UNIVERSITY OF CALGARY

Alpha-Stable, Normal Inverse Gaussian and Multi-Factor Models for Spot and Futures modelling in Natural Gas

by

Thomas Nedunthally

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

December, 2009

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 Masters Thesis entitled "Alpha-Stable, Normal Inverse Gaussian and Multi-Factor Models for Spot and Futures modelling in Natural Gas" submitted by Thomas Nedunthally in partial fulfillment of the requirements for the degree of Master of Science in Applied Mathematics.

Supervisor, Dr. Anatoliy Swishchuk
Department of Mathematics and Statistics
University of Calgary

External Examiner, Dr. Alfred Lehar
Haskayne School of Businesss
University of Calgary

Internal Examiner, Dr. Alexander Badescu
Department of Mathematics and Statistics
University of Calgary

Dec 2, 09
Date

Abstract

In this thesis, we aim to find suitable models for modelling the natural gas futures curve as well as modelling the gas spot prices. We begin introducing natural gas storages, which are key drivers of the futures, spot prices and other derivatives associated with the industry. This also gives us an intuition on the necessity for "convenience yield" in gas models.

We then begin reviewing the literature in natural gas and energy's, spot and futures modelling. A two factor spot model based on Xu [14] is introduced, as well as a new technique, where the underlying futures curve is stripped off at every stage by linear regression. This approach allows us to strictly model the underlying curve, the parameters are estimated by maximum likelihood estimation (MLE).

We also introduce Lévy processes to capture spot dynamics and try to model the futures curve based on the availability of the futures price via the Characteristic Function. We consider OU and CIR type processes with Lévy processes instead of Wiener Processes as the noise driving term. The processes used are alpha-stable Lévy processes as well as the normal inverse gaussian (NIG) Process.

Two different methods of calibration are used, one of them is an estimator based on finding the Empirical Characteristic Function (ECF) of the observations and making use of the availability of the characteristic function of Lévy processes to directly find the minimum of the parameters. We also use MLE to estimate the parameters, and we can see the advantages and disadvantages of the two estimation methods.

We then look at the implication of the parameters estimated on the futures curve and also go ahead to calculate the risk neutral measure of the futures curve.

Acknowledgements

Firstly, I would like to thank Dr. Anatoliy Swishchuk for his patience, encouragement and guidance into learning how to do research.

Also, I thank Dr. Anthony Ware for making good critical points in a few discussions related to this thesis.

I really appreciate the opportunity that Akbar Shahmoradi of Louis Dreyfus Energy Canada gave me as a MITACS intern in natural gas. I learned a lot from him about the fine details of the gas industry.

Then I would also like to thank my parents and sisters (Mathews, Ansy, Ann and Maria) for being loving in general. Also a special note to my housemates Yuval and Sumit for keeping me going through this phase of research by creating a noisy, active and social environment to work in.

Table of Contents

List of Tables

List of Figures

Introduction

The natural gas industry has gone through various phases of deregulations and reforms across many parts of the world. There are different kinds of natural gas storages all over, but the kind that concerns is a salt tavern storage, these are used due to a large number of injections and withdrawals that are possible into these storages every year based on contract terms. Storages are used as a hedging instrument against volatility and possible shortage of natural gas during high demand seasons of the year.

In Chapter 1, we discuss types of natural gas storages and its importance as a hedging instrument. We see its role as a demand-supply cushion for gas, and why modelling expected spot price processes for gas prices are necessary for pricing gas storages.

In Chapter 2, we discuss convenience yield models that are crucial when dealing with modelling commodity prices. A one factor model for natural gas prices would just involve a mean reverting Ornstein-Uhlenbeck(OU) process, a two factor model along the lines of Gibson-Schwartz [1990], which includes a convenience yield model as a second factor, is discussed. Then we look at Carmona [2003], where a stochastic market price of risk based on Runggaldier [2002] is suggested. Then we go on to discuss the natural gas specific spot models such as Pilipovic [34], Xu [12] and Hikspoor [32] as well as exclusive natural gas futures curve modelling that don't attempt to find a consistency between the spot and futures curve.

In Chapter 3, we discuss Levy processes and the Levy Khinchtine formula. Important Levy processes that are used in this thesis are defined, as well as the Levy Ito decompo-

sition. Levy based Ito calculus for discontinuous processes is discussed as well.

In Chapter 4, we go into the alpha stable levy process, that has the advantage of being able to model skewness and kurtosis. We discuss the simulation of stable random variables by the FFT as well as a kernel density estimator of random variables using the FFT. We combine the two to form a risk function based on their characteristic functions for parameter estimation. The natural gas spot price process has been historically deemed to be a mean reverting process and is cyclic in nature over a year due to seasonal variations in supply and demand. Since there are few jumps in natural gas spot prices, so we can comfortably use a stable process to model the spot price.

In Chapter 5, we introduce the normal inverse gaussian (NIG) Process. We first discuss the generalized hyperbolic(GH) class of processes and how NIG is a computationally feasible case of the GH process that still manages to retain properties of assett returns such as semi-heavy tails. We then discuss normal variance-mean mixtures and the simulation of NIG processes via Inverse Gaussian (IG) random variables. Finally, we also estimate the parameters of the NIG process via the empirical characteristic function (ECF).

In Chapter 6, we discuss the important criteria for modelling natural gas spot prices and how the expected value of the spot price process at different times in the future must be consistent with that of the futures curve. We introduce a new two factor model that assigns an affine structure for its seasonality term, the model being an extension to Pilipovic's two factor model. Another approach is that of modelling strictly the underlying curve, which turns out to be the most effective model. We also, introduce two other

models based on an alpha stable process and an NIG process, and take advantage of the fact that Ornstein-Uhlenbeck processes based on stable/NIG processes have an explicit solutions.

In Chapter 7, we discuss the calibration results and procedures for futures matching in natural gas based on the Lévy based models proposed in the previous chapter. We obtain the parameters of a seasonality term from a combination of spot and futures prices. This seasonality parameter is used in the Lévy based OU and CIR processes to match the futures price. We use maximum likelihood and the empirical characteristic function method to estimate the parameters of the models, and see the advantages and disadvantages of each method. We also calibrate the two factor model in the previous section, that strips the underlying futures curve using linear regression and we then use maximum likelihood to estimate the parameters.

Finally, we conclude the thesis, and discuss relevant future research.

Chapter 1

Natural gas storage

Natural gas is currently a 90 billion dollar industry [33], and is a clean and efficient source of energy. It is expected that 900 out of the next 1000 US power plants will use natural gas [33]. With an increasing usage every year and relatively limited resources for transportation and storage, natural gas is one of the most volatile markets in the world, due to supply and demand imbalances.

Seasonal demand for gas is traditionally linked to gas heating of houses, resulting in higher gas demand in winter than in the summer. Storages were owned by companies for balancing the variability in demand of their customers. After the deregulation in US and Europe, natural gas storage is unbundled from the sales and transportation services and is now a distinct seperately charged service. Gas demand in the US, Europe and Asia are increasing every year, the production flexibility is falling. Therefore there is a growing interest to invest in new gas storage facilities. International Energy Agency expect underground storage capacity to double in the next 30 years required an investment of 10-20 billion dollars. Development of new storage facilities, injection and withdrawal rates from compressors, changes pricing.

Natural gas is stored underground in large storage reservoirs. These storages are of three types aquifer, depleted gas reservoirs and salt caverns. Each of these storages have their advantages (and disadvantages) that can be characterized by the following four properties:

1. **Base and Working Gas Capacities**

Storages have to maintain a base level of gas always. The extraction costs increase

when the gas level is low (mostly due to the higher pressure of gas in the storage). Traders speculate on prices also based on the expected rise in gas prices due to low storage levels. Depleted gas and aquifer storages have high base gas requirements unlike salt cavern.

2. **Deliverability**

Refers to the rate of release of gas from reservoirs, deliverability is higher when the storage levels are high.

3. **Injection Capacity**

This is the rate of injecting gas into a storage, injection rates are highest when storage levels are low.

4. **Cycling**

This refers to the nature of operation of the facility. Single cycle storages serve 56 percent of the natural gas demand in eastern US, single cycle refers to the fact that injection gradually takes place through the summer months and and withdrawal during the winter months. Depleted gas reservoirs fall into this category of single cycle storages. Salt Cavern and Aquifer storages have the advantage that they can have about 4-5 cycles a year, and these are important to meet peak gas demand situations.

In order to find the optimal price of a gas storage, there are many approaches in literature to handle the problem. Volatility and mean-reversion in gas spot markets are much larger than in gas forward markets, most value can be created in the spot market. So we consider a spot trading strategy. The basic approach to storage valuation is to calculate the optimal position given the available forward curve and take this position. This "intrinsic value" approach captures the predictable seasonal pattern in gas prices. Additional value (extrinsic value) can be created by reacting to price fluctuations, i.e. based on gas volatility.

To value a storage contract, we follow a spot-based strategy because it captures most of the contract's flexibility. In literature, the first natural gas thesis was that of Bringedal [17] who took a stochastic dual dynamic programming approach to solving the storage formulation. Thompson [1] formulated the storage problem as a partial integro differential equation (PIDE) and finally narrowed down the problem to a control theory problem via the Bellman equation, but the problem with this approach is that the spot price process is taken to be a simple mean reverting OU process. If the spot price model were more complicated, then the Taylor series expansion would yield more parameters and terms and hardens the storage equation. In 2007, Chen and Forsyth [8] used a seasonality term into the one factor OU model and provided an improved numerical scheme for solving the PIDE using a semi-Lagrangian spproach. In [15] they also used a simple one factor regime switching model that switches between an OU process and a Geometric Brownian Motion (GBM) with an upward drift. The upward drift is to intuitively occur during winter months, and during the summer it would follow an OU process. They discuss the implications of this model to solving the storage equation.

In 2008, Boogert and De Jong [4] formulated the problem in a least squares Monte Carlo Approach, so in the style of American Options they price the contract starting backwards in time and solving a (ordinary least squares) OLS Regression problem along each time step. This allows us to have a complicated spot process that is not coupled with the storage equation, which is ideal for the goal of this thesis to incorporate Lévy-based spot price processes.

Chapter 2

Spot models for natural gas and commodities

In this chapter, we shall discuss different spot price models used in commodity pricing. The typical feature of many commodities is that of mean reversion and this is captured by an Ornstein-Uhlenbeck (OU) process . The commonly used OU process is a single factor model with a Wiener process as the driving noise term. Schwartz [13] showed that this is insufficient when modelling the forward curve, as there is a cost of carry that affects the drift of the commodity spot prices. This "drift adjustment" was taken to be stochastic and is the first in the class of "convenience yield" models.

Schwartz also considered a third stochastic factor, which was the interest rate, but this did not yield any qualitative advantage over the two factor model. So in commodity modelling literature, stochastic interest rates are rarely ever considered. Philopivic [12] also proposed a two factor model, which has a long run stochastic mean. Now in the case of natural gas, there is seasonality exhibited in the price dynamics. This is largely due to natural gas being a primary source of heating homes and businesses. Heat usage increases in the winter and goes down in the summer, so due to the market forces of supply and demand, the price of natural gas has a general upward price movement in the winter and downward movement in the summer. This seasonality is also seen in the price of natural gas forwards/futures.

So when modelling natural gas prices, a general commodity pricing model is not adequate. Xu [14] proposed a model that is a generalization of Philipovic's model to include seasonality via a "positioning" term, which is a sum of two sinusoids with different periods, whose parameters are obtained from the forward curve. Chen and Forsyth [2007],

used a one factor regime switching model to simulate the natural gas price that supposedly emulates the two factor convenience yield model from Gibson-Schwartz [1990]. The PDE method of pricing by Chen and Forsyth [26] is much more efficient in computing the value of a storage, and Chen and Forsyth usually look for models that are more analytically tractible. Hikspoor and Jaimungal [2007] proposed a two-factor model with stochastic long run mean reversion and a seasonal component g_t.

Now there is an important test that spot models must withstand, i.e., how well does the model fit the futures curve. Since the futures price is equal to the discounted expected value of the spot price at its expiry, the proposed spot model must be able to match the futures curve when taking its expectation. So in literature, its important that the spot price process chosen has an explicit form for its expection, to determine the futures price.

We shall now discuss different spot price models and its corresponding futures price, based on available literature.

2.1 A one-factor model - Ornstein-Uhlenbeck process

The commonly used process to model natural gas behaviour is the mean reverting Ornstein-Uhlenbeck (OU) process. This is the most popular one factor model in natural gas spot simulation. The OU process is defined by

$$dS_t = \theta(\mu - S_t)dt + \sigma dW_t. \tag{2.1}$$

where θ is the speed of mean reversion, μ is the value that the spot price reverts to, σ is the diffusion term and W_t is a Wiener process. The expectation, variance and covariance are

$$E(S_t) = S_0 e^{-\theta t} + \mu(1 - e^{-\theta t}) \tag{2.2}$$

$$\tag{2.3}$$

$$
\begin{aligned}
Var(S_t) &= \frac{\sigma^2}{2\theta} \\
Cov(S_s, S_t) &= E[(S_s - E[S_s])(S_t - E[S_t])] \\
&= \frac{\sigma^2}{2\theta\, e^{-\theta(s+t)}(e^{2\theta(s\wedge t)} - 1)}
\end{aligned}
$$

This process is used as the standard spot price model for pricing the natural gas storage in Boogert and Jong [23], Chen and Forsyth [26], Thompson, Davison et. al. [1] and Bringedal [17]. The disadvantage of this approach is that the spot price evolution can't be accurately accounted for. But the advantage is the ease of calibration, and the simple form for the futures price, which follows from equation 2.2. The futures price is given by $F(t, T, S_t) = E[S_T|S_t]$.

$$
F(t, T, S_t) = S_t e^{-\theta(T-t)} + \mu(1 - e^{-\theta(T-t)}) \tag{2.4}
$$

where $F(t, T, S_t)$ is the futures price at time t, for a contract expiring at time T given that the spot price is S_t.

Now since it is important to be able to valuate the most accurate price of a natural gas storage, we need more accurate spot models that capture as many of the properties as we can of spot-future dynamics. Due to the least squares Monte Carlo type approach in De Jong [23] for storage evaluation, where the spot price can be isolated from the pricing methodology, we investigate more accurate models.

With respect to storage, another model that is relatively simple but can intuitively simulate the expected spot price process with respect to the futures price is that of Li [16]. It is appropriate for industry practioners, who have to take positions every day with respect to injection and withdrawal of gas from the storage. He takes the spot price to take the following process.

$$S_T = \begin{cases} S_0 exp(-\frac{1}{2}\sigma^2 T + \sigma\sqrt{T}\epsilon_T) & \text{for the valuation month} \\ F_{0,i} exp(-\frac{1}{2}\sigma^2 T + \sigma\sqrt{T}\epsilon_T & \text{for the } i^{th} \text{ month contract} \end{cases} \tag{2.5}$$

Here S_T is the spot price at time T in the future, S_0 is the spot price on the valuation date and σ is the spot price volatility. $F\{0, i\}$ is the price of the futures contract as of today based on a expiry date i. It makes sense that for the first month, the spot price simulated follows todays spot price with a month's worth of simulated noise. Since the futures price is the expected value of the spot price, for every subsequent month he sets the spot price price process for the month, to begin with the futures price expiring that month.

This approach simplifies the expected spot dynamics, takes into account the forward curve and is computationally less expensive, but our goal is to have a spot model that also captures futures dynamics. So we shall move on to more complex models that hope to capture both these features.

2.2 Multi-factor models in commodity pricing

We start off with a variety of two-factor models with their third-factor extensions. The most popular two-factor model is that of Gibson-Schwartz [10]. It was the first in the class of "convenience yield" models. Two factor models attempt to interpret market movements as having two driving noise terms, or two random sources of volatility.

In Carmona [7], a stochastic market price of risk term is introduced to fit the implied convenience yield for different maturities. Before we dive into the Schwartz-Carmona models we shall now discuss the basics of convenience yield and futures prices in markets.

2.2.1 What is convenience yield?

The standard pricing for futures in markets is that the spot price would equal the discounted forward price, i.e.

$$F(t,T) = S_t E[exp(\int_t^T r_s ds)] \tag{2.6}$$

where T is the time of exercise, r is the riskless interest rate S_t is the spot price and $F(t,T)$ is the price of the forward at time t with excercise at time T. In commodities and energy linked assets, the futures price does not work out to be of the form above but has an unobservable quantity called a convenience yield that factors into the above model. In energy linked assets there are storage costs of energy along with the other industrial management costs that influence the assets true value at time T from the purchase of the contract. In other words, on one hand the holder of the contract has the option of consumption flexibility and has no risk in the event of commodity shortage. But the decision to postpone consumption comes with an associated flow of costs implied by the storage expenses. This net flow of services is the convenience yield and is represented by δ where

$$\delta = \text{Benefit of direct access - cost of carry}$$

Pricing is commodity market models is based on the assumption that the spot price exists, which wouldn't be true in the case of assets like electricity which is impossible to store. The forward contract price that includes the convenience yield is found by a no arbitrage argument and the forward price is taken to be

$$F(t,T) = S_t E_Q[exp(\int_t^T (r_s - \delta_s) ds] \tag{2.7}$$

where δ is the convenience yield. Q is the risk neutral measure, so this implies that δ_t can be inferred as a drift correction term in the spot price process. We shall now discuss

the Gibson-Schwartz [1990] two factor model, which is based on the above intuition.

2.2.2 Gibson-Schwartz model

The first spot convenience yield model was introduced by Gibson and Schwartz in 1990. The spot price has the convenience yield δ_t added to the drift and is assumed to be a mean reverting process that drives the geometric brownian motion commodity spot price S_t.

Let (Ω, F, P) be a probability space under a filtration $\{F_t\}_{t\geq 0}$. According to the Gibson-Schwartz model, under the risk-neutral measure Q,

$$dS_t = (r_t - \delta_t)S_t dt + \sigma S_t dW_t^1, \tag{2.8}$$

$$d\delta_t = \kappa(\theta - \delta_t)dt + \gamma dW_t^2, \tag{2.9}$$

where W_1 and W_2 are correlated Wiener processes with $dW_1 dW_2 = \rho dt$.

The spot process is usually taken to be a mean reverting asset in many energy commodities, but in 1990 Gibson [10] argued that the convenience yield influences the spot price process and induces mean reversion to it. Unlike interest rate models, it makes sense that convenience yields can take positive or negative values so the model proposed seems logical.

Schwartz in 1997 compared the one, two and three factor spot models in fitting forward curves. The one factor model just has the mean reverting Ornstein-Uhlenbeck spot price process, the two factor factor model is the above model and the three factor model has stochastic interest rates. His paper, showed that there was no qualitative improvement in assuming a stochastic interest rate, so we shall not include stochastic interest rates in our study.

In Schwartz [1997], it is shown that the forward price for the above spot pricess is

$$F(t,T,S_t) = S_t e^{\int_t^T r_s ds} e^{B(t,T)\delta_t + A(t,T)} \tag{2.10}$$

where

$$\begin{aligned} B(t,T) &= \frac{e^{\kappa T}-1}{\kappa}, \\ A(t,T) &= \frac{\kappa\theta+\rho\sigma\gamma}{\kappa^2}(1-e^{-\kappa(T-t)}-\kappa(T-t), \\ &+ \frac{\gamma^2}{\kappa^3}(2\kappa(T-t)-3+4e^{-\kappa(T-t)}-e^{-2\kappa(T-t)}). \end{aligned}$$

We can see the futures price takes this form, since the spot model in equation (2.4) is of an affine form. Details of affine processes are given in the appendix.

Runggaldier [31] in 2003 suggest that another OU process is used for the market price of risk, it is suggested since it does not affect the "affine structure" of the spot process in equation (2.4). The market price of risk is the risk neutral measure of the spot price process, i.e.

$$dW_t^1 = d\tilde{W}_t^1 - \lambda_t dt. \tag{2.11}$$

And since λ_t follows an OU process, it has the following equation (note that the market price of risk can take positive or negative values):

$$d\lambda_t = \kappa_\lambda(\bar{\lambda}-\lambda_t)dt + \sigma_\lambda dW_t^3 \tag{2.12}$$

In 2004, Carmona [7] used this idea to enlarge the observation equation of the spot price process, for better calibration of the futures curve. The Wiener process W_t^1 of the the spot price process is substituted by equation (2.5), and the extra stochastic factor λ_t is added from equation (2.6).

$$
\begin{aligned}
dS_t &= (r_t - \delta_t - \sigma\lambda_t)S_t dt + \sigma S_t d\tilde{W}_t^1, \\
d\delta_t &= \kappa(\theta - \delta_t)dt + \gamma dW_t^2 \\
d\lambda_t &= \kappa_\lambda(\bar{\lambda} - \lambda_t)dt + \sigma_\lambda dW_t^3 \\
dF(t,T_i) &= (r_t + \sigma\lambda_t + \rho\gamma\frac{e^{-\kappa T_i} - 1}{\kappa}\lambda_t)F(t,T_i)dt \\
&+ \sigma F(t,T_i)d\tilde{W}_t^1 + \gamma F(t,T_i)\frac{e^{-\kappa T_i} - 1}{\kappa}dW_t^2 + \alpha dW_t^{F^i}.
\end{aligned}
$$

The above equation provides another approach to modelling the forward curve, the implications are that the changes in the forward curve as a whole, can be fitted to the above equation for $dF(t,T_i)$. The futures price has an explicit form due to the affine nature of the three factor model of the spot price process described above.

The above discussed models are made for general commodities and are not specific to natural gas or energy. For energy commodities, there are the seasonal forces of supply and demand that create complicated characteristics for the futures curve and the spot price process. We desire there to be a strong seasonal component since it characterizes so much of the movement of natural gas. We shall now discuss models where a seasonality term was introduced. The estimation of parameters of the seasonality term is done through the forward curve usually.

To understand why seperate models are needed for energy, Figure 2.1 is an example of the average spot price of natural gas on the same day over the last eight years. We can see a general trend of high gas prices in the Winter months and lower gas prices in the summer months.

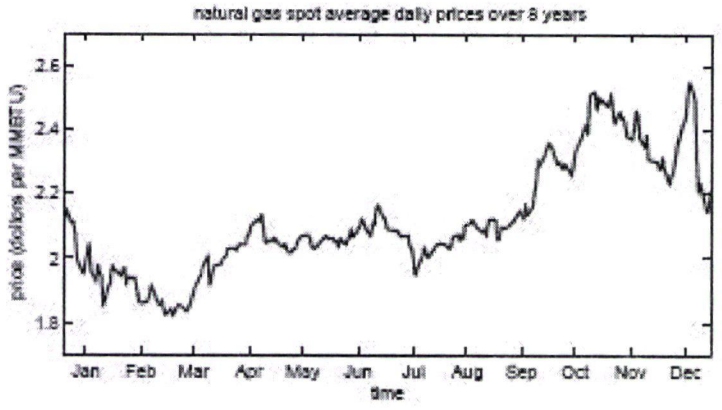

Figure 2.1: Average daily spot prices

2.2.3 The Pilipovic two-factor model for energy

Dragana Pilipovic in 1997 wrote the first book in Energy Risk [12], where she presents a two-factor model for energy taking into consideration the complex dynamics of the futures curve with respect to the spot price. It is seen that from the futures curve, there is an implied long run mean. In Pilipovic [12], she presents the following model where a moving long run mean is allowed, and the spot price being risk adjusted.

$$d\tilde{S}_t = \alpha(L_t - S_t)dt + \sigma S_t d\tilde{W}_t^1 \tag{2.13}$$

$$dL_t = \mu L_t dt + \gamma L_t dW_t^2, \tag{2.14}$$

where $\tilde{W}_t^1$ and W_t^2 are uncorrelated standard brownian motions, and L_t is the long run mean. If W_t^1 is Wiener process driving the above spot process then $\tilde{W}_t^1 = W_t^1 - \lambda dt$.

