MODELING ENERGY SPOT MARKET AND PRICING ENERGY DERIVATIVES: A TECHNICAL ANALYSIS

by

Samir Masood Sheikh

BSc Mathematics, Lahore University of Management Sciences ,

2005

Submitted to the Graduate Faculty of

the Department of Mathematics in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2007

UNIVERSITY OF PITTSBURGH

MATHEMATICS DEPARTMENT

This thesis was presented

by

Samir Masood Sheikh

It was defended on

April 23rd, 2007

and approved by

Dr. John Chadam, Professor

Dr. Xinfu Chen, Professor.

Dr. William Layton, Professor.

Thesis Advisor: Dr. John Chadam, Professor

John Chadam

MODELING ENERGY SPOT MARKET AND PRICING ENERGY DERIVATIVES: A TECHNICAL ANALYSIS

Samir Masood Sheikh, M.S.

University of Pittsburgh, 2007

A data driven approach is utilized to model the energy spot prices using mean reverting diffusion processes with jumps. Initially, the Ornstein Uhlenbeck model is considered to calibrate the parameters using the data without incorporating jumps. After the calibration, a technical analysis of the jump magnitudes is carried out and accordingly a jump term, whose magnitudes are log-normally distributed with the rate of occurrence following a Poisson process, is incorporated into the model. Alternatively, some non-parametric statistics is also employed to analyze the jump process. Finally, an explicit closed-form equation for the price of a forward on energy spot prices is derived and prices are calculated numerically for different times to expiry.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to thank my thesis advisor, Dr John Chadam for his generous help and cooperation throughout my research work. I feel very lucky to work under his supervision and learn through his experience and knowledge in the field of Mathematical Finance. I feel greatly indebted for his remarkable guidance and I feel dignified to admit that without his help, this would not have been possible.

Here, I would also like to thank my peers Qi Mi, Jai Ern and Poning for their considerate feedback and help in some detailed arithmetic calculations and numerical simulations for verification of results.

Lastly, I want to thank my parents, family, friends and my wife for their encouragement and support throughout my academic career.

1.0 INTRODUCTION

Over the years, there has been a profound difference observed between commodity markets and financial markets. Unlike the financial markets, for example the stock market, electricity markets exhibit some intrinsic characteristics that owe their origin to physical and economical constraints such as dynamics of supply and demand, volume constraints on transfer and storage and the local nature of the market. Consequently, the classical approach of assuming log-normal distribution of the asset is not applicable in modeling electricity spot prices and neither do the standard methods for option pricing[1] can be applied directly [1].

The key aspect in a competitive market like electricity, is 'deregulation'. In the year 1996, electricity prices were deregulated resulting in a large inflow of new retail electricity providers, thereby increasing the level of competition existent in the prevailing contemporary electricity production industry at that time. This emergence of new suppliers has continued for the past ten years which accounts for the large volume of the industry today that has led to a significant increase in the spread of fluctuations in the prices creating major concerns for both, the consumers and the producers. Consequently, strong emphasis in the form of research and investment has been observed in structuring the ideal energy derivative[2] securities to "hedge" and isolate the financial risk involved in order to off set the potential of a massive monetary loss, threatening enough to end up in an eventual bankruptcy. As a consequence, formulating accurate models for predicting electricity spot prices has become an emerging challenge.

[1]The most widely used methods for pricing contingent claims on financial assets are the Black-Scholes option pricing formula, the Binomial Tree Pricing method and Monte-Carlo Simulations.

[2]A derivative is an instrument whose price depends on, or is derived from, the price of another asset, called the 'underlying'.

There are two distinctive features observed in energy markets and are specially evident in electricity markets: the mean reverting behavior of spot prices and the existence of jumps or spikes. This anomalous behavior, with respect to other financial markets is attributed to the fact that electricity, in particular, is an extremely difficult and as a consequence, a very expensive asset to store. Hence, markets need to be adjusted and kept in balance on a second-by-second basis. Moreover, as indicated earlier, electricity prices are heavily dependent on physical and economic factors. Due to these prominent differences in the nature of the two markets, various approaches have been employed by researchers over the years to model the electricity prices. These approaches classify into two types of models: spot based models and forward-based models, however, in my thesis I present a spot electricity model for the province of Alberta.[3].

As a consequence of the nature of the market, different approaches have been employed by researchers to model the electricity spot market. Some academicians, like M Davison , C. Marcus and B. Anderson are strong proponents for "hybrid" models[4], which include physical factors like demand and supply, where the ratio of demand to supply is calculated and a certain threshhold is predefined using historical data. If the ratio exceeds the threshold, the prices are expected to sky rocket due to excess demand. On the contrary, if the ratio is lower, standard mean reverting diffusion processes are used [2]. A significantly different approach is adopted by A. Lavassani , A. Sadeghi and A. Ware in [1], where they present single-factor, two-factor and multi-factor models. However, it is believed that this approach, although theoretically strong, has little practical implications. In my thesis work, I adopt a data driven: technical approach, to model the electricity spot market with theoretical validity.

[3]The data was collected from the Industrial Problem Solving Workshop (IPSW) I attended in Vancouver, Canada in June 2006

[4]Also known as regime switching models

2.0 MODEL FOR ENERGY SPOT PRICES

In this chapter, at first an intial model for electricity spot prices is formulated after careful inspection into the data set. Secondly, the parameters for the model are estimated using statistical methods, namely, **Regression** and **Maximum Likelihood Estimation** . Finally, a technical analysis is carried out to calibrate the spikes observed in the data.

