

THE UNIVERSITY OF CALGARY

Stochastic Models for Gas Prices

by

Zhiyong Xu

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES  
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE  
DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF MATHEMATICS AND STATISTICS

CALGARY, ALBERTA

AUGUST, 2004

© Zhiyong Xu 2004

**THE UNIVERSITY OF CALGARY**  
**FACULTY OF GRADUATE STUDIES**

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled “Stochastic Models for Gas Prices” submitted by Zhiyong Xu in partial fulfillment of the requirements for the degree of MASTER OF SCIENCE.

---

Dr. Tony Ware  
Department of Mathematics and  
Statistics

---

Dr. Len Bos  
Department of Mathematics and  
Statistics

---

Dr. Apostolos Serletis  
Department of Economics

---

Date

# Abstract

This thesis focuses on modeling natural gas price behaviour, and in particular mean reversion, seasonality, and the relationship between futures prices and spot prices. Since many of these features can not be adequately captured using only one source of randomness, as well as looking at 1-factor models, we also examine some 2-factor models and compare their effectiveness and explanatory power. By exploiting the relationship between futures prices and spot prices, we are able to recover information which can not be observed directly, and this enables us to perform maximum likelihood estimation for calibration of the models.

# Acknowledgments

This thesis benefited greatly from several people. I would like to express my gratitude to them for their invaluable help.

First and foremost I would like to thank my supervisor, professor Tony Ware, for his boundless enthusiasm and support for this project. I learned so much from him about mathematical finance. His guidance was invaluable in this thesis.

Second, I am grateful to my interim-supervisor, professor Ali Lari-Lavassani, for his early and great support for this project.

Special thanks go to my classmates Lei Xiong and Matthew Emmett, for their sincere assistance in the project.

Finally, I would like to thank my families, especially my sister Lijing Xu. I just can't imagine having done this without their confidence in me and endless support.

# Table of Contents

<b>Approval Page</b>	<b>ii</b>
<b>Abstract</b>	<b>iii</b>
<b>Acknowledgments</b>	<b>iv</b>
<b>Table of Contents</b>	<b>v</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Natural Gas Market and the Behaviour of Gas Prices</b>	<b>4</b>
2.1 Natural Gas Market . . . . .	4
2.2 Natural Gas Spot Prices . . . . .	6
<b>3 Modeling Gas Prices</b>	<b>11</b>
3.1 Basic Stochastic Calculus . . . . .	11
3.1.1 Stochastic Processes . . . . .	11
3.1.2 Wiener Process . . . . .	11
3.1.3 Itô Integral . . . . .	12
3.1.4 Itô Process and Stochastic Differential Equation . . . . .	13
3.1.5 Itô's Lemma . . . . .	13
3.1.6 Some Specific SDEs . . . . .	13
3.2 Mean Reverting Processes . . . . .	15
3.3 Models for Gas Prices . . . . .	18
3.3.1 1-factor Models . . . . .	18
3.3.2 2-factor Models . . . . .	21
<b>4 Futures Prices</b>	<b>24</b>
4.1 Introduction to Forwards and Futures . . . . .	24
4.2 Futures Price Formula . . . . .	25
4.2.1 No Arbitrage Assumption . . . . .	25
4.2.2 Partial Differential Equation Method . . . . .	26
4.2.3 Method of Equivalent Measures . . . . .	30
4.2.4 The Other Futures Price Formulas . . . . .	34
4.2.5 An Adjustment to the Theoretical Futures Price . . . . .	35

<b>5</b>	<b>Calibration of the Models</b>	<b>37</b>
5.1	Maximum Likelihood Estimation . . . . .	37
5.1.1	Characteristic Function and Probability Density Function for Affine Jump-Diffusion Process . . . . .	38
5.2	Calibration for 1-factor Models without Seasonality . . . . .	38
5.2.1	Probability Density Function . . . . .	39
5.2.2	How to Treat the Non-trading Days . . . . .	46
5.2.3	ML Estimation . . . . .	48
5.2.4	Continuous Data and Discrete Data . . . . .	49
5.2.5	How the Length of Data Affects the Results . . . . .	50
5.2.6	Estimation Results for real Data . . . . .	52
5.3	Calibration for 1-factor Models with Seasonality . . . . .	52
5.3.1	Reveal the Seasonal Term $f(t)$ from Futures Prices . . . . .	53
5.3.2	Probability Density Function . . . . .	55
5.3.3	Estimation . . . . .	56
5.3.4	Results . . . . .	57
5.3.5	Futures Matching . . . . .	64
5.4	Calibration for 2-factor Models without Seasonality . . . . .	66
5.4.1	Reveal the Hidden Factor $L_t$ from Futures Prices . . . . .	66
5.4.2	Probability Density Function . . . . .	70
5.4.3	Estimation . . . . .	71
5.4.4	Results . . . . .	73
5.4.5	Futures Matching . . . . .	79
5.5	Calibration for 2-factor Models with Seasonality . . . . .	80
5.5.1	Find the Hidden Factor $L_t$ and Seasonal Term $f(t)$ from Fu- tures Prices . . . . .	80
5.5.2	Probability Density Function . . . . .	84
5.5.3	Estimation . . . . .	85
5.5.4	Results . . . . .	86
5.5.5	Futures Matching . . . . .	93
<b>6</b>	<b>Discussion and Conclusion</b>	<b>95</b>
<b>A</b>	<b>Transform Analysis for Affine Jump-Diffusion State-Vectors</b>	<b>99</b>
A.1	Affine Jump-Diffusion . . . . .	99
A.2	Transform . . . . .	100
A.3	Extended Transform . . . . .	101
<b>B</b>	<b>Deduction of Futures Price Function for 2-factor Model by Affine Process Transform</b>	<b>102</b>

<b>C</b>	<b>The Computation of an Integral</b>	<b>104</b>
<b>D</b>	<b>Continuous PDFs for 2-factor Models without Seasonality</b>	<b>105</b>
D.1	Model 2f-A . . . . .	105
D.2	Model 2f-B . . . . .	107
D.3	Model 2f-C . . . . .	107
	<b>Bibliography</b>	<b>108</b>

# List of Tables

5.1	Volatility on 1-day apart successive trading days vs. volatility on 3-day apart successive trading days . . . . .	47
5.2	Estimation using the subset of prices for model 1f-A . . . . .	50
5.3	Estimation using the subset of prices for model 1f-B . . . . .	50
5.4	The effects of data length for model 1f-A . . . . .	51
5.5	The effects of data length for model 1f-B . . . . .	52
5.6	Estimation results for 1-factor models . . . . .	52
5.7	Estimation of the risk-neutral parameters for 1-factor models with seasonality . . . . .	54
5.8	Estimation result for $f(t)$ in 1-factor models with seasonality . . . . .	57
5.9	Estimation result for model 1f-sn-A . . . . .	58
5.10	Estimation result for model 1f-sn-B . . . . .	60
5.11	Estimation result for model 1f-sn-C . . . . .	62
5.12	Estimation of the risk-neutral parameters for 2-factor unseasonal models . . . . .	68
5.13	Estimation result for model 2f-A . . . . .	73
5.14	Estimation result for model 2f-B . . . . .	75
5.15	Estimation result for model 2f-C . . . . .	77
5.16	Estimation of the risk-neutral parameters for 2-factor models with seasonality . . . . .	81
5.17	Estimation result for $f(t)$ in 2-factor models with seasonality . . . . .	86
5.18	Estimation result for model 2f-sn-A . . . . .	87
5.19	Estimation result for model 2f-sn-B . . . . .	89
5.20	Estimation result for model 2f-sn-C . . . . .	91

# List of Figures

2.1	Henry Hub gas spot prices . . . . .	7
2.2	Average daily spot prices . . . . .	9
2.3	Futures curve . . . . .	9
5.1	Density comparison for model 1f-A . . . . .	42
5.2	Density comparison for model 1f-B . . . . .	45
5.3	Recover $f(t)$ from futures prices for artificial data . . . . .	55
5.4	Reveal $f(t)$ from futures prices for real data . . . . .	55
5.5	Result for model 1f-sn-A . . . . .	59
5.6	Result for model 1f-sn-B . . . . .	61
5.7	Result for model 1f-sn-C . . . . .	63
5.8	Futures matching for 1-factor models with seasonality . . . . .	65
5.9	Recover $L_t$ from futures prices for artificial data . . . . .	68
5.10	Reveal $L_t$ from futures prices for real data . . . . .	69
5.11	Result of model 2f-A . . . . .	74
5.12	Result of model 2f-B . . . . .	76
5.13	Result of model 2f-C . . . . .	78
5.14	Futures matching for 2-factor models without seasonality . . . . .	79
5.15	Recover $L_t$ and $f(t)$ from futures prices for artificial data . . . . .	82
5.16	Reveal $L_t$ and $f(t)$ from futures prices for real data . . . . .	83
5.17	Result for model 2f-sn-A . . . . .	88
5.18	Result for model 2f-sn-B . . . . .	90
5.19	Result for model 2f-sn-C . . . . .	92
5.20	Futures matching for 2-factor models with seasonality . . . . .	94

# Chapter 1

## Introduction

Natural gas, which was considered by early oil drillers to be a nuisance and was vented or burned off at the well site, currently plays a very important role in the energy market. It supplies almost a quarter of the energy needs in North America. And, since natural gas is the cleanest burning fossil fuel, producing primarily carbon dioxide, water vapor and small amounts of nitrogen oxides, recent legislation concerning air pollution control gives the market further growth potential.

Like crude oil and electricity, the natural gas market in north America has been going through a period of deregulation and elimination of stage monopolies. This has been accompanied by the development of the corresponding financial derivatives. Now, the market has become a fairly liberal market. Keener and keener competition drives the need for a better understanding of this market.

This thesis focuses on capturing natural gas price behaviour, in particular, mean reversion, seasonality, and the relationship between futures prices and spot prices, etc. Many of these features can not be adequately captured using only one source of randomness. So we examine multi-factor models and compare their effectiveness and explanatory power. These models serve as the basis for further research in risk management.

In Chapter 2, we begin by briefly introducing the history of natural gas market, together with the development of the natural gas futures and other derivatives. We describe the main characteristics of the behaviours of gas spot prices. Besides some

properties common to many commodities, gas prices also feature mean reversion and seasonality.

In Chapter 3, we start with a review of some basic elements of stochastic calculus, such as Wiener Processes, Itô integrals, Itô processes, Itô's Lemma and stochastic differential equations. We give some examples of specific stochastic differential equations, and concentrate on the introduction of mean reverting processes. Finally, we introduce the leading elements of this thesis, a series of stochastic models for gas spot prices. They are all mean reverting, and can be grouped into two classes: 1-factor models and 2-factor models. For each class, we have a type of nonseasonal model and a type of seasonal model, and for each type, we form the volatility term in 3 different ways.

In Chapter 4, we present the relationship between spot prices and futures prices. Under the no arbitrage assumption, we deduce the formula for futures price by two methods: a partial differential equation method and an equivalent martingale measures method. We demonstrate the deduction using a 2-factor model, and give the futures price formulas for the other models directly.

In Chapter 5, we study the methods for calibration of our models. First, we introduce the basic idea of maximum likelihood estimation. Since most of our models fall into the class of affine processes, we explain how to get conditional probability density functions of the random variables using the affine jump-diffusion process transform theory. In the process of calibrating the relatively simple 1-factor models without seasonality, we study some common issues confronted in calibration, such as how to approximate continuous probability density functions by discrete probability density functions, how the data length affects the calibration results, how to deal

with the non-trading days, and so on. Then we move to the calibration for 1-factor models with seasonality and 2-factor models. An important technique used in this part is to reveal the hidden factors and seasonal term by matching futures prices. We also show how the obtained risk-neutralized parameters can be used to match futures curves.

Finally, we summarize our conclusions and give some discussion in Chapter 6.

The concerned computer programs used in this thesis were written in Matlab and executed on Pentium 4 personal computers.

## Chapter 2

# Natural Gas Market and the Behaviour of Gas Prices

### 2.1 Natural Gas Market

Natural gas is a combustible, gaseous mixture of simple hydrocarbon compounds, usually found in deep underground reservoirs formed by porous rock. Natural gas is a fossil fuel composed almost entirely of methane<sup>1</sup>, but does contain small amounts of other gases, including ethane, propane, butane and pentane.

The prevailing scientific theory is that natural gas was formed millions of years ago when plants and tiny sea animals were buried by sand and rock. Layers of mud, sand, rock, plant and animal matter continued to build up until the pressure and heat from the earth turned them into petroleum and natural gas.

Though natural gas was first discovered in the 1600s, transportation limitations prevented the fuel from being widely used for centuries. The first use of gas energy in North America occurred in 1816, when gaslights illuminated the streets of Baltimore, Md. By 1900, natural gas had been discovered in 17 states of America. During the years following World War II, expansion of the extensive interstate pipeline network occurred, bringing natural gas service to customers all over the continent.

The main uses of natural gas are in heating, in power generation, as a household fuel, and as a chemical feedstock. Currently, natural gas plays a major role in the

---

<sup>1</sup>Methane is composed of a molecule of one carbon atom and four hydrogen atoms.

North American energy profile, where it accounts for about a quarter<sup>2</sup> of total energy consumption.

Over the past two decades, the gas market has not simply grown — it has also experienced radical change. The deregulation of the natural gas industry in North American began with the enactment of the American Natural Gas Policy Act of 1978. Since then, from a market of stable but controlled prices, the natural gas market has emerged as a dynamic, highly competitive business with flexible prices. In this deregulated environment, market participants have found themselves increasingly exposed to price movements and to counterparty performance risk — and this has led to an exponentially increasing need for risk management.

Based on the groundwork laid by the development of the spot and forward market, the New York Mercantile Exchange (NYMEX) launched the world's first natural gas futures contract in April 1990. Since then, futures contracts have been widely used to hedge against price fluctuations in this volatile market. Furthermore, the standardization of contract terms, relatively small contract size, fungibility, lack of requirement for physical delivery, and rigorous performance requirements serve to attract large amounts of financial capital to this market. The participants include not only natural gas producers, marketers, processors, utilities, and end users, but also a host of other speculative market participants. Volume and open interest have grown rapidly, establishing the contract as the fastest growing instrument in Exchange history. In 2002, the daily average volume of the NYMEX market exceeded 97,000 contracts, which involves several times the average daily consumption of gas in North

---

<sup>2</sup>24% for USA and 31% for Canada. Data source: American gas association ([www.aga.org](http://www.aga.org)) and Canadian gas association ([www.cga.ca](http://www.cga.ca)).

America. The futures market successfully adapted itself to the trading practices of the industry, improving them in the process. In turn, the successful futures market then became the foundation for many other forms of derivative trade, such as options and swaps.

In short, the natural gas market has played a very important role in our life, and the corresponding financial markets have also gradually flourished in the last decades. All kinds of problems in this area are attracting more and more researchers to work on them.

## 2.2 Natural Gas Spot Prices

The goal of this thesis is to model the behaviour of gas spot prices<sup>3</sup> mathematically. Before starting with the description of the mathematical models, it is important to keep in mind the actual behaviour of the gas spot prices that we are trying to model.

The prices of natural gas, which is a kind of commodity, behaves somewhat differently to other prices found in financial markets. A number of fundamental price drivers, such as issues of extraction, storage, transport, weather, policies, technological advances, etc., cause extremely complex gas prices behaviour.

Figure 2.1 illustrates the plot of daily gas spot prices<sup>4</sup>, which are measured in dollars per million British thermal units (MMBTU), from January 2, 1992 to December 30, 1999 (2007 observations in total, about 251 observations for every year).

---

<sup>3</sup>Here the spot prices represent natural gas sales contracted for next day delivery.

<sup>4</sup>As the world market for natural gas is fragmented in different regional markets, it is not possible to talk about a world price for natural gas. In this thesis, the prices we use are based on delivery at the Henry Hub in Louisiana.

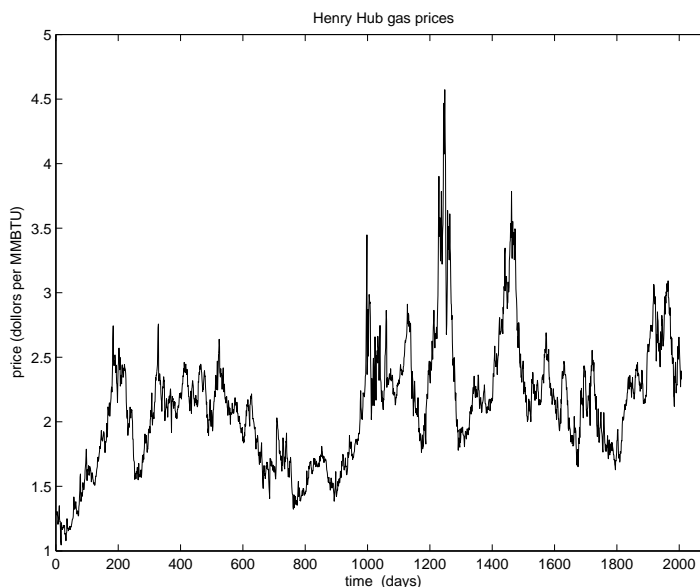


Figure 2.1: Henry Hub gas spot prices

There are several important properties associated with the behaviour of gas spot prices. Besides some common properties for most commodity prices, such as uncertainty and non-negativity, one of the most important characteristics of gas prices is mean reversion, which has been well documented (see [1, 2]). From Figure 2.1, we observe that the prices move up and down frequently, but actually oscillate around an equilibrium level from the point of long term view. When the prices are far away from that level, they tend to move back to that level. This is just the effect of mean reversion, and in this case we say that the prices mean revert to a long term mean.

The mean reversion in natural gas prices appears to be related to their reactions to events such as floods, summer heat waves, and other news-making events, which can create new and unexpected supply-and-demand imbalances in the market. Either a correction on the supply side, to match the demand side, or the actual dissipation

of the event, such as the temperatures reverting to their average seasonal levels, tends to cause the natural gas prices to come back to their typical levels. The speed of mean reversion depends on how quickly it takes for supply and demand to return to a balanced state or for the events to dissipate. The long term mean around which the prices oscillate is determined by the cost of production and the normal level of demand.

Another very important property of gas spot prices we need to capture is strong seasonality. Seasonality results primarily from regular demand fluctuations which are driven by recurring weather related factors. As we know, natural gas is a primary residential and commercial space heating fuel and is used increasingly as a fuel for electricity generation. Thus the cold winter weather results in above average consumption of natural gas. And in some very hot summers, demand increases for power generation to run air conditioning and other cooling devices. On the other hand, the difficulty of storage and the limitation of transmission capacity make the supply side not elastic enough to match the suddenly increased demand side very quickly. Hence, the seasonal imbalances between demand and supply result in the inevitable seasonal fluctuations of gas prices.

The seasonality effects can be seen not only through historical spot prices, but also through futures prices (see Chapter 4). Figure 2.2 is the plot of the daily prices averaged over 8 years. It shows the evident appearance of a winter peak, while a summer peak can also be seen, but is of much smaller magnitude. Figure 2.3 is the futures curve observed on January 1, 1998, with the futures prices 3 years forward. This futures curve reflects the fact that market participants had recognized the seasonal behaviour of spot prices. There are still many other important properties

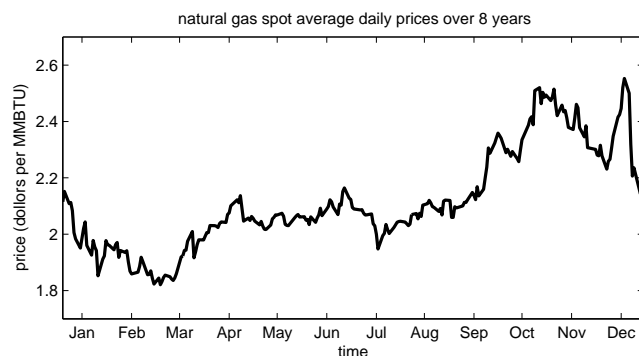


Figure 2.2: Average daily spot prices

Note: Each point on the above curve is the average of the 8 prices with the same date in 8 years.

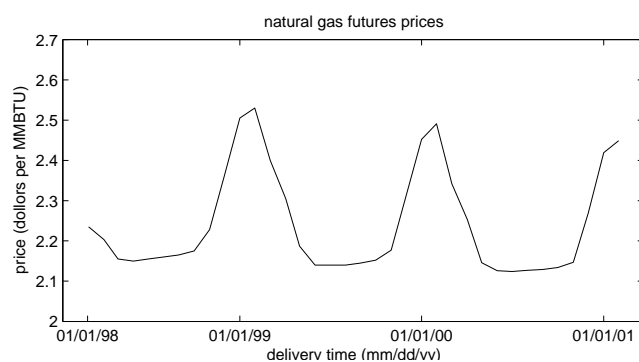


Figure 2.3: Futures curve

in gas price behaviours. For example, gas prices have apparent regionality — they are not the same in different places. Furthermore, occasionally gas prices exhibit sudden and large changes, which may be treated as “jumps”. These features are not modeled directly. In addition, for the sake of simplicity, we ignore the development of the market over time by not allowing the parameters in our models to change from year to year.

In the remainder of this thesis, we will focus on mean reversion and seasonal-

ity. And above all, because of the uncertainty of gas price behaviours, we start our modeling by discussing some mathematical tools about randomness in the next chapter.

# Chapter 3

## Modeling Gas Prices

In this chapter, we introduce a series of mathematical models appropriate for gas prices. Because of the random behaviour of gas prices, stochastic processes are chosen for modeling purposes. Stochastic processes form the basis of almost all the derivative pricing and risk management models in finance markets. They allow us to model possible price evolution through time, and to give probabilities to possible future prices as a function of current prices and other parameters.

### 3.1 Basic Stochastic Calculus

#### 3.1.1 Stochastic Processes

A *stochastic process*  $X = \{X_t\}_{t \in T}$  is a collection of random variables. That is, for each  $t$  in the index set  $T$ ,  $X_t$  is a random variable. We often interpret  $t$  as time and call  $X_t$  the state of the process at time  $t$ . The index set  $T$  can be a countable set, giving us a discrete-time stochastic process, or a non-countable continuous set, giving us a continuous-time stochastic process. Any realization of  $X$  is named a sample path, which can be discrete or continuous.

#### 3.1.2 Wiener Process

One of the most important stochastic processes is the *Wiener process* (or *Brownian motion*); without confusion, hereafter we denote both the process and the state of

the process at time  $t$  by  $W_t$ ). It is defined as a time-continuous process with the properties

1.  $W_0 = 0$  (with probability 1).
2.  $W_t \sim \mathcal{N}(0, t)$  for all  $t \geq 0$ . That is, for each  $t$  the random variable  $W_t$  is normally distributed with mean  $E(W_t) = 0$  and variance  $\text{Var}(W_t) = t$ .
3. All increments  $\Delta W_t := W_{t+\Delta t} - W_t$  on non-overlapping time intervals are independent. That is,  $W_{t_2} - W_{t_1}$  and  $W_{t_4} - W_{t_3}$  are independent for all  $0 \leq t_1 < t_2 \leq t_3 < t_4$ . (This implies that  $W_t$  is a Markov process<sup>1</sup>.)
4.  $W_t$  depends continuously on  $t$ , meaning that any realization of  $W_t$  is a continuous function of  $t$ .

### 3.1.3 Itô Integral

If the random variable  $\sigma(X_t, t)$  is “non-anticipative”, in the sense that it is independent of the future, and “non-explosive”, i.e.,  $E[\int_0^T \sigma(X_t, t)^2 dt] < \infty$ , then the *Itô integral*,

$$\int_0^T \sigma(X_t, t) dW_t, \quad (3.1)$$

can be defined as

$$\lim_{\delta \rightarrow 0} \sum_{k=0}^{n-1} \sigma(X_{\tau_k}, \tau_k) (W_{t_{k+1}} - W_{t_k}), \quad (3.2)$$

where  $0 = t_0 < t_1 < t_2 < \dots < t_n = T$ ,  $\tau_k \in [t_k, t_{k+1}]$  and  $\delta = \max_{k=1,2,\dots,n-1} \{t_{k+1} - t_k\}$ , and the convergence is in the sense of probability.

---

<sup>1</sup>A Markov process is one that only the present value of random variable is relevant for its future motion. That is, it has no memory.

### 3.1.4 Itô Process and Stochastic Differential Equation

An *Itô process*  $X_t$  is governed by a *stochastic differential equation* (SDE) of the form

$$dX_t = a(X_t, t)dt + b(X_t, t)dW_t; \quad (3.3)$$

this together with  $X_{t_0} = X_0$  is a symbolic short form of the integral equation

$$X_t = X_0 + \int_{t_0}^t a(X_s, s)ds + \int_{t_0}^t b(X_s, s)dW_s, \quad (3.4)$$

where  $W_t$  is a Wiener process and  $\int_{t_0}^t b(X_s, s)dW_s$  is an Itô integral. The terms in (3.3) are named as follows:

$a(X_t, t)$ : drift term or drift coefficient,

$b(X_t, t)$ : diffusion term.

### 3.1.5 Itô's Lemma

Itô's lemma is a fundamental tool when analyzing stochastic processes. It can be regarded as the stochastic version of the chain rule. It states that if  $V(X, t)$  is a sufficiently smooth function of a random variable  $X$  and time  $t$ , and  $X_t$  follows an Itô process (3.3), then  $V_t := V(X_t, t)$  follows another Itô process

$$dV_t = \left( \frac{\partial V}{\partial t} + a(X_t, t) \frac{\partial V}{\partial X} + \frac{1}{2} b^2(X_t, t) \frac{\partial^2 V}{\partial X^2} \right) dt + b(X_t, t) \frac{\partial V}{\partial X} dW_t, \quad (3.5)$$

where the dependence on  $(X_t, t)$  is dropped for the sake of brevity.

For proof, see [3]. For a higher dimensional version of Itô's lemma, see Section 4.2.2.

### 3.1.6 Some Specific SDEs

There are some processes described by specific SDEs that are found to be quite useful.

### Generalized Wiener process

One of the simplest case of SDEs is where the drift and diffusion terms are all independent of the information received over time,

$$dS_t = \mu dt + \sigma dW_t, \quad (3.6)$$

where  $W_t$  is a standard Wiener process,  $\mu$  and  $\sigma$  are constant.

However, in practice, this model is seldom used to model the asset prices directly, because  $S_t$  could take negative values in (3.6), which conflicts with the non-negativity of the asset prices.

### Geometric Brownian motion (GBM)

Since being first proposed by Samuelson (see [4]) and then exploited by Black, Scholes (see [5]), and Merton (see [6]), this process has been the most used stochastic process in financial economic theory and in practice<sup>2</sup>. This process captures the movements of stock prices very well. The SDE is given by

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad (3.7)$$

where  $W_t$  is a standard Wiener process,  $\mu$  and  $\sigma$  are constant.

If we define  $X_t = \ln S_t$ , then Itô's Lemma tells us that

$$dX_t = \left(\mu - \frac{\sigma^2}{2}\right)dt + \sigma dW_t, \quad (3.8)$$

---

<sup>2</sup>In 1970's, Fischer Black, Myron Scholes, and Robert Merton made a major breakthrough in the pricing of stock options by developing what has become known as the Black-Scholes model. The model has had a huge influence on the way that traders price and hedge options. It has also been pivotal to the growth and success of financial engineering in the next decades. In 1997, the importance of the model was recognized when Myron Scholes and Robert Merton were awarded the Nobel prize for economics.

which implies that  $X_t \sim \mathcal{N}(X_0 + (\mu - \frac{\sigma^2}{2})t, \sigma^2 t)$ . This is one of the reasons why the GBM model is so popular. We say that  $S_t$  is log-normal. In addition, GBM captures the non-negativity property of stock prices, since  $S_t = \exp(X_t) > 0$ .

