_{GBM}	0.1751
	σ_{GBM}	0.1593
CIR Model	α_{CIR}	406154.478
	β_{CIR}	-2.1487
	σ_{CIR}	66.7345
Vasicek Model	α_{Vas}	439873.2658
	β_{Vas}	-2.3401
	σ_{Vas}	28118.0053

After estimating parameters, daily gold prices in Sri Lanka can be forecasted using Euler- Maruyama approximations of the above four SDE models. Euler- Maruyama approximations of the four SDE models are given below:

- Brownian Motion

$$\boldsymbol{X(t)} = \boldsymbol{X(t-1)}+\mathbf{27722.543}\Delta\boldsymbol{t} + \mathbf{28207.0273}\sqrt{\Delta\boldsymbol{t}}\boldsymbol{\eta_t}$$

➢ Geometric Brownian Motion

$$\boldsymbol{X(t) = X(t-1) + 0.1751X(t-1)\Delta t + 0.1593X(t-1)\sqrt{\Delta t}\eta_t}$$

➢ CIR Model

$$\boldsymbol{X(t) = X(t-1) + (406154.478\text{-}2.1487X(t-1))\Delta t + 66.7345\sqrt{X(t-1)}\sqrt{\Delta t}\eta_t}$$

➢ Vasicek Model

$$\boldsymbol{X(t) = X(t-1) + (439873.2658\text{-}2.3401X(t-1))\Delta t + 28118.0053\sqrt{\Delta t}\eta_t}$$

where $X(t)$ is the gold price at time t with $X(0) = x_0$, $\Delta t = {}^{1}/_{252}$ and $\eta_t \sim N(0, \Delta t)$.

Using the above equations, sample paths can be obtained for the four SDE models. By taking Rs.185099.78 which is the gold price at 10th of October 2016 as the initial gold price, 05 sample paths were obtained for each of the SDE models. Figure 4.4, 4.5, 4.6 and 4.7 represents the sample paths for the Brownian motion, Geometric Brownian motion, CIR and Vasicek model respectively.

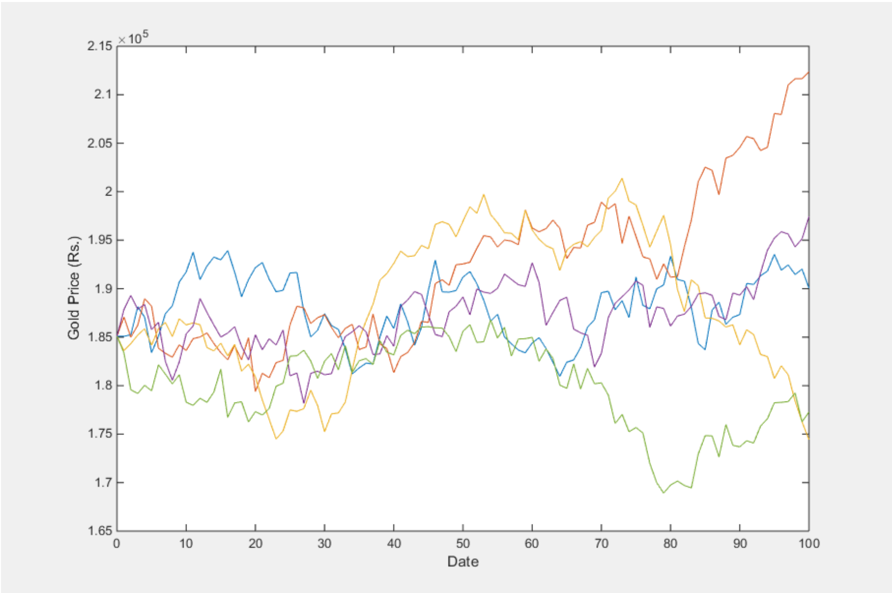


Figure 4.4: The Graph of Five Sample Paths for the Brownian motion Model

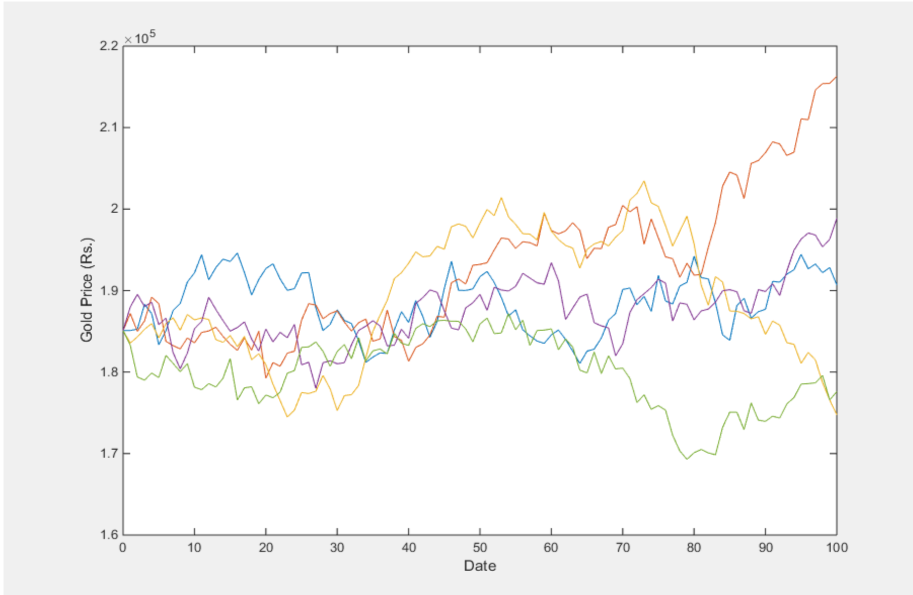


Figure 4.5: The Graph of Five Sample Paths for the Geometric Brownian motion Model

