4-1-2010

Stochastic Modeling and Statistical Analysis

Ling Wu
University of South Florida

Stochastic Modeling and Statistical Analysis

by

Ling Wu

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Mathematics and Statistics
College of Arts and Sciences
University of South Florida

Major Professor: Gangaram S. Ladde, Ph.D.
Kandethody M. Ramachandran, Ph.D.
Wonkuk Kim, Ph.D.
Marcus Mcwaters, Ph.D.

Date of Approval:
April 1, 2010

Keywords: Hybrid System, Nonlinear Models, Option Pricing, Forecasting, ARIMA

Dedication

I dedicate this dissertation to my husband Hu, Xuequn. Without his patience, understanding, support, and most of all love, the completion of this work would not have been possible.

Acknowledgments

I would like to express the deepest appreciation to my major professor Dr. Gangaram S. Ladde, who has supported me throughout my dissertation with his patience, knowledge, encouragement and effort. Without him this dissertation would not have been written. One simply could not wish for a better or friendlier as an individual with demanding and challenging supervisor.

I would like to thank my committee member, Dr. Kandethody M. Ramachandran, Dr. Wonkuk Kim, and Dr. Marcus McWaters, for their support and guidance to complete my research work.

In addition, I would like to express my appreciation to Dr. Christos P. Tsokos for his comments, suggestions to this work.

Table of Contents

List of Tables

List of Figures

Stochastic Modeling and Statistical Analysis

Ling Wu

ABSTRACT

The objective of the present study is to investigate option pricing and forecasting problems in finance. This is achieved by developing stochastic models in the framework of classical modeling approach.

In this study, by utilizing the stock price data, we examine the correctness of the existing Geometric Brownian Motion (GBM) model under standard statistical tests. By recognizing the problems, we attempted to demonstrate the development of modified linear models under different data partitioning processes with or without jumps. Empirical comparisons between the constructed and GBM models are outlined.

By analyzing the residual errors, we observed the nonlinearity in the data set. In order to incorporate this nonlinearity, we further employed the classical model building approach to develop nonlinear stochastic models. Based on the nature of the problems and the knowledge of existing nonlinear models, three different nonlinear stochastic models are proposed. Furthermore, under different data partitioning processes with equal and unequal intervals, a few modified nonlinear models are developed. Again, empirical comparisons between the constructed nonlinear

stochastic and GBM models in the context of three data sets are outlined.

Stochastic dynamic models are also used to predict the future dynamic state of processes. This is achieved by modifying the nonlinear stochastic models from constant to time varying coefficients, and then time series models are constructed. Using these constructed time series models, the prediction and comparison problems with the existing time series models are analyzed in the context of three data sets. The study shows that the nonlinear stochastic model 2 with time varying coefficients is robust with respect different data sets.

We derive the option pricing formula in the context of three nonlinear stochastic models with time varying coefficients. The option pricing formula in the frame work of hybrid systems, namely, Hybrid GBM (HGBM) and hybrid nonlinear stochastic models are also initiated.

Finally, based on our initial investigation about the significance of presented nonlinear stochastic models in forecasting and option pricing problems, we propose to continue and further explore our study in the context of nonlinear stochastic hybrid modeling approach.

Chapter 1
Review and Basic Concepts

1.0 Introduction

Financial mathematics derives and extends the mathematical or numerical models that are suggested by financial economists. Stochastic process is widely used here to obtain the fair price of derivatives of an asset. In this chapter, we first review some financial terminologies and methodologies, in Sections1.1. In Section 1.2, we present the development of stochastic models. General stochastic differential equations and $It\hat{o}-Doob$ formula are reviewed in Section 1.3. Furthermore, the least square estimation method is reviewed to estimate the parameters in Section 1.4. Finally, the maximum likelihood estimation method of time series model (ARMA model) is outlined in Section 1.5.

1.1 Financial Mathematics

1.1.1 Fundamental Concepts

During 1600s, Tulip dealing was big business in Holland. Flower growers and dealers were trading in options to guarantee prices. Until 1700s, options were declared illegal in London. The Investment Securities Act of 1934 created the Securities and Exchange Commission (SEC), and gave the SEC the power to regulate options. In April 26, 1973, the Chicago Board Option Exchange (CBOE) started trading and listed 16 call options on 16 stocks. A few years later, CBOE began trading put option, and ten years later, CBOE began trading Index option. On the first day of trading in 1973, 911 contracts traded. Today, the CBOE's average daily volume consistently exceeds one million contracts per day [4]. The concept of financial derivatives plays an important role in an interconnected financial world.

Definition 1.1.1 Derivatives: Derivatives are financial instruments whose value is derived from the value of something else. They generally take the form of contracts under which the parties agree to payments between them based upon the value of an underlying asset or other data at a particular point in time [2, 4, 19].

The main types of derivatives are futures, forwards, options and swaps. The main use of derivatives is to minimize risk for one party while offering the potential for a high return (at increased risk) to another. In a short term, the main use of derivatives is in risk management. The diverse range of potential underlying assets and payoff alternatives lead to a huge range of derivatives contracts available to be traded in the market. One of the most important derivatives is option. In the following, we define option, and outline different types options.

Definition 1.1.2 Options: Options are financial instruments that convey the right, but not the obligation to engage in a future transaction on some underlying security. Financial instruments are cash, evidence of an ownership interest in an entity, or a contractual right to receive, or deliver, cash or another financial instrument [2, 4, 19].

For example, buying a call option provides the right to buy a specified amount of a security at a set strike price at some time on or before expiration, while buying a put option provides the right to sell. There are 4 kinds of options:
(i) European option: An option that may only be exercised on expiration.
(ii) American option: An option that may be exercised on any trading day on or before expiration.
(iii) Bermuda option: An option that may be exercised only on specified dates on or before expiration.
(iv) Barrier option: Any option with the general characteristic that the underlying security's price must reach some trigger level before the exercise can occur.

Definition 1.1.3 Strike price (K): For an option, the strike price (K) or exercise price, is the key variable in a derivatives contract between two parties. Where the contract requires delivery of the underlying instrument, the trade will be at the strike price, regardless of the spot price (market price S) of the underlying instrument at that time. Strike price is the fixed price at which the owner of an option can purchase, in the case of a call, or the fixed price at which the owner of an option can sell, in the case of a put, the underlying security or commodity [2, 4, 19].

The concepts of payoff for options are defined as below.

Definition 1.1.4 Payoff: The payoff for a call option at time T is Max$\{(S_T - K); 0\}$, or formally $(S_T - K)^+$.The payoff for a put option at time T is Max$\{(K - S_T); 0\}$, or formally $(K - S_T)^+$. T is the maturity time at which the derivative contract expires [2, 4, 19].

In the following, we define the concept of hedge in finance.

Definition 1.1.5 Hedge: A hedge is an investment that is taken out specifically to reduce or cancel out the risk in another investment [2, 4, 19].

Hedging is a strategy designed to minimize exposure to an unwanted business risk, while still allowing the business to profit from an investment activity. Typically, a hedger might invest in a security that he/she believes to be under-priced relative to its "fair value", and combines this with a short sale of a related security or securities. Thus the hedger is indifferent to the movements of the market as a whole, and is interested in only the performance of the 'under-priced' security relative to the hedge.

1.1.2 Option Pricing

Modern option pricing techniques, usually using stochastic calculus, are often considered among the most mathematically complex of all applied areas of finance. In 1959, M. F. M. Osborne wrote a paper "Brownian Motion in the Stock Market" [36]. In 1964, another paper, by A. James Boness, focused on options. In his work, entitled "Elements of a Theory of Stock Option Value", Boness developed a pricing model that made a significant theoretical jump from that of his predecessors [8]. More significantly, his work served as a precursor to that of Fischer Black and Myron Scholes who in 1973 introduced their landmark option pricing model – Black Scholes Model [33]

There are two types of option pricing approaches namely discrete and continuous processes. In the following, we briefly describe the discrete time option pricing process.

Discrete Time Option Pricing Process (Binomial Tree): We suppose that the market is observable at times $0 = t_0 < t_1 < t_2 < \ldots < t_N = T$. On each time period $[t_i, t_{i+1}]$, the stock price follows the binary model. After i time periods, the stock has 2^i possible values. We also suppose that the length of any time period has the same length δt. We define $\{\tilde{S}_k\}_{k\geq 0}$ to be the discounted stock process, such that, $\tilde{S}_k = e^{-kr\delta t} S_k$, where r is the interest rate. Figure 1.1.1 is the binomial tree for a stochastic stock price process.

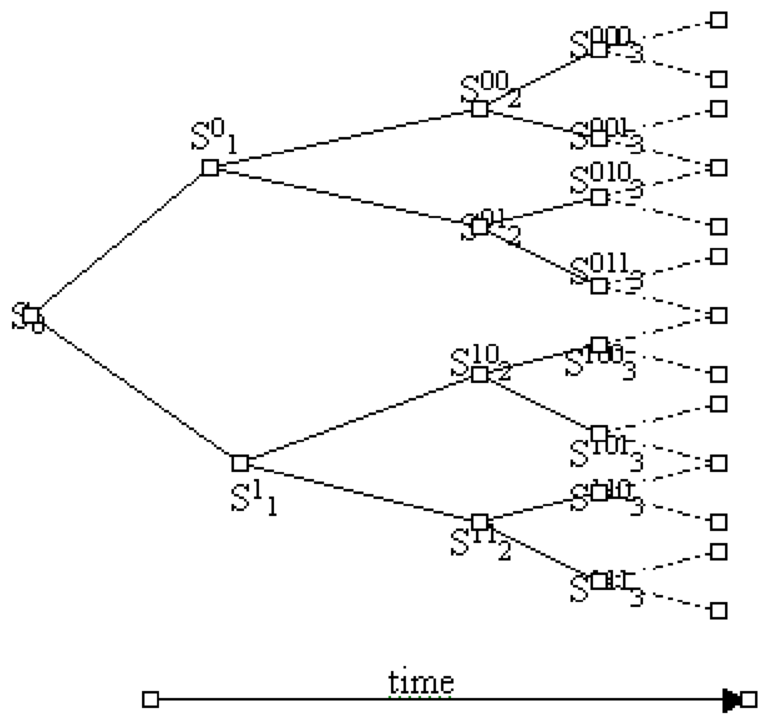


Figure 1.1.1 Binomial Tree for a Stochastic Stock Price Process

In the following, before we state very important result, we first give some definitions:
In probability theory, when we talk about a random variable, we specify a probability triple (Ω, F, P), where Ω is the sample apace, *F* is a collection of subsets of Ω ,also called σ -field, and *P* is the probability of each event $A \in F$.

To specify a stochastic process, we required not only a single σ -field, F, but also an increasing sequence of sub σ - algebras, $F_n \subseteq F_{n+1} \subseteq ... \subseteq F$. The collection $\{F_n\}_{n\geq 0}$ is called a filtration and the quadruple $(\Omega, F, \{F\}_{n\geq 0}, P)$ is called a filtered probability space.

In probability theory, suppose that $(\Omega, F, \{F\}_{n\geq 0}, P)$ is a filtered probability space. The sequence of random variables $\{X_n\}_{n\geq 0}$ is a martingale with respect to P and $\{F_n\}_{n\geq 0}$ if
$E[|X_n|] < \infty,$ and $E[X_{n+1} | F_n] = X_n$, for all n.

Theorem 1.1.1 (The binomial representation theorem) [19]: Suppose that the measure Q is such that the discounted binomial price process $\{\widetilde{S}_n\}_{0\leq n\leq N}$ is a Q-martingale. If $\{\tilde{V}_n\}_{0\leq n\leq N}$ is any other (Q,$\{F_n\}_{n\geq 0}$,)-martingale, then there exists an $\{F_n\}_{n\geq 0}$ – predictable process $\{\phi_n\}_{n\geq 1}$ (portfolio process) such that

$$\tilde{V}_n = \tilde{V}_0 + \sum_{j=0}^{n-1} \phi_{j+1}(\tilde{S}_{j+1} - \tilde{S}_j).$$

Remark 1.1.1 [19]: From Theorem 1.1.1, we know that if $\{\widetilde{V}_i\}_{i\geq 0}$ is the discounted price of a claim (European call or put option), then such a predictable process $\{\phi_i\}_{i\geq 1}$ (portfolio process) arises as the stock holding when we construct out replicating portfolio.
There are three steps to pricing and hedging a claim C_T at time T:
(i) Find a probability measure Q under which the discounted stick price (with its natural filtration) is a martingale.
(ii) Form the discounted value process, $\widetilde{V}_i = e^{-ri\delta t}V_i = E^Q[e^{-rt}C_T \mid F_i]$.
(iii) Find a predictable process $\{\phi_i\}_{1\leq i\leq N}$ such that $\Delta\widetilde{V}_i = \phi_i \Delta\widetilde{S}_i$.

In the following, we present a very fundamental result in the theory of continuous time option pricing process.

Theorem 1.1.2 [19]: The fundamental theorem of continuous option pricing is:
(i) There is a probability measure Q equivalent to P under which the discounted stock price $\{\widetilde{S}_t\}_{t\geq 0}$ is a martingale.

(ii) Under the probability measure Q, suppose that a claim at time T is given by the non-negative random variable $C_T \in F_T$. If $E^Q[C_T^2] < \infty$.

Then, the claim is replicable and the value at time t of any replicating portfolio is given by, $V(S,t) = E^Q[e^{-r(T-t)}C_T \mid F_t]$, in particular, the fair price at time 0 for the option is

$$V_0 = E^Q[e^{-rT}C_T] = E^Q[\widetilde{C}_T].$$

S_t is a stock price process, $\tilde{S}_t$ is the discounted stock price process.

Theorem 1.1.3 Black-Scholes Model [7,19]: Under the following assumptions:

(i) There is no credit risk, we can buy and sell cash bond without credit risk. And there is only market risk, which means the stock price can go up and down arbitrarily.

(ii) The market is maximally efficient, that is, it is infinitely liquid and does not exhibit any friction. This means all relevant information is fully reflected and priced in the stock price, and there are no any other additional costs.

(iii) Continuous trading is possible.

(iv) The time evolution of the asset price is stochastic process and called geometric Brownian motion, the mathematical expression is $dS_t = \mu S_t dt + \sigma S_t dW_t$. S_t is a stock process, μ and σ are constant.

(v) There is no dividend.

(vi) The underlying asset is arbitrarily divisible. And the market is arbitrage free, which means the market prices do not allow for profitable arbitrage.

The value at time t of a European option whose payoff at maturity $C_T = f(S_T)$, is

$V_t = F(t, S_t)$, where $F(t,x) = e^{-r(T-t)} \int_{-\infty}^{\infty} f(xe^{(r-\frac{1}{2}\sigma^2)(T-t)+\sigma y\sqrt{T-t}}) \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy$.

For European call option, suppose that $f(S_T) = (S_T - K)_+$. Then, let $\theta = T - t$,

$$F(t,x) = x\Phi(d_1) - Ke^{-r\theta}\Phi(d_2),$$

where,

$\Phi(.)$ is the standard normal cumulative distribution function,

$$d_1 = \frac{\log\frac{x}{K} + (r + \frac{\sigma^2}{2})\theta}{\sigma\sqrt{\theta}},$$

and $d_2 = d_1 - \sigma\sqrt{\theta}$.

For European put option, suppose that $f(S_T) = (K - S_T)_+$, then let $\theta = T - t$,

$$F(t,x) = Ke^{-r\theta}\Phi(-d_2) - x\Phi(-d_1),$$

where,

$\Phi(.)$ is the standard normal cumulative distribution function,

$$d_1 = \frac{\log\frac{x}{K} + (r + \frac{\sigma^2}{2})\theta}{\sigma\sqrt{\theta}},$$

and $d_2 = d_1 - \sigma\sqrt{\theta}$.

The Black-Scholes formula is based on assumption of log-normal stock diffusion with constant volatility, that is, the stock price process is a stochastic process described by the following stochastic differential equation of the form:

$$dS_t = \mu S_t dt + \sigma S_t dW_t.$$

This has become the universal benchmark for option pricing. But, we are all aware of that it is flawed. The drift and volatility are not a constant. In 1973, Merton first allows the drift and volatility to be a deterministic function of time. Later on, other models allow not only time, but also state dependence of the volatility. This method is called as a local volatility approach.

There are some very famous local volatility models. For example, Merton's model (1973) takes the form $dr_t = \alpha dt + \sigma dW_t$ [33]. Vasicek (1977) deriving an equilibrum model of discount bond price process by using the Ornstein-Uhlenbeck process [41]. It takes the form $dr_t = (\alpha + \beta r_t)dt + \sigma dW_t$. Dothan (1978) used model $dr_t = \sigma r_t dW_t$ in valuing discount bonds [18]. And Brennan and Schwartz (1980) used model $dr_t = (\alpha + \beta r_t)dt + \sigma r_t dW_t$ in deriving a numerical model for convertible bond prices [11]. These are called linear models. Other nonlinear models such as Cox-Ingersoll-Ross model [15] and Black-Karasinski model [6], take nonlinear functions of the asset price at time t as the drift and/or volatility. In the next section, we will introduce how to develop the stochastic process.

1.2 Development of Stochastic Modeling

In this section, by following a real stochastic modeling approach [26], we outline the derivation of stochastic model of stock price. This is based on the basic descriptive statistical approach. It utilizes the Random Walk process to initiate a scope and a development of stochastic models of dynamic processes. Here, a state is a conceptually common term and description of processes in the sciences and engineering is used, for example, a “state” can be “distance” traveled by an object in the physical process, “concentration” of a chemical substance in a chemical process, “number of species” in a biological process, and in social science or this thesis, state is the “price” of an asset in a sociological process.

1.2.1 Conditions of Stochastic Process – Random Walk

Let S_t be a price of a stock at time t. The price of the asset is observed over an interval of [t, t+Δt], where Δt is a small increment in t. Without loss in generality, we assume that Δt is positive. The price process is under the influence of random perturbations. We experimentally observe price process $S_{t_0} = S_t$, $S_{t_1}, S_{t_2}, ..., S_{t_n} = S_{t+\Delta t}$ of a stock at $t_0 = t$, $t_1 = t + \tau$, $t_2 = t + 2\tau, \ldots$, $t_k = t + k\tau, \ \ldots, \ t_n = t + \Delta t$, over the time interval [t, t+Δt], where n belongs to {1, 2, 3, …} and τ=Δt/n. These observations are made under the following conditions:

C_1. The stock price is under the influence of independent and identical random impulses that are taken place at $t_1, t_2, ..., t_k, ..., t_n$.

C_2. The influence of a random impact on the stock price is observed on every time subinterval of length τ.

C_3. For each $k \in I(1,n)=\{1,2,\ldots,k,\ldots n\}$, it is assumed that the stock price is either increased by ΔS_{t_k} or decreased by ΔS_{t_k}. We refer ΔS_{t_k} as a microscopic/local experimental or knowledge-base observed increment to the price of the stock per impact on the subinterval of length τ.

C_4. It is assumed that ΔS_{t_k} is constant for $k \in I(1,n)$ and is denoted by $\Delta S_{t_k} \equiv Z_k = Z$ with $|Z_k| = \Delta S > 0$. Thus, for each $k \in I(1,n)$, there is a constant random increment Z of magnitude ΔS to the price of the stock per impact on the subinterval of length τ.

In short, from the first three conditions, under n independent and identical random impacts, the initial price and n knowledge-base observed random increments Z_k of constant magnitude ΔS in the state at $t_1, t_2, \ldots, t_k, \ldots, t_n$ over the given interval $[t, t+\Delta t]$ of length Δt are:

$$
\begin{aligned}
&S_{t_0} = S_t \\
&S_{t_1} - S_{t_0} = Z_1 \\
&S_{t_2} - S_{t_1} = Z_2 \\
&\ldots\ldots \\
&S_{t_k} - S_{t_{k-1}} = Z_k \\
&\ldots\ldots \\
&S_{t_n} - S_{t_{n-1}} = Z_n
\end{aligned}
$$

Z_k's are defined by

$$
Z_k = \begin{cases} \Delta S, & for \quad positive \quad increment \\ -\Delta S, & for \quad negative \quad increment \end{cases}
$$

The 4th condition implies that they are mutually independent random variables. From the above discussion, the prices S_{t_k} and S_{t_n} are random impacts at the k-th instance and the final time on the price process respectively. Moreover, they are expressed by:

$$S_{t+k\tau} - S_{t_k} = S_t + \sum_{i=1}^{k} Z_i \text{ and } S_{t+\Delta t} = S_t + \sum_{i=1}^{n} Z_i$$

where $\sum_{i=1}^{n} Z_i$ is referred as an aggregate increment to the given price $S \equiv S_t$ of the stock at the given time t over the interval [t, t+Δt] of length Δt.

In this case, the aggregate change of the price of the stock $S_{t+\Delta t} - S_t$ under n observations of the system over the given interval [t, t+Δt] of length Δt is described by

$$S_{t+\Delta t} - S_t = n\frac{\sum_{i=1}^{n} Z_i}{n} = \frac{\Delta t}{\tau} S_n, \tag{1.2.1}$$

where $S_n = \frac{1}{n}\sum_{i=1}^{n} Z_i$. S_n is the sample average of the aggregate price incremental data.

1.2.2 Mean and Variance of Aggregate Change of Price

For each random impact and any real number p, 0<p<1, it is assumed that

$$P\{Z_k = \Delta S > 0\} = p \text{ and } P\{Z_k = -\Delta S \le 0\} = 1 - p = q$$

It is clear that $S_{t_k} - S_{t_0}$ is a discrete-time-real-valued stochastic process which is the sum of k independent Bernouli random variables Z_i , i=1, 2, …, k and k=1, 2, …,n. We note that for each k, $S_{t_k} - S_{t_0}$ is binomial random variable random variable with parameters (k, p). Moreover, the random variable $S_{t_k} - S_{t_0}$ takes values from the set {-kΔS, (1-k)ΔS, …, (2m-k)ΔS, …, kΔS}. The stochastic price process $S_{t_k} - S_{t_0}$ is called a Random Walk process. In particular, for k=n, let m be a number of positive increments ΔS to the price of the stock out of total n changes. (n-m) is the number of negative increment -ΔS to the price of the stock out of total n changes. Furthermore, m∈I(0, n), we have that

$$S_n = \frac{1}{n}[m \frac{\sum_{i \in I_+(0,n)} Z_i}{m} - (n-m) \frac{\sum_{i \in I_-(0,n)} | Z_i |}{n-m}]$$

$$= \frac{1}{n}[m\Delta S - (n-m)\Delta S]$$

$$= \frac{1}{n}[(2m-n)\frac{1}{n}\sum_{i=1}^{n} | Z_i |]$$

$$= \frac{1}{n}[(2m-n)S_n^+], \tag{1.2.2}$$

where $I_+(0,n)$ and $I_-(0,n)$ are denoted by $I_+(0,n) = \{i \in I(0,n) : | Z_i | = Z_i\}$ and $I_-(0,n) = \{i \in I(0,n) : | Z_i | = -Z_i\}$ respectively, and $S_n^+ = \frac{1}{n}\sum_{i=1}^{n} | Z_i |$.

Thus from (1.2.1) and (1.2.2), we get

$$S_{t+\Delta t} - S_t = \frac{1}{n}(2m-n)\frac{S_n^+}{\tau}\Delta t . \tag{1.2.3}$$

Furthermore, in this case, the aggregate change of price of the stock $S(t+\Delta t) - S(t)$ over the time interval of length Δt under n identical random impacts on the stock over the given interval [t, t+Δt] of time is also described by:

$$S_{t+\Delta t} - S_t = \sum_{i=1}^{n} Z_i$$

= total amount of positive increment – total amount of negative increment

$$= m\Delta S - (n-m)\Delta S = (2m-n)S_n^+ = \frac{1}{n}(2m-n)\frac{S_n^+}{\tau}\Delta t .$$

This is identical with the expression in (1.2.3). So, the mean of the aggregate change of the price of the stock $S_{t+\Delta t} - S_t$ over the interval [t, t+Δt] is given by:

$$E[S_{t+\Delta t} - S_t] = \sum_{m=0}^{n} \frac{1}{n}(2m-n)\frac{n!}{(n-m)!} p^m (1-p)^{n-m} \frac{S_n^+}{\tau}\Delta t$$

$$= (p-q)\frac{S_n^+}{\tau}\Delta t \,,$$

and the variance is:

$$Var[S_{t+\Delta t} - S_t] = E[S(t+\Delta t) - S(t) - (p-q)\frac{S_n^+}{\tau}\Delta t]^2$$

$$\begin{aligned} &= \sum_{m=0}^{n} [\frac{1}{n}(2m-n)\frac{S_n^+}{\tau}\Delta t]^2 \,\mathrm{P}(m, np) - [(p-q)\frac{S_n^+}{\tau}\Delta t]^2 \\ &= 4pq\frac{(S_n^+)^2}{\tau}\Delta t. \end{aligned} \tag{1.2.4}$$

S_n^+ / τ (or $\frac{\Delta S}{\tau}$) and $(S_n^+)^2 / \tau$ (or $\frac{(\Delta S)^2}{\tau}$) are microscopic or local stock average increment and sample microscopic or local average square increment per unit time over the uniform length of sample subinterval $[t_{k-1}, t_k]$, k=1,2, …, n of interval [t, t+Δt].

1.2.3 Wiener Process

In reality we note that there are restrictions on ΔS and τ. Similarly, the parameter p cannot be taken arbitrary. Moreover, the price of the stock cannot go to “infinity” on an interval whose length is small. In view of these considerations, for sufficiently large n, the following conditions seem to be natural:

$$S_{t+\Delta t} - S_t = n\Delta S \,,\; \Delta t = n\tau \,,$$

$$4pq = (p+q)^2 - (p-q)^2 = 1 - (p-q)^2 \,,$$

$$\lim_{\tau\to 0}[\frac{(S_n^+)^2}{\tau}] = 2D, \quad \lim_{\Delta S\to 0}\lim_{\tau\to 0}[(p-q)\frac{S_n^+}{\tau}] = C,$$

and $$\lim_{\Delta S\to 0}\lim_{\tau\to 0} 4pq = 1.$$

Here C and D are certain constants. C is called drift, and D is called diffusion coefficient. Moreover, C can be interpreted as the average/mean/expected rate of change of price of the stock per unit time, and D can be interpreted as the mean square rate of price change of the stock per unit time over an interval of length Δt.

From the above discussion, we obtain

$$\lim_{\Delta S\to 0}\lim_{\tau\to 0} E[S_{t+\Delta t} - S_t] = C\Delta t \tag{1.2.5}$$

and $$\lim_{\Delta S\to 0}\lim_{\tau\to 0} Var[S_{t+\Delta t} - S_t] = 2D\Delta t. \tag{1.2.6}$$

Now, we define $y(t,n,\Delta t) = \dfrac{S_{t+\Delta t} - S_t - n(p-q)S_n^+}{\sqrt{4npq(S_n^+)^2}}$. By Central Limit Theorem [37], we conclude that the process $y(t,n,\Delta t)$ is approximated by standard normal random variable for each t.

Moreover, from $n = \dfrac{\Delta t}{\tau}$, we have $y(t,n,\Delta t) = \dfrac{S_{t+\Delta t} - S_t - (p-q)\dfrac{S_n^+}{\tau}\Delta t}{\sqrt{4pq\dfrac{(S_n^+)^2}{\tau}\Delta t}}$,

Hence, $$\lim_{\Delta x\to 0}\lim_{\tau\to 0} y(t,n,\Delta t) = \lim_{\Delta S\to 0}\lim_{\tau\to 0}[\frac{S_{t+\Delta t} - S_t - (p-q)\frac{S_n^+}{\tau}\Delta t\Delta t}{\sqrt{4pq\frac{(S_n^+)^2}{\tau}\Delta t}}]$$

$$= \frac{S_{t+\Delta t} - S_t - C\Delta t}{\sqrt{2D\Delta t}}.$$

For fixed Δt, the random variable $\lim_{\Delta S \to 0} \lim_{\tau \to 0} y(t, n, \Delta t)$ has standard normal distribution. Now, by rearranging the above expression, we get

$$S_{t+\Delta t} - S_t = C\Delta t + \sqrt{2D}\sqrt{\Delta t}[\lim_{\Delta S \to 0} \lim_{\tau \to 0} y(t, n, \Delta t)],$$

and denoting $\sqrt{\Delta t}[\lim_{\Delta S \to 0} \lim_{\tau \to 0} y(t, n, \Delta t)] = \Delta W_t = W_{t+\Delta t} - W_t$, it can be rewritten as:

$$S_{t+\Delta t} - S_t = C\Delta t + \sqrt{2D}\Delta W_t, \tag{1.2.7}$$

where W_t is Wiener process. Thus the aggregate price change of the stock $S_{t+\Delta t} - S_t$ under independent and identical random impacts over the given interval [t, t+Δt] is interpreted as the sum of the average/mean/expected price change of the stock CΔt, and the mean square price change of the stock $\sqrt{2D}\Delta W_t$ due the random environmental perturbations.

If Δt is very small, then its differential $dt = \Delta t$, and the Itô-Doob dS is defined by:

$$dS_t = Cdt + \sqrt{2D}dW_t, \tag{1.2.8}$$

where C and D are the drift and the coefficients, respectively. Equation in (1.2.8) is called the Itô-Doob type stochastic differential equation.

This Random Walk modeling process can be applied to formulate mathematical model in finance. Let S_t be either a rate of price/value of an asset per unit item/size and per unit time. The specific rate of price/value or the rate of price/value is observed over an interval of [t, t+Δt], where Δt is a small increment in t. Without loss in generality, we assume that Δt>0. The process is under the influence of exogenous or endogenous random perturbations of national/international/commerce/trade/monetary/social welfare polices. As the result of this, the S_t is affected by the random environmental perturbations. By following the development of the above Random Walk Model, its mathematical description is as described in (1.2.8).

Furthermore for very small Δt, the Random Walk modeling approach leads to the formulation of mathematical model in finance. Usually it takes the following form:

$$dS_t = \mu dt + \sigma dW_t. \qquad (1.2.9)$$

We note that if S_t is the specific rate of the price/value at time t, then μ (μ=C) is called a measure of the average specific rate (per capital growth/decay rate) of the price/value of the asset at the time t, and σ (σ^2=2D) is called the volatility which measures the standard deviation of the specific rate (per capital growth/decay rate) of the price/value at a time t over an interval of small length Δt=dt.

Remark 1.2.1: Here, for the sake of simplicity, we only assume that ΔS_{t_n} is constant. Actually, it need not be constant. Moreover, it can be any smooth function of t or S_t. The expected value of the increment E[$S_{t+\Delta t}$- S_t] can be replaced by the conditional expected value E[$S_{t+\Delta t}$- S_t| S_t]. C and D may be any smooth function of time t and the state S, satisfying certain conditions. We will discuss this issue in the next section.

1.3 General Stochastic Differential Equations

In this section, we outline the fundamental result that assures to undertake the study of dynamic modeling.

In financial engineering, it is common to model a continuous time price process described by the $It\hat{o} - Doob$ type stochastic differential equation. A general stochastic differential equation takes the form:

$$dS_t = \mu(S_t, t)dt + \sigma(S_t, t)dW_t, \quad S_{t_0} = S_0. \qquad (1.3.1)$$

Here, $t \geq t_0$, W_t is a Brownian motion, and $S_t > 0$, which is the price process.

Under the following smoothness conditions on functions μ and σ, one can establish the existence and uniqueness of the solution of process of (1.3.1).

Theorem 1.3.1 (Existence and Uniqueness Theorem) [23]: Suppose that there exist some constants K, L > 0 such that the functions μ and σ in (1.3.1) satisfy the following conditions

$$\mu^2(S,t)+\sigma^2(S,t)\le K(1+S^2) \qquad \text{(Growth Condition)} \qquad (1.3.2)$$

and

$$|\mu(S_1,t)-\mu(S_2,t)|+|\sigma(S_1,t)-\sigma(S_2,t)|\le L|S_1-S_2| \text{ (Lipschitz Condition)} \quad (1.3.3)$$

Then, it can be shown that the stochastic differential equation in (1.3.1) has a unique solution.

This is very important and well known in financial engineering, because the unique solution of the stochastic equation in (1.3.1) is a stochastic process adapted to Brownian filtration $\{F_t\}_{t\ge 0}$ [23].

These two conditions, growth condition (1.3.2) and Lipschitz condition (1.3.3) (named after Rudolf Lipschitz), are sufficient conditions, not necessary conditions for the existence and a uniqueness. In this dissertation, all models represented by stochastic differential equations must satisfy these two conditions.

In equation (1.3.1), it is known that W_t is a Brownian motion, which is continuous everywhere, but it is not differentiable anywhere. To find the information about the solution of the equation (1.3.1), we need a way to take integral of a stochastic process. In 1951, $It\hat{o}$ Kiyoshithe published his very famous $It\hat{o}$'s formula.

Theorem 1.3.2 ($It\hat{o}-Doob$ Formula) [21]: Let *f* be a function such that $f\in C[[a,b)\vee R,R]$, its partial derivatives $\frac{\partial f}{\partial t}$, $\frac{\partial f}{\partial x}$ and $\frac{\partial^2 f}{\partial x^2}$ exist and are continuous. Then we have

$$f(t,W_t)-f(0,W_0)=\int_0^t\frac{\partial f}{\partial x}(s,W_s)dW_s+\int_0^t\frac{\partial f}{\partial s}(s,W_s)ds+\frac{1}{2}\int_0^t\frac{\partial^2 f}{\partial s^2}(s,W_s)ds. \quad (1.3.4)$$

Moreover, $It\hat{o}-Doob$'s formula in differential form is represented by

$$df(t,W_t)=\frac{\partial}{\partial S}f(t,W_t)dW_t+\frac{\partial}{\partial t}f(t,W_t)dt+\frac{1}{2}\frac{\partial^2}{\partial S^2}f(t,W_t)dt \qquad (1.3.5)$$

This is fundamental result in $It\hat{o}-Doob$ stochastic calculus [10]. We use this formula frequently.

1.4 Least Square Method

The credit for discovery of the method of Least squares is given to Carl Fridrich Gauss who used the procedure in the early part of the nineteenth century [33]. It is the most widely used technique in data analysis. The least square technique can be interpreted as a method of fitting data. The best fit in the least-squares sense is an instance of the model for which the sum of squared residuals has its least value. A residual is the difference between an observed value and the value predicted by the model. Unlike maximum likelihood [37], the least square estimation does not require the distribution assumption. When the parameters appear linearly in an expression, then the estimation problem can be solved in closed form. We recall the formula of the linear model that y is related linearly to the regressor variable x's as:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} + \varepsilon_i \qquad (i = 1, 2, \ldots, n; n \geq k+1) \tag{1.4.1}$$

The ideal conditions of the least square model are

a) ε_i is model error, with mean zero,

b) the ε_i are uncorrelated, and have common variance (homogeneous variance).

1.5 Maximum Likelihood Estimation Method of ARIMA Model

ARIMA(p,d,q) (autoregressive integrated moving average) process provides a very general class of models for modeling and forecasting dynamic phenomena in science and engineering which can be stationarized by applying transformations, namely, difference, logarithm, or other transformations. Here, p stands for the number of autoregressive terms, called autoregressive order; d is the order (or degree) of difference of the time series; and q is the number of lagged forecast errors in the prediction equation, called moving average order. ARIMA(p,d,q) models are ARMA(p,q) models with d*th*-order difference transformation. First, we introduce the difference filter as follows:

$$(1 - B)^d \tag{1.5.1}$$

where B is called backward shift operator, and $Bz_t = z_{t-1}$, $B^m z_t = z_{t-m}$, and $z_t, t = 1, \ldots, n$ is a time series data set. In ARIMA(p,d,q) models, after taking d*th*-order difference transformation, we suppose that the time series is stationary. For stationary time series, ARMA(p,q) models have following form:

$$z_t = \phi_1 z_{t-1} + ... + \phi_p z_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + ... + \theta_q \varepsilon_{t-q}, \tag{1.5.2}$$

that is,

$$(1 - \phi_1 B - \phi_2 B^2 ... + \phi_p B^p) z_t = (1 - \theta_1 B - \theta_2 B^2 - ... - \theta_q B^q) \varepsilon_t,$$

or,

$$\phi(B) z_t = \theta(B) \varepsilon_t, \tag{1.5.3}$$

where $\phi(B)$ and $\theta(B)$ are polynomials of degree p and q in B [10].

Therefore, ARIMA(p,d,q) model can be represented as

$$\phi(B)(1-B)^d z_t = \theta(B) \varepsilon_t, \tag{1.5.4}$$

where d, B, $\phi(B)$ and $\theta(B)$ are as defined above.

Even though, the values of p and q can be determined by the number of significant spikes in PACF (partial auto correlation function) and ACF (auto correlation function) plots respectively. There are several models that are adequately represented by a give time series. Hence, criterions such as AIC (Akaike's information criterion) and BIC (Bayesian information criterion) are used to selecte the best model. In our study, we choose AIC, because BIC penalizes more with larger data sets. AIC was defined by Akaike in 1973 and takes the following form [3]:

$$AIC = -2\ln(L) + 2k, \tag{1.5.5}$$

where, L is maximized value of the likelihood function for the estimated model, k is the number of parameters in the model. If the model errors are assumed to be normally and independently distributed, RSS is the residual sum of square and is defined as $RSS = \sum_{i=1}^{n} \hat{\varepsilon}_i^2$, where n is the number of observations. Maximizing the likelihood, the AIC can be written as $AIC = 2k + n[\ln(\frac{2\pi \times RSS}{n}) + 1]$. After simplification and remove the unaffected constant term, AIC is simplifies to:

$$AIC = 2k + n[\ln(\frac{RSS}{n})]. \tag{1.5.6}$$

The unconditional log-likelihood function of a ARMA(p,q) model is defined by Box, Jenkins, and Reinsel in 1994 as follows [10]:

$$\ln(L) = -\frac{n}{2}\ln(2\pi\sigma_{\varepsilon}^{2}) - \frac{S(\phi,\mu,\theta)}{2\sigma_{\varepsilon}^{2}}, \tag{1.5.7}$$

where, $S(\phi,\mu,\theta)$ is the unconditional sum of residual square, exampled by

$$S(\phi,\mu,\theta) = \sum_{t=-\infty}^{n}[E(\varepsilon_t \mid \phi,\mu,\theta,z)]^2 \approx \sum_{M}^{n}[E(\varepsilon_t \mid \phi,\mu,\theta,z)]^2 , \tag{1.5.8}$$

where, $E(\varepsilon_t \mid \phi,\mu,\theta,z)$ is the conditional expected ε_t, given ϕ,μ,θ,z . M is a large integer such that the backforecast increment $| E(\varepsilon_t \mid \phi,\mu,\theta,z) - E(\varepsilon_{t-1} \mid \phi,\mu,\theta,z) |$ is less than any arbitrary predetermined small value for $t \le -(M+1)$.

Then problem of parameters estimation in ARIMA model reduces to the problem of finding out how to estimate of ϕ, θ and ε_t^2 so that $S(\phi,\mu,\theta)$ has minimum value. For example, the backforecasting for ARMA(1,1):

Given $z_t = \phi z_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}$, we rewrite as $\varepsilon_{t-1} = \dfrac{z_t - \phi z_{t-1} - \varepsilon_t}{\theta}$. If we let $\varepsilon_t = 0$, by giving ϕ and θ, we can recursively solve ε_{t-1}. Then parameters ϕ and θ can be estimated as those value which minimize $S(\phi,\mu,\theta)$.

After obtaining $\hat{\phi},\hat{\mu}$, and $\hat{\theta}$, which maximize the log-likelihood function (1.5.7), the estimator of σ_{ε}^2 is computed by

$$\hat{\sigma}_{\varepsilon}^{2} = \frac{S(\hat{\phi},\hat{\mu},\hat{\theta})}{n}. \tag{1.5.9}$$

Applying (1.5.9) and (1.5.7) in (1.5.5) and reducing the constant in AIC, (1.5.5) is expressed as

$$AIC = n\ln(\hat{\sigma}_{\varepsilon}^{2}) + 2k \tag{1.5.10}$$

Therefore, we can choose the ARIMA model with smallest AIC. The estimated parameters $\phi_1,\ldots,\phi_p$, and $\theta_1,\ldots,\theta_q$ by least square and maximum likelihood are not identical, but the difference is always trivial.

Statistical Model Identification Procedure 1.5.1 [38]: Now, we summarize the development of the ARIMA(p,d,q) model as follows:

i. Transform the original observations $S_t, t = 1,2,...,n$ into $V_t = f(S_t), t = 1,2,...,n$, if necessary.

ii. Seasonal differences chosen if needed using a variation on the Canova-Hansen test [14].

iii. Check for stationarity of $V_t, t = 1,2,...,n$ by determining the order of differencing d, according to KPSS test [22].

iv. Set $p + q \leq 5$, $p \leq 3$ and $q \leq 3$. List all possible set of (p, q).

v. For each possible set of (p, q), applying maximum likelihood method, to estimate the parameters $\phi_1,...,\phi_p$, and $\theta_1,...,\theta_q$ of each model.

vi. Computer AIC for each model.

vii. Choose the model which has the smallest AIC.

Chapter 2
Linear Stochastic Models

2.0 Introduction

Certain stochastic processes are functions of Brownian motion process and have many applications in financial engineering and sciences. Some special processes are solutions of $Itô - Doob$ type stochastic differential equations. Moreover, such processes also describe the stochastic behavior of an asset price in finance [23].

In this chapter, we introduce the well-known linear stochastic models, which are also called GBM (Geometric Brownian Motion) models. By following the historical model building process, we attempt to develop a stochastic model for stock market price system. As the part of the model building process, we employ two stock prices selected from Fortune 500 companies and one stock Index S&P500. The first step in the classical model building approach is to draw a sketch of the data set. The second step is to use a proper knowledge of the dynamic process and the given data set to estimate the parameters in functions.

In section 2.1, we briefly review a basic conceptual model – GBM model. We utilize statistical procedure to sketch a stock price data set and to estimate the parameters in the historical GBM model in Section 2.2. The Q-Q plot of residual error of model in Section 2.2 motivates to seek a modified version of GBM. By using the same modeling procedure, we discuss several different results with different data partitioning processes, in Section 2.3. Again, after studying the Q-Q plots of residual errors of models in Section 2.3, we introduce the different data partitioning schemes combined with jumps in Section 2.4. We give other examples in Section 2.5 and Section 2.6 using the same procedure. A few of conclusions are drawn in Section 2.7. The data sets we applied here are the two stocks selected from Fortune 500 companies and the S&P 500 index. The daily adjusted closing prices can be free downloaded from the website http://finance.yahoo.com.

2.1 GBM Models

In this dissertation, as we noted before, we will be following the classical model building process. For this purpose we use a data about the dynamic process of interest and some prior information about the dynamic process. In our case, we do have a data set about the stock prices selected from Fortune 500 companies and a prior well-known theoretical model – GBM model. Our initial attempt is to use the stock price data, the GBM model and the statistical techniques. To use these three basic components of modeling, first we need to perform the reduction process of converting the GBM model into linear regression equation. This reduction technique is systematically outlined in this section.

We initiate the usage of a classical modeling approach to develop suitable modified stochastic models for the price movement of individual stocks. For this purpose, we first utilize the recent trend in the literature that starts with a conceptual model, and attempt to fit a dataset into it. We begin with utilizing the existing Geometric Brownian Motion (GBM) model and try to fit a dataset into it, and then use the basic statistics to validate the model in the statistical framework. The commonly used benchmark for comparison is the well-known Black-Scholes model which is based on Geometric Brownian motion [32].

S_t is called GBM (Geometric Brownian Motion) process, that is, the solution of the following linear $It\hat{o}-Doob$ type stochastic differential equation

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \tag{2.1.1}$$

where μ and σ are some constants, μ is called drift, σ is called volatility, and W_t is a normalized Brownian motion process. Let K be any number greater than ($\mu^2 + \sigma^2$), and L be any number greater than $|\mu| + |\sigma|$. From this we conclude that equation (2.1.1) satisfies conditions (1.3.2) and (1.3.3). Applying $It\hat{o}-Doob$'s formula applied to $f(S_t) = \ln S_t$, we have

$$S_t = S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t}, \tag{2.1.2}$$

where, W_t is a Brownian motion process as usual. S_t is also called exponential Brownian motion process, since S_t takes the exponential form of W_t. As we already introduced in Chapter 1, one of the most important assumptions in Black-Scholes model [7] is that the stock price process is GBM process.

We want to use the historic stock data set to examine the GBM model (2.1.1), that is, we want to estimate the parameters μ and σ. Here, we try to use the least square method to estimate parameters in the GBM model.

In equation (2.1.1), the error term does not have common variance. It is related to S_t. This means that as the stock price increases, the variance also increases.
With a transformation $V_t = \ln S_t$, using $Itô - Doob$'s formula, we obtain

$$\begin{aligned} dV_t &= \frac{\partial}{\partial S_t}(\ln S_t)dS_t + \frac{1}{2}(\frac{\partial^2}{\partial S_t^2}(\ln S_t))(dS_t)^2 \\ &= \frac{1}{S_t}dS_t + \frac{1}{2}(-\frac{1}{S_t^2})(dS_t)^2 \\ &= \mu dt + \sigma dW_t - \frac{1}{2}\frac{1}{S_t^2}\sigma^2 S_t^2 (dW_t)^2 \\ &= (\mu - \frac{1}{2}\sigma^2)dt + \sigma dW_t. \end{aligned} \tag{2.1.3}$$

Thus,

$$d(\ln S_t) = (\mu - \frac{1}{2}\sigma^2)dt + \sigma dW_t \tag{2.1.4}$$

From the Euler type discretization [24] of the stochastic differential equation (2.1.4), we have

$$\ln S_t - \ln S_{t-1} = (\mu - \frac{1}{2}\sigma^2)\Delta t + \sigma(W_t - W_{t-1}). \tag{2.1.5}$$

Let $y_t = \ln S_t - \ln S_{t-1}$, $\varepsilon_t = W_t - W_{t-1}$, and we know $\Delta t = 1$, equation (2.1.5) can be rewritten as

$$y_t = \mu - \frac{1}{2}\sigma^2 + \sigma\varepsilon_t. \tag{2.1.6}$$

μ and σ are parameters that we want to estimate.

According to the properties of standard Brownian motion process, for each n≥1, and any sequence of time $0 \le t_0 \le t_1 \le ... \le t_i \le ... \le t_n$, the random variables $W_{t_i} - W_{t_{i-1}}$ are independent, $\varepsilon_t = W_t - W_{t-1}$ has the standard normal distribution with mean zero and variance 1. Thus the conditions of least square estimations are satisfied. $\mu - \frac{1}{2}\sigma^2$ can be estimated as the average value of y_t, which is $\ln S_t - \ln S_{t-1}$, σ can be estimated as the standard deviation of $\ln S_t - \ln S_{t-1}$ [35]. We will use this least square estimation in our work for both linear and nonlinear models.