### Square root process

A model close to the one just discussed is the square root process:

$$dS_t = \mu S_t dt + \sigma \sqrt{S_t} dW_t, \quad (3.9)$$

where  $W_t$  is a standard Wiener process,  $\mu$  and  $\sigma$  are constant.

Compared with (3.7), the standard deviation is a function of the square root of  $S_t$ , rather than  $S_t$  itself. This makes the variance increase in a way proportional to  $S_t$ , instead of the square of  $S_t$ . Hence, if the asset price volatility does not change too much when  $S_t$  changes, this model may be more appropriate.

For a discussion of more stochastic models, we refer to [7] and [8].

## 3.2 Mean Reverting Processes

In a commodity market, an important property of spot prices is mean reversion, i.e., prices tend to fluctuate around and drift over time to values determined by the cost of production and the level of demand. To capture the mean reversion property, a class of so-called “mean reverting processes” are widely used. The Itô process that describes mean reverting process takes the form:

$$dS_t = \alpha^*(L^* - S_t)dt + \sigma^* dW_t. \quad (3.10)$$

In the above SDE,  $W_t$  is again a standard Wiener process.  $\alpha^*$  is called “the mean reversion rate” or “the strength of mean reversion”. It governs the speed of mean reversion, and is taken to be strictly positive.  $L^*$  is called “long run mean” or “long term mean”. It is the value around which  $S_t$  tends to oscillate.  $\sigma^*$  is the diffusion term, which governs the magnitude of the noise term. Here, the parameters  $\alpha^*$ ,  $L^*$  and  $\sigma^*$  could be constant, could be time dependent, and could also be random variables. Each choice represents different model and describes different price dynamics.

Notice from the above equation that the drift term is governed by the distance between the current price and the long term mean as well as by the mean reversion rate. If the variable at time  $t$ ,  $S_t$ , is less than the long term mean, the drift term will be positive, resulting in an upward influence on  $S_t$  to approach the higher long term mean value  $L^*$ . Alternatively, if the value of  $S_t$  is greater than the long term mean  $L^*$ , the drift term will be negative, thus exerting a downward influence on  $S_t$ . Over time, this results in a path of  $S_t$  that hovers around the long term mean without getting too far from it. This effect is called “mean reversion”. Note that the greater the mean reversion rate  $\alpha^*$ , the greater is the pull back influence on  $S_t$ , thus the faster  $S_t$  will go back toward the mean.

Here, we give some examples of specific mean reverting models.

The most basic mean reverting process is the *Ornstein-Uhlenbeck process*. This name is due to the paper that first discussed this model by G.E. Uhlenbeck and L.S. Ornstein (see [9]). The SDE is

$$dS_t = \alpha(L - S_t)dt + \sigma dW_t, \quad (3.11)$$

where  $W_t$  is a standard Wiener process,  $\alpha$ ,  $L$  and  $\sigma$  are all positive constant numbers.

Obviously, this model is obtained by simply letting the parameters  $\alpha^*$ ,  $L^*$  and  $\sigma^*$  in (3.10) be all constant.

Since the above stochastic process could take negative values, a little modification is made by replacing the drift term  $\sigma$  in (3.11) with  $\sigma\sqrt{S_t}$ . This ends up with the so-called *Cox-Ingersoll-Ross model* (see [10]). The corresponding SDE is given by

$$dS_t = \alpha(L - S_t)dt + \sigma\sqrt{S_t}dW_t, \quad (3.12)$$

where  $W_t$  is a standard Wiener process,  $\alpha$ ,  $L$  and  $\sigma$  are all positive constant numbers. Notice that the volatility term  $\sigma\sqrt{S_t}$  vanishes when  $S_t$  tends to zero. Provided  $S_0 > 0$  and  $L > 0$ , this guarantees  $S_t \geq 0$  for all  $t \geq 0$  with probability one (see [11]).

One famous mean reverting model is *Schwartz-Ross model*, which was used to model commodity prices (see [12]). The SDE is given by

$$dS_t = \alpha(L - \ln S_t)S_t dt + \sigma S_t dW_t, \quad (3.13)$$

where  $W_t$  is a standard Wiener process,  $\alpha$ ,  $L$  and  $\sigma$  are all positive constant numbers. In this model,  $S_t$  mean reverts to the long term mean  $\hat{L} = e^L$ . Applying the Itô's Lemma for  $X_t = \ln S_t$ , we have

$$dX_t = \alpha\left(L - \frac{\sigma^2}{2\alpha} - X_t\right)dt + \sigma dW_t, \quad (3.14)$$

which is an Ornstein-Uhlenbeck process, and is still in the framework given by (3.10).

Another model for commodity prices, which is more intuitive than the Schwartz-Ross model, involves the reversion to the long term mean of price, instead of the long term mean of the logarithm of the price. This form of mean reverting process was studied by Dixit and Pindyck (see [13]), and is also known as *Dixit-Pindyck model*.

The corresponding SDE is given by

$$dS_t = \alpha(L - S_t)S_t dt + \sigma S_t dW_t, \quad (3.15)$$

where  $W_t$  is a standard Wiener process,  $\alpha$ ,  $L$  and  $\sigma$  are all positive constant numbers.

Actually, some of the parameters  $\alpha^*$ ,  $L^*$  and  $\sigma^*$  in (3.10) can also be supposed to behave stochastically. This will be the case of some 2-factor or 3-factor models.

### 3.3 Models for Gas Prices

In this section, we introduce some gas prices models that we are going to study in the remaining part of this thesis. As shown in the last chapter, mean reversion and seasonality are two main features of gas prices. To capture the behaviour of gas prices, all the models we choose are mean reverting models, and some of them are endowed with seasonal effects.

#### 3.3.1 1-factor Models

The SDE describing generalized 1-factor models<sup>3</sup> is

$$dS_t = \boldsymbol{\alpha}(t, S_t)dt + \boldsymbol{\sigma}(t, S_t)dW_t, \quad (3.16)$$

where  $W_t$  is a standard Wiener process,  $\boldsymbol{\alpha}(t, S_t)$  and  $\boldsymbol{\sigma}(t, S_t)$  are deterministic functions of  $t$  and  $S_t$ .

---

<sup>3</sup>A *factor* represents a market variable that exhibits some form of random behaviour. In these models, there is only one independent random variable, i.e.,  $S_t$ , so they are called 1-factor models.

By selecting different forms of  $\alpha(t, S_t)$  and  $\sigma(t, S_t)$ , we get different 1-factor models. Actually, up to now, the models introduced in the previous sections are all 1-factor models.

Specifically, in this thesis, we consider two classes of 1-factor models:

### **1-factor models without seasonality**

These are perhaps the simplest mean reverting 1-factor models in which we set  $\alpha(t, S_t) = \alpha(L - S_t)$  and  $\sigma(t, S_t) = \sigma S_t^r$ , where  $\alpha$ ,  $L$  and  $\sigma$  are constant, and  $r = 0$  or  $\frac{1}{2}$ . In this case, the SDE is

$$dS_t = \alpha(L - S_t)dt + \sigma S_t^r dW_t. \quad (3.17)$$

When  $r = 0$ , the model is just the Ornstein-Uhlenbeck process introduced above, and when  $r = \frac{1}{2}$ , it is just the Cox-Ingersoll-Ross model. For the sake of convenience, hereafter, we name the model as “model 1f-A” in the case where  $r = 0$ , and “model 1f-B” in the case where  $r = \frac{1}{2}$ .

Here we shall mention that these two models are not the emphases of this thesis. They are too simple, and do not capture the seasonality property of gas prices at all. The reason we study these two relatively simple models is to investigate some significant issues faced in the process of stochastic model calibration.

### **1-factor models with seasonality**

As demonstrated in Chapter 2, seasonality is one of the most remarkable features for gas prices. Hence, to construct a usable model in practice, seasonality has to be included.

It is natural to think that if we strip off the long term trend and seasonal oscillations from the spot prices, then the remaining part of the spot prices may have relatively simpler characteristics. Motivated by this thinking, we consider gas spot price as the sum of a seasonal term  $f(t)$  and an unseasonalized term  $X_t$ , and the latter follows a 1-factor mean reverting process with constant long term mean and time dependent volatility.

In mathematical terms, we let

$$S_t = f(t) + X_t, \quad (3.18)$$

where

$$f(t) = bt + \sum_{i=1}^N \beta_i \cos\left(\frac{2\pi it}{P}\right) + \eta_i \sin\left(\frac{2\pi it}{P}\right), \quad (3.19)$$

and  $X_t$  follows a mean reverting process described by the SDE

$$dX_t = \alpha(L - X_t)dt + \sigma(t)X_t^r dW_t, \quad (3.20)$$

where

$$\sigma(t) = \exp\left(c + \sum_{j=1}^K \lambda_j \cos\left(\frac{2\pi jt}{P}\right) + \omega_j \sin\left(\frac{2\pi jt}{P}\right)\right).$$

In the above equations,  $W_t$  is a Wiener process,  $r = 0, \frac{1}{2}$  or  $1$ , and  $b, \beta_i$ 's,  $\eta_i$ 's,  $\alpha, L, c, \lambda_j$ 's and  $\omega_j$ 's are all constant,  $N$  and  $K$  are positive integers, and  $P$  is the number of trading days in one year<sup>4</sup>.

---

<sup>4</sup>Another popular way to incorporate seasonality into the model is to change the  $L$  in (3.17) into a deterministic function of time  $t$ , and if needed, change  $\sigma$ , as well. In this case, the models can be described as  $dS_t = \alpha(L(t) - S_t)dt + \sigma(t)S_t^r dW_t$ , where we could let  $r = 0, \frac{1}{2}$  or  $1$ . Actually, these models have strong ties with our 1-factor seasonal models. If we let  $r = 0$ , for our corresponding model, we have  $dS = d(f(t)) + dX_t = \alpha(L + \frac{1}{\alpha}f'(t) - X_t)dt + \sigma(t)dW_t = \alpha(L + \frac{1}{\alpha}f'(t) + f(t) - S_t)dt + \sigma(t)dW_t$ . So, in this case, these two models are equivalent, if we let  $L(t) = L + \frac{1}{\alpha}f'(t) + f(t)$ . For discussions about this kind of seasonal models, see [14], Chapter 3.

Notice that we let the volatility term  $\sigma(t)$  be of exponential form in order to ensure that it is positive. The linear term  $bt$  in  $f(t)$  contributes to capture the tendency of the prices. The sine and cosine functions in  $f(t)$  and  $\sigma(t)$  make them move up and down periodically as seasons change. Furthermore, the trigonometric functions corresponding to  $N = 1$  and  $K = 1$  capture the annual seasonality, since the period of these functions is  $P$ , i.e., a year; the trigonometric functions corresponding to  $N = 2$  and  $K = 2$  capture the semiannual seasonality, since the period of these functions is  $P/2$ , i.e., a half year; etc.

Hereafter, we call this model “model 1f-sn-A” for  $r = 0$ , “model 1f-sn-B” for  $r = \frac{1}{2}$  and “model 1f-sn-C” for  $r = 1$ .

### 3.3.2 2-factor Models

The system of SDEs describing generalized 2-factor models is

$$\begin{cases} dX_t = \boldsymbol{\mu}_1(t, X_t, Y_t)dt + \boldsymbol{\sigma}_1(t, X_t, Y_t)dW_t^1 \\ dY_t = \boldsymbol{\mu}_2(t, X_t, Y_t)dt + \boldsymbol{\sigma}_2(t, X_t, Y_t)dW_t^2 \end{cases} \quad (3.21)$$

where  $W_t^i$ 's are standard Wiener processes, for  $i = 1$  and  $2$ , and  $\boldsymbol{\mu}_i(t, X_t, Y_t)$  and  $\boldsymbol{\sigma}_i(t, X_t, Y_t)$  are deterministic functions of  $t, X_t$  and  $Y_t$ , for  $i = 1$  and  $2$ .

In 2-factor models, there are two random variables,  $X_t$  and  $Y_t$ . Specifically, these two variables could be the prices of two correlated assets. For the example given by Bradley [14], he considers a pair of energy prices simultaneously, and investigates the correlation between them. For the two prices, e.g., the natural gas price and electricity price, each one individually obeys a 1-factor mean reverting process.

In another 2-factor framework, one random variable could represent the asset price, with the other random variable representing a parameter involved in the former

one. For example, Pilipovic [15] describes a 2-factor mean reverting model where spot prices revert to a long term mean which is itself a log normally distributed variable. Another example is given by Gibson and Schwartz [1], where the first random variable is the spot price which is assumed to follow a GBM, and the second variable is the instantaneous convenience yield of the spot price which is assumed to follow a mean reverting process. Further works on this model are presented in [12] and [16].

For discussion of more 2-factor models, see [17, 18, 19].

In this thesis, we investigate two classes of 2-factor models as follows:

### 2-factor models without seasonality

The system of SDEs describing these models is

$$\begin{cases} dS_t &= \alpha(L_t - S_t)dt + \sigma S_t^r dW_t^1 \\ dL_t &= \mu(\gamma - L_t)dt + \tau L_t^r dW_t^2 \end{cases} \quad (3.22)$$

where  $\alpha, \sigma, \mu, \gamma$  and  $\tau$  are constant and  $r = 0, \frac{1}{2}$  or 1. For the sake of simplicity, we assume that the Wiener processes  $W_t^1$  and  $W_t^2$  are uncorrelated.

In these models, we enable the spot price long term mean to be a random variable. Compared with the well-known Pilipovic model, the long term mean in these models follows a mean reverting process, instead of a GBM. This idea is based on the thinking that the spot price long term mean, which is usually governed by the cost of production and level of demand, should also stay around a long term equilibrium level, but not have an exponentially increasing trend.

For the remainder of this thesis, we call the model “model 2f-A” for  $r = 0$ , “model

2f-B” for  $r = \frac{1}{2}$  and “model 2f-C” for  $r = 1$ .

### 2-factor models with seasonality

Similar to the 1-factor models with seasonality, we consider the gas spot price as the sum of a seasonal term  $f(t)$  and an unseasonalized term  $X_t$ . The difference is that here  $X_t$  follows a 2-factor mean reverting process.

In mathematical terms, we let

$$S_t = f(t) + X_t, \quad (3.23)$$

where

$$f(t) = bt + \sum_{i=1}^N \beta_i \cos\left(\frac{2\pi it}{P}\right) + \eta_i \sin\left(\frac{2\pi it}{P}\right), \quad (3.24)$$

and  $X_t$  follows a mean reverting process described by the SDEs

$$\begin{cases} dX_t &= \alpha(L_t - X_t)dt + \sigma(t)X_t^r dW_t^1 \\ dL_t &= \mu(\gamma - L_t)dt + \tau L_t^r dW_t^2, \end{cases} \quad (3.25)$$

where

$$\sigma(t) = \exp\left(c + \sum_{j=1}^K \lambda_j \cos\left(\frac{2\pi jt}{P}\right) + \omega_j \sin\left(\frac{2\pi jt}{P}\right)\right).$$

In the above equations,  $W_t^i$ 's are Wiener processes,  $r = 0, \frac{1}{2}$  or  $1$ , and  $b, \beta_i$ 's,  $\eta_i$ 's,  $\alpha, \mu, \gamma, \tau, c, \lambda_j$ 's and  $\omega_j$ 's are all constant,  $N$  and  $K$  are positive integers, and  $P$  is the number of trading days in one year.

Hereafter, we call this model “model 2f-sn-A” for  $r = 0$ , “model 2f-sn-B” for  $r = \frac{1}{2}$  and “model 2f-sn-C” for  $r = 1$ .

# Chapter 4

## Futures Prices

### 4.1 Introduction to Forwards and Futures

A **forward contract** is an agreement where one party promises to buy an asset from another party at some specified time in the future and at some specified price. No money changes hands until the delivery date of the contract. Once the contract is made, it is an obligation to buy (or sell) the asset at the delivery date.

A **futures contract** is very similar to a forward contract. The difference is that futures contracts are usually traded through a formalized exchange, which standardizes the terms of the contracts<sup>1</sup>. The profit from the futures position is calculated every day and the change in this value is paid from one party to the other via the exchange.

A futures vs. forward price bias exists in money markets, where the futures prices are directly related to interest rates. However, in the energy commodity markets, the correlations of the energy futures prices to the interest rates are typically null. Thus the futures and the forward prices are valued in the same manner. They can be used interchangeably, as both reflect the same value. (See [15], page 80.)

---

<sup>1</sup>The most popular natural gas futures contract in North America is traded in NYMEX. It trades in units of 10,000 million British thermal units (mmBtu). The price is based on delivery at the Henry Hub in Louisiana, the nexus of 16 intra- and interstate natural gas pipeline systems that draw supplies from the region's prolific gas deposits. For more details, you can go to <http://www.nymex.com>.

## 4.2 Futures Price Formula

In this section, we show how to find a function that relates the price of futures to the price of the underlying asset and possibly some other market factors. In order to illustrate the approaches clearly, we take a 2-factor model as an example. The model we choose is (3.22), in which we take  $r$  as  $\frac{1}{2}$ . We will introduce several methods to get the futures price  $F(t, T, S_t, L_t)$ , which is a function of the observing time  $t$ , the delivery time  $T$ , the gas spot price  $S_t$  and the corresponding long term mean  $L_t$ .

The SDE describing the model we will work on is given by

$$\begin{cases} dS_t &= \alpha(L_t - S_t)dt + \sigma\sqrt{S_t}dW_t^1 \\ dL_t &= \mu(\gamma - L_t)dt + \tau\sqrt{L_t}dW_t^2, \end{cases} \quad (4.1)$$

where  $\alpha, \sigma, \mu, \gamma$  and  $\tau$  are constant.

### 4.2.1 No Arbitrage Assumption

The pricing of derivatives is based on the so-called “no arbitrage” assumption. This states that prices of different instruments must be related to one another in such a way that they offer no arbitrage opportunities. According to the theory, if market prices ever did offer arbitrage opportunities, arbitrageurs would immediately exploit them, driving prices back to an arbitrage-free state. When arbitrage opportunities do arise, they tend to be small and are soon eliminated.

We know that after decades of deregulation, the natural gas market is quite liquid. So, we think this assumption is reasonable in natural gas futures pricing.

### 4.2.2 Partial Differential Equation Method

In this method, we first construct a risk-free portfolio, and then obtain a partial differential equation (PDE) that is implied by the lack of arbitrage opportunities. Finally, given a boundary condition that futures price must satisfy, the PDE is solved.

- Itô's lemma

Before obtaining the PDE, we need to know an analytical formula that simplifies handling stochastic differentials, Itô's lemma. Since the example we choose is a 2-factor model, here we give the Itô's lemma in 2 dimensions<sup>2</sup>. It states that if we have a smooth function  $V(t, X, Y)$  of time and two random variables, and if  $X_t$  and  $Y_t$  solve SDEs

$$\begin{cases} dX_t = \boldsymbol{\mu}_1(t, X_t, Y_t)dt + \boldsymbol{\sigma}_1(t, X_t, Y_t)dW_t^1 \\ dY_t = \boldsymbol{\mu}_2(t, X_t, Y_t)dt + \boldsymbol{\sigma}_2(t, X_t, Y_t)dW_t^2, \end{cases} \quad (4.2)$$

where  $dW_t^1$  and  $dW_t^2$  are standard Brownian motions in  $\mathbb{R}$  with correlation  $\rho$ , then, with  $V_t = V(t, X_t, Y_t)$ , the increment  $dV_t$  is given by

$$dV_t = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial X}dX_t + \frac{\partial V}{\partial Y}dY_t + \frac{1}{2}\boldsymbol{\sigma}_1^2 \frac{\partial^2 V}{\partial X^2}dt + \frac{1}{2}\boldsymbol{\sigma}_2^2 \frac{\partial^2 V}{\partial Y^2}dt + \rho\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2 \frac{\partial^2 V}{\partial X \partial Y}dt, \quad (4.3)$$

where the dependence on  $(t, X_t, Y_t)$  is dropped for the sake of brevity.

- PDE

We suppose that, at time  $t$  with asset price  $S$  and long term mean  $L$ , the value of a futures contract entered into at time  $t_0 \leq t$  with delivery date  $T \geq t$  is given by the function  $V(t, S, L; t_0, T)$ . Now, we attempt to construct a risk-free portfolio

---

<sup>2</sup>For Itô's lemma in 1 dimension, see Section 3.1.5

containing this futures contract. Since we have two sources of randomness, we must hedge with two other contracts (differing only in the delivery time) to eliminate the risk. We construct a portfolio whose value at time  $t$  will be given by

$$\Pi_t = V(t, S_t, L_t; t_0, T) + \Delta_1 V(t, S_t, L_t; t_0, T_1) + \Delta_2 V(t, S_t, L_t; t_0, T_2). \quad (4.4)$$

(Hereafter, we denote  $V(t, S_t, L_t; t_0, T_1)$  by  $V_1$ ,  $V(t, S_t, L_t; t_0, T_2)$  by  $V_2$ , and  $V(t, S_t, L_t; t_0, T)$  by  $V_t$ .)

By Itô's lemma, we have <sup>3</sup>

$$\begin{aligned} dV_t &= \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S} dS_t + \frac{\partial V}{\partial L} dL_t + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2} dt + \frac{1}{2} \tau^2 L_t \frac{\partial^2 V}{\partial L^2} dt \\ &= \left( \frac{\partial V}{\partial t} + \alpha(L_t - S_t) \frac{\partial V}{\partial S} + \mu(\gamma - L_t) \frac{\partial V}{\partial L} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2} + \frac{1}{2} \tau^2 L_t \frac{\partial^2 V}{\partial L^2} \right) dt \\ &\quad + \frac{\partial V}{\partial S} \sigma \sqrt{S_t} dW_t^1 + \frac{\partial V}{\partial L} \tau \sqrt{L_t} dW_t^2. \end{aligned} \quad (4.5)$$

where the dependence of  $(t, S_t, L_t)$  is dropped for the sake of brevity.

$dV_1$  and  $dV_2$  have similar form.

(Hereafter, we denote  $\frac{\partial}{\partial t} + \alpha(L_t - S_t) \frac{\partial}{\partial S} + \mu(\gamma - L_t) \frac{\partial}{\partial L} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2}{\partial S^2} + \frac{1}{2} \tau^2 L_t \frac{\partial^2}{\partial L^2}$  by  $D$ .)

Thus the change in the value of this portfolio is given by

$$\begin{aligned} d\Pi_t &= dV + \Delta_1 dV_1 + \Delta_2 dV_2 \\ &= (DV + \Delta_1 DV_1 + \Delta_2 DV_2) dt \\ &\quad + \sigma \sqrt{S_t} \left( \frac{\partial V}{\partial S} + \Delta_1 \frac{\partial V_1}{\partial S} + \Delta_2 \frac{\partial V_2}{\partial S} \right) dW_t^1 + \tau \sqrt{L_t} \left( \frac{\partial V}{\partial L} + \Delta_1 \frac{\partial V_1}{\partial L} + \Delta_2 \frac{\partial V_2}{\partial L} \right) dW_t^2. \end{aligned} \quad (4.6)$$

At time  $t$ , we can find  $\Delta_1$  and  $\Delta_2$  such that

$$\frac{\partial V}{\partial S} + \Delta_1 \frac{\partial V_1}{\partial S} + \Delta_2 \frac{\partial V_2}{\partial S} = 0 \quad (1)$$

$$\frac{\partial V}{\partial L} + \Delta_1 \frac{\partial V_1}{\partial L} + \Delta_2 \frac{\partial V_2}{\partial L} = 0 \quad (2),$$

---

<sup>3</sup>Notice that the correlation  $\rho$  of  $W_t^1$  and  $W_t^2$  in (4.1) is 0.

then  $d\Pi_t$  is a deterministic term, thus is risk free. If the risk-free rate of interest is  $r$ , then we have  $d\Pi_t = r\Pi_t dt$ , and hence

$$(DV - rV) + \Delta_1(DV_1 - rV_1) + \Delta_2(DV_2 - rV_2) = 0 \quad (3).$$

Putting (1), (2) and (3) together, we have

$$\begin{bmatrix} DV - rV & DV_1 - rV_1 & DV_2 - rV_2 \\ \frac{\partial V}{\partial S} & \frac{\partial V_1}{\partial S} & \frac{\partial V_2}{\partial S} \\ \frac{\partial V}{\partial V} & \frac{\partial V_1}{\partial V} & \frac{\partial V_2}{\partial V} \\ \frac{\partial V}{\partial L} & \frac{\partial V_1}{\partial L} & \frac{\partial V_2}{\partial L} \end{bmatrix} \begin{bmatrix} 1 \\ \Delta_1 \\ \Delta_2 \end{bmatrix} = 0.$$

Because of the existence of  $\Delta_1$  and  $\Delta_2$ , we have the determinant of the matrix is 0, which implies that the rows are linearly dependent. Hence, there must exist  $\lambda_S$  and  $\lambda_L$ , independent of the delivery time, such that

$$DV - rV - \lambda_S \frac{\partial V}{\partial S} - \lambda_L \frac{\partial V}{\partial L} = 0. \quad (4.7)$$

Expanding the operator  $D$ , and noting that the equation must hold for all possible values of  $S_t$  and  $L_t$ , the above PDE becomes (for  $S > 0$ ,  $L > 0$  and  $t \in (t_0, T)$ ).

$$\begin{aligned} \frac{\partial V}{\partial t} + (\alpha(L - S) - \lambda_S) \frac{\partial V}{\partial S} + (\mu(\gamma - L) - \lambda_L) \frac{\partial V}{\partial L} \\ + \frac{1}{2} \sigma^2 S \frac{\partial^2 V}{\partial S^2} + \frac{1}{2} \tau^2 L \frac{\partial^2 V}{\partial L^2} - rV = 0. \end{aligned} \quad (4.8)$$

In the above PDE,  $(\alpha(L - S) - \lambda_S)$  and  $(\mu(\gamma - L) - \lambda_L)$  are called the risk-adjusted or risk-neutralized drift terms for  $S$  and  $L$ , respectively<sup>4</sup>. Although in general  $\lambda_S$  and  $\lambda_L$  could be arbitrary functions of  $S$ ,  $L$  and  $t$ , for the sake of simplicity, we model  $\lambda_S = (\alpha - \tilde{\alpha})(L - S)$  and  $\lambda_L = \mu(\gamma - L) - \tilde{\mu}(\tilde{\gamma} - L)$ , so that  $(\alpha(L - S) - \lambda_S) = \tilde{\alpha}(L - S)$

---

<sup>4</sup>For this model, if we let  $\lambda_S = \lambda_1 \sigma \sqrt{S}$  and  $\lambda_L = \lambda_2 \tau \sqrt{L}$ , then  $\lambda_1$  and  $\lambda_2$  are the market prices of risk for  $S$  and  $L$ , respectively.

and  $(\mu(\gamma - L) - \lambda_L) = \tilde{\mu}(\tilde{\gamma} - L)$ . Here,  $\tilde{\alpha}$ ,  $\tilde{\mu}$  and  $\tilde{\gamma}$  are the risk-neutralized parameters.