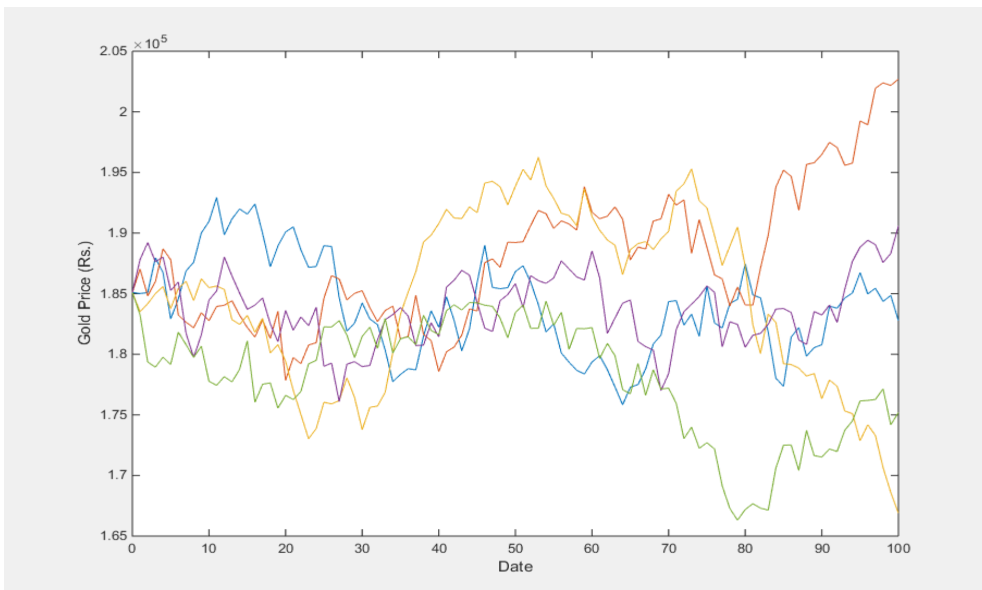


Figure 4.6: The Graph of Five Sample Paths for the CIR Model

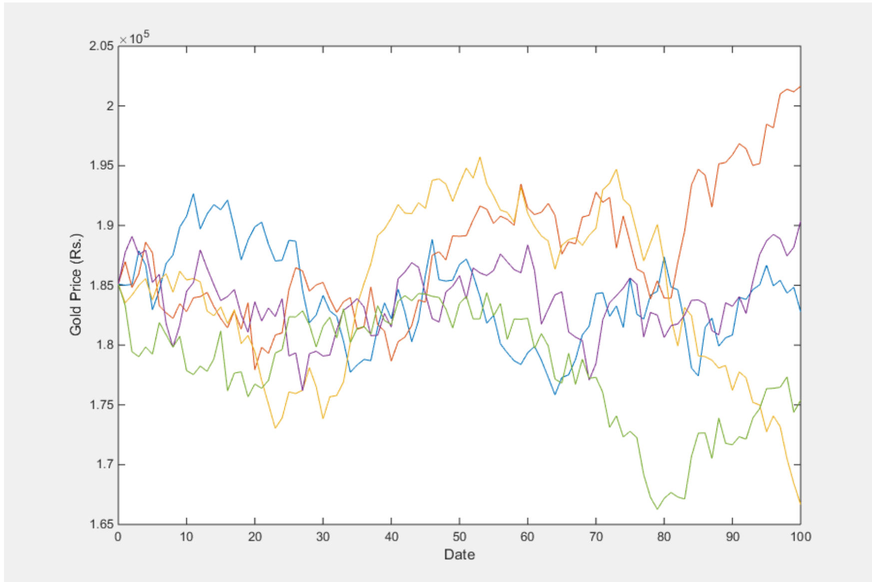


Figure 4.7: The Graph of Five Sample Paths for the Vasicek Model

The Monte Carlo technique is used to simulate the daily gold prices in this study. Considering the law of large numbers, by generating large number of sample paths and taking the average of them, a unique value can be obtained to the gold price. Figure 4.8 represents the convergence of the average gold price on 11th of October 2016.

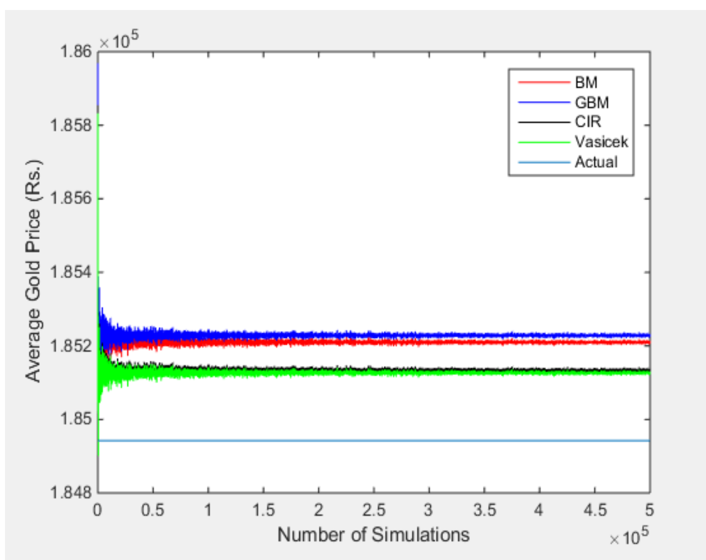


Figure 4.8: The graph of convergence of the forecasted gold prices on 11/10/2016

If we generate few number of sample paths, the average value may vary. But if the number of sample paths is large, the average value of the gold price converges to some fixed value. Figure 4.8 shows that when we increase the number of sample paths from 0 to 500000, average value is converged to some fixed value. Forecasted value for the date 11/10/2016 using Brownian motion is converged to Rs.185212.5. The converged gold prices for Geometric Brownian motion, CIR model and Vasicek model are Rs.185228.35, Rs.185133.7 and Rs.185122.06 respectively. The actual gold price on 11^{th} of October, 2016 is, Rs.184942.16. Similarly, we can obtain the graphs for other 03 days. That graphs are included in appendix 02.