She says that the futures curve should be modelled by first stripping off the seasonality, i.e.,

$F(t,T) = F^{UND}(t,T)$ + seasonality contribution.

F^{UND} is what should be modelled, if there is seasonality in the futures curve, as in natural gas. The spot process being of an exponential affine form, also has an explicit

form for its futures price that Pilipovic [12] shows as:

$$F(t,T) = (S_t - L_t)e^{-(\alpha+\lambda\gamma)(T-t)} + L_t e^{(\mu-\lambda\gamma)(T-t)} \quad (2.15)$$

2.2.4 Xu's generalization of Pilipovic's model

Xu [14] added a seasonality term f(t) to Pilipovic's model, and studied a more general version of equation 2.7. He exclusively studied the natural gas spot-future curve. The following model was proposed by Xu:

$$\begin{aligned}
S_t &= f(t) + X_t \\
dX_t &= \alpha(L_t - X_t)dt + \sigma(t)X_t dW_t^1 \\
dL_t &= \mu(\gamma - L_t)dt + \tau L_t dW_t^2 \\
\sigma(t) &= exp(c + \textstyle\sum_{j=1}^{2} \lambda_j cos2\pi jt + \omega_j sin2\pi jt) \\
f(t) &= bt + \textstyle\sum_{j=1}^{2} (\beta_j cos2\pi jt + \eta_j sin2\pi jt).
\end{aligned}$$

As we can see above, for $\gamma = 0$ and if $\sigma(t)$ is a constant, it would be equivalent to Pilipovic's model. He considered one factor models in the above form where L_t is constant, and studied models with and without seasonality. The models with seasonality terms in it, performed the best. His calibration of the seasonality parameters is particularly interesting and is implemented in this thesis as well in Chapter 6.

2.2.5 Hikspoor and Jaimungal's model

In 2007, Hikspoor et. al. [32] studied a class of models that looked as follows, with long-run mean and a seasonal component g_t.

$$\begin{aligned} S_t &= exp(g_t + X_t) \\ dX_t &= \beta(Y_t - X_t)dt + \sigma_X dW_t \\ dY_t &= \alpha(\phi - Y_t)dt + \sigma_Y dZ_t \\ d[W,Z]_t &= \rho dt. \end{aligned}$$

X_t is a stochastic process which has to be estimated by the observed equation $X_t = log(S_t) - g_t$. Both of these are OU processes and they both have the advantage of being able to estimate the conditional probabilities.

They also considered a three factor model with additional stochastic volatility, with the spot price process given by:

$$\begin{aligned} S_t &= exp(g_t + X_t) \\ dX_t &= \beta(Y_t - X_t)dt + \sigma_X(Z_t)dW_t^1 \\ dY_t &= \alpha(\phi - Y_t)dt + \sigma_Y dW_t^2 \\ dZ_t &= \eta(\mu - Z_t)dt + \sigma_Z dW_t^3 \end{aligned}$$

where $d[W^1, W^2]_t = \rho_{xy}dt,\ d[W^1, W^3]_t = \rho_{xz}dt,\ d[W^2, W^3]_t = \rho_{yz}dt.$

They extend their two and three factor models to include jumps, such that $S_t = exp(g_t + X_t + J_t)$. The jump component J_t satisfies $dJ_t = -\kappa J_{t-}dt + dQ_t$, where Q_t is a compound poisson process.

As can be seen above, rather than including the jump term in X_t, they include it in directly to the spot price dynamics of S_t. The advantage is that when modelling commodity prices such as electricity, the typical behaviour is that of spikes in prices and typically returning to its regular level. So rather than a jump that causes the entire

price evolution to alter itself, a jump is randomly added and the price returns back to its original state.

2.2.6 Eydeland and Wolyniec's model

In Eydeland [35], he proposes a model that attempts to model the entire forward curve. The forward equation is determined by the Schwartz model described previously. It is an HJM (Heath-Jarrow-Morton) type model, where we can think of the underlying forward curve stripped of seasonality to follow an interest rate type model. The forward process generally has the following form:

$$dF(t,T) = \mu(t,T,F(t,T))dt + \sum_j \sigma_j(t,T,F(t,T))dW_t^j \tag{2.16}$$

This is a multifactor equation, and W_t^j are Wiener processes whose correlations can be chosen. For commodities, the forward curve model that could be simplified to,

$$dF(t,T) = F(t,T)\sum_j \sigma_j(t,T)dW_t^j \tag{2.17}$$

The above equations propose a very different approach, here they are not worried about the actual spot price process. They try to capture the dynamics of the futures curve which actually has a very erratic behaviour in commodities due to its dependence on long and short term supply and demand. In this thesis, we also of try to model the gas curve via such an approach, but using simple OLS regression instead.

The models described in this chapter, show of the initial evolution of modelling the futures curve and a corresponding model for the spot price from general commodities to energies to natural gas. Now we shall discuss in detail the modelling of futures prices in natural gas using Lévy processes, although it can be applied to other commodities as well.

Lévy-based gas price models have the advantange of being one factor models with complicated noise driving terms. Since the characteristic functions are available, they can be estimated by the techniques developed in this paper. The densities of the spot price processes fit very nicely when using such classes of one factor models. Before we talk about futures matching, we need to have a little background discussion on Lévy proccesses and Itô's formula for processes of this type.

Chapter 3

Lévy processes

Before we begin discussion of spot models in commodities and natural gas, we shall get acquainted with Lévy processes and the alpha-stable Lévy motion, in particular.

3.1 Stochastic processes

A stochastic process X is a collection of random variables indexed by time where X = $(X_t)_{t\in T}$. That is for each $t \in T$, X_t is a random variable. The time parameter t can be either discrete or continuous, depending on the index set T. If T is a countable set then we have a discrete time stochastic process, and if T is a continuous non-countable set then we can obtain a continous-time stochastic process. Any realization of X is called a sample path. In case of a continuous time stochastic process we often write the index set T as an interval $[0, T]$.

Definition 1. ***(Cádlág function)*** *A function $f : [0,T] \mapsto (R)^d$ is said to be cádlág if it is right-continuous with left limits, i.e. for each $t \in [0,T]$ the limits $f(t-) = lim_{s\mapsto t, s<t} f(s)$ and $f(t+) = lim_{s\mapsto t, s>t} f(s)$ exist and $f(t) = f(t+)$.*

The purpose of this definition is to be able to define jump discontinuities in stock price movements. If t is a discontinuity point we denote $\Delta f(t) = f(t) - f(t-)$, i.e., the jump of f at t.

Definition 2. ***Lévy process*** *A cádlág stochastic process $(X_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, P)$ with values in $(R)^d$ such that $X_0 = 0$ is called a Lévy process if it possesses the following properties:*

1. Independent increments: for every increasing sequence of times $t_0 ... t_n$, the random variables $X_{t_0}, X_{t_1} - X_{t_0}, ... X_{t_n} - X_{t_{n-1}}$ are independant.

2. Stationary increments: the distribution law of $X_{t+h} - X_t$ does not depend on t.

3. Stochastic continuity: $\forall \epsilon > 0, lim_{h \to 0} P(|X_{t+h} - X_t| \geq \epsilon) = 0$.

Now every Lévy process has a general form, in order to introduce and understand the significance of this form, we first discuss the concepts of convolution of measures and infinite divisibility.

3.2 Convolution of measures

Let $M_1((R)^d)$ be the set of all Borel probability measures on $\mathbb{R}^d$. The convolution of two probability measures is defined to be:

$$(\mu_1 * \mu_2)(A) = \int_{\mathbb{R}^d} \mu_1(A - x)\mu_2(dx) \tag{3.1}$$

for each $\mu_i \in M_1(\mathbb{R}^d), i = 1, 2$, and each $A \in B(\mathbb{R}^d)$, where $A - x = y - x, y \in A$.

We note that in the one dimensional case where $d = 1$ above, this just means that if the distribution law of random variables, X and Y are μ_X and μ_Y then the distribution law of the sum of the random variables, i.e. $X + Y$ is given by $\mu_X * \mu_Y$.

Now we define $\mu^{*^n} = \mu * ... * \mu$ (n times) and μ is said to have a *convolution nth root*, if there exists a measure $\mu^{1/n} \in M_1(\mathbb{R}^d)$ for which $(\mu^{1/n})^{*^n} = \mu$.

Definition 3. ***(Infinite divisibility)*** *Let X be a random variable in $\mathbb{R}^d$ with law μ_X. X is said to be infinitely divisible, if for all $n \in (N)$, there exist i.i.d. random variables $Y_1, ..., Y_n$ such that*

$$X \stackrel{d}{=} Y_1 + Y_2 + ... + Y_n. \tag{3.2}$$

The characteristic function of a random variable X is denoted by $\phi_X(u) = E(e^{i(u,X)})$ where $u \in \mathbb{R}^d$. If μ_X is the law of X then $\phi_X(u) = \int_{\mathbb{R}^d} e^{i(u,y)}\mu_X(dy)$.

Proposition 1. *The following statements are equivalent:*

1. *X is infinitely divisible.*
2. *μ_X has a convolution nth root that is itself the law of a random variable.*
3. *ϕ_X has an nth root that is itself the characteristic function of a random variable.*

3.3 The Lévy-Khintchine representation

We shall present the magnificent formula established by Paul Lévy and A. Khintchine which allows us to characterize an infinitely divisible random variable and a Lévy process through its characteristic function. In order to describe this formula, we first define what a *Lévy measure* is.

Definition 4. ***(Lévy measure)*** *Let ν be a Borel measure defined on $\mathbb{R}^d - 0 = x \in \mathbb{R}^d, x \neq 0$. It is called a Lévy measure if*

$$\int_{\mathbb{R}^d - 0} (|y|^2 \wedge 1)\nu(dy) < \infty. \tag{3.3}$$

Theorem 1. *Let $(X_t)_{t \geq 0}$ be a Lévy process on $\mathbb{R}^d$ with a characteristic triplet (b, A, ν), then*

$$\phi_{X_t}(u) = E[e^{i(u,X_t)}] = e^{t\psi(u)}, u \in \mathbb{R}^d \tag{3.4}$$

where

$$\psi(u) = i(b,u) - \frac{1}{2}(u, Au) + \int_{\mathbb{R}^d - 0} [e^{i(u,y)} - 1 - i(u,y)1_{B_1(0)}]\nu(dy). \tag{3.5}$$

with $B_1(0)$, the unit sphere in $\mathbb{R}^d$ centered at 0, or $B_1(0) = |x| < 1, x \in \mathbb{R}^d$.

From the above definition follows the characteristic function of an infinitely divisible random variable.

Theorem 2. *Let X be an infinitely divisible random variable with distribution $\mu \in M_1(\mathbb{R}^d)$. It characteristic function is represented as*

$$\phi_X(u) = e^{\psi(u)}, u \in \mathbb{R}^d \tag{3.6}$$

where $\psi(u)$ is from equation (3.5)

Now that we have established the Lévy-Khintchine formula, it is easy to represent most of the most of the stochastic processes used in financial math in terms of the triplet (b, A, ν). In this thesis, we shall be consistent with the definitions of processes in the upcoming section.

3.4 Popular Lévy Processes

If a Lévy process is represented by the characteristic triplet (b, A, ν), then the following are cases of the triplet.

Standard Brownian Motion

A standard Brownian motion is $\mathbb{R}^d$ is a Lévy process $B = (B_t, t \geq 0)$ with the triplet $(0, I, 0)$, where I is the identity matrix where

1. $B_t \sim N(0, tI)$ for each $t \geq 0$
2. B has continuous sample paths.

This means that if B is a standard Brownian motion then its characteristic function is :

$$\phi_{B_t}(u) = exp(-\frac{1}{2}t|u|^2), u \in \mathbb{R}^d, t \geq 0. \tag{3.7}$$

Brownian motion with drift

If A is a positive symmetric $d \times d$ matrix and σ is a $d \times m$ matrix such that $\sigma\sigma^T = A$. If $b \in \mathbb{R}^d$ and B is a brownian motion in $\mathbb{R}^m$, then a process $C = (C(t), t \geq 0)$ in $\mathbb{R}^d$

defined by

$$C(t) = bt + \sigma B(t) \tag{3.8}$$

is a Lévy process with the triplet $(b, A, 0)$, i.e.

$$\psi_C(u) = i(b, u) - \frac{1}{2}(u, Au), u \in \mathbb{R}^d. \tag{3.9}$$

which can also be recognized as $C_t \sim N(tb, tA)$. C is also called a *Brownian motion with drift.*

Poisson process

A Poisson process is a Lévy process N, of intensity $\lambda > 0$ where $N(t) \sim \pi(\lambda t)$, so

$$P(N(t) = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}. \tag{3.10}$$

for n being positive integers.

Compound Poisson process

Let $(Z(n), n \in \mathbb{N})$ be a sequence of i.i.d. random variables in $\mathbb{R}^d$ with a distribution μ_Z. Also, let N be a Poisson process of intensity λ independent of $Z(n)$, with the process X called a Compound Poisson process defined as:

$$X(t) = Z(1) + ... + Z(N(t)) \tag{3.11}$$

for $t \geq 0$, and each $Y(t) \sim \pi(\lambda t, \mu_Z)$.

Proposition 2. *The compound Poisson process X is a Lévy process with Lévy characteristic* $\psi_X(u) = \int_{\mathbb{R}^d}(e^{i(u,y)} - 1)\lambda\mu_Z(dy)$.

The proof of the above is straightforward can be found in Applebaum [19]. There are many more examples of Lévy processes, our focus is on the alpha-stable Lévy process, which we shall discuss, simulate and estimate in Chapter 3.

Note that the sum of a Brownian motion with drift and a compound Poisson process is a case of Merton's jump diffusion model.

3.5 Martingales

If X is a stochastic process defined on a probability space with filtration $(F_t, t \geq 0)$ and an integrability precondition $E(|X(t)|) < \infty$ for all $t \geq 0$ then we say that X is a martingale if $E(X(t)|F_s) = X(s)$ a.s.

Martingales are important since the above definition implies that the expected value of a process X filtered at a given point of time, is the value of the process at that time. In terms of spot prices of stocks, we usually need to adjust the spot price process by a constant drift to ensure its "risk neutrality" or "martingaleness".

Proposition 3. *If X is a Lévy process with symbol ψ, then for $u \in \mathbb{R}^d$, a process $M_u = (M_u(t), t \geq 0)$ defined by*

$$M_u(t) = exp[i(u, X(t)) - t\psi(u)] \tag{3.12}$$

is a martingale with respect to the natural filtration F^X.

Theorem 3. *Let $N(t, A) = \sharp(0 \leq s \leq t; \Delta X(s) \in A) = \sum_{0 \leq s \leq t} 1_A(\Delta X(s))$. If $A \in B(\mathbb{R}^d - 0)$ then we say that A is bounded below if $0 \notin \bar{A}$. We also write $\mu(A) = E(N(1, A))$ and call it the intensity measure.*

If A is bounded below, then.$(N(t, A), t \geq 0)$ is a Poisson process with intensity $\mu(A)$

Definition 5. *For each $t \geq 0$ and A bounded below, the compensated Poisson random measure is defined by*

$$\widetilde{N}(t, A) = N(t, A) - t\mu(A). \tag{3.13}$$

We note that $\widetilde{N}(t, A)$ is a martingale.

Definition 6. *If f is a Borel measurable function from $\mathbb{R}^d$ to $\mathbb{R}^d$ and let A be bounded below then we define the Poisson integral of f as*

$$\int_A f(x)N(t, dx) = \sum 0 \leq u \leq t f(\Delta X(u)) 1_A(\Delta X(u)). \tag{3.14}$$

We now have all the tools we require to define the Lévy-Itô decomposition as in Applebaum [19].

Theorem 4. ***(The Lévy-Itô decomposition)*** *If X is a Lévy process, then there exists $b \in \mathbb{R}^d$, a Brownian motion B_A with covariance matrix A and a Poisson random measure N on $\mathbb{R}^+ \times (\mathbb{R}^d - 0)$ such that, for each $t \geq 0$,*

$$X(t) = bt + B_A(t) + \int_{|x|<1} x\widetilde{N}(t, dx) + \int_{|x|\geq 1} xN(t, dx). \tag{3.15}$$

3.6 Stochastic Integration and Itô's Formula

For a general Lévy-type stochastic integral, let Y be a process whose stochastic different equation can be written as:

$$dY(t) = G(t)dt + F(t)dB(t) + \int_{|x|<1} H(t,x)\widetilde{N}(dt, dx) + \int_{|x|\geq 1} K(t,x)N(dt, dx) \tag{3.16}$$

Let $Y(t) = Y(0) + Y_c(t) + Y_d(t)$ where Y_c is the continous part of the process and Y_d is the discontinuous part, defined by

$dY_c(t) = G(t)dt + F(t)dB(t)$, and

$dY_d(t) = \int_{|x|<1} H(t,x)\widetilde{N}(dt, dx) + \int_{|x|\geq 1} K(t,x)N(dt, dx)$.

Theorem 5. *If Y is a Lévy-type stochastic integral of the form (3.16), then for each f*

in $L^2((R)^d)$, $t \geq 0$, *we have*

$$\begin{aligned} & f(Y(t)) - f(Y(0)) \\ &= \int_0^t \partial_i f(Y(s-)) dY_c^i(s) + \frac{1}{2} \int_0^t \partial_i \partial_j f(Y(s-)) d[Y_c^i, Y_c^j](s) \\ &\quad + \int_0^t \int_{|x|\geq 1} [f(Y(s-) + K(s,x)) - f(Y(s-))] N(ds, dx) \\ &\quad + \int_0^t \int_{|x|<1} [f(Y(s-) + H(s,x)) - f(Y(s-))] \tilde{N}(ds, dx) \\ &\quad + \int_0^t \int_{|x|<1} [f(Y(s-) + H(s;x)) - f(Y(s-)) \\ &\quad - H^i(s,x) \partial_i f(Y(s-))] v(dx) ds. \end{aligned}$$

In order to obtain Itô's formula for a process of type $df(Y(t)) = b(t)dt + F(t)dW(t)$, we just need to put $H = K = 0$ into the previous theorem.

This concludes our section for Lévy processes, we are now armed with all the tools necessary to describe the kind of stochastic processes that we will deal with in the upcoming sections.

Chapter 4

Alpha-Stable Lévy processes

Introduction

Alpha-stable Lévy motions are flexible and extremely useful in modelling financial markets. We begin with the stable distribution, and its common parametrizations. Then we discuss the alpha-stable levy motion (a Lévy process), which is composed of random variables that have a stable distribution. Since only the characteristic function is known for the stable distribution, we show how to recover the probability density function (p.d.f) using a *fast Fourier transform* (FFT) and simulate random numbers that have a stable distribution. In order to ensure asymptotic convergence, we adjust the number of points simulated using the FFT, in the algorithm. We then show a method on how to estimate parameters of a stable distribution, by a histogram approximation technique called the Kernel Density Estimator that uses the *fast Gauss transform* (FGT) to efficiently compute the p.d.f. Natural gas prices exhibit seasonal variations, so we propose a new alpha-stable levy motion where the location parameter varies with time. We also discuss the calibration procedure for this model.

Now the stable distribution has four parameters, and we shall generally write it as $S(\alpha, \beta, c, \mu)$, where $\alpha \in (0, 2]$ controls the heaviness of the tails.

Also, $\beta \in [-1, 1]$, β makes the distribution skew from the left extreme to the right extreme between -1 and 1. And c is the dispersion term that is similar to the volatility term in the gaussian distribution, whereas μ is the location parameter that is similar to the mean in the gaussian. Infact, the gaussian distribution is just a case of the stable distribution with $S(2, 0, c, \mu)$, where c is proportional to the volatility.

4.1 Stable distributions

Let X be a random variable, X_1 and X_2 be independent copies of X and a, b be any positive constants. X is said to be stable if

$$aX_1 + bX_2 \stackrel{d}{=} cX + d, \tag{4.1}$$

for some positive c and some $d \in \mathbb{R}$. The random variable X, is strictly stable if $d = 0$, for all choices of a and b. X is symmetric stable if it is symmetrically distributed around 0, i.e. $X \stackrel{d}{=} -X$.

To generalize this, it can be shown that X is stable if and only if for all $n > 1$, there exist constants $c_n > 0$ and $d_n \in \mathbb{R}$, such that

$$X_1 + X_2 + ... + X_n \stackrel{d}{=} c_n X + d_n, \tag{4.2}$$

where $X_1, ..., X_n$ are independant, identical copies of X. X is strictly stable if $d_n = 0$ for all n. Nolan [1] (and many other books) show that the only possible choice for $c_n = n^{\frac{1}{\alpha}}$ for some $\alpha \in (0, 2]$. Further definitions and implications can be found in Nolan [1].

Definition 7. *A random variable X is said to be stable if and only if there exists a constant d where* $X_1 + X_2 + ... + X_n \stackrel{d}{=} n^{1/\alpha}X + d$ *where* $\alpha \in (0, 2]$. *It is show in in Nolan [2009] that the characteristic function of such a variable X is given by,*

$$\varphi(u) = E[exp(iuX)] = exp[iu\mu - |cu|^{\alpha}(1 - i\beta sgn(u)\Phi], \tag{4.3}$$

where $sgn(u)$ is simply the sign of u and ϕ is given by

$$\Phi = \begin{cases} -(2/\pi)log(|u|), & \text{if } \alpha = 1 \\ tan(\pi\alpha/2), & \text{otherwise} \end{cases} \tag{4.4}$$

$\mu \in \mathbb{R}$ is the location parameter, $\beta \in [-1, 1]$, is the skewness parameter and is a measure of asymmetry and $c \in [0, \infty)$ is the dispersion parameter.

These stable distributions are symmetric around μ when $\beta = 0$, in this case the distribution is called a symmetric stable distribution with the characteristic function

$$\varphi(u) = exp(-i\mu u - \sigma|u|^{\alpha}) \tag{4.5}$$

where μ and α are the same and $\sigma = |c|^{\alpha}$. Notice that this characteristic function is strikingly similar to that of the normal distribution (and exact when $\alpha = 2$), hence σ is chosen as a parameter for the sake of convention.