Remark 2.1.1: An alternative way to estimate the parameters is as described below. Since $\varepsilon_t = W_t - W_{t-1}$ is standard normal distribution, S_t is log-normally distributed with mean $\mu - \frac{1}{2}\sigma^2$ and variance σ^2. Then we can estimate the drift μ and volatility σ parameters by using the historical price data. Taking the logarithm of S_t, we can estimate $\mu - \frac{1}{2}\sigma^2$ as the average values of $\ln S_t - \ln S_{t-1}$, and can estimate the volatility σ by taking the standard deviation of $\ln S_t - \ln S_{t-1}$ [35]. This is exactly the same as what we have estimated by using least square method.

2.2 GBM Model on Overall Data

In this section, by using fortune 500 companies price dataset, we estimate the parameters $\mu - \frac{1}{2}\sigma^2$ and σ^2 in (2.1.6). This is achieved in the framework of the overall price data of stock X. Suppose we let S_t be the daily adjusted closing values of stock X that we collect form the fortune 500 companies that we mentioned early. A plot of the actual data set is drawn in Figure 2.2.1.

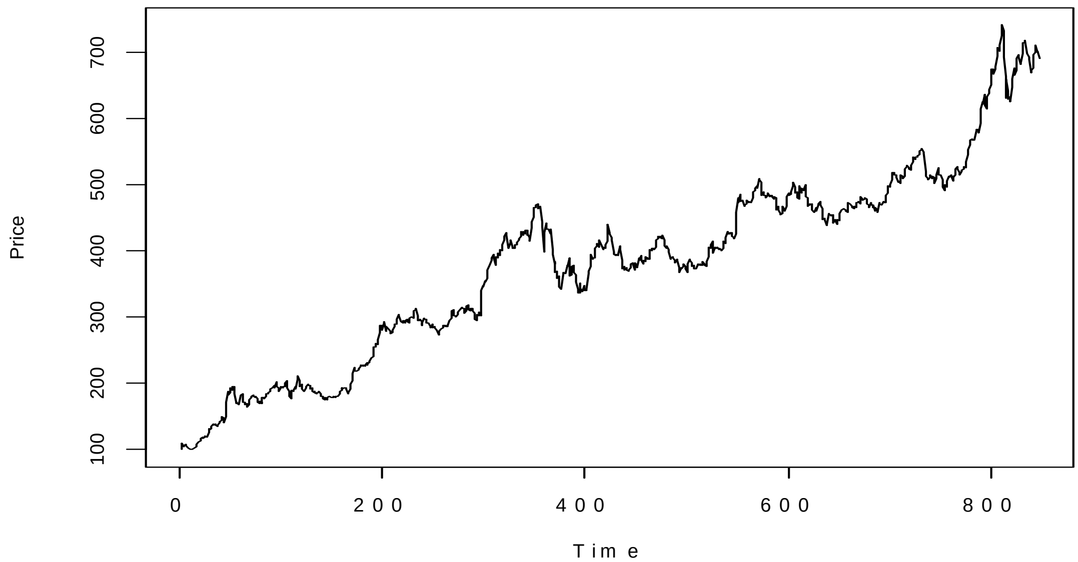


Figure 2.2.1 Daily Adjusted Closing Price Process for Stock X

We pick one stock X over long period (3 ½ years) of time to build its GBW model. The Figure 2.2.1 shows its daily adjusted closing price process from 8/19/2004 to 12/31/2007. Using the least square method described in Section 1.4, the estimates of drift and volatility are as follows

$$\hat{\mu} = 0.002501028, \text{ and } \hat{\sigma} = 0.02107507.$$

Hence the GBM process for the stock X price is the solution of the following linear $It\hat{o} - Doob$ stochastic differential equation:

$$d\hat{S}_t = 0.002501028\hat{S}_t dt + 0.02107507\hat{S}_t dW_t. \quad (2.2.1)$$

The stock price process is

$$\hat{S}_t = S_0 e^{(0.002501028 - \frac{0.02107507^2}{2})t + 0.02107507 W_t} \quad (2.2.2)$$

In equation (2.2.2), S_0 is the initial stock price of the price of the stock process. W_t is Brownian motion, that is, it is a random process. Under direct simulation of the stock price process as we generate the Brownian motion, we get the different values. We first use the Monte Carlo method [34] to predict the stock price process and then calculate the average of the process. This is a very general approach that is used in many areas, such as physics, chemistry, finance etc. Here, we simulate the stock price process 2000 times. Using Monte Carlo method a plot of the stock price process of (2.2.1) is given in Figure 2.2.2. The red curve in Figure 2.2.2 represents the result using Monte Carlo simulation method. The process in blue curve is one resulted from simulation which varies from time to time.

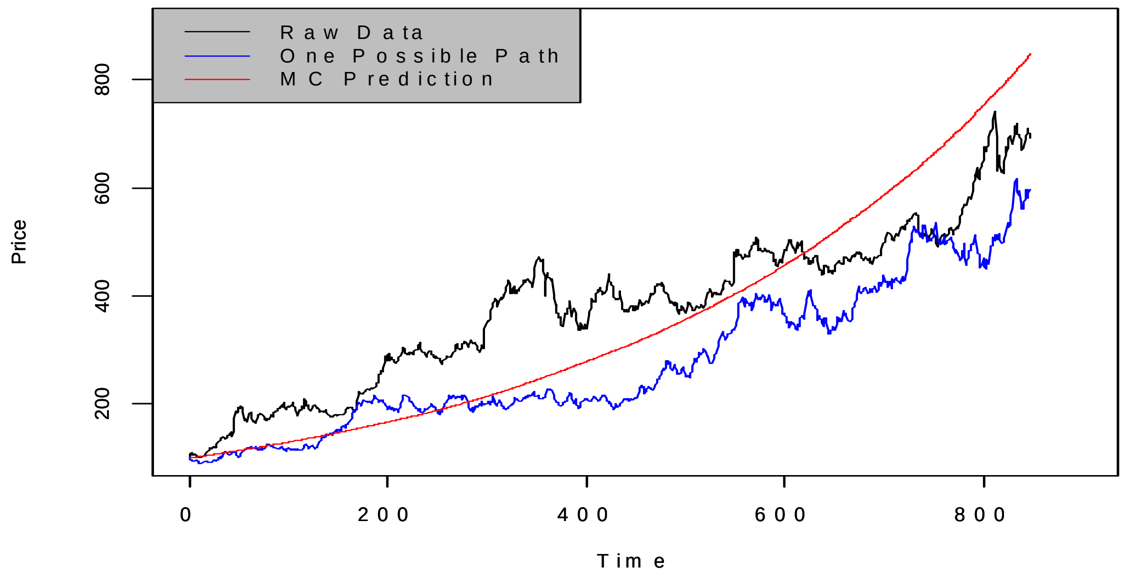


Figure 2.2.2 Prediction and One Possible Path of Stock X's Price Process Using Model (2.2.1)

After we estimate the parameters, $\hat{S}_t$ is estimated by $\ln \hat{S}_t = \ln S_{t-1} + \hat{\mu} - \frac{1}{2}\hat{\sigma}^2$. The basic statistics reflecting the accuracy of model in Equation (2.2.1) are mean of the residuals $\overline{r}$, variance S_r^2 of the residuals, and standard deviation S_r of the residuals, where residual errors are defined as $r_t = S_t - \hat{S}_t$. Table 2.2.1 shows these basic statistics.

Table 2.2.1 Basic Statistics for Model in Equation (2.2.1)

$\overline{r}$	S_r^2	S_r	No. of Parameters
28.29653	8752.84	93.55661	2

To test the homogeneous errors in equation (2.1.6), actually, we assume that the error term is normally distributed. We use Q-Q plot to test it. In statistics, a Q-Q plot ("Q" stands for quantile) is a graphical method for diagnosing differences between the probability distribution of a population from which a random sample has been drawn and a comparison distribution. An example of this kind of difference that can be tested is non-normality of the population distribution. The normal distribution is represented by a straight line. The Q-Q plot is in Figure 2.2.3

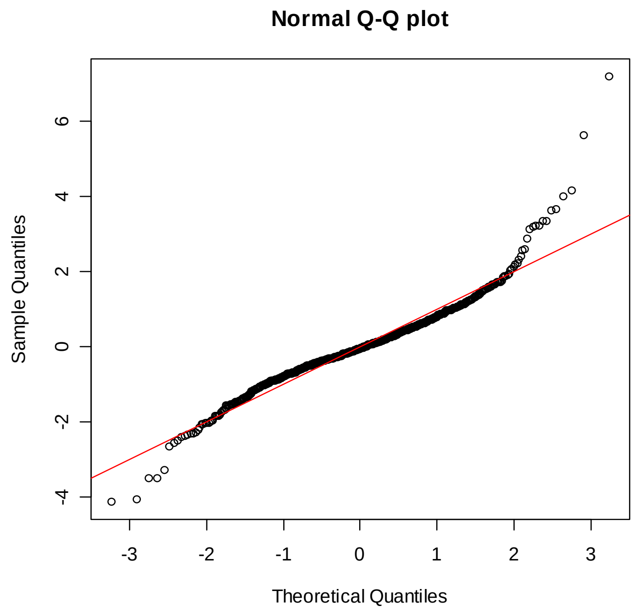


Figure 2.2.3 Q-Q Plot for Model in Equation (2.2.1)

Remark 2.2.1: From the table 2.2.1, the average residual is 28.29653. This means that zero mean condition obviously is not satisfied. Also, the variance is too large. From the Figure 2.2.2, we can see that the prediction line (in red) cannot describe the stock process. Furthermore, we can see a reverse “S” shape in the Q-Q plot. All these observations suggest that we need a more work to get the better model.

2.3 GBM Models under Data Partitioning Schemes without Jumps

The usage of the overall data in estimating the parameters in (2.1.6) suggests to modify the usage of data. The estimated parameters in section 2.2 are not realistic. This has been evidenced by the Q-Q plot test for the homogeneous of error in (2.1.6) over the entire period of the data. As a result of this, it is natural to partition the data, and repeat the procedure outlined in Section 2.2.

In this section, we will use the same stock price process under different data partitioning to develop the GBM models. The data is reorganized half-yearly, quarterly, and monthly to build GBM models on different segments of periods of the overall period of dataset.

If we revise the dataset more closely, we will find some pattern. Figures 2.3.1-2.3.4 show that the daily difference of stock X in 4 quarters from August 2004 to end of year 2007. The daily differences in quarter 2 (Q2) and quarter 3 (Q3) are in the range [-20, 20]. The daily differences in quarter 1 (Q1) and quarter 4 (Q4) are much bigger than those in quarter 2 (Q2) and quarter 3 (Q3). Also in a particular quarter, most of the daily differences follow the similar pattern.

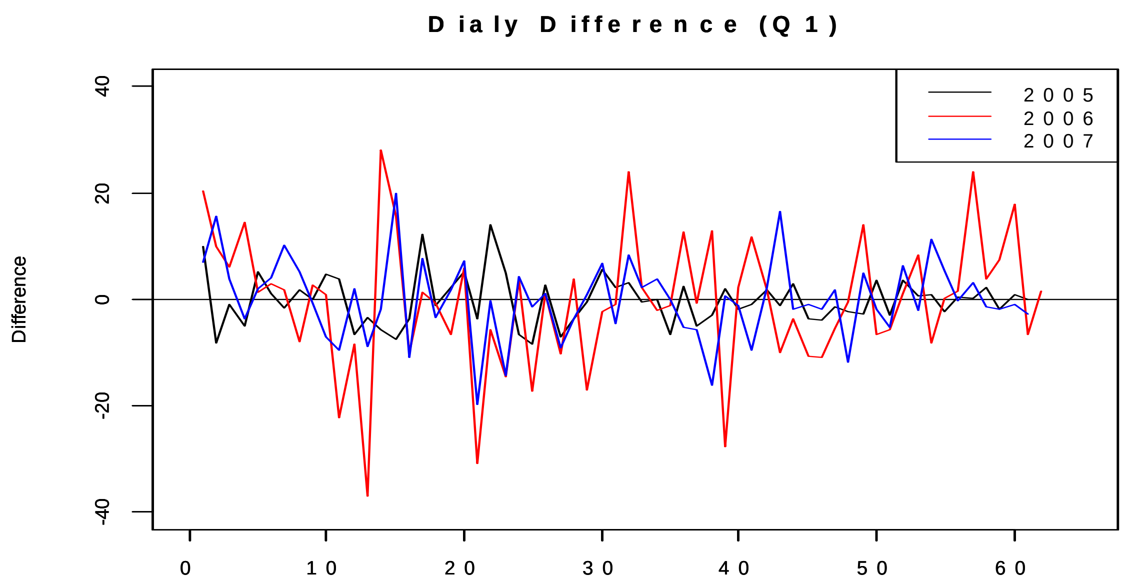


Figure 2.3.1 Daily Difference of Stock X in Q1

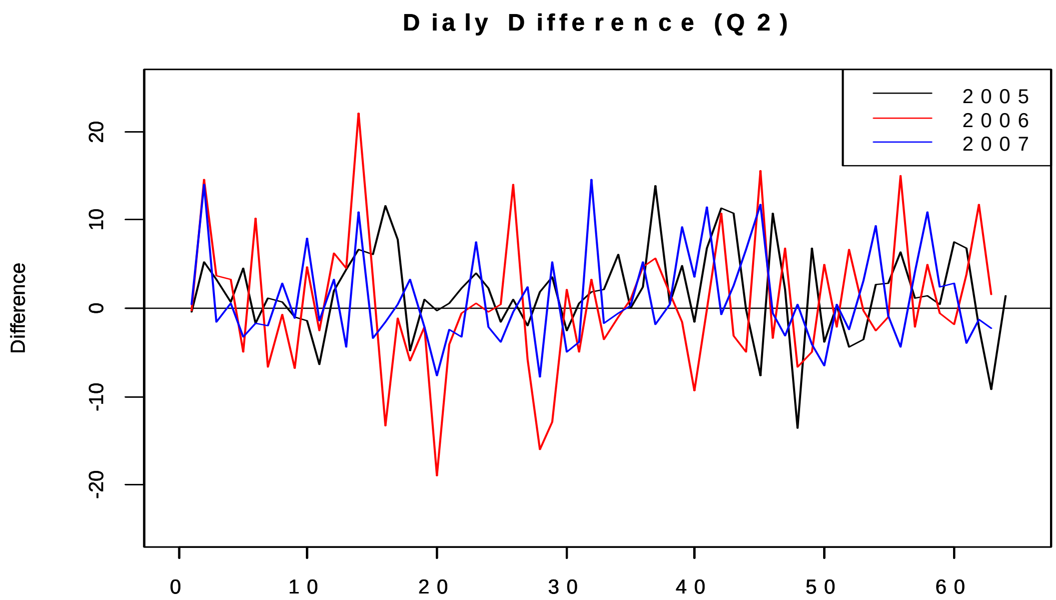


Figure 2.3.2 Daily Difference of Stock X in Q2

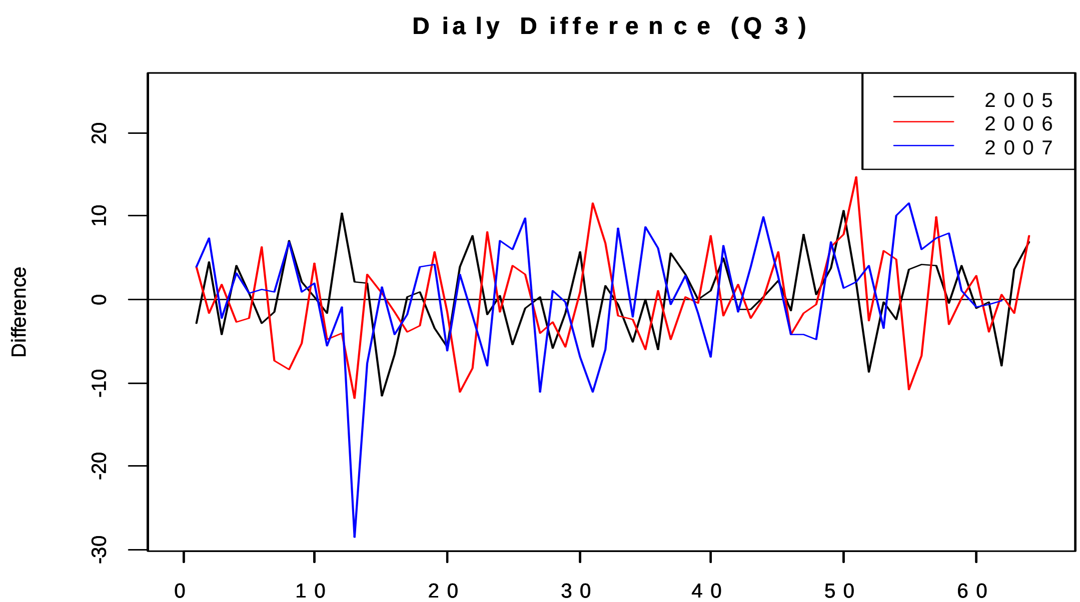


Figure 2.3.3 Daily Difference of Stock X in Q3

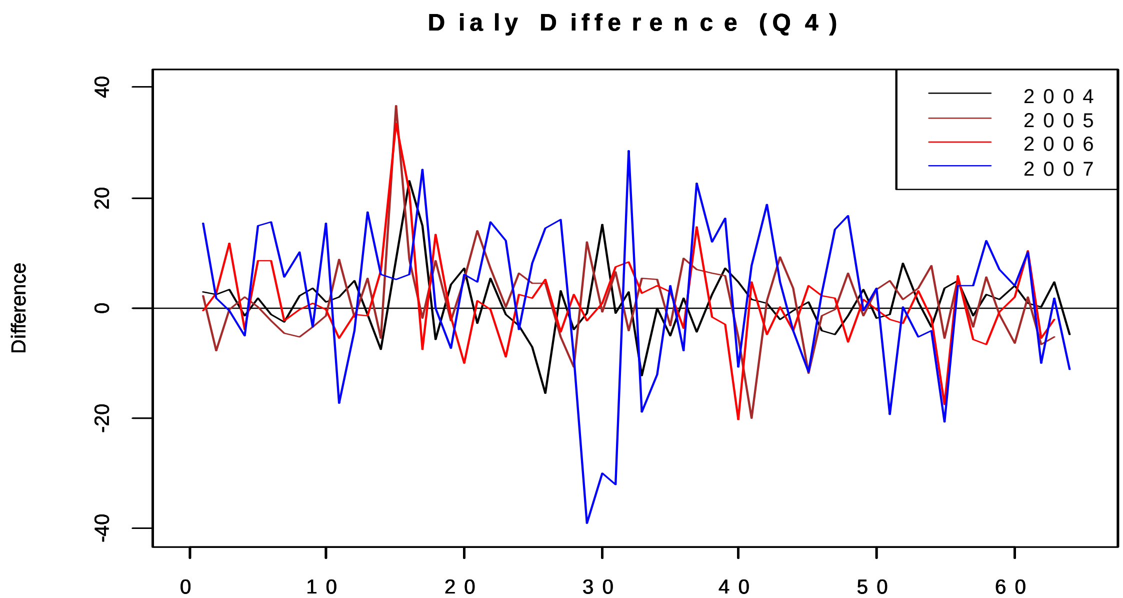


Figure 2.3.4 Daily Difference of Stock X in Q4

Table 2.3.1 also shows the standard deviations in Q1 and Q4 are larger than the standard deviations in Q2 and Q3. This suggests us that we might reorganize the sample dataset into two sub data sets – Q2 and Q3 as subset 1 and Q1 and Q4 as subset 2. Furthermore, we also divide the sample dataset into 4 sub datasets -- 4 quarters. For each subset, we use the same method, which is described in Section 2.2 to develop its GBM model, separately.

Table 2.3.1 Mean and Standard Deviation of Daily Differences of 4 Quarters of Stock X

	2007		2006		2005		2004	
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
Q1	-0.038	7.468	-0.401	12.496	-0.201	4.489	NA	NA
Q2	1.024	5.263	0.466	7.511	1.776	4.963	NA	NA
Q3	0.595	6.573	-0.152	5.569	0.349	4.483	NA	NA
Q4	1.941	13.637	0.930	7.947	1.562	7.659	0.987	5.843

Data Partition Process 2.3.1: From the description of construction of Figures 2.3.1-2.3.4 and Table 2.3.1, we partition the overall data set into two sub datasets.

$[0,t_1),[t_1,t_2),[t_2,t_3),\ [t_3,t_4),[t_4,t_5)...$ represent the quarter year time intervals starting Q1, Q2, Q3 and Q4 etc. The sub dataset 1 contains observations in the Q1 or the Q4. The sub dataset 2 contains observations in the Q2 or the Q3.

GBM Model without Jumps 2.3.1(Half Yearly GBM Model without Jumps): The GBM processes without jumps using Data Partition Process 2.3.1 are the solutions of the following linear $It\hat{o}-Doob$ type stochastic differential equation:

$$\begin{cases} dS_t^{Q_{14}} = \mu^{Q_{14}} S_t^{Q_{14}} dt + \sigma^{Q_{14}} S_t^{Q_{14}} dW_t, \text{ if t is in Q1 or Q4, } S_0 = S_0, \\ dS_t^{Q_{23}} = \mu^{Q_{23}} S_t^{Q_{23}} dt + \sigma^{Q_{23}} S_t^{Q_{23}} dW_t, \text{ if t is in Q2 or Q3.} \end{cases} \tag{2.3.1}$$

$\mu^{Q_{14}}$ and $\mu^{Q_{23}}$ are drifts, and $\sigma^{Q_{14}}$ and $\sigma^{Q_{23}}$ are volatility rates for two sub datasets, respectively.

By following definition [16, 26, 27], the price process is the solution of (2.3.1), it take the form

$$S_t = \begin{cases} S_t^{Q_{14}} = S_0 e^{(\mu^{Q14} - \frac{1}{2}(\sigma^{Q14})^2)t + \sigma^{Q14} W_t} & S_0 = S_0, & 0 \le t < t_1 \\ S_t^{Q_{23}} = S_1 e^{(\mu^{Q23} - \frac{1}{2}(\sigma^{Q23})^2)t + \sigma^{Q23} W_t} & S_1 = \lim\limits_{t \to t_1} S_t^{Q_{14}}, & t_1 \le t < t_3 \\ S_t^{Q_{14}} = S_2 e^{(\mu^{Q14} - \frac{1}{2}(\sigma^{Q14})^2)t + \sigma^{Q14} W_t} & S_2 = \lim\limits_{t \to t_3} S_t^{Q_{23}}, & t_3 \le t < t_5 \\ S_t^{Q_{23}} = S_3 e^{(\mu^{Q23} - \frac{1}{2}(\sigma^{Q23})^2)t + \sigma^{Q23} W_t} & S_3 = \lim\limits_{t \to t_5} S_t^{Q_{23}}, & t_5 \le t < t_7 \\ \qquad ... & \qquad ... & \quad ... \end{cases} \tag{2.3.2}$$

S_0 is the initial value of the price process. There are 4 parameter $\mu^{Q_{14}}$, $\sigma^{Q_{14}}$ and $\mu^{Q_{23}}$, $\sigma^{Q_{23}}$ need to be estimated.

For stock X, the estimated results are as following

$$\begin{cases} d\hat{S}_t^{Q_{14}} = 0.002284141\hat{S}_t^{Q_{14}}dt + 0.02447219\hat{S}_t^{Q_{14}}dW_t \text{, if } t \text{ is in Q1 or Q4.} \\ d\hat{S}_t^{Q_{23}} = 0.002733729\hat{S}_t^{Q_{23}}dt + 0.01671308\hat{S}_t^{Q_{23}}dW_t \text{, if } t \text{ is in Q2 or Q3.} \end{cases} \tag{2.3.3}$$

And the estimated stock X's price process is:

$$\begin{cases} \hat{S}_t^{Q_{14}} = S_0 e^{(0.002284141 - \frac{0.02447219^2}{2})t + 0.02447219W_t} \text{, if } t \in [0, t_1); \\ \hat{S}_t^{Q_{23}} = (\lim_{t \to t_1} \hat{S}_t^{Q_{14}}) e^{(0.002733729 - \frac{0.01671308^2}{2})t + 0.01671308W_t} \text{, if } t \in [t_1, t_3); \\ \hat{S}_t^{Q_{14}} = (\lim_{t \to t_3} S_t^{Q_{23}}) e^{(0.002284141 - \frac{0.02447219^2}{2})t + 0.02447219W_t} \text{, if } t \in [t_3, t_5); \\ \quad \ldots \ldots \end{cases} \tag{2.3.4}$$

The prediction result of stock X's price process of in (2.3.3) is provided in Figure 2.3.5. We see that the blue curve and red curve are very close. Because the two drifts in Equation (2.3.3) are very close. The most different part in Equation (2.3.3) is the volatilities.

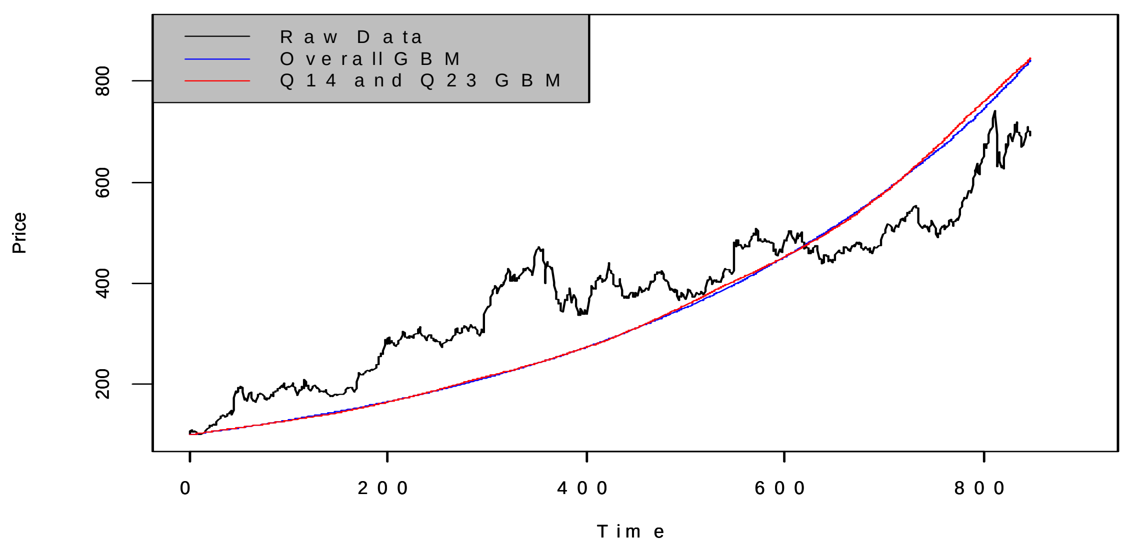


Figure 2.3.5 Comparison on Model (2.2.1) with Model (2.3.3) of Stock X

Table 2.3.2 Basic Statistics for Model in Equation (2.3.2)

Model	$\bar{r}$	S_r^2	S_r	No. of Intervals	No. of Parameters
GBM with Overall Data	28.29653	8752.84	93.55661	1	2
Q14 and Q23 GBM without jumps	29.67727	8836.837	94.00445	8	4

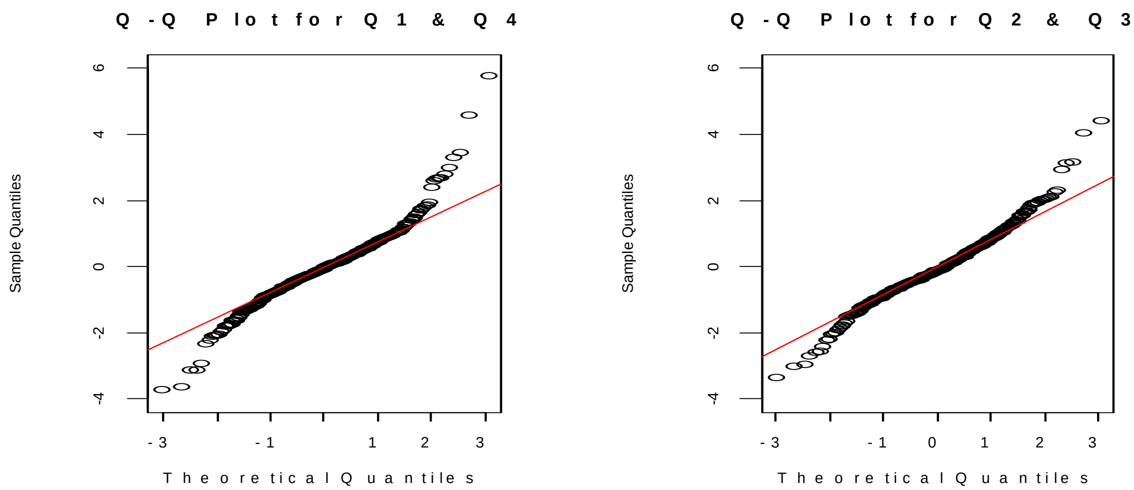


Figure 2.3.6 Q-Q Plot for Model in Equation (2.3.3)

Remark 2.3.1: From Figure 2.3.6, we still can see there are reverse "S" shapes in the two Q-Q plots for both Q1 and Q4, and Q2 and Q3. Table 2.3.2 provides the basic statistics. And from Figure 2.3.5, we don't see an improvement from GBM model on overall data. All these suggest that we need more work to get the better model.

In the following, we try to reorganize the dataset into 4 sub datasets – quarter 1 (Q1), quarter 2 (Q2), quarter 3 (Q3), and quarter 4 (Q4).

Data Partition Process 2.3.2: The time intervals $[0,t_1),[t_1,t_2),[t_2,t_3)$, $[t_3,t_4),[t_4,t_5)$... as defined in data partition process 2.3.1. The sub datasets 1, 2, 3, and 4 contain observations in Q1, Q2 , Q3, and Q4, respectively.

The GBM Model without Jumps 2.3.2 (Quarterly GBM Model without Jumps): The GBM processes without jumps using Data Partition Process 2.3.2 are the solutions of the following linear $It\hat{o}-Doob$ type stochastic differential equation:

$$\begin{cases} dS_t^{Q_1} = \mu^{Q_1} S_t^{Q_1} dt + \sigma^{Q_1} S_t^{Q_1} dW_t \text{, if t is in Q1.} \\ dS_t^{Q_2} = \mu^{Q_2} S_t^{Q_2} dt + \sigma^{Q_2} S_t^{Q_2} dW_t \text{, if t is in Q2.} \\ dS_t^{Q_3} = \mu^{Q_3} S_t^{Q_3} dt + \sigma^{Q_3} S_t^{Q_3} dW_t \text{, if t is in Q3.} \\ dS_t^{Q_4} = \mu^{Q_4} S_t^{Q_4} dt + \sigma^{Q_4} S_t^{Q_4} dW_t \text{, if t is in Q4.} \end{cases} \tag{2.3.5}$$

μ^{Q_1}, μ^{Q_2}, μ^{Q_3} and μ^{Q_4} are drifts, and σ^{Q_1}, σ^{Q_2}, σ^{Q_3} σ^{Q_4} are volatilities for four quarters respectively. By following definition [16, 26, 27], the price process is the solution of Equation (2.3.5), and takes the form

$$S_t = \begin{cases} S_t^{Q_3} = S_0 e^{(\mu^{Q3} - \frac{1}{2}(\sigma^{Q3})^2)t + \sigma^{Q3} W_t} & S_0 = S_0, & 0 \le t < t_1 \\ S_t^{Q_4} = S_1 e^{(\mu^{Q4} - \frac{1}{2}(\sigma^{Q4})^2)t + \sigma^{Q4} W_t} & S_1 = \lim_{t \to t_1} S_t^{Q_3}, & t_1 \le t < t_2 \\ S_t^{Q_1} = S_2 e^{(\mu^{Q1} - \frac{1}{2}(\sigma^{Q1})^2)t + \sigma^{Q1} W_t} & S_2 = \lim_{t \to t_2} S_t^{Q_4}, & t_2 \le t < t_3 \\ S_t^{Q_2} = S_3 e^{(\mu^{Q2} - \frac{1}{2}(\sigma^{Q2})^2)t + \sigma^{Q2} W_t} & S_3 = \lim_{t \to t_3} S_t^{Q_1}, & t_3 \le t < t_4 \\ S_t^{Q_3} = S_4 e^{(\mu^{Q3} - \frac{1}{2}(\sigma^{Q3})^2)t + \sigma^{Q3} W_t} & S_4 = \lim_{t \to t_4} S_t^{Q_2} & t_4 \le t < t_5 \\ \quad \ldots & \quad \ldots & \quad \ldots \end{cases} \tag{2.3.6}$$

There are 8 parameters μ^{Q_1}, μ^{Q_2}, μ^{Q_3}, μ^{Q_4}, σ^{Q_1}, σ^{Q_2}, σ^{Q_3} and σ^{Q_4}. These parameters need to be estimated. For stock X, the estimated results are as following:

$$
\begin{cases}
d\hat{S}_t^{Q_1} = -0.0004185944\hat{S}_t^{Q_1}dt + 0.02459217\hat{S}_t^{Q_1}dW_t\text{, if t is in Q1.}\\
d\hat{S}_t^{Q_2} = 0.003792002\hat{S}_t^{Q_2}dt + 0.01712985\hat{S}_t^{Q_2}dW_t\text{, if t is in Q2.}\\
d\hat{S}_t^{Q_3} = 0.001815337\hat{S}_t^{Q_3}dt + 0.01632730\hat{S}_t^{Q_3}dW_t\text{, if t is in Q3.}\\
d\hat{S}_t^{Q_4} = 0.004238625\hat{S}_t^{Q_4}dt + 0.02424493\hat{S}_t^{Q_4}dW_t\text{, if t is in Q4.}
\end{cases} \tag{2.3.7}
$$

Following the earlier arguments, Figure 2.3.7 exhibits the result of prediction of stock X's price process of (2.3.7). We note that the red curve (quarterly GBM model) is not similar to the blue curve (Overall GBM model) as well as orange curve (Q14 and Q23 GBM model). This is due to obvious reasons

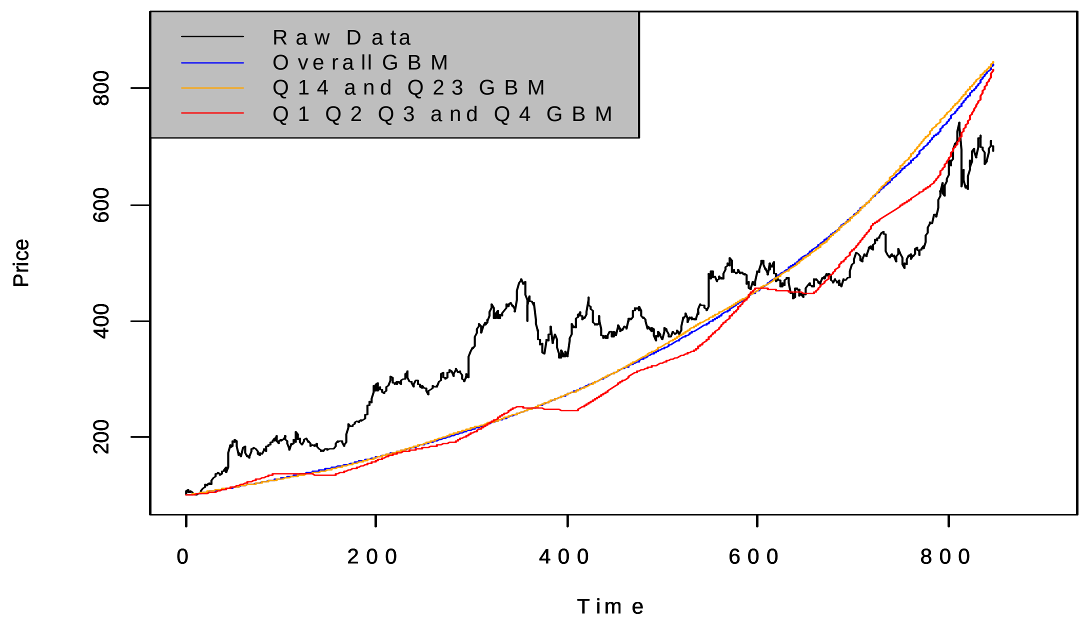


Figure 2.3.7 Comparison on Model (2.2.1), (2.3.3) with Model (2.3.7) of Stock X

Table 2.3.3 provides the statistics for 3 models, namely, overall GBM Model, Q14 and Q23 GBM model and Quarterly GBM model.

Table 2.3.3 Basic Statistics for Model in Equation (2.2.1) (2.3.3) and (2.3.7)

Model	$\overline{r}$	S_r^2	S_r	No. of Intervals	No. of Parameters
GBM with Overall Data	28.29653	8752.84	93.55661	1	2
Q14 and Q23 GBM without jumps	29.67727	8836.837	94.00445	8	4
Quarterly GBM without jumps	53.49948	5570.643	74.63674	14	8

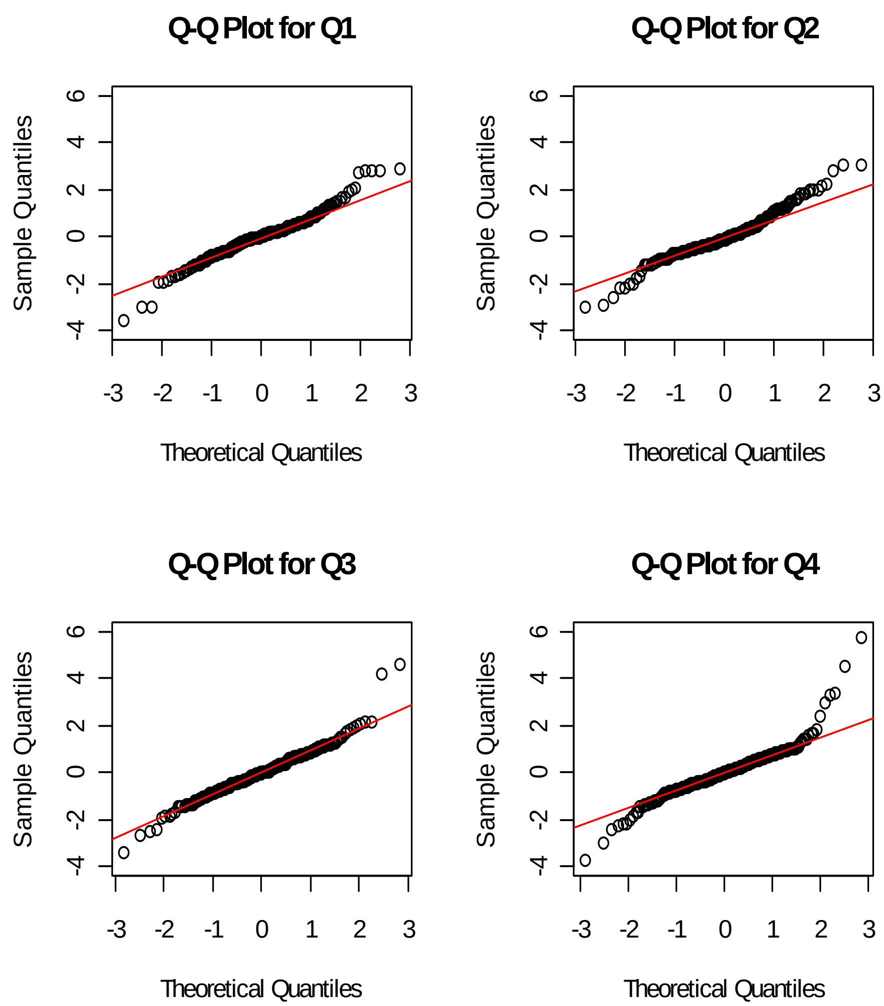


Figure 2.3.8 Q-Q Plot of Model (2.3.7)

Remark 2.3.2: (a) From Figure 2.3.7, we notice the large deviation between the predicted result and the observed data set. (b) From Table 2.3.3, we note that the quarterly partition data set approach gives the least variance with the largest mean of the residuals. (c) Figure 2.3.8 is the Q-Q plots for model (2.3.7). We observe that there is still a reverse "S" shape in the Q-Q plot for Q2 and Q4. In Q1 and Q3, most points fall in the normal distributions and there are a few outliers. (d) Again, after careful review of the Figures 2.3.1-2.3.4, we found some patterns. The daily differences in quarter 1 (Q1) and quarter 4 (Q4) are much larger than those differences with regards to in quarter 2 (Q2) and quarter 3 (Q3). The daily differences do not follow the same pattern in the same quarter in different year, that is, the dynamic of stock price in the same quarter with different year follows different pattern. As a result of this, we develop two kinds of data partitioning schemes, we don't put the observations in different years together.

Data Partition Process 2.3.3: Let $[0,t_1),[t_1,t_2),[t_2,t_3),\ [t_3,t_4),[t_4,t_5)...\ [t_{m-1},t_m)$ be a monthly sub intervals for m month data set. The sub dataset 1 contains observations in the 1st month, the sub dataset 2 contains observations in the 2nd month, the sub dataset 3 contains observations in the 3rd month, …, the sub dataset m contains observations in the m-th month.

GBM Model without Jumps 2.3.3 (Monthly GBM Model without Jumps): Let $[0,t_1),[t_1,t_2),[t_2,t_3),\ [t_3,t_4),[t_4,t_5)...\ [t_{m-1},t_m)$ be the m monthly sub intervals. The GBM process without jumps is the solution of the following linear $It\hat{o}-Doob$ type stochastic differential equation:

$$dS_t^{M_i} = \mu^{M_i} S_t^{M_i} dt + \sigma^{M_i} S_t^{M_i} dW_t,\ S_0 = S_0, \text{ if } t_{i-1} \le t < t_i,\ i = 1,...,m. \tag{2.3.8}$$

Here, μ^{M_i} and σ^{M_i}, $i = 1,...,m$, are monthly drift and volatility coefficients, respectively. Again, by following definition [16, 26, 27], the price process is the solution of (2.4.8), and takes the following form:

$$S_t = \begin{cases} S_t^{M_1} = S_0 e^{(\mu^{M_1} - \frac{1}{2}(\sigma^{M_1})^2)t + \sigma^{M_1} W_t} & S_0 = S_0, \quad t_0 \le t < t_1 \\ S_t^{M_2} = S_1 e^{(\mu^{M_2} - \frac{1}{2}(\sigma^{M_2})^2)t + \sigma^{M_2} W_t} & S_1 = \lim_{t \to t_1} S_t^{M_1}, \quad t_1 \le t < t_2 \\ \quad \cdots & \quad \cdots \qquad \cdots \\ S_t^{M_m} = S_{m-1} e^{(\mu^{M_m} - \frac{1}{2}(\sigma^{M_m})^2)t + \sigma^{M_m} W_t} & S_{m-1} = \lim_{t \to t_{m-1}} S_t^{M_{m-1}}, \quad t_{m-1} \le t < t_m \end{cases} \tag{2.3.9}$$

There are $2 \times m$ parameters μ^{M_i} and $\sigma^{M_i}, i = 1, \ldots, m$, need to be estimated. m is the number of month of stock price process. The methods of estimation parameters are the same. For stock X, the estimated results are as following.

$$\begin{cases} d\hat{S}_t^{M_1} = 0.00321650 \ \hat{S}_t^{M_1} dt + 0.03775822 \ \hat{S}_t^{M_1} dW_t, & if \quad t \in [0, t_1); \\ d\hat{S}_t^{M_2} = 0.01146937 \ \hat{S}_t^{M_2} dt + 0.02181682 \ \hat{S}_t^{M_2} dW_t, & if \quad t \in [t_1, t_2); \\ \qquad \cdots & \cdots \qquad \cdots \\ d\hat{S}_t^{M_{41}} = 0.00000\,259 \, \hat{S}_t^{M_{41}} dt + 0.01499199 \ \hat{S}_t^{M_{41}} dW_t, & if \quad t \in [t_{40}, t_{41}). \end{cases} \tag{2.3.10}$$

All estimated parameters $\hat{\mu}^{M_i}$ and $\hat{\sigma}^{M_i}, i = 1, \ldots, m$, are given in Appendix A1. Figure 2.3.9 is the prediction of stock X's prices process. Table 2.3.4 provides the basic statistics for estimated model corresponding to the original Equation (2.3.8).

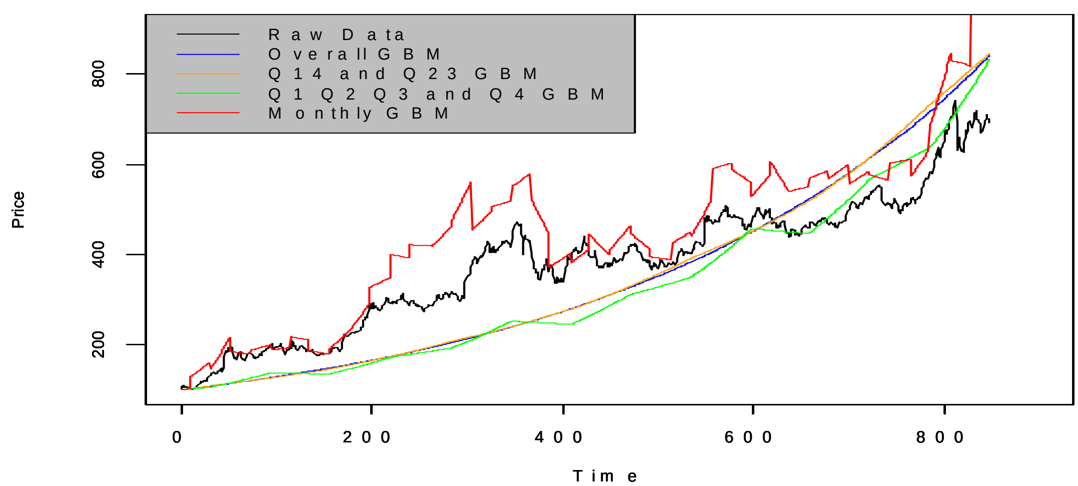


Figure 2.3.9 Comparison on Model (2.2.1), (2.3.3), (2.3.7) with Model (2.3.10) of Stock X

Table 2.3.4 Basic Statistics for Model in Equation (2.2.1) (2.3.3) (2.3.7) and (2.3.10)

Model	$\overline{r}$	S_r^2	S_r	No. of Intervals	No. of Parameters
GBM with Overall Data	28.29653	8752.84	93.55661	1	2
Q14 and Q23 GBM without jumps	29.67727	8836.837	94.00445	8	4
Quarterly GBM without jumps	53.49948	5570.643	74.63674	14	8
Monthly GBM without jumps	-80.10483	11754.25	108.4170	41	82

Remark 2.3.3: (a) From Figure 2.3.9 we can see that the monthly GBM model in red really catches the dynamic of the stock price process. (b) The stock price process shows it is always over predicted. The basic statistics in Table 2.3.4 shows that the variance and standard deviation of the residual are very large in this monthly GBM model.