Actual gold prices and the convergent gold prices from 11^{th} of October 2016 to 14^{th} of October 2016 are given in the table 4.2. First row of the table represents the gold price of 10^{th} of October, 2016 which is used as the initial value. Second column represents the actual gold price and third to sixth columns represents the forecasted values for the Brownian motion, Geometric Brownian motion, CIR model and Vasicek model respectively. According to the table, forecasted daily gold price of Sri Lanka is increasing and the values are much closed to the actual ones.

Table 4.2: Table of Actual and Forecasted Gold Prices from 11/10/2016 to 28/11/2016

Date	Actual data (Rs.)	Brownian Motion (Rs.)	Geometric Brownian Motion(Rs.)	CIR model (Rs.)	Vasicek Model (Rs.)
10/10/2016	185099.78	185099.78	185099.78	185099.78	185099.78
11/10/2016	184942.16	185212.5	185228.35	185133.7	185122.06
12/10/2016	184661.38	185322.9	185358.55	185166.08	185148.44
13/10/2016	184916.99	185430.25	185490.13	185199.81	185174.25
14/10/2016	184741.44	185536.68	185619.04	185231.36	185198.9

Table 4.3 represents the maximum errors of each predicted value using four SDE models under 95% confidence level. From that table, it can be observed that the maximum error is getting large when the date is far from the initial date.

Table 4.3: Table of Maximum Errors of Estimates

Date	Brownian Motion (Rs.)	Geometric Brownian Motion(Rs.)	CIR model (Rs.)	Vasicek Model (Rs.)
11/10/2016	4.9396	5.1545	5.018	4.91
12/10/2016	6.9729	7.2876	7.0597	6.9092
13/10/2016	8.5297	8.9298	8.6335	8.4186
14/10/2016	9.8468	10.3199	9.9214	9.6861

Finding the suitable model among four SDEs to forecast the daily gold prices is a main objective of this study. Forecasting accuracy measures can be used to check the best fitted model among the four SDEs. In this research, root mean square error (RMSE) and mean absolute percentage error (MAPE) are used to check the accuracy. Table 4.4 represents RMSE and MAPE values for the four SDE models.

Table 4.4: Table of forecasting accuracy measures for four SDE models

Model	Root Mean Square Error (RMSE)	Mean Absolute Percentage Error (MAPE)
Brownian Motion	592.987	0.3031%
Geometric Brownian Motion	645.489	0.3293%
CIR Model	390.966	0.1988%
Vasicek Model	369.135	0.187%

According to the table 4.4, Vasicek model has the minimum RMSE value 369.135 and the minimum MAPE value 0.187 %.

As the final step of this research, the MAPEs and RMSEs obtained for four SDE models are compared with the ARIMA (2, 1, 2) model in [14]. ARIMA (2, 1, 2) model can be written as,

$$(1 - \phi_1 B - \phi_2 B^2)(1 - B)X_t = (1 + \theta_1 B + \theta_2 B^2)Z_t$$

By substituting the estimated coefficients to the equation,

$X_t = -0.2099X_{t-1} + 0.4068X_{t-2} + 0.8031X_{t-3} + Z_t - 1.306Z_{t-1} - 0.8385Z_{t-2} + 0.04204.$

Table 4.5 represents the forecasted values for the ARIMA (2, 1, 2) model.

Table 4.5: Forecasted Values for the ARIMA (2, 1, 2) Model

Actual Value (Rs.)	Forecasted Value (Rs.)
184942.16	185540.57
184661.38	184546.92
184916.99	185395.14
184741.44	185166.88

RMSE value and the MAPE value for the ARIMA (2,1,2) model are 441.77 and 0.2187% respectively. Hence the Vasicek model has the minimum RMSE and MAPE value among four SDEs and ARIMA(2,1,2) model.

CHAPTER 05

CONCLUSION AND FURTHER RESEARCH

In this chapter, the main findings are summarized in section 5.1 and general conclusions based on the findings of the studies are described in section 5.2. Furthermore, the strengths and limitations of this study are considered and suggestions for further research are presented in section 5.3 and 5.4.

5.1 Summary

In this research study, the Sri Lankan gold price is forecasted choosing the suitable model among four SDE models, Brownian motion, Geometric Brownian motion, CIR model and Vasicek model. The gold price of Sri Lanka per troy ounce from 01st of October 2015 to 14th of October were used to analyze and predict the daily gold price in Sri Lanka. Parameter estimation of four SDE models were done using maximum likelihood estimation method. Numerical simulations of SDEs are carried out using Euler – Maruyama approximation method. By applying estimated parameters to the four SDE models, gold prices of Sri Lanka is predicted from 11th of October 2016 to 14th of October 2016. To simulate the gold prices, Monte Carlo technique is used. Using two forecasting accuracy measures RMSE and MAPE, the suitable model among four SDE models to forecast the daily gold prices in Sri Lanka is selected. Then, these measures are compared with the forecasting accuracy measures of the ARIMA (2, 1, 2) model in [14]. To analyze the data, MATLAB software is used.

5.2 Conclusion

According to the results of this research study, simulated gold prices of four SDE models are much closer to actual ones. Therefore we can conclude that SDE models can be used to forecast gold prices in Sri Lanka and it will be helpful for investors who are interested in invest their money in gold market. Because they can invest their money in gold market with a law risk.

Among the four SDE models which we considered, Vasicek model has the minimum RMSE and MAPE values. Hence we can conclude that the Vasicek model is the suitable SDE model to forecast the daily gold prices in Sri Lanka from 11th of October 2016 to 14th of October 2016.

In this study, the MAPEs and RMSEs of four SDEs are compared with an ARIMA model which is used to forecast monthly gold prices in Sri Lanka. In literature a model to predict daily gold prices in Sri Lanka could not be found. Therefore that model is used to compare the forecasting accuracy. Comparatively, Vasicek model is better than the ARIMA (2,1,2) model to predict gold prices for a short period according to the values of forecasting accuracy measures. But it can be observed that simulated forecasted values using four SDEs are linearly increasing. The reason for that is using the Monte Carlo simulation to simulate the results. In Monte Carlo simulation, mean of the large number of sample paths is used and the mean of random process is converged to zero when we used a large sample [2]. Hence the predicted daily gold prices have a linear pattern. Because of that, this method is suitable for short runs only.