In order to simulate these random variables, we need their probability density functions (pdf). These can be obtained by taking the inverse fourier transform of the characteristic function and is given by

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{itx} \overline{\varphi_X(t)}\, dt. \tag{4.6}$$

The problem is that the pdf is unknown for a general stable distribution when β is not zero, which unfortunately is the case without skewness. It is for this reason that we discuss the simulation of general (skewed) stable random variables using the fast Fourier transform (FFT). However, in the case of the symmetric stable distribution, the explicit density function is known and can be found in Nolan(2009).

There are three special cases of the stable distribution (from equation (4.3)) which are well known distributions,

$$f_X(x) = \begin{cases} \sqrt{\frac{c}{2\pi}}\, \frac{e^{-c/2x}}{x^{3/2}}, & \text{for } S(1/2, 1, c, 0) \text{ - Lévy} \\ \frac{1}{\pi c\left[1+\left(\frac{x-\mu}{c}\right)^2\right]} & \text{for } S(1, 0, c, \mu) \text{ - Cauchy} \\ \frac{1}{c\sqrt{4\pi}} \exp\left(-\frac{(x-\mu)^2}{4c^2}\right) & \text{for } S(2, 0, c, \mu) \text{ - Normal} \end{cases} \tag{4.7}$$

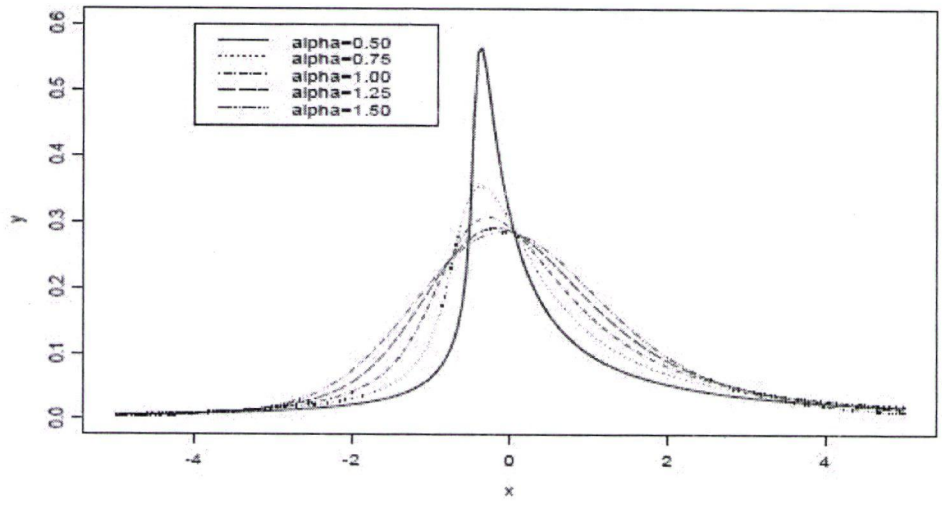

Figure 4.1: Stable densities in the $S(\alpha, 0.5, 1, 0; 0)$ parameterization, $\alpha = 0.5, 0.74, 1, 1.25, 1.5$

In the second and third case, we can see that $\beta = 0$, so they are both also symmetric stable distributions. Now, going back to the skewed stable distribution, in the literature there are two parametrizations that are used commonly and it is important to mention the parametrization one uses. In this paper, only the 1-parameterization of Nolan(2009) is used. The characteristic function for this is precisely that of equation (4.5).

To understand the usefulness of the stable distribution, here is a sample of the the density functions whose shapes it can take.

4.1.1 Simulation of Stable Random Variables

Let X be a random variable, f_X be its density function and φ_X be its characteristic function. We shall now try to approximate f_X from φ_X using the fast Fourier transform (FFT), it is important to setup the problem correctly to utilize the $O(Nlog(N))$ speed of the FFT. Now by definition of the characteristic function,

$$\varphi_X(u) = \int_{-\infty}^{\infty} e^{iux} f_X(x)\, dx \approx \int_{l}^{u} e^{iux} f_X(x)\, dx \approx \sum_{n=0}^{T-1} e^{iux_n} P_n. \tag{4.8}$$

Now we have discretized the above equation and the pdf values are required, so let $P_n = f_X(x_n)\Delta x$, be the sample points starting from the lower bound l and $x_n = l+n(\Delta x)$, and step length $\Delta x = \frac{u-l}{T}$, where T is the total number of points used in the summed discretization.

There are two steps needed in order to recover the probability density function accurately.

1. We shall make the assumption that pdf's have left and right limits, i.e., $\lim_{x\to-\infty} f_X(x) = \lim_{x\to+\infty} f_X(x) = 0$, so u and l are chosen to be sufficiently large, the details of which we shall explain in an algorithm form.

2. Even if our intervals are chosen well, in the case of delta spikes (sharp edges in the pdf) we still need to choose T to be of a large enough value.

So narrowing down our problem, $\varphi_X(u)$ is available for all u and from $\varphi_X(u)$ we need to obtain P_n from which the pdf can be recovered. Equation (8) looks a lot like a Discrete Fourier Transform (DFT), so let us now bring it to a familiar form so that we can apply the FFT.

Let us divide equation (8) by e^{iul} and let $un\Delta x = 2\pi nt/T$. So this gives us,

$$\varphi_X(u)e^{-iul} \approx \sum_{n=0}^{T-1} e^{iun(\Delta x)} P_n = \sum_{n=0}^{T-1} e^{2\pi int/T} P[n] = g[t]. \tag{4.9}$$

In order to remain consistent with Fourier transform notations, we replace P_n by $P[n]$. Thus $g[t] = F^{-1}(P[n])$ or $g[t]$ is the inverse discrete fourier transform of $P[n]$. We input T points of g [t] at the same time to take advantage of the square matrix advantage of the FFT.

So $g[t] = [g[0]g[1]...g[T-1]]$ and $P[n] = [P[0]P[1]...P[T-1]]$. Now, $u(\Delta x) = 2pit/T$, so the u values can be chosen

$$u_t = \frac{2pit}{T(\Delta x)}, t = -\frac{T}{2}, -\frac{T}{2}+1, ..., \frac{T}{2}+1. \tag{4.10}$$

Therefore, $g[t] \approx \varphi_X(u_t)e^{-iu_t l}$. Now when dealing with the stable distributions we often need large intervals chosen for u and l, especially in the case of heavy tail distributions. And when trying to recover the density function, we need to be able to automate the recovery of the pdf, such that u, l and T are obtained. Firstly, we note that the FFT works the fastest when T is a power of 2, i.e, $T = 2^p$ for some p being a positive integer.

The following algorithm calculates the pdf for an interval from -R to R, where R is some sufficiently large number chosen, although any interval can be chosen for u and l. To ensure convergence, we make sure that the pdf recovered has extremes greater than 10^{-4}. L is the log error that can be chosen based on your computational power. In our simulations we have chosen $L = -log(10^{-4}) = 4$. We simply double our interval size if we aren't satisfied with the error.

Algorithm 3.1

Let $a = 1, b = 1, R = 10, L = 4$
while $(a < L)$ OR $(b < L)$
$l = -R, u = R$
$dx = (u - l)/T$
$t = [0, 1, ..., T - 1]$
$st = 2\pi t/t/dx$
$g[t] = \varphi_X(st)exp(-i(st)l)$
$realfft = abs(real((fft(gt, T) - 1/2)/R))$
$a = -log(realfft(1)), b = -log(realfft(T))$
$R = R * 2$
end

Now when α decreases, the tails get heavier. And we know that for sharp edges in the pdf, we need a much higher number of points in the FFT. Experimentally, we see that sharpness happens when $\alpha \in (0, 1]$ and for β close to -1 or 1. But for practical purposes and with respect to natural gas spot modelling, experimentally we can see that $\alpha \in (1.1, 2]$ is sufficient and $\beta \in [0.95, 0.95]$ does the job.

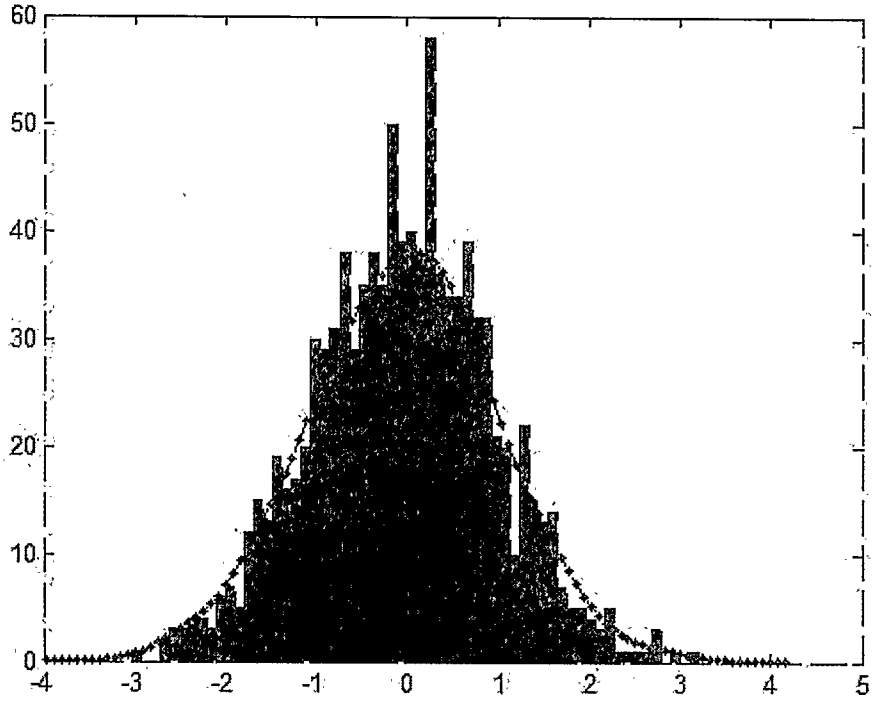

Figure 4.2: Kernel Density Estimate of a sample and its histogram

4.2 Kernel Density Estimation of the Probability Density Function

Kernel density estimation (also called the Parzen [20] window method) allows us to estimate the density function of a random variable *non-parametrically.* In other words, an underlying density function is not assumed. A histogram gives us a rough discontinuous shape of what a density function looks like. What a kernel density tries to do is to smooth the histogram by adding little gaussians at each observed data point, therefore summing up to a smooth looking density function by way of a histogram. Below is a picture of a simulation of 1000 random variables from an alpha-stable distribution with parameters (1.7, -0.5, 2, 0). We can see by looking at the kernel density estimator of the pdf, how close it looks to the actual pdf. Keeping this as our motive, we shall discuss what the kernel density estimator is and its implementation via the fast Fourier transform (FFT).

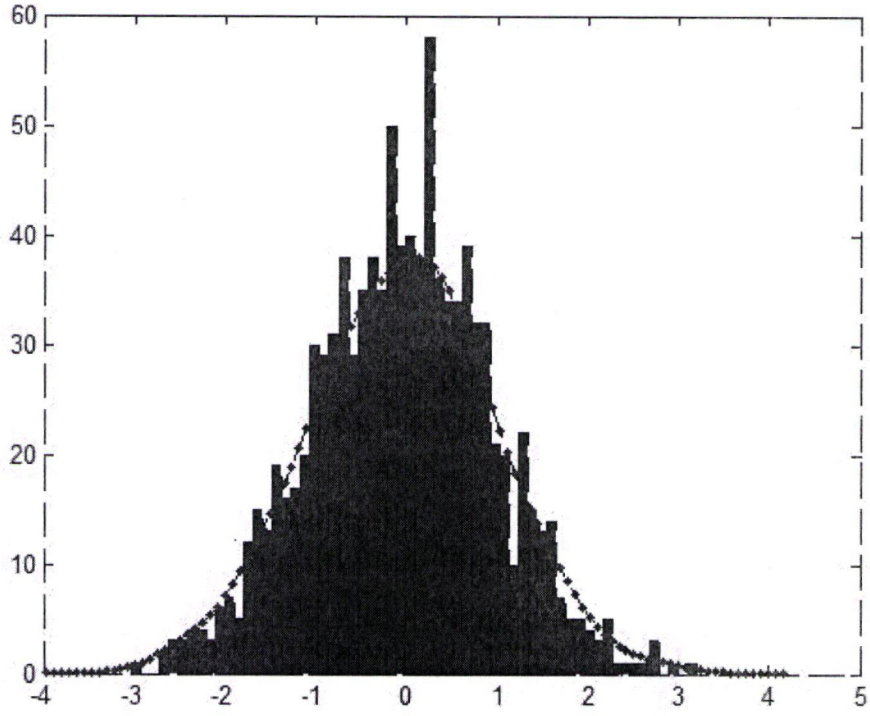

Figure 4.2: Kernel Density Estimate of a sample and its histogram

4.2 Kernel Density Estimation of the Probability Density Function

Kernel density estimation (also called the Parzen [20] window method) allows us to estimate the density function of a random variable *non-parametrically.* In other words, an underlying density function is not assumed. A histogram gives us a rough discontinuous shape of what a density function looks like. What a kernel density tries to do is to smooth the histogram by adding little gaussians at each observed data point, therefore summing up to a smooth looking density function by way of a histogram. Below is a picture of a simulation of 1000 random variables from an alpha-stable distribution with parameters (1.7, -0.5, 2, 0). We can see by looking at the kernel density estimator of the pdf, how close it looks to the actual pdf. Keeping this as our motive, we shall discuss what the kernel density estimator is and its implementation via the fast Fourier transform (FFT).

naive estimator we define the kernel estimator.

Definition 8. *If we have N i.i.d. samples, $(X_i, 0 \leq i \leq N)$ of a random variable X with distribution law f, then the kernel density approximation of its probability density function is*

$$\widehat{f}(x) = \frac{1}{Nh} \sum_{i=1}^{N} K\left(\frac{x - X_i}{h}\right) \tag{4.11}$$

where K is a kernel function and h is a smoothing parameter or the window width.

The most commonly used kernel, which we try to estimate in this thesis as well, is the gaussian function with mean zero and variance one. So,

$K\left(\frac{x-x_i}{h}\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-x_i)^2}{2h^2}}$.

and as could be expected, choosing an optimal bandwidth h is important in finding a good estimator for the kernel density. It is shown in Silverman [21] than optimal value for h is given by

$h_{opt} = \frac{c_1^{-2/5} c_2^{1/5} c_3^{-1/5}}{n^{1/5}}$

where

$c_1 = \int x^2 K(x)\, dx$, $c_2 = \int K(x)^2\, dx$

and

$c_3 = \int (f''(x))^2\, dx$.

Silverman [21] also shows that if we choose a Gaussian kernel with then optimal bandwidth is given by $h_{opt} \approx 1.04 N^{-1/5}$.

If we were to estimate the kernel density in a straightforward way, it would take $O(N^2)$ time, we setup the problem to use the FFT. The following method is adapted from Silverman [22].

Given a function g, let $\widetilde{g}$ denote its fourier transform $\widetilde{g}(s) = (2\pi)^{-1/2} \int e^{ist} g(t) dt$. Now

the empirical characteristic function (ecf) of the samples X_i is defined to be

$$ecf_X(s) = \frac{1}{n}\sum_{j=1}^{n} exp(isX_j) \tag{4.12}$$

So letting $\widetilde{f}(s)$ be the fourier transform of the function f in equation (3.11), we obtain,

$$\widetilde{f}(s) = \widetilde{K}(hs)ecf_X(s). \tag{4.13}$$

by properties of convolutions of Fourier transforms. We used the property that the Fourier transform of the scaled kernel function, $h^{-1}K(h^{-1}t)$ is $\widetilde{K}(hs)$. This expression above is particularly useful when we use a gaussian kernel, since the Fourier transform of a gaussian is a gaussian itself, so we yield an explicit expression,

$$\widetilde{f}(s) = (2\pi)^{-1/2}exp(-\frac{1}{2}h^2s^2)ecf_X(s). \tag{4.14}$$

We once again setup this problem in a similar manner, to use the FFT two times in order to estimate $\widetilde{f}(s)$.

Algorithm 3.2

Let X be our sample r.v.'s, N be the number of samples and $T = 2^m$ where $m > 10$
Choose R such that $X \in [lu]$ where $l = -R$ and $u = l$
$dx = (u - l)/T$
$s = l : dx : u$ (a vector from l to u in intervals of dx
Choose bandwidth, bw (in matlab, do $[pbw] = ksdensity(X, s)$) $p = hist(X, s)$
(histogram of X, evaluated at points s
$fhist = ifft(p, T)$ (inverse fourier transform at points T)
$t = [0, 1, ..., T - 1]$
$st = 2\pi t/t/dx$
$\hat{\varphi}_X = T/Nexp(-\frac{1}{2}bw^2st^2)fhist$ (smoothed empirical characteristic function of X)
$= exp(-\frac{1}{2}bw^2st^2)ecf_X(t)$

In the above algorithm, we have shown a way to compute the risk function discussed in the next section. We shall not show a detailed estimation results in this section since Chapter 7 includes calibration of OU processes based on stable distributions. So estimating the parameters of this distribution is simply a trivial case of an OU process.

4.4 Parameter estimation of stable distributions

Given as estimated density $\hat{f}$ from a sample and a true density f, the most commonly used risk function is the mean integrated square error,

$$R(f, \hat{f}) = \int E[(\hat{f}(x) - f(x))^2]dx \tag{4.15}$$

Now we know that the probability density function f of X is related to its characteristic function ϕ_X, by its inverse fourier transform.

$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{itx} \overline{\phi_X(t)}\, dt.$

and by the definition of $\widetilde{f}$ for the gaussian kernel case in equation (3.13), we establish that

$\hat{f}(x) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{\infty} e^{itx} \overline{\widetilde{f}(t)}\, dt.$

We get the following difference between these two functions.

$$\begin{aligned} \hat{f}(x) - f(x) &= \frac{1}{2\pi}\left(\int_{-\infty}^{\infty} e^{itx}\phi_X(-t)\, dt - \int_{-\infty}^{\infty} e^{itx} exp(-\frac{1}{2}h^2t^2) ecf_X(-t)\, dt\right. \\ &= \int_{-\infty}^{\infty} e^{itx}(\phi_X(-t) - exp(-\frac{1}{2}h^2t^2) ecf_X(-t))dt \end{aligned}$$

We can therefore see that $\phi_X(-t) - exp(-\frac{1}{2}h^2t^2) ecf_X(-t)$ is the fourier transform of the difference of the two functions, based on the optimal bandwidth selection $h \approx 1.04N^{-1/5}$.

Let us define,

$$\hat{D}_X(\theta)(t) = \phi_X(t) - exp(-\frac{1}{2}h^2t^2) ecf_X(t) \tag{4.16}$$

where θ contains the parameters of X. So instead of estimating parameters through a risk function using their density functions, a faster way for parameter estimation through the characteristic function with the following risk function is established.

$$R(X_\theta, \hat{X}) = ||\hat{D}_X(\theta)(t)||_2 \tag{4.17}$$

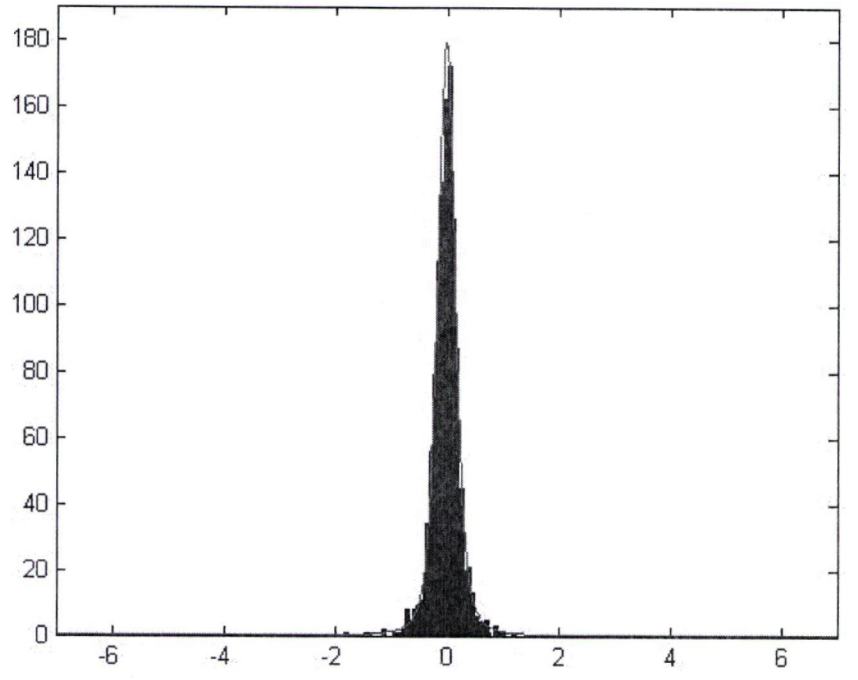

Figure 4.3: Fitted density of Google's returns

So the optimal parameter, in the process of parameter estimation is

$$\theta_{opt} = argmin_{\theta} R(X_{\theta}, \hat{X}) \tag{4.18}$$

Figure 4.3 shows an example of parameter estimation of the log returns of Google stock data using a stable process.

These are Google's returns for the last three years in the histogram, the curve that fits the returns is the stable process whose parameters was found to be (1.5, 0.06, 0.13, 0).

4.4.1 Maximum likelihood estimation

MLE being a standard form of estimation was not used here, but estimation for harder problems using MLE is done in Chapter 7 of this thesis. Now we move on to a description and similar discussion of the NIG process.

Chapter 5

The Normal Inverse Gaussian process

Introduction

The normal inverse gaussian (NIG) distribution is a case of the generalized hyberbolic (GH) distribution. In this chapter we shall discuss GH distributions and its popular cases. We then discuss the NIG process as a more computationally tractible case of a GH distribution, and hence use it for our problem of spot-futures calibration. Then a discussion is made Normal mean-variance mixtures, and how NIG distributions are a case of the variance mixture being an inverse gaussian (IG) distribution. Then we talk about the simulation of IG and NIG processes, since the simulation of the NIG can be trickily done via the IG process. We also shall go into parameter estimation techniques for the NIG process, through maximum likelihood and its empirical characteristic function (ECF).

5.1 Generalized Hyperbolic Distributions

The generalized hyperbolic (GH) distribution was introduced during the study of grains of sand, by Barndorff-Nielsen [25] in 1997. Its probability density function is,

$$\frac{(\gamma/\delta)^{\lambda}}{\sqrt{2\pi}K_{\lambda}(\delta\gamma)}e^{\beta(x-\mu)}\frac{K_{\lambda-1/2}\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)}{\left(\sqrt{\delta^2+(x-\mu)^2}/\alpha\right)^{1/2-\lambda}} \tag{5.1}$$

where μ is the location, β is the asymmetry parameter, δ is the scale, λ and α control the tails. Also $\gamma = \sqrt{\alpha^2-\beta^2}$ and K_{λ} is a modified Bessel function of the second kind [25].