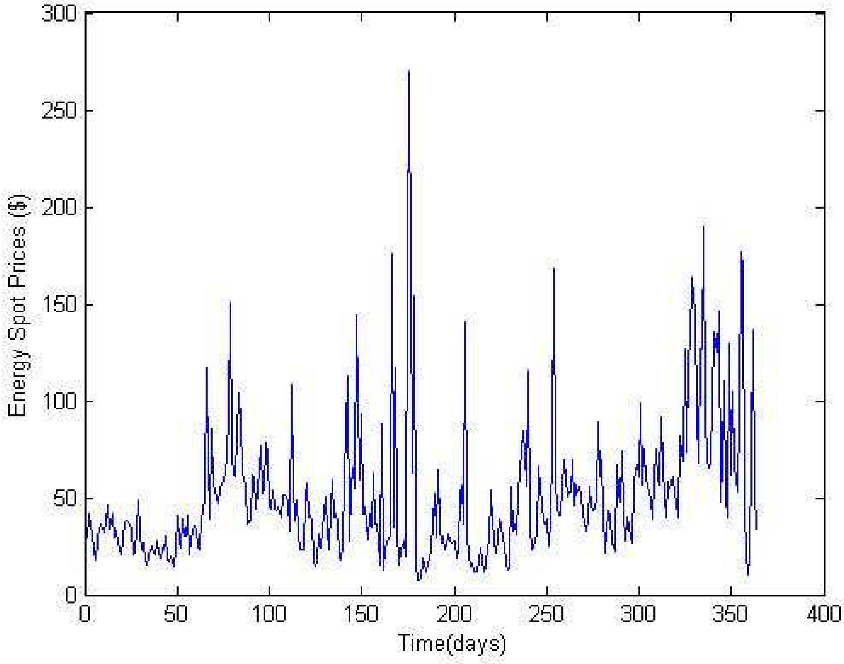
2.1 DATA ANALYSIS

The fundamental approach for determining a suitable model for any data set is to "observe" the data graphically. The main purpose of this procedure is to deduce any periodicity, increasing/decreasing trends and other special features hidden inside the data, which would help in finding the right "ingredients" for constructing a model. In case of electricity markets, as discussed earlier, it is believed that there is a strong evidence for mean reversion with occasional spikes in spot prices, which in general are much more pronounced than in stock markets. However, unlike the stock prices, energy prices tend to revert back a significant amount the next immediate trading day. This special characteristic is commonly observed in energy markets unlike any other consumption asset. This behavior is attributed to the fact that energy is an expensive asset for storage purposes. Therefore, whenever there is an excess demand, resulting in a jump in the price, there are always enough suppliers who are willing to sell causing the prices to fall immediately. This reverting behavior can be observed by simple inspection of the data in both markets. For an initial survey I took on year's on-peak [1] data, where each day's data value was calculated to be the averaged hourly

[1] The on-peak price data refers to electricity prices during the trading hours

energy price for the whole day for the peak hours (See Fig 2.1 below). The justification for choosing this data set is provided in section 2.2.3.

Figure 2.1: Daily On-Peak Energy Spot Prices, June 2000-June 2001



As we observe now, the nature of the price can be seen as a combination of deterministic trend together with random shocks.In other words, the prices tend to oscillate or revert around a mean level, with some extraordinary periods of volatility. These extraordinary periods of high volatility are reflected in the characteristic spikes observed in the market.

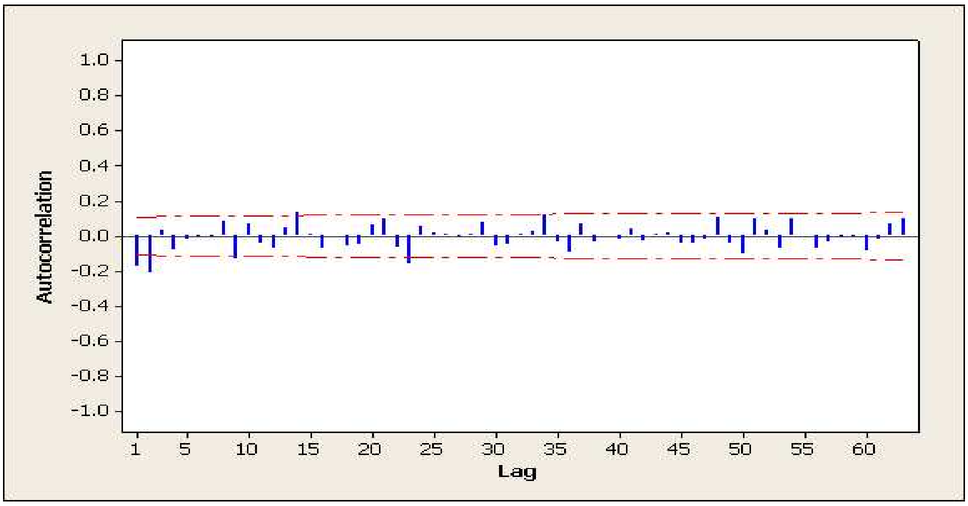
2.1.1 Seasonality

At this point it is imperative to discuss seasonality which is a commonly observed characteristic in energy markets. In order to assess whether there is actually an underlying pattern prevailing in the returns[2] an autocorrelation test can be easily carried out for verification.

[2]Here, the return is defined in the classical sense as, $r_t = \ln \frac{P_t}{P_{t-1}}$

As explained in [3], the evidence of high autocorrelation manifests an underlying seasonality. On the contrary, if the returns are independent, as assumed by the Black-Scholes model, the correlation coefficients would be very close to zero indicating insufficient evidence for an underlying seasonality. So, I compare the intensity of autocorrelation between the results before and after deseasonalization.

Figure 2.2: Autocorrelation test for Returns from 06/20/2000 to 06/20/2001



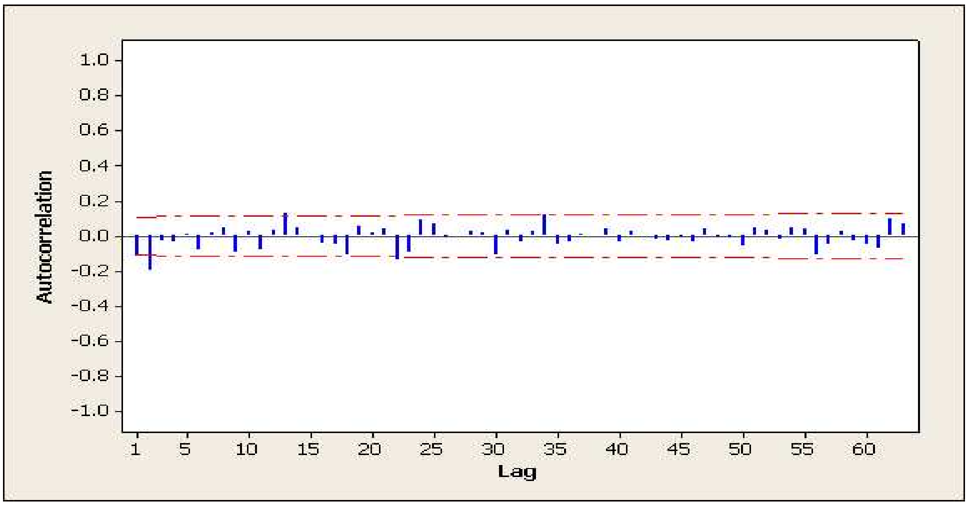
There are many ways to deseasonalise the data. Here, I would follow an approach where the mean of every month is subtracted from the corresponding months returns for the whole year. In particular,

$$DR_t = r_t - \bar{r_d}$$

where DR_t is the defined deseasonalised return at time t, r_t the return at time t and $\bar{r_d}$ is the corresponding mean of each month.