Then (4.8) becomes

$$\frac{\partial V}{\partial t} + \tilde{\alpha}(L - S)\frac{\partial V}{\partial S} + \tilde{\mu}(\tilde{\gamma} - L)\frac{\partial V}{\partial L} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\tau^2 L^2 \frac{\partial^2 V}{\partial L^2} - rV = 0. \quad (4.9)$$

• Solution

We look for a solution  $V(\cdot)$  in linear form, say  $V(t, S_t, L_t; t_0, T) = a(t; t_0, T)S_t + b(t; t_0, T)L_t + c(t; t_0, T)$ . By plugging it into (4.9) and comparing the coefficients of powers of  $S_t$  and  $L_t$  on both sides, we obtain the ODEs

$$\begin{cases} \frac{\partial a}{\partial t} = (r + \tilde{\alpha})a \\ \frac{\partial b}{\partial t} = -\tilde{\alpha}a + (r + \tilde{\mu})b \\ \frac{\partial c}{\partial t} = rc - \tilde{\mu}\tilde{\gamma}b. \end{cases} \quad (4.10)$$

Since under the no arbitrage assumption, at the delivery time  $T$ , the payoff of the futures contract is the difference between spot price and the original futures price decided at time  $t_0$ , say,  $V(T, S_T, L_T; t_0, T) = S_T - F(t_0, T, (S_{t_0}, L_{t_0}))$ , where  $F(\cdot)$  is the price of the futures mentioned above, we have the boundary condition

$$\begin{cases} a(T; t_0, T) = 1 \\ b(T; t_0, T) = 0 \\ c(T; t_0, T) = -F(t_0, T, (S_{t_0}, L_{t_0})). \end{cases} \quad (4.11)$$

Solving the above initial value problem, we get

$$\begin{cases} a = e^{(r+\tilde{\alpha})(t-T)} \\ b = \frac{\tilde{\alpha}}{\tilde{\alpha} - \tilde{\mu}}(e^{(r+\tilde{\mu})(t-T)} - e^{(r+\tilde{\alpha})(t-T)}) \\ c = \frac{\tilde{\mu}\tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}}(e^{(r+\tilde{\alpha})(t-T)} - e^{r(t-T)}) - \frac{\tilde{\alpha}\tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}}(e^{(r+\tilde{\mu})(t-T)} - e^{r(t-T)}) \\ \quad - F(t_0, T, (S_{t_0}, L_{t_0}))e^{r(t-T)}. \end{cases} \quad (4.12)$$

Thus, we have

$$\begin{aligned}
V(t_0, S_{t_0}, L_{t_0}; t_0, T) &= e^{(r+\tilde{\alpha})(t_0-T)} S_{t_0} + \frac{\tilde{\alpha}}{\tilde{\alpha} - \tilde{\mu}} (e^{(r+\tilde{\mu})(t_0-T)} - e^{(r+\tilde{\alpha})(t_0-T)}) L_{t_0} \\
&\quad + \frac{\tilde{\mu}\tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}} (e^{(r+\tilde{\alpha})(t_0-T)} - e^{r(t_0-T)}) - \frac{\tilde{\alpha}\tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}} (e^{(r+\tilde{\mu})(t_0-T)} - e^{r(t_0-T)}) \\
&\quad - F(t_0, T, (S_{t_0}, L_{t_0})) e^{r(t_0-T)}.
\end{aligned} \tag{4.13}$$

When the futures contract is made no money changes hands, which implies that the initial value of the contract is zero, i.e.,  $V(t_0, S_{t_0}, L_{t_0}; t_0, T) = 0$ . Thus, we obtain the futures price

$$\begin{aligned}
F(t_0, T, (S_{t_0}, L_{t_0})) &= e^{\tilde{\alpha}(t_0-T)} S_{t_0} + \frac{\tilde{\alpha}}{\tilde{\alpha} - \tilde{\mu}} (e^{\tilde{\mu}(t_0-T)} - e^{\tilde{\alpha}(t_0-T)}) L_{t_0} \\
&\quad + \frac{\tilde{\mu}\tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}} (e^{\tilde{\alpha}(t_0-T)} - 1) - \frac{\tilde{\alpha}\tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}} (e^{\tilde{\mu}(t_0-T)} - 1),
\end{aligned}$$

which means that we can write<sup>5</sup>

$$\begin{aligned}
F^{\tilde{\theta}}(t, T, (S_t, L_t)) &= e^{\tilde{\alpha}(t-T)} S_t + \frac{\tilde{\alpha}}{\tilde{\alpha} - \tilde{\mu}} (e^{\tilde{\mu}(t-T)} - e^{\tilde{\alpha}(t-T)}) L_t \\
&\quad + \frac{\tilde{\mu}\tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}} (e^{\tilde{\alpha}(t-T)} - 1) - \frac{\tilde{\alpha}\tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}} (e^{\tilde{\mu}(t-T)} - 1).
\end{aligned} \tag{4.14}$$

### 4.2.3 Method of Equivalent Measures

In this section, we introduce a pricing method that utilizes the no arbitrage assumption in a different way. This method is quite different from the PDE method, but leads to the same formula as (4.14). It is more difficult to visualize, but for the computation, it is more flexible and easier to handle.

The main idea of method of the equivalent martingale measures is that we can use the notion of arbitrage to determine a probability measure under which financial

---

<sup>5</sup>Hereafter, we denote the set of risk-neutralized parameters by  $\tilde{\theta}$ . In this example,  $\tilde{\theta} = [\tilde{\alpha}, \tilde{\mu}, \tilde{\gamma}]$ .

assets behave as martingales<sup>6</sup> if they are discounted properly, and then one can calculate arbitrage-free prices easily by evaluating the implied expectations. This method involves the transform of the probability distributions of the underlying assets using the tools provided by the Girsanov theorem. In this thesis, we are not going to give more details about this idea<sup>7</sup>.

Here we just state that the futures price can be considered as the expectation of the spot price in a newly measured world, the risk-neutral world. In the risk-neutral world, investors do not demand any premium to take on extra risk. Thus the expected change in value of all assets will be the risk-free rate of return.

To provide a better understanding of the above statement, we give a simple example showing the relationship between futures price and spot price.

Suppose a stock does not pay any dividends and follows a simple lognormal model, with a rate of return  $\mu$  and a volatility  $\sigma$ :

$$dS_t = \mu S_t dt + \sigma S_t dW_t. \quad (4.15)$$

On the other hand, introduced in almost every elementary book in financial mathematics or financial engineering, the price of a futures contract at time  $t$  with the delivery time  $T$  is given by  $F(t, T, S_t) = S_t e^{r(T-t)}$ , where  $r$  is the risk-free interest rate. If the above relationship between  $F(t, T, S_t)$  and  $S_t$  does not hold, then an arbitrage opportunity appears<sup>8</sup>. Here we notice that  $S_t e^{r(T-t)}$  is not the expectation

---

<sup>6</sup>Suppose at time  $t$  one has information summarized by  $I_t$ . A random variable  $X_t$  satisfies the equality  $\mathbb{E}^P[X_{t+s}|I_t] = X_t$  for all  $s > 0$ , is called a *martingale with respect to the probability  $P$* .

<sup>7</sup>For more details, refer to [20], Chapters 4, 5, 6, 14 and 15.

<sup>8</sup>If  $F(t, T, S_t) > S_t e^{r(T-t)}$ , then at time  $t$ , one can sell a futures contract, and meanwhile borrow money from bank for an amount  $S_t$  and use it to buy a share of stock, then he holds the stock until  $T$ , he performs the futures contract with his stock, gets money  $F(t, T, S_t)$  and returns it to the bank  $S_t e^{r(T-t)}$ , then he can get a profit  $F(t, T, S_t) - S_t e^{r(T-t)}$  without any risk. Oppositely,

of the spot price described by (4.15) in the real world. Actually, it can be considered as the expectation of a spot price following a lognormal model with a rate of return  $r$ , say,

$$dS_t = rS_t dt + \sigma S_t dW_t. \quad (4.16)$$

Here, the spot price is newly measured in the risk-neutral world by simply replacing the real drift rate  $\mu$  with the the risk-neutralized drift rate  $r$ .

Accepting that the futures price can be considered as the expectation of the spot price in risk-neutral world, we move to our problem: derive the futures formula for the 2-factor model (4.1). Similar to the method shown above, we neutralize our model by simply replacing the parameters in (4.1) with the corresponding risk-neutralized parameters. The risk-neutralized model is taken to be

$$\begin{cases} dS_t &= \tilde{\alpha}(L_t - S_t)dt + \sigma\sqrt{S_t}dW_t^1 \\ dL_t &= \tilde{\mu}(\tilde{\gamma} - L_t)dt + \tau\sqrt{L_t}dW_t^2, \end{cases} \quad (4.17)$$

where  $\tilde{\alpha}$ ,  $\tilde{\mu}$ , and  $\tilde{\gamma}$  are the risk-neutralized parameters. For the relationship with real parameters, see Section 4.2.2.

Now, to price the futures contract, our task is to compute the expectation of spot price under the SDEs (4.17). The simplest way is to take the expectation of the two sides of each equation in (4.17). Doing this, we find the randomness terms

---

if  $F(t, T, S_t) < S_t e^{r(T-t)}$ , then at time  $t$ , one can buy a futures contract, and meanwhile borrow a share of stock from someone and sell it in the market immediately to get money  $S_t$ , then he save the money into a bank. At time  $T$ , he performs the futures contract getting a share of stock at the price  $F(t, T, S_t)$  and returns it to that person. Then he can get a risk less profit  $S_t e^{r(T-t)} - F(t, T, S_t)$  again.

are canceled<sup>9</sup>, and it turns out that

$$\begin{cases} d\bar{S}_t &= \tilde{\alpha}(\bar{L}_t - \bar{S}_t)dt \\ d\bar{L}_t &= \tilde{\mu}(\tilde{\gamma} - \bar{L}_t)dt, \end{cases} \quad (4.18)$$

where  $\bar{S}_t$  and  $\bar{L}_t$  are the expectation values of  $S_t$  and  $L_t$  at time  $t$  respectively. Solving the second ODE for  $\bar{L}_t$  with boundary condition that at  $t = t_0$ ,  $\bar{L}_t = L_{t_0}$ , we obtain

$$\bar{L}_t = (L_{t_0} - \tilde{\gamma})e^{\tilde{\mu}(t_0-t)} + \tilde{\gamma}. \quad (4.19)$$

Then we plug it into the first ODE for  $\bar{S}_t$ . With boundary condition that  $t = t_0$ ,  $\bar{S}_t = S_{t_0}$ , we obtain that

$$\begin{aligned} \bar{S}_t &= S_{t_0}e^{\tilde{\alpha}(t_0-t)} + \frac{\tilde{\alpha}}{\tilde{\alpha} - \tilde{\mu}}(L_{t_0} - \tilde{\gamma})(e^{\tilde{\mu}(t_0-t)} - e^{\tilde{\alpha}(t_0-t)}) - \tilde{\gamma}(1 - e^{\tilde{\alpha}(t_0-t)}) \\ &= e^{\tilde{\alpha}(t-T)}S_{t_0} + \frac{\tilde{\alpha}}{\tilde{\alpha} - \tilde{\mu}}(e^{\tilde{\mu}(t_0-t)} - e^{\tilde{\alpha}(t_0-t)})L_{t_0} + \frac{\tilde{\mu}\tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}}(e^{\tilde{\alpha}(t_0-t)} - 1) - \frac{\tilde{\alpha}\tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}}(e^{\tilde{\mu}(t_0-t)} - 1). \end{aligned} \quad (4.20)$$

Since  $F^{\tilde{\theta}}(t, T, (S_t, L_t)) = \bar{S}_t$ , and at initial time  $t_0 = t$ ,  $\bar{S}_{t_0} = S_t$  and  $\bar{L}_{t_0} = L_t$ , we have

$$\begin{aligned} F^{\tilde{\theta}}(t, T, (S_t, L_t)) &= e^{\tilde{\alpha}(t-T)}S_t + \frac{\tilde{\alpha}}{\tilde{\alpha} - \tilde{\mu}}(e^{\tilde{\mu}(t-T)} - e^{\tilde{\alpha}(t-T)})L_t \\ &\quad + \frac{\tilde{\mu}\tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}}(e^{\tilde{\alpha}(t-T)} - 1) - \frac{\tilde{\alpha}\tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}}(e^{\tilde{\mu}(t-T)} - 1), \end{aligned} \quad (4.21)$$

which is the same formula as (4.14).

If we let  $F_t = F^{\tilde{\theta}}(t, T, (S_t, L_t))$ , then it is not hard to verify that under the measure  $\tilde{\theta}$ ,  $F_t$  is a martingale<sup>10</sup>. This is why the method is called “method of equivalent measures”.

---

<sup>9</sup>For more details about why the terms are canceled, see Theorem 3.2.1 in [21].

<sup>10</sup>To show  $F_t$  is a martingale, we need to verify  $E^{\tilde{\theta}}[F_{t^*}|F_t] = F_t$ , for  $t^* \in (t, T]$ . The left hand side involves replacing  $t$ ,  $S_t$  and  $L_t$  with  $t^*$ ,  $\bar{S}_{t^*}$  and  $\bar{L}_{t^*}$ , respectively, in (4.21). And  $\bar{S}_{t^*}$  and  $\bar{L}_{t^*}$  can be obtained from (4.20) and (4.19). The computation is not very hard.

There are also some other ways to get the expectation of  $S_t$  described by (4.17). One of them is using the affine process transform theory, which we outline in Appendix B.

#### 4.2.4 The Other Futures Price Formulas

By applying the above methods for futures pricing, we can obtain futures price formulas for the other models introduced in Chapter 3. For the sake of brevity, here we just show the results:

- 1-factor models without seasonality

$$F^{\tilde{\theta}}(t, T, S_t) = e^{\tilde{\alpha}(t-T)}(S_t - \tilde{L}) + \tilde{L}. \quad (4.22)$$

- 1-factor models with seasonality

$$F^{\tilde{\theta}}(t, T, S_t) = e^{\tilde{\alpha}(t-T)}(S_t - \tilde{L} - f(t)) + \tilde{L} + f(T). \quad (4.23)$$

- 2-factor models without seasonality

$$\begin{aligned} F^{\tilde{\theta}}(t, T, (S_t, L_t)) &= e^{\tilde{\alpha}(t-T)}S_t + \frac{\tilde{\alpha}}{\tilde{\alpha} - \tilde{\mu}}(e^{\tilde{\mu}(t-T)} - e^{\tilde{\alpha}(t-T)})L_t \\ &+ \frac{\tilde{\mu}\tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}}(e^{\tilde{\alpha}(t-T)} - 1) - \frac{\tilde{\alpha}\tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}}(e^{\tilde{\mu}(t-T)} - 1). \end{aligned} \quad (4.24)$$

- 2-factor models with seasonality

$$\begin{aligned} F^{\tilde{\theta}}(t, T, (S_t, L_t)) &= f(T) + e^{\tilde{\alpha}(t-T)}(S_t - f(t)) + \frac{\tilde{\alpha}}{\tilde{\alpha} - \tilde{\mu}}(e^{\tilde{\mu}(t-T)} - e^{\tilde{\alpha}(t-T)})L_t \\ &+ \frac{\tilde{\mu}\tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}}(e^{\tilde{\alpha}(t-T)} - 1) - \frac{\tilde{\alpha}\tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}}(e^{\tilde{\mu}(t-T)} - 1). \end{aligned} \quad (4.25)$$

#### 4.2.5 An Adjustment to the Theoretical Futures Price

We know that in real markets the futures price refers to the price being paid for gas delivery throughout its delivery period, usually from the beginning of one month to the end of that month. This fact motivates us to think about the observed futures price as corresponding to the average theoretical price during its delivery period.

Thus, the adjusted theoretical futures price is given by

$$G^{\tilde{\theta}}(t, T_i, T_{i+1}, X_t) = \frac{1}{T_{i+1} - T_i} \int_{T_i}^{T_{i+1}} F^{\tilde{\theta}}(t, T, X_t) dT, \quad (4.26)$$

where  $F^{\tilde{\theta}}(t, T, X_t)$  is given by the corresponding function in Section 4.2.4;  $G^{\tilde{\theta}}(t, T_i, T_{i+1}, X_t)$  is the average futures prices in its delivery period between  $T_i$  and  $T_{i+1}$ , which is considered as the observed futures price with the starting delivery time  $T_i$ ;  $X_t$  denotes the stochastic factors, which are  $[S_t]$  and  $[S_t, L_t]$  for 1-factor and 2-factor models, respectively.

If we let  $A(x) = \frac{e^{x(t-T_i)} - e^{x(t-T_{i+1})}}{x(T_{i+1} - T_i)}$ , then

for the 1-factor models without seasonality,

$$G^{\tilde{\theta}}(t, T_i, S_t) = A(\tilde{\alpha})(S_t - \tilde{L}) + \tilde{L}; \quad (4.27)$$

for the 1-factor models with seasonality,

$$\begin{aligned} G^{\tilde{\theta}}(t, T_i, (S_t, L_t)) = & \sum_{j=1}^n \left( \beta_j \frac{\sin(P_j T_{i+1}) - \sin(P_j T_i)}{P_j (T_{i+1} - T_i)} + \eta_j \frac{\cos(P_j T_i) - \cos(P_j T_{i+1})}{P_j (T_{i+1} - T_i)} \right) \\ & + \frac{b(T_i + T_{i+1})}{2} + \tilde{L} + (S_t - \tilde{L} - f(t)) A(\tilde{\alpha}); \end{aligned} \quad (4.28)$$

for the 2-factor models without seasonality,

$$\begin{aligned} G^{\tilde{\theta}}(t, T_i, (S_t, L_t)) = & S_t A(\tilde{\alpha}) + \frac{\tilde{\alpha}}{\tilde{\alpha} - \tilde{\mu}} (A(\tilde{\mu}) - A(\tilde{\alpha})) L_t \\ & + \frac{\tilde{\mu} \tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}} (A(\tilde{\alpha}) - 1) - \frac{\tilde{\alpha} \tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}} (A(\tilde{\mu}) - 1); \end{aligned} \quad (4.29)$$

for the 2-factor models with seasonality,

$$\begin{aligned}
G^{\tilde{\theta}}(t, T_i, (S_t, L_t)) = & \sum_{j=1}^n \left( \beta_j \frac{\sin(P_j T_{i+1}) - \sin(P_j T_i)}{P_j(T_{i+1} - T_i)} + \eta_j \frac{\cos(P_j T_i) - \cos(P_j T_{i+1})}{P_j(T_{i+1} - T_i)} \right) \\
& + \frac{b(T_i + T_{i+1})}{2} + (S_t - f(t))A(\tilde{\alpha}) + \frac{\tilde{\alpha}}{\tilde{\alpha} - \tilde{\mu}} (A(\tilde{\mu}) - A(\tilde{\alpha}))L_t \\
& + \frac{\tilde{\mu}\tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}} (A(\tilde{\alpha}) - 1) - \frac{\tilde{\alpha}\tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}} (A(\tilde{\mu}) - 1).
\end{aligned} \tag{4.30}$$

However, in practice, we find this adjustment makes only a small difference and is not implemented in the calibration experiments reported in the next chapter.

# Chapter 5

## Calibration of the Models

### 5.1 Maximum Likelihood Estimation

In this chapter, we discuss how the parameters in the models we have been considering are estimated from historical data. The approach used is the *maximum likelihood* (ML) method. It involves choosing values for the parameters that maximize the chance (or likelihood) of the data occurring.

ML is a popular method of estimating the parameters of stochastic processes when the probability density function (PDF) for the stochastic variable can be written down analytically. It provides a *consistent* approach to parameter estimation problems and ML estimators become *minimum variance unbiased* estimators as the sample size increases.<sup>1</sup>

Suppose that  $X$  is a continuous random variable with PDF  $f(\theta, X)$ , where  $\theta = \{\theta_1, \dots, \theta_k\}$  are  $k$  unknown parameters which need to be estimated. Given a set of observations  $\{X_t\}_{t=1}^n$ , the log likelihood function at the sample is given by

$$\mathcal{L}(\theta, \{X_t\}_{t=1}^n) = \sum_{t=1}^n \ln f(\theta, X_t). \quad (5.1)$$

The ML estimators of  $\theta$  are obtained by maximizing  $\mathcal{L}(\cdot)$ , i.e.,

$$\hat{\theta} = \operatorname{argmax}_{\theta \in \mathbb{R}^k} \mathcal{L}(\theta, \{X_t\}_{t=1}^n). \quad (5.2)$$

---

<sup>1</sup>Any statistic that converges in probability to a parameter is called a consistent estimator of that parameter. “Minimum variance” means that the estimator has the smallest variance among all estimators of the parameter, and “unbiased” means that the mathematical expectation of the estimator of a parameter is equal to the parameter (see [22]).

### 5.1.1 Characteristic Function and Probability Density Function for Affine Jump-Diffusion Process

To apply the ML method, it is useful to have an analytical form of the PDF for the stochastic variable. Since most of our models fall into the class of affine jump-diffusion processes<sup>2</sup>, this can be achieved by using the affine jump-diffusion process transform theory, described in Appendix A.2.

Suppose a stochastic process  $\{\mathbf{X}_t\}_{t \geq 0}$  is an affine jump-diffusion process in  $\mathbb{R}^n$ . According to the theory introduced in [23] (see Appendix A.2), by setting  $\mathbf{u} = i\mathbf{s}$  and  $R(\mathbf{X}_t, t) = 0$  in (A.3), one can obtain a closed-form expression for the characteristic function of  $\mathbf{X}_T$  conditional on  $\mathbf{X}_t$  ( $T > t$ ):

$$\phi^\theta(\mathbf{s}, t, T, \mathbf{X}_t) = E^\theta[\exp(i\mathbf{s} \cdot \mathbf{X}_T) | \mathbf{X}_t] = \Phi^\theta(i\mathbf{s}, t, T, \mathbf{X}_t), \quad i = \sqrt{-1}. \quad (5.3)$$

The characteristic function is unique and contains the same information as the PDF which can be recovered through the inverse Fourier transform. Thus one can exploit knowledge of  $\phi^\theta(\cdot)$  to derive the PDF of  $X_T$  conditional on  $X_t$ :

$$f(\theta, \mathbf{X}_T | \mathbf{X}_t) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{-i\mathbf{s} \cdot \mathbf{X}_T} \phi^\theta(\mathbf{s}, t, T, \mathbf{X}_t) d\mathbf{s}. \quad (5.4)$$

This PDF can be used to construct ML estimators for our models.

## 5.2 Calibration for 1-factor Models without Seasonality

In this class of models, gas spot prices follow the mean reverting process described by (3.17). By setting  $r = 0$  and  $r = \frac{1}{2}$ , respectively, we have two specific models:

---

<sup>2</sup>For the definition of affine jump-diffusion process, refer to Appendix A.1.

**model 1f-A:**

$$dS_t = \alpha(L - S_t)dt + \sigma dW_t. \quad (5.5)$$

**model 1f-B:**

$$dS_t = \alpha(L - S_t)dt + \sigma\sqrt{S_t}dW_t. \quad (5.6)$$

As mentioned before, these two models are not themselves the emphases of this thesis. In this section, we will just take them as examples as we discuss some significant issues faced in the process of calibration for all the models.

### 5.2.1 Probability Density Function

#### Model 1f-A

- In continuous-time framework:

According to the Definition A.1, the price process  $S_t$  described by (5.5) is an affine jump-diffusion process. Thus we can deduce the conditional PDF of  $S_{t+1}$  given  $S_t$  using (5.4).

For (5.5), in (A.2),  $\mathbf{X}_t = S_t$ ,

$$\mathbf{k}_0 = \alpha L, \quad \mathbf{K}_1 = -\alpha, \quad \mathbf{H}_0 = \sigma^2, \quad \mathbf{H}_1 = 0, \quad l_0 = 0, \quad \mathbf{l}_1 = 0, \quad \nu = 0.$$

By (A.4) - (A.6) and (5.3), we have the characteristic function given by

$$\begin{aligned} \phi^\theta(s, t, T, S_t) &= E^\theta[e^{isS_T} | S_t] \\ &= \Phi^\theta(is, t, T, S_t) \\ &= \exp\left(A(t, T, s) + B(t, T, s)S_t\right), \end{aligned} \quad (5.7)$$

where  $A(t, T, s)$  and  $B(t, T, s)$  satisfy the following equations

$$\begin{cases} \partial_t A = -\alpha LB - \frac{1}{2}\sigma^2 B^2 \\ \partial_t B = \alpha B, \end{cases} \quad (5.8)$$

with the boundary conditions

$$\begin{cases} A(T, T, s) = 0, \\ B(T, T, s) = is. \end{cases} \quad (5.9)$$

Solving the initial value problem (5.8)-(5.9), we have

$$\begin{cases} A(t, T, s) = -isL(e^{\alpha(t-T)} - 1) + \frac{1}{4\alpha}\sigma^2s^2(e^{2\alpha(t-T)} - 1) \\ B(t, T, s) = ise^{\alpha(t-T)}. \end{cases} \quad (5.10)$$

From (5.4) and (5.7), the conditional PDF of  $S_{t+1}$  given  $S_t$  is

$$f(\theta, S_T|S_t) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-isS_T} \exp\left(A(t, T, s) + B(t, T, s)S_t\right) ds. \quad (5.11)$$

This integral is not very hard to evaluate in this case. After some computation (see Appendix C), we have

$$f(\theta, S_T|S_t) = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{2\alpha}{\sigma^2(1 - e^{2\alpha(t-T)})}} \exp\left(-\frac{\alpha((1 - e^{\alpha(t-T)})L + e^{\alpha(t-T)}S_t - S_T)^2}{\sigma^2(1 - e^{2\alpha(t-T)})}\right). \quad (5.12)$$

- In discrete-time framework

Actually, if the point  $T$  is close to  $t$ , then (5.8) can be approximated by the following expression,

$$\begin{cases} \frac{A(t, T, s) - A(T, T, s)}{t - T} \approx -\alpha LB(T, T, s) - \frac{1}{2}\sigma^2 B(T, T, s)^2 \\ \frac{B(t, T, s) - B(T, T, s)}{t - T} \approx \alpha B(T, T, s). \end{cases} \quad (5.13)$$

Together with the boundary condition (5.9), the values of  $A(t, T, s)$  and  $B(t, T, s)$  can be approximated by

$$\begin{cases} A(t, T, s) \approx \left(-is\alpha L + \frac{1}{2}\sigma^2s^2\right)(t - T) \\ B(t, T, s) \approx is + i\alpha s(t - T). \end{cases} \quad (5.14)$$

Thus the conditional PDF of  $S_T$  given  $S_t$  can be approximated by

$$\begin{aligned} f(\theta, S_T|S_t) &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{-isS_T} \exp\left(A(t, T, s) + B(t, T, s)S_t\right) ds \\ &\approx \frac{1}{2\pi} \int_{\mathbb{R}} \exp\left(is\left((1 + \alpha(t - T))S_t - \alpha L(t - T) - S_T\right) + \frac{1}{2}\sigma^2(t - T)s^2\right) ds. \end{aligned} \quad (5.15)$$

After some computation (see Appendix C), we have

$$f(\theta, S_T|S_t) \approx \frac{1}{\sqrt{2\pi}\sigma\sqrt{T-t}} \exp\left(-\frac{\left((1 + \alpha(t - T))S_t - \alpha L(t - T) - S_T\right)^2}{2\sigma^2(T-t)}\right). \quad (5.16)$$

Actually, there is an easier way to deduce the above approximate PDF. This time, let's start from the SDE (5.5). We know that the corresponding discrete form of this SDE in Euler scheme is given by

$$S_T - S_t = \alpha(L - S_t)(T - t) + \sigma(W_T - W_t). \quad (5.17)$$

And we know that  $W_T - W_t \sim \mathcal{N}(0, T - t)$ . Therefore, given  $S_t$ , we have  $S_T \sim \mathcal{N}(\alpha(L - S_t)(T - t) + S_t, \sigma^2(T - t))$ . Then the PDF (5.16) is obtained naturally.

### Comparison of two probability density functions

Putting (5.12) and (5.16) together, we find that the latter approximates the former by replacing  $e^{\alpha(t-T)}$  with  $1 + \alpha(t - T)$  and  $e^{2\alpha(t-T)}$  with  $1 + 2\alpha(t - T)$ . If suppose  $\alpha(T - t)$  is not a big number, then this replacement will not make big errors. Figure 5.1 gives a visual sense about how big the difference between the two PDFs can be in a typical setting.