In case of sudden change in the data, the best model among four SDE models may be changed to another model. Because of that, predicting large number of data points using a one SDE model is not suitable. Hence this method is suitable only to forecast smaller number of daily gold prices.

In this research study, several mathematical programs are developed to estimate the parameters of four SDE models, to forecast the daily gold price and to check the best model among four SDEs using the mathematical software MATLAB. These programs can be used for any data set not only for the daily gold prices. If we can update the data set, we can find the best model and forecast the gold price at any time using those mathematical programs.

According to the results obtained in this study, daily gold prices from 11/10/2016 to 14/10/2016 can be predicted using the Vasicek model,

$$X(t) = X(t-1) + \left(439873.2658\text{-}2.3401X(t-1)\right)\Delta t + 28118.0053\sqrt{\Delta t}\eta_t.$$

5.3 Limitations of the Study

In this research study, we only considered four SDE models, Brownian motion, Geometric Brownian motion, CIR model and Vasicek model. Parameters of these four models are not depend on the time. There are many SDEs which has time dependent parameters. If we can consider such SDEs too, we can predict the gold price more accurately.

5.4 Further Research

As a future work of this research, one can compare some statistical models and SDE models to find the most suitable model to forecast the daily gold prices in Sri Lanka.

In this research, four SDE models were considered and the parameters of the models are not depend on time. One can extend this research using some SDE models which have time dependent parameters.

REFERENCES

[1] Ali, A., Muhammad, I.C., Qamar, S., Akhtar, N., Mahmood, T., Hyder, M. & Jamshed, M.T. (2016), International Journal of Asian Social Science, 6(11): 614-624.

[2] Allen, E. (2007), Modelling with Ito Stochastic Differential Equations, ISBN-13 978-1-4020-5952-0 (HB), ISBN-13 978-1-4020-5953-7 (e-book).

[3] Aye, G., Gupta, R., Hammoudeh, S. & Kim, W.J. (2015), Forecasting the Price of Gold Using Dynamic Model Averaging.

[4] Azzutti, A., (2016), Forecasting Gold Price: A Comparative Study.

[5] Basaru, C., Toraman, C. & Bayramoglu, M.F. (2011), Determination of Factors Affecting the Price of Gold: A Study of MGARCH Model, Business and Economics Research Journal Volume 2 Number 4 pp. 37-50.

[6] Dirk, G. B., Joscha B., & Robert C. (2014), Gold Price Forecasts in a Dynamic Model Averaging Framework.

[7] Davis, R., Dedu, V.K. & Bonye, F. (2014) ,Modeling and Forecasting of Gold Prices on Financial Markets, American International Journal of Contemporary Research Vol. 4 No. 3.

[8] Guha, B, & Bandyopadhyay, G. (2016), Gold Price Forecasting Using ARIMA Model, Journal of Advanced Management Science Vol. 4, No. 2.

[9] Guidoum, A.C. & Boukhetala, K. (2016), Estimation of Stochastic Differential Equations with Sim.DiffProc Package Version 3.2.

[10] Ismail, Z., Yahya, A. & Shabri, A. (2009), Forecasting Gold Prices Using Multiple Linear Regression Method, American Journal of Applied Sciences 6 (8): 1509-1514.

[11] Khalid, M., Sultana, M, & Zaidi, F. (2014),Forecasting Gold Price: Evidence from Pakistan Market, Research Journal of Finance and Accounting ,Vol.5, No.3.

[12] Khan, M.M.A. (2013), Forecasting of Gold Prices (Box Jenkins Approach), International Journal of Emerging Technology and Advanced Engineering ,Volume 3, Issue 3.

[13] Phillips, P.C.B. & Yu, J. (2006), Maximum Likelihood and Gaussian Estimation of Continuous Time Models in Finance, Research Collection School of Economics.

[14] Pitigalaarachchi, P. A. A. C., Jayasundara D. D. M. & Chandrasekara N.V. (2016), Modeling and forecasting Sri Lankan Gold Prices. International Journal of Sciences: Basic and Applied Research (IJSBAR), Volume 27, No 3, pp. 247-260.

[15] Sindhu (2013), A study on impact of select factors on the price of Gold, IOSR Journal of Business and Management (IOSR-JBM, Volume 8, Issue 4, pp. 84-93.

[16] Prices: Evidence from Pakistan, The Lahore Journal of Economics 18: 2 pp.1–35.