From the figures we observe that the autocorrelations before and after deseasonalizing the returns are very similar. In both cases, the coefficients are small indicating that our data does not exhibit any seasonality factor. This certifies that there is no need for including any seasonality function into our model.

Figure 2.3: Autocorrelation test for Deseasonalized Returns from 06/20/2000 to 06/20/2001



2.1.2 Initial Model

At first, the simplest model is considered which includes the mean reverting term with Guassian noise, *without* catering for the jumps, though they are evident in the data set. Then using some standard parameter estimation techniques, namely, Ordinary Least Squares Regression (OLSR) and Maximum Likelihood Estimation(MLE), I will carry out a comparison analysis, by running numerous simulations, as to which method gives a relatively accurate parameter estimation to model the data. One thing to notice is that the parameter estimates from both methods should be "reasonably" close.

$$dP(t) = \alpha(\mu - lnP(t))P(t)dt + \sigma P(t)dW(t) \tag{2.1}$$

where,

α = mean reversion rate, μ = related to the long term mean level of the logarithm of Electricity spot prices, σ = related to the volatility in electricity spot prices, and W = Standard Wiener Process

In order to solve this **stochastic differential equation**(SDE) the following substitution is made with the help of *Stochastic Calculus*.

$$z = \ln P$$
$$dz = d(\ln P)$$

Using Ito's Lemma as given in [4], we get

$$d(\ln P) = \left[\frac{1}{P}[\alpha(\mu - \ln P)P] + 0 - \frac{1}{2P^2}\sigma^2(dP)^2\right]dt + \frac{1}{P}\sigma P dW$$
$$dz = \alpha(\mu - \ln P_t) + \sigma dW_t - \frac{\sigma^2}{2}dt$$

$$\Rightarrow dz_t = \alpha(\mu - \frac{\sigma^2}{2\alpha} - z_t)dt + \sigma dW_t \tag{2.2}$$

Notice that the model equation is very similar to what has been known as the **Vasicek Model** for interest rates. Therefore, a similar methodology is used to solve this model. Consider,

$$Y = e^{\alpha t}z$$

Then, by Ito 's Lemma again,

$$d(Y) = d(e^{\alpha t}z) = e^{\alpha t}dz_t + \alpha e^{\alpha t}zdt + 0$$

Using equation (2.2),

$$d(e^{\alpha t}z) = e^{\alpha t}\left[\alpha(\mu - \frac{\sigma^2}{2\alpha} - z)dt + \sigma dW_t\right] + \alpha e^{\alpha t}zdt$$

$$d(e^{\alpha t}z) = \alpha(\mu - \frac{\sigma^2}{2\alpha})e^{\alpha t}dt + \sigma e^{\alpha t}dW_t$$

Integrating both sides w.r.t to time from $t_i \rightarrow t_{i+1}$ yields,

$$e^{\alpha t_{i+1}}z_{i+1} - e^{\alpha t_i}z_i = \alpha(\mu - \frac{\sigma^2}{2\alpha})\int_{t_i}^{t_{i+1}} e^{\alpha s}ds + \sigma\int_{t_i}^{t_{i+1}} e^{\alpha s}dW_s$$

$$z_{i+1} = z_i e^{-\alpha\triangle t} + (\mu - \frac{\sigma^2}{2\alpha})(1 - e^{-\alpha\triangle t}) + \sigma e^{-\alpha t_{i+1}}\int_{t_i}^{t_{i+1}} e^{\alpha s}dW_s \tag{2.3}$$

where $\triangle = t_{i+1} - t_i$

2.2 PARAMETER ESTIMATION

The next step is to use an appropriate method for estimating the parameters α, μ and σ. As mentioned earlier, to address this issue, the most commonly used methods include the Ordinary Least Squares Regression **(OLSR)**, and Maximum Likelihood Estimation **(MLE)**methods are utilized. However, it is expected that since the data exhibits large volatility instances, therefore, the affect of an outlier would tend to make the OLSR method less accurate than the MLE. In order to assert this assumption in a more convincing manner, both methods are applied to estimate the parameters and a few simulations are conducted to validate this hypothesis.

2.2.1 Ordinary Least Squares Regression

The discretized equation has been tailor made for an Autoregression. However, the idea behind converting it into an OLSR model is to subtract the z_i term from both sides of the equation which then gives,

$$z(t_{i+1}) - z(t_i) = z(t_i)(e^{-\alpha\triangle t} - 1) + (\mu - \frac{\sigma^2}{2\alpha})(1 - e^{-\alpha\triangle t}) + \sigma e^{-\alpha t_{i+1}}\int_{t_i}^{t_{i+1}} e^{\alpha s}dW_s \tag{2.4}$$

As we observe, we can now analyze this equation as an algebraic equation given by,

$$Y = mX + c + \epsilon \tag{2.5}$$

where, $m = (e^{-\alpha\triangle t} - 1), c = (\mu - \frac{\sigma^2}{2\alpha})(1 - e^{-\alpha\triangle t})$ and $\epsilon = \sigma e^{-\alpha t_{i+1}} \int_{t_i}^{t_{i+1}} e^{\alpha s} dW_s$.

In order to look at the above equation as a system of Linear equations, take

$$Y = XA + \epsilon$$

where,

$Y = z(t_{i+1}) - z(t_i)$, is an (n-1) dimensional vector containing the difference of the log prices where 'n' is the number of observations.

$X = [(n-1) \times 2]$ matrix with 1's in the first column and the log prices in the second.

$A = (2 \times 1)$ matrix with the first coefficient as the intercept and the second coefficient as the slope of the regression line.

$\epsilon =$ Noise or Residual term.

In particular,

$$\mathrm{Y} = \begin{pmatrix} z_1 - z_0 \\ z_2 - z_1 \\ \cdots \\ \cdots \\ \cdots \\ z_n - z_{n-1} \end{pmatrix}, \mathrm{X} = \begin{pmatrix} 1 & z_0 \\ 1 & z_1 \\ . & \cdots \\ . & \cdots \\ . & \cdots \\ 1 & z_{n-1} \end{pmatrix}, \mathrm{A} = \begin{pmatrix} c \\ m \end{pmatrix}$$

Firstly, we observe that the slope of the regression line must equal the coefficient of the log prices. In particular, as indicated above,

$$m = (e^{-\alpha\triangle t} - 1)$$

Taking natural logarithm of both sides gives,

$$-\alpha\triangle t = \ln(m+1)$$

$$\Rightarrow \alpha = \frac{-\ln(m+1)}{\triangle t} \tag{2.6}$$

Secondly, the deterministic part of equation 2.5 must equal the intercept of the regression line which gives,

$$c = (\mu - \frac{\sigma^2}{2\alpha})(1 - e^{-\alpha \triangle t})$$

$$\Rightarrow \mu = \frac{c}{(1 - e^{-\alpha \triangle t})} + \frac{\sigma^2}{2\alpha} \tag{2.7}$$