Also, the characteristic function for the GH distribution is given by

$$\phi_{GH}(t) = \frac{e^{\mu t}\gamma^{\lambda}}{(\sqrt{\alpha^2-(\beta+t)^2})^{\lambda}} \frac{K_{\lambda}(\delta\sqrt{\alpha^2-(\beta+t)^2})}{K_{\lambda}(\delta\gamma)} \tag{5.2}$$

We say that a random variable X has a GH distribution if $X \sim GH(\lambda, \alpha, \beta, \delta, \mu)$. GH distributions are very flexible and have desirable properties due to its heavy tails which has two parameters (λ and α) to control it. One of the criticisms of the Stable distribution is that the tail can be very heavy and can overestimate the risk of returns, which can also be a good thing for risk managers. The GH distribution typically fits log returns of stock prices quite well, but this can also be attributed to the extra parameter that the GH distribution has over the stable distribution.

The GH is a super class of many popular distributions such as:

1. Hyperbolic distribution with $GH(1, \alpha, \beta, \delta, \mu)$
2. Normal inverse Gaussian (NIG) Distribution with $GH(-1/2, \alpha, \beta, \delta, \mu)$
3. Variance gamma with $GH(\lambda, \alpha, \beta, 0, \mu)$

and the normal distribution, Student-t distribution, among others.

Until recently the GH distribution had major drawbacks due to the computationally intensive nature of calculating K_{λ} in the density and characteristic function. For this reason we use the NIG distribution, that we see in the upcoming section. Its form of the characteristic function and moment generating function, are of interest to us in terms of deriving the futures formula for Natural Gas. Now in order to understand what the NIG distribution is, we first study the Inverse Gaussian distribution.

5.1.1 Inverse Gaussian Distribution

If W_t is a standard brownian motion, then a stochastic process $X_t = \mu t + \sigma W_t$ is brownian motion with drift μ. The subordinator process is defined by

$$T(t) = inf(s < 0; X_t = a), a > 0 \tag{5.3}$$

The density of $T(t)$ turns out to be an Inverse Gaussian(IG) distribution, $IG(\frac{a}{\mu}, \frac{a^2}{sigma^2})$. The density function for the IG distribution is,

$$f(x) = \left[\frac{\lambda}{2\pi x^3}\right]^{1/2} \exp \frac{-\lambda(x-\mu)^2}{2\mu^2 x} \tag{5.4}$$

and has a characteristic function defined by,

$$\phi_I G(t) = e^{\left(\frac{\lambda}{\mu}\right)\left[1-\sqrt{1-\frac{2\mu^2 it}{\lambda}}\right]} \tag{5.5}$$

where λ, $\mu > 0$.

The IG process is just a collection of random variables $X_{IG} = X_t^{IG}, t > 0$ that have an IG distribution, $X_{IG} \sim IG(\mu t, \lambda)$. Now that we know what an IG distribution is, we quickly describe what are normal variance-mean mixture models.

5.1.2 Normal variance-mean mixtures

Normal variance-mean mixtures were discussed in detail by Barndorff-Nielsen and Kent [27] in 1982. If X is a normally distributed random variable with mean zero and variance one, i.e. $X \sim N(0,1)$ and V^2 has a probability density g. Then a random variable Y, of the form

$$Y = a + bV^2 + \sigma VX \tag{5.6}$$

is a normal variance-mean mixture, with "mixing" density g. The probability density function of the random variable Y takes the form.

$$f(x) = \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2 v}} \exp\left(-(x-a-bv)^2/(2\sigma^2 v)\right) g(v)dv \tag{5.7}$$

and its characteristic function is given by,

$$\phi_N VM(t) = \exp(ait)\,\phi_g\left(-bt - i\frac{1}{2}\sigma^2 t^2\right) \tag{5.8}$$

where ϕ_g is the characteristic function of V^2 with density function g.

5.2 Normal Inverse Gaussian Distribution

The normal inverse gaussian distribution is a normal variance-mean mixture where the mixing density g is that of an IG distribution. The formulas we have above directly when evaluated for the density g yield us explicit formulas for the NIG distribution and characteristic function.

The density function of the NIG distribution is

$$\frac{\alpha\delta K_1\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)}{\pi\sqrt{\delta^2+(x-\mu)^2}}\, e^{\delta\gamma+\beta(x-\mu)} \tag{5.9}$$

and has a characteristic function of

$$\psi_{NIG}(t) = e^{i\mu t+\delta(\gamma-\sqrt{\alpha^2-(\beta+it)^2})} \tag{5.10}$$

As mentioned previously, the characteristic function of the NIG has an easily computable and clean form and parameter estimation is done by directly utilizing the FFT based algorithm, developed in the previous section for calibration via the empirical characteristic function.

Finally the NIG process, is a collection of random variables $X_t^{NIG}, t > 0$ that have an NIG distribution.

5.3 Simulation

To simulate the NIG process, we must first learn how to simulate the IG process. So we discuss this next.

5.3.1 Simulation of the IG process

In order to generate random variables that belong to an IG distribution, we perform the following steps.

- $v = N(0,1)$, i.e. generate a number $v \sim N(0,1)$
- $y = v^2$
- $x = \mu + \frac{\mu^2 y}{2\lambda} - \frac{\mu}{2\lambda}\sqrt{4\mu\lambda y + \mu^2 y^2}$
- $u = U(0,1)$, i.e. generate a uniformly distributed random variable, u.
- If $u \leq \frac{\mu}{\mu+x}$ then return x, else, return μ^2/x.

x generated here, is an IG random variable with distribution $IG(\mu, \lambda)$.

To generate an IG process, we consider that a process $X_{IG} = X_t^{IG}, t > 0$ of random variables has the law $IG(\mu t, \lambda)$. Below are the steps for simulating the IG process.

- Generate N random variables, $x_i, i = 1, .., N$ with the distribution $IG(\mu\Delta t, \lambda)$, where $\Delta t = T/N$ and T is the time period you want to simulate.
- Set $X_0^{IG} = 0$
- Set $X_{i\Delta t}^{IG} = X_{(i-1)\Delta t}^{IG} + x_i$

The simulated values, X_t^{IG} form an IG process.

5.3.2 Simulation of the NIG process

To simulate an NIG process of type $NIG(\alpha, \beta, \delta t, 0)$, we think of the definition of NIG as an IG time-changed Brownian motion with drift. The following is the procedure to simulate the NIG process:

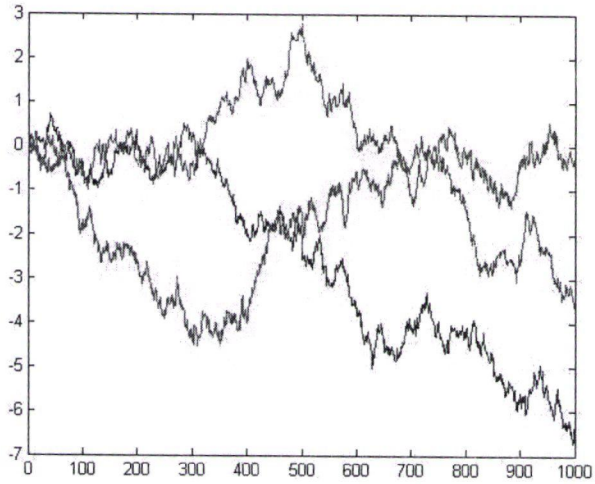

Figure 5.1: NIG sample paths

- Using the above algorithm, we get a series of values X_t^{IG} at times points $n\Delta t, n = 0, 1, 2...,$ for values $a = 1$ and $b = \delta\sqrt{\alpha^2 - \beta^2}$
- Set $dt_{n\Delta t} = X_{n\Delta t}^{IG} - X_{(n-1)\Delta t}^{IG}$
- Simulate the time change of a standard Brownian Motion by

  a. Simulate n random variables with, $\nu_n \sim N(0, 1)$ b. $W_0 = W_{X_0}^{IG} = 0$ c. $W_{n\Delta t} = W_{(n-1)\Delta t} + \sqrt{dt_{n\Delta t}}\nu_n$
- Repeat path with $X_{n\Delta t}^{NIG} = \beta\delta^2 X_{n\Delta t}^{IG} + \delta W_{n\Delta t}$.

5.4 Calibration

The estimation of parameters of the NIG process follows from the previous chapter on the empirical characteristic function (ECF) estimation of the parameters of the alpha-stable process. If we have a sample process $(X_i, i = 0, 1, ..n)$, with n observations, we can determine its ECF by, $ecf_X(s) = \frac{1}{n}\sum_{j=1}^{n} exp(isX_j)$. We also have the characteristic

function of the NIG available in equation 5.10. So as in equation 4.16, our optimal parameter is obtained by

$$\theta_{opt} = argmin_{\theta} R(X_{\theta}, \hat{X}) \tag{5.11}$$

where

$$R(X_{\theta}, \hat{X}) = ||\psi_{NIG}(t) - exp(-\frac{1}{2}h^2t^2)ecf_X(t)||_2$$

The estimation of ecf_X is done in the previous chapter in *algorithm (4.2)*. We do not show the parameter estimation of the NIG process here, since in Chapter 7 we have OU based NIG processes that are calibrated.

Estimating the parameters via the ECF provides the huge advantage of not having to evaluate the Bessel function in the density function. Other methods for estimation in NIG can include the Methods of Moments, Maximum Likelihood Estimation, Bayesian Inference. But all of them have the disadvantage of being computationally expensive, due to the time required to calculate the density function alone.

Chapter 6

Futures Matching in Natural Gas

Introduction

In this section we shall talk about the natural gas futures markets and the effects of natural gas storages on futures prices. Natural gas is produced in different parts of North America, and the centres of gas production have trading hubs where futures contracts on natural gas are bought and sold. Natural gas is produced in various locations and pipelines run all over North America connecting different gas storages. To understand the behaviour of futures prices in natural gas, we have to make a quick dicussion about two types of natural gas storages.

Base Load Storage

These storages are typically large and are meant for a regular base supply of gas for a region. They have a low daily limit on the amount of gas that can be injected and withdrawn from the storage. They provide a steady supply of gas to the market throughout the year.

Peak Load Storages

These storages are smaller and don't operate throughout the year, but have high injection/withdrawal limits, and can make up for the supply of gas during periods of extreme demand (i.e. the peak winter). Storages have to maintain a base level of gas always, otherwise the withdrawal costs increase, due to the high pressure of gas in the storage. Traders speculate on prices also based on the expected rise in gas prices due to low

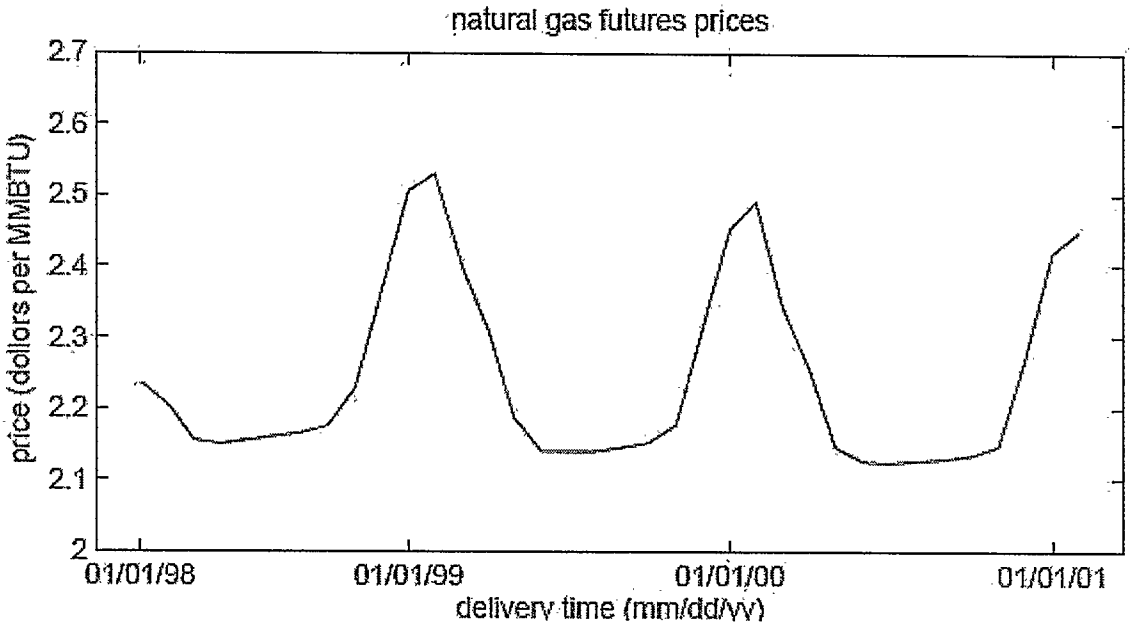

Figure 6.1: Futures curve

storage levels.

One of the most important indexes of futures is that of the Henry Hub (HH), which is located in Louisiana. They are traded over a NYMEX (New York Merchantile) exchange. Natural gas pipelines run all over North America and the Henry Hub (HH) location is almost central to the US and Canada. In Alberta, natural gas futures are traded on the AECO hub. Henry Hub prices are seen as the typical gas prices in the North American market.

Various other pricing points, such as AECO in Alberta, always have to price their contracts lower than that of Henry Hub. This is because gas from Alberta, for example, would be bought in the US only if the price of gas from Alberta along with the transportation costs to the location of delivery would be less than the Henry Hub price (plus local delivery costs). The US imports gas from Canada mainly during cold peak winter months and on the hottest summer days.

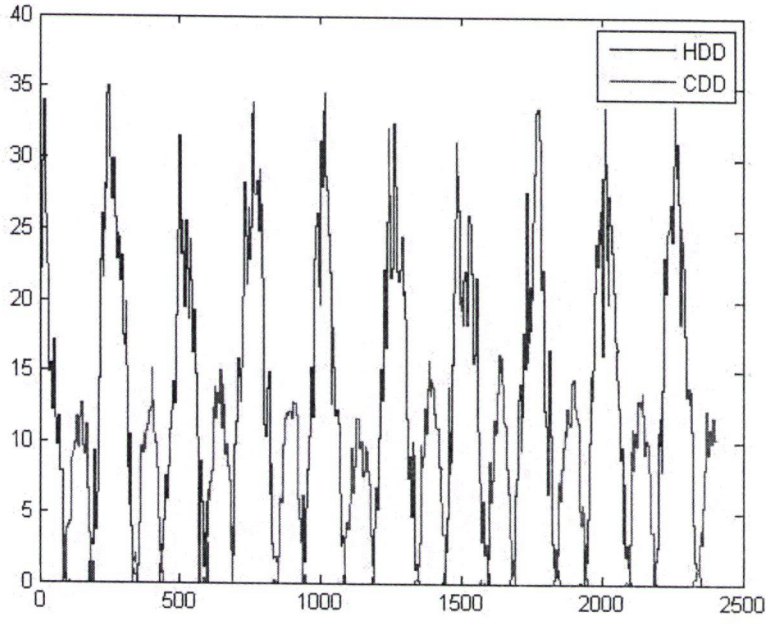

Figure 6.2: HDD and CDD curves

On the hottest days, gas is used for cooling homes (via air conditioners) and on the cold days gas is used for heating homes. It is during these times that the gas storages experience shortage, i.e. the amount of gas produced in the US does not meet its local demand. To meet the demand, natural gas is imported from locations which are geographically closer to the part of the US requiring the gas. For example, in the Northern states of the US, when there is shortage, gas is imported from Alberta or other parts of Canada. In the southern states of the US, gas is usually imported from Mexico.

Another feature of natural gas price behaviour is that when the storage levels (as seen by the EIA[1] inventory numbers) are low, which happens during the winter time, the extraction costs of gas from the storage is higher and this drives the gas prices to spike up during the coldest days.

Now the spot price of natural gas is set at the different trading hubs, and as the

[1]Energy Information Association - US Department of Energy (DOE), contains weekly reports on the gas levels in storages around North America.

name suggests, it is intended for spot delivery. The daily spot price is determined by the supply and demand of gas usage. Futures are traded due to the high uncertainity of these volatile gas spot prices on an intended day of purchase.

6.1 Natural Gas Futures and Spot Price

As mentioned previously, futures contracts in natural gas are traded in different exchanges around North America. Contracts exist on a monthly basis for every year. There are contracts available for the next 10 years, and each year has 12 contracts. There are several contracts of the same maturity being traded as well. In 2002, NYMEX had 97,000 contracts being traded every day. Many of these contracts that have a maturity greater than two or more years are not very liquid.

The futures price, $F(t,T)$ is the price of the contract at time t with the date of maturity being T. Futures contracts have their expiry at the end of a month, in other words T can only take dates that represent the end of the month. For simplicity, we express all the contracts dates available after a day t to be T_i for $i = 1, 2, ...$, where i represents the next available contract month. This means that, $T \in T_1, T_2, ...$ The value of a futures contract at maturity T_i tells about the markets expectation of the spot price at time T_i under the risk neutral measure Q.

$$F(s,t,T) = E_Q[S(T)|S(t) = s] \tag{6.1}$$

Also, $S(t)$ is the spot price at a time t. Now our goal for this thesis is to determine spot price processes that satisfies the above equation.

6.1.1 Handling the Futures Curve

Say we have M available futures contracts with time to maturity T_i where $i = 1, 2, ..., M$. Our spot model has to satisfy the constraint,

$$[F(s,t,T_1)F(s,t,T_2)...F(s,t,T_M)] = E_Q[S(T_1)S(T_2)...S(T_M)] \tag{6.2}$$

This means that, we also need to find a relationship between each futures contract's traded prices and its corresponding spot price for the day. That is,

$$F(t,T_i) = F_S(S(t),t,T_i) \tag{6.3}$$

where F_S is a function of the spot price $S(t)$ at a time t and expiry time T_i.

The relationship between different futures contracts were established the following way. In a continuous time model, we would use the notation

$$\mathbf{F_M}(t) = [F(t,T_1)F(t,T_2)...F(t,T_M)]', \tag{6.4}$$

as a vector (1xM matrix) of a collection of contracts with different maturities traded on a particular day t.

In a discrete time setting which is necessary for calibration, if we are looking at N consecutive observed trading days of futures prices of different maturities and spot prices, then we can use the notation

$$\mathbf{F_M^N} = \begin{bmatrix} F(1,T_1) & F(1,T_2) & ... & F(1,T_M) \\ F(2,T_1) & F(2,T_2) & ... & F(2,T_M) \\ . & . & ... & . \\ . & . & ... & . \\ F(N,T_1) & F(N,T_2) & ... & F(N,T_M) \end{bmatrix}, \tag{6.5}$$

This matrix contains N observed prices on consecutive trading days, for M contracts with maturities T_i.

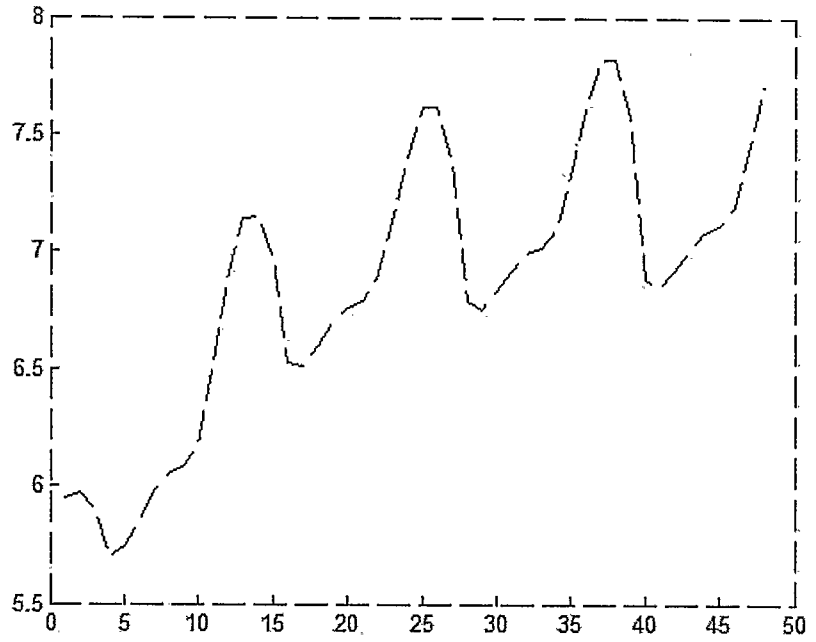

Figure 6.3: Futures curve for June 1st 2007, from 1 to 48 months of contract maturity

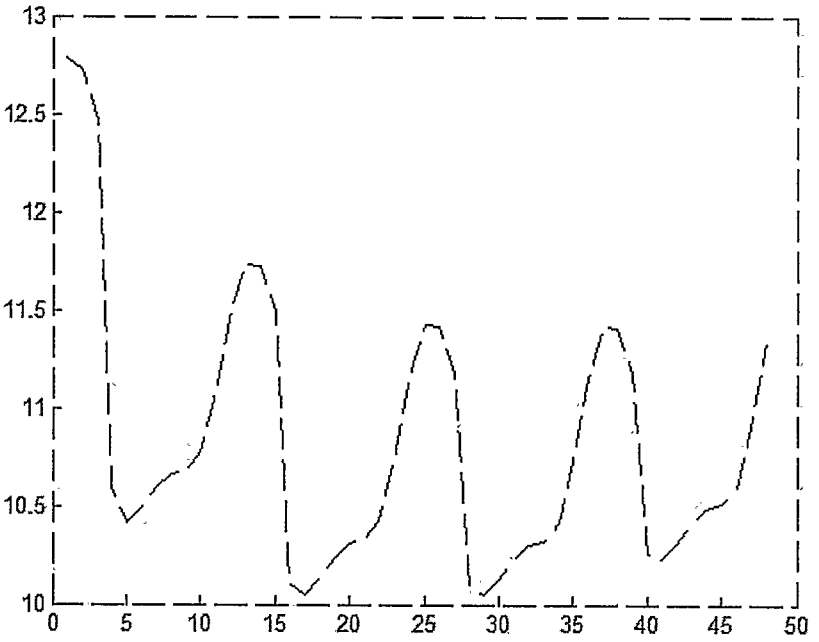

Figure 6.4: Futures curve for June 1st 2008, from 1 to 48 months of contract maturity

In the times of the global economic recession, the gas price as of July 1st 2009 is 3.88/mmbtu USD (US Dollar), which is down drastically from 13.28 USD on the same day in 2008! As a result, we can see that the futures curve at low prices have an upward drift. This is because even though the gas prices are low now, the gas prices are "expected" to pick up soon and so futures have an upward drift. It could also be due to fluctuation in storage capacity issues, but the reasons are not clear for now.