Data Partition Processes 2.3.1-2.3.3 have a common character, that is, the length of time interval in each model is exactly the same. For examples, the length of time interval in Data Partition Process 2.3.1, 2.3.2, and 2.3.3 are two quarters, one quarter, and one month respectively. If there is a big shock in the stock price in one of the intervals, this kind of equal length model cannot incorporate the effects of the big shock. To avoid this problem, we provide a modified data partition process, this allows us to have unequal length of intervals.

Data Partition Process 2.3.4: Let $[0,t_1),[t_1,t_2),[t_2,t_3),\ [t_3,t_4),[t_4,t_5)...\ [t_{n-1},t_n)$ be the data set time decomposition into n time intervals. We suppose all the big shocks come at times $t_1, t_2, ...t_{n-1}$. The sub dataset 1 contains observations in the 1st time interval, that is in $[0,t_1)$, the sub dataset 2 contains observations in the 2nd interval, that is $[t_1,t_2)$, the sub dataset 3 contains observations in the 3rd interval, that is $[t_2,t_3)$, …, the sub dataset n contains observations in the nth interval, that is $[t_{n-1},t_n)$.

GBM Models without Jumps 2.3.4 (Unequal Interval GBM Model without Jumps): By utilizing the above described sub interval decomposition, GBM processe without jumps is the solution of the following linear $It\hat{o}-Doob$ type stochastic differential equation:

$$dS_t^{I_i} = \mu^{I_i} S_t^{I_i} dt + \sigma^{I_i} S_t^{I_i} dW_t,\ S_0 = S_0 \text{ if } t_{i-1} \le t < t_i,\ i = 1,...,n. \tag{2.3.11}$$

μ^{I_i} and σ^{I_i} , $i = 1,...,n$, are the i-*th* drift and i-*th* volatility coefficients, respectively. By following the definition [16, 26, 27], The solution to Equation (2.3.11) is, and takes the following form:

$$S_t = \begin{cases} S_t^{I_1} = S_0 e^{(\mu^{I_1} - \frac{1}{2}(\sigma^{I_1})^2)t + \sigma^{I_1} W_t} & S_0 = S_0, & t_0 \le t < t_1 \\ S_t^{I_2} = S_1 e^{(\mu^{I_2} - \frac{1}{2}(\sigma^{I_2})^2)t + \sigma^{I_2} W_t} & S_1 = \lim_{t \to t_1} S_t^{I_1}, & t_1 \le t < t_2 \\ \quad\quad ... & \quad ... & \quad ... \\ S_t^{I_n} = S_{n-1} e^{(\mu^{I_n} - \frac{1}{2}(\sigma^{I_n})^2)t + \sigma^{I_n} W_t} & S_{n-1} = \lim_{t \to t_{n-1}} S_t^{I_{n-1}}, & t_{n-1} \le t < t_n \end{cases} \tag{2.3.12}$$

There are $2 \times n$ parameters μ^{I_i} and σ^{I_i} , $i = 1,...,n$, and these parameters need to be estimated.

Now the key issue is how to define unequal length of time interval. The basic idea about defining the time intervals is that we want to identify the dates, having the large daily relative difference. So we need to define the threshold first, that is, we need to define the threshold of daily relative difference of stock price. There are two issues that we want to consider. The first issue is that the threshold cannot be either too large or too small. This is because of the fact that if the threshold is too large, then we may have too few intervals, and it cannot incorporate the dynamic of stock price process. Therefore, we cannot have a good model. If the threshold is too small, then we may have too many intervals, that is, for some time intervals, there are few observations so that we cannot reasonably develop a model. The second issue is, after defining the threshold, the lengths of some time intervals are still too long. In this case, we break these time intervals into months, since monthly GBM model shows very good dynamic character. Once we define unequal length of time intervals, we apply the same procedure to estimate parameters.

The method of estimation parameters is as described in chapter 1. For stock X, the estimated results are as

$$\begin{cases} d\hat{S}_t^{I_1} = 0.006501\ \hat{S}_t^{I_1} dt + 0.024052\ \hat{S}_t^{I_1} dW_t, & if \quad t \in [0, t_1); \\ d\hat{S}_t^{I_2} = 0.013391\ \hat{S}_t^{I_2} dt + 0.028731\ \hat{S}_t^{I_2} dW_t, & if \quad t \in [t_1, t_2); \\ \qquad \cdots & \cdots \qquad \cdots \\ d\hat{S}_t^{I_{38}} = 0.00000\,259\ \hat{S}_t^{I_{38}} dt + 0.014992\ \hat{S}_t^{I_{38}} dW_t, & if \quad t \in [t_{37}, t_{38}). \end{cases} \tag{2.3.13}$$

The estimated parameters $\hat{\mu}^{I_i}$ and $\hat{\sigma}^{I_i}, i = 1, \ldots, n$, are given in Appendix A2. Figure 2.3.10 is the predicted stock X's price process. Table 2.3.5 provides the basic statistics for estimated model corresponding to (2.3.13).

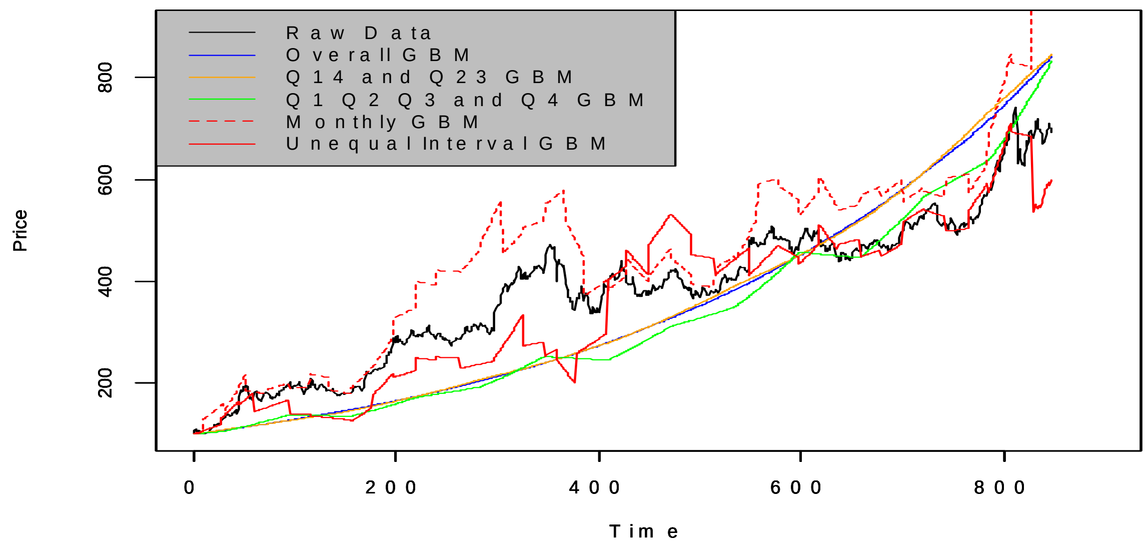


Figure 2.3.10 Comparison on Model (2.3.1), (2.3.3), (2.3.7), (2.3.10) with. Model (2.3.13) of Stock X

Table 2.3.5 Basic Statistics for Model in Equation (2.2.1), (2.3.3), (2.3.7), (2.3.10) and (2.3.13)

Model	$\overline{r}$	S_r^2	S_r	No. of Intervals	No. of Parameters
GBM with Overall Data	28.29653	8752.84	93.55661	1	2

Q14 and Q23 GBM without jumps	29.67727	8836.837	94.00445	8	4
Quarterly GBM without jumps	53.49948	5570.643	74.63674	14	8
Monthly GBM without jumps	-80.10483	11754.25	108.4170	41	82
Unequal Interval GBM without jumps	24.91557	3992.349	63.18504	39	78

Remark 2.3.4: Figure 2.3.10 shows that the Monthly GBM model (dashed red curve) and Unequal interval GBM model (solid red curve) are approximations of the true stock price movements in comparison to other linear models. However, Table 2.3.5 shows all these 5 models have very large residuals. This is largely due to the accumulated errors in models without jumps. When we make a prediction, we only use the stock price at time 0 as the initial value to predict a long time behavior of the stock price. In section 2.4, we will add jumps to this model to reduce the cumulative error.

2.4 GBM Models under Data Partitioning Schemes with Jumps

All models in Section 2.3 are without jumps, that is, we take the left limit of the right endpoint of previous time interval as the initial value of the next time interval. This simplistic approach carries the previous time interval error to next time interval. The cumulated error might be very big. Here, we modify the models of Section 2.3 by adding jumps into the models. The data partition processes and other parameters such as drifts and volatilities remain the same. We will not repeat in this section.

GBM Model with Jumps 2.4.1 (Half Yearly GBM Model with Jumps): Let $[0,t_1),[t_1,t_2),[t_2,t_3),\ [t_3,t_4),[t_4,t_5)...$ be the time intervals as defined in Data Partition Process 2.3.1. By following the argument, the GBM solution process with jumps of (2.3.1) has the following form:

$$
S_t = \begin{cases} S_t^{Q_{14}} = S_0 e^{(\mu^{Q14} - \frac{1}{2}(\sigma^{Q14})^2)t + \sigma^{Q14} W_t} & S_0 = S_0, \quad 0 \le t < t_1 \\ S_t^{Q_{23}} = \phi_1 S_1 e^{(\mu^{Q23} - \frac{1}{2}(\sigma^{Q23})^2)t + \sigma^{Q23} W_t} & S_1 = \lim_{t \to t_1} S_t^{Q_{14}}, \quad t_1 \le t < t_3 \\ S_t^{Q_{14}} = \phi_2 S_2 e^{(\mu^{Q14} - \frac{1}{2}(\sigma^{Q14})^2)t + \sigma^{Q14} W_t} & S_2 = \lim_{t \to t_3} S_t^{Q_{23}}, \quad t_3 \le t < t_5 \\ S_t^{Q_{23}} = \phi_3 S_3 e^{(\mu^{Q23} - \frac{1}{2}(\sigma^{Q23})^2)t + \sigma^{Q23} W_t} & S_3 = \lim_{t \to t_5} S_t^{Q_{23}}, \quad t_5 \le t < t_7 \\ \quad \dots & \quad \dots \qquad \dots \end{cases} \tag{2.4.1}
$$

Here, $\phi_1, \phi_2, \phi_3 \ldots$ are jump coefficients corresponding to jump times $t_1, t_3, t_5, \ldots$, and can be estimated as $\hat{\phi}_1 = \frac{S_{t_1}}{\lim_{t \to t_1} \hat{S}_t^{Q_{14}}}$, $\hat{\phi}_2 = \frac{S_{t_3}}{\lim_{t \to t_3} \hat{S}_t^{Q_{23}}}$, $\hat{\phi}_3 = \frac{S_{t_5}}{\lim_{t \to t_5} \hat{S}_t^{Q_{14}}}$, ….

GBM Model with Jumps 2.4.2 (Quarterly GBM Model with Jumps): Let $[0, t_1), [t_1, t_2), [t_2, t_3), [t_3, t_4), [t_4, t_5) \ldots$ be the time intervals as defined in Data Partition Process 2.3.2. Again, the GBM solution process with jumps of (2.3.2) has the following form:

$$
S_t = \begin{cases} S_t^{Q_3} = S_0 e^{(\mu^{Q_3} - \frac{1}{2}(\sigma^{Q_3})^2)t + \sigma^{Q_3} W_t} & S_0 = S_0, \quad 0 \le t < t_1 \\ S_t^{Q_4} = \phi_1 S_1 e^{(\mu^{Q_4} - \frac{1}{2}(\sigma^{Q_4})^2)t + \sigma^{Q_4} W_t} & S_1 = \lim_{t \to t_1} S_{t_1}^{Q_3}, \quad t_1 \le t < t_2 \\ S_t^{Q_1} = \phi_2 S_2 e^{(\mu^{Q_1} - \frac{1}{2}(\sigma^{Q_1})^2)t + \sigma^{Q_1} W_t} & S_2 = \lim_{t \to t_2} S_t^{Q_4}, \quad t_2 \le t < t_3 \\ S_t^{Q_2} = \phi_3 S_3 e^{(\mu^{Q_2} - \frac{1}{2}(\sigma^{Q_2})^2)t + \sigma^{Q_2} W_t} & S_3 = \lim_{t \to t_3} S_t^{Q_1}, \quad t_3 \le t < t_4 \\ S_t^{Q_3} = \phi_4 S_4 e^{(\mu^{Q_3} - \frac{1}{2}(\sigma^{Q_3})^2)t + \sigma^{Q_3} W_t} & S_4 = \lim_{t \to t_4} S_t^{Q_2} \quad t_4 \le t < t_5 \\ \quad \dots & \quad \dots \qquad \dots \end{cases} \tag{2.4.2}
$$

Here, $\phi_1, \phi_2, \phi_3 \ldots$ are jump coefficients corresponding to the jump time $t_1, t_2, t_3, t_4, \ldots$, and can be estimated as $\hat{\phi}_1 = \frac{S_{t_1}}{\lim_{t \to t_1} \hat{S}_t^{Q_1}}$, $\hat{\phi}_2 = \frac{S_{t_2}}{\lim_{t \to t_2} \hat{S}_t^{Q_2}}$, $\hat{\phi}_3 = \frac{S_{t_3}}{\lim_{t \to t_3} \hat{S}_t^{Q_3}}$, $\hat{\phi}_4 = \frac{S_{t_4}}{\lim_{t \to t_4} \hat{S}_t^{Q_4}}$, ….

Figure 2.4.1 is the result of prediction of stock X's price process of (2.3.3) with jumps and of (2.3.7) with jumps. We see that the red and blue curves are not as smooth as green and orange curves, this is because of the fact that there are jumps in green and orange curves of (2.3.3) and (2.3.7), respectively with respect to jumps. It is obvious that models with jumps provide better predicted results.

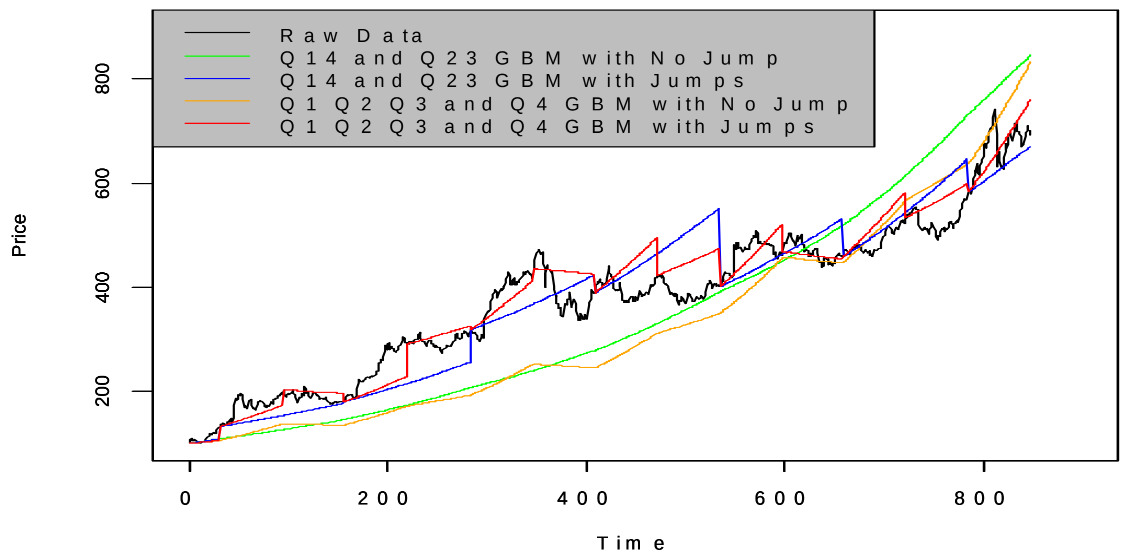


Figure 2.4.1 Comparison of Models (2.3.3), (2.3.7) with and without Jumps of Stock X

Table 2.4.1 provides the basic statistics that reflects the accuracy of model (2.3.1), (2.4.1) and (2.3.5) with jumps.

Table 2.4.1 Basic Statistics for Linear Models

(2.2.1), (2.3.3), (2.3.7), (2.3.10), (2.3.13) and with Jumps (2.3.3), (2.3.7) of Stock X

Model	$\overline{r}$	S_r^2	S_r	No. of Intervals	No. of Parameters
GBM with Overall Data	28.29653	8752.84	93.55661	1	2
Q14 and Q23 GBM without jumps	29.67727	8836.837	94.00445	8	4

Quarterly GBM without jumps	53.49948	5570.643	74.63674	14	8
Monthly GBM without jumps	-80.10483	11754.25	108.4170	41	82
Unequal-Interval GBM without jumps	24.91557	3992.349	63.18504	39	78
Q14 and Q23 GBM with jumps	1.759521	3181.759	56.40708	8	11
Quarterly GBM with jumps	-10.26338	1450.633	38.08717	14	21

Remark 2.4.1: From the Figure 2.4.1 and Table 2.4.1, we notice that models with jumps are much better than models without jumps. Moreover, the quarterly data partition has better result than half yearly data partition.

GBM Models with Jumps 2.4.3 (Monthly GBM Model with Jumps): Let $[0,t_1),[t_1,t_2),[t_2,t_3),\ [t_3,t_4),[t_4,t_5)...\ [t_{m-1},t_m)$ be the m monthly time intervals as defined in data partition process (2.3.3). The GBM solution process with jumps of (2.3.10) takes the form:

$$S_t = \begin{cases} S_t^{M_1} = S_0 e^{(\mu^{M_1} - \frac{1}{2}(\sigma^{M_1})^2)t + \sigma^{M_1} W_t} & S_0 = S_0, \quad t_0 \le t < t_1 \\ S_t^{M_2} = \phi_1 S_1 e^{(\mu^{M_2} - \frac{1}{2}(\sigma^{M_2})^2)t + \sigma^{M_2} W_t} & S_1 = \lim_{t \to t_1} S_t^{M_1}, \quad t_1 \le t < t_2 \\ \quad \dots & \quad \dots \qquad \dots \\ S_t^{M_m} = \phi_{m-1} S_{m-1} e^{(\mu^{M_m} - \frac{1}{2}(\sigma^{M_m})^2)t + \sigma^{M_m} W_t} & S_{m-1} = \lim_{t \to t_{m-1}} S_t^{M_{m-1}}, \quad t_{m-1} \le t < t_m \end{cases} \tag{2.4.3}$$

Here, $\phi_1,\ \phi_2, \phi_3\ \dots$ are jump coefficients and can be estimated as

$$\hat{\phi}_1 = \frac{S_{t_1}}{\lim_{t \to t_1} \hat{S}_t^{M_1}},\ \hat{\phi}_2 = \frac{S_{t_2}}{\lim_{t \to t_2} \hat{S}_t^{M_2}},\ \hat{\phi}_3 = \frac{S_{t_3}}{\lim_{t \to t_3} \hat{S}_t^{M_3}},\ \hat{\phi}_4 = \frac{S_{t_4}}{\lim_{t \to t_4} \hat{S}_t^{M_4}}, \dots.$$

The methods of estimation parameters are the same as we mentioned in Section 2.3. For stock X, the estimated parameters $\hat{\mu}^{M_i}$ and $\hat{\sigma}^{M_i}$, $i = 1,\ldots,m$, are given in Appendix A1, the estimated jump coefficients $\hat{\phi}_1$, $\hat{\phi}_2$, $\hat{\phi}_3$ … $\hat{\phi}_{m-1}$ are provided in Appendix A3. Figure 2.4.2 is the prediction of stock X's prices process. Table 2.4.2 provides the basic statistics of model (2.3.13) with jumps.

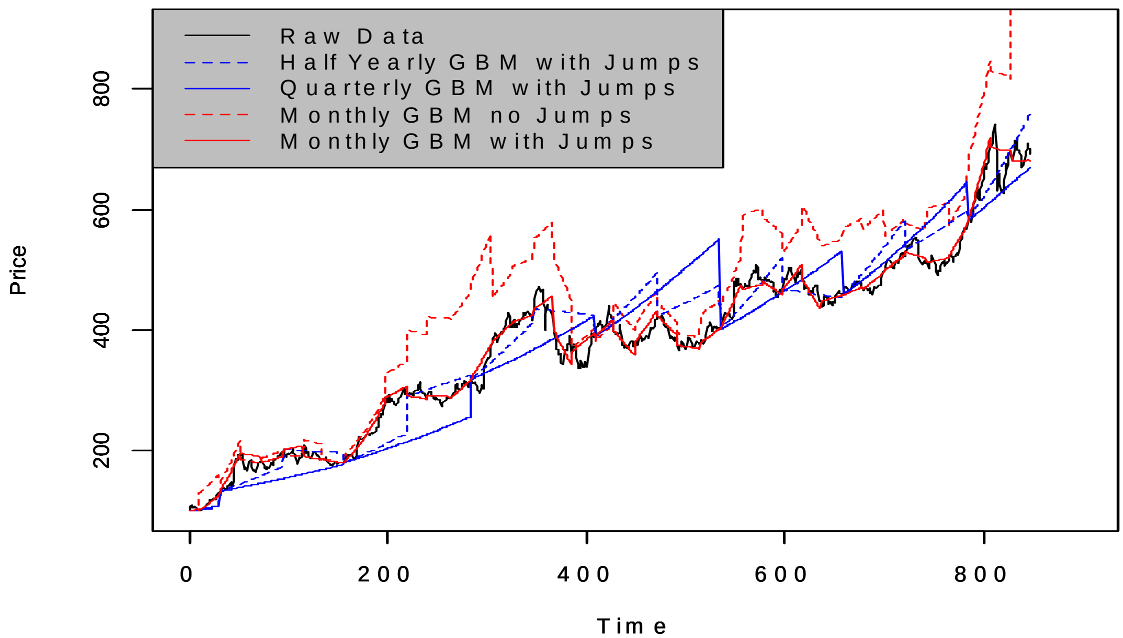


Figure 2.4.2 Comparison of Models (2.3.10) and (2.3.13) with and without Jumps of Stock X

Table 2.4.2 Basic Statistics for Linear Models (2.2.1), (2.3.3), (2.3.7), and with and without Jumps Models (2.3.10), (2.3.13) of Stock X

Model	$\bar{r}$	S_r^2	S_r	No. of Intervals	No. of Parameters
GBM with Overall Data	28.29653	8752.84	93.55661	1	2
Q14 and Q23 GBM without jumps	29.67727	8836.837	94.00445	8	4
Quarterly GBM without jumps	53.49948	5570.643	74.63674	14	8

Monthly GBM without jumps	-80.10483	11754.25	108.4170	41	82
Unequal Interval GBM without jumps	24.91557	3992.349	63.18504	39	78
Q14 and Q23 GBM with jumps	1.759521	3181.759	56.40708	8	11
Quarterly GBM with jumps	-10.26338	1450.633	38.08717	14	21
Monthly GBM with jumps	-1.22683	207.3278	14.3989	41	122

Remark 2.4.2: The solid red curve in Figure 2.4.2 follows the same dynamic pattern as the dashed red curve. The only difference between these two curves is that Monthly GBM with Jumps model doesn't accumulate large error, while the models without jumps do accumulate large errors. This can also be further confirmed from basic statistics in Table 2.4.2. The monthly GBM model with jumps has the least mean, variance, and standard error of residual error.

GBM Model with Jumps 2.4.4 (Unequal Interval GBM Model with Jumps): Let $[0,t_1),[t_1,t_2),[t_2,t_3),\ [t_3,t_4),[t_4,t_5)...\ [t_{n-1},t_n)$ be the n time intervals as defined in Data Partition Process (2.3.4). Similarly, by following definition [16, 26, 27], the GBM solution processes with jumps of (2.3.11) has the following form:

$$S_t = \begin{cases} S_t^{I_1} = S_0 e^{(\mu^{I_1} - \frac{1}{2}(\sigma^{I_1})^2)t + \sigma^{I_1} W_t} & S_0 = S_0, \quad t_0 \le t < t_1 \\ S_t^{I_2} = \phi_1 S_1 e^{(\mu^{I_2} - \frac{1}{2}(\sigma^{I_2})^2)t + \sigma^{I_2} W_t} & S_1 = \lim\limits_{t \to t_1} S_t^{I_1}, \quad t_1 \le t < t_2 \\ \dots & \dots \qquad \dots \\ S_t^{I_n} = \phi_{n-1} S_{n-1} e^{(\mu^{I_n} - \frac{1}{2}(\sigma^{I_n})^2)t + \sigma^{I_n} W_t} & S_{n-1} = \lim\limits_{t \to t_{n-1}} S_t^{I_{n-1}}, \quad t_{n-1} \le t < t_n \end{cases} \tag{2.4.4}$$

There are $2 \times n$ parameters μ^{I_i} and σ^{I_i} , $i = 1,...,n$, and these parameters need to be estimated. n is the number of intervals of stock price process. We adapt the earlier procedure to create unequal intervals, and estimate the drifts and volatilities as in Section 2.3. Here, ϕ_1, ϕ_2, ϕ_3 … are jump coefficients corresponding to jump times at $t_1, t_2, t_3, \ldots$ and can be estimated as

$$\hat{\phi}_1 = \frac{S_{t_1}}{\lim\limits_{t \to t_1} \hat{S}_t^{I_1}},\ \hat{\phi}_2 = \frac{S_{t_2}}{\lim\limits_{t \to t_2} \hat{S}_t^{I_2}},\ \hat{\phi}_3 = \frac{S_{t_3}}{\lim\limits_{t \to t_3} \hat{S}_t^{I_3}},\ \hat{\phi}_4 = \frac{S_{t_4}}{\lim\limits_{t \to t_4} \hat{S}_t^{I_4}}, \ldots.$$

For stock X, we use the estimated parameters $\hat{\mu}^{I_i}$ and $\hat{\sigma}^{I_i}$, $i = 1, \ldots, n$, (Appendix A2), and the estimated jump coefficients $\hat{\phi}_1$, $\hat{\phi}_2$, $\hat{\phi}_3$ … $\hat{\phi}_{m-1}$ (Appendix A4). Figure 2.4.3 is the predicted process of stock X. Table 2.4.3 provides the basic statistics for model in Model (2.3.13) with jumps.

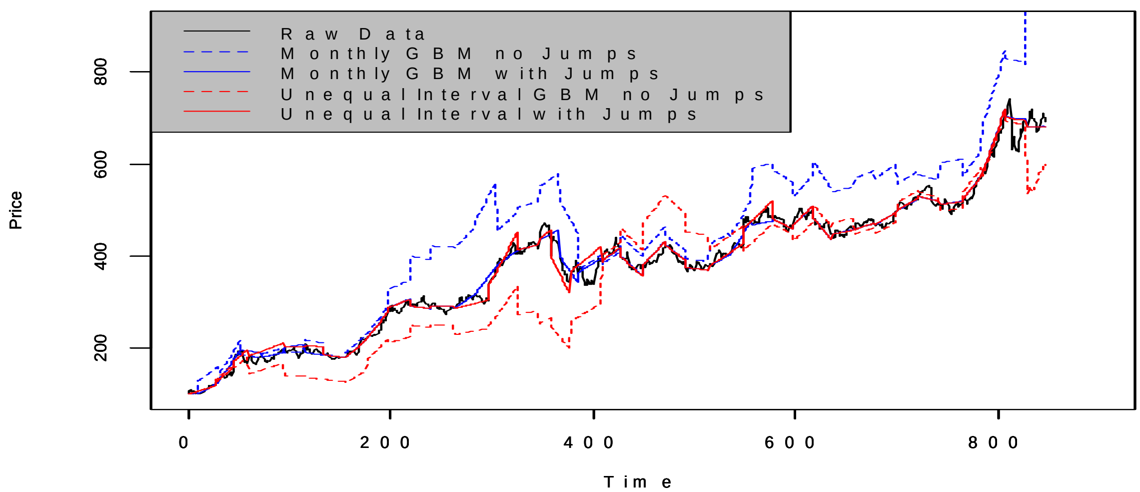


Figure 2.4.3 Comparisons of Models with and without jumps (2.3.10), (2.3.13), (2.4.3) of Stock X

Table 2.4.3 Basic Statistics for Linear Models (2.2.1) and Models with and without Jumps (2.3.3), (2.3.7), (2.3.10), (2.3.13) of Stock X

Model	$\bar{r}$	S_r^2	S_r	No. of Intervals	No. of Parameters
GBM with Overall Data	28.29653	8752.84	93.55661	1	2

Q14 and Q23 GBM without jumps	29.67727	8836.837	94.00445	8	4
Quarterly GBM without jumps	53.49948	5570.643	74.63674	14	8
Monthly GBM without jumps	-80.10483	11754.25	108.4170	41	82
Unequal Interval GBM without jumps	24.91557	3992.349	63.18504	39	78
Q14 and Q23 GBM with jumps	1.759521	3181.759	56.40708	8	11
Quarterly GBM with jumps	-10.26338	1450.633	38.08717	14	21
Monthly GBM with jumps	-1.22683	207.3278	14.3989	41	122
Unequal Interval GBM with jumps	-1.962899	258.1040	16.06562	39	116

Remark 2.4.3: In Table 2.4.3 we remark that overall the Monthly GBM Model with jumps and Unequal Interval GBM model with jumps, relatively provides the least mean and the variance of residual error. Generally speaking, for stock X, the GBM models with jumps perform better than those GBM models without jumps.

2.5 Illustration of GBM Models to Data Set of Stock Y

Before we make conclusions about this chapter, we apply the developed linear stochastic models to the other company's (Y) stock price process. It is more than 22 years and has 5630 observations. Figure 2.5.1 shows its daily adjusted closing price from 9/10/1984 to 12/31/2006.

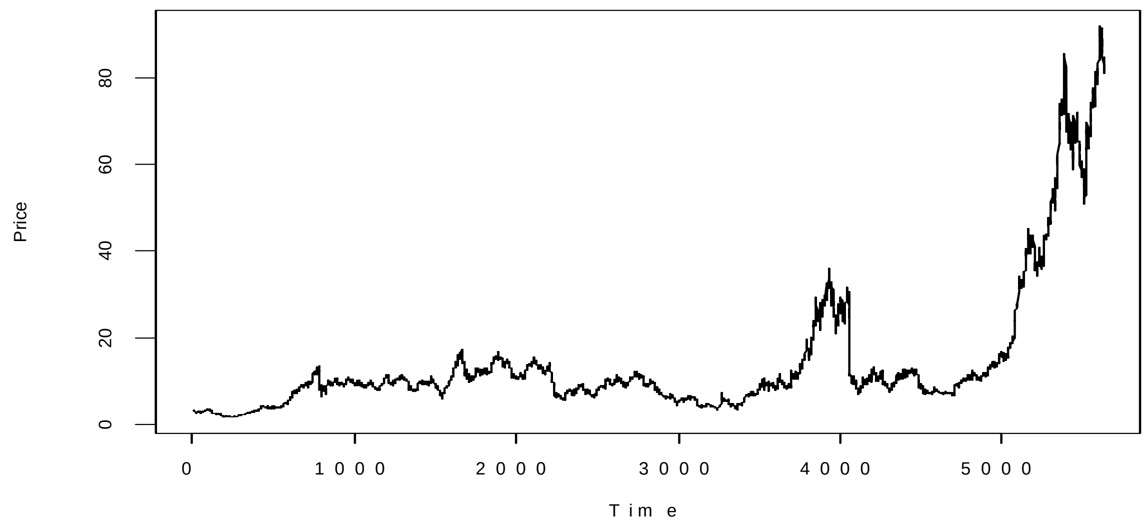


Figure 2.5.1 Daily Adjusted Closing Price for Stock Y

We apply those linear models, under different data portioning process with or without jumps to the price data set of stock Y. The procedures are exactly the same as those applied to stock X in Sections 2.2, 2.3 and 2.4. To minimize the repetition, here we only give Figure 2.5.2 with regard to the best two estimated models and the summary of basic statistics of different linear models of stock Y in Table 2.5.1.

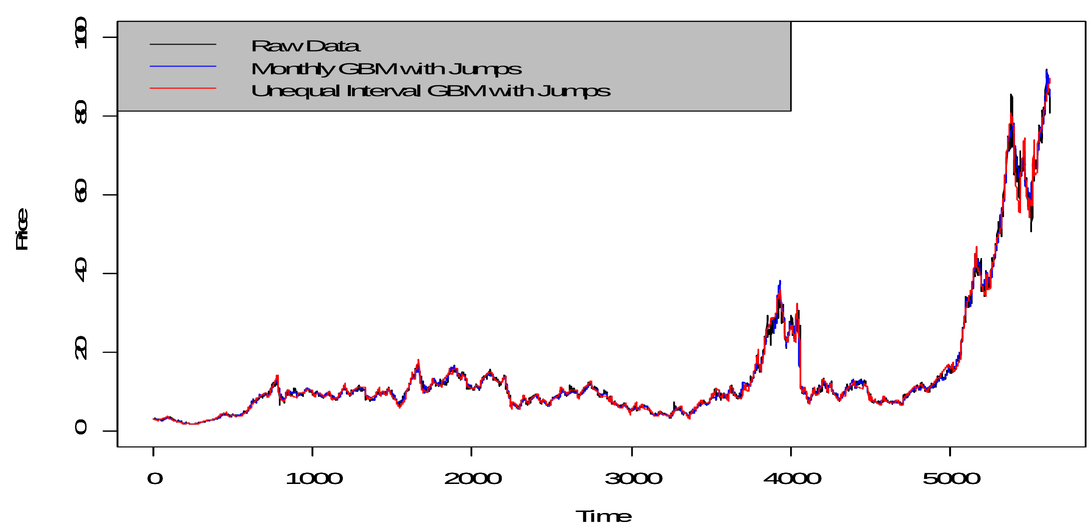


Figure 2.5.2 The Best Two Estimated Models of Stock Y

Table 2.5.1 Basic Statistics for Linear Models without Jumps (2.2.1), with and without Jumps (2.3.3), (2.3.7), (2.3.10), (2.3.13) of Stock Y

Model	$\bar{r}$	S_r^2	S_r	No. of Intervals	No. of Parameters
GBM with Overall Data	-10.22182	211.7418	14.55135	1	2
Q14 and Q23 GBM without jumps	-10.48387	214.6396	14.65058	45	4
Quarterly GBM without jumps	-10.54319	216.0761	14.69953	89	8
Monthly GBM without jumps	-0.5712012	137.0789	11.70807	268	536
Unequal Interval GBM without jumps	-1.461658	77.70724	8.815171	256	512
Q14 and Q23 GBM with jumps	0.993067	26.28088	5.126488	45	48
Quarterly GBM with jumps	0.4321374	12.24818	3.49974	89	96
Monthly GBM with jumps	-0.0098261	1.206479	1.098399	268	803
Unequal Interval GBM with jumps	-0.0124816	1.199703	1.095310	256	767

Remark 2.5.1: From, in Table 2.5.1 we note that for stock Y, two models: the Monthly GBM Model with jumps and Unequal Interval GBM model with jumps, both relatively provide the least mean, variance of residual error. Generally speaking, for stock Y, the GBM models with jumps perform better than those GBM models without jumps.

2.6 Illustration of GBM Models to Data Set of S&P 500 Index

In our previous estimation, we applied the above developed linear stochastic models to two individual stock price data sets of X and Y. In this section, we apply the GBM models to S&P500 Index. It is more than 59 years, and has 14844 observations. Figure 2.6.1 shows its daily adjusted closing price from 1/1/1950 to 12/31/2008.

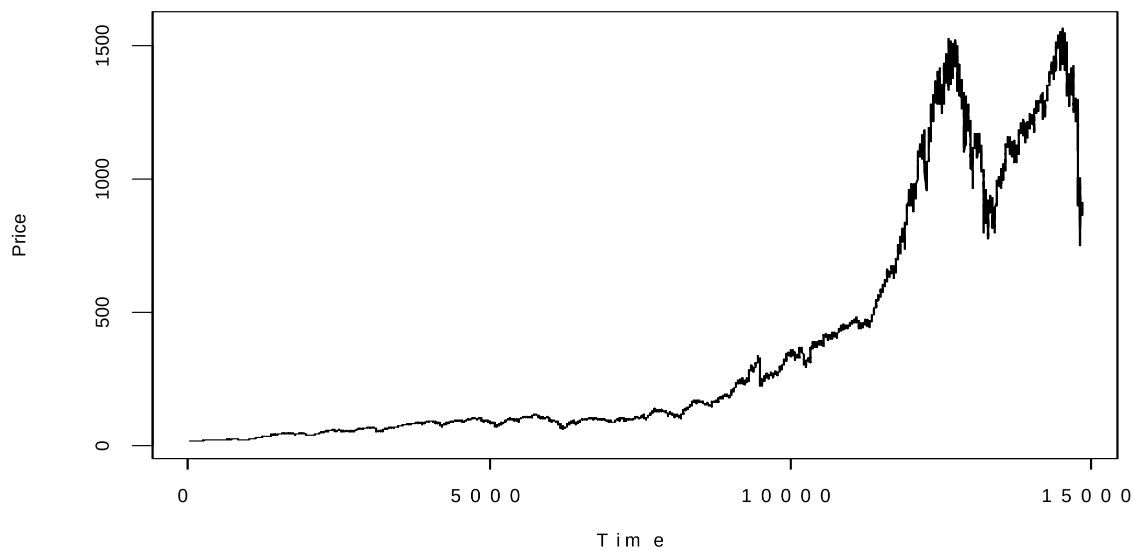


Figure 2.6.1 Daily Adjusted Closing Price for S&P500 Index

We apply same linear models, under different data portioning process with or without jumps to the data set of S&P500 Index. The procedures are exactly the same as those applied to stock X in Sections 2.2, 2.3, 2.4, and section 2.5. To minimize the repetition, here we only give Figure 2.6.2 with regard to the best two estimated models and the summary of basic statistics of different linear models of S&P500 Index in Table 2.6.1.

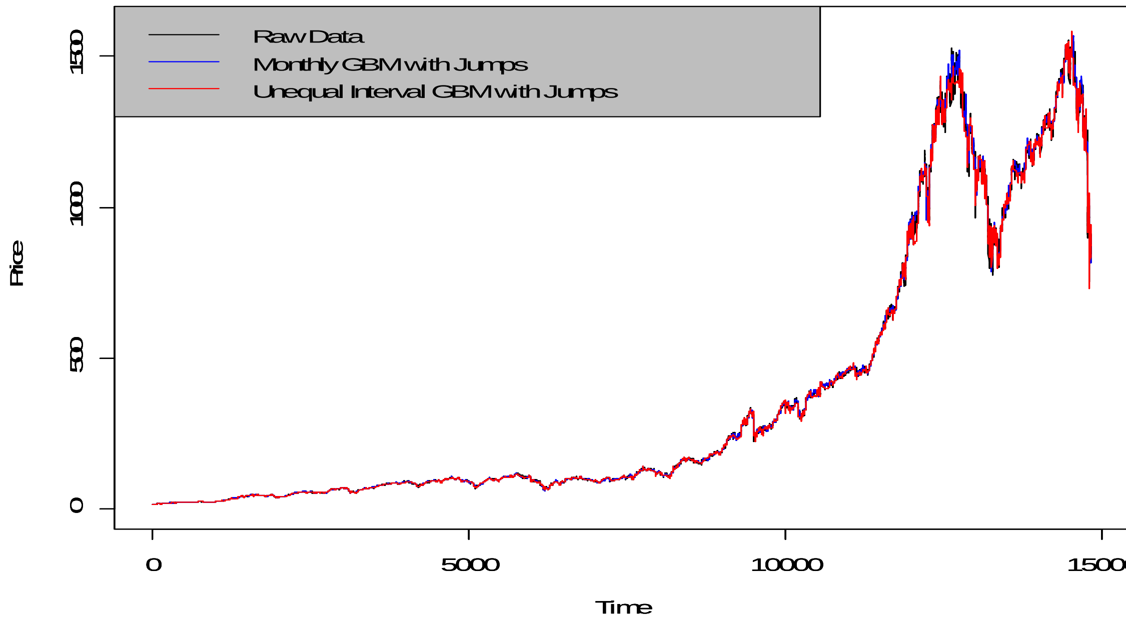


Figure 2.6.2 The Best Two Estimated Models of S&P500 Index

Table 2.6.1 Basic Statistics for Linear Models without Jumps (2.2.1), with and without Jumps (2.3.3), (2.3.7), (2.3.10), (2.3.13) of S&P500 Index

Model	$\overline{r}$	S_r^2	S_r	No. of Intervals	No. of Parameters
GBM with Overall Data	141.3899	55048.98	234.6252	1	2
Q14 and Q23 GBM without jumps	141.6477	55031.77	234.5885	119	4
Quarterly GBM without jumps	142.6970	55408.23	235.3895	236	8
Monthly GBM without jumps	-0.0171440	148.0943	12.1694	708	1416
Unequal Interval GBM without jumps	2.541547	210.3291	14.50273	570	1140

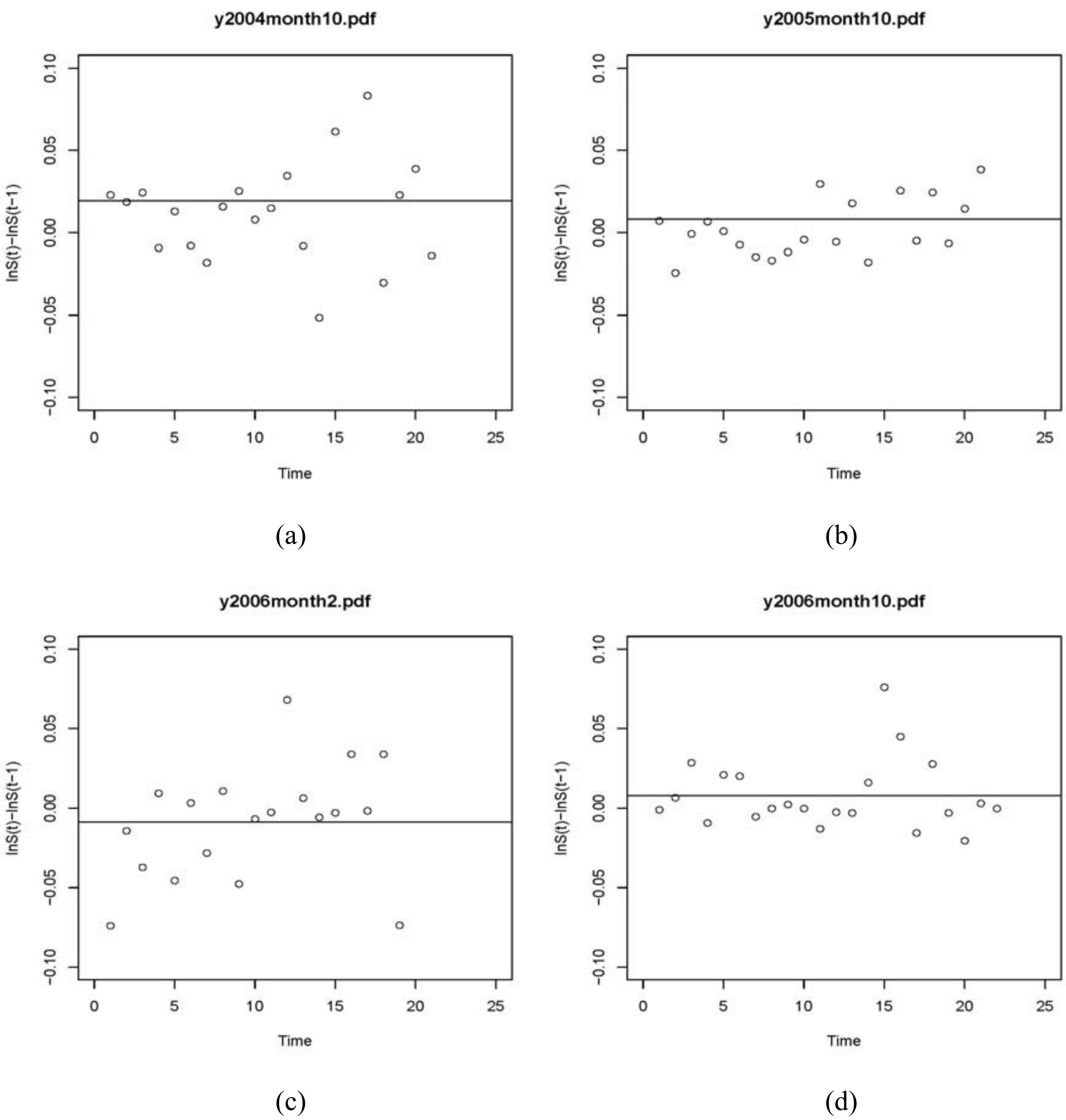
Q14 and Q23 GBM with jumps	-0.2859031	4717.424	68.6835	119	122
Quarterly GBM with jumps	-0.8944856	1954.800	44.21312	236	243
Monthly GBM with jumps	58.30902	486.4579	22.05579	708	2123
Unequal Interval GBM with jumps	2.471564	210.2159	14.49882	570	1709

Remark 2.6.1: Again, from Table 2.6.1 we remark that for S&P500 Index, there are two models: the Monthly GBM Model without jumps and Unequal Interval GBM model with jumps, both relatively, provide the least mean and variance of residual error with the least number of time intervals. Generally speaking, for S&P500 Index, the GBM models with jumps perform better than those GBM models without jumps.

2.7 Conclusions and Comments

In this chapter, by employing classical model building process, we develop the modified version of GBM models under different data partitioning processes and coupled with or without jumps. The main focus was how to modify the existing GBM model in order to have a best fit with least mean and variance of residual error. Based on the study of three data sets in Chapter 2, one can immediately draw a couple of conclusions. (i) The first one is the usage of GBM model of overall dataset might not give us a good fit. Data partitioning improves the result. (ii) Also we show that models with jumps perform much better than the ones without jumps. This improvement is largely due to the accumulated errors in the model without jumps. Moreover, the environmental random perturbations cause to modify parameters in GBM model. In the next chapters, we will focus on models with jumps using monthly data partitioning and unequal interval data partitioning process, since models with these two data partitioning with jumps have less mean and variance of residual error.

The GBM process is the solution of a linear stochastic differential equation. Because the drift and volatility rate functions are linear. From the equation (2.1.6), we know that $y_t = \ln S_t - \ln S_{t-1}$ is expected to have a random pattern around the $\hat{\mu} - \frac{1}{2}\hat{\sigma}^2$. Moreover, we would like to see the values in the neighborhood of the line y= $\hat{\mu} - \frac{1}{2}\hat{\sigma}^2$. Figure 2.7.1 is a residual plot of monthly GBM model of Stock X.



(a) (b) (c) (d)

Figure 2.7.1 Some Residual Plots of Stock X

In Figure 2.7.1(a), we see that the residual values start out close to the line, then deviate from it. In Figure 2.7.1(b), there are a lot of runs of many negative residuals in a row. In Figure 2.7.1(c), we see there is a trend of the residuals. The magnitude of the residuals gets bigger as time goes on. Moreover, in Figure 2.7.1(b) and (d), we see the number of positive points are much larger than the number of negative points. From these observations and the Q-Q plots for model (2.2.1) (Figure 2.2.3), (2.3.3) (Figure 2.3.6) and (2.3.7) (Figure 2.3.8) suggest that the linear GBM model and its generalized models are inadequate to represent the stock price models. All these indicate that the linear model might not be good enough to fit the dataset. To build more precise models for competitive business processes, even a small difference is important. In Chapter 3, we find a remedy to partially solve the cited limitations by developing the nonlinear stochastic models.