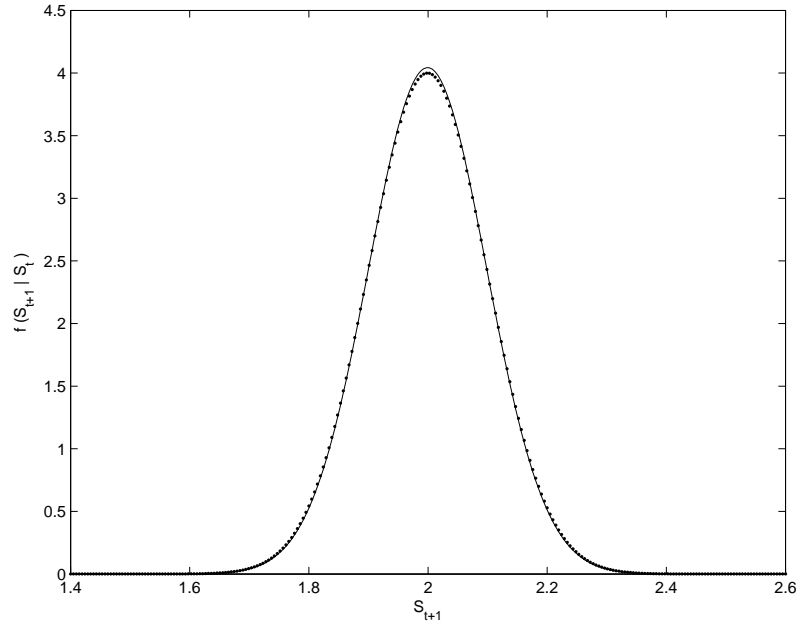


Figure 5.1: Density comparison for model 1f-A

Note: The solid curve is the graph of PDF in continuous-time framework, which is given by (5.12). The dotted curve is the graph of PDF in discrete-time framework, which is given by (5.16).  $S_t = 2$ ,  $T - t = 1$ ,  $\alpha = 0.02$ ,  $L = 2$  and  $\sigma = 0.1$ .

### Model 1f-B

- In continuous-time framework:

According to the Definition A.1, the price process  $S_t$  described by (5.6) is also an affine jump-diffusion process. Thus we can deduce the conditional PDF of  $S_{t+1}$  given  $S_t$  using (5.4).

For (5.6), in (A.2),  $\mathbf{X}_t = S_t$ ,

$$\mathbf{k}_0 = \alpha L, \quad \mathbf{K}_1 = -\alpha, \quad \mathbf{H}_0 = 0, \quad \mathbf{H}_1 = \sigma^2, \quad l_0 = 0, \quad \mathbf{1}_1 = 0, \quad \nu = 0.$$

By (A.4) - (A.6) and (5.3), we have the characteristic function given by (5.7), where  $A(t)$  and  $B(t)$  satisfy the following equations

$$\begin{cases} \partial_t A = -\alpha L B \\ \partial_t B = \alpha B - \frac{1}{2} \sigma^2 B^2, \end{cases} \quad (5.18)$$

with the same boundary conditions as (5.9).

Solving this Riccati equation system, we have

$$\begin{cases} A(t, T, s) = -\frac{2\alpha L}{\sigma^2} \ln\left[1 + is \frac{\sigma^2}{2\alpha} (e^{\alpha(t-T)} - 1)\right] \\ B(t, T, s) = \frac{ise^{\alpha(t-T)}}{1 + is \frac{\sigma^2}{2\alpha} (e^{\alpha(t-T)} - 1)}. \end{cases} \quad (5.19)$$

The conditional PDF of  $S_{t+1}$  given  $S_t$  is

$$f(\theta, S_T | S_t) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-isS_T} \exp\left(A(t, T, s) + B(t, T, s)S_t\right) ds. \quad (5.20)$$

The explicit form of the above integral is not easy to find in this case, so we can only compute it numerically.

- In discrete-time framework

Similarly, if the point  $T$  is close to  $t$ , then (5.18) can be approximated by the following expression,

$$\begin{cases} \frac{A(t, T, s) - A(T, T, s)}{t - T} \approx -\alpha L B(T, T, s) \\ \frac{B(t, T, s) - B(T, T, s)}{t - T} \approx \alpha B(T, T, s) - \frac{1}{2}\sigma^2 B(T, T, s)^2. \end{cases} \quad (5.21)$$

Together with the boundary condition (5.9), the values of  $A(t, T, s)$  and  $B(t, T, s)$  can be approximated by

$$\begin{cases} A(t, T, s) \approx (-is\alpha L)(t - T) \\ B(t, T, s) \approx is + (i\alpha s + \frac{1}{2}\sigma^2 s^2)(t - T). \end{cases} \quad (5.22)$$

Thus the conditional PDF of  $S_T$  given  $S_t$  can be approximated by

$$\begin{aligned} f(\theta, S_T | S_t) &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{-isS_T} \exp\left(A(t, T, s) + B(t, T, s)S_t\right) ds \\ &\approx \frac{1}{2\pi} \int_{\mathbb{R}} \exp\left(is\left((1 + \alpha(t - T))S_t - \alpha L(t - T) - S_T\right) + \frac{1}{2}\sigma^2 S_t(t - T)s^2\right) ds. \end{aligned} \quad (5.23)$$

After some computation (see Appendix C), we have

$$f(\theta, S_T | S_t) \approx \frac{1}{\sqrt{2\pi}\sigma\sqrt{S_t(T - t)}} \exp\left(-\frac{\left((1 + \alpha(t - T))S_t - \alpha L(t - T) - S_T\right)^2}{2\sigma^2 S_t(T - t)}\right). \quad (5.24)$$

Also, there is an easier way to deduce the above approximate PDF. For the SDE (5.6), we know that the corresponding discrete form of this SDE in Euler scheme is given by

$$S_T - S_t = \alpha(L - S_t)(T - t) + \sigma\sqrt{S_t}(W_T - W_t). \quad (5.25)$$

Given  $S_t$ , from  $W_T - W_t \sim \mathcal{N}(0, T - t)$ , we have  $S_T \sim \mathcal{N}(\alpha(L - S_t)(T - t) + S_t, \sigma^2 S_t(T - t))$ . Then the PDF (5.24) is obtained naturally.

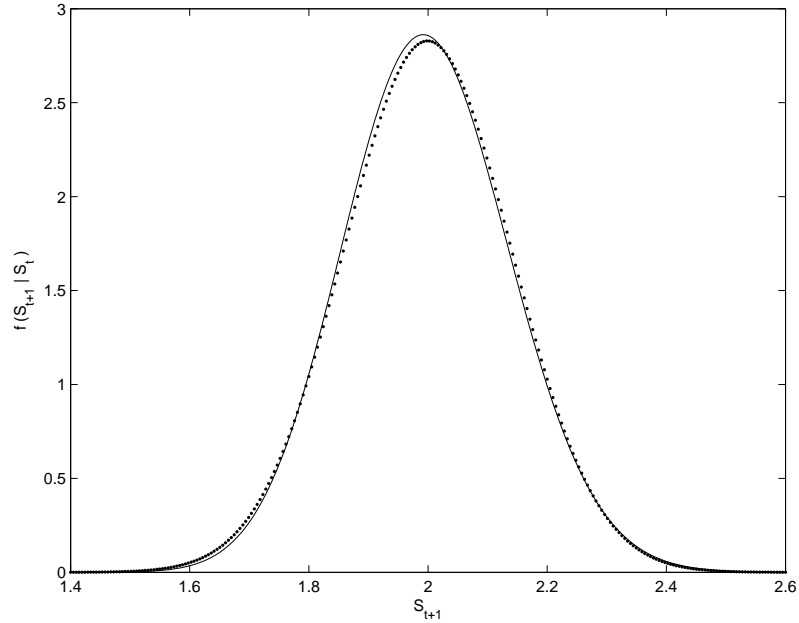


Figure 5.2: Density comparison for model 1f-B

Note: The solid curve is the graph of PDF in continuous-time framework, which is given by (5.20). The dotted curve is the graph of PDF in discrete-time framework, which is given by (5.24).  $S_t = 2$ ,  $T - t = 1$ ,  $\alpha = 0.02$ ,  $L = 2$  and  $\sigma = 0.1$ .

### Comparison of two probability density functions

Actually, the function (5.24) approximates the function (5.20) by approximating  $A$  and  $B$  in (5.19) using  $A$  and  $B$  in (5.22). For  $A$ , this is done by replacing  $e^{\alpha(t-T)}$  with  $1 + \alpha(t - T)$ , and then  $\ln(1 + is\frac{\sigma^2(t-T)}{2})$  with  $is\frac{\sigma^2(t-T)}{2}$ ; for  $B$ , this is done by replacing  $e^{\alpha(t-T)}$  with  $1 + \alpha(t - T)$ , and then  $\frac{1}{1 + is\sigma^2(t-T)/2}$  with  $1 - is\frac{\sigma^2(t-T)}{2}$ , and finally omitting the higher degree term  $\alpha\sigma^2(t - T)^2$ . If suppose  $\alpha$ ,  $\sigma$  and  $(T - t)$  are sufficiently small, then the above replacements will not make big errors. Figure 5.2 gives a visual sense about how big the difference between the two PDFs can be in a typical setting.

### 5.2.2 How to Treat the Non-trading Days

In real markets, gas is not traded every day. Usually, there are only 251, not 365, trading days in a year. On the remaining days, trade is suspended. The question of how to deal with these non-trading days becomes a problem in the process of our calibration.

For example, Friday, July 21, 1995 is a trading day with the price  $S_t$ . Then gas is not traded until 3 days later, i.e., Monday, July 24, 1995. The new price on Monday is thus the price next to  $S_t$ . Now, the question arises: when we apply the maximum likelihood estimation method, should we use  $f(\theta, S_{t+3}|S_t)$  or something else?

For the above question, we are in a dilemma. On the one hand, the world is still running as usual on those non-trading days. Some factors which are supposed to affect the behaviour of gas prices, such as weather, political events, etc., are still behaving in the same way. It seems, therefore, that we should not treat non-trading days differently from trading days. This encourages us to just simply use  $f(\theta, S_{t+3}|S_t)$ . On the other hand, since trading is suspended on non-trading days, some important factors, like the emotion of traders, will not affect the price in the same way as they do on trading days. From this point of view, it also makes sense if we treat non-trading days distinctively.

In order to decide how to deal with it, we do some data analysis to find whether on non-trading days the price behaves differently from it does on trading days or not. We compute the variance of the changes between successive prices, and try to see if the variances are different for those successive prices be separated by different numbers of days. Out of 2,007 prices, there are 2,006 pairs of successive prices, 1583

pairs separated by 1 day, 6 pairs separated by 2 days, 354 pairs separated by 3 days, 53 pairs separated by 4 day and 10 pairs separated by 5 days. Here, we compare the volatility on successive trading days (using a range of definitions corresponding to different possible models) when the trading days are separated by 24 hours and 72 hours in real time.

Changes	1-day Std.	3-day Std.	Ratio
$S_{i+1} - S_i$	0.0822	0.1011	1.2304
$\frac{S_{i+1} - S_i}{S_i}$	0.0341	0.0420	1.2333
$\frac{S_{i+1} - S_i}{\sqrt{S_i}}$	0.0520	0.0637	1.2243
$\ln \frac{S_{i+1}}{S_i}$	0.0343	0.0419	1.2202

Table 5.1: Volatility on 1-day apart successive trading days vs. volatility on 3-day apart successive trading days

Note: In the first column,  $S_{i+1}$  is the successive spot price to  $S_i$ . The second and third columns are the standard deviations of the random variable in the first column, which are used to measure the volatilities. The fourth column is the ratio of the third column to the second column.

If time progressed regardless of trading activity, we would expect (according to our diffusion models) the ratios in Table 5.1 to be close to  $\sqrt{3}$ . In fact, they are closer to 1 (although still significantly different). Thus, we are encouraged for the sake of simplicity to make the assumption in the remainder of the thesis that, as far as our models are concerned, time stands still on non-trading days. For the above example, instead of using  $f(S_{t+3}|S_t)$ , we use  $f(S_{t+1}|S_t)$  in our computation, even though the price  $S_{t+1}$  occurred 3 days after  $S_t$  did.

### 5.2.3 ML Estimation

Suppose we have the observations of spot prices  $\{S_t\}_{t=1}^n$ . Following the general procedure of the ML method, we first let

$$\mathcal{L} = \log \prod_{t=1}^{n-1} f(\theta, S_{t+1}|S_t). \quad (5.26)$$

As demonstrated in the previous section, the conditional PDF in continuous-time framework can be fairly well approximated by the corresponding conditional PDF in discrete-time framework, if  $\alpha$ ,  $\sigma$  and  $(T - t)$  are not very big. This is exactly the case in the gas market, as will be shown below. Hence, for the simplification of computation, we just use the approximate PDFs given by (5.16) and (5.24). Then, by setting  $\partial\mathcal{L}/\partial\theta = 0$ , we can get estimates of each parameter:

$$\begin{cases} \frac{\partial\mathcal{L}}{\partial\alpha} = 0 \Rightarrow \sum_{t=1}^{n-1} \frac{(L - S_t)(\alpha(L - S_t) + S_t - S_{t+1})}{\sigma^2 H_t^2} = 0 \\ \frac{\partial\mathcal{L}}{\partial L} = 0 \Rightarrow \sum_{t=1}^{n-1} \frac{\alpha(\alpha(L - S_t) + S_t - S_{t+1})}{\sigma^2 H_t^2} = 0 \\ \frac{\partial\mathcal{L}}{\partial\sigma} = 0 \Rightarrow \sum_{t=1}^{n-1} \left( -\frac{1}{\sigma} + \frac{(\alpha(L - S_t) + S_t - S_{t+1})^2}{\sigma^3 H_t^2} \right) = 0, \end{cases} \quad (5.27)$$

where  $H_t = 1$  for model 1f-A, and  $H_t = \sqrt{S_t}$  for model 1f-B.

Solving the above equations, we obtain

$$\begin{aligned} \alpha &= \frac{DE - CF}{C^2 - BE}, \\ L &= -\frac{B}{C} - \frac{D}{\alpha C}, \\ \sigma &= \sqrt{\frac{1}{n-1} \sum_{t=1}^{n-1} \frac{(\alpha(L - S_t) + S_t - S_{t+1})^2}{H_t^2}}, \end{aligned} \quad (5.28)$$

where

$$B = \sum_{t=1}^{n-1} \frac{S_t^2}{H_t^2}, \quad C = -\sum_{t=1}^{n-1} \frac{S_t}{H_t^2}, \quad D = \sum_{t=1}^{n-1} \frac{S_t(S_{t+1} - S_t)}{H_t^2},$$

$$E = \sum_{t=1}^{n-1} \frac{1}{H_t^2}, \quad F = \sum_{t=1}^{n-1} \frac{S_t - S_{t+1}}{H_t^2}.$$

#### 5.2.4 Continuous Data and Discrete Data

If we suppose gas prices follow the models introduced in the previous chapter, which are described by SDEs, then we are actually supposing prices change continuously. However, we can only get gas prices at discrete times in real market. In this thesis, we just use daily gas prices. Thus, a question is raised: whether we can calibrate those models just using a subset of the realization of those SDEs.

In this section, we do an empirical test.

First, setting the value of  $S_0$ , we simulate 200,000 values one by one using the discrete forms of the SDEs. For model 1f-A, the recursion formula is

$$S_{t+\Delta t} = S_t + \alpha(L - S_t)\Delta t + \sigma\sqrt{\Delta t}\varepsilon_t. \quad (5.29)$$

For model 1f-B, the recursion formula is

$$S_{t+\Delta t} = S_t + \alpha(L - S_t)\Delta t + \sigma\sqrt{S_t}\sqrt{\Delta t}\varepsilon_t. \quad (5.30)$$

In the above formulas,  $\varepsilon_t \sim \mathcal{N}(0, 1)$ , and we just set  $\Delta t = 0.01$ .

Second, we select the values with integer time subscript, say 2,001 values, which is just 1 percent of the original simulated values, as a subset of the whole.

Finally, we try to estimate the values of the parameters using the above selected subset and see whether the results recover the originally given parameters.

All together, we have the above recovery test for 50 different paths. With the obtained 50 estimation results, we examine the mean and standard deviation<sup>3</sup> of estimation results. The results are shown in Table 5.2 and Table 5.3 as follows:

	Given values	Mean of Est.	Std. of Est.
$\alpha$	0.0200	0.0205	0.0056
$L$	2.0000	1.9969	0.0938
$\sigma$	0.1000	0.0991	0.0018

Table 5.2: Estimation using the subset of prices for model 1f-A

	Given values	Mean of Est.	Std. of Est.
$\alpha$	0.0200	0.0220	0.0056
$L$	2.0000	1.9766	0.1425
$\sigma$	0.1000	0.0990	0.0016

Table 5.3: Estimation using the subset of prices for model 1f-B

From the empirical test results shown above, we can see that the estimation results obtained from the subset of the whole data recover the true values reasonably well. Encouraged by this conclusion, we can perform the calibration of our models just using the daily prices with some confidence.

### 5.2.5 How the Length of Data Affects the Results

In statistics, sample size is a very important factor influencing the results of estimation. In this section, we empirically study the effects of data length on the calibration.

The method we use is quite simple: (1) Given a set of parameters, simulate sample paths with different lengths. For each length, generate 100 paths; (2) For each sample path, obtain the corresponding estimated parameters. (3) Put the estimation results

---

<sup>3</sup>Giving a set of observations of random variable  $X$ ,  $\{X_t\}_{t=1}^n$ , we estimate its expectation by the mean  $E(X) = \frac{1}{n} \sum_{t=1}^n X_t$ , and the standard deviation by  $\text{std}(X) = \sqrt{\frac{1}{n-1} \sum_{t=1}^n (X_t - E(X))^2}$ , which is the best unbiased estimate.

corresponding to the same data length together and compute the mean bias and standard deviation. (4) Compare the estimation accuracy for different data lengths.

The results are shown in Table 5.4 and Table 5.5. We can see that when the data length increases, both the bias of the estimation and the standard deviation tend to decrease. In fact, the results are by and large consistent with what we would expect to see theoretically, i.e., that the bias decreases like  $\frac{1}{N}$  and the Std. like  $\frac{1}{\sqrt{N}}$ .

Data length	$\alpha$ (given value: 0.02)			L (given value: 2.00)			$\sigma$ (given value: 0.10)		
	Mean	Bias	Std.	Mean	Bias	Std.	Mean	Bias	Std.
100	.0719	.0519	.0530	1.990	0.010	.4180	.0982	.0018	.0079
500	.0282	.0082	.0119	2.060	0.060	.2020	.1000	.0000	.0029
1000	.0242	.0042	.0074	2.003	0.003	.1537	.1000	.0000	.0022
1500	.0226	.0026	.0056	2.005	0.005	.1318	.1000	.0000	.0019
2000	.0216	.0016	.0047	1.997	0.003	.1069	.0996	.0004	.0017
3000	.0210	.0010	.0035	2.015	0.015	.0991	.0999	.0001	.0013
5000	.0209	.0009	.0023	1.994	0.006	.0718	.0999	.0001	.0010
10000	.0204	.0004	.0022	1.994	0.006	.0485	.0998	.0002	.0008
20000	.0202	.0002	.0013	1.994	0.006	.0386	.1000	.0000	.0004

Table 5.4: The effects of data length for model 1f-A

Data length	$\alpha$ (given value: 0.02)			L (given value: 2.00)			$\sigma$ (given value: 0.10)		
	Mean	Bias	Std.	Mean	Bias	Std.	Mean	Bias	Std.
100	.0668	.0468	.0481	2.281	0.281	1.357	.0981	.0019	.0066
500	.0291	.0091	.0120	1.966	0.034	.3152	.0995	.0005	.0032
1000	.0242	.0042	.0074	1.978	0.022	.1973	.0999	.0001	.0023
1500	.0229	.0029	.0060	1.972	0.028	.1614	.0997	.0003	.0020
2000	.0223	.0023	.0051	1.998	0.002	.1738	.0998	.0002	.0016
3000	.0212	.0012	.0037	2.008	0.008	.1317	.1001	.0001	.0015
5000	.0203	.0003	.0029	2.018	0.018	.0946	.1001	.0001	.0010
10000	.0205	.0005	.0021	1.993	0.007	.0685	.1000	.0000	.0007
20000	.0201	.0001	.0013	2.000	0.000	.0504	.1000	.0000	.0005

Table 5.5: The effects of data length for model 1f-B

### 5.2.6 Estimation Results for real Data

For the real data, i.e., the Henry Hub prices from January 2, 1992 to December 30, 1999, the estimation results are shown in Table 5.6.

	model 1f-A	model 1f-B
$\alpha$	0.0181	0.0139
$L$	2.1441	2.1526
$\sigma$	0.0870	0.0549

Table 5.6: Estimation results for 1-factor models

## 5.3 Calibration for 1-factor Models with Seasonality

In this class of models, gas spot prices follow the mean reverting process described by equations (3.18) - (3.20). By setting  $r = 0$ ,  $r = \frac{1}{2}$  and  $r = 1$ , respectively, we have three specific models: model 1f-sn-A, model 1f-sn-B and model 1f-sn-C.

These models extend the previous, relatively simpler, 1-factor models by incorporating seasonality. We first strip off the seasonal term  $f(t)$  from the observed futures prices, then with the known values of the unseasonalized term,  $\{X_t\}_{t=1}^n$ , we get the other parameters by performing ML method.

### 5.3.1 Reveal the Seasonal Term $f(t)$ from Futures Prices

Calibration is essentially a process of matching the information observed from the market. The better the information is matched, the more successful the calibration is. As we have shown in Chapter 2, seasonality can be easily seen from the natural gas spot prices and futures prices. Hence, to reveal the seasonal term  $f(t)$ , a natural idea is to exploit the information in both spot prices and futures prices.

We denote the historical spot prices by  $\{S_t\}_{t=1}^n$ , and the futures prices by  $\{F_{t,T_{ii}}|t = 1, 2, \dots, n; i = 1, 2, \dots, m\}$ , where  $T_{ii}$  is the  $i$ th delivery time after  $t$ ,  $m$  is the number of futures prices we can observe at time  $t$ . These are the data we can obtain from the real market.

On the other hand, under the assumption that the spot price follows the process described by equations (3.18) - (3.20), we have the theoretical futures price function  $F^{\tilde{\theta}}(t, T, S_t)$  given by (4.23), where  $t$  is the observing time,  $T$  is the delivery time,  $\tilde{\theta}$  is the set of active parameters, i.e.,  $[b, \beta_1, \beta_2, \dots, \beta_N, \eta_1, \eta_2, \dots, \eta_N, \tilde{\alpha}, \tilde{L}]$ .

To match the real market data to our models, we need to find some appropriate parameters to make the theoretical futures prices and the actual futures prices as close as possible. If the distance between two vectors is defined in Euclidean space, then the estimation of  $\tilde{\theta}$  can be obtained by minimizing the sum of the squares of the differences between  $F^{\tilde{\theta}}(t, T_{ii}, S_t)$  and  $F_{t,T_{ii}}$ , for  $t = 1, 2, \dots, n; i = 1, 2, \dots, m$ . In

mathematical terms,

$$\tilde{\theta} = \underset{\tilde{\theta}}{\operatorname{argmin}} \sum_{t=1}^n \sum_{i=1}^m (F^{\tilde{\theta}}(t, T_{ti}, S_t) - F_{t, T_{ti}})^2. \quad (5.31)$$

With the obtained  $\tilde{\theta}$ ,  $f(t)$ , and hence  $\{X_t\}_{t=1}^n$  are both easily computed via (3.19) and (3.18).

Table 5.7 shows the estimation of the risk-neutral parameters for artificial data<sup>4</sup> and real data.

	For artificial data			For real data
	Given	Mean of Est.	Std. of Est.	Est.
$\tilde{\alpha}$	0.0070	0.0070	1.4569e-13	0.0073
$\tilde{L}$	1.7000	1.7000	1.8859e-11	1.6663

Table 5.7: Estimation of the risk-neutral parameters for 1-factor models with seasonality

From Table 5.7, we see that the estimation of the risk-neutral parameters for artificial data recovers the given values accurately. This implies that our method works and the programs are doing their job correctly, and therefore we believe that the estimation for real data is reliable. Theoretically, the values in the ‘‘Std. of Est.’’ column could be arbitrary small, if we set the tolerance of the Matlab optimization function small enough. For the sake of clarity, hereafter we just put 0 there if the absolute value of a number is smaller than 1e-10.

For the estimated values of the parameters in  $f(t)$ , see Table 5.8 in Section 5.3.4.

Figure 5.3 and Figure 5.4 show the recovery of  $f(t)$  for artificial data and real data.

---

<sup>4</sup>For each type of model, no matter 1-factor or 2-factor, with seasonality or without seasonality, the artificial data we use include 50 paths of simulated  $\{S_t\}_{t=1}^{2000}$ , and the corresponding futures prices; for 2-factor models,  $\{L_t\}_{t=1}^{2000}$ , as well.

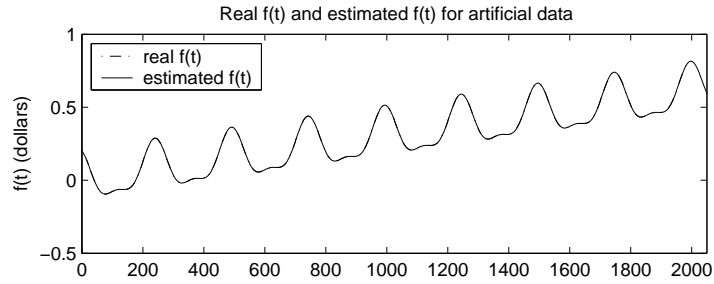


Figure 5.3: Recover  $f(t)$  from futures prices for artificial data

Note: From the figure, we see that the estimated  $f(t)$  exactly covers the true function, which implies that our method works and the programs are doing their job correctly. The corresponding parameters for the true  $f(t)$  are those in the first column of Table 5.8; the corresponding parameters for estimated  $f(t)$  are those in the second column of Table 5.8.

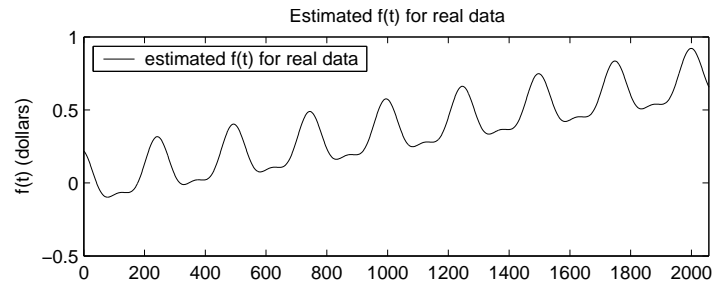


Figure 5.4: Reveal  $f(t)$  from futures prices for real data

Note: The corresponding parameters for  $f(t)$  are those in the last column of Table 5.8.