[17] http://www.investopedia.com/terms/t/troyounce.asp

[18] https://bebusinessed.com/history/the-history-of-gold/

[19] https://wikipedia.org

APPENDIX 01

Parameter Estimation of Four SDE Models

```
%This program compute the parameters of four SDE models Brownian motion, Geometric Brownian motion, CIR model and Vasicek model according to the Sri Lankan gold price using maximum likelihood method
clc;
format long;
gold=xlsread('gold rate.xlsx'); %Read the excel file which contains gold price data
N=length(gold)-1;
dt=1/252;                %The length of the time step

%Parameters of Brownian motion using maximum likelihood method
alpha_bm=(gold(end)-gold(1))/(dt*N);

SS1=0;
for i=1:N
    SS1=SS1+(gold(i+1)-gold(i)-alpha_bm*dt)^2;
End

sigma_bm=sqrt(SS1/(N*dt));

%Parameters of Geometric Brownian motion using maximum likelihood method
SS2=0;
```

```
for j=1:N
    SS2=SS2+(gold(j+1)-gold(j))/gold(j);
End

beta_gbm=SS2/(dt*N);

SS22=0;
for k=1:N
    SS22=SS22+((gold(k+1)-
(1+beta_gbm*dt)*gold(k))/gold(k))^2;
end

sigma_gbm=sqrt(SS22/(N*dt));


%Parameters of CIR model using maximum likelihood method
alpha_cir=(N*(gold(end)-gold(1))-(sum(gold)-
gold(end))*SS2)/(dt*(N^2-(sum(gold)-
gold(end))*(sum(1./gold)-1/gold(end))));

beta_cir=(gold(end)-gold(1)-alpha_cir*N*dt)/(dt*(sum(gold)-
gold(end)));

SS3=0;
for m=1:N
    SS3=SS3+(gold(m+1)-alpha_cir*dt-
(1+beta_cir*dt)*gold(m))^2/gold(m);
End

sigma_cir=sqrt(SS3/(N*dt));
```

```
%Parameters  of  Vasicek  model  using  maximum  likelihood
method
SS41=0;
for b=1:N
    SS41=SS41+(gold(b+1)-gold(b))*gold(b);
end
sum_1=(((sum(gold)-gold(end))*SS41)-((gold(end)-
gold(1))*(sum(gold.^2)-gold(end)^2)))/(dt*((sum(gold)-
gold(end))^2-N*(sum(gold.^2)-gold(end)^2)));

alpha_vas=(((sum(gold)-gold(end))*SS41)-((gold(end)-
gold(1))*(sum(gold.^2)-gold(end)^2)))/(dt*((sum(gold)-
gold(end))^2-N*(sum(gold.^2)-gold(end)^2)));

beta_vas=(gold(end)-gold(1)-sum_1*N*dt)/(dt*(sum(gold)-
gold(end)));

SS4=0;
for p=1:N
    SS4=SS4+(gold(p+1)-alpha_vas*dt-
(1+beta_vas*dt)*gold(p))^2;
End

sigma_vas=sqrt(SS4/(N*dt));

%Create the table of parameters.
```

```
parameter={'alpha(BM)';'sigma(BM)';'beta(GBM)';'sigma(GBM)'
;'alpha(CIR)';'beta(CIR)';'sigma(CIR)';'alpha(Vasicek)';'be
ta(Vasicek)';'sigma(Vasicek)'};

MLE=[alpha_bm;sigma_bm;beta_gbm;sigma_gbm;alpha_cir;beta_ci
r;sigma_cir;alpha_vas;beta_vas;sigma_vas];

T1=table(parameter,MLE);

fprintf('Parameters of four SDE models using maximum
likelihood estimation method \n');

disp(T1); %Display the table of parameters
```

Simulation and Forecasting Accuracy Measures

```
%Written by WMHN Weerasinghe
%This program simulate the SDEs using Euler  Maruyama
method
%and test the forecasting accuracy using two tests RMSE and
MAPE

%Simulation
parameters;    %Run the program parameters

y=xlsread('gold_forecast.xlsx'); %Read the excel file which
contains the data use to forecast
m=length(y);
n=1:10:500000;
```

```
BM=zeros(length(n),m);    %Vector of generating values of Brownian motion

GBM=zeros(length(n),m);  %Vector of generating values of Geometric Brownian motion

CIR=zeros(length(n),m); %Vector of generating values of CIR model

Vasicek=zeros(length(n),m); %Vector of generating values of Vasicek model
x0=y(1);    %Initial value

BM(:,1)=x0;
GBM(:,1)=x0;
CIR(:,1)=x0;
Vasicek(:,1)=x0;

for p=1:length(n)

A=randn([n(p),m-1]); %A matrix of random numbers

B=zeros(1,m-1);

     for b=1:m-1

     B(b)=mean(A(:,b));      %Calculate the mean of each column of A

     end
```

```
    for c=2:m

  BM(p,c)=BM(p,c-1)+alpha_bm*dt+sigma_bm*sqrt(dt)*B(c-1);

  GBM(p,c)=GBM(p,c-1)+beta_gbm*GBM(p,c-1)*dt+
               sigma_gbm*GBM(p,c-1)*sqrt(dt)*B(c-1);

  CIR(p,c)=CIR(p,c-1)+((alpha_cir)+beta_cir*CIR(p,c-1))*dt+
               sigma_cir*sqrt(CIR(p,c-1))*sqrt(dt)*B(c-1);
   Vasicek(p,c)=Vasicek(p,c-1)+((alpha_vas)+
          beta_vas*Vasicek(p,c-
          1))*dt+sigma_vas*sqrt(dt)*B(c-1);

    end
end

for z=2:m

    figure;

    plot(n,BM(:,z)','r-',n,GBM(:,z)','b-',n,CIR(:,z)',
     'k-',n,Vasicek(:,z)','g-',n,y(z)*ones(1,length(n)));

    legend('BM','GBM','CIR','Vasicek','Actual');

end
x=1:length(gold)+length(y);
figure;
```

```
plot(x,[gold' y']);

title('Gold price of Sri Lanka from 01st of October 2015 to
28th of October 2016');
xlabel('Date');
ylabel('Gold Price');
figure;

subplot(2,2,1);
plot(1:length(y),BM(length(n),:),1:length(y),y,'*');
title('Plot of past gold prices and forecasted gold prices
using BM model');

subplot(2,2,2);
plot(1:length(y),GBM(length(n),:),1:length(y),y,'*');
title('Plot of past gold prices and forecasted gold prices
using GBM model');

subplot(2,2,3);
plot(1:length(y),CIR(length(n),:),1:length(y),y,'*');
title('Plot of past gold prices and forecasted gold prices
using CIR model');

subplot(2,2,4);
plot(1:length(y),Vasicek(length(n),:),1:length(y),y,'*');
title('Plot of past gold prices and forecasted gold prices
using Vasicek model');
actual_data=y;
model1=BM(length(n),:)';
model2=GBM(length(n),:)';
```

```
model3=CIR(length(n),:)';
model4=Vasicek(length(n),:)';
figure;
t=0:length(y)-1;
plot(t,model1','k-',t,model2','g-',t,model3','b-
',t,model4',
'c-',t,y,'r-');
legend('BM','GBM','CIR','Vasicek','actual');

T1=table(actual_data,model1,model2,model3,model4);
disp(T1);




A11=(y'-BM(length(n),:));
B11=(y'-GBM(length(n),:));
C11=(y'-CIR(length(n),:));
D11=(y'-Vasicek(length(n),:));

figure;
subplot(2,2,1);
plot(1:length(y),A11,'*');
subplot(2,2,2);
plot(1:length(y),B11,'*');
subplot(2,2,3);
plot(1:length(y),C11,'*');
subplot(2,2,4);
plot(1:length(y),D11,'*');
```

```
%Forecasting accuracy

RMSE1=sqrt(sum(abs(A11).^2)/length(y));
RMSE2=sqrt(sum(abs(B11).^2)/length(y));
RMSE3=sqrt(sum(abs(C11).^2)/length(y));
RMSE4=sqrt(sum(abs(D11).^2)/length(y));

MAPE1=(sum(abs(A11)./BM(length(n),:))/length(y))*100;
MAPE2=(sum(abs(B11)./GBM(length(n),:))/length(y))*100;
MAPE3=(sum(abs(C11)./CIR(length(n),:))/length(y))*100;
MAPE4=(sum(abs(D11)./Vasicek(length(n),:))/length(y))*100;

model={'Brownian Motion'; 'Geometric Brownian Motion';
'CIR';'Vasicek'};

RMSE=[RMSE1;RMSE2;RMSE3;RMSE4];

MAPE=[MAPE1;MAPE2;MAPE3;MAPE4];

T2=table(model,RMSE,MAPE);

disp(T2);
```