Chapter 3
Nonlinear Stochastic Models

3.0 Introduction

In Chapter 2, we initiated the development of stochastic modeling by using the classical modeling procedure in a systematic way. We made an attempt to modify the GBM model. The developed modified GBM models raised the issue about the stochastic linear models of stock price processes. This was eluded in Section 2.7. There are many nonlinear stochastic models that describe the stochastic behavior of asset price in finance. In this chapter, we will focus on the nonlinear stochastic models. In Chapter 2, we have already seen that modified GBM models with monthly and unequal interval data partitioning process with jumps have better results in terms of minimum mean and variance of residual error, even though, we needed to estimate more parameters. Here, we will just focus on monthly and unequal interval data partitioning processes with jumps. In Sections 3.1, 3.2 and 3.3, we develop three different nonlinear stochastic models to our three datasets. In each section, we will first introduce the nonlinear stochastic model. We then develop the monthly and unequal interval nonlinear models with jumps based on each data set. Furthermore, we analyze and compare the nonlinear models with corresponding modified GBM models. In Sections 3.4 and 3.5, we illustrate nonlinear stochastic models in the context of data sets stock Y and S&P 500 Index respectively. Finally, conclusions are drawn in Section 3.6.

3.1 Stochastic Nonlinear Dynamic Model 1 (Black-Karasinski Model)

Black-Karasinski (BK) model [6] describes a short-term interest rate process. It takes the following form

$$dS_t = (\alpha \ln S_t + \beta + \frac{\sigma^2}{2})S_t dt + \sigma S_t dW_t \qquad (3.1.1)$$

where, α, β and σ are parameters and W_t is Brownian motion.

To test the existence of a unique solution, let K be any number greater than $(\alpha M_1 + \beta + \frac{\sigma^2}{2})^2 + \sigma^2$, and L be any number greater than $|\alpha M_2| + |\beta| + \frac{\sigma^2}{2} + |\sigma|$, where M_1 and M_2 are sufficiently large constants such that $M_1 \geq \ln S_t$ and $M_2 \geq | S_{t_2} \ln \frac{S_{t_1}}{S_{t_2}} + \ln S_{t_1} |$.

It is obvious that equation (3.1.1) satisfy the conditions (1.3.2) and (1.3.3). S_t is the unique solution of (3.1.1). Even though, The BK model usually describes a short-term interest rate process, it may also be applied to the short-term stock price process.

We note that the volatility function is linear and drift function is nonlinear. In order to derive the regression equation, we use the following transformation $V_t = \ln S_t$ and apply $It\hat{o} - Doob$ differential formula (1.3.5) to obtain,

$$\begin{aligned}
dV_t &= \frac{\partial}{\partial S_t}(\ln S_t)dS_t + \frac{1}{2}(\frac{\partial^2}{\partial S_t^{\,2}}(\ln S_t))(dS_t)^2 \\
&= \frac{1}{S_t}dS_t + \frac{1}{2}(-\frac{1}{S_t^{\,2}})(dS_t)^2 \\
&= \frac{1}{S_t}((\alpha \ln S_t + \beta + \frac{\sigma^2}{2})S_t dt + \sigma S_t dW_t) - \frac{1}{2}\frac{1}{S_t^{\,2}}\sigma^2 S_t^{\,2}(dW_t)^2 \\
&= (\alpha \ln S_t + \beta)dt + \sigma dW_t
\end{aligned}$$

Then,

$$dV_t = (\alpha V_t + \beta)dt + \sigma dW_t \tag{3.1.2}$$

By using the Euler type discretization process [24], stochastic differential equation (3.1.2) can be reduced to

$$V_t - V_{t-1} = (\alpha V_{t-1} + \beta)\Delta t + \sigma(W_t - W_{t-1})\,. \tag{3.1.3}$$

From $\varepsilon_t = W_t - W_{t-1}$ and $\Delta t = 1$, equation (3.1.3) can be rewritten as

$$V_t = (\alpha + 1)V_{t-1} + \beta + \sigma\varepsilon_t \tag{3.1.4}$$

where α, β and σ are as defined in (3.1.1). By applying the least square regression method [35] and using above cited data sets, we can estimate these parameters.

Nonlinear Stochastic Model 3.1.1 (Monthly Nonlinear Model 1 with Jumps): Let $[0,t_1),[t_1,t_2),[t_2,t_3),\ [t_3,t_4),[t_4,t_5)...\ [t_{m-1},t_m)$ be the m monthly time intervals as defined in Data Partition Process (2.3.3). The nonlinear stochastic model is described by following stochastic differential equation:

$$dS_t^{M_i} = (\alpha^{M_i} \ln S_t^{M_i} + \beta^{M_i} + \frac{(\sigma^{M_i})^2}{2})S_t^{M_i} dt + \sigma^{M_i} S_t^{M_i} dW_t,\ S_0 = S_0$$

$$\text{if } t_{i-1} \le t < t_i,\ i = 1,...,m. \tag{3.1.5}$$

α^{M_i}, β^{M_i}, and $\sigma^{M_i}, i = 1,...,m$ are parameters. These parameters need to be estimated. By following definition [16, 26, 27], the solution of (3.1.5) takes the form

$$S(t) = \begin{cases} S_1(t,t_0,S_0), & t_0 \le t < t_1 \\ \phi_1 S_2(t,t_1,S_1), & t_1 \le t < t_2, \quad S_1 = \lim\limits_{t \to t_1^-} S_1(t,t_0,S_0) \\ \quad ... & \quad ... \quad\quad\quad\quad ... \\ \phi_{m-1} S_m(t,t_{m-1},S_{m-1}), & t_{m-1} \le t < t_m, \quad S_{m-1} = \lim\limits_{t \to t_{m-1}^-} S_{m-1}(t,t_{m-2},S_{n-2}) \end{cases} \tag{3.1.6}$$

Here, S_0 is the initial value of the stock price process. $\phi_1, \phi_{2,} ..., \phi_{m-1}$ are jumps. These jumps are estimated by $\hat{\phi}_1 = \dfrac{S_{t_1}}{\lim\limits_{t \to t_1} \hat{S}_1}, \hat{\phi}_2 = \dfrac{S_{t_2}}{\lim\limits_{t \to t_2} \hat{S}_2}, ..., \hat{\phi}_{m-1} = \dfrac{S_{t_{m-1}}}{\lim\limits_{t \to t_{m-1}} \hat{S}_{m-1}}$.

The estimated parameters in Monthly Nonlinear Stochastic Model (3.1.1) of stock X are presented in Table 3.1.1. The AIC (Akaike's information criterion) criterion [3] defined in (1.5.6). Here, we use AIC as the criterion whenever we need to compare different models. The preferred model is the model with the lowest AIC value.

Table 3.1.1 Estimated Parameters in Model 3.1.1 of Stock X

Interval Index	Monthly GBM Model with Jumps			Monthly Nonlinear Model 1 with Jumps			
	$\hat{\mu}$	$\hat{\sigma}$	AIC	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}$	AIC
1	0.003217	0.037758	24.99217	-0.93259	4.347902	0.025463	20.74442
2	0.011469	0.021817	42.50585	0.010445	-0.03801	0.021799	44.52993
3	0.01924	0.041526	83.16599	-0.00148	0.025789	0.041526	85.15694
4	-0.00152	0.037226	82.10229	-0.27051	1.398991	0.034175	80.71667
5	0.002806	0.019117	57.44675	-0.05069	0.26617	0.018992	59.16818
6	0.001162	0.029425	72.08337	-0.3736	1.966097	0.026578	69.95869

7	-0.0017	0.027966	67.90378	-0.32282	1.700644	0.02585	66.93144
8	-0.00177	0.012578	39.10143	-0.23453	1.217925	0.011507	37.27279
9	0.009641	0.020969	63.75357	-0.03145	0.175542	0.020858	65.44605
10	0.01115	0.016296	61.93046	0.089916	-0.48049	0.015445	61.48279
11	0.002921	0.021646	83.43298	-0.33056	1.873092	0.019655	81.20437
12	-0.00096	0.016438	66.92268	-0.2209	1.257716	0.015708	67.08642
13	-0.00017	0.013734	66.01481	-0.19786	1.119452	0.013038	65.65693
14	0.004932	0.015047	67.11831	-0.14534	0.835089	0.014276	66.79189
15	0.008161	0.02977	98.04718	0.027813	-0.15265	0.029718	99.9177
16	0.004194	0.01872	87.811	-0.18623	1.118515	0.017263	86.41869
17	0.001243	0.013153	74.5625	-0.22105	1.33551	0.012174	73.30687
18	0.002672	0.033797	109.8641	-0.32346	1.974105	0.030375	107.9387
19	-0.00865	0.035981	101.9126	-0.39163	2.309854	0.028169	93.75399
20	0.003501	0.025925	105.6795	-0.07197	0.426261	0.025685	107.2806
21	0.003837	0.019752	83.25841	-0.26766	1.615277	0.017746	81.18764
22	-0.00514	0.018633	90.55914	-0.24664	1.46339	0.016451	86.83802
23	0.005609	0.01691	85.99189	-0.12762	0.767304	0.016503	87.09986
24	-0.00399	0.011988	65.98114	-0.04066	0.240058	0.011892	67.73263
25	-0.00081	0.014567	81.26636	-0.55654	3.301192	0.01207	74.65778
26	0.003112	0.015259	75.33679	-0.1616	0.969446	0.01435	74.72001
27	0.007971	0.022015	103.9302	-0.05786	0.35943	0.021685	105.1114
28	0.000937	0.014354	85.01112	-0.15455	0.956602	0.01374	85.10525
29	-0.0025	0.012368	73.32039	-0.11528	0.707825	0.012047	74.33965
30	0.004407	0.016734	87.08846	-0.41656	2.58325	0.013689	81.1518
31	-0.00564	0.016148	80.05819	-0.38312	2.351359	0.013406	74.47795
32	0.000967	0.013749	83.07253	-0.2835	1.7344	0.012763	81.86601
33	0.001484	0.011091	69.15906	-0.39337	2.423194	0.009461	64.82723
34	0.00256	0.011958	79.16138	-0.01537	0.097133	0.011955	81.16398
35	0.002363	0.009975	71.65205	-0.21045	1.315889	0.00921	70.31184
36	-0.00107	0.014438	88.62695	-0.07007	0.438718	0.014276	90.13775

37	0.000524	0.012519	88.16149	-0.34231	2.134238	0.011438	85.94986
38	0.00511	0.00978	65.80358	-0.02092	0.136596	0.009749	67.64959
39	0.009694	0.015521	108.0988	-0.02538	0.173118	0.015449	109.955
40	-0.00057	0.027629	125.7364	-0.15358	0.99984	0.026446	125.9705
41	2.59E-06	0.014992	96.70551	-0.28643	1.874382	0.013882	95.61875

Figures 3.1.1- 3.1.3 are the plots of predicted value of Monthly Nonlinear Model 1 of stock X with Jumps with observations ranging from 1 to 300, 300 to 600 and 600 to 848 respectively.

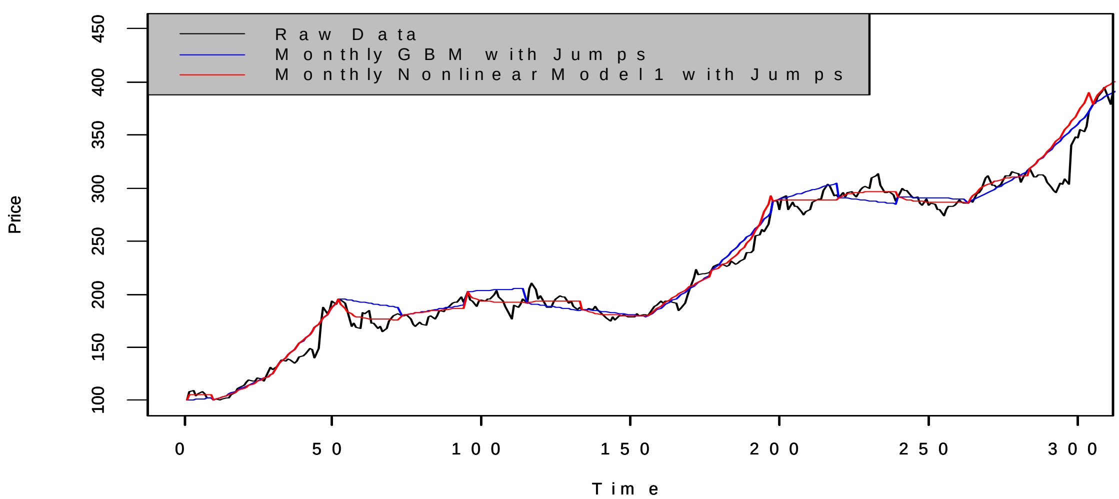


Figure 3.1.1 Comparison of Model 2.4.3 with Model 3.1.1 of Stock X (Observations 1-300)

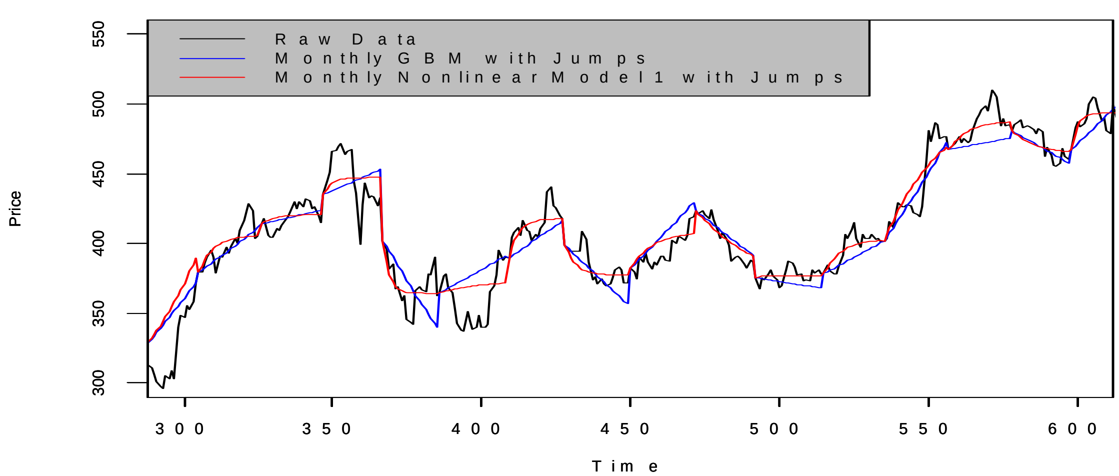


Figure 3.1.2 Comparison of Model 2.4.3 with Model 3.1.1 of Stock X (Observations 300-600)

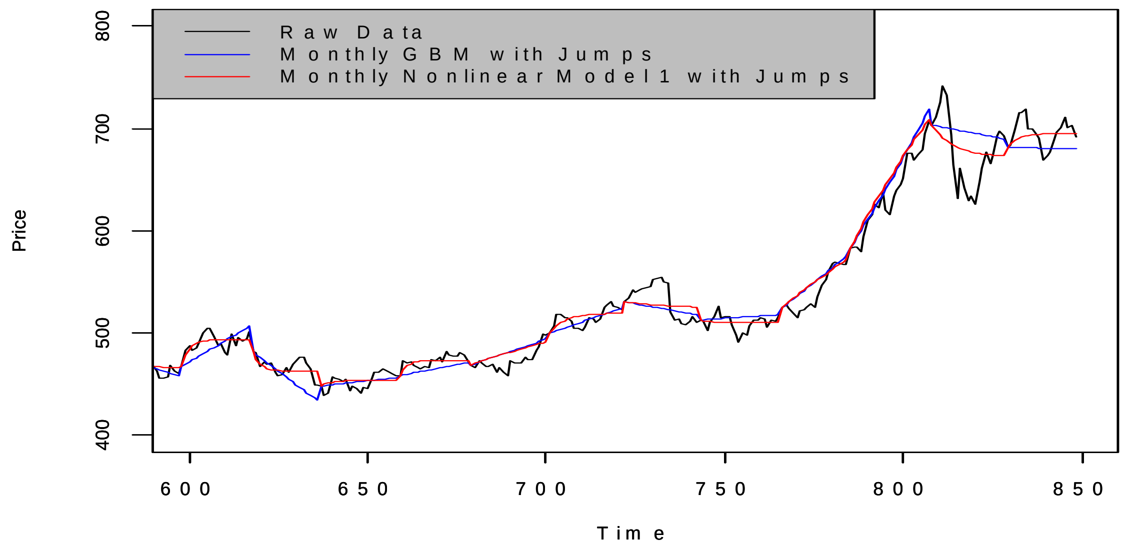


Figure 3.1.3 Comparison of Model 2.4.3 with Model 3.1.1 of Stock X (Observations 600-848)

Table 3.1.2 shows the overall basic statistics of Monthly GBM Model and Monthly Nonlinear Model 3.1.1.

Table 3.1.2 Basic Statistics of Model 3.1.1 of Stock X

Model	$\bar{r}$	S_r^2	S_r	No. of Intervals	No. of Parameters
Monthly GBM with Jumps	-1.242020	207.264	14.39667	41	122
Monthly Nonlinear Model 1 with Jumps	-1.928296	141.1754	11.88173	41	163

Remark 3.1.1: From Table 3.1.2 we can see that overall, the Monthly Nonlinear Model 3.1.1 with Jumps has less variance of the residual error. From the Table 3.1.1 and Figures 3.1.1 – 3.1.3, we remark that for some months, GBM Model is better than Nonlinear Model 3.1.1 in terms of AIC. For example, in the 2nd, 3rd, 5th, 9th, 12th, 15th month etc, GBM model has less AIC than Nonlinear Model 3.1.1. There are 17 out of 41 months (41%), that GBM model has less AIC than Nonlinear Model 1. We further note that the Nonlinear Model 3.1.1 has 3 parameters and the GBM model has 2 parameters.

Nonlinear Stochastic Model 3.1.2 (Unequal Interval Nonlinear Model 1 with Jumps): Let $[0,t_1),[t_1,t_2),[t_2,t_3), [t_3,t_4),[t_4,t_5)... [t_{n-1},t_n)$ be the n time intervals as defined in Data Partition Process 2.3.4. The nonlinear stochastic differential equation is described by:

$$dS_t^{I_i} = (\alpha^{I_i} \ln S_t^{I_i} + \beta^{I_i} + \frac{(\sigma^{I_i})^2}{2})S_t^{I_i}dt + \sigma^{I_i} S_t^{I_i} dW_t,$$

$$S_0 = S_0, \text{ if } t_{i-1} \le t < t_i,\ i = 1,...,n. \tag{3.1.7}$$

α^{I_i}, β^{I_i}, and σ^{I_i} , $i = 1,...,n$, are parameters which can be estimated by the method as described above.

By following definition [16, 26, 27], the solution of (3.1.7) is given by:

$$S(t) = \begin{cases} S_1(t,t_0,S_0), & t_0 \le t < t_1 \\ \phi_1 S_2(t,t_1,S_1), & t_1 \le t < t_2, \quad S_1 = \lim_{t\to t_1^-} S_1(t,t_0,S_0) \\ \quad ... & \quad ... \qquad\qquad ... \\ \phi_{n-1} S_m(t,t_{n-1},S_{n-1}), & t_{n-1} \le t < t_n, \quad S_{n-1} = \lim_{t\to t_{n-1}^-} S_{n-1}(t,t_{n-2},S_{n-2}) \end{cases} \tag{3.1.8}$$

S_0 is the initial value of the stock price process. $\phi_1, \phi_{2,} ..., \phi_{n-1}$ are jumps. These jumps are estimated by: $\hat{\phi}_1 = \dfrac{S_{t_1}}{\lim_{t\to t_1} \hat{S}_1}, \hat{\phi}_2 = \dfrac{S_{t_2}}{\lim_{t\to t_2} \hat{S}_2}, ..., \hat{\phi}_{n-1} = \dfrac{S_{t_{n-1}}}{\lim_{t\to t_{n-1}} \hat{S}_{n-1}}.$

The parameters of stochastic model (3.1.7) are presented in Table 3.1.3. Furthermore, the AIC for both GBM and nonlinear model are also included in Table 3.1.3.

Table 3.1.3 Estimated Parameters in Model 3.1.2 of Stock X

Interval Index	Unequal Interval GBM Model with Jumps			Unequal Interval Nonlinear Model 1 with Jumps			
	$\hat{\mu}$	$\hat{\sigma}$	AIC	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}$	AIC
1	0.006621	0.024521	53.16678	-0.06936	0.331231	0.024124	54.29912
2	0.013391	0.028731	52.68493	-0.28443	1.411569	0.023982	48.97729
3	0.009891	0.055846	66.41971	-0.48125	2.509755	0.042146	61.5785
4	0.001149	0.030116	76.19733	-0.41257	2.131512	0.026637	72.7967
5	0.009663	0.019472	36.22862	-0.24579	1.292223	0.016518	34.07066

6	-0.00134	0.026352	49.0021	-0.70337	3.70968	0.021988	45.88822
7	0.005189	0.038095	50.92597	-0.27218	1.436821	0.035407	51.11064
8	-0.00261	0.015574	80.28809	-0.13185	0.68589	0.014794	78.59092
9	0.008553	0.017281	92.29718	-0.03304	0.185145	0.017021	93.15485
10	0.00884	0.022777	99.00801	-0.20055	1.137652	0.020094	95.10469
11	-0.00053	0.014418	40.64565	-0.69145	3.933416	0.011311	36.36812
12	7.88E-05	0.015385	117.7922	-0.12929	0.733401	0.014874	117.2144
13	0.000596	0.015815	94.64219	-0.36491	2.091368	0.014159	90.2432
14	0.016518	0.033196	71.0436	-0.35994	2.137839	0.02067	61.21107
15	0.009311	0.013909	44.40171	-0.14129	0.856515	0.012957	44.72932
16	-0.00076	0.016247	90.84893	-0.25676	1.548757	0.015142	89.6624
17	0.004541	0.025039	60.94867	-0.41581	2.549001	0.019093	56.78929
18	-0.01211	0.036582	105.6407	-0.0507	0.290682	0.036318	107.3503
19	0.000358	0.028831	120.3625	-0.20871	1.226962	0.027295	119.5099
20	0.010446	0.022202	84.47223	-0.31116	1.870249	0.015605	74.16229
21	-0.00389	0.02337	57.79999	-0.10737	0.642615	0.022928	59.33579
22	-0.0026	0.019065	78.89382	-0.26636	1.579654	0.017477	77.70388
23	0.002168	0.015421	115.5262	-0.15501	0.931566	0.014359	113.3126
24	-0.00417	0.014351	64.01523	-0.2814	1.66841	0.012898	62.11432
25	0.003034	0.012816	66.33031	-0.52939	3.146627	0.01112	62.77867
26	0.003202	0.014617	99.57111	-0.20284	1.224103	0.013662	98.0609
27	0.007015	0.020763	116.315	-0.37375	2.313722	0.015966	106.9751
28	-0.00327	0.012388	57.10168	-0.88324	5.459753	0.005951	36.7315
29	1.37E-05	0.015938	96.23449	-0.116	0.715536	0.015467	96.88494
30	-0.00115	0.017478	101.1193	-0.18984	1.169378	0.016698	100.9772
31	-0.00045	0.014767	97.74304	-0.37187	2.273761	0.013167	94.06035
32	2.02E-05	0.01084	100.6316	-0.36305	2.234232	0.009826	96.81189
33	0.007892	0.011226	46.63383	-0.07094	0.445688	0.011061	48.32431
34	0.0025	0.008967	101.7473	-0.04038	0.255452	0.008882	103.1137
35	-0.00407	0.016643	80.90792	-0.77061	4.807935	0.008814	59.38289

36	0.003471	0.010888	113.8111	0.010153	-0.06017	0.01088	115.7619
37	0.008705	0.015797	75.73326	-0.11488	0.745197	0.015096	76.28284
38	0.010025	0.015063	59.36543	-0.13148	0.86973	0.014223	60.01012
39	-0.00136	0.022851	199.8702	-0.21082	1.373688	0.021244	196.409

Figures 3.1.4 - 3.1.6 are the plots of predicted value of Unequal Interval Nonlinear Model 3.1.2 of stock X with observations ranging from 1 to 300, 300 to 600 and 600 to 848 respectively.

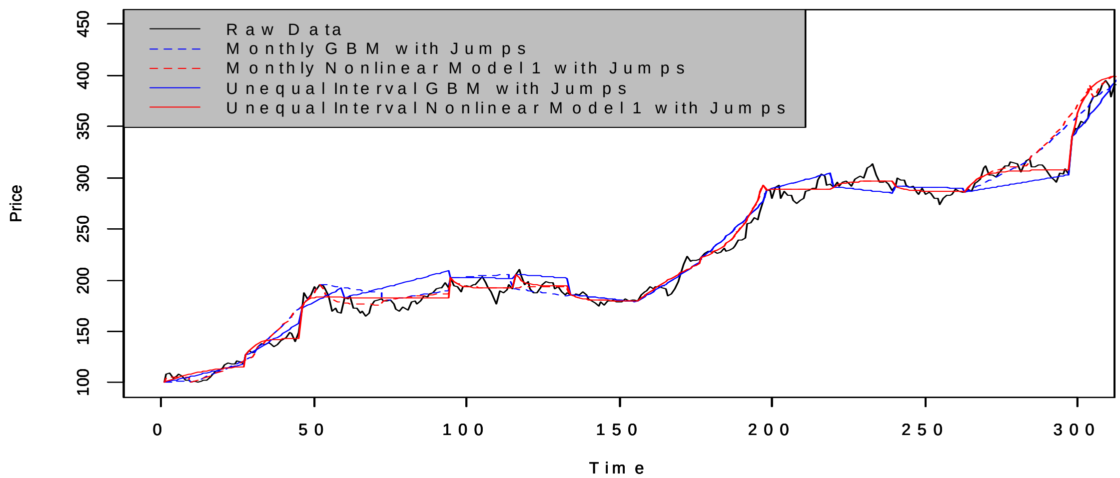


Figure 3.1.4 Comparison of Model 2.4.3, 2.4.4, 3.1.1 with Model 3.1.2 of Stock X (Observations 1-300)

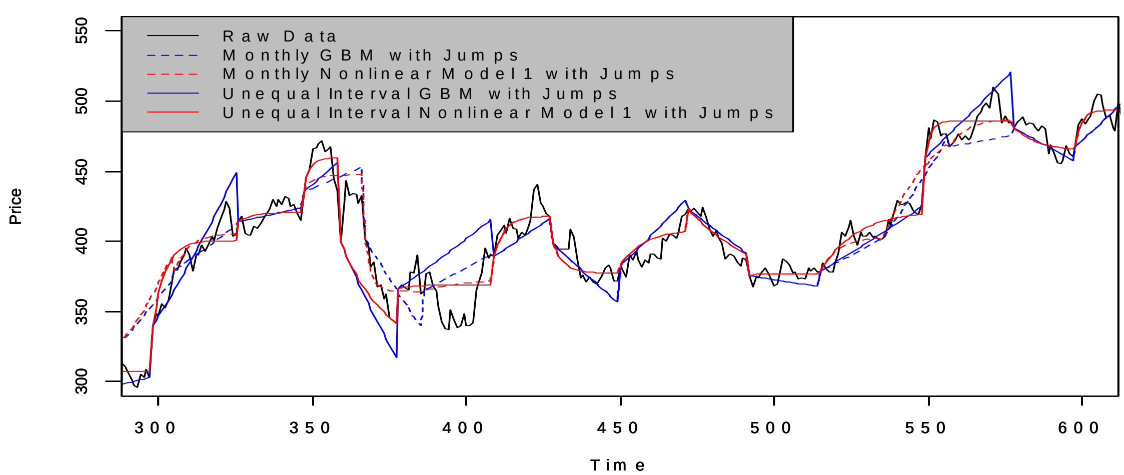


Figure 3.1.5 Comparison of Model 2.4.3, 2.4.4, 3.1.1 with Model 3.1.2 of Stock X (Observations 300-600)

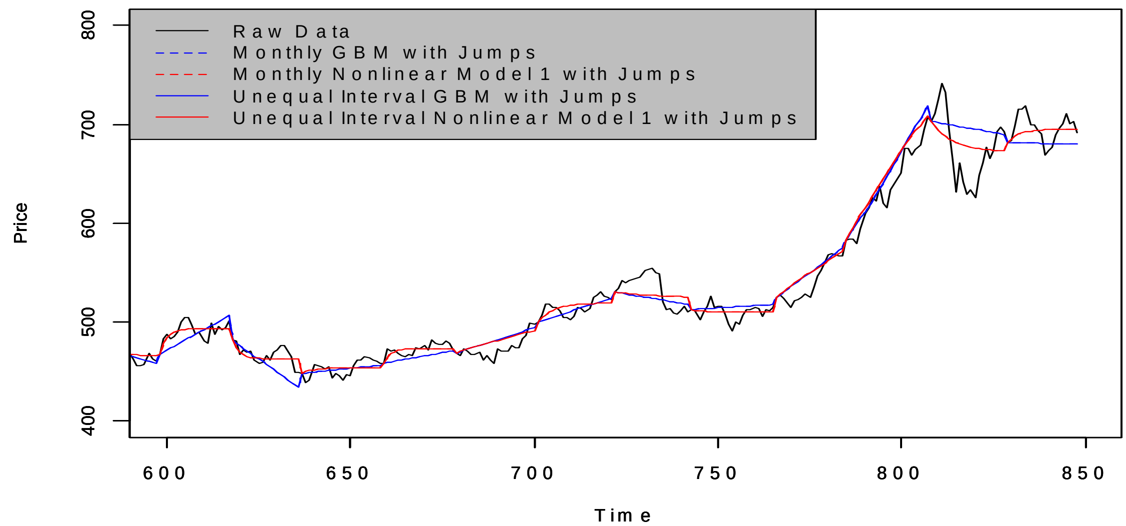


Figure 3.1.6 Comparison of Model 2.4.3, 2.4.4, 3.1.1 with Model 3.1.2 of Stock X (Observations 600-848)

Table 3.1.4 shows the overall basic statistics of monthly GBM model 2.4.3 with jumps, Monthly Nonlinear Model 3.1.1 with Jumps, Unequal Interval GBM model 2.4.4 with Jumps and Unequal Interval Nonlinear Model 3.1.2 with jumps.

Table 3.1.4 Basic Statistics of Models 2.4.3, 3.1.1, 2.4.4 and 3.1.2 of Stock X

Model	$\overline{r}$	S_r^2	S_r	No. of Intervals	No. of Parameters
Monthly GBM with Jumps	-1.242020	207.264	14.39667	41	122
Monthly Nonlinear Model 1 with Jumps	-1.928296	141.1754	11.88173	41	163
Unequal Interval GBM with Jumps	1.962899	258.1040	16.06562	39	116
Unequal Interval Nonlinear Model 1 with Jumps	0.5315015	131.2354	11.4558	39	155

Remark 3.1.2: From Table 3.1.4, we note that, the Unequal Interval Nonlinear Model 3.1.2 with Jumps has least mean and variance of the residual error. From the Table 3.1.3 and Figures 3.1.4 – 3.1.6 we conclude that on some intervals, the GBM model is better than Nonlinear Model 3.1.2. In addition the GBM model is better than Nonlinear Model 3.1.2 in terms of AIC. For example, on the 7th, 9th, 15th, 21st, 29th, 33rd, … intervals, the GBM model has less AIC than Nonlinear Model 3.1.2. There are 12 out of 39 intervals (31%), on which the GBM model has less AIC than Nonlinear Model 3.1.2.

3.2 Stochastic Nonlinear Dynamic Model 2

This nonlinear stochastic model 2 [26] is described by the following $It\hat{o}-Doob$ differential equation

$$dS_t = (\alpha S_t + \beta S_t^N + \frac{N}{2}\sigma^2 S_t^{2N-1})dt + \sigma S_t^N dW_t \qquad (3.2.1)$$

where, α, β, N and σ are parameters; moreover $0 \le N \le 1.2, N \ne 1$, and W_t is Brownian motion. It is easy to verify that rate functions in (3.2.1) satisfies the conditions for existence and uniqueness of solution [23,28]. We note that the volatility and drift functions are nonlinear functions of S_t. In order to derive the regression equation, we use the following transformation $V_t = \frac{S_t^{1-N}}{1-N}$ and apply $It\hat{o}-Doob$ differential formula to obtain

$$\begin{aligned}
dV_t &= \frac{\partial}{\partial S_t}(\frac{S_t^{1-N}}{1-N})dS_t + \frac{1}{2}(\frac{\partial^2}{\partial S_t^2}(\frac{S_t^{1-N}}{1-N}))(dS_t)^2 \\
&= S_t^{-N} dS_t + \frac{1}{2}(S_t^{-N})'(dS_t)^2 \\
&= S_t^{-N}(\alpha S_t + \beta S_t^N + \frac{N}{2}\sigma^2 S_t^{2N-1})dt + \sigma S_t^N dW_t) - \frac{N}{2} S_t^{-N-1}\sigma^2 S_t^{2N}(dW_t)^2 \\
&= (\alpha S_t^{1-N} + \beta)dt + \sigma dW_t
\end{aligned}$$

Then,

$$dV_t = (\alpha(1-N)V_t + \beta)dt + \sigma dW_t \qquad (3.2.2)$$

Again by using Euler type discretization process [24], stochastic differential equation (3.2.2) can be reduced to

$$V_t - V_{t-1} = (\alpha(1-N)V_{t-1} + \beta)\Delta t + \sigma(W_t - W_{t-1}) \tag{3.2.3}$$

From $\varepsilon_t = W_t - W_{t-1}$ and $\Delta t = 1$, equation (3.2.3) can be rewritten as

$$V_t = ((1+\alpha(1-N))V_{t-1} + \beta) + \sigma\varepsilon_t \tag{3.2.4}$$

where α, β and σ are as defined in (3.2.1). For given N, by applying the least square regression method [35] and using above cited data set, these parameters can be estimated, analogously. N is estimated by the value, under which the model has least variance of residual error.

Nonlinear Stochastic Model 3.2.1 (Monthly Nonlinear Model 2 with Jumps): Let $[0,t_1),[t_1,t_2),[t_2,t_3),\ [t_3,t_4),[t_4,t_5)...\ [t_{m-1},t_m)$ be the m monthly time intervals as defined in stochastic model 3.1.1. The nonlinear stochastic model 3.2.1 with jumps takes the following form of nonlinear $It\hat{o} - Doob$ type stochastic differential equation:

$$dS_t^{M_i} = (\alpha^{M_i} S_t^{M_i} + \beta^{M_i} (S_t^{M_i})^{N^{M_i}} + \frac{N^{M_i}}{2}(\sigma^{M_i})^2 (S_t^{M_i})^{2N^{M_i}-1})dt + \sigma^{M_i} (S_t^{M_i})^{N^{M_i}} dW_t\,,$$

$$S_0 = S_0,\ \text{if}\ t_{i-1} \le t < t_i,\ i = 1,...,m. \tag{3.2.5}$$

α^{M_i}, β^{M_i}, and $\sigma^{M_i}, i = 1,...,m$ are parameters. These parameters are estimated as described above. As before, following definition [16, 26, 27], the solution of (3.2.5) takes the form:

$$S(t) = \begin{cases} S_1(t,t_0,S_0), & t_0 \le t < t_1 & \\ \phi_1 S_2(t,t_1,S_1), & t_1 \le t < t_2, & S_1 = \lim\limits_{t\to t_1^-} S_1(t,t_0,S_0) \\ \dots & \dots & \dots \\ \phi_{m-1} S_m(t,t_{m-1},S_{m-1}), & t_{m-1} \le t < t_m, & S_{m-1} = \lim\limits_{t\to t_{m-1}^-} S_{m-1}(t,t_{m-2},S_{n-2}) \end{cases} \tag{3.2.6}$$

Again, S_0 is the initial value of the stock price process. $\phi_1, \phi_{2,} ..., \phi_{m-1}$ are jumps, and can.

$$\hat{\phi}_1 = \frac{S_{t_1}}{\lim\limits_{t\to t_1} \hat{S}_1}, \hat{\phi}_2 = \frac{S_{t_2}}{\lim\limits_{t\to t_2} \hat{S}_2}, ..., \hat{\phi}_{m-1} = \frac{S_{t_{m-1}}}{\lim\limits_{t\to t_{m-1}} \hat{S}_{m-1}}.$$

Table 3.2.1 gives estimated parameters by applying Monthly Nonlinear Stochastic Model 3.2.1 of stock X.

Table 3.2.1 Estimated Parameters of Model 3.2.1 of Stock X

Interval Index	Monthly Nonlinear Model 2 with Jumps			
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}$	$\hat{N}$
1	-0.93097	98.58582	2.685764	0
2	0.015167	-0.40094	2.562583	0
3	0.006386	0.702974	2.43304	0.2
4	-0.26424	46.59828	6.056676	0
5	-0.04732	9.06878	3.426131	0
6	3.111272	-1.65455	0.014153	1.12
7	-0.32461	63.03162	5.104841	0
8	-0.23319	41.98315	2.08146	0
9	0.162624	-0.05328	0.007219	1.2
10	-0.43324	0.148857	0.005128	1.2
11	-0.32678	94.48073	5.651904	0
12	-3.1569	4.702766	0.02342	0.93
13	-0.19667	56.36633	3.738925	0
14	0.726705	-0.2303	0.004546	1.2
15	0.227674	-0.4653	0.06296	0.87
16	0.934944	-0.28125	0.005202	1.2
17	1.106195	-0.33042	0.003641	1.2
18	-0.30599	136.999	13.06886	0
19	1.943516	-0.59744	0.008646	1.2
20	-0.06734	25.27551	9.241028	0
21	1.339686	-0.4007	0.005306	1.2
22	1.223408	-0.37344	0.004986	1.2
23	-0.11434	46.91534	6.455683	0
24	-0.04108	15.00859	4.799298	0
25	-0.55537	209.269	4.546397	0
26	0.805277	-0.24258	0.004332	1.2
27	0.291471	-0.08414	0.006396	1.2

28	0.770865	-0.22354	0.00398	1.2
29	-0.11439	53.08718	5.661491	0
30	2.085035	-0.60321	0.003964	1.2
31	1.903402	-0.55777	0.003924	1.2
32	-0.27964	126.9381	5.734823	0
33	-0.39253	185.8715	4.464326	0
34	-0.00631	4.182385	5.648348	0
35	-0.20829	108.206	4.736808	0
36	0.348545	-0.09964	0.004068	1.2
37	1.709708	-0.49133	0.003289	1.2
38	0.107688	-0.02919	0.002777	1.2
39	-0.01282	14.14102	9.792642	0
40	-0.15131	101.7773	17.81769	0
41	1.431732	-0.38677	0.003751	1.2

Figures 3.2.1- 3.2.3 are the plots of predicted value of Monthly Nonlinear Model 3.2.1 of stock X with observation ranging from 1 to 300, 300 to 600 and 600 to 848 respectively.

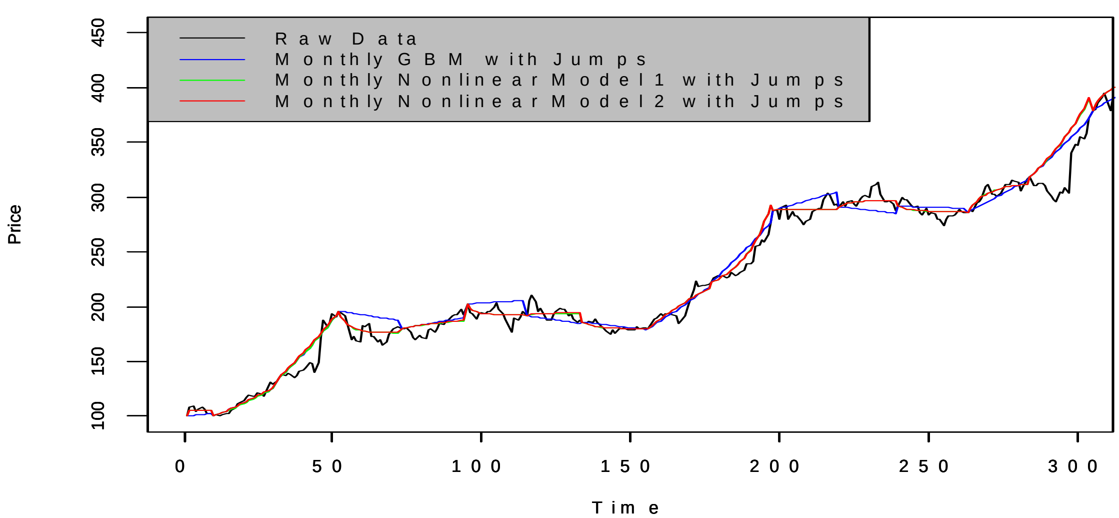


Figure 3.2.1 Comparison of Model 2.4.3, 3.1.1 with Model 3.2.1 of Stock X
(Observations 1-300)

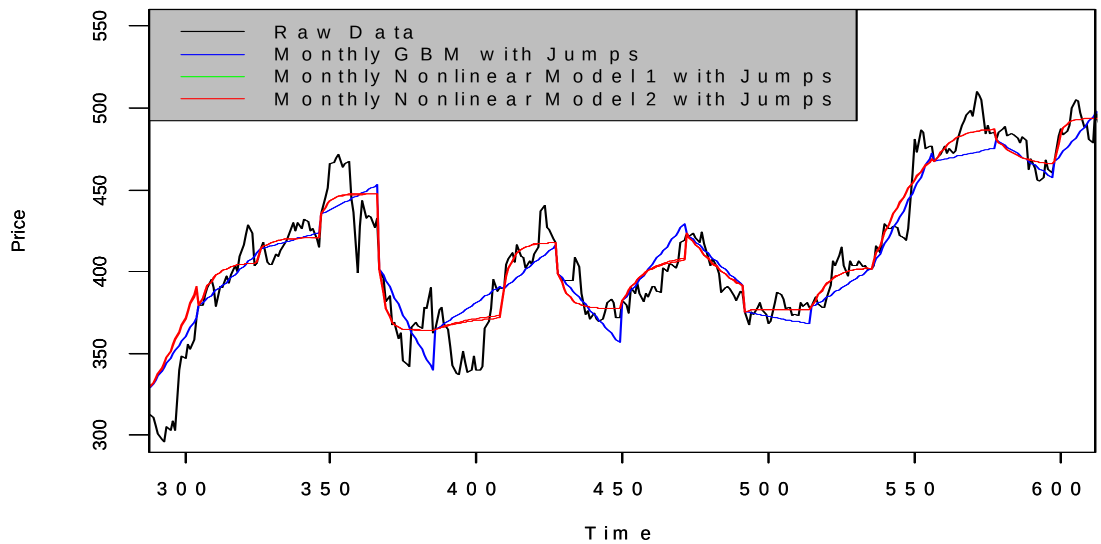


Figure 3.2.2 Comparison of Model 2.4.3, 3.1.1 with Model 3.2.1 of Stock X

(Observations 300-600)

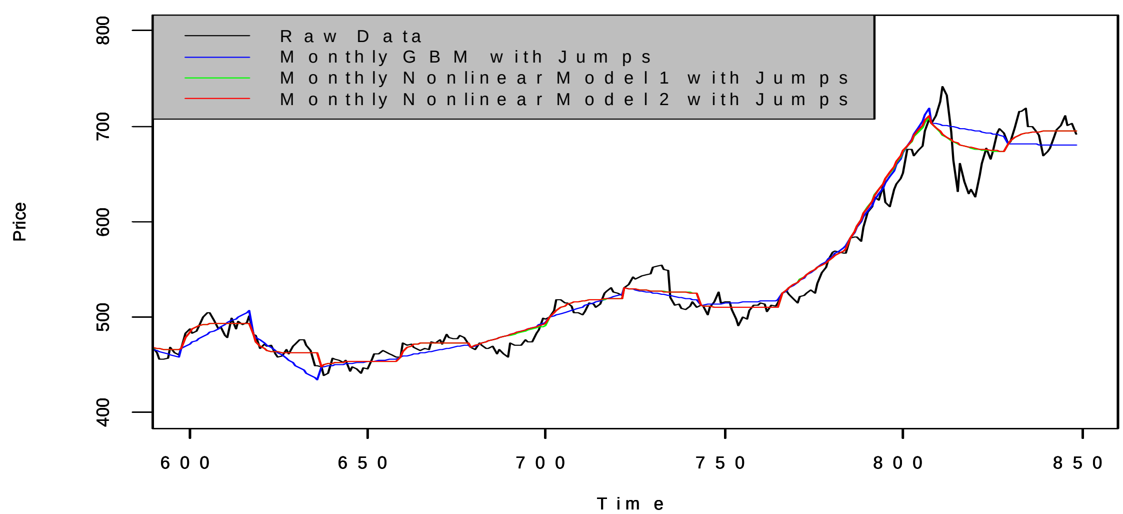


Figure 3.2.3 Comparison of Model 2.4.3, 3.1.1 with Model 3.2.1 of Stock X

(Observations 600-848)

Table 3.2.2 shows the overall basic statistics of Monthly GBM Model 2.4.3, Monthly Nonlinear Model 3.1.1 and Monthly Nonlinear Model 3.2.1 with Jumps of Stock X.

Table 3.2.2 Basic Statistics of Models 2.4.3, 3.1.1 and 3.2.1 of Stock X

Model	$\overline{r}$	S_r^2	S_r	No. of Intervals	No. of Parameters
Monthly GBM with Jumps	-1.242020	207.264	14.39667	41	122
Monthly Nonlinear Model 1 with Jumps	-1.928296	141.1754	11.88173	41	163
Monthly Nonlinear Model 2 with Jumps	-2.090806	143.2248	11.96765	41	204

Remark 3.2.1: From Table 3.2.2, we observe that under the same data partition process, Monthly Nonlinear Model 3.2.1 with Jumps has less variance than Monthly GBM with Jumps. Overall, the Monthly Nonlinear Model 3.1.1 with Jumps has less variance of the residual error than Monthly GBM Model and Monthly Nonlinear Model 3.2.1 with Jumps.