### 5.3.2 Probability Density Function

Here we give the conditional PDF of  $X_{t+1}$  given  $X_t$  in the discrete-time framework, i.e., derived from the approximate SDEs. This is the form of the conditional PDF that we will use in the calibrations below.

$$f(\theta, X_{t+1}|X_t) \approx \frac{1}{\sqrt{2\pi}\sigma(t)H_t} \exp\left(-\frac{(\alpha(L - X_t) + X_t - X_{t+1})^2}{2\sigma(t)^2 H_t^2}\right), \quad (5.32)$$

where

$$H_t = \begin{cases} 1 & \text{for model 1f-sn-A,} \\ \sqrt{X_t} & \text{for model 1f-sn-B,} \\ X_t & \text{for model 1f-sn-C.} \end{cases} \quad (5.33)$$

### 5.3.3 Estimation

Suppose to begin with that we have obtained the values of  $\{X_t\}_{t=1}^n$ . Following the general procedure of the ML method, we first let  $\mathcal{L} = \log \prod_{t=1}^{n-1} f(\theta, X_{t+1}|X_t)$ . The estimation of the parameters for  $X_t$  can be obtained as follows:

$$\left\{ \begin{array}{l} \frac{\partial \mathcal{L}}{\partial \alpha} = 0 \Rightarrow \sum_{t=1}^{n-1} \frac{(L - X_t)(\alpha(L - X_t) + X_t - X_{t+1})}{\sigma(t)^2 H_t^2} = 0 \\ \frac{\partial \mathcal{L}}{\partial L} = 0 \Rightarrow \sum_{t=1}^{n-1} \frac{\alpha(\alpha(L - X_t) + X_t - X_{t+1})}{\sigma(t)^2 H_t^2} = 0 \\ \frac{\partial \mathcal{L}}{\partial c} = 0 \Rightarrow \sum_{t=1}^{n-1} \left( -1 + \frac{(\alpha(L - X_t) + X_t - X_{t+1})^2}{\sigma(t)^2 H_t^2} \right) = 0 \\ \frac{\partial \mathcal{L}}{\partial \lambda_j} = 0 \Rightarrow \sum_{t=1}^{n-1} \left( -1 + \frac{(\alpha(L - X_t) + X_t - X_{t+1})^2}{\sigma(t)^2 H_t^2} \right) \cos\left(\frac{2\pi jt}{P}\right) = 0 \\ \frac{\partial \mathcal{L}}{\partial \omega_j} = 0 \Rightarrow \sum_{t=1}^{n-1} \left( -1 + \frac{(\alpha(L - X_t) + X_t - X_{t+1})^2}{\sigma(t)^2 H_t^2} \right) \sin\left(\frac{2\pi jt}{P}\right) = 0. \end{array} \right. \quad (5.34)$$

Because of the existence of a time dependent function  $\sigma(t)$ , it is hard to give an explicit solution to the above equation system. The best we can do is to get a numerical solution<sup>5</sup>.

---

<sup>5</sup>For the equation  $f(x) = 0$ , where  $x \in \mathbb{R}^n$  and  $f(x) : x \rightarrow \mathbb{R}^n$ , if the solution exists and unique, then this solution can be obtained by solving an equivalent optimization problem:  $x = \underset{x \in \mathbb{R}^n}{\operatorname{argmin}} \|f(x)\|$ , where  $\|\cdot\|$  is defined as the Euclidean length of a vector. In practice, we solve this optimization problem using the Matlab function “fminsearch”, which finds the minimum of a scalar function of several variables.

### 5.3.4 Results

Since the three 1-factor models with seasonality share the same futures price function given by (4.23), the parameters for  $f(t)$  have the same estimation values for these three models. The results are shown in Table 5.8.

	For artificial data			For real data
	Given	Mean of Est.	Std. of Est.	Est.
$b$	0.0003	0.0003	0	0.0003
$\beta_1$	0.1500	0.1500	0	0.1639
$\beta_2$	0.0500	0.0500	0	0.0558
$\eta_1$	-0.0500	-0.0500	0	-0.0468
$\eta_2$	-0.0300	-0.0300	0	-0.0295

Table 5.8: Estimation result for  $f(t)$  in 1-factor models with seasonality

Note: Again, the recovering of the given parameters implies that our method works and the programs are doing their job correctly, and therefore we believe that the estimation for real data is reliable. PS: In the third column, 0 denotes some number with the absolute value less than 1e-10. The reason for the extreme accuracy of the parameters for  $f(t)$  is that they are estimated directly from matching the futures prices.

The other parameters have different estimation values for each model. The results are shown below. In Tables 5.9-5.11, we give the results for model 1f-sn-A, model 1f-sn-B and model 1f-sn-C, respectively. In each case we report the estimated values of the parameters, and also the results of robustness analysis carried out using simulated data. 50 simulated paths involving 2,000 data points were used. More information about the estimations is provided in Figures 5.5 - 5.7, where the goodness of fit, the shape of the objective function near the optimal values of  $\tilde{\theta}$ , and the behaviour of the seasonally varying terms  $f(t)$  and  $\sigma(t)$  are shown, together with a sample simulation to give a visual impression of the effectiveness of the model.

**Model 1f-sn-A**

	For artificial data			For real data
	Given	Mean of Est.	Std. of Est.	Est.
$\alpha$	0.0200	0.0218	0.0046	0.0205
$L$	1.9000	1.9090	0.0788	1.8599
$c$	-2.6000	-2.5981	0.0151	-2.5617
$\lambda_1$	0.4000	0.4044	0.0209	0.4120
$\lambda_2$	0.0800	0.0823	0.0241	0.0834
$\omega_1$	-0.2000	-0.1969	0.0223	-0.1970
$\omega_2$	-0.0600	-0.0620	0.0210	-0.0644

Table 5.9: Estimation result for model 1f-sn-A

Note: From the table, we see that the estimation of the parameters for artificial data recovers the given values well. This implies that our method works and the programs are doing their job correctly, and therefore, we believe that the estimation for real data is reliable.

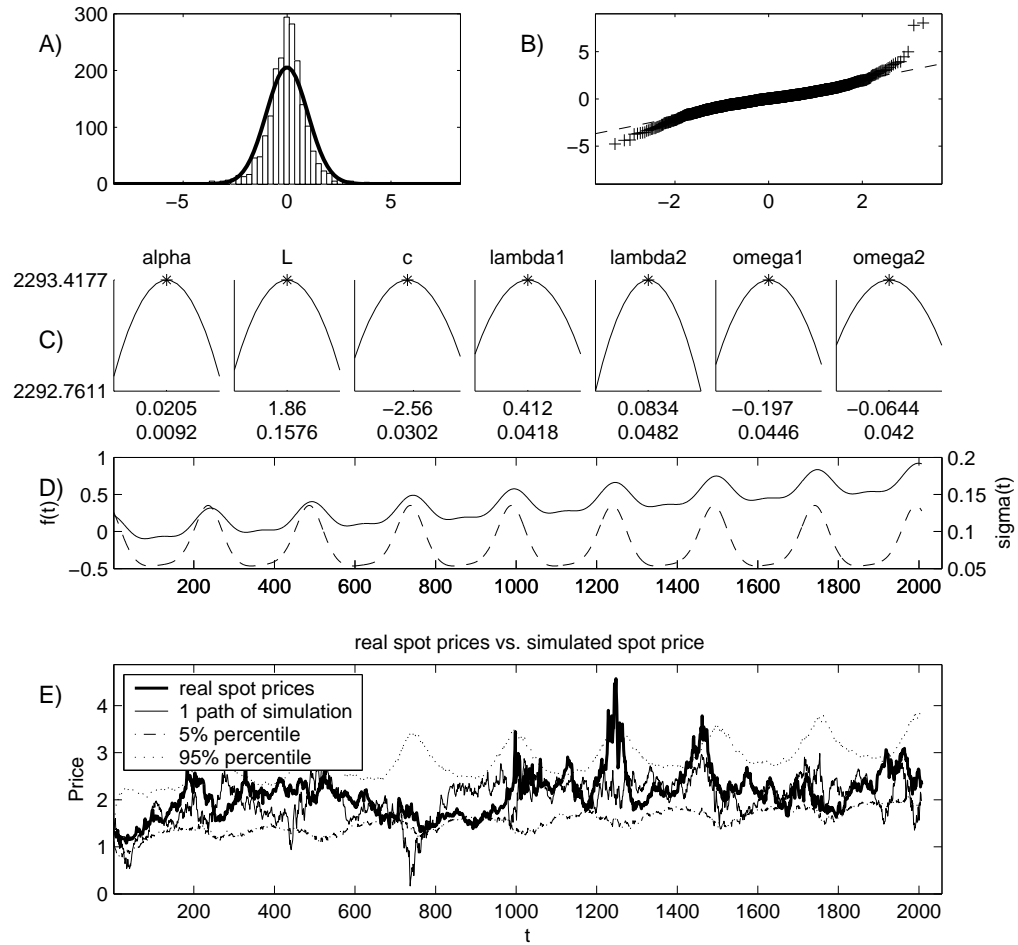


Figure 5.5: Result for model 1f-sn-A

Note: A) is the histogram of  $Y_t = \frac{\alpha(L - X_t) + X_t - X_{t+1}}{\sigma(t)}$ , with the graph of the standard normal distribution PDF  $f(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$  superimposed; B) is the quantile-quantile plot of  $Y_t$  vs. the standard normal distribution; C) are the intersecting slices of log likelihood function with respect to the parameters estimated by ML method. Under the plots, the numbers in the first row are the estimates of the parameters, the numbers in the second row are the widths of the plots, which are just 2 times the third column of Table 5.9. Our estimates are all on the top of slices, which implies that we successfully maximize the likelihood; D) is the graphs of  $f(t)$  (solid curve) and  $\sigma(t)$  (dashed curve); E) is one simulated spot price path vs. the actual spot price path.

**Model 1f-sn-B**

	For artificial data			For real data
	Given	Mean of Est.	Std. of Est.	Est.
$\alpha$	0.0180	0.0191	0.0043	0.0178
$L$	1.9000	1.9159	0.0858	1.8667
$c$	-2.9000	-2.9006	0.0144	-2.8970
$\lambda_1$	0.4000	0.4020	0.0242	0.3846
$\lambda_2$	0.0800	0.0824	0.0238	0.0787
$\omega_1$	-0.2000	-0.2031	0.0242	-0.1977
$\omega_2$	-0.0200	-0.0206	0.0239	-0.0233

Table 5.10: Estimation result for model 1f-sn-B

Note: Refer to the note in Table 5.9.

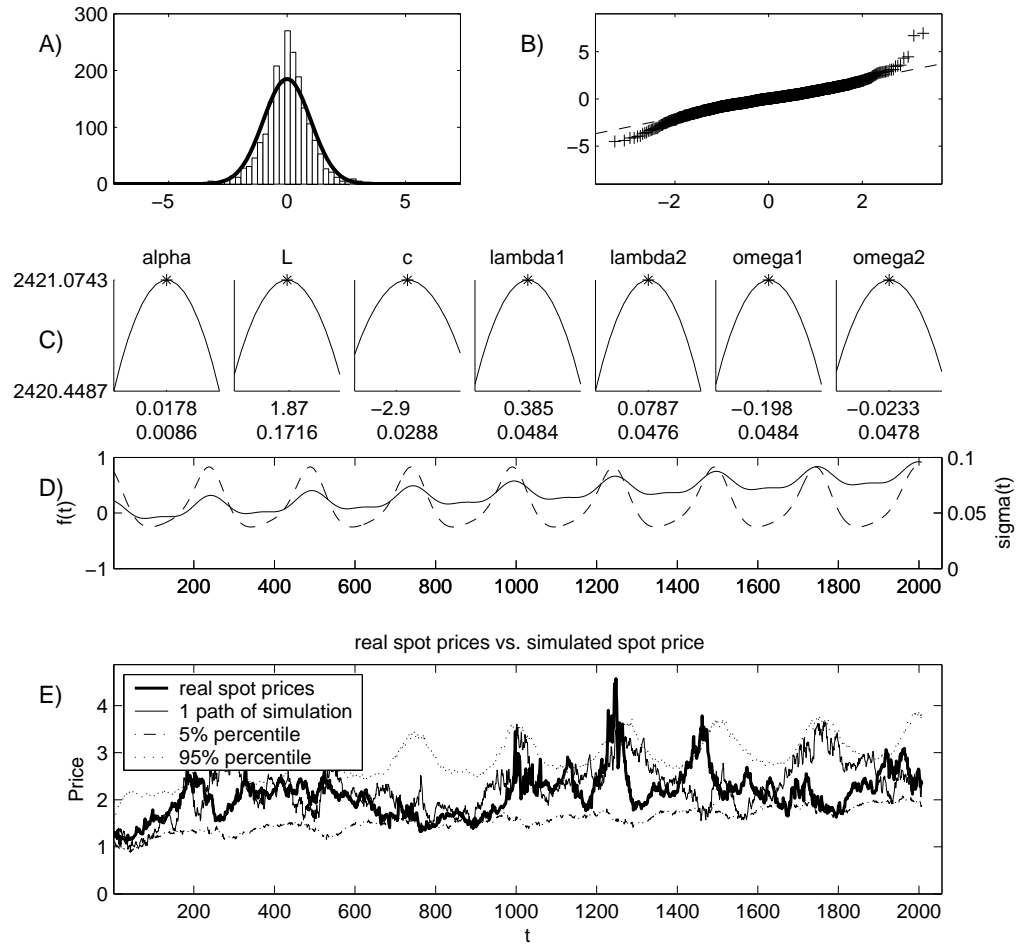


Figure 5.6: Result for model 1f-sn-B

Note: A) is the histogram of  $Y_t = \frac{\alpha(L - X_t) + X_t - X_{t+1}}{\sigma(t)\sqrt{X_t}}$ , with the graph of the standard normal distribution PDF  $f(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$  superimposed; For the description of other plots, please refer to the note of Figure 5.5.

**Model 1f-sn-C**

	For artificial data			For real data
	Given	Mean of Est.	Std. of Est.	Est.
$\alpha$	0.0160	0.0181	0.0043	0.0163
$L$	1.9000	1.9055	0.0900	1.8792
$c$	-3.2000	-3.2033	0.0127	-3.1999
$\lambda_1$	0.4000	0.4033	0.0227	0.3802
$\lambda_2$	0.0800	0.0777	0.0210	0.0773
$\omega_1$	-0.2000	-0.2040	0.0252	-0.2042
$\omega_2$	0.0150	0.0218	0.0201	0.0147

Table 5.11: Estimation result for model 1f-sn-C

Note: Refer to the note in Table 5.9.

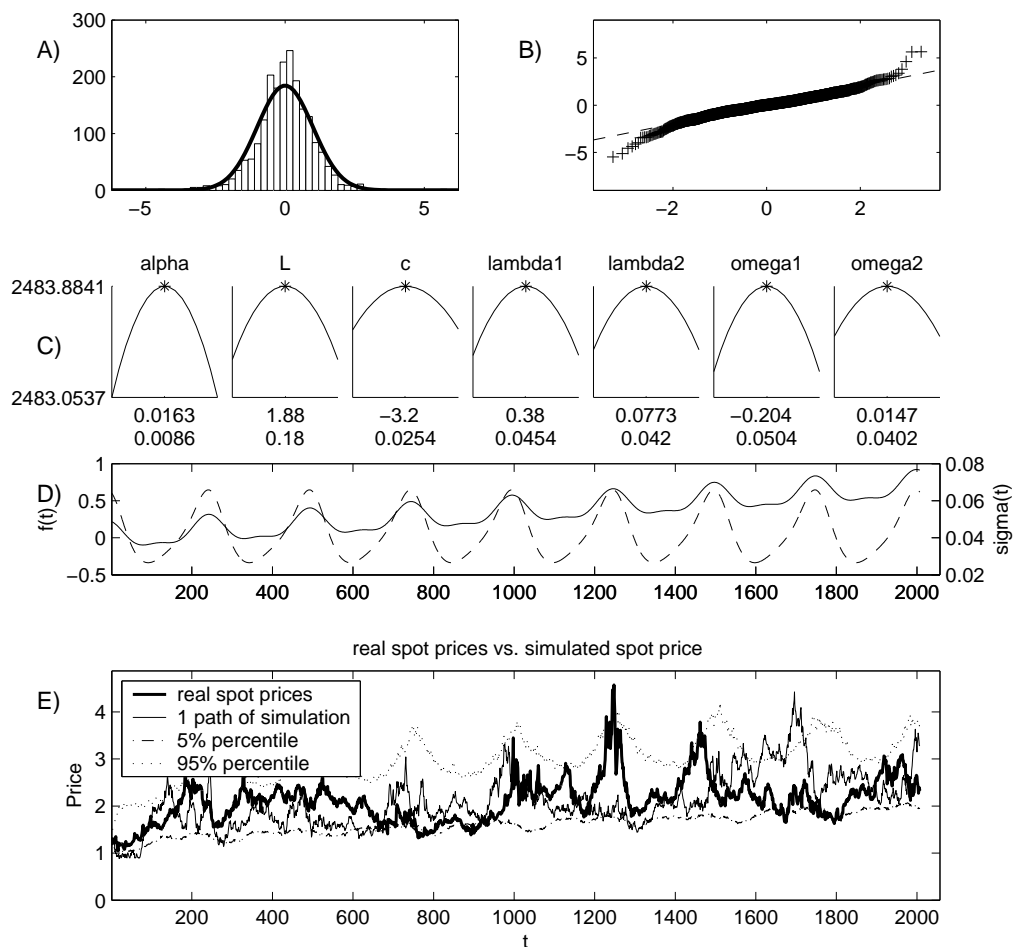


Figure 5.7: Result for model 1f-sn-C

Note: A) is the histogram of  $Y_t = \frac{\alpha(L - X_t) + X_t - X_{t+1}}{\sigma(t)X_t}$ , with the graph of the standard normal distribution PDF  $f(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$  superimposed; For the description of other plots, please refer to the note of Figure 5.5.

### 5.3.5 Futures Matching

Once the spot price model has been defined, one important test of the model is how well it fits the observed futures curves. For the models in this thesis, we treat the spot price as the key variable, and the futures price is just considered as an endogenous function of the model's parameters. But this does not really mean that futures price is subject to and computed from spot price. Actually, the futures market plays a more important role in the natural gas industry. If the spot price model fails to capture the evolution of futures prices, then we have to say it is a poor model.

Actually, when we calibrate our spot models, we have considered futures prices as the inputs. In this process of stripping off  $f(t)$  from spot prices, futures prices have been matched as well as possible. The corresponding minimum obtained in (5.31) is 519.137. The middle plot in Figure 5.8 shows how well they are matched for a typical futures curve.

On the other hand, we obtained the so-called risk-neutral parameters when we match futures prices. Since the futures price is supposed to be the expectation of spot price in the risk-neutral world (see Chapter 4), we would expect that our spot price model with the risk-neutral parameters plugged in could really achieve it. The plot in the bottom of the Figure 5.8 shows how well the futures prices (on the same date as before) are matched from this point of view.

As shown in Figure 5.8, it seems our 1-factor models with seasonality do match the shape of the futures curve qualitatively, but are not quantitatively very accurate.

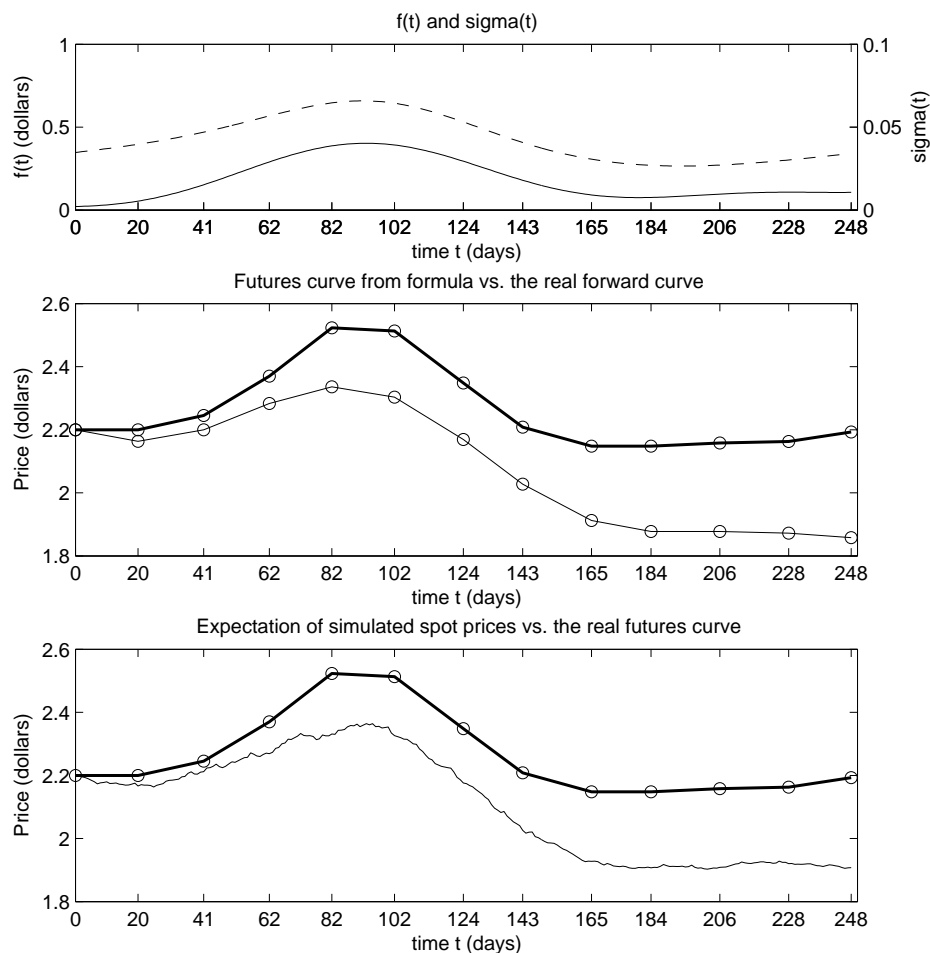


Figure 5.8: Futures matching for 1-factor models with seasonality

Note: In the middle plot, the bold curve is the real futures curve, which is the collection of 12 futures prices with different maturities, being observed on Aug 1, 1993. The thinner curve is graph of futures function given by (4.23). In the lower plot, the bold curve is the real futures curve. The thinner curve is the mean values of 250 paths of  $S_t$  simulated by equations (3.18)-(3.20), where the risk-neutral parameters shown in Table 5.7 are plugged in, and  $r = 1$ . For comparison, the upper plot shows the seasonal term  $f(t)$  (solid curve), and the seasonal volatility function  $\sigma(t)$  (dashed curve).

## 5.4 Calibration for 2-factor Models without Seasonality

In these models, gas spot prices follow one of the mean reverting processes described by (3.22). By setting  $r = 0$ ,  $r = \frac{1}{2}$  and  $r = 1$ , respectively, we have three specific models:

**model 2f-A:**

$$\begin{cases} dS_t = \alpha(L_t - S_t)dt + \sigma dW_t^1 \\ dL_t = \mu(\gamma - L_t)dt + \tau dW_t^2, \end{cases} \quad (5.35)$$

**model 2f-B:**

$$\begin{cases} dS_t = \alpha(L_t - S_t)dt + \sigma\sqrt{S_t}dW_t^1 \\ dL_t = \mu(\gamma - L_t)dt + \tau\sqrt{L_t}dW_t^2, \end{cases} \quad (5.36)$$

**model 2f-C:**

$$\begin{cases} dS_t = \alpha(L_t - S_t)dt + \sigma S_t dW_t^1 \\ dL_t = \mu(\gamma - L_t)dt + \tau L_t dW_t^2. \end{cases} \quad (5.37)$$

### 5.4.1 Reveal the Hidden Factor $L_t$ from Futures Prices

At any time, the information we can observe from real market includes spot prices and futures prices<sup>6</sup>. However, for the above 2-factor models, to implement maximum likelihood estimation, we need to know the values of  $\{L_t\}_{t=1}^n$ , the long term mean of spot prices, which are unobservable in the market. Therefore, we are faced with a task to reveal the hidden factors from the observable information. Since  $L_t$  appears in the futures price function, it can be estimated by matching the actual futures prices with the theoretical futures prices. In this process, we find the values of  $\{L_t\}_{t=1}^n$  and the risk-neutralized parameters simultaneously. The algorithm is similar to the one

---

<sup>6</sup>For the prices of other financial instruments, like options, swap, etc., they are not under our consideration in this thesis.

used to find  $f(t)$  in 1-factor models with seasonality, but more complicated.

Now, let's explain the method mathematically.

As before, at time  $t$ , we denote the observed spot price by  $S_t$ , and the futures prices by  $F_{t,T_{ii}}$  ( $i = 1, 2, 3, \dots, m$ ), where  $T_{ii}$  is the  $i$ th delivery time after  $t$ , and  $m$  is the number of futures prices we can observe at time  $t$ .

On the other hand, we also have the theoretical futures prices  $F^{\tilde{\theta}}(t, T_{ii}, (S_t, L_t)) = A_i + B_i L_t$  given by (4.24), where  $t$  is the observing time,  $T_{ii}$  is the delivery time,  $\tilde{\theta}$  is the set of active parameters, say,  $[\tilde{\alpha}, \tilde{\sigma}, \tilde{\mu}, \tilde{\gamma}, \tilde{\tau}]$ ,  $A_i = e^{\tilde{\alpha}(t-T_{ii})} S_t + \frac{\tilde{\mu}\tilde{\gamma}}{\tilde{\alpha}-\tilde{\mu}} (e^{\tilde{\alpha}(t-T_{ii})} - 1) - \frac{\tilde{\alpha}\tilde{\gamma}}{\tilde{\alpha}-\tilde{\mu}} (e^{\tilde{\mu}(t-T_{ii})} - 1)$ ,  $B_i = \frac{\tilde{\alpha}}{\tilde{\alpha}-\tilde{\mu}} (e^{\tilde{\mu}(t-T_{ii})} - e^{\tilde{\alpha}(t-T_{ii})})$ .

First, suppose  $\tilde{\theta}$  is given. At a fixed time  $t$ , we compute  $L_t$ 's value by minimizing the sum of the squares of the differences between  $\{F_{t,T_{ii}}\}_{i=1}^m$  and  $\{F^{\tilde{\theta}}(t, T_{ii}, (S_t, L_t))\}_{i=1}^m$ . This  $L_t$ 's value is thought of as the most possible value hiding behind these observable information. Since a fixed  $\tilde{\theta}$  corresponds to a unique  $L_t$ 's value in this process,  $L_t$  can be regarded as a function of  $\tilde{\theta}$ , i.e.,

$$L_t(\tilde{\theta}) = \operatorname{argmin}_{L_t} \sum_{i=1}^m (F^{\tilde{\theta}}(t, T_{ii}, (S_t, L_t)) - F_{t,T_{ii}})^2. \quad (5.38)$$

By setting the derivative with respect to  $L_t$  to 0, (5.38) is simplified to give

$$L_t(\tilde{\theta}) = \frac{\sum_{i=1}^m (B_i F_{t,T_{ii}} - A_i B_i)}{\sum_{i=1}^m B_i^2}. \quad (5.39)$$

After finding  $L_t$  as a function of  $\tilde{\theta}$ , we then free  $\tilde{\theta}$  to minimize the sum of the differences between  $\{F_{t,T_{ii}}\}_{i=1}^m$  and  $\{F^{\tilde{\theta}}(t, T_{ii}, (S_t, L_t))\}_{i=1}^m$  for all  $t$ . In this way, we get the optimal estimation of  $\tilde{\theta}$ , i.e.

$$\tilde{\theta} = \operatorname{argmin}_{\tilde{\theta}} \sum_{t=1}^n \sum_{i=1}^m (F^{\tilde{\theta}}(t, T_{ii}, (S_t, L_t(\tilde{\theta}))) - F_{t,T_{ii}})^2. \quad (5.40)$$

With the obtained  $\tilde{\theta}$ ,  $\{L_t\}_{t=1}^n$  can then be computed by (5.39).

Table 5.12 shows the estimation of the risk-neutral parameters for artificial data and real data.