Finally, we need to formulate an equation for the volatility parameter, σ. Here, we rearrange the equation(2.5) in terms of the residual ϵ as follows,

$$\epsilon = Y - mX - c$$

The idea now is to minimize the variance of the residual error ϵ, which can be achieved by using elementary calculus optimization method, and set it equal to the variance from the model which would be calculated later by taking the second moment of equation(2.3).

$$Var(\epsilon) = \mathbb{E}(\epsilon^2) - [\mathbb{E}(\epsilon)]^2$$

Here, I claim that in order to attain minimum variance $\mathbb{E}(\epsilon) = \mathbb{E}(\mathbf{Y} - m\mathbf{X} - c) = 0$. The proof for this assumption will be seen shortly as we proceed.

$$\begin{aligned} Var(\epsilon) &= \mathbb{E}[(Y - mX - c)^2] \\ &= \mathbb{E}[(Y - mX)^2 - 2c(Y - mX) + c^2] \\ &= \mathbb{E}[Y^2 - 2mXY + m^2X^2 - 2cY + 2mcX + c^2] \end{aligned}$$

Now,by linearity of the expectation operator

$$Var(\epsilon) = G(m, c) = \mathbb{E}(Y^2) - 2mE(XY) + m^2\mathbb{E}(X^2) - 2cE(Y) + 2mcE(X) + c^2 \tag{2.8}$$

Notice that in the above equation the X's and Y's are known entities that come from the data. Hence, the variance is a **function of m and c alone**. As discussed earlier, with the use of one variable calculus for optimization we get,

$$\frac{\partial G}{\partial m} = -2\mathbb{E}(XY) + 2mE(X^2) + 2cE(X) = 0$$

$$\frac{\partial G}{\partial c} = -2\mathbb{E}(Y) + 2mE(X) + 2c = 0$$

At this point, observe that the **second equation gives the proof** for the claim which was made earlier. In order to see that closely, a simple rearrangement of the equation yields,

$$-2\mathbb{E}(Y) + 2mE(X) + 2c = 0$$

$$\Rightarrow \mathbb{E}(Y) - mE(X) - c = 0$$

$$\Rightarrow \mathbb{E}(Y - mX - c) = \mathbb{E}(\epsilon) = 0$$

Reverting back to finding the optimal values for **m** and **c** . From the above, a linear system of two equations is set up as follows.

$$\mathbb{E}(X^2)m + \mathbb{E}(X)c = \mathbb{E}(XY)$$
$$[\mathbb{E}(X)m + 1c = \mathbb{E}(Y)\mathbb{E}(X)] \times \mathbb{E}(X)$$

The system is now solved simultaneously by multiplying the second equation by the factor $\mathbb{E}(\mathbf{X})$ and subtracting from the first yielding the equation

$$m[\underbrace{\mathbb{E}(X^2) - (\mathbb{E}(X))^2}_{Variance(X)}] = [\underbrace{\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)}_{Covariance(X,Y)}]$$

$$\Rightarrow m = \frac{Cov(X,Y)}{\sigma_X^2} \tag{2.9}$$

$$\Rightarrow c = \mathbb{E}(Y) - \frac{Cov(X,Y)\mathbb{E}(X)}{\sigma_X^2} \tag{2.10}$$

Now, substituting these **optimal** values of 'm' and 'c' into equation (2.8) to get the minimum variance gives,

$$Var_{min}(\epsilon) = \mathbb{E}(Y^2) - 2\frac{Cov(X,Y)}{\sigma_X^2}\mathbb{E}(XY) + \left(\frac{Cov(X,Y)}{\sigma_X^2}\right)^2 \mathbb{E}(X^2) - 2\mathbb{E}[Y] - \left(\frac{Cov(X,Y)\mathbb{E}(X)}{\sigma_X^2}\right)\mathbb{E}(Y) +$$
$$2\frac{Cov(X,Y)}{\sigma_X^2}\left[\mathbb{E}(Y) - \frac{Cov(X,Y)\mathbb{E}(X)}{\sigma_X^2}\right]\mathbb{E}(X) + \left[\mathbb{E}(Y) - \left(\frac{Cov(X,Y)\mathbb{E}(X)}{\sigma_X^2}\right)\right]^2$$

$$= \mathbb{E}(Y^2) - 2[\mathbb{E}(Y)]^2 + [\mathbb{E}(Y)]^2 - \overbrace{2\frac{Cov(X,Y)}{\sigma_X^2}\mathbb{E}(XY)} + \frac{Cov(X,Y)^2}{(\sigma_X^2)^2}\mathbb{E}(X^2) + \overbrace{2\mathbb{E}(X)\mathbb{E}(Y)\frac{Cov(X,Y)}{\sigma_X^2}}$$
$$+ \overline{2\mathbb{E}(X)\mathbb{E}(Y)\frac{Cov(X,Y)}{\sigma_X^2}} - 2\frac{Cov(X,Y)^2}{(\sigma_X^2)^2}[\mathbb{E}(X)]^2 - \overline{2\mathbb{E}(X)\mathbb{E}(Y)\frac{Cov(X,Y)}{\sigma_X^2}} + \frac{Cov(X,Y)^2}{(\sigma_X^2)^2}[\mathbb{E}(X)]^2$$

After combining the terms who have been marked with the same symbols above them, the equation is reduced to,

$$= \underbrace{[\mathbb{E}(Y^2) - [\mathbb{E}(Y)]^2}_{\sigma_Y^2} - 2\tfrac{Cov(X,Y)}{\sigma_X^2}[\underbrace{\mathbb{E}(XY) - \mathbb{E}(Y)\mathbb{E}(X)}_{Cov(X,Y)}] + \tfrac{Cov(X,Y)^2}{(\sigma_X^2)^2}[\underbrace{\mathbb{E}(X^2) - [\mathbb{E}(X)]^2}_{\sigma_X^2}]$$

$$\Rightarrow Var_{min}(\epsilon) = \sigma_Y^2 - \frac{[Cov(X,Y)]^2}{\sigma_X^2} \tag{2.11}$$

As stated earlier, the second moment of the model [**See equation(2.3)**] is now calculated and set equal to equation(2.11) get the equation for σ as follows.

$$\epsilon_{model} = \sigma e^{-\alpha t_{i+1}} \int_{t_i}^{t_{i+1}} e^{\alpha s} dW_s$$

$$Var(\epsilon_{model}) = \mathbb{E}(\epsilon_{model}^2)$$

$$\Rightarrow Var(\epsilon_{model}) = \sigma^2 e^{-2\alpha t_{i+1}} \left(\int_{t_i}^{t_{i+1}} e^{\alpha s} dW_s \right)^2 \tag{2.12}$$

In order to evaluate the integral above, we make use of Ito's isometry theorem, [5] and [6].