Nonlinear Stochastic Model 3.2.2 (Unequal Interval Nonlinear Model 2 with Jumps): Let $[0,t_1),[t_1,t_2),[t_2,t_3),\ [t_3,t_4),[t_4,t_5)...\ [t_{n-1},t_n)$ be the n time intervals defined in stochastic model 3.1.2. The nonlinear stochastic model 3.2.2 with jumps takes the following form:

$$dS_t^{I_i} = (\alpha^{I_i} S_t^{I_i} + \beta^{I_i}(S_t^{I_i})^{N^{I_i}} + \frac{N^{I_i}}{2}(\sigma^{I_i})^2 (S_t^{I_i})^{2N^{I_i}-1})dt + \sigma^{I_i}(S_t^{I_i})^{N^{I_i}} dW_t,$$

$$S_0 = S_0,\ \text{if}\ t_{i-1} \le t < t_i,\ i = 1,...,n. \tag{3.2.7}$$

α^{I_i}, β^{I_i}, and σ^{I_i} , $i = 1,...,n$, are parameters and can be estimated as described before.
By following definition [16, 26, 27], the solution of (3.2.7) takes the form

$$S(t) = \begin{cases} S_1(t,t_0,S_0), & t_0 \le t < t_1 \\ \phi_1 S_2(t,t_1,S_1), & t_1 \le t < t_2, \quad S_1 = \lim_{t \to t_1^-} S_1(t,t_0,S_0) \\ \cdots & \cdots \qquad\qquad \cdots \\ \phi_{n-1} S_m(t,t_{n-1},S_{n-1}), & t_{n-1} \le t < t_n, \quad S_{n-1} = \lim_{t \to t_{n-1}^-} S_{n-1}(t,t_{n-2},S_{n-2}) \end{cases} \tag{3.2.8}$$

Here, S_0 is the initial value of the stock price process. $\phi_1,\phi_2,...,\phi_{n-1}$ are jumps and can be estimated as $\hat{\phi}_1 = \frac{S_{t_1}}{\lim_{t\to t_1^-} \hat{S}_1}, \hat{\phi}_2 = \frac{S_{t_2}}{\lim_{t\to t_2^-} \hat{S}_2}, ...\hat{\phi}_{n-1} = \frac{S_{t_{n-1}}}{\lim_{t\to t_{n-1}^-} \hat{S}_{n-1}}$.

Table 3.2.3 gives estimated parameters with regard to Unequal Interval Nonlinear Stochastic Model 3.2.2 of stock X.

Table 3.2.3 Estimated Parameters of Model 3.2.2 of Stock X

Interval Index	Unequal Interval Nonlinear Model 2 with Jumps			
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}$	$\hat{N}$
1	-0.06298	7.518172	2.580372	0
2	-0.2657	38.08205	3.391367	0
3	-0.44948	82.83006	7.48971	0
4	-0.41314	72.45667	4.66966	0
5	-0.23482	45.135	3.059264	0
6	-0.75107	106.8789	3.141901	0.06
7	-0.27195	53.40573	6.840905	0
8	0.656408	-0.23191	0.005193	1.2
9	0.172972	-0.05645	0.005866	1.2
10	-0.18794	54.75371	5.654817	0
11	-0.68875	203.5307	3.345713	0
12	0.646214	-0.20781	0.004773	1.2
13	-0.36426	112.3249	4.348781	0
14	-0.32689	124.4562	7.29626	0
15	-0.13146	56.60131	5.240336	0
16	-0.25507	106.2544	6.30183	0
17	-0.40544	186.3563	8.661997	0
18	0.239471	-0.07623	0.010924	1.2
19	1.039288	-0.32073	0.008402	1.2
20	-0.29476	120.2958	6.163061	0
21	0.531576	-0.16061	0.00685	1.2
22	-0.26539	99.89832	6.775433	0
23	-0.14965	60.98254	5.730283	0
24	1.402875	-0.42858	0.003933	1.2
25	-0.52427	199.9783	4.240122	0

26	-0.1963	82.03641	5.598122	0
27	-0.34739	169.7543	7.612169	0
28	4.412648	-1.28169	0.001728	1.2
29	-0.1149	54.87902	7.372088	0
30	0.945964	-0.27597	0.004852	1.2
31	-0.37018	167.4691	5.936408	0
32	-0.36119	170.0127	4.617445	0
33	-0.05807	31.62257	5.295368	0
34	-0.03837	21.49953	4.633074	0
35	3.847553	-1.10474	0.002529	1.2
36	0.01336	-5.18139	5.645615	0
37	0.58191	-0.15902	0.00418	1.2
38	-0.12128	90.81585	9.907542	0
39	1.050197	-0.2853	0.005779	1.2

Figures 3.2.4, 3.2.5 and 3.2.6 are the plots of predicted value of Unequal Interval Nonlinear Model 3.2.2 of stock X with Jumps with observations ranging from 1 to 300, 300 to 600 and 600 to 848 respectively.

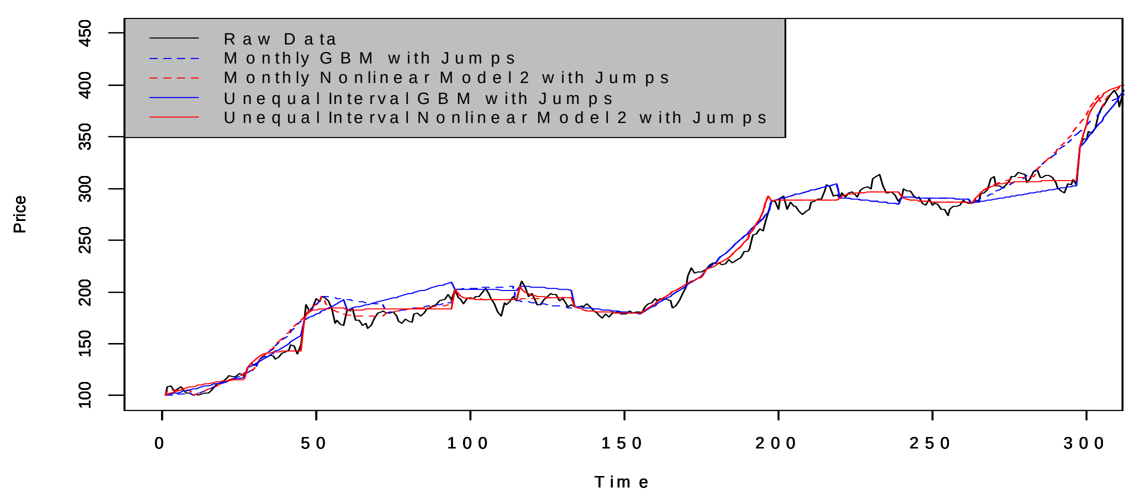


Figure 3.2.4 Comparison of Model 2.4.3, 2.4.4, 3.2.1 with Model 3.2.2 of Stock X (Observations 1-300)

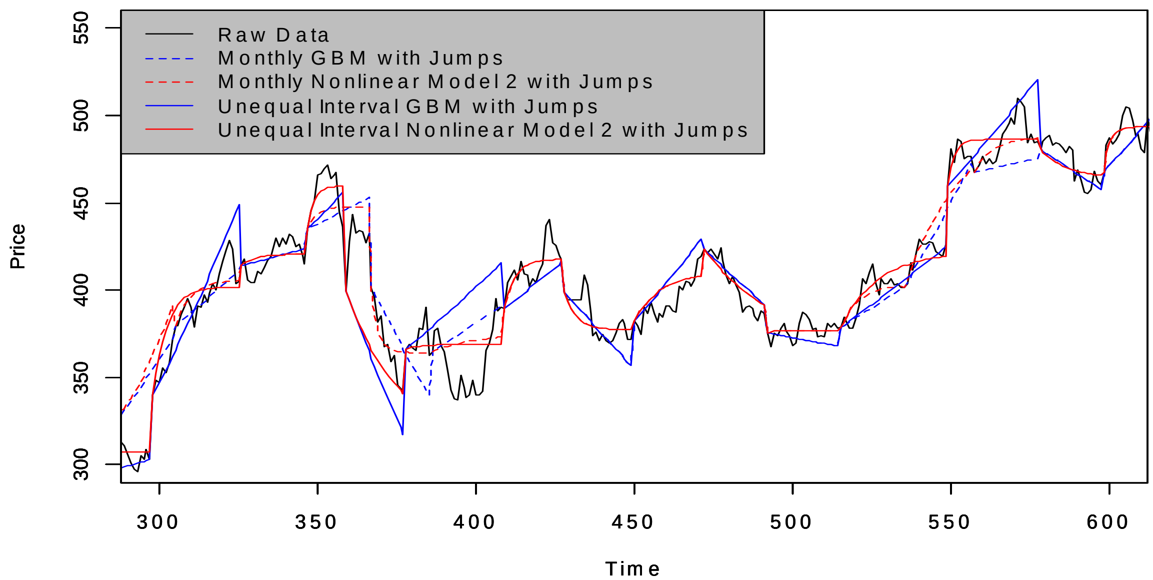


Figure 3.2.5 Comparison of Model 2.4.3, 2.4.4, 3.2.1 with Model 3.2.2 of Stock X

(Observations 300-600)

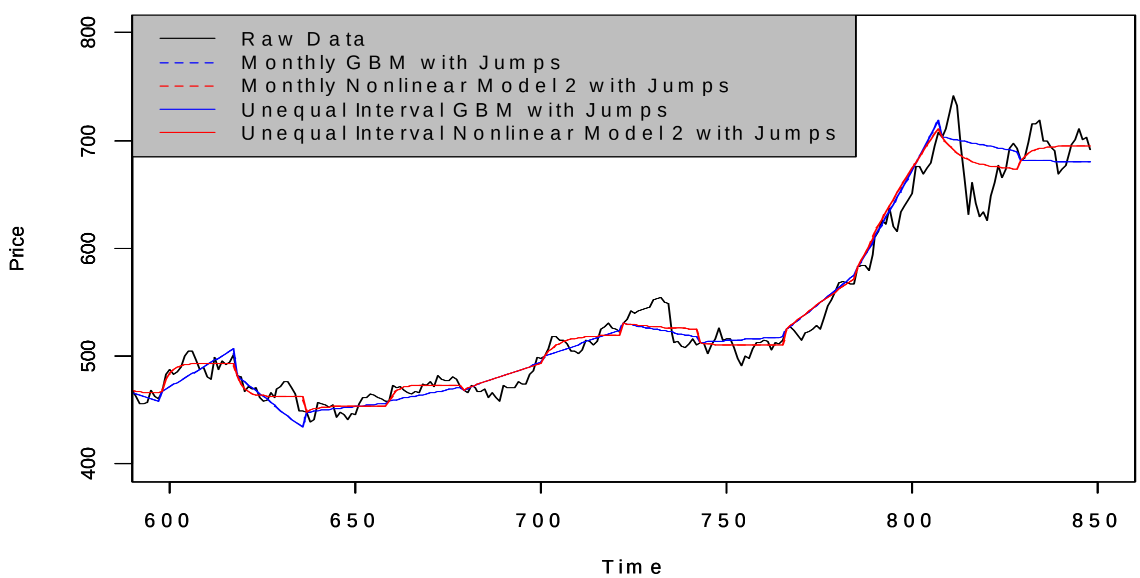


Figure 3.2.6 Comparison of Model 2.4.3, 2.4.4, 3.2.1 with Model 3.2.2 of Stock X

(Observations 600-848)

Table 3.2.4 shows the overall basic statistics of Monthly GBM model 2.4.3, Nonlinear Model 3.1.1 and 3.1.2 with Jumps, Unequal GBM model 2.4.4, and Unequal Nonlinear Model 3.2.1 and 3.2.2 with Jumps.

Table 3.2.4 Basic Statistics of Models 2.4.3, 3.1.1, 3.1.2, 2.4.4, 3.2.1 and 3.2.2 of Stock X

Model	$\overline{r}$	S_r^2	S_r	No. of Intervals	No. of Parameters
Monthly GBM with Jumps	-1.242020	207.264	14.39667	41	122
Monthly Nonlinear Model 1 with Jumps	-1.928296	141.1754	11.88173	41	163
Monthly Nonlinear Model 2 with Jumps	-2.090806	143.2248	11.96765	41	204
Unequal GBM with Jumps	-1.962899	258.1040	16.06562	39	116
Unequal Nonlinear Model 1 with Jumps	-0.5315015	131.2354	11.4558	39	155
Unequal Nonlinear Model 2 with Jumps	-0.6097021	131.3068	11.45892	39	194

Remark 3.2.2: From Table 3.2.4 we observe that under the same data partition processes, Nonlinear Models 3.1.1 and 3.2.1 have less variance of the residual error than Monthly GBM model, and Nonlinear Models 3.1.2 and 3.2.2 also have less variance of the residual than Unequal GBM model. Overall, the nonlinear models 3.1.2 and 3.2.2 under the Unequal Interval data partition process have less variance of the residual error than the nonlinear models 3.1.1 and 3.2.1under the monthly data partition process.

3.3 Stochastic Nonlinear Dynamic Model 3

This nonlinear stochastic model 3 [26] is described by the following $It\hat{o}-Doob$ differential equation

$$dS_t = (\alpha S_t + \beta S_t^2 + \sigma^2 S_t)dt + \sigma S_t dW_t \tag{3.3.1}$$

where, α, β and σ are parameters and W_t is Brownian motion. It is easy to check that rate functions in (3.3.1) satisfies the conditions for existence and uniqueness of solution [23,28]. In order to derive a regression equation, we use the following transformation $V_t = \frac{-1}{S_t}$ and applying $It\hat{o} - Doob$ differential formula to obtain

$$
\begin{aligned}
dV_t &= \frac{\partial}{\partial S_t}(\frac{-1}{S_t})dS_t + \frac{1}{2}(\frac{\partial^2}{\partial S_t{}^2}(\frac{-1}{S_t}))(dS_t)^2 \\
&= S_t^{-2}dS_t + \frac{1}{2}(S_t^{-2})'(dS_t)^2 \\
&= S_t^{-2}((\alpha S_t + \beta S_t^2 + \sigma^2 S_t)dt + \sigma S_t dW_t) - \frac{1}{2}S_t^{-3}\sigma^2 S_t^2 (dW_t)^2 \\
&= (\alpha S_t^{-1} + \beta)dt + \sigma S_t^{-1} dW_t
\end{aligned}
$$

Then,

$$dV_t = (-\alpha V_t + \beta)dt - \sigma V_t dW_t \tag{3.3.2}$$

Again, the Euler type discretized version of (3.3.2) is as follows

$$V_t - V_{t-1} = (-\alpha V_{t-1} + \beta)\Delta t - \sigma V_{t-1}(W_t - W_{t-1}) \,. \tag{3.3.3}$$

From the definition of V, we note that $y_t = \frac{V_t - V_{t-1}}{V_{t-1}} = \frac{S_{t-1}}{S_t} - 1$, $\varepsilon_t = W_t - W_{t-1}$, and $\Delta t = 1$.

With this notation, equation (3.3.3) can be rewritten as

$$y_t = (-\alpha + \beta \frac{1}{V_{t-1}}) - \sigma \varepsilon_t \,. \tag{3.3.4}$$

Then, parameters α, β and σ can be estimated using least square method [35].

Nonlinear Stochastic Model 3.3.1 (Monthly Nonlinear Model 3 with Jumps): Let $[0,t_1),[t_1,t_2),[t_2,t_3), [t_3,t_4),[t_4,t_5)...[t_{m-1},t_m)$ be the m monthly time intervals as defined in stochastic model 3.1.1. The nonlinear stochastic model 3.3.1 takes the form of following nonlinear $It\hat{o} - Doob$ type stochastic differential equation:

$$
\begin{aligned}
&dS_t^{M_i} = (\alpha^{M_i} S_t^{M_i} + \beta^{M_i} (S_t^{M_i})^2 + (\sigma^{M_i})^2 S_t^{M_i})dt + \sigma^{M_i} S_t^{M_i} dW_t \,, \\
&S_0 = S_0 \text{, if } t_{i-1} \le t < t_i \,,\ i = 1,...,m \,.
\end{aligned} \tag{3.3.5}
$$

α^{M_i}, β^{M_i}, and $\sigma^{M_i}, i = 1,...,m$ are parameters and are estimated as described above. Thus, the solution of equation (3.3.5) is given by

$$S(t) = \begin{cases} S_1(t, t_0, S_0), & t_0 \le t < t_1 \\ \phi_1 S_2(t, t_1, S_1), & t_1 \le t < t_2, \qquad S_1 = \lim_{t \to t_1^-} S_1(t, t_0, S_0) \\ \quad ... & \quad ... \qquad\qquad ... \\ \phi_{m-1} S_m(t, t_{m-1}, S_{m-1}), & t_{m-1} \le t < t_m, \quad S_{m-1} = \lim_{t \to t_{m-1}^-} S_{m-1}(t, t_{m-2}, S_{n-2}) \end{cases} \tag{3.3.6}$$

where S_0 is the initial value of the stock price process. $\phi_1, \phi_{2,}..., \phi_{m-1}$ are jumps. These jumps are estimated as $\hat{\phi}_1 = \dfrac{S_{t_1}}{\lim\limits_{t \to t_1} \hat{S}_1}, \hat{\phi}_2 = \dfrac{S_{t_2}}{\lim\limits_{t \to t_2} \hat{S}_2}, ..., \hat{\phi}_{m-1} = \dfrac{S_{t_{m-1}}}{\lim\limits_{t \to t_{m-1}} \hat{S}_{m-1}}$.

The estimated parameters of Monthly Nonlinear Stochastic Model 3.3.1 of stock X are recorded in Table 3.3.1.

Table 3.3.1 Estimated Parameters of Model 3.3.1 of Stock X

Interval Index	Monthly Nonlinear Model 3 with Jumps		
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}$
1	0.922514	-0.0087161	0.025344
2	0.004041	6.17E-05	0.021409
3	0.023693	-4.18E-05	0.039823
4	0.261409	-0.0014855	0.034493
5	0.050772	-0.0002666	0.018895
6	0.37251	-0.0019314	0.026447
7	0.318094	-0.0016404	0.025629
8	0.229969	-0.0012777	0.011528
9	0.041487	-0.0001638	0.020631
10	-0.07527	0.00036332	0.015148
11	0.330384	-0.0011435	0.019702

12	0.219936	-0.0007408	0.015711
13	0.196132	-0.0006846	0.013041
14	0.153618	-0.0004918	0.014218
15	-0.0255	0.00010249	0.028538
16	0.190213	-0.0004688	0.017376
17	0.222578	-0.0005293	0.012198
18	0.312192	-0.0006983	0.031076
19	0.381663	-0.001048	0.028307
20	0.071652	-0.0001924	0.025384
21	0.269717	-0.0006458	0.017525
22	0.246151	-0.0006521	0.016515
23	0.12528	-0.0003065	0.016356
24	0.03558	-9.80E-05	0.011957
25	0.552865	-0.0014675	0.012074
26	0.167816	-0.0004166	0.014242
27	0.06565	-0.0001331	0.021147
28	0.158273	-0.0003247	0.013793
29	0.111985	-0.0002415	0.01214
30	0.422469	-0.0008563	0.013585
31	0.381775	-0.0008249	0.013505
32	0.278473	-0.0006137	0.012737
33	0.392775	-0.0008295	0.009392
34	0.01369	-2.39E-05	0.011873
35	0.210991	-0.0004062	0.009164
36	0.071916	-0.0001373	0.014521
37	0.345019	-0.0006763	0.011447
38	0.025629	-3.84E-05	0.009695
39	0.033435	-3.82E-05	0.01535
40	0.150713	-0.0002246	0.026663
41	0.287854	-0.0004142	0.013927

Figures 3.3.1, 3.3.2 and 3.3.3 are the plots of predicted value of Monthly Nonlinear Model 3.3.1 of stock X with observations ranging from 1 to 300, 300 to 600 and 600 to 848 respectively.

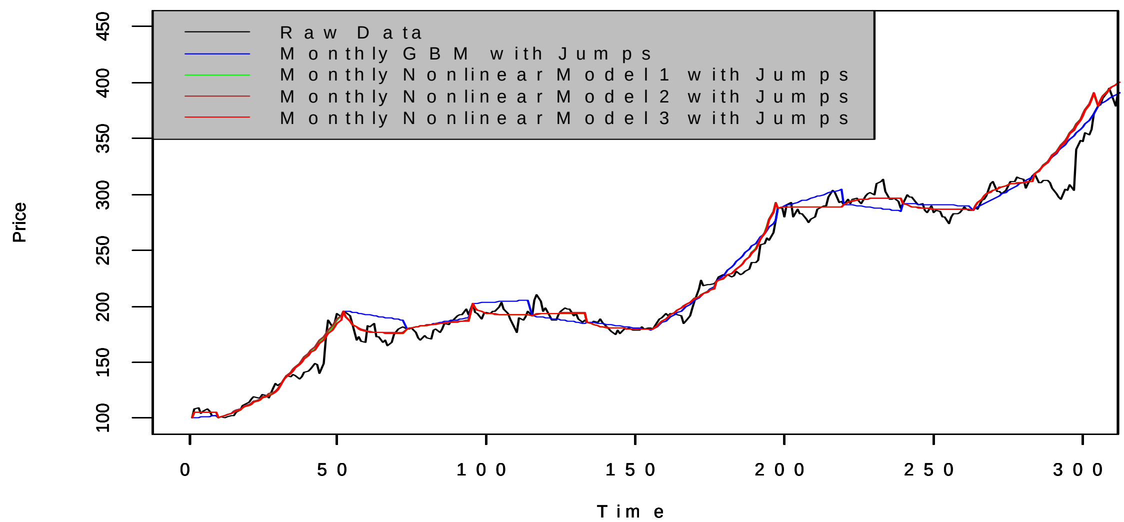


Figure 3.3.1 Comparison of Model 2.4.3, 3.1.1, 3.2.1 with Model 3.3.1 of Stock X (Observations 1-300)

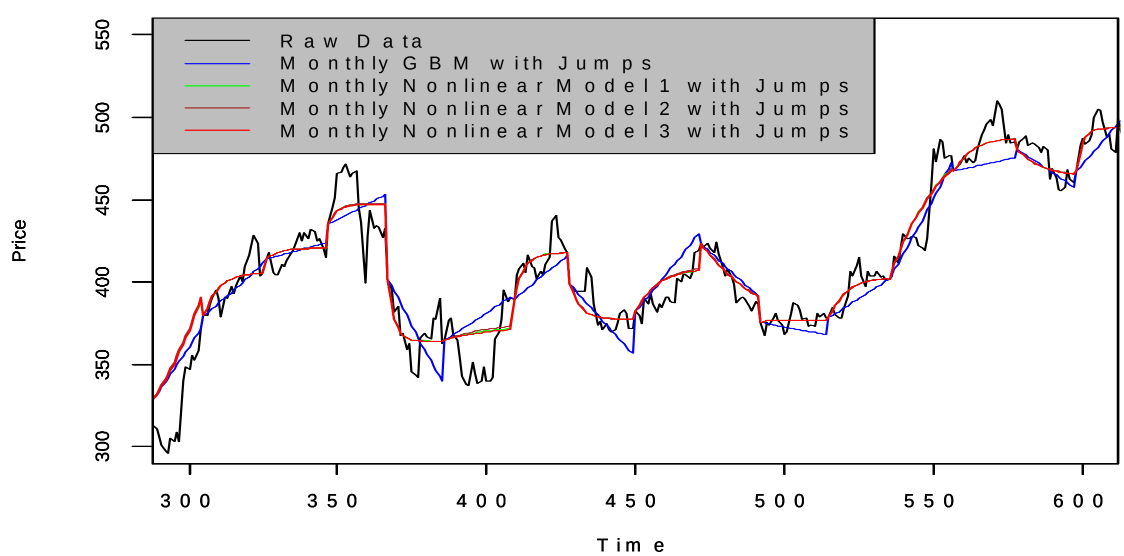


Figure 3.3.2 Comparison of Model 2.4.3, 3.1.1, 3.2.1 with Model 3.3.1 of Stock X (Observations 300-600)

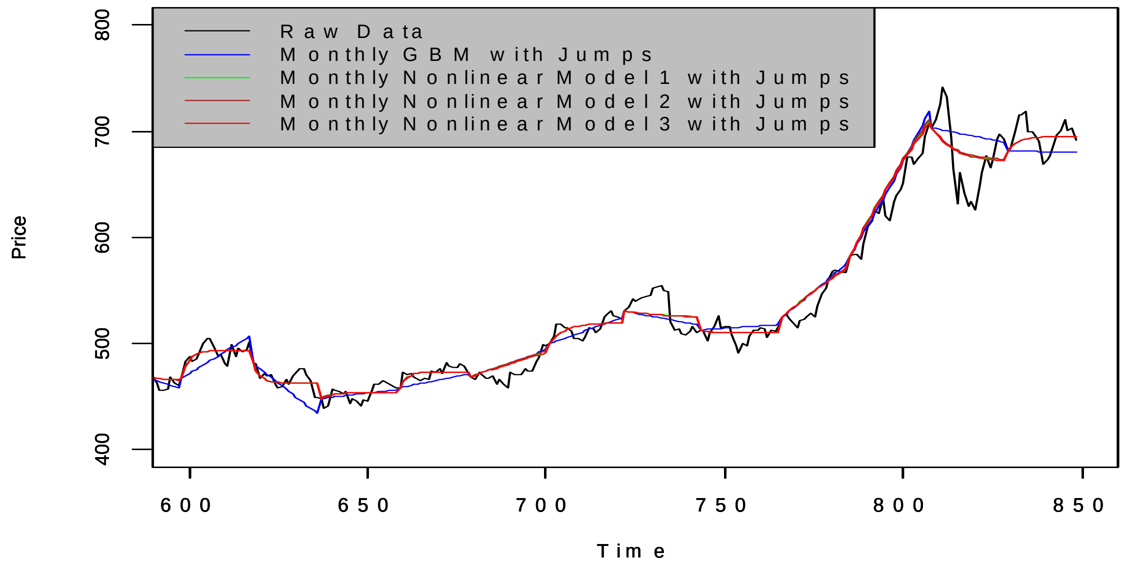


Figure 3.3.3 Comparison of Model 2.4.3, 3.1.1, 3.2.1 with Model 3.3.1 of Stock X (Observations 600-848)

Table 3.3.2 shows the overall basic statistics of Monthly GBM Model 2.4.3, Monthly Nonlinear Model 3.1.1, 3.2.1, and Monthly Nonlinear Model 3.3.1 with Jumps.

Table 3.3.2 Basic Statistics of Models 2.4.3, 3.1.1, 3.2.1 and 3.3.1 of Stock X

Model	$\overline{r}$	S_r^2	S_r	No. of Intervals	No. of Parameters
Monthly GBM with Jumps	-1.242020	207.264	14.39667	41	122
Monthly Nonlinear Model 1 with Jumps	-1.928296	141.1754	11.88173	41	163
Monthly Nonlinear Model 2 with Jumps	-2.090806	143.2248	11.96765	41	204
Monthly Nonlinear Model 3 with Jumps	-1.731151	139.2792	11.80166	41	163

Remark 3.3.1: From Table 3.3.2 we remark that overall, the Monthly Nonlinear Model 3.3.1 with Jumps has least variance of the residual error than Monthly GBM Model 2.4.3, Nonlinear Model 3.1.1 and 3.2.1.

Nonlinear Stochastic Model 3.3.2 (Unequal Interval Nonlinear Model 3 with Jumps): Let $[0,t_1),[t_1,t_2),[t_2,t_3),\ [t_3,t_4),[t_4,t_5)...\ [t_{n-1},t_n)$ be the n time intervals as defined stochastic model 3.1.2. The nonlinear stochastic model 3.3.2 is described by:

$$dS_t^{M_i} = (\alpha^{M_i} S_t^{M_i} + \beta^{M_i} (S_t^{I_i})^2 + (\sigma^{I_i})^2 S_t^{I_i})dt + \sigma^{I_i} S_t^{I_i} dW_t,$$
$$S_0 = S_0, \text{ if } t_{i-1} \le t < t_i,\ i = 1,...,n. \tag{3.3.7}$$

where, α^{I_i} , β^{I_i} , and σ^{I_i} , $i = 1,...,n$, are parameters as defined and estimated. The solution of equation (3.3.7) is represented by:

$$S(t) = \begin{cases} S_1(t,t_0,S_0), & t_0 \le t < t_1 \\ \phi_1 S_2(t,t_1,S_1), & t_1 \le t < t_2, \qquad S_1 = \lim\limits_{t \to t_1^-} S_1(t,t_0,S_0) \\ \quad ... & \quad ... \qquad\qquad ... \\ \phi_{n-1} S_m(t,t_{n-1},S_{n-1}), & t_{n-1} \le t < t_n, \quad S_{n-1} = \lim\limits_{t \to t_{n-1}^-} S_{n-1}(t,t_{n-2},S_{n-2}) \end{cases} \tag{3.3.8}$$

S_0 is the initial value of the stock price process. $\phi_1, \phi_{2,}..., \phi_{n-1}$ are jumps and can be estimated as

$$\hat{\phi}_1 = \frac{S_{t_1}}{\lim\limits_{t \to t_1^-} \hat{S}_1}, \hat{\phi}_2 = \frac{S_{t_2}}{\lim\limits_{t \to t_2^-} \hat{S}_2}, ...\hat{\phi}_{n-1} = \frac{S_{t_{n-1}}}{\lim\limits_{t \to t_{n-1}^-} \hat{S}_{n-1}}.$$

The estimated parameters of Unequal Interval Nonlinear Stochastic Model 3.3.2 of stock X are recorded in Table 3.3.3.

Table 3.3.3 Estimated Parameters of Model 3.3.2 of Stock X

Interval Index	Monthly Nonlinear Model 3 with Jumps		
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}$
1	0.071226	-0.0006014	0.023908
2	0.29716	-0.0020805	0.023786
3	0.471856	-0.0025641	0.042281

4	0.409774	-0.0023388	0.026435
5	0.251018	-0.0013076	0.016441
6	0.700165	-0.0035874	0.021907
7	0.270348	-0.0013796	0.034988
8	0.130079	-0.0007162	0.014833
9	0.040871	-0.0001544	0.01686
10	0.210273	-0.000724	0.020032
11	0.689476	-0.0023335	0.011359
12	0.129436	-0.0004453	0.014872
13	0.364516	-0.0011824	0.014153
14	0.362752	-0.0009541	0.020535
15	0.147044	-0.0003428	0.01289
16	0.257365	-0.0006181	0.015308
17	0.422811	-0.0009202	0.019309
18	0.043869	-0.0001438	0.036822
19	0.213002	-0.0005966	0.027236
20	0.322719	-0.0007918	0.015437
21	0.106569	-0.0002685	0.022847
22	0.262288	-0.0006972	0.017497
23	0.153312	-0.0003764	0.014273
24	0.278518	-0.0007413	0.012968
25	0.527504	-0.0013832	0.011062
26	0.200628	-0.0004803	0.013615
27	0.36656	-0.0007506	0.015936
28	0.88476	-0.0018291	0.005951
29	0.114746	-0.0002405	0.01549
30	0.19202	-0.0004058	0.016691
31	0.368976	-0.0008158	0.013194
32	0.358247	-0.0007612	0.009788

33	0.074069	-0.0001383	0.010944
34	0.041831	-7.49E-05	0.008841
35	0.766581	-0.001496	0.008805
36	-0.00713	2.00E-05	0.010855
37	0.123638	-0.000189	0.01503
38	0.137906	-0.0001851	0.014094
39	0.21229	-0.0003142	0.021393

Figures 3.3.4, 3.3.5 and 3.3.6 are plots of predictions of Unequal Interval Nonlinear Model 3.3.2 of stock X with observations ranging from 1 to 300, 300 to 600 and 600 to 848 respectively.

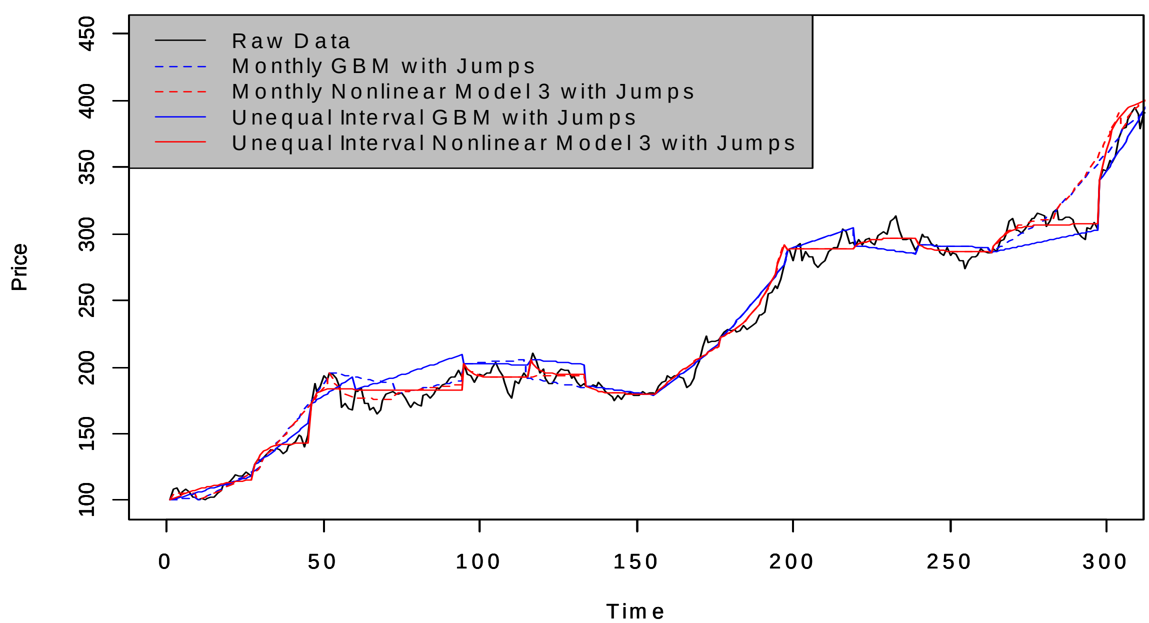


Figure 3.3.4 Comparison on Model 2.4.3, 2.4.4, 3.3.1 with Model 3.3.2 of Stock X (Observations 1-300)

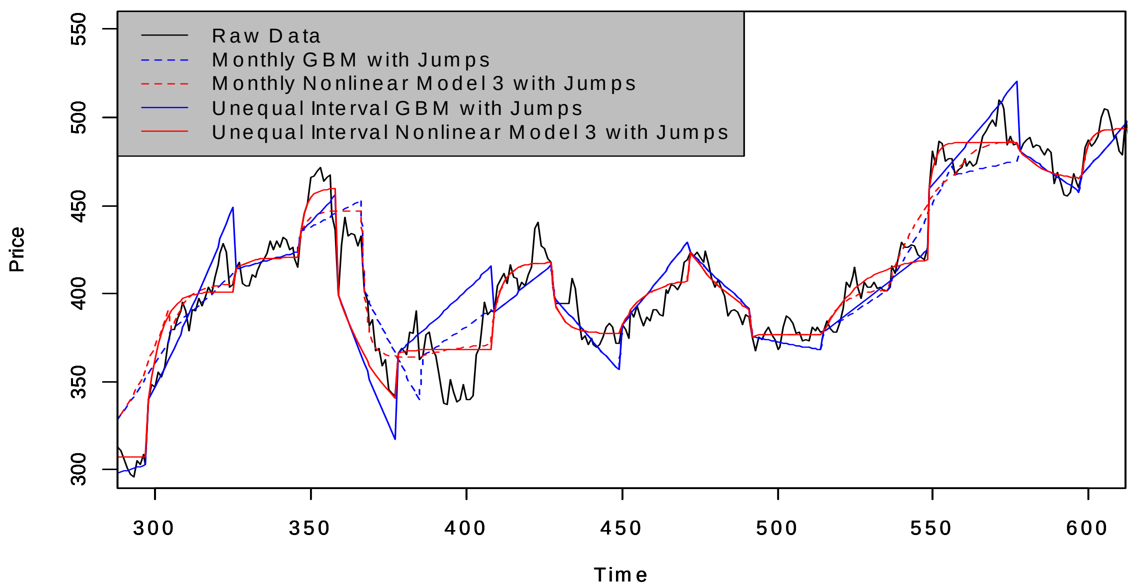


Figure 3.3.5 Comparison on Model 2.4.3, 2.4.4, 3.3.1 with Model 3.3.2 of Stock X
(Observations 300-600)

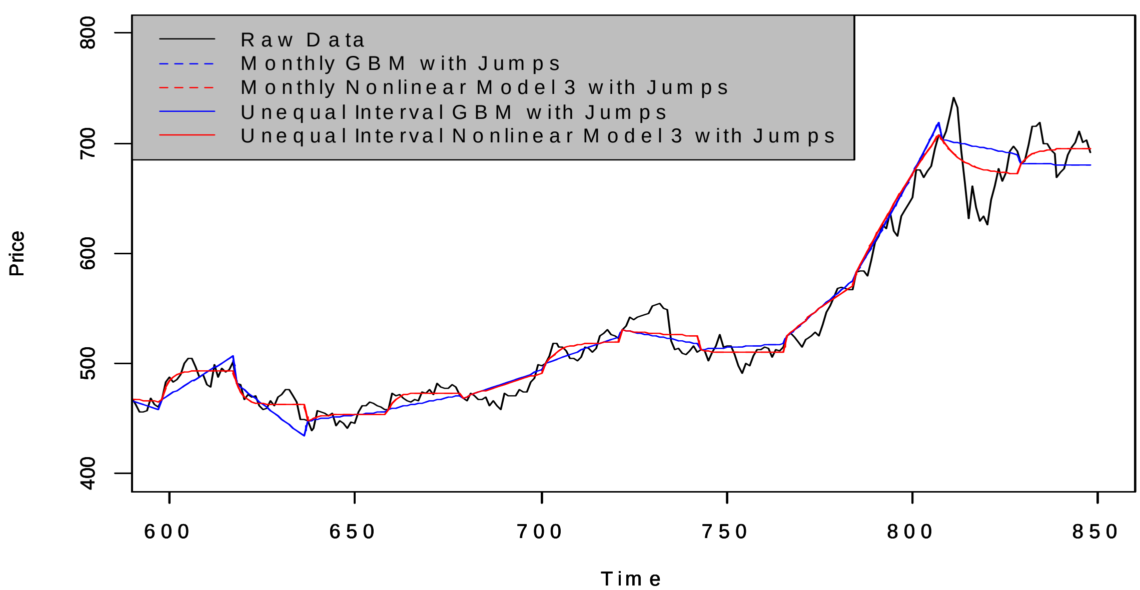


Figure 3.3.6 Comparison on Model 2.4.3, 2.4.4, 3.3.1 with Model 3.3.2 of Stock X
(Observations 600-848)

Table 3.3.4 shows the overall basic statistics of Monthly GBM Model 2.4.3, Monthly Nonlinear Model 3.1.1, 3.2.1 and 3.3.1 with jumps, and Unequal Nonlinear Model 3.1.2, 3.2.2 and 3.3.2 with Jumps.

Table 3.3.4 Basic Statistics for Models 2.4.3, 3.1.1, 3.2.1, 3.3.1, 3.1.2, 3.2.2 and 3.3.2 of Stock X

Model	$\overline{r}$	S_r^2	S_r	No. of Intervals	No. of Parameters
Monthly GBM with Jumps	-1.22683	207.3278	14.3989	41	122
Monthly Nonlinear Model 1 with Jumps	-1.928296	141.1754	11.88173	41	163
Monthly Nonlinear Model 2 with Jumps	-2.090806	143.2248	11.96765	41	204
Monthly Nonlinear Model 3 with Jumps	-1.731151	139.2792	11.80166	41	163
Unequal Interval GBM with Jumps	-1.962899	258.1040	16.06562	39	116
Unequal Interval Nonlinear Model 1 with Jumps	-0.531502	131.2354	11.4558	39	155
Unequal Interval Nonlinear Model 2 with Jumps	-0.609702	131.3068	11.45892	39	194
Unequal Interval Nonlinear Model 3 with Jumps	-0.402368	132.16	11.49609	39	155

Remark 3.3.2: From Table 3.3.4 we conclude that overall, the Nonlinear Model 3.1.2 with Unequal Interval has less variance among all Models (Monthly and Unequal Intervals) . With Monthly data partitioning, Nonlinear Model 3.3.1 with Jumps has the least mean and variance of residual error. With Unequal Interval data partitioning, all Nonlinear Models 3.1.2, 3.2.2 and 3.3.2 have less mean and variance of residual error than GBM (linear) model.

3.4 Illustration of Nonlinear Stochastic Models to Data Set of Stock Y

In this section, we apply the Monthly Nonlinear Models 1, 2 and 3 with jumps, that is, Nonlinear Model 3.1.1, 3.2.1 and 3.3.1 to stock Y. We also apply the Unequal Interval Nonlinear Models 1, 2 and 3 with jumps, that is, Nonlinear Model 3.1.2, 3.2.2 and 3.3.2 to stock Y. To minimize the repetition, here we only give the summary of these 6 models in Table 3.4.1.

The price data set of stock Y is relative larger than the price data set of stock X. There are 5630 observations over the past 22 years from September 1984 to December 2006. The Monthly Nonlinear Models have 268 monthly intervals, and the Unequal Interval Models have 256 intervals with the daily relative difference = 3.5% as the threshold.

Table 3.4.1 Basic Statistics for Models of Stock Y

Model	$\bar{r}$	S_r^2	S_r	No. of Intervals	No. of Parameters
Monthly GBM with Jumps	-0.009826	1.206479	1.098399	268	803
Monthly Nonlinear Model 1 with Jumps	0.020068	1.469688	1.212307	268	1071
Monthly Nonlinear Model 2 with Jumps	-0.002057	1.155363	1.074878	268	1339
Monthly Nonlinear Model 3 with Jumps	0.026632	1.295051	1.138003	268	1071
Unequal Interval GBM with Jumps	-0.012482	1.199703	1.095310	256	767
Unequal Interval Nonlinear Model 1 with Jumps	-0.004258	0.606470	0.778762	256	1023
Unequal Interval Nonlinear Model 2 with Jumps	-0.011478	0.603385	0.776779	256	1279
Unequal Interval Nonlinear Model 3 with Jumps	0.006577	0.612513	0.782632	256	1023

Remark 3.4.1: Table 3.4.1 shows the overall basic statistics of Stock Y with respect to all stated Monthly and Unequal Interval Nonlinear Models. Under the Monthly data partitioning, Nonlinear Model 2 has the least mean and variance of residual error. Under the Unequal Interval data partitioning, all stated Nonlinear Models 1, 2 and 3 have less mean and variance of residual error than the GBM model. Moreover, Nonlinear Model 2 with Unequal Interval has the least variance and standard deviation of residual error among all models. Furthermore, this unequal data partitioning process has less number of subintervals than the monthly data partitioning process.

3.5 Illustration of Nonlinear Stochastic Models to Data Set of S&P 500 Index

In this section, we apply the Monthly Nonlinear Models 1, 2 and 3, that is, Nonlinear Model 3.1.1, 3.2.1 and 3.3.1 of S&P 500 Index. We also apply the Unequal Interval Nonlinear Models 1, 2 and 3, that is, Nonlinear Model 3.1.2, 3.2.2 and 3.3.2 on S&P 500 Index. Again, to minimize the repetition, here we only give the summary of these 6 models in Table 3.5.1. Since the dataset is too large, here we only provide the summary of the models. The dataset of SP500 Index is larger than the previous datasets of stocks X and Y. There are 14844 observations over the past 59 years starting from January 1950 to December 2008. The Monthly Nonlinear Models have 708 monthly intervals, and the Unequal Interval Models have 570 intervals with the daily relative difference = 0.8% as the threshold.

Table 3.5.1 Basic Statistics for Models of S&P 500 Index

Model	$\overline{r}$	S_r^2	S_r	No. of Intervals	No. of Parameters
Monthly GBM with Jumps	58.30902	486.4579	22.05579	708	2123
Monthly Nonlinear Model 1 with Jumps	4.027517	281.7153	16.78438	708	2831
Monthly Nonlinear Model 2 with Jumps	4.084275	330.1271	16.78393	708	3539
Monthly Nonlinear Model 3 with Jumps	4.274907	282.7780	16.81600	708	2831

Unequal Interval GBM with Jumps	2.471564	210.2159	14.49882	570	1709
Unequal Interval Nonlinear Model 1 with Jumps	0.6186245	79.46592	8.914366	570	2279
Unequal Interval Nonlinear Model 2 with Jumps	0.5835638	78.5180	8.861039	570	2849
Unequal Interval Nonlinear Model 3 with Jumps	0.6590607	79.5725	8.920342	570	2279

Remark 3.5.1: Table 3.5.1 shows the overall basic statistics of S&P 500 Index for all stated Monthly and Unequal Interval Nonlinear Models. Under the Monthly data partitioning, Nonlinear Model 1 has the least mean and variance of residual error. Under the Unequal Interval data partitioning, all stated Nonlinear Models 1, 2 and 3 have less mean and variance of residual error than the GBM model. Nonlinear Model 2 with Unequal Interval has the least mean and the variance of residual error among all stated models. Furthermore, this unequal interval nonlinear model 2 has the least variance of residual error and the number of intervals.

3.6 Conclusions and Comments

In this chapter, we presented three nonlinear stochastic models. By using classical model building process, we developed the modified version of nonlinear stochastic models under equal and unequal data partitioning processes with jumps. Based on our study, in the following, we draw a few important conclusions.

(a) The Following Table 3.6.1 provides the summary of results for all three datasets. It shows that Nonlinear Model 2 ranks No.1 in both monthly and unequal interval data partitioning models of 2 out 3 data sets (stock X, stock Y and S&P 500 Index).

Table 3.6.1 Summary of Models in Chapter 3

Stock	Monthly Interval				Unequal Interval			
	Rank1	Rank2	Rank3	Rank4	Rank1	Rank2	Rank3	Rank4
X	Non.3	Non.2	Non.1	GBM	Non.1	Non.2	Non.3	GBM
Y	Non.2	GBM	Non.3	Non.1	Non.2	Non.1	Non.3	GBM
S&P500	Non.2	Non.1	Non.3	GBM	Non.2	Non.1	Non.3	GBM

The three data sets in our study, both stocks X and Y are from world Fortune 500 companies; S&P 500 Index is a stock Index. From Table 3.6.1, we notice that for two data sets (stock Y and S&P 500 Index) with both Monthly Interval and Unequal Interval data partitioning processes, Nonlinear Model 2 is the best model. These two data sets (stock Y and S&P 500 Index) share a common characteristic. Comparing to data set stock X (848 observations), both of these two dataset are very large, having 5630 and 14844 observations, respectively.