6.1.2 An Affine Seasonality Term

In this section, we discuss an extension to Pilipovic's model and another model. Now we can see in section 2.2.3, the general rule as followed by Xu [14] as well, is that $F(t,T) = F_{UND}(t,T) + seasonality$ or

$$F(t,T) = F_{UND}(t,T) + f(t,T) \tag{6.6}$$

where $f(t,T)$ is a seasonality term whose parameters are calibrated by

$$f(t,T) = \sum_{i=1}^{2}[u_i sin(2\pi r f c((T-t) - t_i^C)) + v_i cos(2\pi r f c((T-t) - t_i^C))] \tag{6.7}$$

Now from observations during calibration, it was seen that this seasonality was certainly not constant (results for this is shown in the next chapter). But an affine form for its coefficient does a reasonable job in trying to capture the "amplitude" of the seasonality. That is

$$f(t,T,S_t) = m(t,T,S_t)f(t,T) \tag{6.8}$$

where $m(t,T,S_t) = a + bS_t$, i.e. has an affine structure. Now another observation made was that

$$f(t,T,S_t) = m(t,T,F(t,T))f(t,T) \tag{6.9}$$

where $m(t,T,F(t,T)) = a + b(max(F(t,[T_M...T_{M-12}])) - min(F(t,[T_M...T_{M-12}])))$. Either way, the idea is that the seasonality term is initially estimated in equation 6.26 and

is taken as $f(t,T)$, but its coefficient is not constant, i.e.,

$$F(t,T,S_t) = F_{UND}(t,T,S_t)+f(t,T,S_t) = F_{UND}(t,T,S_t)+m(t,T,F(t,T))f(t,T) \quad (6.10)$$

It is convenient to write it in the above way, since the futures price in formulation is related to the spot price S_t.

6.1.3 Two-Factor models for gas (Pilipovic and Xu extensions)

Pilipovic [34], uses the following two factor model

$$dS_t = \alpha(L_t - S_t)dt + \sigma S_t dW_t^1 \quad (6.11)$$

$$dL_t = \mu L_t dt + \gamma L_t dW_t^2, \quad (6.12)$$

In Pilipovic [34], she simplifies her model further which actually helps a great deal when doing parameter estimation, as the number of parameters reduce and there is no ambiguity. Equation 2.5 can be reduced to,

$$F_{UND}(t,T,S_t) = (S_t - L_t)e^{-(\alpha')(T-t)} + L_t e^{(\mu')(T-t)} \quad (6.13)$$

where $\alpha' = \alpha + \lambda\gamma$ and $\mu' = \mu - \lambda\gamma$ Therefore, we have a futures model where,

$$F(t,T,S_t) = (a + bS_t)f(t,T) + (S_t - L_t)e^{-(\alpha')(T-t)} + L_t e^{(\mu')(T-t)} \quad (6.14)$$

There are only two parameters (seen in Chapter 7) to be recovered in total, thus simplifying the problem.

Now as can see, the long run mean implied by the spot price process just follows a geometric Brownian motion. We believe that the long run mean should be mean reverting as well, we back the claim since upon calibration we saw that two parameters are not sufficient to estimate the system.

Instead we actually use the idea of Xu [14], where the long run mean is also mean reverting, i.e.

$$dS_t = \alpha(L_t - S_t)dt + \sigma S_t^r dW_t^1 dL_t = \mu(\gamma - L_t)dt + \tau L_t^r dW_t^2 \tag{6.15}$$

where $r = 0, 0.5, 1$. The solution to the above system is given in Xu [14] by,

$$\begin{aligned} F_U ND^\theta(t,T,S_t,L_t) &= e^{\alpha(t-T)}S_t + \frac{\alpha}{\alpha-\mu}(e^{\mu(t-T)} - e^{\alpha(t-T)})L_t & (6.16)\\ &+ \frac{\mu\gamma}{\alpha-\mu}(e^{\alpha(t-T)} - 1) - \frac{\alpha\gamma}{\alpha-\gamma}(e^{\mu(t-T)} - 1) & (6.17) \end{aligned}$$

so the futures curve could be modelled as,

$$F(t,T,S_t) = (a + bS_t)f(t,T) + F_U ND^\theta(t,T,S_t,L_t) \tag{6.18}$$

The futures equation here requires the estimation of three parameters (as seen in Chapter 7) and works well in calibration. In this thesis, the multifactor model that we calibrate is the above.

6.1.4 General Structure of Lévy-Based Models

This time instead of working backwards, we start off with a reasonable choice of process for the spot price. In this thesis we shall deal with one factor Lévy-based stochastic models of the following form

$$dX_t = \lambda(b - X_t)dt + X_t^r dL_t \tag{6.19}$$

$$S_t = f(t) + exp(X_t) \tag{6.20}$$

where L_t is a Lévy process, $f(t)$ is the same seasonality parameter as the previous section and S_t is the natural gas spot price process.

If $r = 0$, it is an OU process and for $r = 0.5$ it is a Cox-Ingersoll-Ross type process. Now for each of these models we would like to get the futures price and its characteristic

function, but these are unknown to the author for $r = 0.5$ and $r = 1$. So in the upcoming sections we only consider the above model for $r = 0$.

For an OU type model we simplify the equation for the sake of parameter convergence during estimation,

$$dX_t = -\lambda X_t dt + dL_t \tag{6.21}$$

This is the form also considered in Cont and Tankov [24], basically the mean reverting indicator gets mixed into the "location" term for the Lévy process.

For the cases where $r = 1$ and $r = 0.5$ since we don't have the futures price, we calibrate it only for the spot process as models of such type don't a known explicit characteristic function when there is an underlying Lévy process that is non-Wiener. Unfortunately such are the cases that is of interest to us. The approach we take in calibrating such models is that of using the simulation equation or the corresponding difference equation of equation 6.34,

$$\begin{aligned} X_{t+1} - X_t &= \lambda - X_t h + X_t^r L_1 \\ \Rightarrow X_{t+1} - X_t(1 - \lambda h) &= X_t^r L_1 \\ \Rightarrow \frac{X_{t+1} - X_t(1 - \lambda h)}{X_t^r} &= L_1 \end{aligned}$$

But L_1 above is just a Lévy distribution, so it has a known characteristic function. Therefore the characteristic function for the simulation equation is

$$exp(iu\frac{X_{t+1} - X_t(1 - \hat{\lambda})}{X_t^r}) = exp(\psi_L(u)) \tag{6.22}$$

where $\psi_L(u)$ is the characteristic function of the chosen Lévy process. In the section on calibration in the next chapter we discuss how the above expression is exploited to actually retrieve the parameters for equation 6.34.

Now we shall look at the explicit solutions and futures price inferred from the models of the following OU type Lévy models. As far as naming conventions go, we shall name our model in equation 6.34, in the following form

$$model - 1 -' distribution' - r$$

So if we were looking at L_t being an alpha-stable noise driving term with $r = 0.5$, then the model would be called $model - 1 - alpha - 0.5$. Similarily for an NIG process, if $r = 1$, we call it $model - 1 - NIG - 1$.

6.1.5 Model-1-Alpha-0: An alpha-stable based OU futures model

Stable distributions were dicussed in the previous chapter. In this section we shall introduce a stable Lévy based spot price process and derive the futures price formula.

$$dX_t = -\lambda X_t dt + dL_t \tag{6.23}$$

$$\tag{6.24}$$

where L_t is of a stable distribution $(\alpha, \beta, \sigma, \mu)$, $f(t)$ is the same seasonality parameter as the previous section and S_t is the natural gas spot price process.

The futures price of a contract where L_t is a symmetric alpha-stable process (i.e. $\beta = 0$) is:

$$F(t,T) = E[S(T)|S(t)] = E[f(T) + exp(X_T)|X(t)] = f(T) + E[exp(X_T)|X(t)] \tag{6.25}$$

Now from Cont and Tankov [24], the solution to X_t is of the form,

$$E[exp(iuX_T)|X(t)] = exp(iuX_t exp(-\lambda(T-t)) + \int_t^T \psi_{L_T}(-ie^{\lambda(s-T)}ds) \tag{6.26}$$

where L_t in the above equation can actually be any Lévy process, but in our case we shall solve it for a Stable Lévy process and in the next section we do the same for a normal inverse gaussian Process. Solving the above integral gives us,

$$\int_t^T \psi_{L_T}(ue^{\lambda(s-T)})ds = \frac{iu\mu\,\alpha\,\left(1-\mathrm{e}^{-\lambda\,(T-t)}\right)-(|cu|)^{\alpha}\left(1-\mathrm{e}^{-\lambda\,\alpha\,(T-t)}\right)}{\alpha\,\lambda} \tag{6.27}$$

where L_t is an alpha-stable Lévy process. Starting at a time 0, the characteristic function of X_t is given by

$$\psi(t,u,X_t) = E[exp(iuX_t)] = exp(iuX_0exp^{-\lambda t} + \frac{iu\mu\,\alpha\,\left(1-\mathrm{e}^{-\lambda t}\right)-(|cu|)^{\alpha}\left(1-\mathrm{e}^{-\alpha\,\lambda\,t}\right)}{\alpha\,\lambda}) \tag{6.28}$$

Now with respect to calibration its useful to have a transitional density function, or a conditional density. This can be easily obtained by finding the dynamics of $X_{t+1}-X_t$, which can be easily obtained from the characteristic function evaluation above.

So, the transitional characteristic function of X_t is

$$\begin{aligned} E[exp(iu(X_{t+1}-X_t)] &= exp(iuX_0e^{-\lambda t}(e^{-\lambda}-1) \\ &\quad + \frac{iu\mu\,\alpha\,\mathrm{e}^{-\lambda t}\left(-\mathrm{e}^{-\lambda}+1\right)-(|cu|)^{\alpha}\,\mathrm{e}^{-\alpha\,\lambda\,t}\left(1-\mathrm{e}^{-\alpha\,\lambda}\right)}{\alpha\,\lambda}) \end{aligned}$$

The above equation also tells us the characteristic function of the log returns spot price process if we set $f(t)=0$, i.e., $X_{t+1}-X_t = log(S_{t+1})-log(S_t) = log(\frac{S_{t+1}}{S_t})$.

For the futures price, we do not take into account skewness since in practice we see impulsiveness in noise but there is no skewness exhibited in the natural gas spot process. We shall not include the proof here, since it is straight forward, but the solution to the futures process for a symmetric alpha-stable Lévy process, L_t, is found by simply setting $u=-i$ in the characteristic function, i.e.

$$F(t,T) = f(T)+E[exp(X_T)|X(t)] = f(T)+\psi(t,-i,X_t) \tag{6.29}$$

$$= f(T)+exp(X_te^{-\lambda(T-t)} \tag{6.30}$$

$$+\frac{\mu\,\alpha\,\left(1-\mathrm{e}^{-\lambda\,(T-t)}\right)-(|c|)^{\alpha}\left(1-\mathrm{e}^{-\lambda\,\alpha\,(T-t)}\right)}{\alpha\,\lambda}) \tag{6.31}$$

6.1.6 Model-1-NIG-0: An NIG process OU based futures model

The NIG process has the advantage of having semi-heavy tails which is also gives good accuracy when looking at densities for gas prices for shorter periods of time. Now as the progression went in the previous section, we look are going to look at the equation for X_t. In this setting we have,

$$dX_t = -\lambda X_t dt + dL_t$$

where L_t is an NIG process with parameters $(\alpha, \beta, \delta, \mu)$. Now that we have a new process for our model, solving the same integral in equation 6.37, it works out to be

$$\int_t^T \psi_{L_T}(ue^{\lambda(s-T)}ds = \frac{-iu\mu\left(\mathrm{e}^{-\lambda(T-t)}-1\right)}{\lambda} - \delta\sqrt{\alpha^2-\beta^2}\,(T-t) + P_t^{NIG}(u,T) \quad (6.32)$$

where

$$P_t^{NIG}(u,T) = \int_t^T \delta\sqrt{\alpha^2-(\beta+iue^{\lambda(s-T)})^2}ds \quad (6.33)$$

The explicit integral solution of $P_t^{NIG}(u,T)$ is not known to us, so we shall simply numerically integrate it. Therefore, the characteristic function of X_t is given by

$$\begin{aligned}\psi(t,u,X_t) &= E[exp(iuX_t)] = exp(iuX_0exp^{-\lambda t} + \frac{-iu\mu\left(\mathrm{e}^{-\lambda(t)}-1\right)}{\lambda} \\ &\quad -\delta\sqrt{\alpha^2-\beta^2}\,(t) + P_0^{NIG}(u,t))\end{aligned}$$

And the transitional characteristic function of X_t is

$$\begin{aligned}E[exp(iu(X_{t+1}-X_t)] &= exp(iuX_0e^{-\lambda t}(e^{-\lambda}-1) \\ &\quad +\frac{i\mu\,u\mathrm{e}^{-\lambda t}\left(\mathrm{e}^{-\lambda}-1\right)}{\lambda} \\ &\quad +\delta\sqrt{\alpha^2-\beta^2} + P_0^{NIG}(u,t+1) - P_0^{NIG}(u,t)\end{aligned}$$

Also as in the previous section, the futures price is obtained the same way by setting $u=-i$ in the characteristic function, i.e.

$$F(t,T) = f(T) + \psi(t,-i,X_t) \tag{6.34}$$

$$= f(T) + exp(X_t e^{-\lambda(T-t)} \tag{6.35}$$

$$+\frac{\mu\left(1-\mathrm{e}^{-\lambda(T-t)}\right)}{\lambda} - \delta\sqrt{\alpha^2-\beta^2}\,(T-t) + P_t^{NIG}(-i,T)) \tag{6.36}$$

In the upcoming section, we calibrate all the models described in this section.

Chapter 7

Futures and Spot Model Calibration

Introduction

In natural gas futures models the ultimate test is to be able to see if the formulated futures price from the spot price actually captures the dynamics of the futures curve. From empirical work done in this thesis, many models were implemented, but what was observed was that models described in literature with parameter estimation results were typically estimated the futures prices of contracts for more than 1 year expiry.

The relation between futures prices in the first year of expiry, specially in the first four to five months are very uncertain and is extremely difficult to model. This is because they are not only dependent on the spot price but also on natural gas storage levels, weather, hurricane forecasts, among others. In academic literature these factors are conveniently ignored to find a satisfactory analytical model that describes future prices based on seasonality relations and spot prices.

7.1 Choosing the Spot Price

Natural gas in North America has many different hubs and pricing points, but Louisana being a major producing region has the Henry Hub associated with it. Being the major supplier of eastern US (where the population is very high), it is considered to be nationally and internationally an indicator of gas prices. But over the last five years, the Henry Hub is no longer the major producing region for gas and so the other commonly used proxy for gas prices is the front month prices for the futures contracts traded in the NYMEX

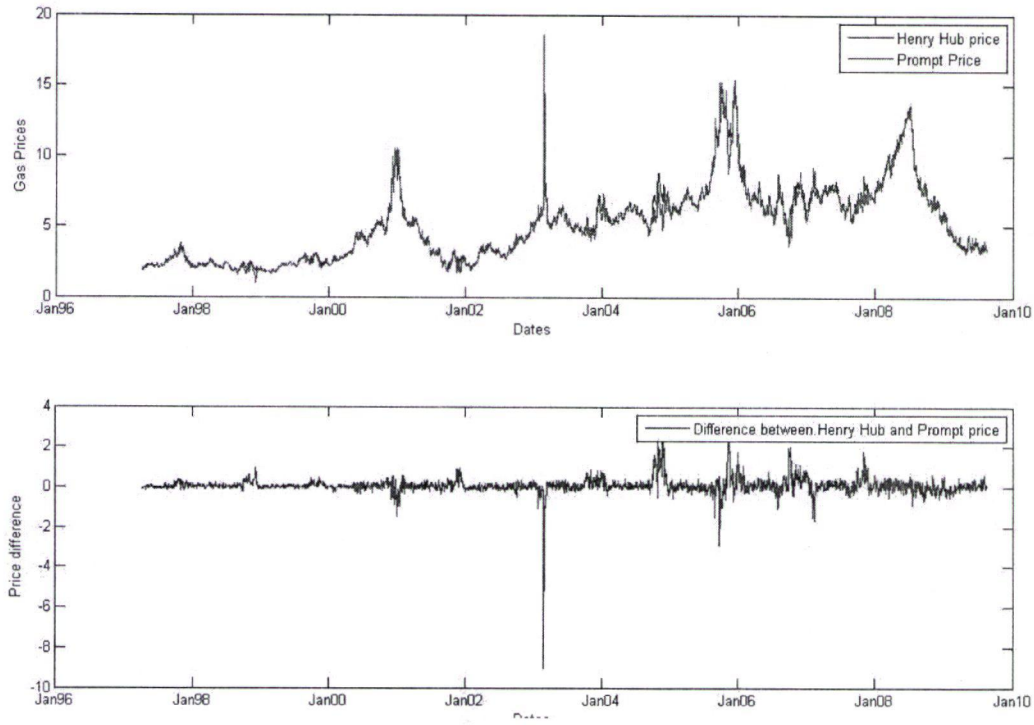

Figure 7.1: Prompt and Henry Hub prices

exchange.

That is, if we are in the month of September, the gas prices for the month are taken to be the futures prices for October contract being traded in September. This is known as the prompt price, we can see below the prompt price, spot price and their differences over the last 10 years.

When testing futures models with respect to the spot price process, it is ideal to calibrate using one or the other. As can be seen, there are several times in which the Henry Hub price and the futures price has a gap, thus actually violating the usual rule of spot convergence of futures contracts, which says that $F(T,T) = S(T)$. This is usually used as a boundary condition in computing the futures price PDE by Pilipovic [12], Xu [14], etc.

So it should be considered that in the natural gas market this convergence rule doesn't always hold, since there is no "real" spot price for natural gas. For calibration purposes

however, we use mainly Henry Hub prices, but prompt prices are also mentioned if there was a significant difference in the model.

7.2 Approaches to Modelling the Futures Curve

Modelling the futures curve is similar to modeling the bond yield curve, the seasonality can be filtered out and the underlying curve can be obtained. Atleast that is the target of modelling the futures curve.

All of the models in Chapter 2, assume a spot price process for gas and want to find a corresponding futures model. If the models were to fit, the parameters recovered from the futures model should be close to the parameters in the spot price calibration of the gas prices.

However, this is usually not implemented and Pilipovic [34] describes it as an "ultimate test" for model goodness. This approach is also the one used in model 6.1.2, 6.1.3 and 6.1.4 in this thesis.

Another way is to not worry about the implications of an underlying spot price process and to model the dynamics of the futures curve by itself. This approach is taken in Eydeland and Wolyniec [35] in Chapter 2, as well as model 6.1.1 in this thesis. Although this is the approach taken in 6.1.1, the parameters were seen to be related to the spot process and so a formulation of the spot price process was achieved.

The calibration procedure for both models are similar except that when a spot process is assumed, although it gives the parameters obtained an economic intuition, it also limits the flexibility of capturing the futures curve. In this chapter we provide the implementation, for all the models in Chapter 6. Now we talk about estimating the seasonality term that we have talked about a lot in the previous chapters.

7.3 Modelling the Underlying Futures Curve

Now in natural gas futures modelling, a seasonality term is subracted from the spot price process to obtained the underlying noise. This is given by $S_t = f(t) + X_t$, all of our spot models talk about modelling X_t. Now seasonality terms are modelled in most literature so far, where f(t) is of the form

$$f(t) = \sum_{i=1}^{2} [u_i sin(2\pi r f c(t - t_i^C)) + v_i cos(2\pi r f c(t - t_i^C))] \tag{7.1}$$

where t_i^C is a centering term for the function. When looking at futures curve seasonality we shall also use the notation, $f(t,T) = f(T-t)$, since we are looking at an entire range of T's or times to expiry. So a corresponding futures model for a spot process of the form above, is $F(t,T) = f(t,T) + X(t,T)$, where $X(t,T)$ is the entire underlying futures curve.

If we know where the center is, we can do a linear regression to estimate one of the futures curves. We now describe two ways of estimating this seasonality term, $f(t)$. The first one can be used by using any of the prescribed models, but in reality many of these models do not work when looking at the futures curve for close maturities. The second one is a new way to obtain the underlying curve by Pilipovic [64].

7.4 Calibration of Futures Models

The one key phenomenon that has been overlooked by authors in past literature has been that the seasonality term $f(t)$ is determined and is used as a deterministic term in the spot process. However from our work, it has been seen that this is not true and infact the seasonality term keeps compressing itself by a constant factor, and even the relations between futures prices of different expiry change as well. This is a serious error

to overlook, and was not addressed by past papers in the subject.

In this section, we calibrate the model proposed in 6.1.2. The model is given by:

7.4.1 One Factor Models - Modelling the Futures Curve by Spot Calibration

The essence of this approach is that the futures curve should represent the expected value of the spot prices. We shall describe the steps involved with this procedure.

Step 1. Estimate seasonality, $f(t)$, from a combination of the historical futures prices and spot prices for the past T years. This procedure shall be explained in this section.

Step 2. We simply subtract the seasonality $f(t)$ from the spot prices and find the parameters θ of our spot model. That is, $S_t = f(t) + exp(X_t)$. So essentially we are trying to calibrate $X_t = ln(S_t - f(t))$.

Step 3. Take parameters θ and calculate the futures curve $F(\theta, t, T_i)$.

Step 4. Estimate the parameters of the actual futures curve by our chosen model of $F(\theta, t, T_i)$, these parameters are our risk neutral parameters. The difference in our retrieved parameters tell us the markets price of risk.

DATA For our data we used four years of spot and futures prices from January 5^{th}, 2004 until January 2^{nd}, 2008.

Step 1, which is shown first, is the same seasonality term $f(t)$ that is used for the rest of the models. So all of our calibration methods are basically for the stochastic process X_t. This was the approach taken by Pilipovic [12] and Xu [14] to model the futures curve. The advantage of this approach is that we get to experiment with models using Lévy processes as driving noise terms, and the calibration of such models is hard since explicit probability density functions are unknown. But we do all our manipulations in the frequency domain, and make use of a combination of optimization techniques with the characteristic function to calibrate these models.

7.4.2 Capturing the Seasonality

The whole point of finding the seasonality term is so that we have more data to work with. As explained earlier since there is no real spot price for gas, every month we have the futures dynamics for a certain number of months from the front month. We want to in a sense "normalize" the futures prices. In order to capture seasonality, we follow the approach followed by Xu [14]. We saw that it gives a reasonable underlying seasonality upon estimating the parameters.

In our world of discrete observations, we denote the historical spot prices by $\{S_t\}_{t=1}^n$ and the futures prices by $F(t, T_{ti}|t = 1, 2, ..., n; i = 1, 2, ..., m)$. Where $t = 1, 2..., n$ represents our observation dates and T_{ti} represents the next m contracts whose prices are available from time t. As in Xu [14], if the spot model is given by $S_t = f(t) + X_t$ and $dX_t = \alpha(L - X_t)dt + \sigma S_t^r dW_t$, where $r = 0, \frac{1}{2}, 1$ then the futures price is given by

$$F_\theta(t, T_{ti}, S_t) = e^{\alpha(t-T_{ti})}(S_t - L - f(t)) + L + f(T_{ti}) \tag{7.2}$$

and $f(t) = \sum_{i=1}^2 [u_i sin(2rfc\pi t) + v_i cos(2\pi rfct)]$.