Theorem 1. *If f belongs to* $H_2[0,T]$*, the space of random functions defined for all t in* $[0,T]$*, and* $\int_0^T \mathbb{E}(f(t))^2 dt < \infty$*, then*

$$\mathbb{E}\left[\int_0^T f(t) dW(t)\right] = 0 \tag{2.13}$$

$$\textit{and, } \mathbb{E}\left[\left(\int_0^T f(t) dW(t)\right)^2\right] = \int_0^T \mathbb{E}\left[f(t)\right]^2 dt \tag{2.14}$$

It is evident that the theorem's conditions hold true in our case since $\int_0^t \mathbb{E}(e^{\alpha s})^2 ds < \infty$. In particular, the explicit solution is,

$$\int_0^t \mathbb{E}(e^{\alpha s})^2 ds = \frac{1}{2\alpha}[e^{2\alpha t} - 1] < \infty$$

for $\alpha \neq 0$ and finite $\mathbf{t}$.

Therefore, using Ito's isometry the solution to our integral equation (2.12) is given by,

$$\begin{aligned} \sigma^2 e^{-2\alpha t_{i+1}} \left(\int_{t_i}^{t_{i+1}} e^{\alpha s} dW_s\right)^2 &= \sigma^2 e^{-2\alpha t_{i+1}} \int_{t_i}^{t_{i+1}} e^{2\alpha s} ds \\ &= \sigma^2 e^{-2\alpha t_{i+1}} \frac{1}{2\alpha} [e^{2\alpha s}]_{s=t_i}^{s=t_{i+1}} \\ &= \sigma^2 e^{-2\alpha t_{i+1}} \frac{1}{2\alpha} [e^{2\alpha t_{i+1}} - e^{2\alpha t_i}] \end{aligned}$$

$$\Rightarrow Var(\epsilon_{model}) = \frac{\sigma^2}{2\alpha}\left(1 - e^{-2\alpha \triangle t}\right) \tag{2.15}$$

Now putting together equations (2.12) and (2.14), the equation for the volatility parameter σ is given by,

$$\frac{\sigma^2}{2\alpha}[1 - e^{-2\alpha\triangle t}] = \sigma_Y^2 - \frac{[Cov(X,Y)]^2}{\sigma_X^2}$$

$$\Rightarrow \sigma^2 = \frac{2\alpha}{(1 - e^{-2\alpha\triangle t})}\left[\sigma_Y^2 - \frac{[Cov(X,Y)]^2}{\sigma_X^2}\right] \tag{2.16}$$

In summary, here are the three equations for the three parameters α, μ and σ in the order they would be estimated.

$$\alpha = \frac{-\ln(m+1)}{\triangle t} \tag{2.17}$$

$$\sigma^2 = \frac{2\alpha}{(1 - e^{-2\alpha\triangle t})}\left[\sigma_Y^2 - \frac{[Cov(X,Y)]^2}{\sigma_X^2}\right] \tag{2.18}$$

$$\mu = \frac{c}{(1 - e^{-\alpha\triangle t})} + \frac{\sigma^2}{2\alpha} \tag{2.19}$$

2.2.2 Maximum Likelihood Estimation (MLE)

The Maximum Likelihood Estimation (MLE) is a popular statistical method to estimate and make inferences about the unknown parameters of the **underlying probability distribution** from a given data set. It is typically believed to yield better estimates than the regression method primarily because of two reasons.

- Significantly Less volatile to outliers in the data set.
- Takes into all moments for parameter estimation.

According to our initial model formulation for the logarithm of electricity spot prices given by equation (2.3), we assume that the underlying **conditional**[3]**distribution** of the

[3]The data untill the previous time step is used in the computation making it a conditional distribution

log-prices with unknown parameters α, μ and σ , given by equation (2.3) is **Gaussian** at each time step with mean ν_i and variance ω_i^2 given by,

$$\nu_{i+1} = \mathbb{E}\left[z_{i+1})|z_i\right] \tag{2.20}$$

$$\omega_{i+1}^2 = \sigma^2 e^{-2\alpha t_i}(\int_{t_i}^{t_{i+1}} e^{\alpha\tau} dW_\tau)^2 \tag{2.21}$$

Since each z_i term is known, the calculation shown below can be easily extended with $z_0 \longrightarrow z_i$

$$\nu_i = \mathbb{E}\left[z_{i+1}|z_i\right] = \mathbb{E}\left[z_i e^{-\alpha t_i} + (\mu - \frac{\sigma^2}{2\alpha})(1 - e^{-\alpha t_i}) + \sigma e^{-\alpha t_i}\int_0^{t_i} e^{\alpha\tau} dW_\tau\right]$$

By linearity of the expectation operator,

$$\nu_i = \mathbb{E}[z_0 e^{-\alpha t_i}] + \mathbb{E}[(\mu - \frac{\sigma^2}{2\alpha})(1 - e^{-\alpha t_i})] + \sigma e^{-\alpha t_i}\mathbb{E}\left[\int_0^{t_i} e^{\alpha\tau} dW_\tau\right]$$

Now, from the properties of the standard Wiener process $\mathbb{E}[\int_0^{t_i} e^{\alpha\tau} dW_\tau] = 0$ and the expectation of the remaining (deterministic) part is just itself. This yields,

$$\nu_i = z_0 e^{-\alpha t_i} + (\mu - \frac{\sigma^2}{2\alpha})(1 - e^{-\alpha t_i}) \tag{2.22}$$

To determine the variance ω_i^2 of the underlying probability distribution we calculate the second moment for equation(2.3) as follows using the Ito's isometry (Theorem 1).

$$\begin{aligned}
\omega_i^2 &= \sigma^2 e^{-2\alpha t_i}(\int_0^{t_i} e^{\alpha\tau} dW_\tau)^2 \\
&= \sigma^2 e^{-2\alpha t_i}\left(\int_0^{t_i} e^{2\alpha\tau} d\tau\right) \\
&= \frac{\sigma^2 e^{-2\alpha t_i}}{2\alpha}\left[e^{2\alpha t_i} - 1\right]
\end{aligned}$$

$$\Rightarrow \omega_i^2 = \frac{\sigma^2}{2\alpha}[1 - e^{-2\alpha t_i}] \tag{2.23}$$

To summarize our findings, it is concluded that our data at each time step follows a **Gaussian** distribution with mean ν_i and variance ω_i^2 given by equations (2.22) and (2.23). In other words, the log-prices $z_i's \sim N(\nu_i, \omega_i^2)$ for i=1,2....N, where N is the number of data points.