<u>Sample Paths Generation</u>

```
%This program generates sample paths for the SDE models,
Brownian motion, geometric Brownian motion, CIR model and
Vasicek model.

parameters;
```

```
y=xlsread('gold_forecast.xlsx'); %Read the excel file which contains the data use to forecast

m=101;
n=5;
BM=zeros(n,m);   %Vector of generating values of Brownian motion
GBM=zeros(n,m); %Vector of generating values of Geometric Brownian motion
CIR=zeros(n,m); %Vector of generating values of CIR model
Vasicek=zeros(n,m); %Vector of generating values of Vasicek model


x0=y(1);     %Initial value

BM(:,1)=x0;
GBM(:,1)=x0;
CIR(:,1)=x0;
Vasicek(:,1)=x0;

for p=1:n
     A=randn([n,m-1]); %A matrix of random numbers

     for c=2:m

  BM(p,c)=BM(p,c-1)+alpha_bm*dt+sigma_bm*sqrt(dt)*A(p,c-1);

  GBM(p,c)=GBM(p,c-1)+beta_gbm*GBM(p,c-1)*dt+
                sigma_gbm*GBM(p,c-1)*sqrt(dt)*A(p,c-1);
```

```
  CIR(p,c)=CIR(p,c-1)+((alpha_cir)+beta_cir*CIR(p,c-1))*dt+
          sigma_cir*sqrt(CIR(p,c-1))*sqrt(dt)*A(p,c-1);

  Vasicek(p,c)=Vasicek(p,c-1)+
          ((alpha_vas)+beta_vas*Vasicek(p,c-1))*dt+
          sigma_vas*sqrt(dt)*A(p,c-1);

     end

end
figure

for i_1=1:n
    hold on;
    plot(0:m-1,BM(i_1,:));
end

figure;

for i_2=1:n
   hold on;
    plot(0:m-1,GBM(i_2,:));
end

figure;

for i_3=1:n
   hold on;
    plot(0:m-1,CIR(i_3,:));
```

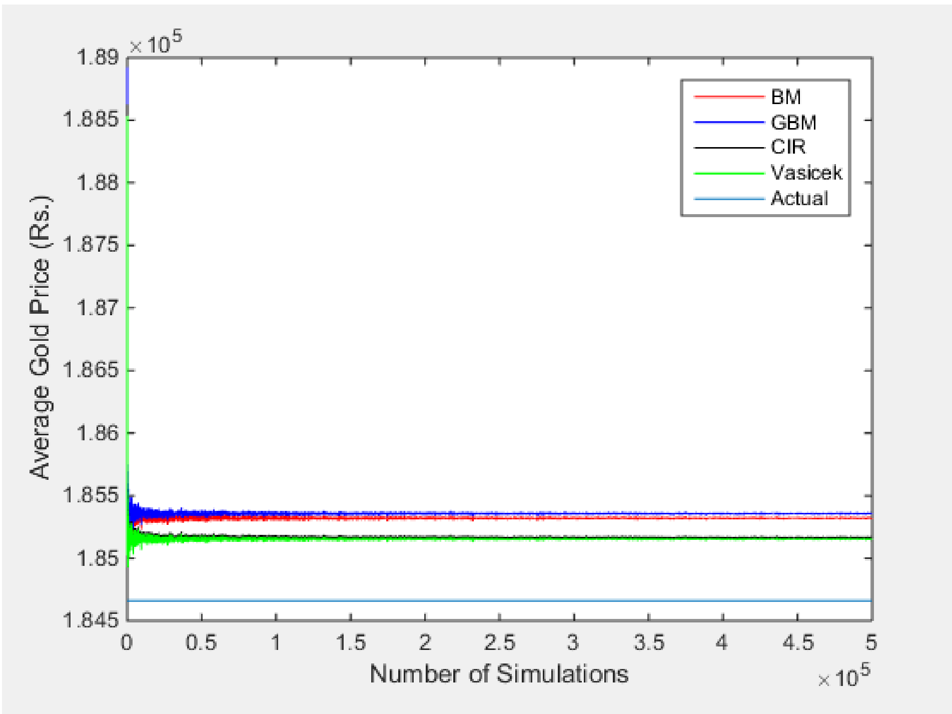
```
end
figure;

for i_4=1:n
    hold on;
    plot(0:m-1,Vasicek(i_4,:));
end
```