We consider 252 trading days in a year, so $rfc = \frac{1}{252}$ above and the parameter $\theta = [\alpha, L, u_1, u_2, v_1, v_2]$. In order to estimate the parameters we find the least squares error between the observed futures curve and the proposed equation above, i.e.,

$$\theta = arg_\theta min \sum_{t=1}^{n} \sum_{i=1}^{m} (F^0(t, T_{ti}, S_t) - F(t, T_{ti}))^2. \tag{7.3}$$

The parameters estimated for α and L are given below

$$L = 2.39$$

$$\alpha = 0.0016$$

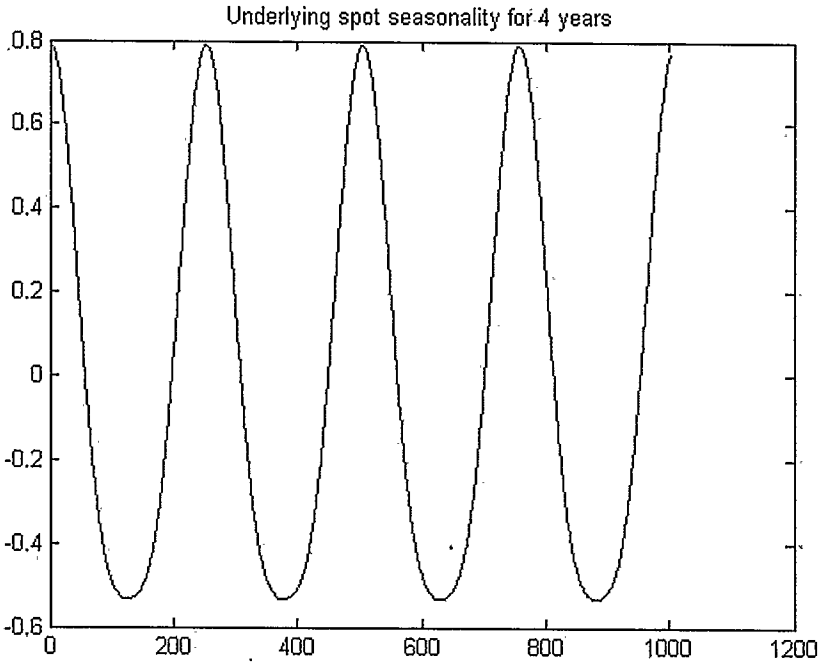

Figure 7.2: Seasonality of the spot price and its estimated artificial value

The parameters for the seasonality term which all of our models share is given below.

$$
\begin{aligned}
b &= 0.0031 \\
u_1 &= 0.6622 \\
u_2 &= 0.1284 \\
v_1 &= -0.0181 \\
v_2 &= 0.0068
\end{aligned}
$$

Seasonality simulated and estimated, i.e., our $f(t)$ is

Spot price and seasonality, i.e., S_t and $f(t)$ is shown below

The price process we want to calibrate, $X_t = log(S_t - f(t))$

Now that we have our seasonality term to calibrate our models we work with, we follow **Step 2** and our goal is to calibrate and find an appropriate model for X_t. So in each model we perform steps 2 to 4 and see what results.

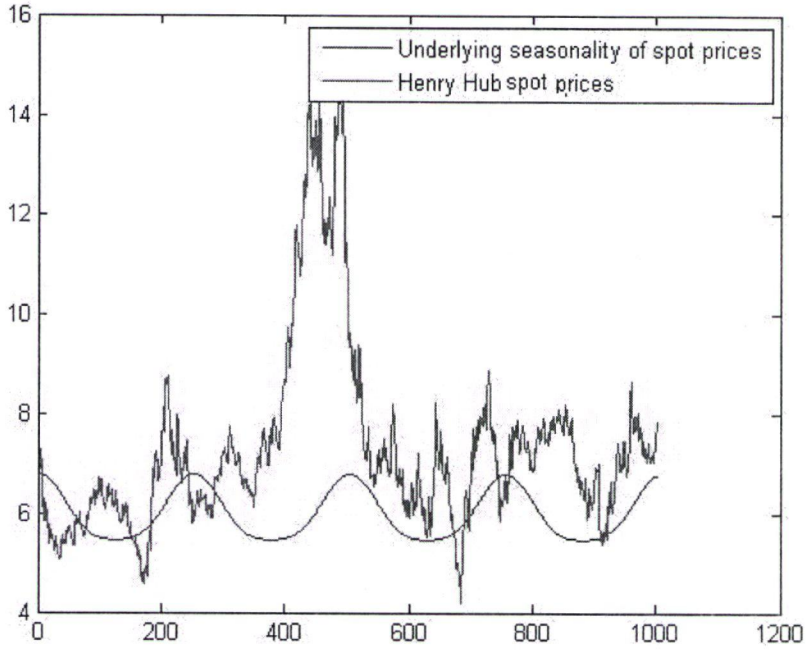

Figure 7.3: Seasonality of the spot price plotted on the same time scale as the seasonality

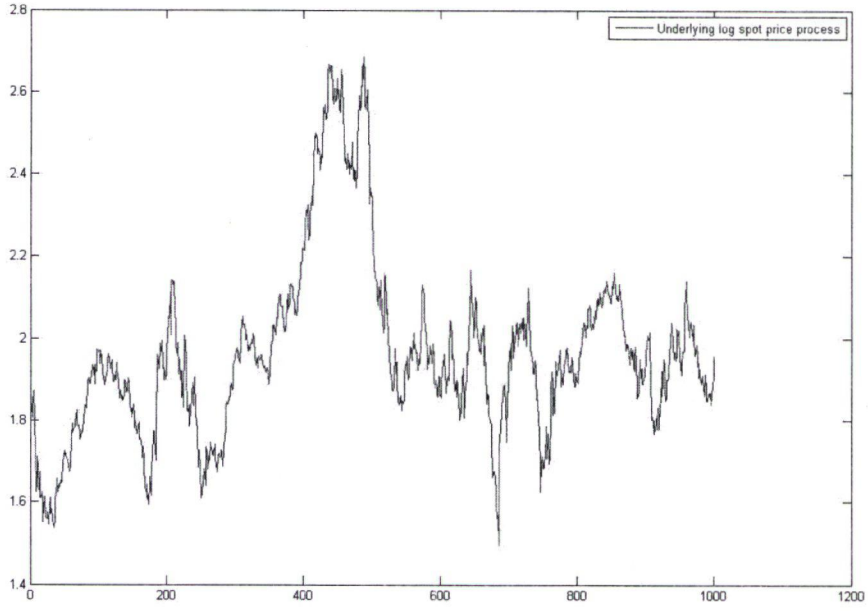

Figure 7.4: The underlying process X_t that we want to calibrate

7.4.3 Calibration of 1-'distribution'-r Type Models

As discussed in the previous chapter, for our model of type

$$dX_t = \lambda(b - X_t)dt + X_t^r dL_t$$

we have to understand that b indicates the mean level to which it reverts to. So when the equation is simplified to

$dX_t = -\lambda X_t dt + dL_t$

and the location term of L_t looks similar to $\lambda b dt$. In other words, when simulating the equation, to make the mean reversion seem reasonable we have to ensure that the "location term" is chosen for some defined mean b and a rate of mean reversion λ. So when simplifying the simulating equation from Chapter 5 in equation (6.36), and taking our observed equation to be

$X_{t+1} - X_t = \lambda - X_t h + X_t^r L_1$ and as shown in Chapter 5, we now set Y_t to take the form:

$$Y_t(\hat{\lambda}) = \frac{X_{t+1} - X_t(1 - \hat{\lambda})}{X_t^r} = L_1(\theta_L) \tag{7.4}$$

where θ_L are the parameters of the Lévy process we consider.

In other words $Y_t(\hat{\lambda}) \sim L_1(\theta_L)$, i.e., if $f(Y_t(\hat{\lambda}))$ is the probability density function of $Y_t(\hat{\lambda})$ then

$$\begin{aligned} f(Y_t(\hat{\lambda})) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iux}\varphi_{Y_t(\hat{\lambda})}(u)du \\ &= \mathcal{F}[E(exp(\psi_{Y_t(\hat{\lambda})}(u)))] \\ &= \mathcal{F}[E(exp(\psi_L(u)))] \end{aligned}$$

Here $\mathcal{F}$ is the fourier transform. Therefore we know what the probability density function of Y_t is, since the characteristic function of the Lévy process L_1 is known.

Armed with the knowledge of our density function and characteristic function, we can use two approaches to obtain the parameters.

1 Maximum Likelihood Estimation - an obvious choice and an unbiased estimator when the density is available, our goal is to set up an appropriate error function, that allows us to estimate all the parameters of the system.

2 Empirical Characteristic Function - This procedure was discussed in the Chapter 4, and it can be used to directly estimate the parameters via an established error function.

7.4.4 Maximum Likelihood Estimation of 1-factor Lévy-based Models

If we have a set of values $X_1, ..., X_n$ that belong to a distribution P with parameters $\theta = (\theta_1, ...\theta_n)$, and if the probability density function can either be quickly and accurately estimated or has an explicit form, then the maximum likelihood estimator becomes a minimum variance unbiased estimator as the sample size increases.

If the probability density function of the distribution L is given by $f(\theta, Y_t)$, then the "most likely" parameters of the observed variables are given by

$\mathcal{L}^*(\theta, \{X_t\}_{t=1}^n) = \prod_{i=1}^n f_\theta(X_i)$

The more commonly used version that we use in this thesis as well is the log likelihood function, i.e. the log of the above equation,

$$\mathcal{L}(\theta, \{X_t\}_{t=1}^n) = \sum_{i=1}^{n} ln(f_\theta, X_i) \tag{7.5}$$

Then we obtain the maximum likelihood estimator by maximizing $L(.)$,

$$\hat{\theta} = arg_0 max \mathcal{L}(\theta, \{X_t\}_{t=1}^n) \tag{7.6}$$

In our case, our observed values are X_t, but they are transformed to

$$Y_t(\hat{\lambda}) = \frac{X_{t+1} - X_t(1-\hat{\lambda})}{X_t^r} = L_1(\theta_L)$$

Since $\hat{\lambda}$ is unknown, what we do is slightly modify the estimation equation, and our parameters are $\theta = (\hat{\lambda}, \theta_L)$. Therefore we obtain our estimate for θ by maximizing,

$$\hat{\theta} = arg_\theta max \mathcal{L}(\theta_L, \{Y_t(\hat{\lambda})\}_{t=1}^n)$$
$$= arg_\theta max \mathcal{L}(\theta_L, \{\frac{X_{t+1} - X_t(1-\hat{\lambda})}{X_t^r}\}_{t=1}^n)$$

We can see in the next two sections the comparisons and efficiency of estimating the models of type 6.34 via this method of the maximum likelihood estimation.

7.4.5 Empirical Characteristic Function Method

Similar to the method in Chapters 4 and 5 for calibrating alpha-stable and NIG processes, we formulate the problem in equation 6.34. If we have 'n' observations, i.e., $\{X_t\}_{t=1}^n$

$$\begin{aligned} ecf_Y(u) &= \frac{1}{n}\sum_{t=1}^{n} exp(iuY_n(\hat{\lambda}) \\ &= \frac{1}{n}\sum_{t=1}^{n} exp(iu\frac{X_{t+1} - X_t(1-\hat{\lambda})}{X_t^r}) \end{aligned}$$

As mentioned in Chapter 4, the empirical characteristic function is approximated by taking the fourier transform of the histogram, as in section 4.3.

If our parameters are defined as in the previous section where $\theta = (\hat{\lambda}, \theta_L)$, our error function is obtained by finding the maximum of

$$\hat{\theta} = arg_\theta max || \frac{1}{n}\sum_{t=1}^{n} exp(iu\frac{X_{t+1} - X_t(1-\hat{\lambda})}{X_t^r}) - exp(\psi_X(u)) ||_2 \qquad (7.7)$$

This method as we can see, directly exploits the characteristic function and is computationally inexpensive.

7.4.6 Model 1-alpha-r

In this model in order to calibrate, we look at an equation with: $dX_t = -\lambda X_t dt + X_t^r dL_t$ with L_t being alpha-stable with parameters $L(\alpha, 0, c, \lambda k dt)$. The location term for the

Table 7.1: MLE estimation for alpha-stable-OU process with r=0

MLE estimation for alpha-stable-OU process with r=0					
Parameters	Real Values	Simulated 1000	Simulated 500	Simulated 1500	Simulated 2000
λ	.07410	.12320	.33910	.10120	.13100
α	1.73110	1.72270	1.78730	1.74120	1.73720
c	.02080	.02090	.02090	.02100	.02140
k	.05570	.06580	.14580	.05890	.05940

alpha-stable process has that form since, we want k to be what X_t mean reverts to. But for parameter stability during estimation, we set

$\mu = \lambda k dt$

and the parameters that we need to estimate are

$\theta = (\lambda, \alpha, c, \mu)$.

We calibrate for the maximum likelihood and the empirical characteristic function, their estimated parameters. These parameters are used to create simulations of the equation and then we try to retrieve these parameters again. We do this to ensure that the parameters that we have estimated are reasonably accurate and that the programs are doing their job properly. We also show sample paths for the models, and compare it to the spot price process.

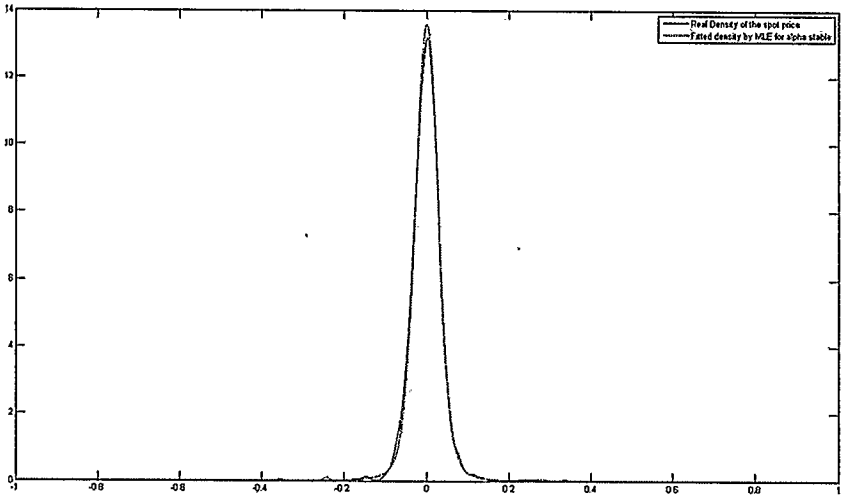

Figure 7.5: Maximum Likelihood Density estimate for model-1-alpha-0

Table 7.2: Empirical Characteristic Function(ECF) estimation for alpha-stable process with r=0

Empirical Characteristic Function(ECF) estimation for alpha-stable process with r=0					
Parameters	Real Values	Simulated 1000	Simulated 500	Simulated 1500	Simulated 2000
λ	0.0741	0.1577	0.0998	0.1104	0.1100
α	1.7311	1.7403	1.7724	1.7477	1.7349
c	0.0208	0.0208	0.0218	0.0219	0.0214
k	0.0557	0.0171	0.1011	0.0620	0.0609

In the process of calibration, the values that we find as our "real parameters" are those we satisfactorily obtained via both the ECF and MLE method. One observation we find is that the MLE estimator is not as good as the ECF in spotting spikes, as evident by observing the densities in each figure.

Now we take the "real values" in the above two tables and use it into the futures equation for the alpha-stable process, given in equation (6.46). We then look at the risk neutral values of the equation that is obtained by fitting the futures equation into the

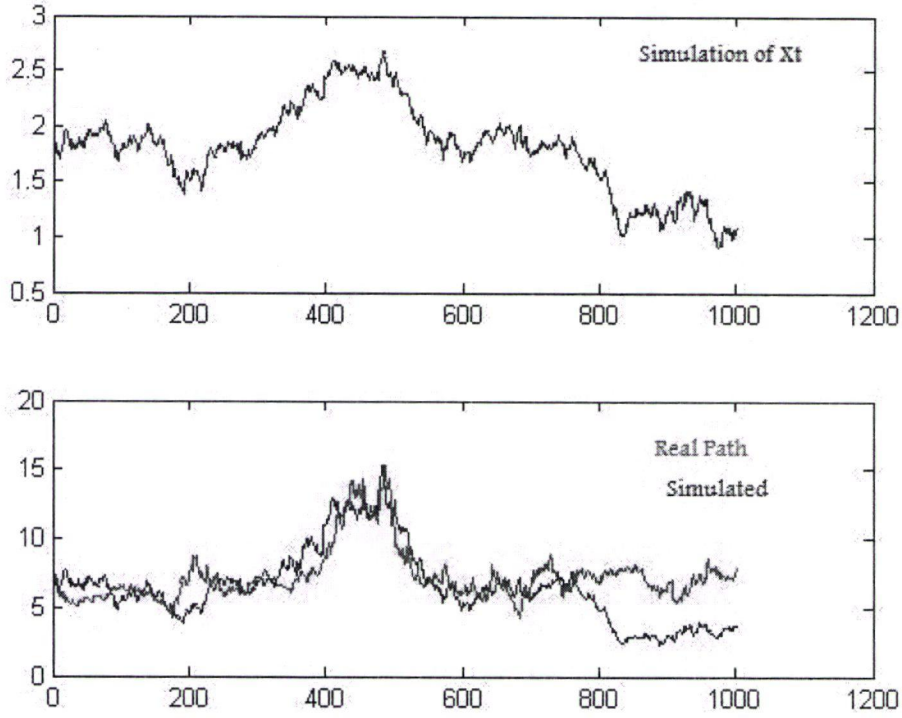

Figure 7.6: A sample path of X_t and S_t calculated underneath, in red is the true spot price process

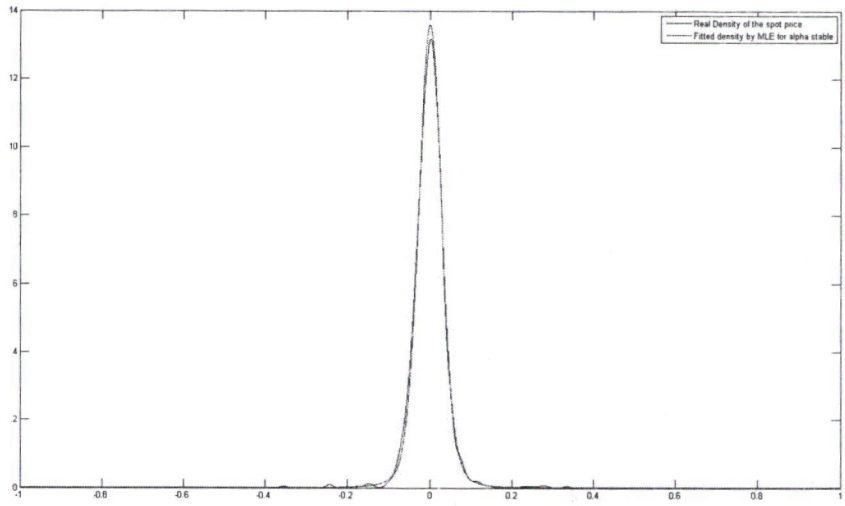

Figure 7.7: ECF density estimate for model-1-alpha-0

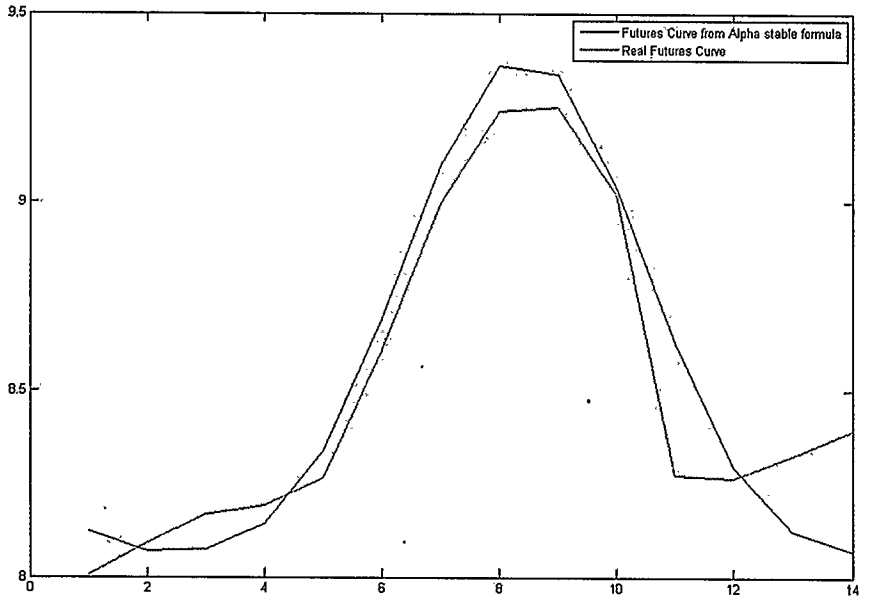

Figure 7.8: Futures curve for model-1-alpha-0 from the estimated parameters

actual futures curve.

Note that in the figure 7.8, we take contracts from the 5th month to the 16th month from the current observation date. This is because the first 4 months or so of the contract are highly volatile and do not follow seasonality rules (to an extent), especially in the last 4 years. In the case of risk neutral parameters, we see the contracts from the first month till the 15th month. The risk neutral parameters obtained are $\lambda = 0.03$, $\alpha = 1.87$, $c = 0.12$, $k = 0.06$. Now we also estimate the parameters and see sample simulations for $r = 0.5{,}1$ in equation 6.34. We notice that for $r = 1$, the process is quite erratic. In both cases we don't show the futures curve, since we don't have an explicit expression for it, but it is interesting to see the performance of calibration of non-Gaussian CIR-type and the other "exponential Lévy motion" type process.

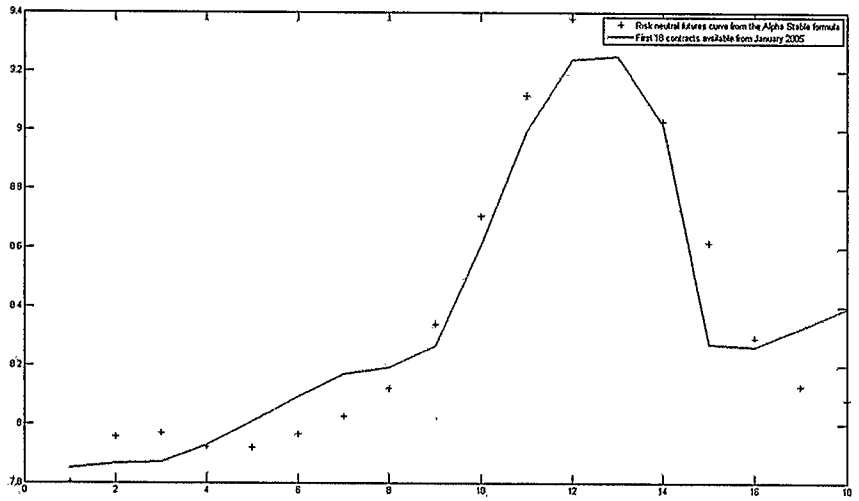

Figure 7.9: Futures curve for model-1-alpha-0 from the risk neutral parameters

Table 7.3: MLE estimation for NIG-OU process with r=0

MLE estimation for NIG-OU process with r=0					
Parameters	Real Values	Simulated 1000	Simulated 500	Simulated 1500	Simulated 2000
λ	0.580	0.641	0.172	0.455	0.610
α	31.654	33.233	26.953	37.932	34.403
β	3.744	4.746	-1.834	1.633	3.344
δ	0.039	0.037	0.035	0.046	0.038
μ	0.727	0.768	0.680	0.665	0.714

7.4.7 Model 1-NIG-r

Similar to last section, we use the same simulation equation and L_t has parameters $(\lambda, \alpha, \beta, \delta, \mu)$. Also we make, $\mu = \lambda k dt$ and the parameters that we need to estimate are $\theta = (\lambda, \alpha, \beta, \delta, \mu)$.