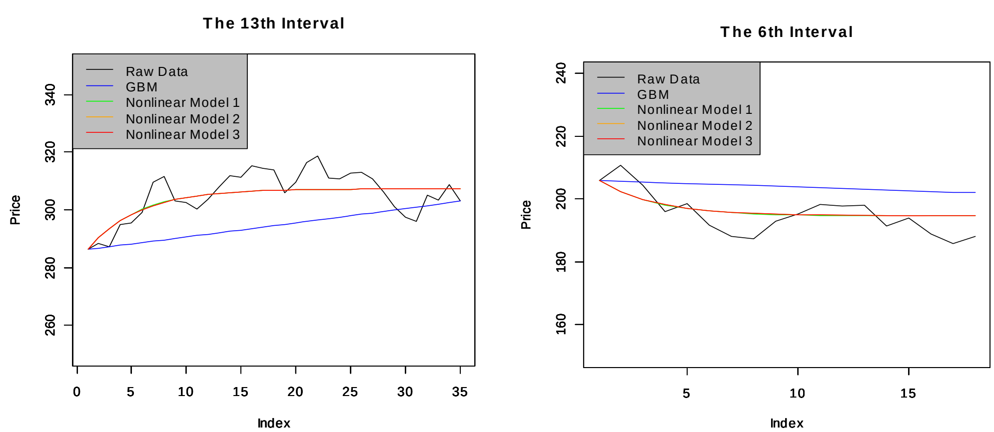
(b) Tables 3.3.4, 3.4.1 and 3.5.1 show the overall basic statistics of different models of stocks X, Y and S&P 500 Index. For the monthly data partitioning, Nonlinear Model 2 is better than GBM Model for Stock Y and S&P 500 Index, and Nonlinear Model 3 is better than GBM model for stock X. For the unequal interval data partitioning process, Nonlinear Model 2 is better than GBM Model for Stock Y and S&P 500 Index, and Nonlinear Model 3 is better than GBM model for stock X.

(i) For monthly data partitioning, all three nonlinear models are better than GBM model for all three price data sets. The Nonlinear Model 2 is better than GBM model and Nonlinear Models 1 and 3.

(ii) Under unequal interval data partitioning process and for all three stock data sets, all non-linear models are better than GBM model.

(iii) The unequal data partitioning approach is superior than the monthly data partitioning approach.

(iv) Under both equal and unequal data partitioning approach, the Nonlinear Model 2 is the best for stock Y and S&P 500 Index, and Nonlinear Model 1 is best for stock X.



(a) (b)

Figure 3.6.1 The Predicted Value of Stock X

(c) Again, from three tables 3.3.4, 3.4.1 and 3.5.1, we observe that the performance of Nonlinear Models 1, 2, and 3 are very similar. The predicted values for a particular interval are in Figure 3.6.1(a) and (b). In Figure 3.6.1 (a) and (b), we notice that the 3 red curve, orange curve and green curve are overlapped on each other. The blue curve represents the predicted value using GBM model. Furthermore, from the plot (the 13^{th} interval and the 6^{th} interval), we conclude that in this particular intervals Nonlinear Models estimates the stock price better than the GBM model. The reason is in that particular time interval, every possible environmental information often leads to wild movements in stock price. The drift and volatility are not constant any more in that particular time interval. Hence the nonlinear model can describe the stock price process much better than the GBM model.

(d) So far, we focused our attention to build stochastic models for stock price data sets. This modeling approach can also be used for any other type of data sets. Furthermore, our preceding stochastic modeling analysis of stock price confirms that a stock price process is nonlinear and non stationary stochastic models. However, the next important problem in modeling is to predict the future dynamic state of processes, in particular, stock market price. The study of this problem is focused in the next chapter.

Chapter 4
Nonlinear Stochastic Models with Time Varying Coefficients

4.0 Introduction

Stochastic dynamic models described in Chapters 2 and 3 were applicable to piece-wise time-invariant dynamic processes. In this chapter, based on our study of three stochastic nonlinear models, we generalize our stochastic modeling dynamic process by using the nonlinear stochastic differential equations with time varying coefficients. We focus our attention to only nonlinear models 1 and 2 that have been exhibited better than model 3 in Chapter 3.

Corresponding to nonlinear time invariant models 1 and 2, we present nonlinear stochastic models with time varying coefficient in Section 4.1 and 4.2, respectively. Using these nonlinear time varying models, we derive corresponding time series models. These time series models are tested by the three data sets, stock X, Y and S&P 500 Index. Furthermore, they are compared to the existing time series models [10, 12, 13, 38, 39] in Section 4.3. Finally, conclusions are drawn in Section 4.4.

4.1 Nonlinear Stochastic Dynamic Model 1 with Time Varying Coefficients

In Chapters 3, nonlinear stochastic dynamic models with constant coefficients were investigated. In this section, we assume that the rate parameters in the nonlinear stochastic dynamic model 1 (in Section 3.1, Chapter 3) are functions of time.

Nonlinear Stochastic Model 1 on Overall Data: Let us consider a stochastic nonlinear model corresponding to equation (3.1.1) as

$$dS_t = (\alpha_t \ln S_t + \beta_t + \frac{\sigma_t^2}{2})S_t dt + \sigma_t S_t dW_t \ , \ S(0) = S_0 \qquad (4.1.1)$$

where, parameters α, β and σ are time varying smooth functions [6, 26]. We note that the existence and uniqueness of solution process of (4.1.1) follows by following similar arguments used in Section 3.1.

By following the arguments and using the transformation $V_t = \ln S_t$, we obtain

$$dV_t = (\alpha_t V_t + \beta_t)dt + \sigma_t dW_t. \tag{4.1.2}$$

To estimate the time varying parameters α, β and σ, we first use a numerical integration applied (4.1.2) as:

$$\int_{t-k}^{t} dV_s = \int_{t-k}^{t} (\alpha_s V_s + \beta_s)ds + \int_{t-k}^{t} \sigma_s dW_s$$

$$V_t - V_{t-k} = \int_{t-k}^{t-k+1} (\alpha_s V_s + \beta_s)ds + \ldots + \int_{t-1}^{t} (\alpha_s V_s + \beta_s)ds + \int_{t-k}^{t-k+1} \sigma_s dW_s + \ldots + \int_{t-1}^{t} \sigma_s dW_s$$

$$\approx \alpha_k V_{t-k} + \alpha_{k-1} V_{t-k+1} + \ldots + \alpha_1 V_{t-1} + \beta_k + \ldots + \beta_1 + \sigma_k \varepsilon_{t-k+1} + \ldots + \sigma_1 \varepsilon_t,$$

where, k is any positive integer and $\varepsilon_{t-i} = W_{t-i} - W_{t-i-1} \sim N(0,1)$, for $i = 0,1,\ldots,k-1$.

By denoting $\beta = \beta_k + \beta_{k-1} \ldots + \beta_1$, and rearranging terms in the equation, we have the following equation

$$V_t = (\alpha_k + 1)V_{t-k} + \alpha_{k-1} V_{t-k+1} + \ldots + \alpha_1 V_{t-1} + \beta + \sigma_k \varepsilon_{t-k+1} + \ldots + \sigma_1 \varepsilon_t. \tag{4.1.3}$$

This is exactly a time series ARIMA model with order (p,q), where p=k and q=k-1. The constant term β can be eliminated by taking the first order difference filter (d=1). Obviously, we notice that when k=1, we have the constant coefficients case (3.1.4) in Chapter 3. If k=2, equation (4.1.3) is equivalent to ARIMA(2,1). If we assume that k=2 and $\sigma_2 = 0$, then equation (4.1.3) is equivalent to ARIMA(2,0).

Under the transformation $V_t = \ln S_t$ and following the Statistical Model Identification Procedure 1.5.1 described in Section 1.5, the AICs of ARIMA models of three data sets (stock X, Y and S&P 500 Index) are presented in Table 4.1.1.

Table 4.1.1 AIC of Time Varying Coefficients Nonlinear Model 1
of Different Models of Three Datasets: Stock X, Stock Y and S&P 500 Index

Model	Stock X	Stock Y	S&P 500 Index
	AIC	AIC	AIC
(3,1,2)	-4122.57	-22693.15	-96126.87
(3,1,1)	-4124.70	-22694.84	-96126.3
(3,1,0)	-4124.20	-22687.87	-96128.11
(2,1,3)	-4127.76	-22692.57	-96126.82
(2,1,2)	-4125.36	-22685.91	-96128.74
(2,1,1)	-4126.42	-22682.35	-96130.18
(2,1,0)	-4126.20	-22683.14	-96126.53
(1,1,3)	-4124.86	-22694.39	-96127.68
(1,1,2)	-4124.19	-22682.03	-96130.20
(1,1,1)	-4126.21	-22681.05	-96120.59
(1,1,0)	-4128.10	-22683.05	-96082.23
(0,1,3)	-4124.30	-22687.57	-96129.77
(0,1,2)	-4126.20	-22682.90	-96130.02
(0,1,1)	-4127.92	-22683.06	-96086.05

From Table 4.1.1, we notice that for stock X, ARIMA model (1,1,0) gives us the minimum AIC, that is, a mix model of a first order autoregressive with a first difference filter. The model is written as

$$(1-0.0872B)(1-B)V_t = 0.02110687\varepsilon_t .$$

After expanding the autoregressive operator and the difference filter, we have

$$(1-1.0872B-0.0872B^2)V_t = 0.02110687\varepsilon_t$$

which implies

$$V_t = 1.0872V_{t-1} + 0.0872V_{t-2} + 0.02110687\varepsilon_t .$$

By letting $\varepsilon_t = 0$, we have the one day ahead forecasting formula of V_t of stock X as

$$\hat{V}_t = 1.0872V_{t-1} + 0.0872V_{t-2} . \tag{4.1.4}$$

Then, by applying the inverse transformation of "ln", we get $\hat{S}_t = \exp(\hat{V}_t)$. The residual error $r_i = S_i - \hat{S}_i$ is computed, and its basic statistics is recorded in Table 4.1.2.

Similarly, for data set Stock Y, the fitted ARIMA model (3,1,1) gives us the minimum AIC. The model is

$$(1 + 0.6575B + 0.0156B^2 + 0.0523B^3)(1 - B)V_t = (0.03220248 + 0.6638B)\varepsilon_t .$$

By following above argument, we have

$$V_t = 0.3425V_{t-1} + 0.6419V_{t-2} - 0.0367V_{t-3} + 0.0523V_{t-4} + 0.03220248\varepsilon_t + 0.6638\varepsilon_{t-1}$$

.

By letting $\varepsilon_t = 0$, we obtain the one day ahead forecasting formula of V_t of stock Y as

$$\hat{V}_t = 0.3425V_{t-1} + 0.6419V_{t-2} - 0.0367V_{t-3} + 0.0523V_{t-4} + 0.6638\varepsilon_{t-1} . \qquad (4.1.5)$$

Again, by applying the inverse transformation of "ln", we get $\hat{S}_t = \exp(\hat{V}_t)$. The residual error $r_i = S_i - \hat{S}_i$ is computed, and its basic statistics are recorded in Table 4.1.2.

For data set S&P 500 Index, the fitted ARIMA models (1,1,2) gives us the minimum AIC, and the model is

$$(1 + 0.2297B)(1 - B)V_t = (0.009490522 + 0.2787B - 0.0438B^2)\varepsilon_t$$

By following above argument, we have

$$V_t = 0.7703V_{t-1} + 0.2297V_{t-2} + 0.009490522\varepsilon_t + 0.2787\varepsilon_{t-1} - 0.0438\varepsilon_{t-2}$$

By letting $\varepsilon_t = 0$, we obtain the one day ahead forecasting formula of V_t of S&P 500 Index as

$$\hat{V}_t = 0.7703V_{t-1} + 0.2297V_{t-2} + 0.2787\varepsilon_{t-1} - 0.0438\varepsilon_{t-2} \qquad (4.1.6)$$

Again, by applying the inverse transformation of "ln", we get $\hat{S}_t = \exp(\hat{V}_t)$. The residual error $r_i = S_i - \hat{S}_i$ is computed, and its basic statistics are recorded in Table 4.1.2.

Table 4.1.2 Basic Statistics of Time Varying Coefficients Nonlinear Model 1 of Three Data Sets: Stock X, Stock Y and S&P 500 Index

Data Set	Model	Mean of residual	Variance of residual	Standard deviation of residual
Stock X	(1,1,0)	0.628727	57.38475	7.575272
Stock Y	(3,1,1)	0.015286	0.344827	0.587220
S&P 500 Index	(1,1,2)	0.058922	46.90737	6.848895

Remark 4.1.1: For data set stock X, from Tables 3.3.4 and 4.1.2, the nonlinear stochastic model 1 with time varying coefficients has the minimum variance of residual error. This is the same as for stock Y (Tables 3.4.1 and 4.1.2) and S&P 500 Index (Tables 3.5.1 and 4.1.2). We note that the nonlinear stochastic model 1 with time varying coefficients is applied to overall data set. The study in Chapter 3 is with regard to the unequal interval data partitioning process.

In the following we apply the unequal interval Data Partitioning Process 2.3.4 for nonlinear stochastic model 1 with time varying coefficients.

Nonlinear Stochastic Model 4.1.1 (Unequal Interval Nonlinear Model 1 with Time Varying Coefficients): Let $[0,t_1),[t_1,t_2),[t_2,t_3),\ [t_3,t_4),[t_4,t_5)...\ [t_{n-1},t_n)$ be the n time intervals as defined in Data Partition Process 2.3.4. The nonlinear stochastic differential equation is described by:

$$dS_t^{I_i} = (\alpha_t^{I_i} \ln S_t^{I_i} + \beta_t^{I_i} + \frac{(\sigma_t^{I_i})^2}{2})S_t^{I_i} dt + \sigma_t^{I_i} S_t^{I_i} dW_t ,$$

$$S_0 = S_0 \text{, if } t_{i-1} \le t < t_i ,\ i = 1,...,n . \tag{4.1.7}$$

$\alpha^{I_i}, \beta^{I_i}$ and σ^{I_i} , $i = 1,...,n$, are time varying parameters.

As before, by imitating the time series definition process, we arrive at

$$V_t^{I_i} = (\alpha_k^{I_i} + 1)V_{t-k}^{I_i} + \alpha_{k-1}^{I_i} V_{t-k+1}^{I_i} + ... + \alpha_1^{I_i} V_{t-1}^{I_i} + \beta + \sigma_k^{I_i} \varepsilon_{t-k+1}^{I_i} + \sigma_{k-1}^{I_i} \varepsilon_{t-k+1}^{I_i} + ... + \sigma_1^{I_i} \varepsilon_t^{I_i} . \tag{4.1.8}$$

Furthermore, S_t and $\hat{\phi}_i, i = 1,2,...,n-1$ are defined analogous to (3.1.8).

Stochastic Model Identification Procedure 4.1.1: In the following, we present a modified version of Statistical Model Identification Procedure 1.5.1 [10,12,38]. It is as follows:

i. By following the Data Partition Process 2.3.4, the entire data set is decomposed into n sub data sets.

ii. For every sub data set, use the transformation $V_t^{I_i} = \ln S_t^{I_i}$, $i = 1,...,n$.

iii. For every sub data sets, repeat steps ii – v in Stochastic Model Identification Procedure 1.5.1.

iv. For every sub data set, and for each possible set of (p, q), compute the predicted value $\hat{V}_t^{I_i(p,q)}$, and then compute the predicted value $\hat{S}_t^{I_i(p,q)}$, by using the inverse of "ln" transformation, that is, $\hat{S}_t^{I_i(p,q)} = \exp(\hat{V}_t^{I_i(p,q)})$.

v. For every sub data set and for each possible models, compute the residual error $r_t^{(p,q)} = S_t - \hat{S}_t^{I_i(p,q)}, t_{i-1} \le t < t_i, i = 1,2,...,n.$.

vi. For all possible set of (p, q), compute mean, variance and standard deviation of overall residual error $r_t^{(p,q)}, 1 \le t \le T.$ The model provides the smallest variance of residual is the fitted model.

Table 4.1.3, 4.1.4 and Table 4.1.5 exhibit the basic statistics of the residuals using different value of k with unequal interval data partitioning of three datasets: Stock X, Y and S&P 500 Index respectively. Here the thresholds of daily relative difference for three data sets are set to 5%, 4.5% and 2%, respectively, and the corresponding number of intervals are 10, 66 and 87.

Table 4.1.3 Basic Statistics of Stochastic Models 4.1.1 with Different Set of (p, q) Under Log-Transformation with Unequal Data Partition, threshold=5% of Stock X

Model	Mean of Residual	Variance of Residual	Standard Deviation of Residual	Number of Intervals
(3,1,2)	0.535531	44.01918	6.634695	10
(3,1,1)	0.477112	46.09286	6.789173	10
(3,1,0)	0.584083	47.90231	6.92115	10

(2,1,3)	0.571359	43.62668	6.60505	10
(2,1,2)	0.588372	44.5358	6.673515	10
(2,1,1)	0.648873	47.5926	6.898739	10
(2,1,0)	0.652255	48.23812	6.945367	10
(1,1,3)	0.479333	45.94483	6.778261	10
(1,1,2)	0.558626	46.4712	6.816979	10
(1,1,1)	0.562882	46.79851	6.840944	10
(1,1,0)	0.636871	48.66251	6.975852	10
(0,1,3)	0.640242	48.47799	6.962613	10
(0,1,2)	0.64442	48.24343	6.945749	10
(0,1,1)	0.640242	48.47799	6.962613	10

Table 4.1.4 Basic Statistics of Stochastic Models 4.1.1 with Different Set of (p, q) Under Log-Transformation with Unequal Data Partition, threshold=4.5% of Stock Y

Model	Mean of Residual	Variance of Residual	Standard Deviation of Residual	Number of intervals
(3,1,2)	0.013137	0.284444	0.533333	66
(3,1,1)	0.010776	0.303817	0.551196	66
(3,1,0)	0.011263	0.305855	0.553041	66
(2,1,3)	0.012679	0.287968	0.536626	66
(2,1,2)	0.00985	0.296015	0.544072	66
(2,1,1)	0.010492	0.306466	0.553593	66
(2,1,0)	0.012846	0.308965	0.555846	66
(1,1,3)	0.010388	0.302722	0.550202	66
(1,1,2)	0.010206	0.306028	0.553198	66
(1,1,1)	0.013911	0.308552	0.555475	66
(1,1,0)	0.013396	0.311613	0.558223	66
(0,1,3)	0.013494	0.31157	0.558185	66
(0,1,2)	0.013209	0.30873	0.555635	66
(0,1,1)	0.013494	0.31157	0.558185	66

Table 4.1.5 Basic Statistics of Stochastic Models 4.1.1 with Different Set of (p, q) Under Log-Transformation with Unequal Data Partition, threshold=2% of S&P 500 Index

Model	Mean of Residual	Variance of Residual	Standard Deviation of Residual	Number of intervals
(3,1,2)	0.116613	39.73163	6.303303	87
(3,1,1)	0.107649	41.31667	6.427805	87
(3,1,0)	0.108677	42.04793	6.484438	87
(2,1,3)	0.130816	37.94544	6.159987	87
(2,1,2)	0.118426	40.47064	6.361654	87
(2,1,1)	0.105609	41.83156	6.467732	87
(2,1,0)	0.107766	42.51392	6.52027	87
(1,1,3)	0.112032	41.55797	6.446547	87
(1,1,2)	0.115318	41.9422	6.47628	87
(1,1,1)	0.113373	42.86889	6.547434	87
(1,1,0)	0.110322	43.49516	6.595086	87
(0,1,3)	0.105709	43.30008	6.580279	87
(0,1,2)	0.105485	42.55877	6.523708	87
(0,1,1)	0.105709	43.30008	6.580279	87

From Table 4.1.3 and Table 4.1.5, we can see that the model (2,1,3) has minimum variance and standard deviation of residuals, for stock X and S&P 500 Index. From Table 4.1.4 we see that the model (3,1,2) is the best model which provides the minimum variance of the residual. We further note that ARIMA model (2,1,3) is the best for three all data sets.

Remark 4.1.2: For stock X, we compare Table 3.3.4, 4.1.2, with Table 4.1.3, we notice that nonlinear model 1 with time varying coefficients under unequal interval data partitioning process provides least variance and standard deviation of residual error. Similarly, for stock Y, comparing Table 3.4.1, 4.1.2 with Table 4.1.4; for S&P 500 Index, comparing Table 3.5.1, 4.1.2 with Table 4.1.5, we have the same conclusion.

4.2 Nonlinear Stochastic Dynamic Model 2 with Time Varying Coefficients

In this section, we assume that the rate parameters in the nonlinear stochastic dynamic model 2 (in Section 3.2, Chapter 3) are not constants, that is, the rates, α, β and σ are functions of time, and N is still a constant.

Nonlinear Stochastic Model 2 on Overall Data: Let us consider a stochastic nonlinear model corresponding to equation (3.2.1) as

$$dS_t = (\alpha_t S_t + \beta_t S_t^N + \frac{N}{2}\sigma_t^2 S_t^{2N-1})dt + \sigma_t S_t^N dW_t \ , \ S(0) = S_0 \qquad (4.2.1)$$

where, coefficients α, β and σ are time varying smooth functions [26]. We note that the existence and uniqueness of solution process of (4.2.1) follows by following similar arguments used in Section 3.2.

By following the arguments and using the transformation $V_t = \frac{S_t^{1-N}}{1-N}$, we obtain

$$dV_t = (\alpha_t (1-N)V_t + \beta_t)dt + \sigma_t dW_t \,. \qquad (4.2.2)$$

To estimate the time varying parameters α, β and σ, we first use a numerical integration applied to (4.2.2) as follows:

$$\int_{t-k}^{t} dV_s = \int_{t-k}^{t} (\alpha_s (1-N)V_s + \beta_s)ds + \int_{t-k}^{t} \sigma_s dW_s$$

$$V_t - V_{t-k} = \int_{t-k}^{t-k+1} (\alpha_s (1-N)V_s + \beta_s)ds + ... + \int_{t-1}^{t} (\alpha_s (1-N)V_s + \beta_s)ds + \int_{t-k}^{t-k+1} \sigma_s dW_s + ... + \int_{t-1}^{t} \sigma_s dW_s$$

$$\approx (1-N)\alpha_k V_{t-k} + ... + (1-N)\alpha_1 V_{t-1} + \beta_k + \beta_{k-1} ... + \beta_1 + \sigma_k \varepsilon_{t-k+1} + ... + \sigma_1 \varepsilon_t \,,$$

where, k is any positive integer and $\varepsilon_{t-i} = W_{t-i} - W_{t-i-1} \sim N(0,1), \text{ for } i = 0,1,...,k-1.$

By denoting $\beta = \beta_k + \beta_{k-1} ... + \beta_1$ and rearranging terms in the equation, we have the following equation

$$V_t = ((1-N)\alpha_k + 1)V_{t-k} + (1-N)\alpha_{k-1} V_{t-k+1} + ... + (1-N)\alpha_1 V_{t-1} + \beta + \sigma_k \varepsilon_{t-k} + ... + \sigma_1 \varepsilon_t \,. \qquad (4.2.3)$$

This is also a time series ARIMA model with order (p,q), where p equals k and q equals k-1. The constant term β can be eliminated by taking the first order difference filter (d=1).

Similarly, we notice that when k=1, we have the constant coefficients case (3.2.4) in Chapter 3. If k=2, equation (4.2.3) is equivalent to ARIMA(2,1). If we assume that k=2 and $\sigma_2 = 0$, then equation (4.2.3) is equivalent to ARIMA(2,0).

Stochastic Model Identification Procedure 4.2.1: The difference between the nonlinear model 1 and 2 is that nonlinear model 2 has a parameter N that can not be estimated directly. In the following, we present a modified version of Statistical Model Identification Procedure 1.5.1. It is as follows:

i. Let $0 \le N \le 1.2$ and N≠1.

ii. For each value of N, say $N = \hat{N}$, use the transformation, $V_t = \dfrac{S_t^{1-\hat{N}}}{1-\hat{N}}, t = 1,2,...,T$.

iii. Follow the Stochastic Model Identification Procedures 1.5.1 (ii-vii).

iv. By knowing the best model $(p,q)^{(\hat{N})}$ for each value of N, compute the predicted value of price process by applying the inverse transformation of $\hat{S}_t^{(\hat{N})} = ((1-\hat{N})\hat{V}_t)^{\frac{1}{1-\hat{N}}}$.

v. Computer the residual error $r_t = S_t - \hat{S}_t^{(\hat{N})}, t = 1,2,...,T$.

vi. Repeat the steps (ii-v) for each given $N = \hat{N} \in [0,1) \cup (1,1.2]$.

vii. The value $\hat{N}$ and the corresponding model provides the smallest variance of residual error ($r_t, t = 1,2,...,T$) is the estimated N and fitted model.

We apply the Stochastic Model Identification Procedure 4.2.1 to three data sets and the result is exhibited in Table 4.2.1.

Table 4.2.1 Basic Statistics of Time Varying Coefficients Nonlinear Model 2 of Three Data Sets: Stock X, Stock Y and S&P 500

Data Set	Model	$\hat{N}$	Mean of residual	Variance of residual	Standard deviation of residual
Stock X	(3,1,2)	0.07	0.623089	56.59861	7.523205
Stock Y	(2,1,2)	0.03	0.013675	0.340876	0.583846
S&P 500 Index	(3,1,2)	0.03	0.071906	46.27549	6.802609

Table 4.2.1 shows basic statistics of time varying coefficients of nonlinear stochastic model 2 for three Data Sets: Stock X, Stock Y and S&P 500. From the table, we can see that for stock X, ARIMA model (3,1,2) gives us the minimum variance of residual. The model is

$$(1-0.5837B-0.9055B^2+0.0841B^3)(1-B)V_t=(4.942671+0.6804B+0.999B^2)\varepsilon_t.$$

After expanding the autoregressive operator and the difference filter, we have

$$(1-1.5837B-0.3218B^2+0.8214B^3+0.0841B^4)V_t=(4.942671+0.6804B+0.999B^2)\varepsilon_t$$

which implies

$$V_t=1.5837V_{t-1}+0.3218V_{t-2}-0.8214V_{t-3}-0.0841V_{t-4}+4.942671\varepsilon_t+0.6804\varepsilon_{t-1}+0.999\varepsilon_{t-2}.$$

By letting $\varepsilon_t=0$, we have the one day ahead forecasting formula of V_t of stock X as

$$\hat{V}_t=1.5837V_{t-1}+0.3218V_{t-2}-0.8214V_{t-3}-0.0841V_{t-4}+0.6804\varepsilon_{t-1}+0.999\varepsilon_{t-2}. \qquad (4.2.4)$$

Then, by applying the inverse transformation, $\hat{S}_t=((1-\hat{N})\hat{V}_t)^{\frac{1}{1-\hat{N}}}$, $\hat{N}=0.07$ the residual error $r_i=S_i-\hat{S}_i$ is computed, and its basic statistics are recorded in Table 4.2.1.

Similarly, for data set Stock Y, the fitted ARIMA model (2,1,2) gives us the minimum variance of residual. The model is

$$(1+1.2999B-0.6948B^2)(1-B)V_t=(0.5260228-1.3207B+0.7420B^2)\varepsilon_t$$

By following above argument, we have

$$V_t=-0.2999V_{t-1}+1.9947V_{t-2}-0.6448V_{t-3}+0.5260228\varepsilon_t-1.3207\varepsilon_{t-1}+0.7420\varepsilon_{t-2}.$$

By letting $\varepsilon_t=0$, we obtain the one day ahead forecasting formula of V_t of stock Y as

$$\hat{V}_t=-0.2999V_{t-1}+1.9947V_{t-2}-0.6448V_{t-3}-1.3207\varepsilon_{t-1}+0.7420\varepsilon_{t-2}. \qquad (4.2.5)$$

Again, by applying the inverse transformation, $\hat{S}_t=((1-\hat{N})\hat{V}_t)^{\frac{1}{1-\hat{N}}}$, $\hat{N}=0.03$ the residual error $r_i=S_i-\hat{S}_i$ is computed, and its basic statistics are recorded in Table 4.2.1.

For data set S&P 500 Index, the fitted ARIMA models (3,1,2) gives us the minimum variance of residual, and the model is

$$(1+0.3286B+0.1991B^2+0.0385B^3)(1-B)V_t = (5.517246-0.3932B-0.2477B^2)\varepsilon_t$$

By following above argument, we have

$$V_t = 0.6714V_{t-1}+0.1295V_{t-2}+0.1606V_{t-3}-0.0385V_{t-4}+5.517246\varepsilon_t-0.3932\varepsilon_{t-1}-0.2477\varepsilon_{t-2}$$

By letting $\varepsilon_t = 0$, we obtain the one day ahead forecasting formula of V_t of S&P 500 Index as

$$\hat{V}_t = 0.6714V_{t-1}+0.1295V_{t-2}+0.1606V_{t-3}-0.0385V_{t-4}-0.3932\varepsilon_{t-1}-0.2477\varepsilon_{t-2} \quad (4.2.6)$$

Again, by applying the inverse transformation, $\hat{S}_t = ((1-\hat{N})\hat{V}_t)^{\frac{1}{1-\hat{N}}}$, $\hat{N}=0.03$ the residual error $r_i = S_i - \hat{S}_i$ is computed, and its basic statistics are recorded in Table 4.2.1.

Remark 4.2.1: For data set stock X, comparing Table 3.3.4, 4.1.1 and Table 4.2.1, nonlinear stochastic model 2 with time varying coefficients has the minimum variance of residual error. This is the same as for stock Y (Table 3.4.1, 4.1.1 and Table 4.2.1) and S&P 500 Index (Table 3.5.1, 4.1.1and Table 4.2.1).

In the following we apply the unequal interval Data Partitioning Process 2.3.4 for nonlinear stochastic model 2 with time varying coefficients.

Nonlinear Stochastic Model 4.2.1 (Unequal Interval Nonlinear Model 2 with Time Varying Coefficients): Let $[0,t_1),[t_1,t_2),[t_2,t_3)$, $[t_3,t_4),[t_4,t_5)\ldots[t_{n-1},t_n)$ be the n time intervals as defined in Data Partition Process 2.3.4. The nonlinear stochastic differential equation is described by:

$$dS_t^{I_i} = (\alpha_t^{I_i}S_t^{I_i}+\beta_t^{I_i}(S_t^{I_i})^N+\frac{N}{2}(\sigma_t^{I_i})^2(S_t^{I_i})^{2N-1})dt+\sigma_t^{I_i}(S_t^{I_i})^N dW_t\ ,$$

$$S_0 = S_0,\ \text{if}\ t_{i-1}\le t<t_i,\ i=1,\ldots,n. \quad (4.2.7)$$

α^{I_i},β^{I_i} and σ^{I_i} , $i=1,\ldots,n$, are time varying parameters, N is constant.[26,27,29].

As before, by imitating the time series definition process, we arrive at

$$V_t^{I_i} = ((1-N)\alpha_k^{I_i}+1)V_{t-k}^{I_i}+(1-N)\alpha_{k-1}^{I_i}V_{t-k+1}^{I_i}+\ldots+(1-N)\alpha_1^{I_i}V_{t-1}^{I_i}$$
$$+\beta+\sigma_k^{I_i}\varepsilon_{t-k+1}^{I_i}+\ldots+\sigma_1^{I_i}\varepsilon_t^{I_i} \quad (4.2.8)$$

Furthermore, S_t and $\hat{\phi}_i, i=1,2,\ldots,n-1$ are defined analogous to (3.2.8).

Stochastic Model Identification Procedure 4.2.2: In the following, we present a modified version of Statistical Model Identification Procedure 1.5.1. It is as follows:

i. By following the Data Partition Process 2.3.4, the entire data set is decomposed into n sub data sets.

ii. For each sub data set, follow steps i-ii of the Stochastic Model Identification Procedures 4.2.1.

iii. For each sub data set, follow the Stochastic Model Identification Procedures 1.5.1 steps ii-v.

iv. Using estimated parameters in step iii and compute the residual error $r_t^{(\hat{N})(p,q)} = S_t - \hat{S}_t^{(\hat{N})(p,q)}, t = 1,2,...,T$ for all possible (p,q).

v. Repeat steps ii-iv for $N = \hat{N} \in [0,1) \cup (1,1.2]$.

vi. For, a given set of (p, q), we compute overall sum of squared error for every value of $\hat{N}$ by $RSS^{(\hat{N})(p,q)} = \sum_{t=1}^{T} \left(r_t^{(\hat{N})(p,q)} \right)^2$.

vii. For the given (p, q) in step vi, we find the best N, corresponding to the minimum RSS.

viii. Repeat steps vi – vii for all possible model (p,q), we find the best N's with respect to the minimum RSS.

ix. From viii we choose the model with corresponding $\hat{N}$, which provides the smallest RSS.

Table 4.2.2, 4.2.3 and Table 4.2.4 show the basic statistics of the residual error using different set of (p, q) with unequal interval data partitioning of three datasets: Stock X, Y and S&P 500 Index, respectively. Here the thresholds of daily relative difference for three data sets are set to 5%, 4.5% and 2%, respectively, such that the sub intervals have enough observations to estimate the parameters. The residual error is defined as well as $r_i = S_i - \hat{S}_i$ for all observations.

Table 4.2.2 Basic Statistics of Stochastic Models 4.2.1 with Different Set of (p, q) under Transformation $\frac{S_t^{1-N}}{1-N}$ with Unequal Data Partition, threshold=5% of Stock X

Model	Mean of Residual	Variance of Residual	Standard Deviation of Residual	Number of intervals
(3,1,2)	0.48359	42.8776	6.548099	10
(3,1,1)	0.531298	45.89411	6.774519	10

(3,1,0)	0.571061	47.81292	6.914689	10
(2,1,3)	0.527294	41.97917	6.479133	10
(2,1,1)	0.631136	47.1654	6.867707	10
(2,1,0)	0.636309	48.16542	6.940131	10
(1,1,3)	0.520179	45.68867	6.75934	10
(1,1,2)	0.532497	46.36265	6.809013	10
(1,1,0)	0.619884	48.62252	6.972985	10
(0,1,3)	0.623101	48.43584	6.959586	10
(0,1,2)	0.627815	48.17912	6.941118	10
(0,1,1)	0.623101	48.43584	6.959586	10

Table 4.2.3 Basic Statistics of Stochastic Models 4.2.1 with Different Set of (p, q) under Transformation $\frac{S_t^{1-N}}{1-N}$ with Unequal Data Partition, threshold=4.5% of Stock Y

Model	Mean of Residual	Variance of Residual	Standard Deviation of Residual	Number of intervals
(3,1,2)	0.013494	0.280789	0.529895	66
(3,1,1)	0.012033	0.301285	0.548894	66
(3,1,0)	0.011385	0.305035	0.552300	66
(2,1,3)	0.012897	0.281243	0.530324	66
(2,1,0)	0.012724	0.308477	0.555407	66
(1,1,3)	0.009512	0.301548	0.549134	66
(1,1,2)	0.011650	0.305205	0.552454	66
(1,1,0)	0.012968	0.311111	0.557773	66
(0,1,3)	0.012975	0.311157	0.557814	66
(0,1,2)	0.013045	0.308202	0.555160	66
(0,1,1)	0.012975	0.311157	0.557814	66

Table 4.2.4 Basic Statistics of Stochastic Models 4.2.1 with Different Set of (p, q) under Transformation $\frac{S_t^{1-N}}{1-N}$ with Unequal Data Partition, threshold=2% of S&P 500 Index

Model	Mean of Residual	Variance of Residual	Standard Deviation of Residual	Number of intervals
(3,1,1)	0.084646	41.14148	6.414162	87
(3,1,0)	0.093201	41.90288	6.473243	87
(2,1,3)	0.114031	37.68062	6.138454	87
(2,1,0)	0.091409	42.38462	6.510347	87
(1,1,3)	0.099771	41.4015	6.43440	87
(1,1,2)	0.093127	41.76005	6.462202	87
(1,1,0)	0.095328	43.36592	6.585281	87
(0,1,3)	0.08993	43.17434	6.570719	87
(0,1,2)	0.090155	42.4362	6.514308	87
(0,1,1)	0.08993	43.17434	6.570719	87

Remark 4.2.2: For stock X, we compare Table 3.3.4, 4.2.1, with Table 4.2.2, we notice that nonlinear model 2 with time varying coefficients under unequal interval data partitioning process provides least variance and standard deviation of residual error. Similarly, for stock Y, comparing Table 3.4.1, 4.2.1 with Table 4.2.3; for S&P 500 Index, comparing Table 3.5.1, 4.2.1 with Table 4.2.4, we have the same conclusion.

4.3 Prediction and Comparison on Overall Data Sets

In Sections 4.1 and 4.2, using nonlinear continuous time varying stochastic models, we derived time series models. In this section, we compare our study of Sections 4.1 and 4.2 with the existing time series models, namely, k-th moving average model, k-th weighted and k-th exponential weighted moving average models [13,38,39]. A comparative study is made in the context of three overall data sets. In fact, the following models are compared with each other.

- Time series model (ARIMA) [10,12]
- k-th moving average model (Shi's model 1) [13,38,39]
- k-th weighted moving average model (Shi's model 2) [13,38,39]

- k-th exponential weighted moving average model (Shi's model 3) [13,38,39]
- Nonlinear Models with constant coefficients, Chapter 3
- Nonlinear Stochastic Model 1 on Overall Data Set, Section 4.1
- Nonlinear Stochastic Model 2 on Overall Data Set, Section 4.2

We summary the results for stock X, stock Y and S&P 500 Index in Table 4.3.1 , Table 4.3.2 and Table 4.3.3 respectively.

Table 4.3.1 Comparison Cited Models in Section 4.3 for Stock X

Model	Mean of Residual	Variance of Residual	Standard Deviation of Residual
ARIMA	0.6385010	57.39102	7.575686
k-th Moving Average Model	0.6342918	57.03750	7.552318
k-th Weighted Moving Average Model	0.6359891	57.14087	7.559158
k-th Exponential Weighted Moving Average Model	0.8944923	64.64898	8.040459
Nonlinear Models with Constant Coefficients	-0.6097021	131.2354	11.45580
Nonlinear Model 1 with Time Varying Coefficients	0.628727	57.38475	7.575272
Nonlinear Model 2 with Time Varying Coefficients	0.623089	56.59861	7.523205

Remark 4.3.1: For stock X, five models perform pretty much close to each other. They are ARIMA model, k-th moving average model, k-th weighted moving average model, nonlinear stochastic models 1 and 2 with time varying coefficients. Among these models, nonlinear stochastic model 2 with time varying coefficients has the least variance and standard deviation.

Table 4.3.2 Comparison Cited Models in Section 4.3 for Stock Y

Model	Mean of Residual	Variance of Residual	Standard Deviation of Residual
ARIMA	0.00725343	0.3419923	0.5848011
k-th Moving Average Model	0.00748872	0.3418268	0.5846595
k-th Weighted Moving Average Model	0.00741209	0.3411141	0.5840498
k-th Exponential Weighted Moving Average Model	0.01503370	0.3773675	0.6143024
Nonlinear Models with Constant Coefficients	-0.01147757	0.6033852	0.776779
Nonlinear Model 1 with Time Varying Coefficients	0.015286	0.344827	0.587220
Nonlinear Model 2 with Time Varying Coefficients	0.013675	0.340876	0.583846

Remark 4.3.2: Like stock X, for stock Y, five models perform pretty much close to each other. They are also ARIMA model, k-th moving average model, k-th weighted moving average model, nonlinear stochastic models 1 and 2 with time varying coefficients. Among these models, nonlinear stochastic model 2 with time varying coefficients has the least variance and standard deviation.

Table 4.3.3 Comparison Cited Models in Section 4.3 for S&P 500 Index

Model	Mean of Residual	Variance of Residual	Standard Deviation of Residual
ARIMA	0.07225937	46.2575	6.801286
k-th Moving Average Model	0.08731528	59.17083	7.692258

k-th Weighted Moving Average Model	0.07014848	46.25595	6.801173
k-th Exponential Weighted Moving Average Model	0.09544027	56.68555	7.52898
Nonlinear Models with Constant Coefficients	0.5835638	78.5180	8.861039
Nonlinear Model 1 with Time Varying Coefficients	0.058922	46.90737	6.848895
Nonlinear Model 2 with Time Varying Coefficients	0.071906	46.27549	6.802609

Remark 4.3.3: In Table 4.3.3, for S&P 500 Index, four models perform pretty much close to each other. There are ARIMA model, k-th weighted moving average model, nonlinear stochastic models 1 and 2 with time varying coefficients. Among these models, k-th weighted moving average model has least variance and standard deviation of residual error. We note that our nonlinear model 2 is reasonably close to linear weighted model.

From above discussion, we draw a few conclusions:

- For all three datasets, nonlinear stochastic models with time vary coefficient have less variance and standard deviation of residual than the nonlinear models with constant coefficients.
- Nonlinear stochastic model 2 with time varying coefficients has the least variance and standard deviation of residual among all models for two data sets, namely, stocks X and Y.
- Dr. Shi's k-th weighted moving average model [38] has the least variance and standard deviation of residual among all models for one dataset – S&P 500 Index. We remark the standard deviation of nonlinear stochastic model 2 is larger 0.01954, that is, about 0.04% larger than Dr. Shi's k-th weighted moving average model.

Knowing the performance of nonlinear stochastic models with constant coefficients, we present Tables 4.3.4, 4.3.5 and 4.3.6 for remaining six models, namely, ARIMA model, k-th moving average model, k-th weighted moving average model, k-th exponential weighted moving average model, nonlinear stochastic models 1 and 2 with time varying coefficients. These tables contain the actual and forecasted values for three data sets.

Table 4.3.4 Actual and Predicted Price for Stock X

t	Actual Value	Predicted Value					
		ARIMA	k-th	k-th Weighted	k-th Exp. Weighted	Nonlinear Model 1	Nonlinear Model 2
848	685.19	690.5668	683.002	687.8663	688.95	690.5245	688.9735
849	685.33	684.6626	669.0717	678.8684	680.7026	684.642	684.7898
850	657	685.3417	675.2173	683.2993	684.0164	685.3423	685.6791
851	649.25	654.6294	634.8802	645.5674	648.7488	654.5775	655.224
852	631.68	648.5779	613.5021	636.6892	640.2909	648.5692	647.6274
853	653.2	630.1105	601.3531	622.4698	625.3764	630.1352	630.6804
854	646.73	654.9494	649.3253	655.5306	655.3039	655.0986	654.9312
855	638.25	646.2265	659.2641	650.889	649.8597	646.1798	647.0794
856	653.82	637.5815	630.9644	632.4086	633.6497	637.5274	635.8812
857	637.65	655.0054	655.3506	657.6908	656.6988	655.16	654.9123
858	615.95	636.504	639.1148	636.6878	637.1356	636.3224	637.4456
859	600.79	614.2723	587.4148	601.4316	605.362	614.1321	612.8671
860	600.25	599.5327	562.7062	588.4312	591.9794	599.4835	599.1939
861	584.35	600.2053	576.8988	595.1241	597.1468	600.2029	601.133
862	548.62	583.0321	566.9171	576.6775	579.0589	582.9814	583.1959
863	574.49	545.2863	505.6626	528.6499	533.91	545.4584	544.1784
864	566.4	576.4021	550.7285	573.7575	574.4719	576.695	574.846
865	555.98	565.8405	575.6357	571.1122	570.1095	565.7497	566.9464
866	550.52	555.2441	549.3242	548.4349	550.4871	555.1378	553.4377
867	548.27	550.1295	532.6924	545.7815	547.0418	550.0745	549.7962
868	564.3	548.1087	536.8028	545.8932	546.8686	548.0856	546.3439
869	515.9	565.4388	572.9851	569.0178	567.9075	565.6321	565.2503
870	495.43	513.1611	496.6788	502.3305	506.0323	512.4411	512.5852
871	506.8	493.965	434.953	471.6746	478.8219	493.7756	491.6517
872	501.71	507.562	478.7335	505.019	506.1324	507.7206	505.5978

Table 4.3.5 Actual and Predicted Price for Stock Y

t	Actual Value	Predicted Value					
		ARIMA	k-th	k-th Weighted	k-th Exp. Weighted	Nonlinear Model 1	Nonlinear Model 2
5631	83.8	84.82902	86.77076	85.95206	85.54406	84.87698	84.7722
5632	85.66	84.00733	87.0601	84.66872	84.41756	83.74178	83.90168
5633	85.05	85.66035	87.22635	86.23666	86.03169	85.51144	85.77717
5634	85.47	85.15029	86.28798	85.54065	85.43127	85.16966	85.14353
5635	92.57	85.52184	85.80557	85.64303	85.62554	85.30488	85.46407
5636	97	92.53652	97.42856	94.78931	93.93159	92.758	92.42399
5637	95.8	97.25259	107.4367	100.8205	99.56975	96.75397	97.02509
5638	94.62	96.03867	102.1023	97.29214	96.84044	95.48694	96.09939
5639	97.1	94.72032	94.26917	94.38129	94.52896	94.60442	95.00356
5640	94.95	97.146	97.17758	97.85285	97.66013	97.20943	97.29875
5641	89.07	95.14956	95.76419	94.89111	94.93671	94.90279	95.07715
5642	88.5	89.20548	84.0502	86.72816	87.58486	88.98043	89.02443
5643	86.79	88.50399	81.45928	86.72795	87.39757	88.74691	88.18962
5644	85.7	86.87246	83.05968	85.62383	85.99651	86.91702	86.35764
5645	86.7	85.75254	83.14882	84.53992	84.89636	85.66368	85.31255
5646	86.25	86.74345	85.69431	86.68557	86.72493	86.83622	86.48165
5647	85.38	86.34851	86.61938	86.425	86.38849	86.1994	86.26917
5648	85.94	85.45339	84.8209	84.94762	85.08256	85.36444	85.54056
5649	85.55	85.9905	85.31589	85.98491	86.02736	85.99153	86.15072
5650	85.73	85.63656	85.58075	85.61227	85.6082	85.54971	85.74118
5651	84.74	85.80065	85.76531	85.66177	85.67719	85.70833	85.81162
5652	84.75	84.82062	84.19112	84.5681	84.68931	84.76484	84.6882
5653	83.94	84.80058	83.82978	84.54765	84.62527	84.73972	84.61102
5654	84.15	84.01544	83.17927	83.60801	83.7308	83.99208	83.83675
5655	86.15	84.19971	83.51475	84.04115	84.10393	84.12795	84.05082

Table 4.3.6 Actual and Predicted Price for S&P 500 Index

t	Actual Value	Predicted Value					
		ARIMA	k-th	k-th Weighted	k-th Exp. Weighted	Nonlinear Model 1	Nonlinear Model 2
14845	931.8	900.8314	922.4829	909.6238	905.9727	902.5525	900.9227
14846	927.45	929.2406	970.9482	944.2084	939.3075	932.9931	929.0251
14847	934.7	925.1714	953.0097	932.0198	929.1136	925.6639	925.1775
14848	906.65	933.7874	944.5406	936.7494	935.8587	935.8555	933.9431
14849	909.73	906.9402	895.3576	899.2609	901.1965	904.6217	907.0545
14850	890.35	910.8315	889.0191	904.3632	906.3103	911.6803	911.0914
14851	870.26	891.0103	870.3993	884.6703	886.3622	888.6707	891.0485
14852	871.79	873.1138	843.0015	860.8976	865.2716	870.6269	873.2177
14853	842.62	873.4106	852.11	868.7196	870.5928	872.5018	873.1921
14854	843.74	844.7466	810.8334	832.4831	837.2048	840.9775	845.0018
14855	850.12	846.7019	835.1808	841.1104	843.4486	845.4405	846.3571
14856	805.22	850.2321	836.0299	848.9394	849.8067	849.7196	850.1915
14857	840.24	808.6526	790.8013	796.2332	801.0583	803.0576	808.5679
14858	827.5	841.9541	837.8833	841.4893	841.914	844.4779	841.5228
14859	831.95	825.9884	813.0403	827.1792	827.3513	824.1323	826.4057
14860	836.57	834.0853	855.8398	834.8218	834.2957	833.7654	833.6063
14861	845.71	836.0695	841.6322	838.2363	838.3141	835.7427	835.7924
14862	874.09	844.725	838.8276	847.1184	846.0498	846.1652	845.2083
14863	845.14	872.1378	915.9827	886.4861	881.3798	874.839	871.5874
14864	825.88	844.9611	849.1547	841.3719	842.5598	842.5651	845.1329
14865	825.44	828.8902	812.1273	818.3743	822.268	826.9521	828.9131
14866	838.51	825.9724	788.507	816.5074	819.2051	825.8489	826.2782
14867	832.23	837.8741	832.4872	841.8058	840.0885	839.0741	838.1304
14868	845.85	832.6112	852.194	834.7648	833.7577	831.2401	832.112
14869	868.6	845.546	846.6754	849.1995	848.8302	847.0508	845.5924

Then we compute the basic statistics for residual errors using different predicted models for the three data sets. Table 4.3.7, 4.3.8 and 4.3.9 contain the results.