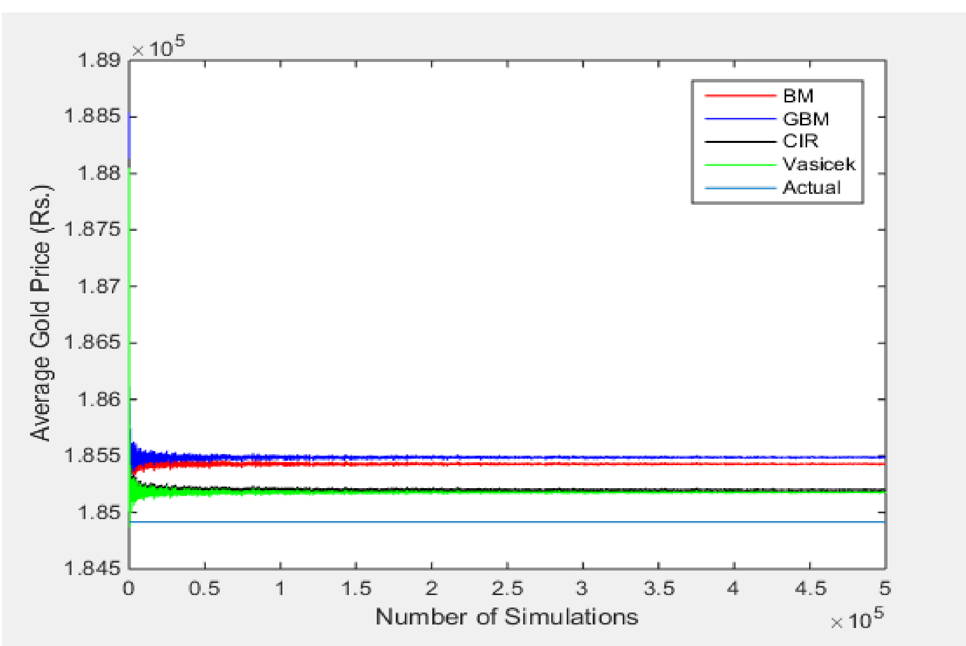
APPENDIX 02

Monte Carlo simulations of the forecasted gold prices from 12th of October 2016 to 14th of October 2016.

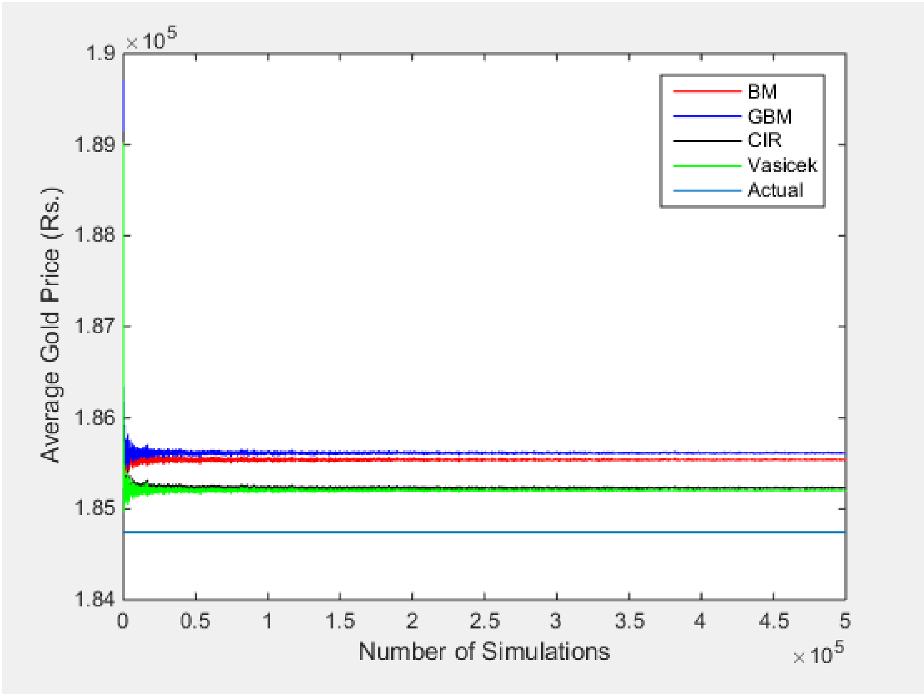
12/10/2016



13/10/2016



14/10/2016



$\times 10^5$
1.9
1.89
1.88
1.87
1.86
1.85
1.84
Average Gold Price (Rs.)
BM
GBM
CIR
Vasicek
Actual
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
Number of Simulations
$\times 10^5$

APPENDIX 03

Gold Price of Sri Lanka in Rupees from 01/10/2015 to 14/10/2016

Date	Gold Price (Rs.)
01.10.2015	157574.104
02.10.2015	157134.843
05.10.2015	160768.1198
06.10.2015	160487.2945
07.10.2015	161959.6559
08.10.2015	160774.8253
09.10.2015	160728.0625
12.10.2015	162490.3533
13.10.2015	162442.4899
14.10.2015	164685.7591
15.10.2015	166316.0396
16.10.2015	166228.3504
19.10.2015	165650.7861
20.10.2015	164663.7863
21.10.2015	166008.0883
22.10.2015	164343.1347
23.10.2015	164931.795
26.10.2015	164243.6864
28.10.2015	164680.2028
29.10.2015	163635.9117
30.10.2015	161538.0497
02.11.2015	160814.9254
03.11.2015	160128.1341
04.11.2015	158333.5129
05.11.2015	156819.7196
06.11.2015	156934.9254
09.11.2015	154711.1945
11.11.2015	154970.0821
12.11.2015	154546.4881
13.11.2015	153732.0576
16.11.2015	155136.5138
17.11.2015	154112.8495
18.11.2015	151885.0426
19.11.2015	153455.7704
20.11.2015	154174.6454
23.11.2015	152459.4916
24.11.2015	152738.933
26.11.2015	153252.5517
27.11.2015	153094.9603
30.11.2015	150926.2789
01.12.2015	153616.6656
02.12.2015	153187.3727
03.12.2015	150522.2986
04.12.2015	152087.5964
07.12.2015	155407.1923
08.12.2015	153197.9607
09.12.2015	154163.515