	For artificial data			For real data
	Given	Mean of Est.	Std. of Est.	Est.
$\tilde{\alpha}$	0.0120	0.0120	0	0.0118
$\tilde{\mu}$	0.0050	0.0050	0	0.0054
$\tilde{\gamma}$	2.0000	2.0000	0	2.0561

Table 5.12: Estimation of the risk-neutral parameters for 2-factor unseasonal models

Figure 5.9 and Figure 5.10 show the recovery of  $L_t$ 's values for artificial data and real data. We see from Figure 5.10 that the values for  $L_t$  recovered from the real

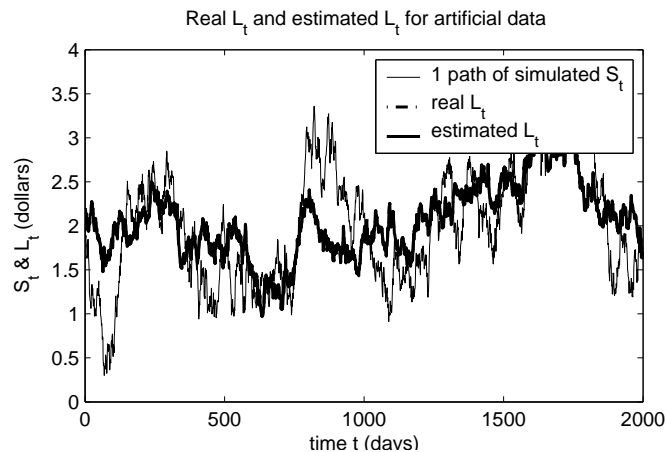


Figure 5.9: Recover  $L_t$  from futures prices for artificial data

Note: From the figure, we see that the estimated  $L_t$ 's values exactly cover the true values, which implies that our method works and the programs are doing their job correctly.

data are perhaps more volatile than might be expected if we are to interpret them as the evolution of the long term mean, but otherwise this interpretation seems quite

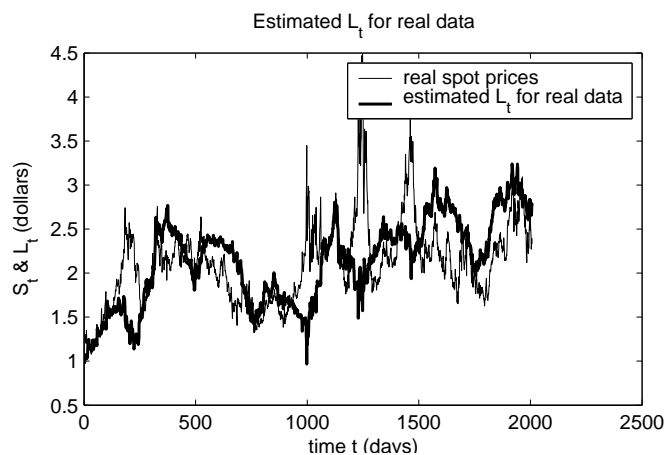


Figure 5.10: Reveal  $L_t$  from futures prices for real data

reasonable—the values obtained are not far from the mean value of  $S_t$  and appear to be ‘leading’ the evolution of  $S_t$  as would be expected from the mean reverting nature of the model. Later in this chapter we will introduce seasonality into these models, and this will allow us to recover much more realistic values of the long term mean. Before this we turn our attention to the implications of the above estimates for the estimation of the real world processes for  $S_t$  and  $L_t$ .

### 5.4.2 Probability Density Function

For these 2-factor models without seasonality, two are affine jump-diffusion processes, and so the conditional PDFs can be obtained by the transform analysis for affine jump-diffusion process (see Section 5.1.1). However, the related computation is somewhat complicated. In particular, when we need to evaluate the obtained PDFs we can only apply numerical methods to get approximate results. (For the deduction of PDFs in continuous-time framework, see Appendix D.) Hence in the calibration of these models we just use the approximate PDFs from the discrete-time framework, instead of those from the continuous-time framework.

We deduce the PDFs of  $[S_{t+1}, L_{t+1}]$  conditional on  $[S_t, L_t]$  in the discrete-time framework—the approximate PDFs—directly from the discrete form of the SDEs satisfied by  $S_t$  and  $L_t$ :

$$\begin{cases} S_{t+1} - S_t = \alpha(L_t - S_t) + \sigma H_{1t}(W_{t+1}^1 - W_t^1) \\ L_{t+1} - L_t = \mu(\gamma - L_t) + \tau H_{2t}(W_{t+1}^2 - W_t^2). \end{cases} \quad (5.41)$$

where

$$\begin{aligned} H_{1t} &= \begin{cases} 1 & \text{for model 2f-A,} \\ \sqrt{S_t} & \text{for model 2f-B,} \\ S_t & \text{for model 2f-C,} \end{cases} \\ H_{2t} &= \begin{cases} 1 & \text{for model 2f-A,} \\ \sqrt{L_t} & \text{for model 2f-B,} \\ L_t & \text{for model 2f-C.} \end{cases} \end{aligned} \quad (5.42)$$

Given  $S_t$  and  $L_t$ , with the fact that  $(W_{t+1}^1 - W_t^1) \sim \mathcal{N}(0, 1)$ , we have  $S_{t+1} \sim \mathcal{N}(\alpha(L_t - S_t) + S_t, \sigma^2 H_{1t}^2)$ . Similarly, we have  $L_{t+1} \sim \mathcal{N}(\mu(\gamma - L_t) + L_t, \tau^2 H_{2t}^2)$ .

Since  $(W_{t+1}^1 - W_t^1)$  and  $(W_{t+1}^2 - W_t^2)$  are independent, the joint conditional PDFs of  $[S_{t+1}, L_{t+1}]$  given  $[S_t, L_t]$  can be approximated by the product of the conditional PDFs of each variable, say,

$$\begin{aligned} f(\theta, [S_{t+1}, L_{t+1}]|[S_t, L_t]) &\approx f(\theta, S_{t+1}|[S_t, L_t]) \cdot f(\theta, L_{t+1}|[S_t, L_t]) \\ &= \frac{1}{\sqrt{2\pi\sigma}H_{1t}} \exp\left(-\frac{(\alpha(L_t - S_t) + S_t - S_{t+1})^2}{2\sigma^2 H_{1t}^2}\right) \\ &\quad \cdot \frac{1}{\sqrt{2\pi\tau}H_{2t}} \exp\left(-\frac{(\mu(\gamma - L_t) + L_t - L_{t+1})^2}{2\tau^2 H_{2t}^2}\right). \end{aligned} \quad (5.43)$$

### 5.4.3 Estimation

Suppose to begin with that we know the values of  $\{S_t\}_{t=1}^n$  and  $\{L_t\}_{t=1}^n$ . Following the general procedure of the ML method, if we let

$$\mathcal{L} = \log \prod_{t=1}^{n-1} f(\theta, [S_{t+1}, L_{t+1}]|[S_t, L_t]), \quad (5.44)$$

then by setting  $\partial\mathcal{L}/\partial\theta = 0$ , we can get estimates of each parameter:

$$\left\{ \begin{aligned} \frac{\partial\mathcal{L}}{\partial\alpha} = 0 &\Rightarrow \sum_{t=1}^{n-1} \frac{(L_t - S_t)(\alpha(L_t - S_t) + S_t - S_{t+1})}{\sigma^2 H_{1t}^2} = 0 \\ \frac{\partial\mathcal{L}}{\partial\sigma} = 0 &\Rightarrow \sum_{t=1}^{n-1} \left( -\frac{1}{\sigma} + \frac{(\alpha(L_t - S_t) + S_t - S_{t+1})^2}{\sigma^3 H_{1t}^2} \right) = 0 \\ \frac{\partial\mathcal{L}}{\partial\mu} = 0 &\Rightarrow \sum_{t=1}^{n-1} \frac{(\gamma - L_t)(\mu(\gamma - L_t) + L_t - L_{t+1})}{\tau^2 H_{2t}^2} = 0 \\ \frac{\partial\mathcal{L}}{\partial\gamma} = 0 &\Rightarrow \sum_{t=1}^{n-1} \frac{\mu(\mu(\gamma - L_t) + L_t - L_{t+1})}{\tau^2 H_{2t}^2} = 0 \\ \frac{\partial\mathcal{L}}{\partial\tau} = 0 &\Rightarrow \sum_{t=1}^{n-1} \left( -\frac{1}{\tau} + \frac{(\mu(\gamma - L_t) + L_t - L_{t+1})^2}{\tau^3 H_{2t}^2} \right) = 0. \end{aligned} \right. \quad (5.45)$$

Solving the above equations, we obtain

$$\begin{aligned}
\alpha &= \frac{\sum_{t=1}^{n-1} \left( (S_t - S_{t+1})(S_t - L_t) / H_{1t}^2 \right)}{\sum_{t=1}^{n-1} \left( (S_t - L_t)^2 / H_{1t}^2 \right)}, \\
\sigma &= \sqrt{\frac{1}{n-1} \sum_{t=1}^{n-1} \frac{(\alpha(L_t - S_t) + S_t - S_{t+1})^2}{H_{1t}^2}}, \\
\mu &= \frac{DE - CF}{C^2 - BE}, \\
\gamma &= -\frac{B}{C} - \frac{D}{\mu C}, \\
\tau &= \sqrt{\frac{1}{n-1} \sum_{t=1}^{n-1} \frac{(\mu(\gamma - L_t) + L_t - L_{t+1})^2}{H_{2t}^2}},
\end{aligned} \tag{5.46}$$

where

$$\begin{aligned}
B &= \sum_{t=1}^{n-1} \frac{L_t^2}{H_{2t}^2}, & C &= -\sum_{t=1}^{n-1} \frac{L_t}{H_{2t}^2}, & D &= \sum_{t=1}^{n-1} \frac{L_t(L_{t+1} - L_t)}{H_{2t}^2}, \\
E &= \sum_{t=1}^{n-1} \frac{1}{H_{2t}^2}, & F &= \sum_{t=1}^{n-1} \frac{L_t - L_{t+1}}{H_{2t}^2}.
\end{aligned}$$

In practice, these formulas are applied to the values of  $L_t$  obtained from the futures prices as described in the previous section.

#### 5.4.4 Results

##### Model 2f-A

Table 5.13 shows the estimation of the parameters in model 2f-A for both artificial data and real data.

	For artificial data			For real data
	Given	Mean of Est.	Std. of Est.	Est.
$\alpha$	0.0170	0.0172	0.0034	0.0173
$\sigma$	0.0900	0.0898	0.0014	0.0870
$\mu$	0.0060	0.0080	0.0034	0.0056
$\gamma$	2.3000	2.3684	0.1774	2.2957
$\tau$	0.0500	0.0502	0.0007	0.0465

Table 5.13: Estimation result for model 2f-A

Note: From the table, we see that the estimation of the parameters for artificial data recovers the given values fairly well. This implies that our method works and the programs are doing their job correctly, and therefore we believe that the estimation for real data is reliable.

More information about the calibration is given in Figure 5.11, where the goodness of fit and the shape of the objective function near the optimal values of  $\tilde{\theta}$  are shown, together with a sample simulation to give a visual impression of the effectiveness of the model.

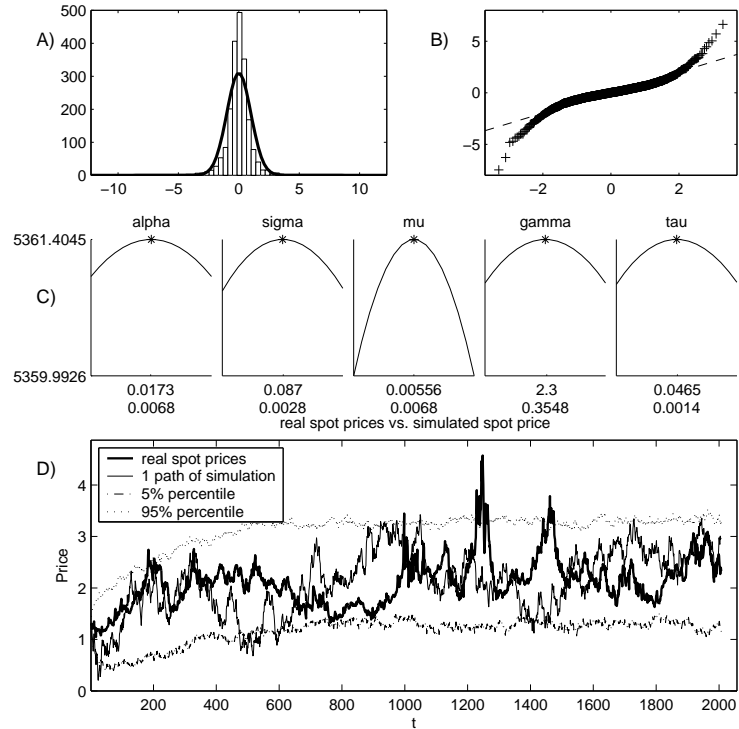


Figure 5.11: Result of model 2f-A

Note: A) is the histogram of  $Y_t = \frac{\alpha(L_t - S_t) + S_t - S_{t+1}}{\sigma}$ , with the graph of the standard normal distribution PDF  $f(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$  superimposed; B) is the quantile-quantile plot of  $Y_t$  vs. the standard normal distribution; C) are the intersecting slices of log likelihood function with respect to the parameters estimated by ML method. Under the plots, the numbers in the first row are the estimates of the parameters, the numbers in the second row are the widths of the plots, which are just 2 times the third column of Table 5.13. Our estimates are all on the top of slices, which implies that we successfully maximize the likelihood; D) is one simulated spot price path vs. the actual spot price path.

**Model 2f-B**

Table 5.14 shows the estimation of the parameters in model 2f-B for both artificial data and real data. More information about the calibration is given in Figure 5.12.

	For artificial data			For real data
	Given	Mean of Est.	Std. of Est.	Est.
$\alpha$	0.0150	0.0156	0.0036	0.0148
$\sigma$	0.0550	0.0551	0.0010	0.0549
$\mu$	0.0060	0.0083	0.0028	0.0063
$\gamma$	2.3000	2.3027	0.1624	2.2772
$\tau$	0.0300	0.0299	0.0005	0.0335

Table 5.14: Estimation result for model 2f-B

Note: Refer to the note in Table 5.13.

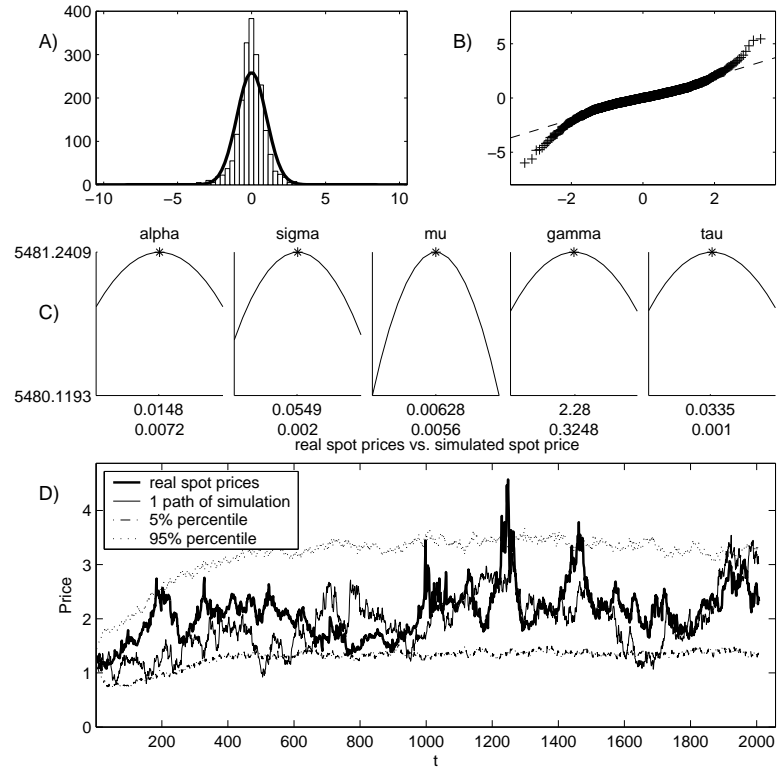


Figure 5.12: Result of model 2f-B

Note: A) is the histogram of  $Y_t = \frac{\alpha(L_t - S_t) + S_t - S_{t+1}}{\sigma\sqrt{S_t}}$ , with the graph of the standard normal distribution PDF  $f(x) = \frac{1}{\sqrt{2\pi}}\exp(-\frac{x^2}{2})$  superimposed; For the description of other plots, please refer to the note of Figure 5.11.

**Model 2f-C**

Table 5.15 shows the estimation of the parameters in model 2f-C for both artificial data and real data. More information about the calibration is given in Figure 5.13.

	For artificial data			For real data
	Given	Mean of Est.	Std. of Est.	Est.
$\alpha$	0.0140	0.0151	0.0034	0.0140
$\sigma$	0.0360	0.0360	0.0006	0.0360
$\mu$	0.0080	0.0093	0.0033	0.0078
$\gamma$	2.2000	2.1854	0.1461	2.2264
$\tau$	0.0250	0.0250	0.0004	0.0253

Table 5.15: Estimation result for model 2f-C

Note: Refer to the note in Table 5.13.

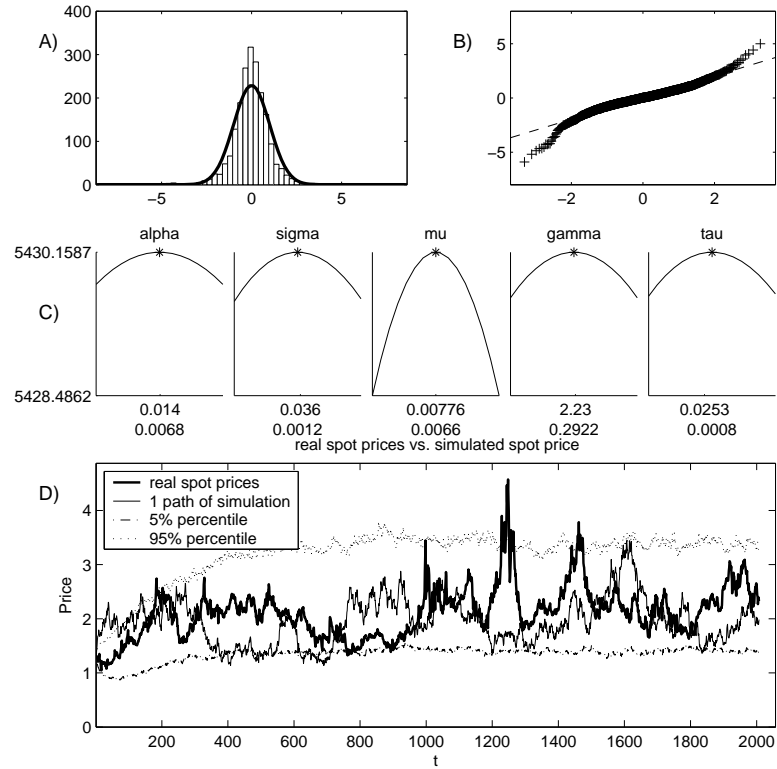


Figure 5.13: Result of model 2f-C

Note: A) is the histogram of  $Y_t = \frac{\alpha(L_t - S_t) + S_t - S_{t+1}}{\sigma S_t}$ , with the graph of the standard normal distribution PDF  $f(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$  superimposed; For the description of other plots, please refer to the note of Figure 5.11.

### 5.4.5 Futures Matching

Again, we check whether our models fit the futures market well. Actually, when we calibrated our models, we considered futures prices as the inputs. In this process of “excavating” the unobservable factor  $L_t$ , futures prices have been matched as well as possible. The corresponding minimum obtained in (5.40) is 426.249. The upper plot in Figure 5.14 shows how well they are matched for a typical futures curve.

The lower plot in Figure 5.14 shows how well the expectations of these spot prices match the futures prices in the risk-neutral world.

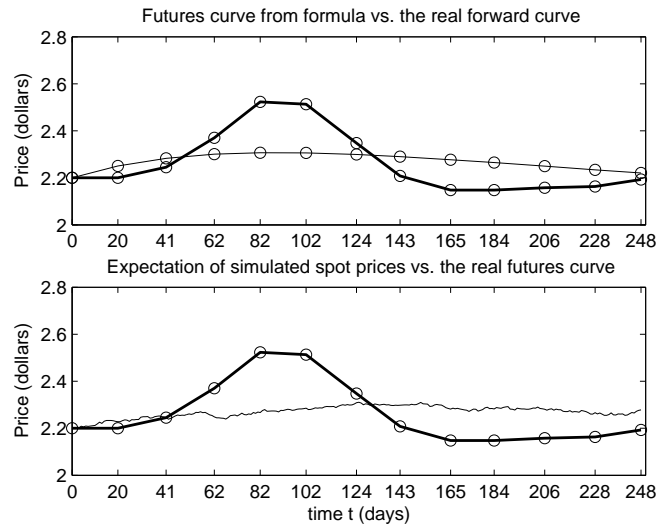


Figure 5.14: Futures matching for 2-factor models without seasonality

Note: The bold curves in both of the plots are the real futures curve, which is the same as that in Figure 5.8. The thinner curve in the upper plot is graph of futures function given by (4.24). The thinner curve in the lower plot is the mean values of 250 paths of  $S_t$  simulated by SDEs (3.22), where the risk-neutral parameters shown in Table 5.12 are plugged in, and  $r = 1$ .

As shown in Figure 5.14, since seasonality is not captured, it seems our 2-factor models without seasonality do not match the futures prices very well.

## 5.5 Calibration for 2-factor Models with Seasonality

In this class of models, gas spot prices follow the mean reverting process described by (3.23)-(3.25). By setting  $r = 0$ ,  $r = \frac{1}{2}$  and  $r = 1$ , respectively, we have three specific models: model 2f-sn-A, model 2f-sn-B and model 2f-sn-C.

As the extension of the previous, relatively simpler, 2-factor models without seasonality, these models' calibration follows almost the same main ideas of the former ones. First, we “excavate” the hidden things, including the seasonal term  $f(t)$  and the long-term mean factor  $L_t$  by matching the actual futures prices with the theoretical futures prices. In this process, we can obtain  $X_t$ , parameters for  $f(t)$  and some risk-neutralized parameters. Next, with the known  $X_t$  and  $L_t$ 's values, we get parameters in these two stochastic processes by the ML method.

### 5.5.1 Find the Hidden Factor $L_t$ and Seasonal Term $f(t)$ from Futures Prices

For this class of models, the way we find hidden things is the same as that in the 1-factor models with seasonality and the 2-factor models without seasonality.

If we let  $A_i = f(T_{ti}) + e^{\tilde{\alpha}(t-T_{ti})}(S_t - f(t)) + \frac{\tilde{\mu}\tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}}(e^{\tilde{\alpha}(t-T_{ti})} - 1) - \frac{\tilde{\alpha}\tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}}(e^{\tilde{\mu}(t-T_{ti})} - 1)$  and  $B_i = \frac{\tilde{\alpha}}{\tilde{\alpha} - \tilde{\mu}}(e^{\tilde{\mu}(t-T_{ti})} - e^{\tilde{\alpha}(t-T_{ti})})$ , then by (4.25) we have the futures price function

$$F^{\tilde{\theta}}(t, T_{ti}, (S_t, L_t)) = A_i + B_i L_t.$$

Notice that the futures price function involves not only  $[\tilde{\alpha}, \tilde{\mu}, \tilde{\gamma}]$  and  $L_t$ , but also  $f(t)$ , hence, we can get all of them by matching futures prices.

Here, we let  $\tilde{\theta} = [\tilde{\alpha}, \tilde{\mu}, \tilde{\gamma}, b, \beta_1, \dots, \beta_N, \eta_1, \dots, \eta_N]$ . In the same way as before,  $L_t$

can be defined as a function of  $\tilde{\theta}$ ,

$$L_t(\tilde{\theta}) = \frac{\sum_{i=1}^m (B_i F_{t, T_{ti}} - A_i B_i)}{\sum_{i=1}^m B_i^2}, \quad (5.47)$$

Then we free  $\tilde{\theta}$  to obtain the optimal one by

$$\tilde{\theta} = \arg \min_{\tilde{\theta} \in \mathbb{R}^{n_0}} \sum_{t=1}^n \sum_{i=1}^m (F^{\tilde{\theta}}(t, T_{ti}, (S_t, L_t(\tilde{\theta}))) - F_{t, T_{ti}})^2, \quad (5.48)$$

where  $n_0$  is the vector length of  $\tilde{\theta}$ . With the obtained  $\tilde{\theta}$ ,  $\{L_t\}_{t=1}^n$ ,  $\{f(t)\}_{t=1}^n$ , and hence  $\{X_t\}_{t=1}^n$  are all easily computed.

Table 5.16 shows the estimation of the risk-neutral parameters for artificial data and real data.

	For artificial data			For real data
	Given	Mean of Est.	Std. of Est.	Est.
$\tilde{\alpha}$	0.0110	0.0110	0	0.0111
$\tilde{\mu}$	0.0020	0.0020	0	0.0018
$\tilde{\gamma}$	2.0000	2.0000	0	2.0028

Table 5.16: Estimation of the risk-neutral parameters for 2-factor models with seasonality

Note: From the table, we see that the estimation of the risk-neutral parameters for artificial data recovers the given values accurately. This implies that our method works and the programs are doing their job correctly, and therefore we believe that the estimation for real data is reliable.

For the estimated values of the parameters for  $f(t)$ , see Table 5.17 in Section 5.5.4.

Figure 5.15 and Figure 5.16 show the recovering of  $L_t$ 's values for artificial data and real data.

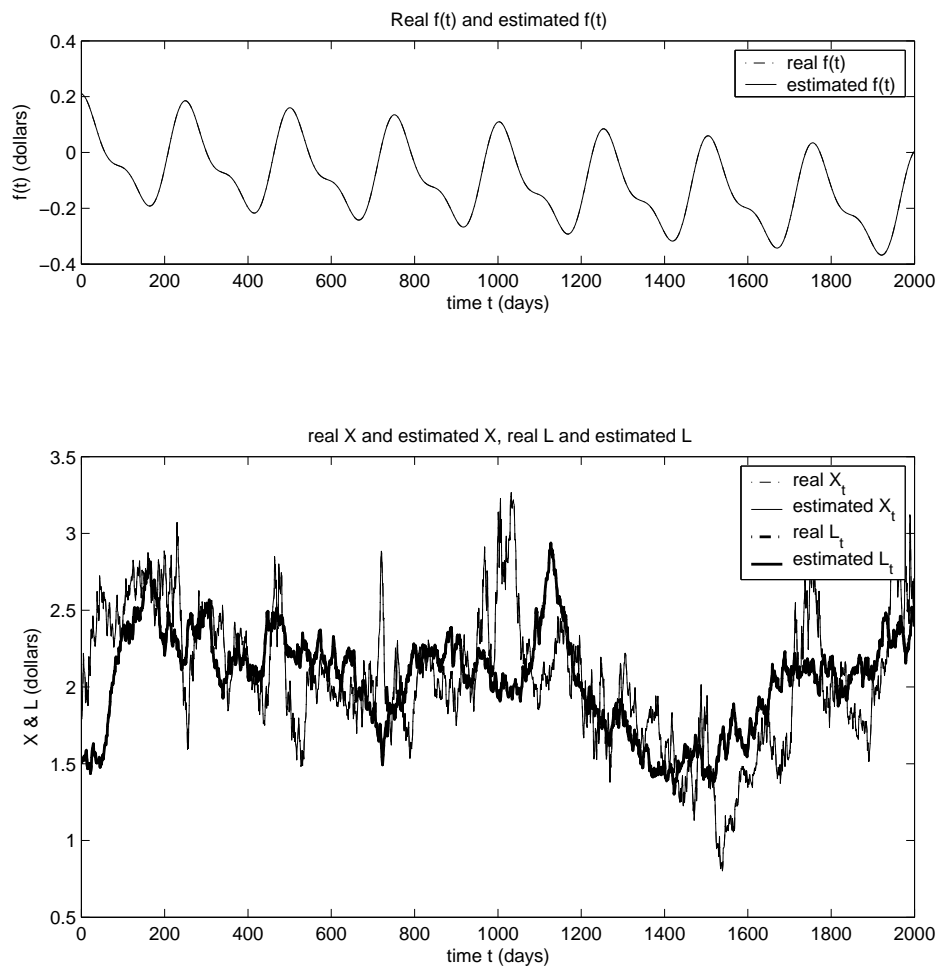


Figure 5.15: Recover  $L_t$  and  $f(t)$  from futures prices for artificial data  
 Note: From the figure, we see that the estimated  $X_t$ ,  $L_t$  and  $f(t)$  exactly cover the true values, which implies that our method works and the programs are doing their job correctly.