Table 4.3.7 Basic Statistics by Using Different Predicted Models for Stock X

Stat.	ARIMA	k-th	k-th Weighted	k-th Exp. Weighted	Nonlinear Model 1	Nonlinear Model 2
Mean	6.979054	-9.14724	1.737706	3.528108	6.95156	6.495499
Var.	312.6866	868.8794	461.1024	406.0678	312.5214	327.0268
S.D.	17.68295	29.47676	21.4733	20.15112	17.67828	18.08388

Table 4.3.8 Basic Statistics by Using Different Predicted Models for Stock Y

Stat.	ARIMA	k-th	k-th Weighted	k-th Exp. Weighted	Nonlinear Model 1	Nonlinear Model 2
Mean	0.031855	0.247946	0.096411	0.067641	-0.0583	-0.0523
Var.	5.767871	16.25066	6.582993	5.9614	5.6971	5.741183
S.D.	2.401639	4.03121	2.565734	2.441598	2.38686	2.396077

Table 4.3.9 Basic Statistics by Using Different Predicted Modelsfor S&P 500 Index

Stat.	ARIMA	k-th	k-th Weighted	k-th Exp. Weighted	Nonlinear Model 1	Nonlinear Model 2
Mean	1.73382	-0.96296	0.579749	1.01513	1.362997	1.717323
Var.	406.8795	850.3763	463.5857	430.691	423.539	404.9443
S.D.	20.17126	29.16121	21.53104	20.7531	20.58006	20.12323

Table 4.3.10 Summary of Predictions for Three Data Sets

Data Set	ARIMA	k-th	k-th Weighted	k-th Exp. Weighted	Nonlinear Model 1	Nonlinear Model 2
Stock X	5	5	5	3	3	4
Stock Y	4	4	5	1	6	5
S&P 500	5	8	5	1	2	4

Remark 4.3.4: From Table 4.3.7, 4.3.8 and 4.3.9, we conclude that the statistics of using different model to predict a future value, the nonlinear stochastic model 1 with time varying coefficients model has less standard deviation of residual for two data sets stock X and stock Y. For S&P 500 Index, our nonlinear stochastic model 2 with time varying coefficients has the least standard deviation of residual. Table 4.3.10 summarizes the frequency of best model when predicting three data sets using different models. We further note that Table 4.3.10 summarizes the frequency of the best performance of models under three data sets predicted values. This summary in the context of Table 4.3.7, 4.3.8 and 4.3.9 suggests that nonlinear stochastic model 2 with time varying coefficients is robust with respect different data sets.

4.4 Conclusions

In Section 4.3, we studied prediction and comparison about the performance of presented and existing models. This was based on three overall data sets. So far the formulations of stochastic nonlinear Models 4.1.1 and 4.2.1 with time varying coefficient were utilized for the data fitting. We note that the performance of these models in the framework of data fitting is superior than the existing time series models nonlinear stochastic models 1 and 2 for overall data. Due to the nature of these models, the forecasting problem is open. This problem will be part of our future research plan.

Chapter 5
European Option Pricing

5.0 Introduction

In this chapter, we investigate the option pricing problem in the frame of nonlinear stochastic models described in Chapters 3 and 4. By employing the nonlinear stochastic models of stock price process, the formulas of option pricing are derived. In particular, we derive the European call and put option pricing formulas of nonlinear stochastic models 1, 2 and 3. These results are presented in Sections 5.1, 5.2 and 5.3.

5.1 European Option Pricing for Nonlinear Stochastic Model 1

The probabilistic approach to pricing options will result in a price expressed as the discounted expected value of a claim with respect to a probability measure. The solution process of stochastic differential equation in (1.3.1) is a stochastic process adapted to Brownian filtration $\{F_t\}_{t\geq 0}$. Under conditions (1.3.2) and (1.3.3) it has a unique solution process [23, 28]. We recall that (4.1.1) has a unique solution of equation (4.1.1).

The nonlinear stochastic model 1 (Section 4.1) with time varying coefficients, takes the following form

$$dS_t = (\alpha_t \ln S_t + \beta_t + \frac{\sigma_t^2}{2})S_t dt + \sigma_t S_t dW_t,\ S(0) = S_0,$$

where, coefficients α, β and σ are time varying smooth functions, and W_t is Brownian motion. In Section 4.1, by using the transformation $Y_t = \ln S_t$, equation (4.1.1) is transformed into linear form

$$dY_t = (\alpha_t Y_t + \beta_t)dt + \sigma_t dW_t.$$

The solution to this stochastic differential equation is

$$Y_t = \phi_t Y_{t_0} + \int_{t_0}^{t} \phi_{t,s} \beta_s ds + \int_{t_0}^{t} \phi_{t,s} \sigma_s dW_s \ ,$$

where $\phi_t = \exp(\int_{t_0}^{t} \alpha_s ds)$, and $\phi_{t,s} = \exp(\int_{s}^{t} \alpha_u du)$.

Then Y_t can be written as

$$Y_t = Y_{t_0} \exp(\int_{t_0}^{t} \alpha_s ds) + \int_{t_0}^{t} \left(\exp(\int_{s}^{t} \alpha_u du) \right) \beta_s ds + \int_{t_0}^{t} \left(\exp(\int_{s}^{t} \alpha_u du) \right) \sigma_s dW_s \ . \qquad (5.1.1)$$

Then by using the inverse transformation of "ln", we obtain

$$S_t = \exp\left(Y_{t_0} \exp(\int_{t_0}^{t} \alpha_s ds) + \int_{t_0}^{t} \left(\exp(\int_{s}^{t} \alpha_u du) \right) \beta_s ds + \int_{t_0}^{t} \left(\exp(\int_{s}^{t} \alpha_u du) \right) \sigma_s dW_s \right) \qquad (5.1.2)$$

Remark 5.1.1: For nonlinear stochastic model 1 with constant coefficients (3.1.1) and $t_0 = 0$, (5.1.1) and (5.1.2) reduce to

$$Y_t = \ln S_t = e^{\alpha t} Y_0 + \beta \int_{0}^{t} e^{\alpha(t-s)} ds + \sigma \int_{0}^{t} e^{\alpha(t-s)} dW_s$$

$$= e^{\alpha t} Y_0 + \frac{\beta}{\alpha}(e^{\alpha t} - 1) + \sigma e^{\alpha t} \int_{0}^{t} e^{-\alpha s} dW_s \qquad (5.1.3)$$

and

$$S_t = S_0^{e^{\alpha t}} e^{\frac{\beta}{\alpha}(e^{\alpha t} - 1) + \sigma e^{\alpha t} \int_0^t e^{-\alpha s} dW_s} \ , \qquad (5.1.4)$$

respectively.

Now, let V be the European option for a stock with respect to nonlinear stochastic model 1 with time varying coefficients (4.1.1). $V(S,t)$ is the value of the option at time t, where S_t is the stock price defined in (5.1.2). The strike price K and maturity time T are as defined in Section 1.1, and r is fixed interest rate. Applying to Theorem 1.1.2, we have

$$V(S,t) = E^Q[e^{-r(T-t)}C_T \mid F_t], \text{ where } C_T = \begin{cases} \max\{(S_T - K),0\}, & call \\ \max\{(K - S_T),0\}, & put \end{cases} \tag{5.1.5}$$

There is no a simple formula to compute the value of $V(S,t)$. To compute the numerical value of $V(S,t)$, we use equations (5.1.2) to simulate the value of S_T and then compute the expected value in (5.1.5).

From (5.1.1), knowing S_t, T and $Y_T = \ln S_T$, let $t_0 = t, S_{t_0} = S_t$ and $\theta = T - t$, for nonlinear stochastic model 1 with constant coefficients, we note that Y_T is normally distributed with

$$E[Y_T] = E\left(e^{\alpha(T-t)} \ln S_t + \frac{\beta}{\alpha}(e^{\alpha(T-t)} - 1) + \sigma \int_t^T e^{\alpha(T-s)} dW_s \right)$$

$$= e^{\alpha\theta} \ln S_t + \frac{\beta}{\alpha}(e^{\alpha\theta} - 1) + \sigma \int_t^T e^{\alpha(T-s)} E(dW_s)$$

$$= e^{\alpha\theta} \ln S_t + \frac{\beta}{\alpha}(e^{\alpha\theta} - 1),$$

and

$$Var(Y_T) = Var\left(e^{\alpha(T-t)} \ln S_t + \frac{\beta}{\alpha}(e^{\alpha(T-t)} - 1) + \sigma \int_t^T e^{\alpha(T-s)} dW_s \right)$$

$$= \sigma^2 \int_t^T e^{2\alpha(T-s)} Var(dW_s)$$

$$= \sigma^2 \int_t^T e^{2\alpha(T-s)} ds$$

$$= \frac{\sigma^2}{2\alpha}\left(e^{2\alpha\theta} - 1\right).$$

Thus, for European call option, (5.1.5) reduces to

$$V(S,t) = E^Q[e^{-r(T-t)}C_T \mid F_t] \quad = E^Q\left[e^{-r(T-t)} \left(e^{e^{\alpha\theta} \ln S_t + \frac{\beta}{\alpha}(e^{\alpha\theta} - 1) + \sqrt{\frac{\sigma^2}{2\alpha}\left(e^{2\alpha\theta} - 1\right)} Z} - K \right)_+ \right], \tag{5.1.6}$$

where Z is standard normal random variable.

First, we establish for the range of values of Z the integrand is non-zero.

$e^{e^{\alpha\theta}\ln S_t+\frac{\beta}{\alpha}(e^{\alpha\theta}-1)+\sqrt{\frac{\sigma^2}{2\alpha}\left(e^{2\alpha\theta}-1\right)}Z}-K>0$ is equivalent to $Z>\dfrac{\ln K-e^{\alpha\theta}\ln S_t-\frac{\beta}{\alpha}(e^{\alpha\theta}-1)}{\sqrt{\frac{\sigma^2}{2\alpha}\left(e^{2\alpha\theta}-1\right)}}$.

By letting $d=\dfrac{\ln K-e^{\alpha\theta}\ln S_t-\frac{\beta}{\alpha}(e^{\alpha\theta}-1)}{\sqrt{\frac{\sigma^2}{2\alpha}\left(e^{2\alpha\theta}-1\right)}}$, (5.1.6) reduces to

$$V(S,t)=\int_d^\infty e^{-r\theta}\left(e^{e^{\alpha\theta}\ln S_t+\frac{\beta}{\alpha}(e^{\alpha\theta}-1)+\sqrt{\frac{\sigma^2}{2\alpha}\left(e^{2\alpha\theta}-1\right)}Z}-K\right)\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}dz$$

$$=e^{-r\theta}\int_d^\infty e^{e^{\alpha\theta}\ln S_t+\frac{\beta}{\alpha}(e^{\alpha\theta}-1)+\sqrt{\frac{\sigma^2}{2\alpha}\left(e^{2\alpha\theta}-1\right)}Z}\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}dz-e^{-r\theta}K\int_d^\infty\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}dz$$

$$=e^{e^{\alpha\theta}\ln S_t+\frac{\beta}{\alpha}(e^{\alpha\theta}-1)-r\theta}\int_d^\infty\frac{e^{-\frac{z^2}{2}+\sqrt{\frac{\sigma^2}{2\alpha}\left(e^{2\alpha\theta}-1\right)}Z}}{\sqrt{2\pi}}dz-Ke^{-r\theta}\Phi(-d)$$

$$=e^{e^{\alpha\theta}\ln S_t+\frac{\beta}{\alpha}(e^{\alpha\theta}-1)-r\theta+\frac{\sigma^2}{4\alpha}\left(e^{2\alpha\theta}-1\right)}\int_d^\infty\frac{e^{-\frac{\left(z-\sqrt{\frac{\sigma^2}{2\alpha}\left(e^{2\alpha\theta}-1\right)}\right)^2}{2}}}{\sqrt{2\pi}}dz-Ke^{-r\theta}\Phi(-d)$$

$$=e^{e^{\alpha\theta}\ln S_t+\frac{\beta}{\alpha}(e^{\alpha\theta}-1)-r\theta+\frac{\sigma^2}{4\alpha}\left(e^{2\alpha\theta}-1\right)}\Phi\left(\sqrt{\frac{\sigma^2}{2\alpha}\left(e^{2\alpha\theta}-1\right)}-d\right)-Ke^{-r\theta}\Phi(-d). \qquad (5.1.7)$$

Similarly, the formula corresponding to a European put option is

$$V(S,t)=E^Q[e^{-r(T-t)}C_T\mid F_t]\quad=E^Q\left[e^{-r(T-t)}\left(K-e^{e^{\alpha\theta}\ln S_t+\frac{\beta}{\alpha}(e^{\alpha\theta}-1)+\sqrt{\frac{\sigma^2}{2\alpha}\left(e^{2\alpha\theta}-1\right)}Z}\right)_+\right], \qquad (5.1.8)$$

where Z is standard normal random variable.

Again, first we establish for the range of values of Z the integrand is non-zero.

$$K - e^{e^{\alpha\theta}\ln S_t + \frac{\beta}{\alpha}(e^{\alpha\theta}-1) + \sqrt{\frac{\sigma^2}{2\alpha}\left(e^{2\alpha\theta}-1\right)}Z} > 0 \text{ is equivalent to } Z < \frac{\ln K - e^{\alpha\theta}\ln S_t - \frac{\beta}{\alpha}(e^{\alpha\theta}-1)}{\sqrt{\frac{\sigma^2}{2\alpha}\left(e^{2\alpha\theta}-1\right)}} = d .$$

From the above discuss, (5.1.8) reduces to

$$V(S,t) = \int_{-\infty}^{d} e^{-r\theta}\left(K - e^{e^{\alpha\theta}\ln S_t + \frac{\beta}{\alpha}(e^{\alpha\theta}-1) + \sqrt{\frac{\sigma^2}{2\alpha}\left(e^{2\alpha\theta}-1\right)}z}\right)\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}dz$$

$$V(S,t) = Ke^{-r\theta}\Phi(d) - e^{e^{\alpha\theta}\ln S_t + \frac{\beta}{\alpha}(e^{\alpha\theta}-1) - r\theta + \frac{\sigma^2}{4\alpha}\left(e^{2\alpha\theta}-1\right)}\Phi\left(d - \sqrt{\frac{\sigma^2}{2\alpha}\left(e^{2\alpha\theta}-1\right)}\right). \qquad (5.1.9)$$

Illustration 5.1.1: In the following, we outline an illustration to exhibit the usefulness of the resented result. Suppose the yearly interest rate $r = 6.5\%$, by applying (5.1.7) and (5.1.9), the call and put option price are computed and recorded in Table 5.1.1 for three data sets. Similarly, the call and put option price of GBM model are computed and recorded in Table 5.1.2 for three data sets.

Table 5.1.1 Call and Put Option Price of Nonlinear Model 1

T	Stock X S_0=691.48 K=700		Stock Y S_0=84.84 K=90		S&P 500 Index S_0=903.25 K=910	
	call	put	call	put	call	put
20	18.17	32.71	2.96	7.76	13.56	17.27
60	27.53	52.99	6.63	10.76	27.60	25.29
100	30.55	65.45	9.21	12.71	38.22	29.96
200	30.75	83.54	13.83	15.93	59.61	36.77

Table 5.1.2 Call and Put Option Price of GBM Model

T	Stock X S_0=691.48 K=700		Stock Y S_0=84.84 K=90		S&P 500 Index S_0=903.25 K=910	
	call	put	call	put	call	put
20	23.18	29.25	2.96	7.81	13.64	17.21
60	44.42	45.63	6.69	10.92	27.91	25.16
100	59.71	56.09	9.39	12.99	38.80	29.77
200	89.00	73.46	14.48	16.54	61.05	36.52

5.2 European Option Pricing for Nonlinear Stochastic Model 2

The nonlinear stochastic model 2 (Section 4.2) with time varying coefficients, takes the following form

$$dS_t = (\alpha_t S_t + \beta_t S_t^N + \frac{N}{2}\sigma_t^2 S_t^{2N-1})dt + \sigma_t S_t^N dW_t,\ S(0) = S_0,$$

where, coefficients α, β and σ are time varying smooth functions, N is a constant $0 \le N \le 1.2, N \ne 1$, and W_t is Brownian motion. An argument about the existence and uniqueness of solutions of this equation can be reformulated.

In Section 4.2, using the transformation $Y_t = \dfrac{S_t^{1-N}}{1-N}$, equation (4.2.1) was transformed into linear form

$$dY_t = ((1-N)\alpha_t Y_t + \beta_t)dt + \sigma_t dW_t.$$

The solution to this stochastic differential equation is as follows

$$Y_t = \phi_t Y_{t_o} + \int_{t_0}^{t} \phi_{t,s}\beta_s ds + \int_{t_0}^{t} \phi_{t,s}\sigma_s dW_s,$$

where, $\phi_t = \exp\left(\int_{t_0}^{t} (1-N)\alpha_s ds\right)$, and $\phi_{t,s} = \exp\left(\int_{s}^{t} (1-N)\alpha_u du\right)$.

Then Y_t can be written as

$$Y_t = Y_{t_0} \exp\left(\int_{t_0}^{t}(1-N)\alpha_s ds\right) + \int_0^t \left(\exp\left(\int_s^t (1-N)\alpha_u du\right)\right)\beta_s ds + \int_{t_0}^{t}\left(\exp\left(\int_s^t (1-N)\alpha_u du\right)\right)\sigma_s dW_s$$

. (5.2.1)

Then by using the inverse transformation of $Y_t = \dfrac{S_t^{1-N}}{1-N}$

$$S_t = (1-N)$$

$$\left\{\frac{S_{t_0}^{1-N}}{1-N}\exp\left(\int_{t_0}^{t}(1-N)\alpha_s ds\right) + \int_{t_0}^{t}\left(\exp\left(\int_s^t (1-N)\alpha_u du\right)\right)\beta_s ds + \int_{t_0}^{t}\left(\exp\left(\int_s^t (1-N)\alpha_u du\right)\right)\sigma_s dW\right\}^{\frac{1}{1-N}} \tag{5.2.2}$$

Remark 5.2.1: For nonlinear stochastic model 2 with constant coefficients (3.2.1) and $t_0 = 0$, (5.2.1) and (5.2.2) reduce to

$$\begin{aligned} Y_t = \frac{S_t^{1-N}}{1-N} &= e^{(1-N)\alpha t}V_0 + \beta\int_0^t e^{(1-N)\alpha(t-s)}ds + \sigma\int_0^t e^{(1-N)\alpha(t-s)}dW_s \\ &= \frac{S_0^{1-N}}{1-N}e^{(1-N)\alpha t} + \frac{\beta}{\alpha(1-N)}(e^{(1-N)\alpha t}-1) + \sigma\int_0^t e^{(1-N)\alpha(t-s)}dW_s \end{aligned} \tag{5.2.3}$$

and

$$S_t = \left(S_0^{1-N}e^{(1-N)\alpha t} + \frac{\beta}{\alpha}(e^{(1-N)\alpha t}-1) + (1-N)\sigma\int_0^t e^{(1-N)\alpha(t-s)}dW_s\right)^{\frac{1}{1-N}}, \tag{5.2.4}$$

respectively.

Now, let V be the European option on a stock with respect to nonlinear stochastic model 2 with time varying coefficients. $V(S,t)$ is the value of the option at time t, where S_t is the stock price process defined in (5.2.2). The strike price K and maturity time T are as defined in Section 1.1. r is fixed interest rate. Applying to Theorem 1.1.2, we have

$$V(S,t) = E^Q[e^{-r(T-t)}C_T \mid F_t], \text{ where } C_T = \begin{cases}\max\{(S_T-K),0\}, & call \\ \max\{(K-S_T),0\}, & put\end{cases} \tag{5.2.5}$$

There is no a simple formula to compute the value of $V(S,t)$. To compute the numerical value of $V(S,t)$, we use equations (5.2.2) to simulate the value of S_T and then compute the expected value in (5.2.5).

For nonlinear stochastic model 2 with constant coefficients, from (5.2.1), knowing S_t and T, letting $t_0 = t, S_{t_0} = S_t$ and $\theta = T - t$, we note that Y_T is normally distributed with

$$E[Y_T] = E\left(e^{(1-N)\alpha(T-t)} \frac{S_t^{1-N}}{1-N} + \frac{\beta}{(1-N)\alpha}(e^{(1-N)\alpha(T-t)} - 1) + \sigma \int_t^T e^{(1-N)\alpha(T-s)} dW_s \right)$$

$$= e^{(1-N)\alpha\theta} \frac{S_t^{1-N}}{1-N} + \frac{\beta}{(1-N)\alpha}(e^{(1-N)\alpha\theta} - 1) + E\left(\sigma \int_t^T e^{(1-N)\alpha(T-s)} dW_s \right)$$

$$= e^{(1-N)\alpha\theta} \frac{S_t^{1-N}}{1-N} + \frac{\beta}{(1-N)\alpha}(e^{(1-N)\alpha\theta} - 1),$$

and

$$Var(Y_T) = Var\left(e^{(1-N)\alpha(T-t)} \frac{S_t^{1-N}}{1-N} + \frac{\beta}{(1-N)\alpha}(e^{(1-N)\alpha(T-t)} - 1) + \sigma \int_t^T e^{(1-N)\alpha(T-s)} dW_s \right)$$

$$= \sigma^2 \int_t^T e^{2(1-N)\alpha(T-s)} ds$$

$$= \frac{\sigma^2}{2(1-N)\alpha}(e^{2(1-N)\alpha\theta} - 1).$$

Hence, for European call option, (5.2.5) reduces to

$$V(S,t) = E^Q[e^{-r(T-t)} C_T \mid F_t]$$

$$= E^Q\left[e^{-r\theta} \left(\left(S_t^{1-N} e^{(1-N)\alpha\theta} + \frac{\beta}{\alpha}(e^{(1-N)\alpha\theta} - 1) + (1-N)\sqrt{\frac{\sigma^2}{2(1-N)\alpha}} \left(e^{2(1-N)\alpha\theta} - 1\right) Z \right)^{\frac{1}{1-N}} - K \right)_+ \right], \tag{5.2.6}$$

where Z is standard normal random variable.

First, we establish for the range of values of Z the integrand is non-zero.

$$\left(S_t^{1-N}e^{(1-N)\alpha\theta}+\frac{\beta}{\alpha}(e^{(1-N)\alpha\theta}-1)+(1-N)\sqrt{\frac{\sigma^2}{2(1-N)\alpha}}\left(e^{2(1-N)\alpha\theta}-1\right)Z\right)^{\frac{1}{1-N}}-K>0$$

is equivalent to $Z>\dfrac{K^{1-N}-S_t^{1-N}e^{(1-N)\alpha\theta}-\dfrac{\beta}{\alpha}(e^{(1-N)\alpha\theta}-1)}{(1-N)\sqrt{\dfrac{\sigma^2}{2(1-N)\alpha}\left(e^{2(1-N)\alpha\theta}-1\right)}}$.

Setting $d=\dfrac{K^{1-N}-S_t^{1-N}e^{(1-N)\alpha\theta}-\dfrac{\beta}{\alpha}(e^{(1-N)\alpha\theta}-1)}{(1-N)\sqrt{\dfrac{\sigma^2}{2(1-N)\alpha}\left(e^{2(1-N)\alpha\theta}-1\right)}}$, (5.2.6) reduces to

$$V(S,t)$$

$$=\int_d^\infty e^{-r\theta}\left(\left(S_t^{1-N}e^{(1-N)\alpha\theta}+\frac{\beta}{\alpha}(e^{(1-N)\alpha\theta}-1)+(1-N)\sqrt{\frac{\sigma^2}{2(1-N)\alpha}}\left(e^{2(1-N)\alpha\theta}-1\right)Z\right)^{\frac{1}{1-N}}-K\right)\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}dz$$

$$=e^{-r\theta}\int_d^\infty\left(S_t^{1-N}e^{(1-N)\alpha\theta}+\frac{\beta}{\alpha}(e^{(1-N)\alpha\theta}-1)+(1-N)\sqrt{\frac{\sigma^2}{2(1-N)\alpha}}\left(e^{2(1-N)\alpha\theta}-1\right)Z\right)^{\frac{1}{1-N}}\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}dz$$

$$-Ke^{-r\theta}\Phi(-d) \tag{5.2.7}$$

Similarly, the formula corresponding to a European put option is

$$V(S,t)=E^Q[e^{-r(T-t)}C_T\mid F_t]$$

$$=E^Q\left[e^{-r\theta}\left(\left(K-S_t^{1-N}e^{(1-N)\alpha\theta}+\frac{\beta}{\alpha}(e^{(1-N)\alpha\theta}-1)+(1-N)\sqrt{\frac{\sigma^2}{2(1-N)\alpha}}\left(e^{2(1-N)\alpha\theta}-1\right)Z\right)^{\frac{1}{1-N}}\right)_+\right], \tag{5.2.8}$$

where Z is standard normal random variable.

We establish for the range of values of Z the integrand is non-zero.

$$K-\left(S_t^{1-N}e^{(1-N)\alpha\theta}+\frac{\beta}{\alpha}(e^{(1-N)\alpha\theta}-1)+(1-N)\sqrt{\frac{\sigma^2}{2(1-N)\alpha}}\left(e^{2(1-N)\alpha\theta}-1\right)Z\right)^{\frac{1}{1-N}}>0$$

is equivalent to $Z < \dfrac{K^{1-N} - S_t^{1-N} e^{(1-N)\alpha\theta} - \dfrac{\beta}{\alpha}(e^{(1-N)\alpha\theta} - 1)}{(1-N)\sqrt{\dfrac{\sigma^2}{2(1-N)\alpha}\left(e^{2(1-N)\alpha\theta} - 1\right)}} = d$.

$V(S,t)$

$$= \int_{-\infty}^{d} e^{-r\theta} \left(K - \left(S_t^{1-N} e^{(1-N)\alpha\theta} + \frac{\beta}{\alpha}(e^{(1-N)\alpha\theta} - 1) + (1-N)\sqrt{\frac{\sigma^2}{2(1-N)\alpha}\left(e^{2(1-N)\alpha\theta} - 1\right)} Z \right)^{\frac{1}{1-N}} \right) \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} dz$$

$$= Ke^{-r\theta}\Phi(d) -$$

$$e^{-r\theta} \int_{-\infty}^{d} \left(S_t^{1-N} e^{(1-N)\alpha\theta} + \frac{\beta}{\alpha}(e^{(1-N)\alpha\theta} - 1) + (1-N)\sqrt{\frac{\sigma^2}{2(1-N)\alpha}\left(e^{2(1-N)\alpha\theta} - 1\right)} Z \right)^{\frac{1}{1-N}} \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} dz$$

(5.2.9)

Illustration 5.2.1: In the following, we illustration the usefulness of the above presented result. In (5.2.7) and (5.2.9), α, β, σ and N are estimated from observations. Other parameters such as the yearly interest rate is set to $r = 6.5\%$, by applying (5.2.7) and (5.2.9), the call and put option price are computed and recorded in Table 5.2.1 for three data sets.

Table 5.2.1 Call and Put Option Price of Nonlinear Model 2

T	Stock X S_0=691.48 K=700		Stock Y S_0=84.84 K=90		S&P 500 Index S_0=903.25 K=910	
	call	put	call	put	call	put
20	13.41	13.37	0.13	3.46	9.17	15.67
60	31.89	15.32	2.06	1.65	17.98	23.99
100	47.57	15.26	5.13	0.86	24.00	29.53
200	81.73	13.25	14.61	0.20	34.77	39.12

5.3 European Option Pricing for Nonlinear Stochastic Model 3

Similarly, the nonlinear stochastic model 3, with time varying coefficients, takes the following form

$$dS_t = (\alpha_t S_t + \beta_t S_t^2 + \sigma_t^2 S_t)dt + \sigma_t S_t dW_t\,,\ S(0) = S_0\,, \tag{5.3.1}$$

where, coefficients α, β and σ are time varying smooth functions, and W_t is Brownian motion. We note that the existence and uniqueness of solution process of (5.3.1) is justied in Section 3.3.

In Section 3.3 (Chapter 3), by using the transformation $Y_t = -S_t^{-1}$, equation (5.3.1) was transformed into linear form

$$dY_t = (-\alpha_t Y_t + \beta_t)dt - \sigma_t Y_t dW_t\,.$$

The solution to this stochastic differential equation is

$$Y_t = \phi_t Y_{t_0} + \int_{t_0}^{t} \phi_{t,s}\beta_s ds\,,$$

where, $\phi_t = \exp\left(\int_{t_0}^{t} (-\alpha_s - \frac{1}{2}\sigma_s^2)ds + \int_{t_0}^{t} \sigma_s dW_s \right)$,

and $\phi_{t,s} = \exp\left(\int_{s}^{t} (-\alpha_u - \frac{1}{2}\sigma_u^2)du - \int_{s}^{t} \sigma_u dW_u \right)$.

Then Y_t can be written as

$$Y_t = \frac{Y_{t_0}}{\exp\left(\int_{t_0}^{t} (\alpha_s + \frac{1}{2}\sigma_s^2)ds + \int_{t_0}^{t} \sigma_s dW_s \right)} + \int_{t_0}^{t} \frac{\beta_s}{\exp\left(\int_{s}^{t} (-\alpha_u - \frac{1}{2}\sigma_u^2)du - \int_{s}^{t} \sigma_u dW_u \right)} ds \tag{5.3.2}$$

Then by using the inverse transformation of $Y_t = -S_t^{-1}$

$$S_t = \left(\frac{Y_{t_0}}{\exp\left(\int_{t_0}^{t} (\alpha_s + \frac{1}{2}\sigma_s^2)ds + \int_{t_0}^{t} \sigma_s dW_s \right)} + \int_{t_0}^{t} \frac{\beta_s}{\exp\left(\int_{s}^{t} (-\alpha_u - \frac{1}{2}\sigma_u^2)du - \int_{s}^{t} \sigma_u dW_u \right)} ds \right)^{-1} \tag{5.3.3}$$

Remark 5.3.1: For nonlinear stochastic model 3 with constant coefficients (3.3.1) and $t_0 = 0$, (5.3.2) and (5.3.3) reduce to

$$Y_t = -S_t^{-1} = Y_0 e^{(-\alpha - \frac{\sigma^2}{2})t - \sigma\int_0^t dW_s} + \beta\int_0^t e^{(-\alpha - \frac{\sigma^2}{2})(t-s) - \sigma\int_s^t dW_\mu} ds$$

$$= e^{-(\alpha + \frac{\sigma^2}{2})t - \sigma W_t}\left(Y_0 + \beta\int_0^t e^{(\alpha + \frac{\sigma^2}{2})s + \sigma W(s)} ds \right) \qquad (5.3.4)$$

and

$$S_t = \frac{e^{(\alpha + \frac{\sigma^2}{2})t + \sigma W_t}}{\frac{1}{S_0} - \beta\int_0^t e^{(\alpha + \frac{\sigma^2}{2})s + \sigma W(s)} ds}, \qquad (5.3.5)$$

respectively.

Similarly, let V be the European option on a stock with respect to nonlinear stochastic model 3 with time varying coefficients. $V(S,t)$ is the value of the option at time t, where S_t is the stock price process defined in (5.3.3). The strike price K and maturity time T are as defined in Section 1.1. r is fixed interest rate. Applying to Theorem 1.1.2, we have

$$V(S,t) = E^Q[e^{-r(T-t)} C_T \mid F_t], \text{ where } C_T = \begin{cases} \max\{(S_T - K), 0\}, & call \\ \max\{(K - S_T), 0\}, & put \end{cases} \qquad (5.3.6)$$

There is no a simple formula to compute the value of $V(S,t)$.

Illustration 5.3.1: In the following, we illustrate the usefulness of the above presented result. To compute the numerical value of $V(S,t)$, we use equations (5.3.3) or (5.3.5) to simulate the value of S_T and then compute the expected value in (5.3.6). Suppose the yearly interest rate $r = 6.5\%$, the call and put option price are computed and recorded in Table 5.3.1 for three data sets.

Table 5.3.1 Call and Put Option Price of Nonlinear Model 3

T	Stock X S_0 =691.48 K=600		Stock Y S_0 =84.84 K=70		S&P 500 Index S_0 =903.25 K=800	
	call	put	call	put	call	put
5	58.92	5.18	16.16	0.56	101.56	0.02
10	2.73	65.75	15.39	0.58	92.16	0.06
20	0	286.2	11.57	1.03	56.24	0.91
60	0	525.13	0.11	17.25	0	210.66
100	0	556.44	0	38.04	0	431.09

Chapter 6
Option Pricing for Hybrid Models

6.0 Introduction

We studied GBM models and nonlinear stochastic models under the different data partitioning processes in Chapter 2, 3 and 4. In Chapter 5, we derived the European option pricing formulas for three nonlinear stochastic models, and apply to three data sets. In this chapter, we first derive the European call and put option pricing formulas in Section 6.1 for Hybrid GBM Models. In Section 6.2, we present option pricing formulas for hybrid nonlinear stochastic models.

6.1 Option Pricing for Hybrid GBM Models

In 2003, G.Yin, et proposed a hybrid GBM model (HGBM). In HGBM model, drift and volatility are not deterministic functions anymore. They are perturbed by stochastic process such as a Markov Chain.
By following development of a class of stochastic hybrid GBM system [16,44]:

$$
\begin{aligned}
& dS = \mu(t,\eta(t))Sdt + \sigma(t,\eta(t))SdW_t, S(t_k) = S_k, t \neq t_k, \\
& S_k = G(S_{k-1}(t_k^-, t_{k-1}, S_{k-1}, \eta_{k-1}), \eta_k), \ \ S(t_0) = S_0, \\
& \eta_{k+1} = M(S, \eta_k), \eta(t_0) = \eta_0, k \in I(1,\infty),
\end{aligned}
\tag{6.1.1}
$$

where, S is a continuous price of the stock, $\mu(t,\eta(t))$ and $\sigma(t,\eta(t))$ are drift and volatility governed by the underlying discrete events that can be modeled by a stochastic process $\eta(t)$ with a finite state. Figure 6.1.1 illustrate system switching from state k to state k+1 when at time t_k, event occurs, a jump $\phi_{\eta(t)}$ also occurs here.

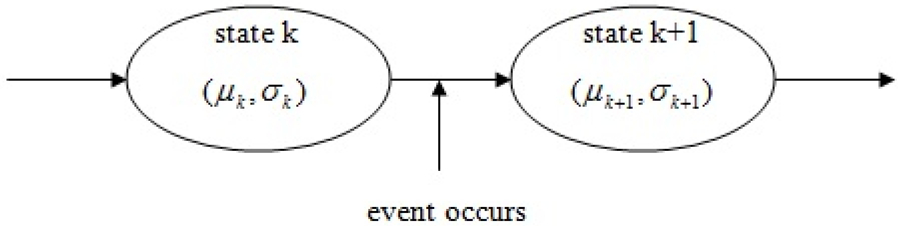


Figure 6.1.1 State Switch Illustration of Hybrid GBM

In Chapter 2, we develop several modified GBM models which are HGBM models. The solution process of these HGBM models takes the general form:

$$
=\begin{cases}
S_0 e^{(\mu_1-\frac{1}{2}\sigma_1^2)t+\sigma_1 W_t} & 0 \le t < t_1 \\
\phi_1 S_0 e^{(\mu_1-\frac{1}{2}\sigma_1^2)t_1+\sigma_1 W_t+(\mu_2-\frac{1}{2}\sigma_2^2)(t-t_1)+\sigma_2(W_t-W_{t_1})} & t_1 \le t < t_2 \\
\ldots & \ldots \\
\phi_{m-1}\ldots\phi_1 S_0 e^{(\mu_{m-1}-\frac{1}{2}\sigma_{m-1}^2)(t-t_{m-1})+\sigma_{m-1}(W_t-W_{t_{m-1}})+\ldots+\sigma_1 W_t+(\mu_m-\frac{1}{2}\sigma_m^2)(t_1-0)+\sigma_m(W_{t_1}-0)} & t_{m-1} \le t < t_m
\end{cases}
$$

Let $\Delta t_1 = t_1 - 0, \Delta t_2 = t_2 - t_1, \ldots, \Delta t_{m-1} = t_{m-1} - t_{m-2}, \Delta t_m = t - t_m$, $\Delta W_1 = W_{t_1} - W_0, \Delta W_2 = W_{t_2} - W_{t_1}, \ldots, \Delta W_{m-1} = W_{t_{m-1}} - W_{t_{m-2}}, \Delta W_m = W_t - W_{t_{m-1}}$, the price process is represented as following

$$
S_t = \begin{cases}
S_0 e^{(\mu_1-\frac{1}{2}\sigma_1^2)t+\sigma_1 W_t} & 0 \le t < t_1 \\
\phi_1 S_0 e^{(\mu_1-\frac{1}{2}\sigma_1^2)\Delta t_1+\sigma_1 \Delta W_1} e^{(\mu_2-\frac{1}{2}\sigma_2^2)(t-t_1)+\sigma_2(W_t-W_{t_1})} & t_1 \le t < t_2 \\
\ldots & \ldots \\
S_0 \prod_{i=1}^{m-1}\phi_i \left(e^{(\mu_m-\frac{1}{2}\sigma_m^2)(t-t_{m-1})+\sigma_m(W_t-W_{t_{m-1}})+\sum_{i=1}^{m-1}\left((\mu_i-\frac{1}{2}\sigma_i^2)\Delta t_i+\sum_{i=1}^{m}\sigma_i \Delta W_i\right)} \right) & t_{m-1} \le t < t_m
\end{cases} \quad (6.1.2)
$$

where, $[0,t_1),[t_1,t_2),[t_2,t_3), [t_3,t_4),[t_4,t_5)\ldots[t_{m-1},t_m)$ be any one of data partition processes which are defined in Chapter 2.

Now, let V be the European option on a stock with respect to hybrid GBM model (6.1.1). $V(S,t)$ is the value of the option at time t, where S_t is the stock price process defined in (6.1.2). The strike price K and maturity time T are as defined in Section 1.1. r is fixed interest rate. Applying to Theorem 1.1.2, we have

$$V(S,t) = E^Q[e^{-r(T-t)}C_T \mid F_t], \text{ where } C_T = \begin{cases} \max\{(S_T - K),0\}, & call \\ \max\{(K - S_T),0\}, & put \end{cases} \tag{6.1.3}$$

For hybrid GBM model, from (6.1.3), by knowing S_t and $\theta = T - t$, we note that

$$Y_T = \ln\left(\frac{S_T}{S_0 \prod_{i=1}^{m-1} \phi_i}\right) \text{ is normal distributed with}$$

$$E[Y_T] = E\left((\mu_m - \frac{1}{2}\sigma_m^2)(T - t_{m-1}) + \sigma_m (W_t - W_{t_{m-1}}) + \sum_{i=1}^{m-1}\left((\mu_i - \frac{1}{2}\sigma_i^2)\Delta t_i + \sum_{i=1}^{m}\sigma_i W_i \right)\right)$$

$$= (\mu_m - \frac{1}{2}\sigma_m^2)(T - t_{m-1}) + \sum_{i=1}^{m-1}\left((\mu_i - \frac{1}{2}\sigma_i^2)\Delta t_i \right) = \sum_{i=1}^{m}\left((\mu_i - \frac{1}{2}\sigma_i^2)\Delta t_i \right),$$

and

$$Var(Y_T) = Var\left((\sum_{i=1}^{m}(\mu_i - \frac{1}{2}\sigma_i^2)\Delta t_i + \sum_{i=1}^{m}\sigma_i \Delta W_i \right)$$

$$= \sigma_m^2 (T - t_{m-1}) + \sum_{i=1}^{m-1}\sigma_i^2 \Delta t_i = \sum_{i=1}^{m}\sigma_i^2 \Delta t_i \ .$$

Hence, for European call option, (6.1.3) reduces to

$$V(S,t) = E^Q[e^{-r(T-t)}C_T \mid F_t]$$

$$= E^Q\left[e^{-r(T-t)}\left(S_0 \prod_{i=1}^{m-1}\phi_i e^{\sum_{i=1}^{m}(\mu_i - \frac{1}{2}\sigma_i^2)\Delta t_i + \sqrt{\sum_{i=1}^{m}\sigma_i^2 \Delta t_i}\, Z} - K \right)_+ \right], \tag{6.1.4}$$

where Z is standard normal random variable.

First, we establish for the range of values of Z the integrand is non-zero.

$$S_0\prod_{i=1}^{m-1}\phi_i e^{\sum_{i=1}^{m}(\mu_i-\frac{1}{2}\sigma_i^2)\Delta t_i+\sqrt{\sum_{i=1}^{m}\sigma_i^2\Delta t_i}Z}-K>0 \text{ is equivalent to } Z>\frac{\ln\frac{K}{S_0\prod_{i=1}^{m-1}\phi_i}-\sum_{i=1}^{m}(\mu_i-\frac{1}{2}\sigma_i^2)\Delta t_i}{\sqrt{\sum_{i=1}^{m}\sigma_i^2\Delta t_i}}.$$

Let $d=\dfrac{\ln\frac{K}{S_0\prod_{i=1}^{m-1}\phi_i}-\sum_{i=1}^{m}(\mu_i-\frac{1}{2}\sigma_i^2)\Delta t_i}{\sqrt{\sum_{i=1}^{m}\sigma_i^2\Delta t_i}}$, (6.1.4) reduces to

$$V(S,t)=\int_d^\infty e^{-r\theta}\left(S_0\prod_{i=1}^{m-1}\phi_i e^{\sum_{i=1}^{m}(\mu_i-\frac{1}{2}\sigma_i^2)\Delta t_i+\sqrt{\sum_{i=1}^{m}\sigma_i^2\Delta t_i}Z}-K\right)\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}dz$$

$$=e^{-r\theta}\int_d^\infty S_0\prod_{i=1}^{m-1}\phi_i e^{\sum_{i=1}^{m}(\mu_i-\frac{1}{2}\sigma_i^2)\Delta t_i+\sqrt{\sum_{i=1}^{m}\sigma_i^2\Delta t_i}Z}\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}dz-e^{-r\theta}K\int_d^\infty\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}dz$$

$$=e^{\sum_{i=1}^{m}(\mu_i-\frac{1}{2}\sigma_i^2)\Delta t_i-r\theta}S_0\prod_{i=1}^{m-1}\phi_i\int_d^\infty\frac{e^{-\frac{z^2}{2}+\sqrt{\sum_{i=1}^{m}\sigma_i^2\Delta t_i}Z}}{\sqrt{2\pi}}dz-Ke^{-r\theta}\Phi(-d)$$

$$=e^{\sum_{i=1}^{m}(\mu_i-\frac{1}{2}\sigma_i^2)\Delta t_i+\sum_{i=1}^{m}\sigma_i^2\Delta t_i-r\theta}S_0\prod_{i=1}^{m-1}\phi_i\int_d^\infty\frac{e^{-\frac{(z-\sqrt{\sum_{i=1}^{m}\sigma_i^2\Delta t_i})^2}{2}}}{\sqrt{2\pi}}dz-Ke^{-r\theta}\Phi(-d)$$

$$=e^{\sum_{i=1}^{m}(\mu_i+\frac{1}{2}\sigma_i^2)\Delta t_i-r\theta}S_0\left(\prod_{i=1}^{m-1}\phi_i\right)\Phi(\sqrt{\sum_{i=1}^{m}\sigma_i^2\Delta t_i}-d)-Ke^{-r\theta}\Phi(-d). \tag{6.1.5}$$

Similarly, the corresponding formula for a European put option is

$$V(S,t)=E^Q[e^{-r(T-t)}C_T\mid F_t]$$

$$=E^Q\left[e^{-r(T-t)}\left(K-S_0\prod_{i=1}^{m-1}\phi_i e^{\sum_{i=1}^{m}(\mu_i-\frac{1}{2}\sigma_i^2)\Delta t_i+\sqrt{\sum_{i=1}^{m}\sigma_i^2\Delta t_i}Z}\right)_+\right], \tag{6.1.6}$$

where Z is standard normal random variable.

First, we establish for the range of values of Z the integrand is non-zero.

$$K - S_0 \prod_{i=1}^{m-1} \phi_i e^{\sum_{i=1}^{m}(\mu_i - \frac{1}{2}\sigma_i^2)\Delta t_i + \sqrt{\sum_{i=1}^{m}\sigma_i^2 \Delta t_i} Z} > 0$$

is equivalent to $Z < \dfrac{\ln \dfrac{K}{S_0 \prod_{i=1}^{m-1} \phi_i} - \sum_{i=1}^{m}(\mu_i - \frac{1}{2}\sigma_i^2)\Delta t_i}{\sqrt{\sum_{i=1}^{m}\sigma_i^2 \Delta t_i}} = d$.

(6.1.6) reduces to

$$V(S,t) = \int_{-\infty}^{d} e^{-r\theta} \left(K - S_0 \prod_{i=1}^{m-1} \phi_i e^{\sum_{i=1}^{m}(\mu_i - \frac{1}{2}\sigma_i^2)\Delta t_i + \sqrt{\sum_{i=1}^{m}\sigma_i^2 \Delta t_i} Z} \right) \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} dz$$

$$V(S,t) = Ke^{-r\theta}\Phi(d) - e^{\sum_{i=1}^{m}(\mu_i - \frac{1}{2}\sigma_i^2)\Delta t_i - r\theta} S_0 \left(\prod_{i=1}^{m-1} \phi_i \right) \Phi(d - \sqrt{\sum_{i=1}^{m}\sigma_i^2 \Delta t_i}) . \tag{6.1.7}$$

Illustration 6.1.1: In the following, we illustrate the usefulness of the above presented result. Suppose the yearly interest rate $r = 6.5\%$, by applying (6.1.5) and (6.1.7), the call and put option price of three data sets are computed and recorded in Tables 6.1.1 and 6.1.2 for Hybrid GBM models 2.4.3 and 2.4.4, respectively.