10.12.2015	153780.3309
11.12.2015	152829.6352
14.12.2015	154150.6484
15.12.2015	152803.6777
16.12.2015	152638.9139
17.12.2015	153312.7431
18.12.2015	151484.4273
21.12.2015	153315.7149
22.12.2015	154716.2058
23.12.2015	154189.7187
28.12.2015	154134.3139
29.12.2015	154247.0284
30.12.2015	153982.5198
31.12.2015	153015.7719
01.01.2016	152822.509
04.01.2016	153507.5012
05.01.2016	155184.4858
06.01.2016	155207.84
07.01.2016	158331.7341
08.01.2016	158796.2628
11.01.2016	158874.4732
12.01.2016	157582.4629
13.01.2016	155826.0485
14.01.2016	157105.9157
18.01.2016	156989.6548
19.01.2016	156956.5374
20.01.2016	157114.0797
21.01.2016	158324.0206
22.01.2016	158188.5505
25.01.2016	158333.17
26.01.2016	159846.0911

27.01.2016	161285.5816
28.01.2016	161164.864
29.01.2016	160529.2603
01.02.2016	161311.9
02.02.2016	162226.4738
03.02.2016	162419.2105
05.02.2016	166259.5578
08.02.2016	167758.2265
09.02.2016	171661.2504
10.02.2016	171579.165
11.02.2016	173622.545
12.02.2016	178058.1478
15.02.2016	176002.7446
16.02.2016	172935.1131
17.02.2016	173040.231
18.02.2016	173903.15
19.02.2016	176602.11
23.02.2016	174999.9904
24.02.2016	176730.9793
25.02.2016	177616.3693
26.02.2016	178082.7255
29.02.2016	176876.9562
01.03.2016	179291.175
02.03.2016	176636.1159
03.03.2016	178804.8162
04.03.2016	181787.1656
08.03.2016	183199.09
09.03.2016	180597.7545
10.03.2016	180130.7807
11.03.2016	184523.7413
14.03.2016	180600.9755

15.03.2016	176673.225
16.03.2016	177317.591
17.03.2016	180702.425
18.03.2016	181760.09
21.03.2016	180321.09
23.03.2016	179084.1829
24.03.2016	175202.5113
28.03.2016	174234.5413
29.03.2016	175485.3305
30.03.2016	178063.875
31.03.2016	176679.7005
01.04.2016	176953.1105
04.04.2016	175334.955
05.04.2016	176270.305
06.04.2016	176787.6255
07.04.2016	176392.62
08.04.2016	178018.69
11.04.2016	179975.73
12.04.2016	180600.9755
15.04.2016	176709.2
18.04.2016	177846.01
19.04.2016	177140.9
20.04.2016	180378.65
22.04.2016	179910.2555
25.04.2016	177546.7811
26.04.2016	178406.5005
27.04.2016	179169.89
28.04.2016	178709.41
29.04.2016	183342.99
03.05.2016	186285.745
04.05.2016	186575.2044

05.05.2016	186595.6066
06.05.2016	186511.8118
09.05.2016	187537.0032
10.05.2016	184255.8889
11.05.2016	184993.0894
12.05.2016	185651.9775
13.05.2016	184737.3963
16.05.2016	186248.433
17.05.2016	186488.904
18.05.2016	186112.6838
19.05.2016	183309.0463
20.05.2016	183083.8625
24.05.2016	181939.725
25.05.2016	178937.275
26.05.2016	179549.425
27.05.2016	177261.15
30.05.2016	175220.65
31.05.2016	176874.9125
01.06.2016	177780.4877
02.06.2016	177328.5583
03.06.2016	176673.9345
06.06.2016	181326.405
07.06.2016	181560.878
08.06.2016	181823.125
09.06.2016	183973.4659
10.06.2016	184399.1276
13.06.2016	184745.9643
14.06.2016	185493.7931
15.06.2016	185884.79
16.06.2016	187842.075
17.06.2016	185016.4683

20.06.2016	185573.4025
21.06.2016	186180.0418
22.06.2016	184273.7825
23.06.2016	183850.1875
24.06.2016	185933.7988
27.06.2016	192731.4988
28.06.2016	191889.775
29.06.2016	191809.8875
30.06.2016	191170.7875
01.07.2016	192876.7488
04.07.2016	196820.2863
05.07.2016	195225.1486
07.07.2016	198798.2002
08.07.2016	196973.525
11.07.2016	198984.5113
12.07.2016	197169.6125
13.07.2016	193876.63
14.07.2016	194670.5863
15.07.2016	193065.5738
18.07.2016	193022.725
20.07.2016	193507.86
21.07.2016	190912.2425
22.07.2016	193789.2
25.07.2016	191945.4625
26.07.2016	192345.5463
27.07.2016	192163.3588
28.07.2016	194877.7789
29.07.2016	195048.3024
01.08.2016	196728.9775
02.08.2016	196529.3
03.08.2016	198660.6372

04.08.2016	197765.6975
05.08.2016	198085.3093
08.08.2016	194674.48
09.08.2016	194271.168
10.08.2016	196370.72
11.08.2016	195700.96
12.08.2016	194986.0123
15.08.2016	194729.1975
16.08.2016	195495.9825
18.08.2016	197093.5725
19.08.2016	196010.325
22.08.2016	194094.09
23.08.2016	194542.23
24.08.2016	194651.355
25.08.2016	192829.8546
26.08.2016	193043.0764
29.08.2016	191658.3833
30.08.2016	192720.5204
31.08.2016	191475.738
01.09.2016	190839.0033
02.09.2016	191121.7462
05.09.2016	192579.2945
06.09.2016	192866.5032
07.09.2016	196240.1892
08.09.2016	195801.9492
09.09.2016	194602.9175
13.09.2016	193267.6899
14.09.2016	191694.3008
15.09.2016	192861.6922
19.09.2016	191707.1159
20.09.2016	191849.9963

21.09.2016	191775.8014
22.09.2016	194711.5757
23.09.2016	194834.86
26.09.2016	195314.8651
27.09.2016	195402.5499
28.09.2016	193956.3162
29.09.2016	194065.7428
30.09.2016	194305.1365
03.10.2016	193448.7159
04.10.2016	192048.5351
05.10.2016	186601.3864
06.10.2016	185938.9861
07.10.2016	184526.5768
10.10.2016	185099.7832
11.10.2016	184942.1631
12.10.2016	184661.3795
13.10.2016	184916.9855
14.10.2016	184741.44