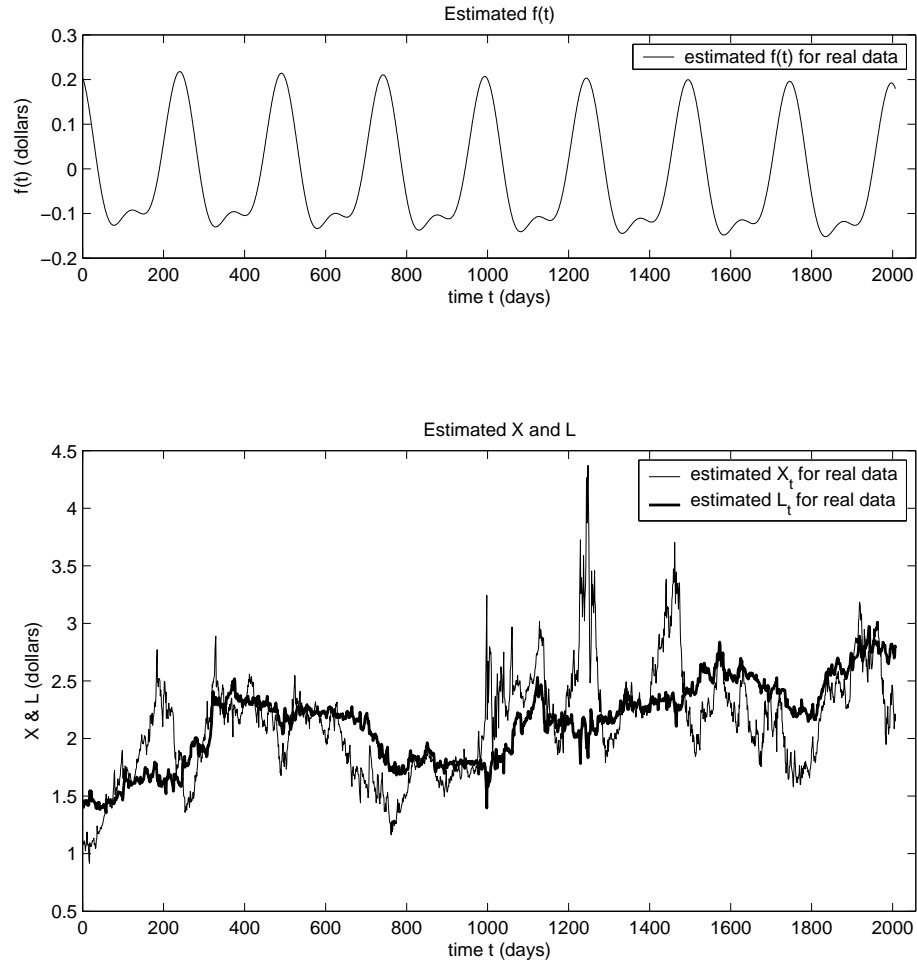


Figure 5.16: Reveal  $L_t$  and  $f(t)$  from futures prices for real data

We see from Figure 5.16 that the values for  $L_t$  recovered from the real data are much more realistic than those in Figure 5.10 when interpreted as the long term mean.

### 5.5.2 Probability Density Function

For the sake of simplicity, we just give the the approximate joint conditional PDF of  $[X_{t+1}, L_{t+1}]$  given  $[X_t, L_t]$  deduced in discrete-time framework, which is used in our calibration. The method by which we obtain it is the same as before.

$$f(\theta, [X_{t+1}, L_{t+1}]|[X_t, L_t]) \approx \frac{1}{\sqrt{2\pi}\sigma(t)H_{1t}} \exp\left(-\frac{(\alpha(L_t - X_t) + X_t - X_{t+1})^2}{2\sigma(t)^2 H_{1t}^2}\right) \cdot \frac{1}{\sqrt{2\pi}\tau H_{2t}} \exp\left(-\frac{(\mu(\gamma - L_t) + L_t - L_{t+1})^2}{2\tau^2 H_{2t}^2}\right), \quad (5.49)$$

where  $H_{1t}$  and  $H_{2t}$  are redefined as

$$H_{1t} = \begin{cases} 1 & \text{for model 2f-sn-A,} \\ \sqrt{X_t} & \text{for model 2f-sn-B,} \\ X_t & \text{for model 2f-sn-C,} \end{cases} \quad (5.50)$$

$$H_{2t} = \begin{cases} 1 & \text{for model 2f-sn-A,} \\ \sqrt{L_t} & \text{for model 2f-sn-B,} \\ L_t & \text{for model 2f-sn-C.} \end{cases}$$

### 5.5.3 Estimation

Applying the general procedure of the ML method, we get estimates of the parameters in (3.23)-(3.25). We find for parameters involving  $L_t$ , i.e.,  $\mu$ ,  $\gamma$  and  $\tau$ , the results from the 2-factor models without seasonality (see Section 5.4.3) are usable. The corresponding estimates are the same as those in (5.46),

$$\begin{aligned}\mu &= \frac{DE - CF}{C^2 - BE}, \\ \gamma &= -\frac{B}{C} - \frac{D}{\mu C}, \\ \tau &= \sqrt{\frac{1}{n-1} \sum_{t=1}^{n-1} \frac{(\mu(\gamma - L_t) + L_t - L_{t+1})^2}{H_{2t}^2}},\end{aligned}\tag{5.51}$$

where  $B$ ,  $C$ ,  $D$ ,  $E$  and  $F$  are the same as those in (5.46).

For parameters involving  $X_t$ , i.e.,  $\alpha$ ,  $c$ ,  $\lambda_1, \dots, \lambda_K$ ,  $\omega_1, \dots, \omega_K$ , the results from the 1-factor models with seasonality are usable (see Section 5.3.3). The corresponding estimates are obtained by numerically solving some similar equations in (5.34), say,

$$\left\{ \begin{aligned} \sum_{t=1}^{n-1} \frac{(L_t - X_t)(\alpha(L_t - X_t) + X_t - X_{t+1})}{\sigma(t)^2 H_{1t}^2} &= 0 \\ \sum_{t=1}^{n-1} \left( -1 + \frac{(\alpha(L_t - X_t) + X_t - X_{t+1})^2}{\sigma(t)^2 H_{1t}^2} \right) &= 0 \\ \sum_{t=1}^{n-1} \left( -1 + \frac{(\alpha(L_t - X_t) + X_t - X_{t+1})^2}{\sigma(t)^2 H_{1t}^2} \right) \cos\left(\frac{2\pi jt}{P}\right) &= 0 \\ \sum_{t=1}^{n-1} \left( -1 + \frac{(\alpha(L_t - X_t) + X_t - X_{t+1})^2}{\sigma(t)^2 H_{1t}^2} \right) \sin\left(\frac{2\pi jt}{P}\right) &= 0, \end{aligned} \right.\tag{5.52}$$

where  $\sigma(t)$  and  $P$  are defined in (3.23)-(3.25).

The parameters in the seasonal term  $f(t)$  have already been estimated in the process of revealing  $L_t$ 's values.

### 5.5.4 Results

Since the three 2-factor models with seasonality share the same futures price function given by (4.25), the parameters for  $f(t)$  have the same estimation values for these three models. The results are shown in Table 5.17.

	For artificial data			For real data
	Given	Mean of Est.	Std. of Est.	Est.
$b$	-0.0001	-0.0001	0	-1.4476e-5
$\beta_1$	0.1500	0.1500	0	0.1483
$\beta_2$	0.0600	0.0600	0	0.0574
$\eta_1$	-0.0500	-0.0500	0	-0.0527
$\eta_2$	-0.0300	-0.0300	0	-0.0292

Table 5.17: Estimation result for  $f(t)$  in 2-factor models with seasonality

Note: From the table, we see that the estimation of the parameters for artificial data recovers the given values fairly well. This implies that our method works and the programs are doing their job correctly, and therefore we believe that the estimation for real data is reliable. PS: The reason for the extreme accuracy of the parameters for  $f(t)$  is that they are directly estimated through matching futures prices.

The other parameters have different estimation values for each model. The results are shown below. In Tables 5.18-5.20, we give the results for model 2f-sn-A, model 2f-sn-B and model 2f-sn-C, respectively. In each case we report the estimated values of the parameters, and also the results of robustness analysis carried out using simulated data. More information about the estimations is provided in Figures 5.17 - 5.19, where the goodness of fit, the shape of the objective function near the optimal values of  $\tilde{\theta}$ , and the behaviour of the seasonally varying terms  $f(t)$  and  $\sigma(t)$  are shown, together with a sample simulation to give a visual impression of the effectiveness of the model.

**Model 2f-sn-A**

Table 5.18 shows the estimation of the parameters in model 2f-sn-A for both artificial data and real data.

	For artificial data			For real data
	Given	Mean of Est.	Std. of Est.	Est.
$\alpha$	0.0240	0.0252	0.0045	0.0244
$c$	-2.6000	-2.6025	0.0180	-2.5616
$\lambda_1$	0.4000	0.4021	0.0203	0.4105
$\lambda_2$	0.0800	0.0777	0.0267	0.0790
$\omega_1$	-0.2000	-0.1983	0.0202	-0.1982
$\omega_2$	-0.0700	-0.0683	0.0231	-0.0660
$\mu$	0.0050	0.0068	0.0030	0.0048
$\tau$	0.0300	0.0299	0.0004	0.0329
$\gamma$	2.3000	2.2833	0.1297	2.2860

Table 5.18: Estimation result for model 2f-sn-A

Note: From the table, we see that the estimation of the parameters for artificial data recovers the given values fairly well. This implies that our method works and the programs are doing their job correctly, and therefore we believe that the estimation for real data is reliable.

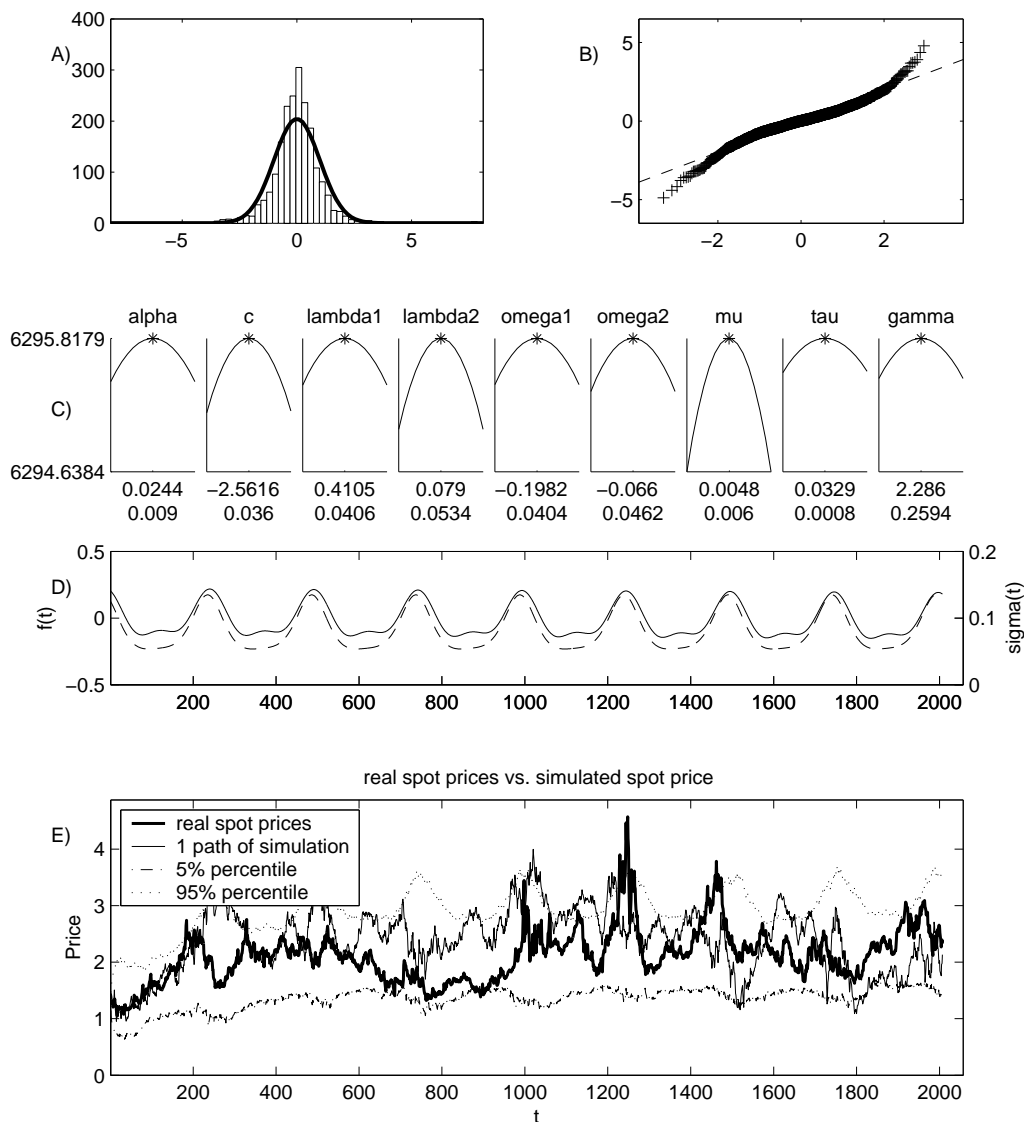


Figure 5.17: Result for model 2f-sn-A

Note: A) is the histogram of  $Y_t = \frac{\alpha(L_t - X_t) + X_t - X_{t+1}}{\sigma(t)}$ , with the graph of the standard normal distribution PDF  $f(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$  superimposed; B) is the quantile-quantile plot of  $Y_t$  vs. the standard normal distribution; C) are the intersecting slices of log likelihood function with respect to the parameters estimated by ML method. Under the plots, the numbers in the first row are the estimates of the parameters, the numbers in the second row are the widths of the plots, which are just 2 times the third column of Table 5.18. Our estimates are all on the top of slices, which implies that we successfully maximize the likelihood; D) is the graphs of  $f(t)$  (solid curve) and  $\sigma(t)$  (dashed curve); E) is one simulated spot price path vs. the actual spot price path.

**Model 2f-sn-B**

Table 5.19 shows the estimation of the parameters in model 2f-sn-B for both artificial data and real data.

	For artificial data			For real data
	Given	Mean of Est.	Std. of Est.	Est.
$\alpha$	0.0220	0.0230	0.0048	0.0221
$c$	-3.0000	-3.0030	0.0169	-2.9894
$\lambda_1$	0.3800	0.3787	0.0212	0.3791
$\lambda_2$	0.0750	0.0743	0.0198	0.0748
$\omega_1$	-0.1800	-0.1828	0.0229	-0.1826
$\omega_2$	-0.0300	-0.0333	0.0224	-0.0317
$\mu$	0.0050	0.0075	0.0037	0.0050
$\tau$	0.0230	0.0228	0.0004	0.0230
$\gamma$	2.3000	2.2866	0.1419	2.2793

Table 5.19: Estimation result for model 2f-sn-B

Note: Please refer to the note of Table 5.18.

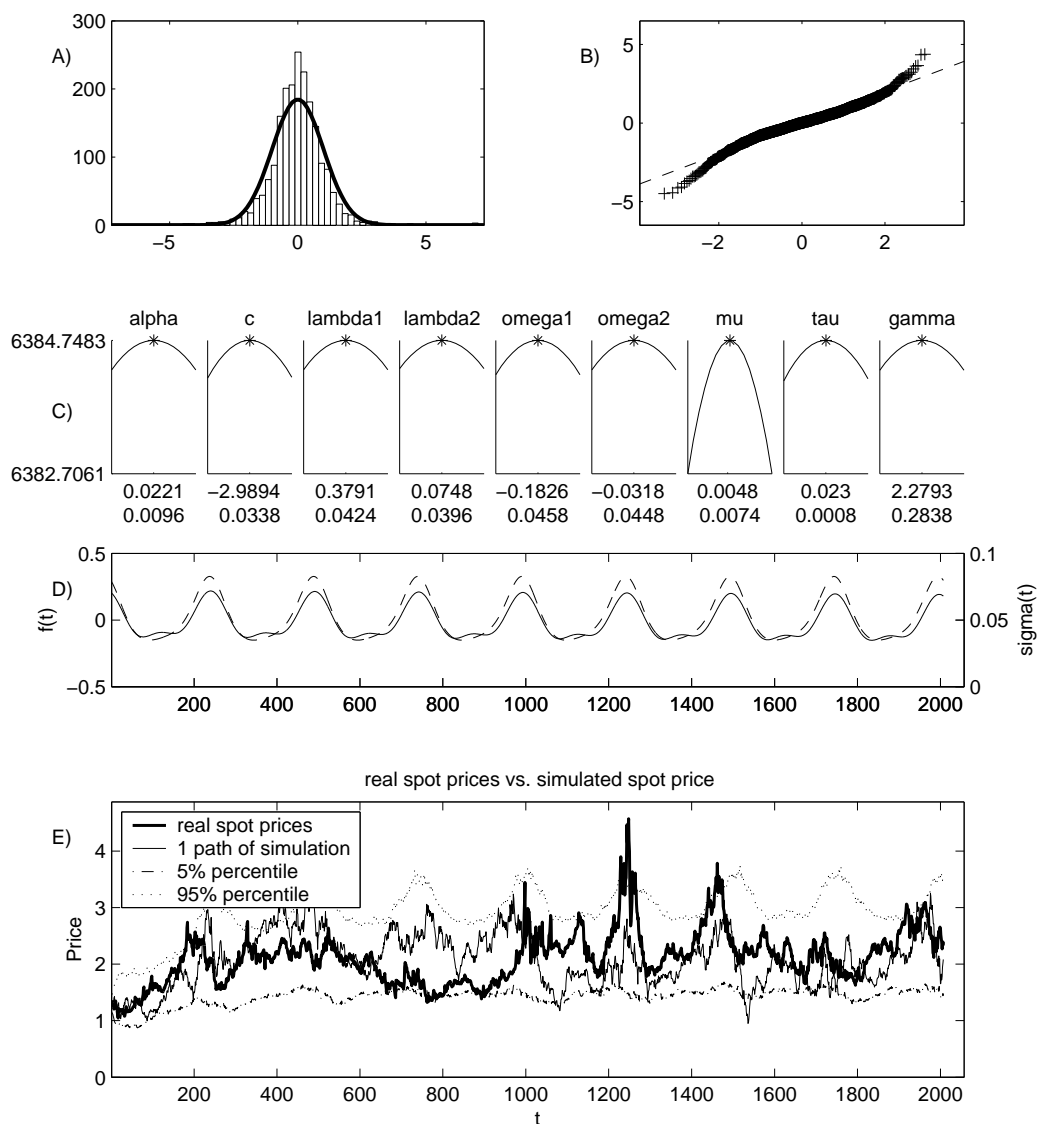


Figure 5.18: Result for model 2f-sn-B

Note: A) is the histogram of  $Y_t = \frac{\alpha(L_t - X_t) + X_t - X_{t+1}}{\sigma(t)\sqrt{X_t}}$ , with the graph of the standard normal distribution PDF  $f(x) = \frac{1}{\sqrt{2\pi}}\exp(-\frac{x^2}{2})$  superimposed; For the description of other plots, please refer to the note of Figure 5.17.

**Model 2f-sn-C**

Table 5.20 shows the estimation of the parameters in model 2f-sn-C for both artificial data and real data.

	For artificial data			For real data
	Given	Mean of Est.	Std. of Est.	Est.
$\alpha$	0.0210	0.0222	0.0043	0.0212
$c$	-3.4000	-3.4033	0.0160	-3.3913
$\lambda_1$	0.3600	0.3646	0.0232	0.3639
$\lambda_2$	0.0700	0.0699	0.0180	0.0724
$\omega_1$	-0.1700	-0.1674	0.0217	-0.1693
$\omega_2$	0.0030	0.0027	0.0240	0.0030
$\mu$	0.0050	0.0069	0.0035	0.0054
$\tau$	0.0160	0.0160	0.0002	0.0164
$\gamma$	2.3000	2.3162	0.1611	2.2652

Table 5.20: Estimation result for model 2f-sn-C

Note: Please refer to the note of Table 5.18.

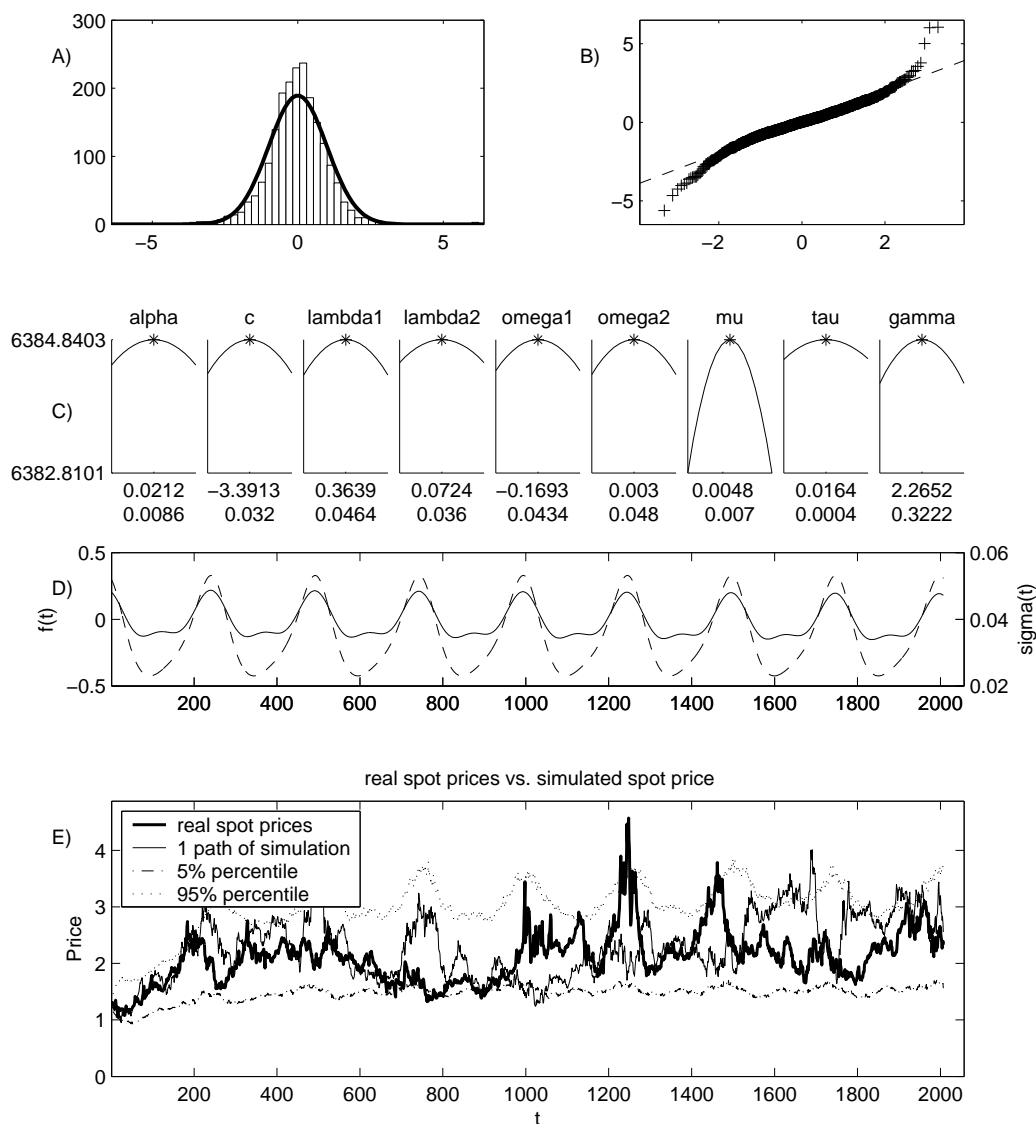


Figure 5.19: Result for model 2f-sn-C

Note: A) is the histogram of  $Y_t = \frac{\alpha(L_t - X_t) + X_t - X_{t+1}}{\sigma(t)X_t}$ , with the graph of the standard normal distribution PDF  $f(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$  superimposed; For the description of other plots, please refer to the note of Figure 5.17.

### 5.5.5 Futures Matching

Again, we check whether our models fit the futures market well.

Similarly, in this process of “excavating” the unobservable factor  $L_t$  and seasonal term  $f(t)$ , futures prices have been matched. The corresponding minimum obtained in (5.48) is 113.145, which is the smallest among all the models in this thesis. The middle plot in Figure 5.20 shows how well they are matched for a typical futures curve.

Also, we simulate lots of paths of spot prices by one of our models with the risk-neutral parameters plugged in. The lower plot in Figure 5.20 shows how well the expectations of these spot prices match the futures prices.

As shown in Figure 5.20, it seems our 2-factor models with seasonality match the futures prices fairly well. Among all of the models introduced in this thesis, this class of models are the best in this aspect.

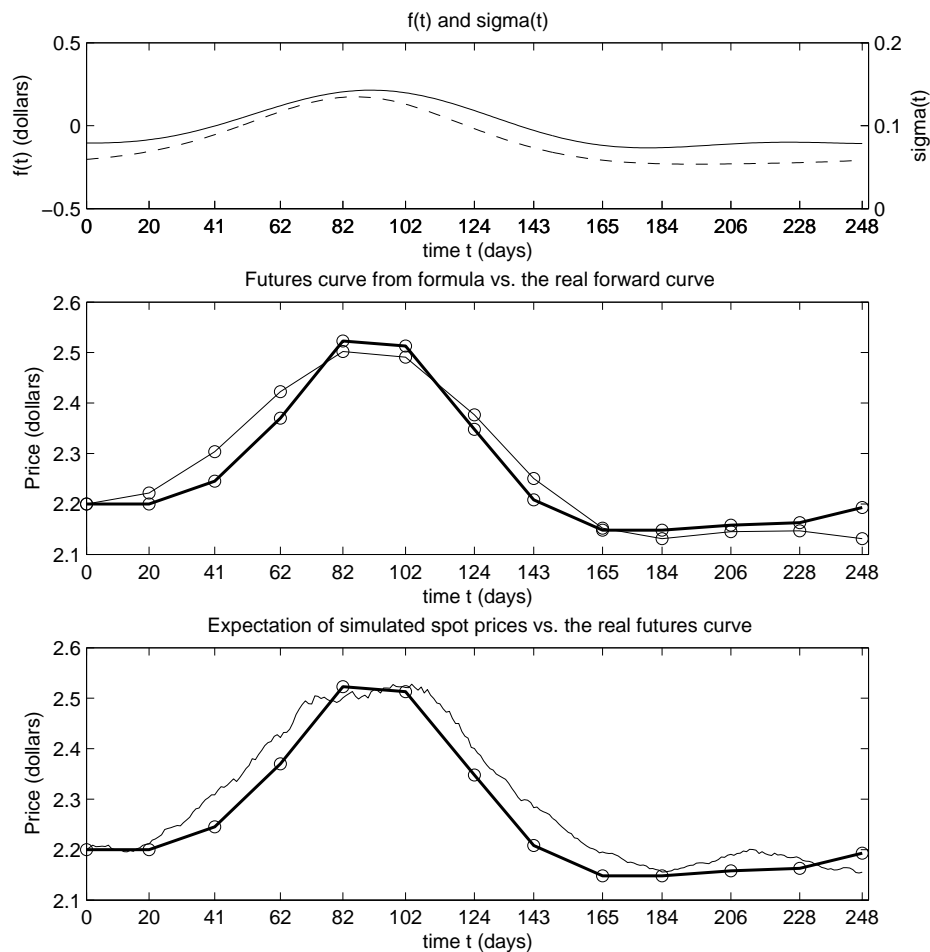


Figure 5.20: Futures matching for 2-factor models with seasonality

Note: In the middle plot, the bold curve is the real futures curve, which is the same as that in Figure 5.8 and Figure 5.14. The thinner curve is graph of futures function given by (4.25). In the lower plot, the bold curve is the real futures curve. The thinner curve is the mean values of 250 paths of  $S_t$  simulated by equations (3.23)-(3.25), where the risk-neutral parameters shown in Table 5.16 are plugged in, and  $r = 1$ . For comparison, the upper plot shows the seasonal term  $f(t)$  (solid curve), and the seasonal volatility function  $\sigma(t)$  (dashed curve).