Table 6.1.1 Call and Put Option Price of Hybrid GBM Model 2.4.3

T	Stock X S_0=691.48 K=700		Stock Y S_0=84.84 K=90		S&P 500 Index S_0=903.25 K=910	
	call	put	call	put	call	put
20	22.19	12.99	2.06	5.30	11.57	106.96
60	95.38	12.97	5.99	6.04	34.26	130.94
100	165.04	8.68	9.56	6.32	50.87	148.54
200	373.42	3.76	18.20	6.37	80.60	180.52

Table 6.1.2 Call and Put Option Price of Hybrid GBM Model 2.4.4

T	Stock X S_0=691.48 K=700		Stock Y S_0=84.84 K=90		S&P 500 Index S_0=903.25 K=910	
	call	put	call	put	call	put
20	28.75	15.66	4.71	4.07	44.70	48.50
60	107.67	8.25	8.25	5.35	51.33	87.26
100	185.16	6.60	11.96	5.86	64.30	106.90
200	486.84	1.34	20.79	6.37	81.72	136.09

6.2 Option Pricing for Hybrid Nonlinear Stochastic Models

By following development of a class of stochastic hybrid dynamic system [16]:

$$dS = F_0(t,S,\eta(t))dt + F_e(t,S,\eta(t))dW_t, S(t_k) = S_k, t \neq t_k,$$
$$S_k = G(S_{k-1}(t_k^-, t_{k-1}, S_{k-1}, \eta_{k-1}), \eta_k), \; S(t_0) = S_0,$$
$$\eta_{k+1} = M(S,\eta_k), \eta(t_0) = \eta_0, k \in I(1,\infty), \tag{6.2.1}$$

where,

S is a continuous price of the stock,

for nonlinear model 1, $F_0 = (\alpha_{\eta(t)} \ln S_t + \beta_{\eta(t)} + \frac{\sigma^2_{\eta(t)}}{2})S_t$, and $F_e = \sigma_{\eta(t)} S_t$,

for nonlinear model 2, $F_0 = \alpha_{\eta(t)} S_t + \beta_{\eta(t)} S_t^N + \frac{N}{2}\sigma^2_{\eta(t)} S_t^{2N-1}$, and $F_e = \sigma_{\eta(t)} S_t^N$,

and

for nonlinear model 3, $F_0 = \alpha_{\eta(t)} S_t + \beta_{\eta(t)} S_t^2 + \sigma^2_{\eta(t)} S_t$, and $F_e = \sigma_{\eta(t)} S_t$,

$F_0(t,S,\eta(t))$ and $F_e(t,S,\eta(t))$ are governed by the underlying discrete events that can be modeled by a stochastic process $\eta(t)$ with a finite state.

Figure 6.2.1 illustrate system switching from state k to state k+1 when at time t_k, event occurs, a jump $\phi_{\eta(t)}$ also occurs here.

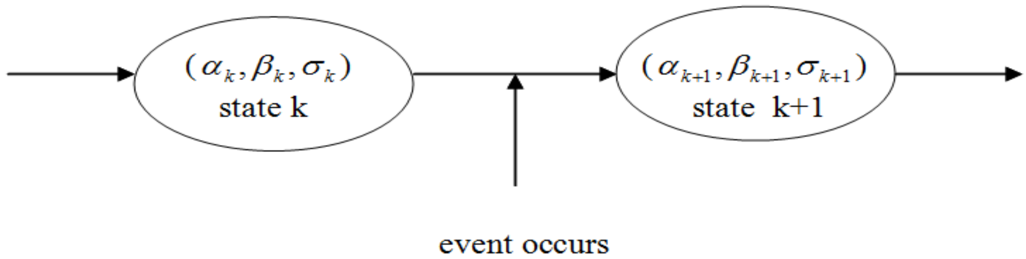


Figure 6.2.1 State Switch Illustration of Hybrid Nonlinear Stochastic Model

In Chapter 3, we develop several nonlinear stochastic models which are hybrid nonlinear stochastic models. The solution process of these hybrid stochastic models takes the following form:

$$S_t = \begin{cases} S_1(t, t_0, S_0) & S_0 = S_0, & 0 \le t < t_1 \\ \phi_1 S_2(t, t_1, S_1) & S_1 = \lim\limits_{t \to t_1^-} S_t, & t_1 \le t < t_2 \\ \ldots & \ldots & \ldots \\ \phi_{m-1} S_m(t, t_{m-1}, S_{m-1}) & S_{m-1} = \lim\limits_{t \to t_{m-1}^-} S_t, & t_{m-1} \le t < t_m \end{cases} \tag{6.2.2}$$

where, $[0, t_1), [t_1, t_2), [t_2, t_3), [t_3, t_4), [t_4, t_5) \ldots [t_{m-1}, t_m)$ be any one of data partition processes which are defined in Chapter 2.

6.2.1 Hybrid Nonlinear Stochastic Model 1

The solution process of the hybrid nonlinear stochastic models 3.1.1 and 3.1.2 is given by,

$$S_t = \begin{cases} S_0^{e^{\alpha_1 t}} e^{\frac{\beta_1}{\alpha_1}(e^{\alpha_1 t} - 1) + \sigma_1 \int_0^t e^{\alpha_1 (t-s)} dW_s} & S_0 = S_0, & 0 \le t < t_1 \\ \phi_1 S_1^{e^{\alpha_2 (t-t_1)}} e^{\frac{\beta_2}{\alpha_2}(e^{\alpha_2 (t-t_1)} - 1) + \sigma_2 \int_{t_1}^t e^{\alpha_2 (t-s)} dW_s} & S_1 = \lim\limits_{t \to t_1^-} S_t, & t_1 \le t < t_2 \\ \ldots & \ldots & \ldots \\ \phi_{m-1} S_{m-1}^{e^{\alpha_m (t-t_{m-1})}} e^{\frac{\beta_m}{\alpha_m}(e^{\alpha_m (t-t_{m-1})} - 1) + \sigma_m \int_{t_{m-1}}^t e^{\alpha_m (t-s)} dW_s} & S_{m-1} = \lim\limits_{t \to t_{m-1}^-} S_t, & t_{m-1} \le t < t_m \end{cases}, \tag{6.2.3}$$

Recursively, we have

$$\ln S_t = e^{\alpha_m (t-t_{m-1})+\alpha_{m-1}\Delta t_{m-1}+...+\alpha_1 \Delta t_1} \ln S_0$$

$$+\ln\phi_{m-1} + e^{\alpha_m (t-t_{m-1})} \ln\phi_{m-2} + e^{\alpha_m (t-t_{m-1})+\alpha_{m-1}\Delta t_{m-1}} \ln\phi_{m-3} + ... + e^{\alpha_m (t-t_{m-1})+\alpha_{m-1}\Delta t_{m-1}+...+\alpha_3 \Delta t_3} \ln\phi_1$$

$$+\frac{\beta_m}{\alpha_m}(e^{\alpha_m (t-t_{m-1})}-1) + \frac{\beta_{m-1}}{\alpha_{m-1}} e^{\alpha_m (t-t_{m-1})}(e^{\alpha_{m-1}\Delta t_{m-1}}-1) + ... + \frac{\beta_1}{\alpha_1} e^{\alpha_m (t-t_{m-1})+\alpha_{m-1}\Delta t_{m-1}+...+\alpha_2 \Delta t_2}(e^{\alpha_1 \Delta t_1}-1)$$

$$+\sigma_m \int_{t_{m-1}}^{t} e^{\alpha_m (t-s)} dW_s + \sigma_{m-1} e^{\alpha_m (t-t_{m-1})} \int_{t_{m-2}}^{t_{m-1}} e^{\alpha_{m-1}(t-s)} dW_s + ... + \sigma_1 e^{\alpha_m (t-t_{m-1})+\alpha_{m-1}\Delta t_{m-1}+...+\alpha_2 \Delta t_2} \int_{t_0}^{t_1} e^{\alpha_1 (t-s)} dW_s$$

Let $\Delta t_1 = t_1 - 0, \Delta t_2 = t_2 - t_1, ..., \Delta t_{m-1} = t_{m-1} - t_{m-2}, \Delta t_m = t - t_m$,

$$\ln S_t = e^{\sum_{i=1}^{m} \alpha_i \Delta t_i} \ln S_0 + \ln\phi_{m-1} + \left(\sum_{i=2}^{m-1} e^{\sum_{j=2}^{i} \alpha_{m-j+2}\Delta t_{m-j+2}} \ln\phi_{m-i} \right)$$
$$+\frac{\beta_m}{\alpha_m}(e^{\alpha_m (t-t_{m-1})}-1) + \left(\sum_{i=1}^{m-1} \frac{\beta_{m-i}}{\alpha_{m-i}} (e^{\alpha_{m-i}\Delta t_{m-i}}-1) e^{\sum_{j=1}^{i} \alpha_{m-j+1}\Delta t_{m-j+1}} \right)$$
$$+\sigma_m \int_{t_{m-1}}^{t} e^{\alpha_m (t-s)} dW_s + \left(\sum_{i=1}^{m-1} \sigma_{m-i} e^{\sum_{j=1}^{i} \alpha_{m-j+1}\Delta t_{m-j+1}} \int_{t_{m-i-1}}^{t_{m-i}} e^{-\alpha_{m-i}(t_{m-i}-s)} dW_s \right). \quad (6.2.4)$$

Now, let V be the European option on a stock with respect to hybrid nonlinear model 1. $V(S,t)$ is the value of the option at time t, where S_T is the stock price process defined in (6.2.4). The strike price K and maturity time T are as defined in Section 1.1. r is fixed interest rate. Applying to Theorem 1.1.2, we have

$V(S,t) = E^Q[e^{-r(T-t)} C_T \mid F_t]$, where $C_T = \begin{cases} \max\{(S_T - K),0\}, & call \\ \max\{(K - S_T),0\}, & put \end{cases}$

For hybrid nonlinear model 1, by knowing S_t and $\theta = T - t$, we note that $\ln S_T$ is normal distributed with

$$E[\ln S_T] = e^{\sum_{i=1}^{m} \alpha_i \Delta t_i} \ln S_0 + \ln\phi_{m-1} + \left(\sum_{i=2}^{m-1} e^{\sum_{j=2}^{i} \alpha_{m-j+2}\Delta t_{m-j+2}} \ln\phi_{m-i} \right)$$
$$+\frac{\beta_m}{\alpha_m}(e^{\alpha_m (T-t_{m-1})}-1) + \left(\sum_{i=1}^{m-1} \frac{\beta_{m-i}}{\alpha_{m-i}} (e^{\alpha_{m-i}\Delta t_{m-i}}-1) e^{\sum_{j=1}^{i} \alpha_{m-j+1}\Delta t_{m-j+1}} \right) = \mu(T), \quad (6.2.5)$$

and

$$Var(\ln S_T) = Var\left(\sigma_m \int_{t_{m-1}}^{T} e^{\alpha_m (t-s)} dW_s + \left(\sum_{i=1}^{m-1} \sigma_{m-i} e^{\sum_{j=1}^{i} \alpha_{m-j+1} \Delta t_{m-j+1}} \int_{t_{m-i-1}}^{t_{m-i}} e^{-\alpha_{m-i}(t_{m-i}-s)} dW_s \right) \right)$$

$$= \frac{\sigma_m^2}{2\alpha_m}(e^{2\alpha_m (T-t_{m-1})} - 1) + \left(\sum_{i=1}^{m-1} \frac{\sigma_{m-i}^2}{2\alpha_{m-i}} e^{2\sum_{j=1}^{i} \alpha_{m-j+1} \Delta t_{m-j+1}} (e^{2\alpha_{m-i} \Delta t_{m-i}} - 1) \right) = \sigma^2(T) \qquad (6.2.6)$$

Hence, for European call option,

$V(S,t) = E^Q[e^{-r(T-t)} C_T \mid F_t] \quad = E^Q\left[e^{-r(T-t)}\left(e^{\mu(T)+\sigma(T)Z} - K\right)_+\right],$

where Z is standard normal random variable.

First, we establish for the range of values of Z the integrand is non-zero.

$e^{\mu(T)+\sigma(T)Z} - K > 0$ is equivalent to $Z > \dfrac{\ln K - \mu(T)}{\sigma(T)} = d$,

where, $\mu(T)$ and $\sigma(T)$ are defined in (6.2.5) and (6.2.6) respectively.

Now we compute a European call option as

$$V(S,t) = \int_{d}^{\infty} e^{-r\theta}\left(e^{\mu(T)+\sigma(T)Z} - K\right)\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} dz = e^{\mu(T)-r\theta}\Phi(\sigma(T) - d) - Ke^{-r\theta}\Phi(-d) . \qquad (6.2.7)$$

Similarly, the corresponding formula for a European put option is

$$V(S,t) = \int_{-\infty}^{d} e^{-r\theta}\left(K - e^{\mu(T)+\sigma(T)Z}\right)\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} dz = Ke^{-r\theta}\Phi(d) - e^{\mu(T)-r\theta}\Phi(d - \sigma(T)) . \qquad (6.2.8)$$

Illustration 6.2.1: In the following, we illustrate the usefulness of the above presented result. Suppose the yearly interest rate $r = 6.5\%$, by applying (6.2.7) and (6.2.8), the call and put option price of three data sets are computed and recorded in Tables 6.2.1 and 6.2.2 for Hybrid nonlinear models 3.1.1 and 3.1.2, respectively.

Table 6.2.1 Call and Put Option Price of Hybrid Nonlinear Model 3.1.1

T	Stock X S_0 =691.48 K=700		Stock Y S_0 =84.84 K=90		S&P 500 Index S_0 =903.25 K=910	
	call	put	call	put	call	put
20	3.54	16.17	0	3.27	1.98	35.93
60	8.03	10.40	0	9.75	61.27	0.64
100	0	92.73	0	12.58	170.64	0
200	0	169.69	0	21.40	262.25	0

Table 6.2.2 Call and Put Option Price of Hybrid Nonlinear Model 3.1.2

T	Stock X S_0 =691.48 K=700		Stock Y S_0 =84.84 K=90		S&P 500 Index S_0 =903.25 K=910	
	call	put	call	put	call	put
20	9.27	7.66	0	12.76	0.40	31.67
60	9.10	8.86	0	15.74	3.08	31.02
100	0	123.62	0	19.09	61.14	0.36
200	0	171.43	0	21.37	192.99	0

6.2.2 Hybrid Nonlinear Stochastic Model 2

Similarly, the solution process of the hybrid nonlinear stochastic models 3.2.1 and 3.2.2 is given by,

$$
S_t = \begin{cases} \left(S_0^{1-N} e^{(1-N)\alpha_1 t} + \frac{\beta_1}{\alpha_1}(e^{(1-N)\alpha_1 t} - 1) + \sigma_1 (1-N) \int_0^t e^{(1-N)\alpha_1 (t-s)} dW_s \right)^{\frac{1}{1-N}} & S_0 = S_0, \quad 0 \le t < t_1 \\ \phi_1 \left(S_1^{1-N} e^{(1-N)\alpha_2 (t-t_1)} + \frac{\beta_2}{\alpha_2}(e^{(1-N)\alpha_2 (t-t_1)} - 1) + \sigma_2 (1-N) \int_{t_1}^t e^{(1-N)\alpha_2 (t-s)} dW_s \right)^{\frac{1}{1-N}} & S_1 = \lim_{t \to t_1^-} S_t, \quad t_1 \le t < t_2 \\ \ldots & \ldots \quad \ldots \\ \phi_{m-1} \left(S_{m-1}^{1-N} e^{(1-N)\alpha_m (t-t_{m-1})} + \frac{\beta_m}{\alpha_m}(e^{(1-N)\alpha_m (t-t_{m-1})} - 1) + \sigma_m (1-N) \int_{t_{m-1}}^t e^{(1-N)\alpha_m (t-s)} dW_s \right)^{\frac{1}{1-N}} & S_{m-1} = \lim_{t \to t_{m-1}^-} S_t, \quad t_{m-1} \le t < t_m \end{cases} \tag{6.2.9}
$$

Let $\Delta t_1 = t_1 - 0, \Delta t_2 = t_2 - t_1, ..., \Delta t_{m-1} = t_{m-1} - t_{m-2}, \Delta t_m = t - t_m$, recursively, we have

$$
\begin{aligned}
\frac{S_t^{1-N_{m+1}}}{1-N_{m+1}} = {} & \phi_{m-1}^{1-N_m} \ldots \phi_1^{1-N_2} \frac{S_0^{1-N_1}}{1-N_1} e^{(1-N_m)\alpha_m (t-t_{m-1}) + (1-N_{m-1})\alpha_{m-1}\Delta t_{m-1} + \ldots + (1-N_1)\alpha_1 \Delta t_1} \\
& + \phi_{m-1}^{1-N_m} \ldots \phi_1^{1-N_2} \frac{\beta_1}{\alpha_1 (1-N_1)} e^{(1-N_m)\alpha_m (t-t_{m-1}) + (1-N_{m-1})\alpha_{m-1}\Delta t_{m-1} + \ldots + (1-N_2)\alpha_2 \Delta t_2} (e^{(1-N_1)\alpha_1 \Delta t_1} - 1) \\
& + \phi_{m-1}^{1-N_m} \ldots \phi_2^{1-N_3} \frac{\beta_2}{\alpha_2 (1-N_2)} e^{(1-N_m)\alpha_m (t-t_{m-1}) + (1-N_{m-1})\alpha_{m-1}\Delta t_{m-1} + \ldots + (1-N_3)\alpha_3 \Delta t_3} (e^{(1-N_2)\alpha_2 \Delta t_2} - 1) + \ldots \\
& + \phi_{m-1}^{1-N_m} \frac{\beta_{m-1}}{\alpha_{m-1}(1-N_{m-1})} e^{(1-N_m)\alpha_m (t-t_{m-1})} (e^{(1-N_{m-1})\alpha_{m-1}\Delta t_{m-1}} - 1) + \frac{\beta_m}{\alpha_m (1-N_m)} (e^{(1-N_m)\alpha_m (t-t_{m-1})} - 1) \\
& + \phi_{m-1}^{1-N_m} \ldots \phi_1^{1-N_2} \sigma_1 e^{(1-N_m)\alpha_m (t-t_{m-1}) + (1-N_{m-1})\alpha_{m-1}\Delta t_{m-1} + \ldots + (1-N_2)\alpha_2 \Delta t_2} \int_0^{t_1} e^{(1-N_1)\alpha_1 (t_1 - s)} dW_s \\
& + \phi_{m-1}^{1-N_m} \ldots \phi_2^{1-N_3} \sigma_2 e^{(1-N_m)\alpha_m (t-t_{m-1}) + (1-N_{m-1})\alpha_{m-1}\Delta t_{m-1} + \ldots + (1-N_3)\alpha_3 \Delta t_3} \int_{t_1}^{t_2} e^{(1-N_2)\alpha_2 (t_2 - s)} dW_s + \ldots \\
& + \phi_{m-1}^{1-N_m} \sigma_{m-1} e^{(1-N_m)\alpha_m (t-t_{m-1})} \int_{t_{m-2}}^{t_{m-1}} e^{(1-N_{m-1})\alpha_{m-1}(t_{m-1}-s)} dW_s + \sigma_m \int_{t_{m-1}}^{t} e^{(1-N_m)\alpha_m (t-s)} dW_s
\end{aligned}
$$

Hence,

$$
\frac{S_t^{1-N_{m+1}}}{1-N_{m+1}} = \frac{S_0^{1-N_1}}{1-N_1} \left(\prod_{i=1}^{m-1} \phi_i^{1-N_{i+1}} \right) e^{\sum_{i=1}^{m} (1-N_i)\alpha_i \Delta t_i} + \frac{\beta_m}{\alpha_m (1-N_m)} \left(e^{(1-N_m)\alpha_m \Delta t_m} - 1 \right)
$$

$$+\sum_{i=1}^{m-1}\left(\left(\prod_{j=i}^{m-1}\phi_j^{1-N_{j+1}}\right)\frac{\beta_i}{\alpha_i(1-N_i)}e^{\sum_{j=i+1}^{m}(1-N_i)\alpha_i\Delta t_i}(e^{(1-N_i)\alpha_i\Delta t_i}-1)\right)$$

$$+\sigma_m\int_{t_{m-1}}^{t}e^{(1-N_m)\alpha_m(t-s)}dW_s+\sum_{i=1}^{m-1}\left(\left(\prod_{j=i}^{m-1}\phi_j^{1-N_{j+1}}\right)\sigma_i e^{\sum_{j=i+1}^{m}(1-N_j)\alpha_j\Delta t_j}\int_{t_{i-1}}^{t_i}e^{(1-N_i)\alpha_i(t_i-s)}dW_s\right)$$

(6.2.10)

Now, let V be the European option on a stock with respect to hybrid nonlinear model 2. $V(S,t)$ is the value of the option at time t, where S_T is the stock price process defined in (6.2.10). The strike price K and maturity time T are as defined in Section 1.1. r is fixed interest rate. Applying to Theorem 1.1.2, we have

$$V(S,t)=E^Q[e^{-r(T-t)}C_T \mid F_t], \text{ where } C_T=\begin{cases}\max\{(S_T-K),0\}, & call\\ \max\{(K-S_T),0\}, & put\end{cases}$$

For hybrid nonlinear model 2, by knowing S_t and $\theta=T-t$, we note that $Y_T=\dfrac{S_T^{1-N_m}}{1-N_m}$ is normal distributed with

$$E[Y_T]=\frac{S_0^{1-N_1}}{1-N_1}\left(\prod_{i=1}^{m-1}\phi_i^{1-N_{i+1}}\right)e^{\sum_{i=1}^{m}(1-N_i)\alpha_i\Delta t_i}+\frac{\beta_m}{\alpha_m(1-N_m)}\left(e^{(1-N_m)\alpha_m\Delta t_m}-1\right)$$

$$+\sum_{i=1}^{m-1}\left(\left(\prod_{j=i}^{m-1}\phi_j^{1-N_{j+1}}\right)\frac{\beta_i}{\alpha_i(1-N_i)}e^{\sum_{j=i+1}^{m}(1-N_i)\alpha_i\Delta t_i}(e^{(1-N_i)\alpha_i\Delta t_i}-1)\right)=\mu(T), \qquad (6.2.11)$$

and

$$Var(Y_T)=Var\left(\sum_{i=1}^{m-1}\left(\left(\prod_{j=i}^{m-1}\phi_j^{1-N_{j+1}}\right)\sigma_i e^{\sum_{j=i+1}^{m}(1-N_j)\alpha_j\Delta t_j}\int_{t_{i-1}}^{t_i}e^{(1-N_i)\alpha_i(t_i-s)}dW_s\right)+\sigma_m\int_{t_{m-1}}^{T}e^{(1-N_m)\alpha_m(T-s)}dW_s\right)$$

$$=\sum_{i=1}^{m-1}\left(\left(\prod_{j=i}^{m-1}\phi_j^{2(1-N_{j+1})}\right)^2\frac{\sigma_i^2 e^{2\sum_{j=i+1}^{m}(1-N_j)\alpha_j\Delta t_j}(e^{2(1-N_i)\alpha_i\Delta t_i}-1)}{2(1-N_i)\alpha_i}\right)$$

$$+\frac{\sigma_m^2}{2(1-N_m)\alpha_m}(e^{2(1-N_m)\alpha_m(T-t_{m-1})}-1)=\sigma^2(T). \tag{6.2.12}$$

Hence, for European call option,

$$V(S,t)=E^Q[e^{-r(T-t)}C_T \mid F_t] \quad = E^Q\left[e^{-r(T-t)}\left(\left((\mu(T)+\sigma(T)Z)(1-N_m)\right)^{\frac{1}{1-N_m}}-K\right)_+\right],$$

where Z is standard normal random variable.

First, we establish for the range of values of Z the integrand is non-zero.

$\left((\mu(T)+\sigma(T)Z)(1-N_m)\right)^{\frac{1}{1-N_m}}-K>0$ is equivalent to $Z>\dfrac{\dfrac{K^{1-N_m}}{1-N_m}-\mu(T)}{\sigma(T)}=d$,

where, $\mu(T)$ and $\sigma(T)$ are defined in (6.2.11) and (6.2.12) respectively.

Now we compute a European call option as

$$V(S,t)=\int_d^{\infty} e^{-r\theta}\left(\left((\mu(T)+\sigma(T)Z)(1-N_m)\right)^{\frac{1}{1-N_m}}-K\right)\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}dz$$

$$=\int_d^{\infty} e^{-r\theta}\left((\mu(T)+\sigma(T)Z)(1-N_m)\right)^{\frac{1}{1-N_m}}\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}dz-Ke^{-r\theta}\Phi(-d). \tag{6.2.13}$$

Similarly, the corresponding formula for a European put option is

$$V(S,t)=\int_{-\infty}^{d} e^{-r\theta}\left(K-\left((\mu(T)+\sigma(T)Z)(1-N_m)\right)^{\frac{1}{1-N_m}}\right)\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}dz$$

$$=Ke^{-r\theta}\Phi(d)-\int_{-\infty}^{d} e^{-r\theta}\left((\mu(T)+\sigma(T)Z)(1-N_m)\right)^{\frac{1}{1-N_m}}\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}dz. \tag{6.2.14}$$

Illustration 6.2.2: In the following, we illustrate the usefulness of the above presented result. Suppose the yearly interest rate $r=6.5\%$, by applying (6.2.13) and (6.2.14), the call and put option price of three data sets are computed and recorded in Tables 6.2.3 and 6.2.4 for Hybrid nonlinear models 3.2.1 and 3.2.2, respectively.

Table 6.2.3 Call and Put Option Price of Hybrid Nonlinear Model 3.2.1

T	Stock X S_0 =691.48 K=700		Stock Y S_0 =84.84 K=90		S&P 500 Index S_0 =903.25 K=910	
	call	put	call	put	call	put
20	29.06	33.76	6.01	9.49	3.49	1.45
60	55.40	10.04	8.56	13.13	4.15	2.08
100	74.16	9.68	6.30	11.76	27.22	32.74
200	114.6	9.06	13.57	20.97	53.77	59.20

Table 6.2.4 Call and Put Option Price of Hybrid Nonlinear Model 3.2.2

T	Stock X S_0 =691.48 K=700		Stock Y S_0 =84.84 K=90		S&P 500 Index S_0 =903.25 K=910	
	call	put	call	put	call	put
20	92.56	105.79	17.06	22.34	4.33	0.59
60	50.09	19.55	40.90	44.69	4.98	0.58
100	61.06	22.46	19.68	23.50	5.02	1.20
200	84.54	29.72	16.38	21.94	9.23	14.66

6.2.3 Hybrid Nonlinear Stochastic Model 3

The solution process of the hybrid nonlinear stochastic models 3.3.1 and 3.3.2 is given by,

$$S_t = \begin{cases} \dfrac{e^{(\alpha_1+\frac{\sigma_1^2}{2})t+\sigma_1 W_t}}{\dfrac{1}{S_0}-\beta_1\int_0^t e^{(\alpha_1+\frac{\sigma_1^2}{2})s+\sigma_1 W(s)}ds} & S_0 = S_0, \qquad 0 \le t < t_1 \\ \dfrac{\phi_1 e^{(\alpha_2+\frac{\sigma_2^2}{2})(t-t_1)+\sigma_2(W_t - W_{t_1})}}{\dfrac{1}{S_1}-\beta_1\int_{t_1}^t e^{(\alpha_2+\frac{\sigma_2^2}{2})s+\sigma_2 W(s)}ds} & S_1 = \lim_{t\to t_1^-} S_t, \qquad t_1 \le t < t_2 \\ \dots & \dots \qquad \dots \\ \dfrac{\phi_{m-1} e^{(\alpha_m+\frac{\sigma_m^2}{2})(t-t_{m-1})+\sigma_m(W_t - W_{t_{m-1}})}}{\dfrac{1}{S_{m-1}}-\beta_m\int_{t_{m-1}}^t e^{(\alpha_m+\frac{\sigma_m^2}{2})s+\sigma_m W(s)}ds} & S_{m-1} = \lim_{t\to t_{m-1}^-} S_t, \quad t_{m-1} \le t < t_m \end{cases} . \tag{6.2.15}$$

Now, let V be the European option on a stock with respect to hybrid nonlinear stochastic model 3. $V(S,t)$ is the value of the option at time t, where S_T is the stock price process defined in (6.2.15). The strike price K and maturity time T are as defined in Section 1.1. r is fixed interest rate. Applying to Theorem 1.1.2, we have

$$V(S,t) = E^Q[e^{-r(T-t)}C_T \mid F_t], \text{ where } C_T = \begin{cases} \max\{(S_T - K),0\}, & call \\ \max\{(K - S_T),0\}, & put \end{cases} \tag{6.2.16}$$

Illustration 6.2.2: In the following, we illustrate the usefulness of the above presented result. There is no a simple formula to compute the value of $V(S,t)$. To compute the numerical value of $V(S,t)$, we use equations (6.2.15) to simulate the value of S_T and then compute the expected value in (6.2.16). Suppose the yearly interest rate $r = 6.5\%$, by applying (6.2.16), the call and put option price of three data sets are computed and recorded in Tables 6.2.5 and 6.2.6 for Hybrid nonlinear models 3.3.1 and 3.3.2, respectively.

Table 6.2.5 Call and Put Option Price of Hybrid Nonlinear Model 3.3.1

T	Stock X S_0=691.48 K=700		Stock Y S_0=84.84 K=90		S&P 500 Index S_0=903.25 K=910	
	call	put	call	put	call	put
20	420.52	0	35.75	0	56.62	0
60	298.64	0	32.59	0	55.61	0
100	279.83	0	30.44	0	54.63	0
200	291.04	0	31.85	0	53.65	0

Table 6.2.6 Call and Put Option Price of Hybrid Nonlinear Model 3.3.2

T	Stock X S_0=691.48 K=700		Stock Y S_0=84.84 K=90		S&P 500 Index S_0=903.25 K=910	
	call	put	call	put	call	put
20	353.42	0	45.34	0	79.12	0
60	295.16	0	50.79	0	78.17	0
100	338.64	0	53.46	0	77.24	0
200	315.10	0	54.45	0	76.21	0

Chapter 7
Future Research Plan

The nonlinear stochastic modeling approach initiated in this work for solving forecasting and option pricing problems generates several interesting research problems in the financial engineering.

7.1 Data Smoothing Transformation

We note that a stochastic differential equation describes the continuous stock price process. The data sets we apply in our study are daily stock prices. In our future research, we want to explore the smoothing functions approach for better prediction and forecasting results.

7.1.1 Nonlinear Stochastic Model 1

In the following, a preliminary study with regard to nonlinear stochastic model 1 is presented. Here, we apply the smoothing function $Z_j = \frac{1}{n}\sum_{i=j}^{j+n-1} S_i$, $j = 1,2,\ldots,T-n+1$. Table 7.1.1 contains the result of AIC when we use value $n = 3$, and then apply to the Nonlinear Stochastic Model 1 using overall data set (Section 4.1). The basic statistics of the residual errors of fitted model are recorded in Table 7.1.2.

Table 7.1.1 AIC of Time Varying Coefficients Nonlinear Model 1 (n=3)
of Different Models of Three Datasets: Stock X, Stock Y and S&P 500 Index

	Stock X	Stock Y	S&P500 Index
(3, 1, 2)	-5925.74	-34022.21	-128409.7
(3, 1, 1)	-5667.26	-32847.02	-123984.7

(3, 1, 0)	-5600.38	-32401.51	-122989.2
(2, 1, 3)	-5925.82	-34016.82	-128410.7
(2, 1, 2)	-5927.59	-34019.93	-128409.8
(2, 1, 1)	-5575.21	-32161.64	-122653.9
(2, 1, 0)	-5566.89	-32090.21	-122485.6
(1, 1, 3)	-5927.38	-34019.00	-128404.0
(1, 1, 2)	-5929.58	-34019.86	-128373.8
(1, 1, 1)	-5549.04	-31995.15	-121906.5
(1, 1, 0)	-5528.05	-31892.23	-121206.7
(0, 1, 3)	-5929.52	-34019.89	-128377.2
(0, 1, 2)	-5926.42	-34021.24	-128345.8
(0, 1, 1)	-5335.44	-30982.72	-118845.7

Table 7.1.2 Basic Statistics of Time Varying Coefficients Nonlinear Model 1 (n=3) of Three Data Sets: Stock X, Stock Y and S&P500 Index

	Model	mean	variance	Standard deviation
Stock X	(1, 1, 2)	0.590782	59.550900	7.716923
Stock Y	(3, 1, 2)	0.012320	0.387723	0.622674
S&P 500 Index	(2, 1, 3)	0.045248	47.82333	6.915441

7.1.2 Nonlinear Stochastic Model 2

We repeat the smoothing transformation approach with regard to nonlinear stochastic model 2. The Stochastic Model Identification Procedure 4.2.1 is applied to obtain the time series model corresponding to the nonlinear stochastic model 2. Table 7.1.3 exhibits the basic statistics of the residual errors of nonlinear model 2 with $n = 3$.

Table 7.1.3 Basic Statistics of Time Varying Coefficients Nonlinear Model 2 (n=3) of Three Data Sets: Stock X, Stock Y and S&P500 Index

	Model	N	mean	variance	Standard deviation
Stock X	(1, 1, 2)	0.04	0.632399	57.25794	7.566898
Stock Y	(2, 1, 3)	0	0.014074	0.341886	0.58471
S&P 500 Index	(2, 1, 2)	0.02	0.067176	46.31976	6.805862

From this preliminary study, comparing the results in Tables 7.1.2, 7.1.3, 4.3.1, 4.3.2 and 4.3.3, we propose to utilize the smoothing function $Z_j = \frac{1}{n}\sum_{i=j}^{j+n} S_i$, $j = 1,2,...,T-n+1$, and also other smoothing linear and nonlinear functions to investigate forecasting problem.

7.2 Forecasting Problem

We recall that, in Section 4.3, we studied prediction problem and comparison about the performance of presented and existing models. This was based on three overall data sets. We simply attempted to use the formulations of stochastic nonlinear Models 4.1.1 and 4.2.1 with time varying coefficient for the data fitting problem. We further note that the performance of these models in the framework of data fitting is superior than the existing time series models and nonlinear stochastic models 1 and 2. The forecasting in the frame work of these models is open research problem. This problem will be also addressed in the future.

7.3 Option Pricing Problem

We observe that the parameters in our option pricing illustrations in Chapter 5 and 6 are estimated from stock price data sets. In practice, these implied parameters are computed from the historical option pricing data set. In our future research, we attempt to find the historical option pricing data (if available) and then apply to develop modified option pricing models.

References

[1] B. K. Øksendal, Stochastic differential Equations: An Introduction with Applications, 5th Edition, Berlin, New York: Springer, (1998).

[2] Fred D. Arditti, Derivatives, A Comprehensive Resource for Options, Futures, Interest Rates Swaps, and Mortgage Securities, Harvard Business School Press, (1996).

[3] Hirotugu Akaike, A new look at the statistical model identification, IEEE Transactions on Automatic Control, 19 (6) (1974), pp. 716–723.

[4] J. Marc Allaire, James B. Bittman, Riva Aidus Hemond, Elliot Kazt, Harrison Roth, Anthony J. Saliba, Gary L. Trennepohl, C. R. Tucker, and James W. Yates, Jr. Options, Essential Concepts and Trading Strategies, Second Edition, Richard D. IRWIN, Inc. (1995).

[5] Turan G. Bali, An Empirical Comparison of Continuous Time Models of The Short Term Interest Rate, The Journal of Futures Markets, Vol.19, No.7, (1999), pp.777-797.

[6] Fisher Black, and Piotr Karasinski, Bond and Option Pricing When Short Rates are Lognormal, Financial Analysts Journal (July-August) (1991), 52-59.

[7] Fischer Black, and Myron Scholes, The Pricing of Options and Corporate Liabilities, Journal of Political Economy 81 (3), (1973). 637–654.

[8] A. James Boness, Elements of a Theory of Stock Option Value, The Journal of Political Economy, April (1964), pp. 163-175.

[9] George E. P. Box, Gwilym M. Jenkins and Gregory C. Reinsel, Time Series Analysis, Forecasting and Control, Holden-Day, San Francisco. (1970).

[10] George E. P. Box, Gwilym M. Jenkins and Gregory C. Reinsel, Time Series Analysis, Forecasting and Control, Third Edition, Prentice-Hall International, Inc. (1994).

[11] M. J. Brennan and E. S. Schwartz, Conditional Predictions of Bond Prices and Returns, Journal of Finance 35 (1980), pp. 405-417.

[12] P. J. Brockwell and R. A. Davis, Introduction to Time Series and Forecasting, Springer, New York, (1996).

[13] Robert Goodell Brown, Smoothing, Forecasting and Prediction of Discrete Time Series, Prentice-Hall, Inc. Englewood Cliffs, N.J. (1963).

[14] F. Canova and B.E. Hansen, Are seasonal patterns constant over time? A test for seasonal stability. Journal of Business and Economic Statistics, ***13***, (1995), pp. 237-252.

[15] J. C. Cox, J. E. Ingersoll and S.A. Ross, A Theory of the Term Structure of Interest Rates, Econometrica 53(1985), pp.385–407.

[16] J. Chandra and G.S.Ladde, Stability Analysis of Stochastic Hybrid Systems, International Journal of Hybrid Systems, Vol.4 (2004), pp. 179-198.

[17] K.C. Chan, G. Andrew Karolyi, Francis A. Longstaff, and Anthony B. Sanders, An Empirical Comparison of Alternative Models of the Short-Term Interest Rate, The Journal of Finance, Vol. XLVII, N0.3, (1992) 1209-1227.

[18] Uri L. Dothan, On the Term Structure of Interest Rates, Journal of Financial Economics 6(1978), pp. 59-69.

[19] Alison Etheridge, A course in Financial Calculus, Cambridge, (2002).

[20] R. J. Elliott, and P. E. Kopp, Mathematics of Financial Markets, Springer Verlag, New York, NY, (1998).

[21] Kiyoshi Itō, On stochastic differential equations. Memoirs, American Mathematical Society 4 (1951), 1–51.

[22] Denis Kwiatkowski, Peter C. B. Phillips, Peter Schmidt and Youngcheol Shin, Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root, Journal of Econometrics 54 (1992) pp.159-178.

[23] Masaaki Kijima, Stochastic Processes with Applications to Finance, Chapman & Hall/CRC, (2003).

[24] Peter E. Kloeden and Eckhard Platen, Numerical Solution of Stochastic Di_erencial Equations, Springer-Verlag, (1992).

[25] A. G. Ladde, and G. S. Ladde, Dynamic Processes Under Random Environment, Bulletin of the Marathwada Mathematical Society, Vol.8, No.2 (2007), pp. 96-123.

[26] A. G. Ladde, and G. S. Ladde, An Introduction to Differential Equations-II: Stochastic Modeling, methods and Analysis, in press.

[27] G. S. Ladde, Hybrid Systems: Convergence and Stability Analysis of Stochastic Large-Scale Approximation Schemes, Dynamic Systems and Applications, Vol.30 (2004), pp.487-512.

[28] G. S. Ladde, and V. Lakshmikantham, Random Differential Inequalities, Academic Press, New York, (1980), MR 84b: 60081.

[29] G. S. Ladde, and M. Sambandham, Stochastic Versus Deterministic Systems of Differential Equations, Marcel Dekker, Inc., New York, (2004).

[30] G. S. Ladde and Ling Wu, Development of Modified Geometric Brownian Motion Models by Using Stock Price Data and Basic Statistics, Nonlinear Analysis 71 (2009) e1203–e1208.

[31] G. S. Ladde and Ling Wu, Development of Nonlinear Stochastic Models by Using Stock Price Data and Basic Statistics, Special Issue, Communications in Applied Analysis (In processing).

[32] M. E. Lindelöf, Sur l'application de la méthode des approximations successives aux équations différentielles ordinaires du premier ordre, Comptes rendus hebdomadaires des séances de l'Académie des sciences, Vol. 114, (1894), pp. 454–457.

[33] Robert C. Merton, Theory of rational option pricing, The bell Journal of Economics and Management Science, Vol. 4, No. 1, (1973), 141-183.

[34] N. Metropolis, and S. Ulam, The Monte Carlo Method. Journal of the American Statistical Association 44 (247) (1949), pp. 335–341.

[35] Raymond H. Myers, Classical and Modern Regression with Applications, Duxbury, (1989).

[36] M. F. M. Osborne, Brownian Motion in the Stock Market, Operations Research, March-April (1959), pp. 145-173.

[37] Vijay K. Rohatgi and A. K. Md. Ehsanes Saleh, An Introduction to Probability and Statistics, Second Edition, Wiley, New York, (2000).

[38] Shou Hsing Shih, Forecasting Models for Economic and Environmental Applications, Doctoral dissertation, University of South Florida, (2008).

[39] John W. Tukey, Exploratory Data Analysis, Addison-Wesley Publishing Company, (1977).

[40] Y K Tse, Some International Evidence on the Stochastic Behavior of Interest Rates, Journal of International Money and Finance, Vol. 14, No.5 (1995), pp. 721-738.

[41] Oldrich Vasicek, An Equilibrium Characterisation of the Term Structure, journal of Financial Economics 5 (1977), pp. 177-188.

[42] W. N. Venables and B. D. Ripley, Modern Applied Statistics with S, Fourth Edition, Springer, (2002).

[43] Simon Vine, Options: Trading Strategy and Risk Management, Hoboken, Weiey, N.J. (2005).

[44] G. Yin, Q. Zhang, and K. Yin, Constrained Stochastic Estimation Algorithms for a Class of Hybrid Stock Market Models, Journal of Optimization Theory and Applications, Vol. 118, No.1, (2003) 157-182.

Appendices

Appendix A1: The Estimated Parameters of Stock X Applying Monthly GBM Model

	2004		2005	
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$
January			0.001161554	0.02942549
February			-0.001702899	0.02796643
March			-0.001766466	0.01257827
April			0.009640864	0.02096884
May			0.01115012	0.01629551
June			0.002920542	0.02164556
July			-0.0009630533	0.01643769
August	0.003216502	0.03775822	-0.0001724213	0.01373443
September	0.01146937	0.02181682	0.00493249	0.01504678
October	0.01924003	0.04152636	0.008160907	0.02977017
November	-0.001520906	0.03722648	0.004194011	0.01871986
December	0.002805678	0.01911732	0.001242512	0.01315277

Appendix A1: (Continued)

	2006		2007	
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$
January	0.002671674	0.03379703	0.004406722	0.01673394
February	-0.008647245	0.03598111	-0.005636942	0.01614764
March	0.003500880	0.02592455	0.0009669722	0.01374949
April	0.003836704	0.01975170	0.001483809	0.01109096
May	-0.005141313	0.01863313	0.002560355	0.01195841
June	0.005608799	0.01690952	0.002363476	0.009974502
July	-0.003991524	0.0119881	-0.001067062	0.01443759
August	-0.0008110875	0.01456686	0.0005236452	0.01251902
September	0.003111817	0.01525902	0.005110094	0.009779767
October	0.007971126	0.02201493	0.009694167	0.0155215
November	0.0009373092	0.01435362	-0.0005707263	0.0276292
December	-0.002497903	0.01236763	2.591382e-06	0.01499199

Appendix A2: The Estimated Parameters of Stock X Applying Unequal Interval GBM Model

Index	$\hat{\mu}$	$\hat{\sigma}$
1	0.006621	0.024521
2	0.013391	0.028731
3	0.009891	0.055846
4	0.001149	0.030116
5	0.009663	0.019472
6	-0.00134	0.026352
7	0.005189	0.038095
8	-0.00261	0.015574
9	0.008553	0.017281
10	0.00884	0.022777
11	-0.00053	0.014418
12	7.88E-05	0.015385
13	0.000596	0.015815
14	0.016518	0.033196
15	0.009311	0.013909
16	-0.00076	0.016247
17	0.004541	0.025039
18	-0.01211	0.036582
19	0.000358	0.028831

Index	$\hat{\mu}$	$\hat{\sigma}$
20	0.010446	0.022202
21	-0.00389	0.02337
22	-0.0026	0.019065
23	0.002168	0.015421
24	-0.00417	0.014351
25	0.003034	0.012816
26	0.003202	0.014617
27	0.007015	0.020763
28	-0.00327	0.012388
29	1.37E-05	0.015938
30	-0.00115	0.017478
31	-0.00045	0.014767
32	2.02E-05	0.01084
33	0.007892	0.011226
34	0.0025	0.008967
35	-0.00407	0.016643
36	0.003471	0.010888
37	0.008705	0.015797
38	0.010025	0.015063
39	-0.00136	0.022851

Appendix A3: The Estimated Jump Coefficient of Stock X Applying Monthly GBM Model

	2004	2005	2006	2007
Jan to Feb	N/A	0.9333078	0.8899062	0.9496149
Feb to Mar	N/A	1.0128700	1.0826370	1.0392360
Mar to Apr	N/A	1.0101639	0.9906786	1.0042266
Apr to May	N/A	1.0139198	0.9524189	0.9933268
May to June	N/A	1.0276019	1.0794418	1.0103405
June to July	N/A	0.9493318	0.9786493	1.0100002
July to Aug	N/A	1.0248249	0.9630886	0.9895443
Aug to Sep	0.9751628	0.9868456	1.0288616	1.0153643
Sep to Oct	1.0456927	1.0070887	0.9986621	1.0084123
Oct to Nov	1.0101994	1.0134436	0.9816198	0.9663877
Nov to Dec	0.9644825	1.0051667	1.0108246	0.9845458
Dec to next Jan	1.0668690	1.0269237	1.0249850	N/A

Appendix A4: The Estimated Jump Coefficient of Stock X Applying Unequal Interval GBM Model with Jumps

Index	Jump Coefficient
1	N/A
2	0.938119
3	0.929323
4	1.058691
5	1.040035
6	0.99704
7	1.09544
8	0.995219
9	0.937131
10	0.973401
11	1.090801
12	0.936869
13	0.999014
14	0.92379
15	1.086789
16	1.082488
17	0.908315
18	1.146248
19	0.855107
20	1.000371

Index	Jump Coefficient
21	1.015838
22	1.017131
23	0.994225
24	1.073491
25	0.940836
26	0.994028
27	0.961578
28	1.123955
29	0.996198
30	0.924911
31	1.079305
32	0.936505
33	0.998823
34	1.008491
35	1.07922
36	0.96916
37	0.952614
38	0.988914
39	1.096968

About the Author

Ling Wu got her undergraduate education at one of the highly respected higher educational institution, namely, DaLian University of Technology in China. In 2004, she came to the United States pursue her further higher education at University of South Florida, Tampa, Florida. Her passion and interest for mathematics and statistics motivated to earn her M.A. Degree in Mathematics with the concentration in statistics in 2006. Thereafter, she continued her Ph.D, and wrote her dissertation, entitled "Stochastic Modeling and Statistical Analysis" at University of South Florida. In her dissertation, Ling shows her creativity, intellectual independence, and exhibited her both analytical and computational abilities. She is sharp and energetic young lady. As a person, she always extends helping hands to other people.