We show the MLE results and futures curve forecast of the NIG futures curve.

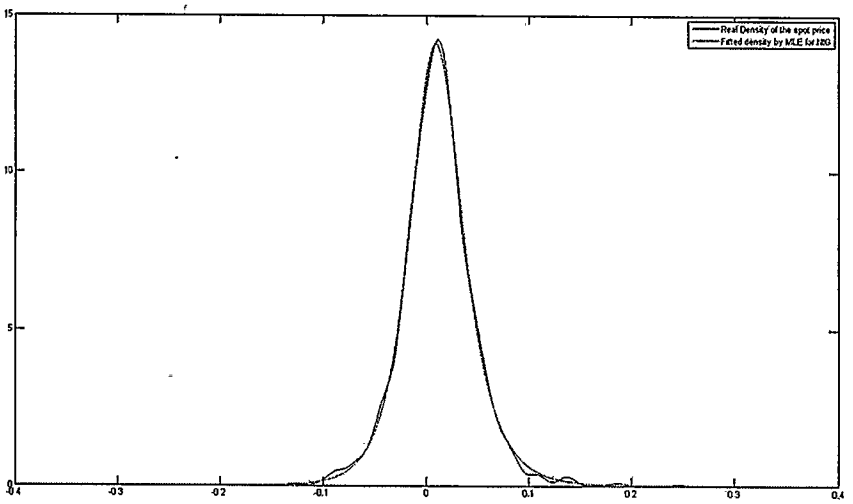

Figure 7.10: Maximum Likelihood Estimate for density of model-1-NIG-0

We once again see how the ECF method handles spikes well, infact we get to choose our intervals. One trick used in this thesis is to find first the MLE estimate and then find the "local" ECF estimate by making the interval smaller, so that the spikes in the density are given higher importance.

Model-1-NIG-0.5

Here we show the calibration results for model 9.34, for $r = 0.5$.

Table 7.4: Empirical Characteristic Function(ECF) estimation for NIG-OU process with r=0

Empirical Characteristic Function(ECF) estimation for NIG-OU process with r=0					
Parameters	Real Values	Simulated 1000	Simulated 500	Simulated 1500	Simulated 2000
λ	0.580	0.616	0.132	0.580	0.577
α	31.654	36.030	29.567	35.301	32.310
β	3.744	5.528	-1.481	2.970	3.955
δ	0.039	0.041	0.041	0.044	0.041
μ	0.727	0.508	0.598	0.890	0.729

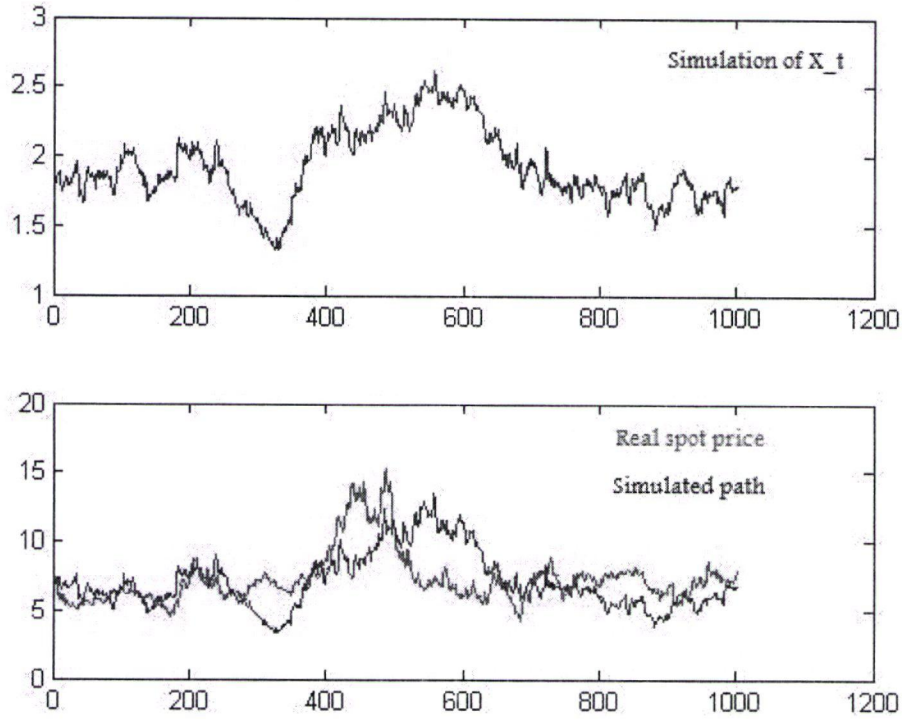

Figure 7.11: A sample path of X_t and S_t calculated underneath, in red is the true spot price process

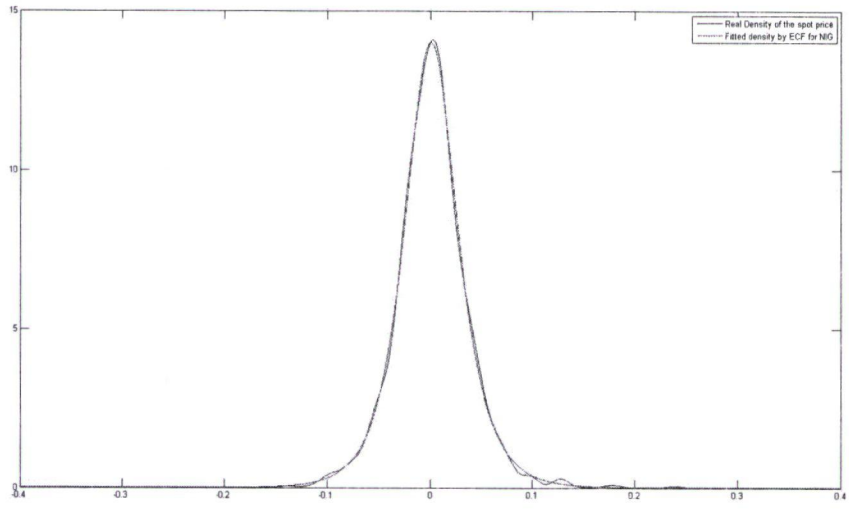

Figure 7.12: ECF density estimate for model-1-NIG-0

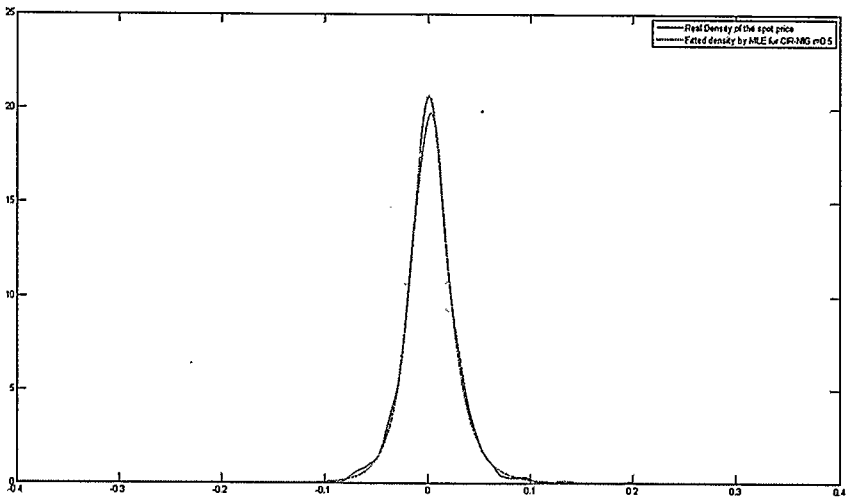

Figure 7.13: Maximum Likelihood Estimate for density of model-1-NIG-0.5

Table 7.5: Empirical Characteristic Function(ECF) estimation for NIG-CIR process with r=0.5

Empirical Characteristic Function(ECF) estimation for NIG-CIR process with r=0.5		
Parameters	Real Values	Simulated 1000
λ	0.165	0.247
α	59.188	55.000
β	4.894	4.959
δ	0.020	0.018
μ	0.003	0.004

7.4.8 Calibration of Two-Factor Models

In this section we develop a new technique to calibrate a two factor model, where only the underlying curve is taken into consideration. The advantage of this approach is that, most yield curve modelling techniques would work when trying to capture the dynamics of a pure yield curve type process.

Rather than the previous approach of having a single seasonality curve, we strip off the seasonality at every observation of the futures curve. For such a "stripping" procedure,

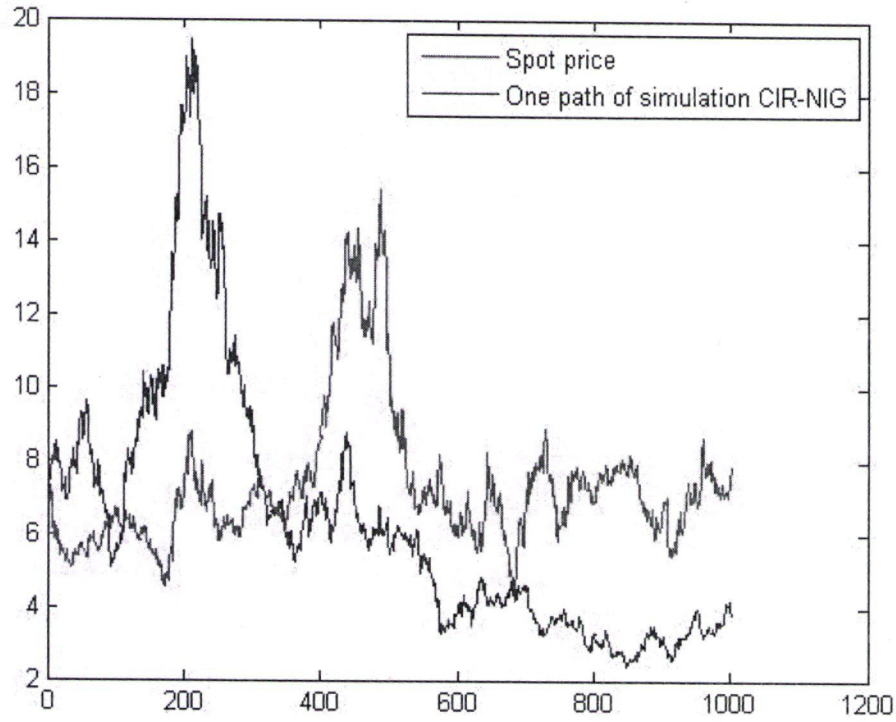

Figure 7.14: A sample path of X_t and S_t calculated underneath, in red is the true spot price process

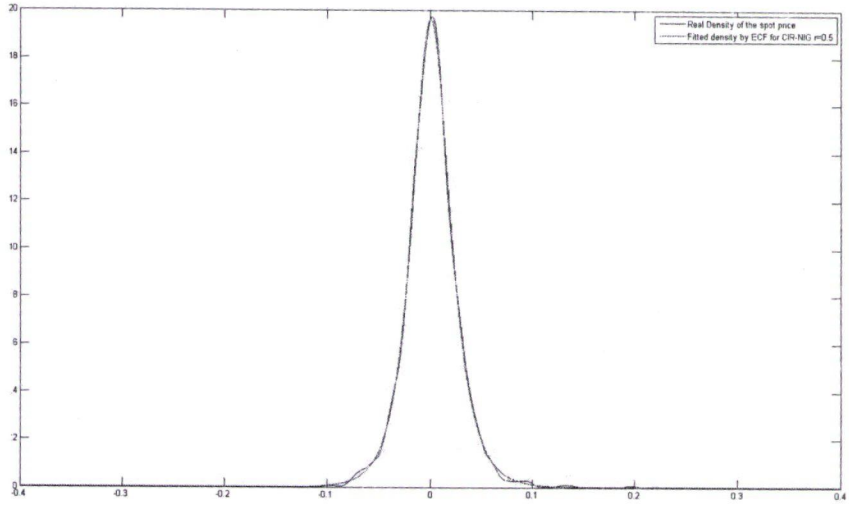

Figure 7.15: ECF density estimate for model-1-NIG-0

Table 7.6: Empirical Characteristic Function(ECF) estimation for NIG-CIR process with r=0.5

Empirical Characteristic Function(ECF) estimation for NIG-CIR process with r=0.5		
Parameters	Real Values	Simulated 1000
λ	0.165	0.284
α	59.188	51.070
β	4.894	5.705
δ	0.020	0.019
μ	0.003	0.010

we need to make sure that we have a consistent and computationally inexpensive method of determining the underlying curve. This will be truly consistent in modelling the futures curve by

$$F(t,T) = F_{UND}(t,T) + seasonality \tag{7.8}$$

The seasonality is justifiably ignored, as we are only concerned about the underlying dynamics. We now show two ways of doing such a procedure, and choose the most appropriate method of finding the underlying curve.

7.4.9 Extracting the Underlying Curve using Log Polynomials

A simple way to extract the seasonality is given by the following regression equation

$$F(t,T) = \sum_{i=1}^{2}[u_i sin(2rfc\pi((T-t)-t_i^C)))+v_i cos(2\pi rfc(T-t-t_i^C))]+\mu+\mathrm{b}log(T-t) \tag{7.9}$$

We use this as an example, to strip seasonality, and obtain parameters for $f(t,T)$. From this we also see how the underlying seasonality curve looks like, which is $f(t,T)$, as well the underlying curve itself which is $X(t,T) = \mu + \underline{l}og(T-t)$. This is estimated by the following, where $F_\theta(t,T_i)$ is the above equation and $F(t,T_i)$ actually refers to the observed futures prices, for maturities $(T_1,T_2,...,T_M)$. M is the number of months and θ represents the parameters in the above equation. Let N be the number of observations

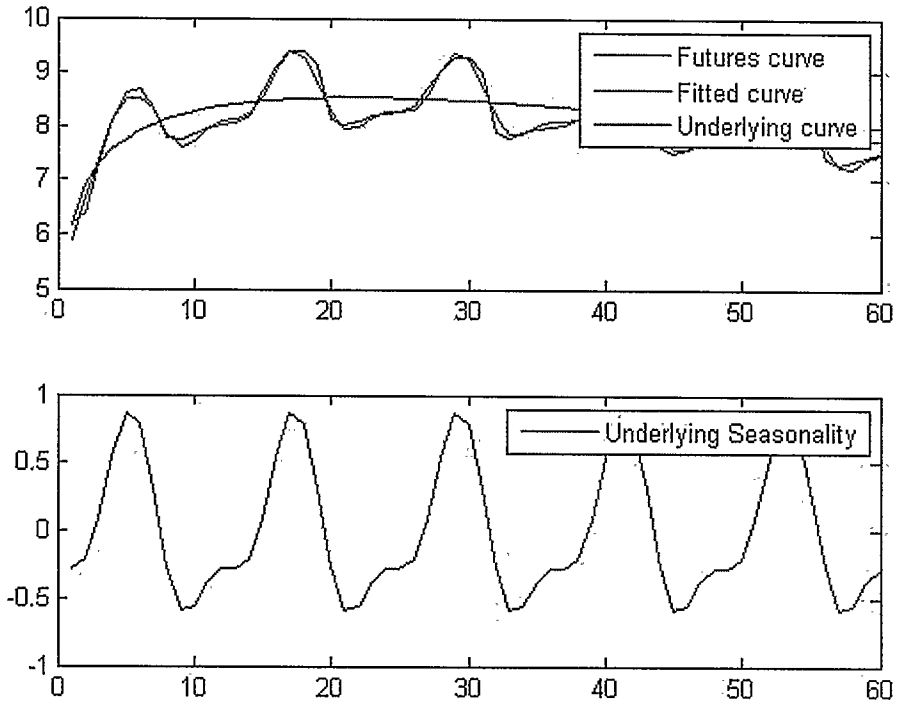

Figure 7.16: Seasonality and underlying curve using regression as of Jan 2007

over time, i.e., $t = 1, 2, 3, ..., N$

$$\hat{\theta} = argmin_{\theta} \sum_{i=}^{M} (F_{\theta}(t, T_i) - F(t, T_i))^2 \tag{7.10}$$

We can see slight errors in the actual fit, but these errors are seen to be constant over a few months, and can be used an bootstrapped during forecasting.

Calibrated parameters: $u_1 = -0.3330$, $v_1 = 0.1318$, $u_2 = -0.5271$, $v_2 = -0.2532$, $\mu = 6.2273$, $b = 1.1081$. R-squared statistic = 0.9629, F-statistic = 229.5818

In place of the equation $F(t, T)$ above, we can use any of the futures models in this thesis in practice, but its actual abilities (with respect to parameter convergence) are not very consistent. Regression gives us a reliable, convergent scheme to obtain the underlying curve.

7.4.10 Extracting the Underlying Curve using Exponential Seasonality

Pilipovic [34] came up with a slick way to estimate the underlying curve, although the problem is that fitting the underlying curve with an existing model is hard, but could be useful for traders to know the change in curves over different days. If $F(t,T) = f(t,T) + X(t,T)$ is our futures equation, in the previous subsection we directly estimated $F(t,T)$ and recovered $f(t,T)$. Over here the equation below estimates $X(t,T)$, and we estimate the seasonality term by $f(t,T) = F(t,T) + X(t,T)$. The equation is of the following form

$$\begin{aligned} X(t,T) &= \beta_1 exp(-\gamma_1(rfc((T-t) - t_1^C))^2) + \beta_2 exp(-\gamma_2(rfc((T-t) - t_2^C)).^2) \\ &\quad + \beta_3 exp(-\gamma_3(rfc((T-t) - t_3^C))^2)); \end{aligned}$$

Again we show what results when implementing the equation above, and see a satisfactory underlying curve.

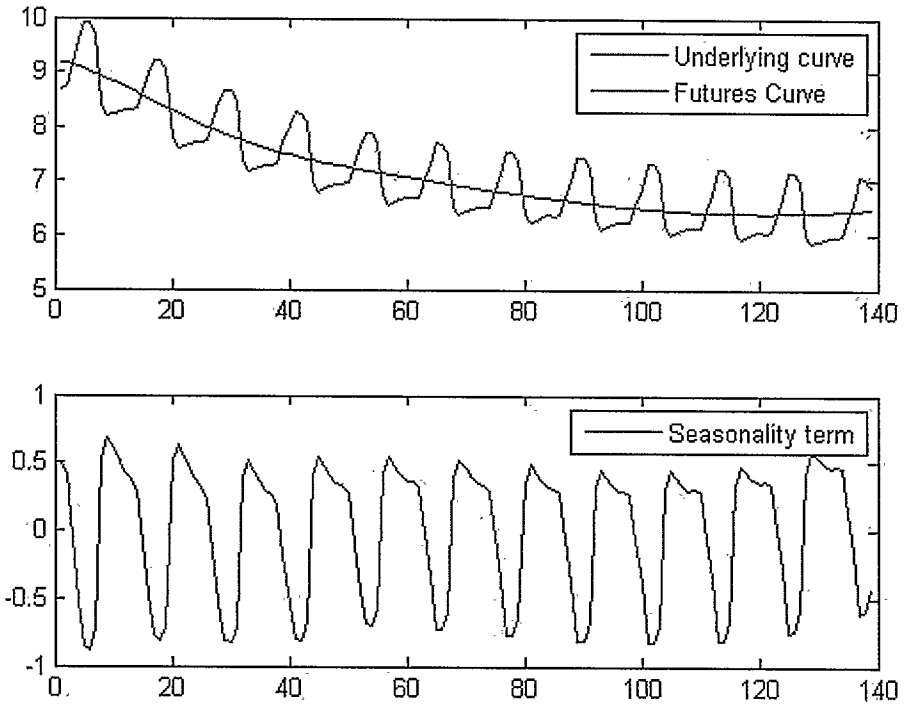

Figure 7.17: Seasonality and underlying curve using exponential seasonality as of Jan 2005

Calibrated parameters: $\beta_1 = 3.9031$, $\gamma_1 = 0.0099$, $t_1^C =$ -5.1825

$\beta_2 = 3.8450$, $\gamma_2 =$ -0.0025, $t_2^C =$ -5.1756

$\beta_3 = 1.5332$, $\gamma_3 = 0.1846$, $t_3^C =$ -5.4838

But the seasonality term recovered isn't consistent, and its structure seems to change well beyond our comfort. Crucial to futures modelling is also to be able to capture the underlying seasonality term, that tells us the relative prices of all contracts along the years. If it is not attainable easily, then it could not be used for a scientific trading strategy.

7.4.11 Extracting the Underlying Curve using a modified Nelson-Siegel Yield Curve Model

In the previous subsections we have seen what the underlying curve looks like. We borrow literature in yield curve modelling, like that of Nelson-Siegel [36], to see if it performs well at fitting the underlying curve when seasonality terms are added to it. To our surprise, it seemed to give the best picture of the underlying yield curve. The following is the equation that defines it

$$X(\tau_i) = l + s(\frac{1 - e^{-\lambda\tau_i}}{\lambda\tau_i}) + c(\frac{1 - e^{-\lambda\tau_i}}{\lambda\tau_i} - e^{-\lambda\tau_i}) + v(\tau_i) \tag{7.11}$$

where $\tau_i = T - t$, and l, s, c, v are parameters. Now Nelson-Siegel suggests to use $\lambda = 0.0609$, so that the equation is now be estimation via regression. So we use this approach and set the forward curve to equal

$$F(t, T_i) = \sum_{i=1}^{2}[u_i sin(2rfc\pi(T - t) + v_i cos(2\pi rfc(T - t)] + X(T_i - t) \tag{7.12}$$

The following diagram shows the underlying curve as fit by the Nelson-Siegel with seasonality procedure.

7.4.12 Modelling the Underlying Futures Dynamics

We take the two factor model of Xu [14], but use a very different calibration technique, since we are only modelling the underlying futures curve. Now we use the regression based method introduced in this chapter to strip off the seasonality. Our futures equation is supposed to model the underlying curve, which we denote for brevity as F', since we exclusively working with this curve. Once again we consider m futures contracts, give by $F'(t, T_{ti})$ for $i = 1, ..., m$.