2.2.2.1 Likelihood function After establishing the parameter equations for the underlying probability distribution we can define the **Likelihood function** $\boldsymbol{L(\theta)}$, where, $\vec{\theta} = [\alpha, \mu, \sigma^2]$, which basically allows us to determine the unknown parameters based on the information about the underlying probability distribution and the outcome (data). Formally, with the known data points $z_1, z_2, ...z_N$ having a continous probability distribution (Gaussian), we can compute the probability density function associated with our observed data points as

$$f_\theta(z_1, z_2, ...z_N|\theta) = \prod_{i=1}^{N} f(z_i|z_{i-1}, \nu_i, \omega_i^2)$$

As a function of θ with all the z_i's known, this is the likelihood function,

$$L(\theta) = \prod_{i=1}^{N} f(z_i|z_{i-1}, \nu_i, \omega_i^2)$$

or more conveniently as,

$$L(\alpha, \mu, \sigma^2) = \prod_{i=1}^{N} (\frac{1}{\sqrt{2\pi\omega_i^2}}) e^{-\frac{(z_i-\nu_i)^2}{2\omega_i^2}} \tag{2.24}$$

To avoid a tedious algebraic manipulation we take the logarithm for the likelihood function. Since the logarithm is a **continuous strictly increasing function** over the range of the likelihood, the values which maximize the likelihood will also maximize its logarithm.

$$\hat{L}(\alpha, \mu, \sigma^2) = \ln L(\alpha, \mu, \sigma^2) = -\frac{N}{2} \ln 2\pi - \sum_{i=1}^{N} \left[\frac{\ln \omega_i^2}{2} + \frac{(z_i - \nu_i)^2}{2\omega_i^2} \right] \tag{2.25}$$

In order to find the parameter equations which maximize this log-likelihood function we take the partial derivatives with respect to two of the parameters μ and σ^2 which would convert the multi-variable log-likelihood function into a single variable function in terms of the parameter α. The motivation behind this manipulation is to avoid indulging into tedious algebraic derivative calculation of the function $\mathbf{\hat{L}}(\alpha, \mu, \boldsymbol{\sigma^2})$ with respect to α. Now, since the first two terms in equation (2.25) are independent of the parameter μ, we have

$$0 = \frac{\partial(\ln L)}{\partial\mu} = 0 - 0 - \frac{\partial}{\partial\mu}[\sum_{i=1}^{N}\frac{(z_i-\nu_i)^2}{2\omega_i^2}]$$

$$0 = -\frac{1}{2}\sum_{i=1}^{N}2\frac{(z_i-\nu_i)}{\omega_i^2}[-(1-e^{-\alpha t_i})]$$

$$0 = \sum_{i=1}^{N}\frac{(z_i-\nu_i)}{\omega_i^2}(1-e^{-\alpha t_i})$$

$$\Rightarrow \qquad \sum_{i=1}^{N}\frac{z_i}{\omega_i^2}(1-e^{-\alpha t_i}) = \sum_{i=1}^{N}\frac{\nu_i}{\omega_i^2}(1-e^{-\alpha t_i})$$

Substituting for ν_i from equation (2.22) yields,

$$\sum_{i=1}^{N}\frac{z_i}{\omega_i^2}(1-e^{-\alpha t_i}) = \sum_{i=1}^{N}\frac{[z_0e^{-\alpha t_i}+(\mu-\frac{\sigma^2}{2\alpha})(1-e^{-\alpha t_i})](1-e^{-\alpha t_i})}{\omega_i^2}$$

$$\Rightarrow (\mu-\frac{\sigma^2}{2\alpha}) = \frac{\sum_{i=1}^{N}\frac{(1-e^{-\alpha t_i})(z_i-z_0e^{-\alpha t_i})}{\omega_i^2}}{\sum_{i=1}^{N}\frac{(1-e^{-\alpha t_i})^2}{\omega_i^2}}$$

Substituting for ω_i^2 from equation (2.23) gives,

$$(\mu-\frac{\sigma^2}{2\alpha}) = \frac{\sum_{i=1}^{N}\frac{(1-e^{-\alpha t_i})(z_i-z_0e^{-\alpha t_i})}{\frac{\sigma^2}{2\alpha}(1-e^{-2\alpha t_i})}}{\sum_{i=1}^{N}\frac{(1-e^{-\alpha t_i})^2}{\frac{\sigma^2}{2\alpha}(1-e^{-2\alpha t_i})}}$$

Notice now that the term $\frac{\sigma^2}{2\alpha}$ is independent of the index i and therefore gets eliminated. Moreover, $(1-e^{-2\alpha t_i}) = (1-e^{-\alpha t_i})(1+e^{-\alpha t_i})$, which simplifies the above expression to give a **function entirely in terms of** α , say $\hat{\mathbf{h}}(\alpha)$ explicitly given by,

$$\hat{h}(\alpha) = (\mu-\frac{\sigma^2}{2\alpha}) = \frac{\sum_{i=1}^{N}\frac{(z_i-z_0e^{-\alpha t_i})}{(1+e^{-\alpha t_i})}}{\sum_{i=1}^{N}\frac{(1-e^{-\alpha t_i})}{(1+e^{-\alpha t_i})}} \tag{2.26}$$

Notice here that the function $\hat{\mathbf{h}}(\alpha)$ equals the term $(\mu - \frac{\sigma^2}{2\alpha})$. The aim here is to avoid calculating the complicated quantity $\frac{\partial L}{\partial \alpha}$ by converting the multi-variable log-likelihood function given by equation (2.25) into a single variable function of α so that the problem of finding the optimal parameters α_{opt}, μ_{opt} and σ_{opt} is converted to a one variable maximization problem. Therefore, a similar algebraic manipulation for $\frac{\partial L}{\partial \sigma^2}$ is carried out as follows.

$$0 = \frac{\partial L}{\partial \sigma^2} = -\frac{N}{2\sigma^2} - \frac{1}{2}\left[\sum_{i=1}^{N} \frac{2(z_i - v_i)\frac{1}{2\alpha}(1 - e^{-\alpha t_i})}{\omega_i^2} - \sum_{i=1}^{N} \frac{(z_i - v_i)^2 \frac{1}{2\alpha}(1 - e^{-2\alpha t_i})}{\omega_i^4}\right]$$