## Chapter 6

### Discussion and Conclusion

The natural gas market has been going through a period of deregulation and the elimination of stage monopolies in North America. It has grown rapidly and has become a fairly liberal market. The keen competition between market players drives the need to understand the dynamics of the gas market for risk management and asset pricing.

Due to some physical constraints such as the local nature of markets, volume constraints on transfer and storage, different elasticities in supply and demand, etc., natural gas markets present very specific features. This makes it a challenging task for researchers and practitioners to model their dynamics adequately.

In this thesis, we have focused on issues of modeling the behaviour of natural gas spot prices, in particular, unpredictable fluctuations, mean reversion, seasonality, etc. In this process, we examined a range of stochastic mean reverting models.

The first models we considered were two 1-factor models, differing only in the way they form the volatility term. By studying these two relatively simple models, we addressed some issues which would be faced in the process of fitting models via maximum likelihood estimation to actual gas prices. We introduced how to get conditional PDFs in the framework of affine jump-diffusion transform theory. Then we demonstrated that the PDFs obtained in this way could be fairly well approximated by some simpler PDFs derived in discrete-time framework. Using the approximate PDFs, the calibrations of the other complex models were extremely

simplified. By looking at the volatilities of prices, we found that prices behave differently on non-trading days, which drove us to treat them distinctively in the process of calibration. By comparing the estimation results from continuous data and discrete data, we got the confidence to do the calibration using daily data. We also studied the influences of data length to the estimation results.

After investigating the above issues, we moved to some practicable models.

We examined a class of 1-factor models incorporated with seasonality. In the process of calibration, we introduced a key technique developed in this thesis — recovering the unobservable information by exploiting the relationship between futures prices and spot prices. Afterward, we studied how well our models could match the futures prices behaviour, and found that the 1-factor models did match the shape of the futures curve qualitatively, but were not quantitatively very accurate.

Then we moved to study a class of 2-factor models without seasonality. Similarly, we used the technique mentioned above to reveal values of the hidden factor and then applied the maximum likelihood estimation. The futures matching results showed that, although these 2-factor models performed a little bit better than 1-factor models with seasonality did on the whole, they could not even capture the shape of the futures curves.

Finally, we combined the previous two class of models together and developed some 2-factor seasonal models. For these models, both seasonal terms and hidden factor values could be revealed by exploiting the relationship between futures prices and spot prices, which enabled us to perform the maximum likelihood estimation easily. These models matched futures prices behaviour very successfully. Not only the number measuring the distance between the actual futures prices and the model-

based theoretical prices was much less than those of the previous two models, but also the shape of futures curves could be perfectly matched.

It seems our 2-factor seasonal models have successfully captured most of the gas price features we had attempted to model, i.e., the mean reversion, seasonality, and they also match futures prices fairly well. However, there are still some weaknesses in these models. One evidence of their imperfection is given by the Q-Q plots of the random terms in these models. The shape of these Q-Q plots shows that the distribution of the actual spot prices is not very consistent with what our models suppose them to be, especially in the tails. Hence, more work is still required.

In future research, there are mainly two directions we can try. We can exploit the results we have achieved. For example, the risk-neutral parameters obtained through matching futures prices can be used to construct models for the risk-neutral world, and therefore can be used for the pricing of many other derivatives, such as swaps, options, etc. The calibrated spot price models we have obtained can also be used in risk management directly.

On the other hand, there are still areas in which we can improve the effectiveness of our models. We can compare the long term mean estimates with other calibration results, for example, Kalman filter based methods, to see whether our approach can be improved by exploiting their ideas. Another possible direction is to construct the models in some other ways. For example, instead of simply assuming that the volatility is a deterministic function of time, we can suppose that the volatility changes stochastically. This will end up with some 3-factor models. We can also incorporate the concept of “jumps” into the model to capture those sudden big changes in prices. Both of these ideas have been proposed as a means of dealing

with the discrepancies uncovered by the Q-Q plots (see [24]). Since one of the main use of natural gas is to generate electricity, considering the strong correlation between them, another area of future research would be to try to model the behaviour of them simultaneously. Furthermore, instead of modeling spot prices directly, we can try to model futures prices first, and then investigate the relationship between them, as is often done in interest rate model.

As a fundamental problem in finance and some other related fields, modeling the behaviour of prices is very important and worthy of more research. But anyway, the best model does not exist, what we ought to do is just keep making efforts to find a better one, then a better one, better and better...

# Appendix A

## Transform Analysis for Affine Jump-Diffusion State-Vectors

The content of this part is from Duffie, Pan and Singleton (see [23]).

### A.1 Affine Jump-Diffusion

Suppose  $\mathbf{X}$  is a *strong Markov process*<sup>1</sup> with realizations  $\{\mathbf{X}_t, 0 \leq t < \infty\}$  in some state space  $D \subset \mathbb{R}^n$ . Under certain regularity conditions<sup>2</sup>,  $\mathbf{X}$  uniquely solves the stochastic differential equation (SDE):

$$d\mathbf{X}_t = \mu(\mathbf{X}_t, t)dt + \sigma(\mathbf{X}_t, t)d\mathbf{W}_t + d\mathbf{Z}_t, \quad (\text{A.1})$$

where  $\mu : (D, \mathbb{R}_+) \mapsto \mathbb{R}^n$ ,  $\sigma : (D, \mathbb{R}_+) \mapsto \mathbb{R}^{n \times n}$ ;  $\mathbf{W}$  is an Standard Brownian motion in  $\mathbb{R}^n$ ;  $\mathbf{Z}$  is a pure jump process with arrival intensity  $\lambda(\mathbf{X}_t, t)$ , for some  $\lambda : (D, \mathbb{R}_+) \mapsto \mathbb{R}^n$  and jump amplitude distribution  $\nu_t$  on  $\mathbb{R}^n$  and  $\nu_t$  only depends on time  $t$ .

**Definition A.1** *We say  $\mathbf{X}$  satisfying (A.1) is an **affine jump-diffusion process***

---

<sup>1</sup>For the definition, see [25].

<sup>2</sup>For more details, see [23].

if <sup>3</sup>

$$\left\{ \begin{array}{l} \mu(\mathbf{x}, t) = \mathbf{k}_0(t) + \mathbf{K}_1(t)\mathbf{x}, \\ \sigma(\mathbf{x}, t)\sigma(\mathbf{x}, t)' = \mathbf{H}_0(t) + \sum_{k=1}^n \mathbf{H}_1^{(k)}(t) \cdot \mathbf{x}_k, \\ \lambda(\mathbf{x}, t) = l_0(t) + \mathbf{l}_1(t) \cdot \mathbf{x}, \end{array} \right. \quad (\text{A.2})$$

where for each  $0 \leq t < \infty$ ,  $\mathbf{k}_0(t) \in \mathbb{R}^n$ ,  $\mathbf{K}_1(t) \in \mathbb{R}^{n \times n}$ ,  $\mathbf{H}_0(t) \in \mathbb{R}^{n \times n}$  and is symmetric,  $\mathbf{H}_1(t) \in \mathbb{R}^{n \times n \times n}$  and  $\mathbf{H}_1^{(k)}$  is symmetric<sup>4</sup>,  $l_0(t) \in \mathbb{R}$ ,  $\mathbf{l}_1(t) \in \mathbb{R}^n$ .

## A.2 Transform

Given an initial condition  $\mathbf{X}_0$ , the above ‘‘coefficient’’ tuple  $\theta = (\mathbf{k}_0, \mathbf{K}_1, \mathbf{H}_0, \mathbf{H}_1, l_0, \mathbf{l}_1, \nu_t)$  determines a transform  $\Phi^\theta : \mathbb{C}^n \times [0, \infty) \times [0, \infty) \times D \mapsto \mathbb{C}$  of  $\mathbf{X}_T$  conditional on  $\mathbf{X}_t$  defined by

$$\Phi^\theta(\mathbf{u}, t, T, \mathbf{X}_t) = E^\theta[e^{\mathbf{u} \cdot \mathbf{X}_T} | \mathbf{X}_t], \quad (\text{A.3})$$

where  $E^\theta$  denotes expectation under the distribution of  $\mathbf{X}_T$  determined by  $\theta$ .

In [23], it is proved that supposing  $\theta = (\mathbf{k}_0, \mathbf{K}_1, \mathbf{H}_0, \mathbf{H}_1, l_0, \mathbf{l}_1, \nu_t)$  is ‘‘well-behaved’’<sup>5</sup> at  $(\mathbf{u}, T)$ , then the transform  $\Phi^\theta$  of  $\mathbf{X}$  defined by (A.3) exists and is given by:

$$\Phi^\theta(\mathbf{u}, t, T, \mathbf{X}_t) = \exp(\mathbf{A}(\mathbf{u}, t, T) + \mathbf{B}(\mathbf{u}, t, T) \cdot \mathbf{X}_t), \quad (\text{A.4})$$

where  $\mathbf{A}(\cdot)$  and  $\mathbf{B}(\cdot)$  satisfy the following complex-valued Riccati equations,

$$\left\{ \begin{array}{l} \frac{\partial \mathbf{A}(\mathbf{u}, t, T)}{\partial t} = -\mathbf{k}_0(t)' \mathbf{B} - \frac{1}{2} \mathbf{B}' \mathbf{H}_0(t) \mathbf{B} + l_0(t)(\varphi(\mathbf{B}, t) - 1), \\ \frac{\partial \mathbf{B}(\mathbf{u}, t, T)}{\partial t} = -\mathbf{K}_1(t)' \mathbf{B} - \frac{1}{2} \mathbf{B}' \mathbf{H}_1(t) \mathbf{B} + \mathbf{l}_1(t)(\varphi(\mathbf{B}, t) - 1), \end{array} \right. \quad (\text{A.5})$$

<sup>3</sup>For column vectors  $\mathbf{a}$  and  $\mathbf{b}$ , the operation  $\mathbf{a} \cdot \mathbf{b}$  is the scalar product of  $\mathbf{a}$  and  $\mathbf{b}$ .

<sup>4</sup>Let  $\mathbf{H} \in \mathbb{R}^{n \times n \times n}$ ,  $\mathbf{H}^k$  is the matrix reduced by fixing the third index of  $\mathbf{H}$  to be  $k$ .

<sup>5</sup>For details, see [23].

with the boundary conditions

$$\begin{cases} \mathbf{A}(\mathbf{u}, T, T) = 0, \\ \mathbf{B}(\mathbf{u}, T, T) = \mathbf{u}, \end{cases} \quad (\text{A.6})$$

where  $\varphi(\mathbf{c}, t) = \int_{\mathbb{R}^n} \exp(\mathbf{c} \cdot \mathbf{z}) d\nu_t(\mathbf{z})$  whenever the integral is well defined. This “jump transform”  $\varphi$  determines the jump-size distribution.

### A.3 Extended Transform

We define the “extended” transform  $\Psi^\theta : \mathbb{R}^n \times \mathbb{C}^n \times [0, \infty) \times [0, \infty) \times D \mapsto \mathbb{C}$  of  $\mathbf{X}_T$  conditional on  $\mathbf{X}_t$  by

$$\Psi^\theta(\mathbf{v}, \mathbf{u}, t, T, \mathbf{X}_t) = E^\theta[(\mathbf{v} \cdot \mathbf{X}_T) e^{\mathbf{u} \cdot \mathbf{X}_T} | \mathbf{X}_t]. \quad (\text{A.7})$$

The extended transform  $\Psi^\theta$  can be computed by differentiation of the transform  $\Phi^\theta$ . As proved in [23], we have

$$\Psi^\theta(\mathbf{v}, \mathbf{u}, t, T, \mathbf{X}_t) = \Phi^\theta(\mathbf{u}, t, T, \mathbf{X}_t)(\mathbf{C}(t) + \mathbf{D}(t) \cdot \mathbf{X}_t), \quad (\text{A.8})$$

where  $\Phi^\theta$  is given by (A.4), and  $\mathbf{C}(\cdot)$  and  $\mathbf{D}(\cdot)$  satisfy the linear ordinary differential equations

$$\begin{cases} \frac{\partial \mathbf{C}(\cdot)}{\partial t} = -\mathbf{k}_0(t)' \mathbf{D} - \mathbf{B}' \mathbf{H}_0(t) \mathbf{D} - l_0(t) \nabla \varphi(\mathbf{B}) \mathbf{D} \\ \frac{\partial \mathbf{D}(\cdot)}{\partial t} = -\mathbf{K}_1(t)' \mathbf{D} - \mathbf{B}' \mathbf{H}_1(t) \mathbf{D} - l_1(t) \nabla \varphi(\mathbf{B}) \mathbf{D}, \end{cases} \quad (\text{A.9})$$

with the boundary conditions

$$\begin{cases} \mathbf{C}(T) = 0, \\ \mathbf{D}(T) = \mathbf{v}, \end{cases} \quad (\text{A.10})$$

where  $\nabla \varphi(c)$  is the gradient of  $\varphi(c)$  with respect to  $c \in \mathbb{C}^n$ .

## Appendix B

### Deduction of Futures Price Function for 2-factor Model by Affine Process Transform

The SDEs (4.1) describing the 2-factor model can be rewritten as

$$d \begin{bmatrix} S_t \\ L_t \end{bmatrix} = \begin{bmatrix} 0 \\ \mu\gamma \end{bmatrix} dt + \begin{bmatrix} -\alpha & \alpha \\ 0 & -\mu \end{bmatrix} \begin{bmatrix} S_t \\ L_t \end{bmatrix} dt + \begin{bmatrix} \sigma\sqrt{S_t} & 0 \\ 0 & \tau\sqrt{L_t} \end{bmatrix} \begin{bmatrix} dW_t^1 \\ dW_t^2 \end{bmatrix}. \quad (\text{B.1})$$

This is an affine-jump diffusion process, and in (A.2),  $\mathbf{X}_t = [S_t, L_t]'$ ,

$$\mathbf{k}_0 = \begin{bmatrix} 0 \\ \mu\gamma \end{bmatrix}, \mathbf{K}_1 = \begin{bmatrix} -\alpha & \alpha \\ 0 & -\mu \end{bmatrix}, \mathbf{H}_0 = \mathbf{0}_{2 \times 2}, \mathbf{H}_1 = \begin{bmatrix} (\sigma^2, 0)' & (0, 0)' \\ (0, 0)' & (0, \tau^2)' \end{bmatrix},$$

$$l_0 = 0, \mathbf{l}_1 = \mathbf{0}_{2 \times 1}, \nu = \mathbf{0}_{2 \times 1}.$$

By (A.7)-(A.10), we have the futures price given by

$$\begin{aligned} F^{\tilde{\theta}}(t, T, (S_t, L_t)) &= E^{\tilde{\theta}}[S_T | (S_t, L_t)] \\ &= \Psi^{\tilde{\theta}}((1, 0)', (0, 0)', t, T, (S_t, L_t)') \\ &= \Phi^{\tilde{\theta}}((0, 0)', t, T, (S_t, L_t)')(C(t) + D_1(t)S_t + D_2(t)L_t), \end{aligned} \quad (\text{B.2})$$

where  $\Phi^{\tilde{\theta}}((0, 0)', t, T, (S_t, L_t)')$  is computed by (A.4)-(A.6), and  $C(t)$  and  $D_i(t)$ 's satisfy the following equations:

$$\begin{cases} \partial_t C = -\tilde{\mu}\tilde{\gamma}D_2 \\ \partial_t D_1 = \tilde{\alpha}D_1 \\ \partial_t D_2 = -\tilde{\alpha}D_1 + \tilde{\mu}D_2, \end{cases} \quad (\text{B.3})$$

with the boundary conditions

$$\begin{cases} C(T) = 0 \\ D_1(T) = 1 \\ D_2(T) = 0. \end{cases} \quad (\text{B.4})$$

Since  $\mathbf{u} = (0, 0)'$ , it is easy to get  $(B_1, B_2) = (0, 0)$ . And, solving the above initial value problem (B.3) - (B.4), we have  $C = \frac{\tilde{\mu}\tilde{\gamma}}{\tilde{\alpha}-\tilde{\mu}}(e^{\tilde{\alpha}(t-T)} - 1) - \frac{\tilde{\alpha}\tilde{\gamma}}{\tilde{\alpha}-\tilde{\mu}}(e^{\tilde{\mu}(t-T)} - 1)$ ,  $D_1 = e^{\tilde{\alpha}(t-T)}$  and  $D_2 = \frac{\tilde{\alpha}}{\tilde{\alpha}-\tilde{\mu}}(e^{\tilde{\mu}(t-T)} - e^{\tilde{\alpha}(t-T)})$ . Hence, the futures price is given by

$$\begin{aligned} F^{\tilde{\theta}}(t, T, (S_t, L_t)) = & e^{\tilde{\alpha}(t-T)} S_t + \frac{\tilde{\alpha}}{\tilde{\alpha} - \tilde{\mu}} (e^{\tilde{\mu}(t-T)} - e^{\tilde{\alpha}(t-T)}) L_t \\ & + \frac{\tilde{\mu}\tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}} (e^{\tilde{\alpha}(t-T)} - 1) - \frac{\tilde{\alpha}\tilde{\gamma}}{\tilde{\alpha} - \tilde{\mu}} (e^{\tilde{\mu}(t-T)} - 1), \end{aligned} \quad (\text{B.5})$$

which is the same as (4.14) and (4.21).

# Appendix C

## The Computation of an Integral

For the integral

$$I = \int_{\mathbb{R}} \exp(iks - \frac{\lambda^2 s^2}{2}) ds, \quad (\text{C.1})$$

if we let  $x = \lambda s$ , then we have

$$\begin{aligned} I &= \frac{1}{\lambda} \int_{\mathbb{R}} \exp(i\frac{k}{\lambda}x - \frac{x^2}{2}) dx \\ &= \frac{1}{\lambda} \int_{\mathbb{R}} \exp(-\frac{(x - i\frac{k}{\lambda})^2}{2} - \frac{k^2}{2\lambda^2}) dx \\ &= \frac{1}{\lambda} \exp(-\frac{k^2}{2\lambda^2}) \int_{\mathbb{R}} \exp(-\frac{(x - i\frac{k}{\lambda})^2}{2}) dx \\ &= \frac{1}{\lambda} \exp(-\frac{k^2}{2\lambda^2}) \int_{\mathbb{R}} \exp(-\frac{x^2}{2}) dx \\ &= \frac{1}{\lambda} \exp(-\frac{k^2}{2\lambda^2}) \sqrt{2\pi}. \end{aligned} \quad (\text{C.2})$$

# Appendix D

## Continuous PDFs for 2-factor Models without Seasonality

### D.1 Model 2f-A

The SDEs (5.35) describing Model 2f-A can be rewritten as

$$d \begin{bmatrix} S_t \\ L_t \end{bmatrix} = \begin{bmatrix} 0 \\ \mu\gamma \end{bmatrix} dt + \begin{bmatrix} -\alpha & \alpha \\ 0 & -\mu \end{bmatrix} \begin{bmatrix} S_t \\ L_t \end{bmatrix} dt + \begin{bmatrix} \sigma & 0 \\ 0 & \tau \end{bmatrix} \begin{bmatrix} dW_t^1 \\ dW_t^2 \end{bmatrix}. \quad (\text{D.1})$$

Thus, in (A.2),  $\mathbf{X}_t = [S_t, L_t]'$ ,

$$\mathbf{k}_0 = \begin{bmatrix} 0 \\ \mu\gamma \end{bmatrix}, \mathbf{K}_1 = \begin{bmatrix} -\alpha & \alpha \\ 0 & -\mu \end{bmatrix}, \mathbf{H}_0 = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \tau^2 \end{bmatrix}, \mathbf{H}_1 = \mathbf{0}_{2 \times 2 \times 2},$$

$$l_0 = 0, \mathbf{l}_1 = \mathbf{0}_{2 \times 1}, \nu = \mathbf{0}_{2 \times 1}.$$

By (A.4) - (A.6) and (5.3), we have the characteristic function given by

$$\begin{aligned} \phi^\theta(s, t, T, (S_t, L_t)) &= \Phi^\theta(is, t, T, (S_t, L_t)) \\ &= \exp\left(A(t) + B_1(t)S_t + B_2(t)L_t\right), \end{aligned} \quad (\text{D.2})$$

where  $A(t)$  and  $B_i(t)$  satisfy the following equations,

$$\begin{cases} \partial_t A = -\mu\gamma B_2 - \frac{1}{2}(\sigma^2 B_1^2 + \tau^2 B_2^2) \\ \partial_t B_1 = \alpha B_1 \\ \partial_t B_2 = -\alpha B_1 + \mu B_2, \end{cases} \quad (\text{D.3})$$

with the boundary conditions

$$\begin{cases} A(T) = 0, \\ B_1(T) = is_1 \\ B_2(T) = is_2. \end{cases} \quad (\text{D.4})$$

Solving the initial value problem (D.3)-(D.4), we have

$$\begin{cases} A = (e^{\alpha(t-T)} - 1)P_1 + (e^{\mu(t-T)} - 1)P_2 \\ \quad + (e^{2\alpha(t-T)} - 1)P_3 + (e^{(\alpha+\mu)(t-T)} - 1)P_4 + (e^{2\mu(t-T)} - 1)P_5 \\ B_1 = e^{\alpha(t-T)}Q_1 \\ B_2 = e^{\alpha(t-T)}Q_2 + e^{\mu(t-T)}Q_3, \end{cases} \quad (\text{D.5})$$

where

$$\begin{aligned} Q_1 &= is_1, \\ Q_2 &= -\frac{is_1\alpha}{\alpha-\mu}, \\ Q_3 &= \frac{is_1\alpha}{\alpha-\mu} + is_2, \\ P_1 &= -\frac{\mu\gamma Q_2}{\alpha}, \\ P_2 &= -\gamma Q_3, \\ P_3 &= -\left(\frac{\sigma^2 Q_1^2}{4\alpha} + \frac{\tau^2 Q_2^2}{4\alpha}\right), \\ P_4 &= -\frac{Q_2 Q_3 \tau^2}{\alpha+\mu}, \\ P_5 &= -\frac{Q_3^2 \tau^2}{4\mu}. \end{aligned} \quad (\text{D.6})$$

Thus, from (5.4), the conditional PDF of  $[S_{t+1}, L_{t+1}]$  given  $[S_t, L_t]$  is

$$\begin{aligned} &f(\theta, [S_{t+1}, L_{t+1}]|[S_t, L_t]) \\ &= \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-is \cdot [S_{t+1}, L_{t+1}]} \exp\left(A(t, t+1, s) + B_1(t, t+1, s)S_t + B_2(t, t+1, s)L_t\right) ds. \end{aligned} \quad (\text{D.7})$$

In the above expression,  $s = [s_1, s_2]$ , and the integral is a double integral which is not easy to compute. Even to get some approximate result by numerical methods is hard to achieve efficiently.

## D.2 Model 2f-B

Following all the procedures similar to those in model 2f-A, we find the conditional PDF of  $[S_{t+1}, L_{t+1}]$  given  $[S_t, L_t]$  is also of the form (D.7), where

$$\begin{cases} \partial_t A = -\mu\gamma B_2 \\ \partial_t B_1 = \alpha B_1 - \frac{1}{2}\sigma^2 B_1^2 \\ \partial_t B_2 = -\alpha B_1 + \mu B_2 - \frac{1}{2}\tau^2 B_2^2, \end{cases} \quad (\text{D.8})$$

with the same boundary conditions as (D.4).

This time the equation system (D.8) is a Riccati equation system, which is not easy to solve. The best we can do is just compute  $A$ ,  $B_1$  and  $B_2$  numerically. Furthermore, in the computation of the conditional PDF (D.7), a numerical method again has to be applied in order to figure out a double integral.

## D.3 Model 2f-C

This model is not an affine jump-diffusion process, so we can not obtain its PDF by the same method as above.

# Bibliography

- [1] R. Gibson and E. Schwartz, *Stochastic Convenience Yield and the Pricing of Oil Contingent Claims*, Journal of Finance, vol.45, no 3, pp. 959-976, July 1990.
- [2] H. Bessembinder, J. F. Coughenour, M. M. Smoller and P. J. Seguin, *Mean Reversion in Equilibrium Asset Prices: Evidence from the Futures Term Structure*, Journal of Finance, vol.50, no 1, pp. 361-375, 1995.
- [3] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, 1999.
- [4] P. Samuelson, *Rational Theory of Warrant Pricing*, Industrial Management Review, J. 6, pp. 13-31, 1965.
- [5] F. Black and M. Scholes, *The Pricing of Options and Corporate Liabilities*, The Journal of Political Economy, 81(3), pp. 637-654, May-June 1973.
- [6] R. Merton, *Theory of Rational Option Pricing*, The Bell Journal of Economics and Management Science, 4(1), pp. 141-183, Spring 1973.
- [7] Y. K. Kwok, *Mathematical Models of Financial Derivatives*, 1998.
- [8] J. C. Hull, *Options, Futures, and Other Derivatives*, 4th edition, 1998.
- [9] G. E. Uhlenbeck and L. S. Ornstein, *On the Theory of Brownian Motion*, Physical Review vol.36, September 1930. (It was reprinted in N. Wax, eds.,

- Selected Papers on Noise and Stochastic Processes, Dover Pub., pp. 93-111, 1954.)
- [10] J. C. Cox, J. E. Ingersoll, and S. A. Ross, *A Theory of Term Structure of Interest Rates*, *Econometrica*, vol. 53, pp. 385-407, 1985.
- [11] J. Tykesson, *Some Aspects of Levy Processes in Finance*, Master's thesis, 2003.
- [12] E. S. Schwartz, *The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging*, *Journal of Finance*, vol.52, n. 3, pp. 923-973, July 1997. (It was with reference to an unpublished manuscript of Ross.)
- [13] A.K. Dixit and R. Pindyck, *Investment under Uncertainty*, 1994.
- [14] B. Tifenbach, *Numerical Methods for Modeling Energy Spot Prices*, Master's thesis, 2000.
- [15] D. Pilipovic, *Energy Risk: Valuing and Managing Energy Derivatives*, 1998.
- [16] J. E. Hilliard and J. Reis, *Valuation of Commodity Futures and Options under Stochastic Convenience Yields, Interest Rates, and Jump Diffusions in the Spot*, *Journal of Finance and quantitative analysis*, 33, pp. 66-86, 1998.
- [17] A. Lari-Lavassani, A. A. Sadeghi and T. Ware, *Modeling and Implementing Mean Reverting Price Processes in Energy Markets*, Electronic Publications of the International Energy Credit Association ([www.ieca.net](http://www.ieca.net)), 2001.
- [18] M. T. Barlow, Y. Gusev and M. Lai. *Parameter Estimation of Multifactor Models in Power Markets*, (preprint).

- [19] L. Clewlow and C. Strickland, *Energy Derivatives: Pricing and Risk Management*, Lacima Publications, 2000.
- [20] S. N. Neftci, *An Introduction to the Mathematics of Financial Derivatives*, 1996.
- [21] B. Øksendal, *Stochastic Differential Equations: an Introduction With Applications*, 5th edition, 1998.
- [22] V. Robert and T. Allen, *Introduction to Mathematical Statistics*, 5th Edition, Prentice Hall, Englewood Cliffs, New Jersey, 1995.
- [23] D. Duffie, J. Pan, and K. J Singleton, *Transform Analysis and Asset Pricing for Affine Jump-Diffusions*, *Econometrica*, 68, pp. 1343-1376, 2000.
- [24] R. Cont and T. Peter, *Financial Modelling with Jump Processes*, Chapman & Hall/CRC, 2004.
- [25] P. E. Protter, *Stochastic Integration and Differential Equations*, 2nd Edition, 2003.