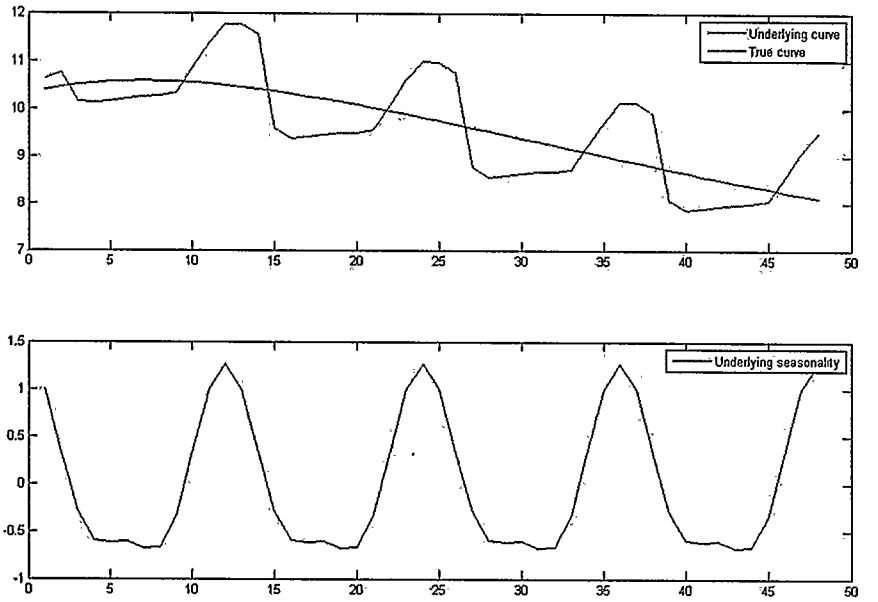

Figure 7.18: Underlying curve using Nelson-Siegel yield curve model with seasonality as of Jan 2006

$$dS_t = \tilde{\alpha}(L_t - S_t)dt + \sigma S_t^{\tau} dW_t^1$$

$$dL_t = \tilde{\mu}(\tilde{\gamma} - L_t)dt + \tau L_t^{\tau} dW_t^2$$

So the theoretical futures price is given by,

$F'^{\tilde{\theta}}(t, T_{ti}, S_t, L_t) = A_i + B_i L_t$

where

$A_i = e^{\tilde{\alpha}(t-T)} S_t + \frac{\tilde{\mu}\tilde{\gamma}}{\tilde{\alpha}-\tilde{\mu}}(e^{\tilde{\alpha}(t-T)} - 1) - \frac{\tilde{\alpha}\tilde{\gamma}}{\tilde{\alpha}-\tilde{\gamma}}(e^{\tilde{\mu}(t-T)} - 1)$

and

$B_i = \frac{\tilde{\alpha}}{\tilde{\alpha}-\tilde{\mu}}(e^{\tilde{\mu}(t-T)} - e^{\tilde{\mu}(t-T)})$

Filtering the Long Run Mean L_t from the Futures Prices

Now our set of parameters that we have are $\tilde{\theta} = [\tilde{\alpha}, \tilde{\mu}, \tilde{\gamma}, \tau, \sigma]$. As we can see in the above theoretical futures price equation, the parameters only depend on $\tilde{\alpha}, \tilde{\mu}, \tilde{\gamma}$. Now for

a set of parameters, there is a unique L_t that is hidden in our data. This can be obtained by finding the minimum of the following.

$$L_t(\tilde{\theta}) = argmin_{L_t} \sum_{i=1}^{m} (F'^{\tilde{\theta}}(t, T_{ti}, S_t, L_t) - F'(t, T_{ti}))^2 \tag{7.13}$$

Substituting our equation $F'^{\tilde{\theta}}(t, T_{ti}, S_t, L_t)$ in the above, and realizing that we are simply calculating the distance in euclidian space, $\mathbb{R}^m$. We obtain

$$\begin{aligned} L_t(\tilde{\theta}) &= argmin_{L_t} ||A + BL_t - F'(t,T)||_2 \\ &= argmin_{L_t} ||A - F(t,T) + BL_t||_2 \\ &= \frac{||BF'(t,T) - AB||_1}{||B||_2} \end{aligned}$$

In our setting this is equivalent to,

$$L_t(\tilde{\theta}) = \frac{\sum_{i=1}^{m} (B_i F'(t, T_{ti}) - A_i B_i)}{\sum_{i=1}^{m} B_i^2} \tag{7.14}$$

So we recover our L_t values by minimizing the following equation, with the above equation of L_t replaced in our theoretical futures equation.

$$\tilde{\theta} = argmin_{\tilde{\theta}} \sum_{t=1}^{n} \sum_{i=1}^{m} (F'^{\tilde{\theta}}(t, T_{ti}, S_t, L_t) - F'(t, T_{ti}))^2 \tag{7.15}$$

We find the values of the parameters $\tilde{\theta}$ and plug it into equation (7.13). These parameters are our risk neutral parameters.

We see that in both cases the risk neutral parameters are very close to each other, but the overal error gets minimized the most in the Nelson-Siegel type yield curve model. So we choose it as our risk neutral parameters, and from now onwards take the underlying curve observed by the market as the Nelson-Siegel curve.

Now using the parameters above, we calculate the long run mean L_t. Note that there is a very negligible difference in the long run plot irrespective of the model we choose

Table 7.7: Risk Neutral values for Futures curve, using Nelson-Siegel and Log Polynomials

Risk Neutral values for Futures curve, using Nelson-Siegel and Log Polynomials		
Parameters	Nelson-Siegel	Log Polynomials
$\tilde{\alpha}$	0.005885	0.0064
$\tilde{\mu}$	3.2497	3.2775
$\tilde{\gamma}$	0.000637	0.0006
errors	0.3864	0.4401

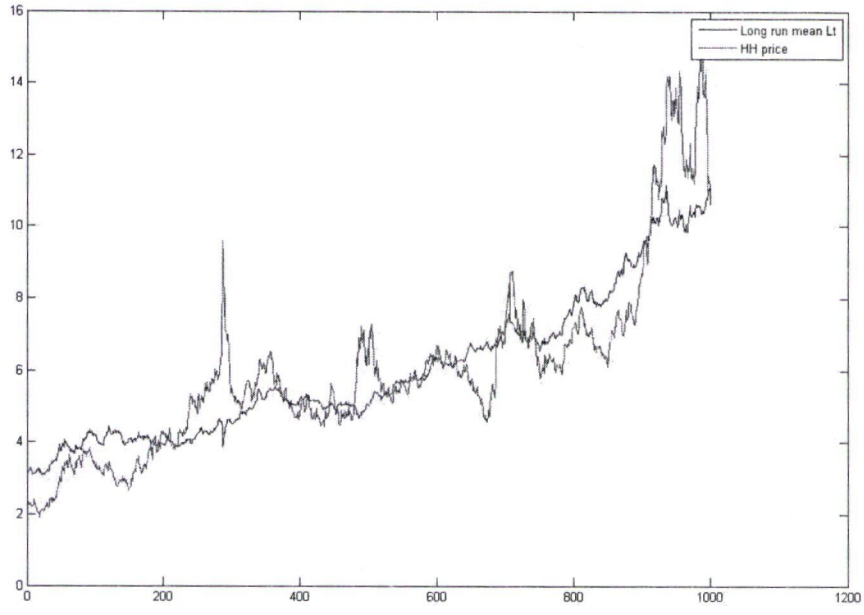

Figure 7.19: Long Run Mean Filtered vs the Spot Price

but the difference is in forecasting the yield curve. The data used in our two factor calibrations are that of January 2002 to January 2006.

We can see the long run mean filtered really does seem to "lead the evolution" of the spot price, as observed by Xu [14]. This is consistent with his observations of the long run mean behaviour. The parameters obtained here seem meaningful so far, surprisingly, γ is our rate of mean reversion and μ the mean that the long run mean reverts to. We can see an upward drift, and a weak mean reversion for the long run mean process is seen in the parameter $\gamma = 0.0006$ for $\mu = 3.24$. Now for the big test to see, how close are our parameters in the real world are with respect to the risk neutral parameters calculated above.

7.4.13 Parameter estimation of two factor model

The two factor model can be estimated via the simulation equation, that is given by

$$S_{t+1} - S_t = \alpha(L_t - S_t) + \sigma S_t^r (W_{t+1}^1 - W_t^1) \tag{7.16}$$

$$L_{t+1} - L_t = \mu(\gamma - L_t) + \tau L_t^r (W_{t+1}^2 - W_t^2) \tag{7.17}$$

Now $W_{t+1}^i - W_t^i \sim N(0,1)$, so the probability densities for each S_t and L_t, can easily be calculated. If θ are our parameters then,

$$f(\theta, [S_{t+1}, L_{t+1}]|[S_t, L_t]) = f(\theta, S_{t+1}[S_t, L_t]) f(\theta, L_{t+1}|L_t)$$

From the simulation equation, we have

$$f(\theta, S_{t+1}[S_t, L_t]) \sim N(S_{t+1} - S_t - \alpha(L_t - S_t), \sigma S_t^r) \tag{7.18}$$

and

$$f(\theta, L_{t+1}|L_t) \sim N(L_{t+1} - L_t - \mu(\gamma - L_t), \tau L_t^r) \tag{7.19}$$

Table 7.8: Estimated Parameters of the Two-Factor Model

Parameters	Sim. Mean	Sim. Std. Dev.	Real Est.
α	0.0295	0.0085	0.0262
σ	0.2588	0.0056	0.2581
μ	-0.0017	0	-0.0017
γ	1.3147	0	1.3147
τ	0.0764	0	0.0764

So since our conditional density function is known, we can easily use maximum likelihood to estimate the parameters. Xu [14] figured out explicit expressions for each of these parameters via setting the partial derivatives of the parameters to zero through the likelihood function.

In this thesis we only show results for r = 0, where both factors are OU processes. The other cases of 'r' are just as easy to calibrate. Below are the parameters recovered

As we can see above the parameters are recovered well (since the standard deviations of the simulated values are small), this was done assuming the L_t that we filtered from the Futures curve, and simulating values of S_t. Below are sample paths with the above parameters, with L_t fixed,

As we can see, the dynamics of the spot price are really very nicely captured. Now we look at the futures curve.

7.4.14 Fitting the Futures Curve from the risk neutral parameters

Using the risk neutral parameters we obtained from Table 7.7, we simulate the futures curve. The long run mean successfully tells us if the curve is in contango or backwardation, but then again this value is obtained from the futures curve.

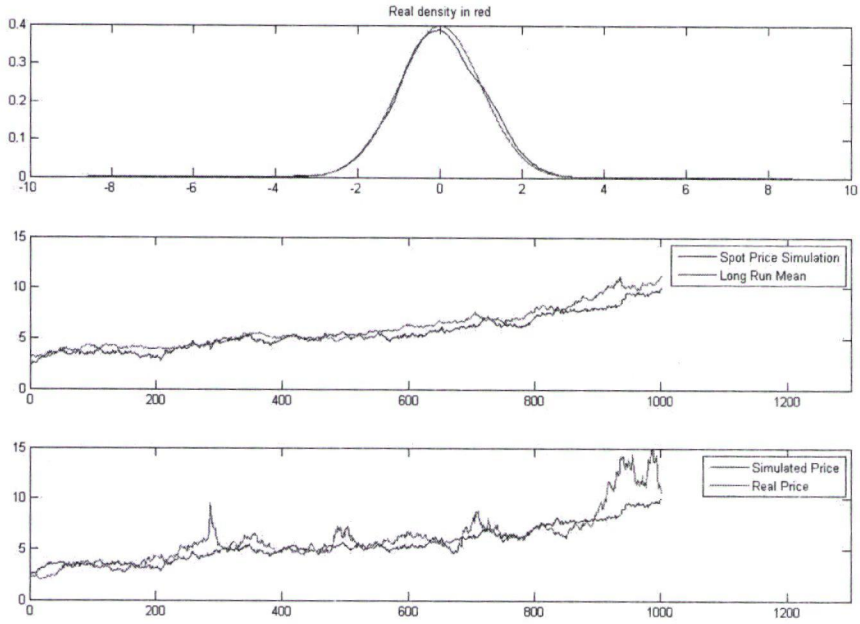

Figure 7.20: Two Factor Simulations from Long Run Mean L_t

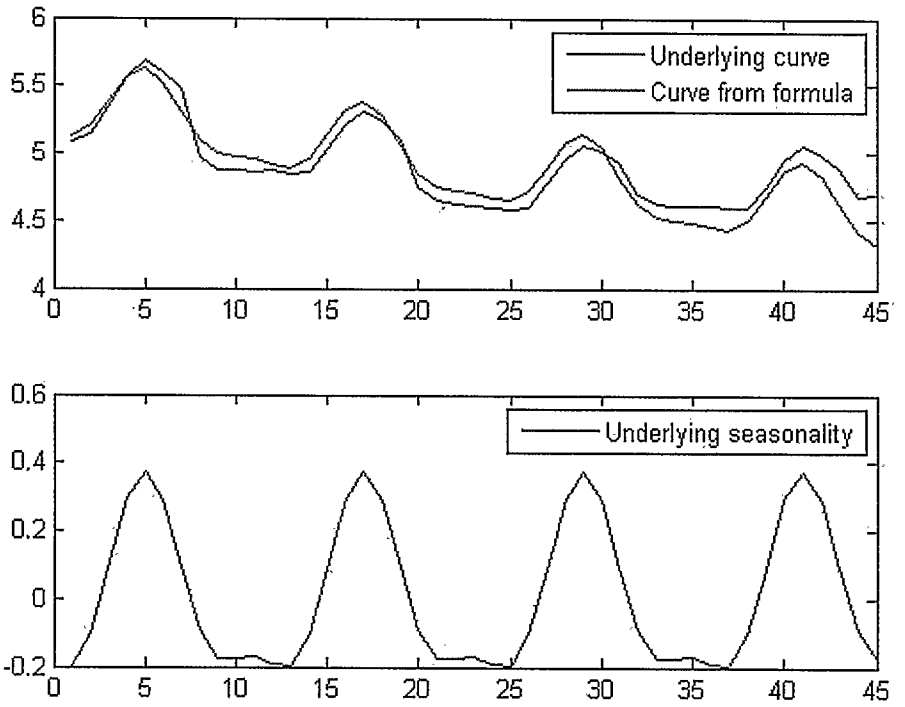

Figure 7.21: Futures Curve from Two-Factor Model in Backwardation from Formula

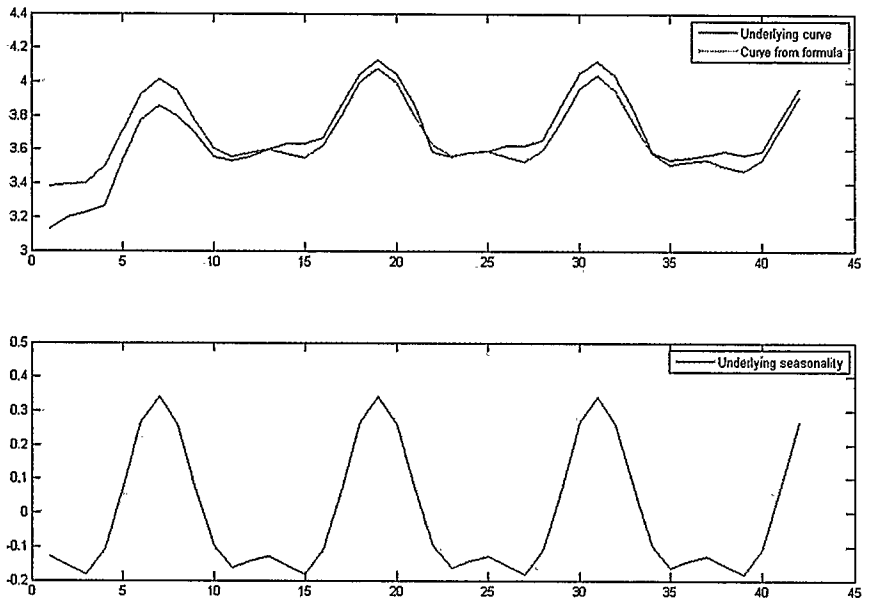

Figure 7.22: Futures Curve from Two-Factor Model in Contango from Formula

Chapter 8

Conclusion

One of the aims of this thesis was to introduce Levy processes into spot modelling of gas prices, we concluded that the one factor OU process with the NIG process was the most effective in the class of one factor models. Even though we did not price options in this thesis, it is straightforward to use Monte Carlo methods to price options irrespective of our process choice. Levy based one factor models can also have their options priced via the fast Fourier transform method of Carr and Madan [37], since we know the explicit characteristic function. We also showed how the empirical characteristic function method of estimating these models work well when there are centered spikes in the density function.

We then also showed different ways of stripping off seasonality from the spot prices, and used the Nelson-Siegel yield curve function to obtain the underlying seasonality. This method of estimation of a two factor gas futures model is different from other literature. Rather than trying to find meaning in the yield curve from Nelson-Siegel directly, we instead look at the implications of those curve shapes by calibrating it for the two factor model. This also gives us an economic intuition for each of the parameters obtained in our estimation.

Modelling gas futures is extremely hard, and over the last two to three years, the market dynamics have become a lot more erratic. The modelling techniques developed in this thesis does work until 2006, after which was the hurricane Katrina, which was enough of a large period to skew data during parameter estimation. It is in the authors belief that it would not be possible to do so without considering the futures price also as

a function of storage levels.

Hopefully this thesis does justice to futures modelling in gas, without having to deal with modelling time spreads seperately and establishing the relationship between different contracts in such a manner. Natural gas is among the worlds most volatile markets, and it would be expected that newer models would have to be developed routinely in order to accomodate for the complexity in market behaviour.

References

[1] M. Thompson, M. Davison, and H. Rasmussen. Natural gas storage valuation and optimization: a real options application. Naval Research Logistics, July 2003. Submitted.

[2] H. Ahn, A. Danilova, and G. Swindle. Storing ARB. Wilmott Magazine, 1, 2002.

[3] S. G. Bjerksund, P. and F. Vagstad. Gas storage valuation: Price modelling v. optimization methods. Discussion Paper 2008/2, NHH Dept. of Finance Management Science, October 2008.

[4] A. Boogert and C. de Jong. Gas storage valuation using a Monte Carlo method. Birkbeck Working Papers in Economics and Finance 0704, Birkbeck, School of Economics, Mathematics and Statistics, January 2007.

[5] S. Borovkova and H. Geman. Seasonal and stochastic effects in commodity forward curves. Review of Derivatives Research, 9(2):167186, 2006.

[6] R. Carmona and M. Ludkovski. Gas storage and supply guarantees: An optimal switching approach. Submitted to Management Science.

[7] R. Carmona and M. Ludkovski. Spot convenience yield models for the energy markets. In G. Yin and Y. Zhang, editors, AMS Mathematics of Finance, volume 351 of Contemporary Mathematics, pages 6580, 2004.

[8] Z. Chen and P. A. Forsyth. A semi-Lagrangian approach for natural gas storage valuation and optimal operation. Submitted to Quantitative Finance, 2007.

[9] L. Clewlow and C. Strickland. Energy Derivatives. Lacima Group, 2000.

[10] R. Gibson and E. S. Schwartz. Stochastic convenience yield and the pricing of oil contingent claims. Journal of Finance, 45(3):959976, 1990.

[11] J. Gray and P. Khandelwal. Towards a realistic gas storage model. Commodities

Now, June 2004.

[12] D. Pilipovic. Energy Risk: Valuing and Managing Energy Derivatives. McGraw-Hill, 1997.

[13] E. S. Schwartz. The stochastic behavior of commodity prices: implications for valuation and hedging. Journal of Finance, 52(3):923973, July 1997.

[14] J. Xu. Stochastic models for gas prices. Masters thesis, University of Calgary, 2004.

[15] Z. Chen and P. Forsyth, Implications of a Regime-Switching Model on Natural Gas Storage Valuation and Optimal Operation. September 13, 2007.

[16] Yun Li. Natural Gas Storage evaluation. Masters thesis, Geogia Tech. 2006.

[17] B. Bringedal. Valuation of Gas Storage, a real options approach. NTNU, 2003.

[18] G.E.Uhlenbeck and L.S.Ornstein: "On the theory of Brownian Motion", Phys.Rev. 36:823-41, 1930.

[19] D. Applebaum. Levy processes and stochastic calculus. 2004

[20] E. Parzen (1962). On estimation of a probability density function and mode, Ann. Math. Stat. 33, pp. 10651076.

[21] B. W. Silverman, Density Estimation for Statistics and Data Analysis, Chapman and Hall, 1986.

[22] B. W. Silverman,Kernel Density Estimation using the Fast Fourier Transform, Applied Statistics, Royal Statistical Society, (1982), Vol. 33.

[23] R. S. Pindyck. The dynamics of commodity spot and futures markets: a primer. The Energy Journal, 22(3):129, 2001.

[24] Cont and Tankov. Financial modelling with Jump Processes.

[25] Barndorff-Nielsen, O. E. (1977). Exponentially decreasing distributions for the logarithm of particle size. Proceedings of the Royal Society of London A, 353:401419.

[26] O. E. Barndorff-Nielsen, Thomas Mikosch, Sidney I. Resnick. Lévy processes

[27] O.E. Barndorff-Nielsen, J. Kent and M. Srensen (1982): Normal variance-mean mixtures and z-distributions.International Statistical Review, 50, pp. 145 - 159.

[28] A. Cartea and T. Williams. UK gas markets: The market price of risk and applications to multiple interruptible supply contracts. Birkbeck Working Papers in Economics and Finance BWPEF 0608, Birkbeck School of Economics, Mathematics and Statistics, 2006.

[29] Z. Chen and P. A. Forsyth. Stochastic models of natural gas prices and applications to natural gas storage valuation. Working paper.

[30] Canadian Natural Gas 2005. Published by Natural Resources Canada. http://www2.nrcan.gc.ca/es/erb/prb/english/View.asp?x=447, 2005.

[31] W.J. Runggaldier, Estimation via stochastic Fltering in Financial market models, Preprint, 2003.

[32] S. Hikspoors and S. Jaimungal. Energy Spot Price Models and Spread Options Pricing. [IJTAF, Vol.10, Nov.2007, pp.1111-1135]

[33] US Department of Energy. http://www.energy.gov/energysources/naturalgas.htm

[34] D. Pilipovic. Energy Risk: Valuing and Managing Energy Derivatives. McGraw-Hill, 2007.

[35] A. Eydeland and K. Wolyniec (2003). Energy and Power Risk Management. New Developments in Modeling Pricing and Hedging. Wiley.

[36] C.R. Nelson and A.F. Siegel (1987), "Parsimonious Modeling of Yield Curves," Journal of Business, 60, 473-489. [37] P. Carr, D. B. Madan (1998), "Option Valuation Using the Fast Fourier Transform," Journal of Computational Finance 2, 61-73.