Now, from equation (2.23), we know that

$$\omega_i^2 2\alpha = \sigma^2[1 - e^{-2\alpha t_i}]$$

and equivalently,

$$\frac{(1 - e^{-2\alpha t_i})}{2\alpha} = \frac{\omega_i^2}{\sigma^2}$$

Using the above two equivalent forms, we get

$$0 = -\frac{N}{2\sigma^2} - \frac{1}{\sigma^2}\sum_{i=1}^{N} \frac{(z_i - \nu_i)(1 - e^{-\alpha t_i})}{(1 - e^{-2\alpha t_i})} + \frac{1}{2\sigma^2}\sum_{i=1}^{N} \frac{(z_i - \nu_i)^2}{\omega_i^2}$$

$$0 = -\frac{N}{2} - \sum_{i=1}^{N} \frac{(z_i - \nu_i)(1 - e^{-\alpha t_i})}{(1 - e^{-2\alpha t_i})} + \frac{1}{2}\sum_{i=1}^{N} \frac{(z_i - \nu_i)^2}{\frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t_i})}$$

$$\frac{1}{2(\frac{\sigma^2}{2\alpha})}\sum_{i=1}^{N} \frac{(z_i - \nu_i)^2}{(1 - e^{-2\alpha t_i})} = N + \sum_{i=1}^{N} \frac{(z_i - \nu_i)}{(1 + e^{-\alpha t_i})}$$

$$\frac{\sigma^2}{2\alpha} = \frac{\sum_{i=1}^{N} \frac{(z_i - \nu_i)^2}{(1 - e^{-2\alpha} t_i)}}{2[N + \sum_{i=1}^{N} \frac{(z_i - \nu_i)}{(1 + e^{-\alpha t_i})}]}$$

Observe that equation(2.22) in terms of the function $\hat{h}(\alpha)$ can be written as another function entirely in terms of α, say $\hat{\mathbf{g}}(\alpha)$ for each index i as,

$$\hat{g}_i(\alpha) = \nu_i = z_0 e^{-\alpha t_i} + \hat{h}(\alpha)(1 - e^{-\alpha t_i}) \tag{2.27}$$

Finally, using the above equation in the preceding algebra yields the function, $\hat{K}(\alpha)$ we ought to seek given explicitly by,

$$\hat{K}(\alpha) = \frac{\sigma^2}{2\alpha} = \frac{\sum_{i=1}^{N} \frac{\hat{g}_i(\alpha)^2}{(1-e^{-2\alpha t_i})}}{2[N + \sum_{i=1}^{N} \frac{\hat{g}_i(\alpha)}{(1+e^{-\alpha t_i})}]} \tag{2.28}$$

At this point, it is important to rewrite equations(2.22) and (2.23) in terms of the new functions defined above and (2.25) in order to see clearly how the multi-variable problem has been converted to single-variable optimization problem.

$$\omega_i^2 = \hat{K}(\alpha)[1 - e^{-2\alpha t_i}] \tag{2.29}$$

$$\nu_i = z_0 e^{-\alpha t_i} + \hat{h}(\alpha)(1 - e^{-\alpha t_i}) \tag{2.30}$$

Plugging the above two equations in equation(2.25), yields the single variable log-likelihood function given by,

$$\tilde{L}(\alpha) = -\frac{N}{2} \ln 2\pi - \frac{1}{2} \sum_{i=1}^{N} \ln \left(\hat{K}(\alpha)(1 - e^{-2\alpha t_i}) \right)^2 + \frac{[z_i - (z_0 e^{-\alpha t_i} + \hat{h}(\alpha)(1 - e^{-\alpha t_i}))]^2}{2[\hat{K}(\alpha)(1 - e^{-2\alpha t_i})]^2} \tag{2.31}$$

2.2.3 Simulation Results for Initial Model

To recall, my current model does not incorporate the jumps. Therefore, in order to proceed further towards calibrating the jumps, first the parameters are estimated using, both, **Regression** and **MLE** against an initial data set as discussed earlier, to ensure the validity of our arithmetic. In particular, a filtered data set is considered with one year daily "on pea" electricity prices. This extraction is carried out due to the following reasons.

- My ultimate goal is to price some **forwards contracts** based on the underlying model with different dates of expiry. As most derivative trading is done during the peak hours in all markets, this data set promises to yield pragmatic estimates for the parameters.

- After careful analysis of the data set, it was observed that the"on-peak" spot prices exhibit the major variation in the prices as compared to the off-peak hours when the prices are generally stable.

In order to simulate the prices using the regression model, the following discretized scheme was developed with the corresponding results followed.

$$
\begin{aligned}
dz &= \alpha(\mu - \frac{\sigma^2}{2\alpha} - z)dt + \sigma dW \\
z_{i+1} - z_i &= \alpha(\mu - \frac{\sigma^2}{2\alpha} - z_i)dt + \sigma\sqrt{\Delta t}\phi_i \\
z_{i+1} &= \alpha(\mu - \frac{\sigma^2}{2\alpha})dt(1 - \alpha dt)z_i + \sigma\sqrt{\Delta t}\phi_i
\end{aligned}
$$

After simulating the log-prices as above, the prices were simulated by simply exponentiating the log-prices as,

$$P_{i+1} = e^{z_{i+1}}$$

We notice by observing the simulated prices (see Fig (2.4)) that the model captures the high frequency and low-volatility variation in the data set reasonably well. However, it is clear that in order to account for the relatively infrequent high volatility (jumps) in the data, a **jump process** needs to be incorporated into the model. This assertion is formally explained by carrying out the normality tests for the returns[4]. Notice also in the given table below Table (2.1) that the parameter estimates for regression and MLE methods are fairly close which validates the correctness of the arithmetic details shown earlier.

[4]Here, I define the "return" as in the classical definition; $r_t = ln(\frac{P_{t+1}}{P_t})$. Note that this is also referred as the difference in Log Prices in the paper.

Figure 2.4: Simulated Energy Spot Prices versus Real Prices, June 2000-June 2001

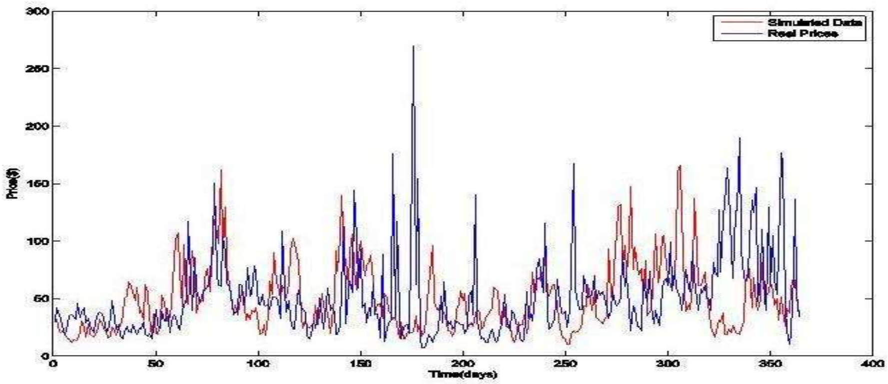


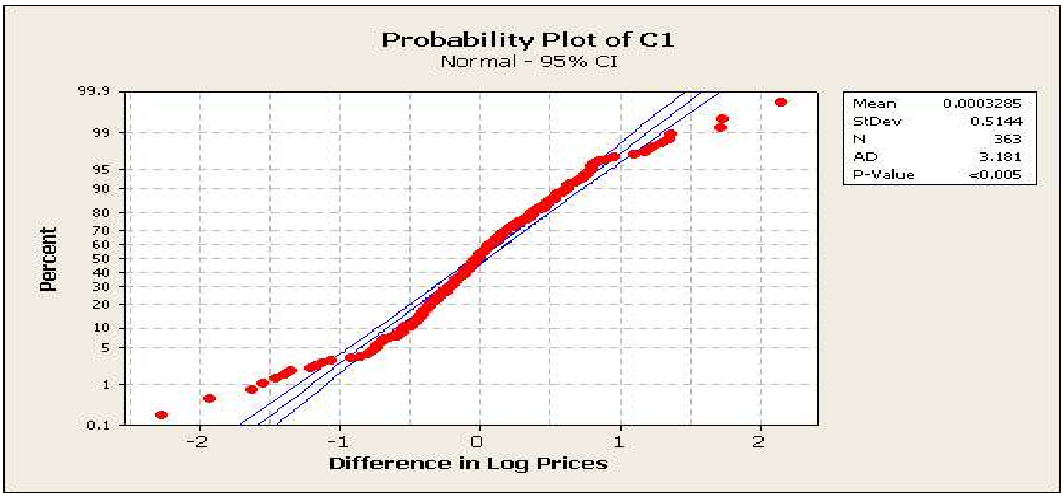
Table 2.1: Parameter Estimates for one-year data

Parameter	Reg	MLE
α	158.49	158.49
μ	4.112	4.114
σ^2	118.42	118.41

2.3 NORMALITY TESTS

In the preceding simulation there is evidence that our model fails to capture the jumps in the returns. In other words, the rare events of abnormally high spot prices are more frequent than predicted by the Normal distribution. This observation is verified by analyzing the Normal probability plots[5] shown in Fig(2.5).

Figure 2.5: Probability Plot for returns of electricity prices from 06/20/2000 to 06/20/2001



The p-value given by the Minitab probability plot toolbox is much lower than α, which is 0.05. This provides evidence for rejecting the null-hypothesis that the data is normally distributed and we can observe this directly from the fat tails. For instance, corresponding to a probability of 0.01 we have returns which are higher than 1.5; instead if the data were perfectly Normally distributed, the red dotted points would lie close to the central line and hence the probability of such a return should be virtually zero. Therefore, it is concluded that we need to add a **jump process** to our model to account for the fat tails, that is,

[5]The probability plots are obtained using the probability plot tool in Minitab statistical package. The indicated p-value indicates acceptance/rejection of the null hypothesis that the data is normally distributed. If p-value higher than α we accept the null hypothesis, else we reject.

abnormally high returns.

2.4 CALIBRATION OF JUMPS

The key issue in calibrating the jumps is to formulate a suitable definition of a jump. I define the jumps in the following two ways.

1. **Residual Deviation**

   Calculate the residual, $\epsilon_i = Y_i - mX_i - c$. Compute the standard deviation of the **absolute** values of the residual and define the upper and lower bounds for truncating the jumps as 3 standard deviations [6] above and below the regression line at each point, respectively. Formally,

$$\sigma_{rd} = \sqrt{\frac{\sum_{i=1}^{N}(\epsilon_i - \bar{\epsilon})^2}{N}} \tag{2.32}$$

$$U_i^{rd} = mX_i + c + 3\sigma_{rd} \tag{2.33}$$

$$L_i^{rd} = mX_i + c - 3\sigma_{rd} \tag{2.34}$$

2. **Absolute Deviation**

   Compute the absolute change in the log-prices, $\delta_i = z_{i+1} - z_i$. Calculate the mean and standard deviation for the differences and define the upper and lower bounds for truncating the jumps as 3 standard deviations above and below the mean level, respectively. In particular,

[6] A normally distributed data set includes 99.9within 3 standard deviations from its mean level. Therefore, this criteria for separating the jumps has been adopted

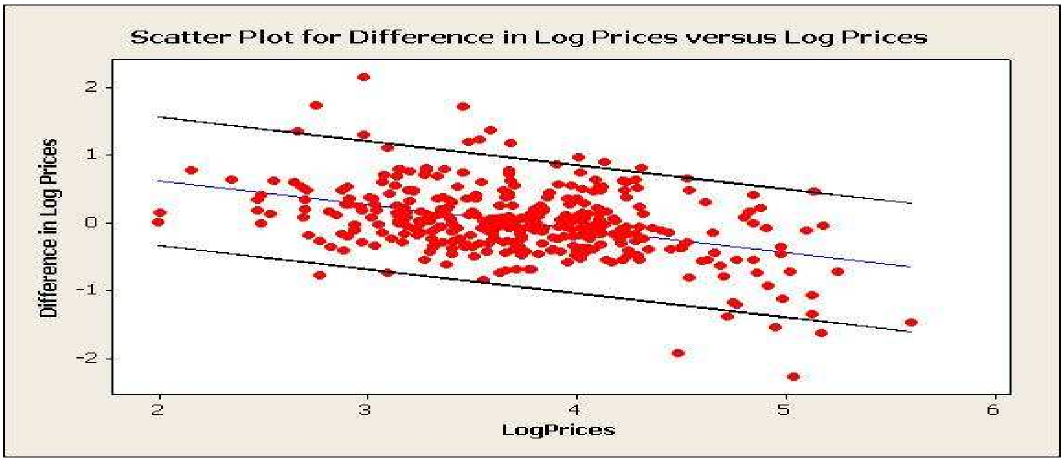
$$\mu_{AC} = \frac{\sum_{i=1}^{N} \delta_i}{N} \tag{2.35}$$

$$\sigma_{AC} = \sqrt{\frac{\sum_{i=1}^{N} (\delta_i - \mu_{AC})^2}{N}} \tag{2.36}$$

$$U_i^{AC} = \mu_{AC} + 3\sigma_{AC} \tag{2.37}$$

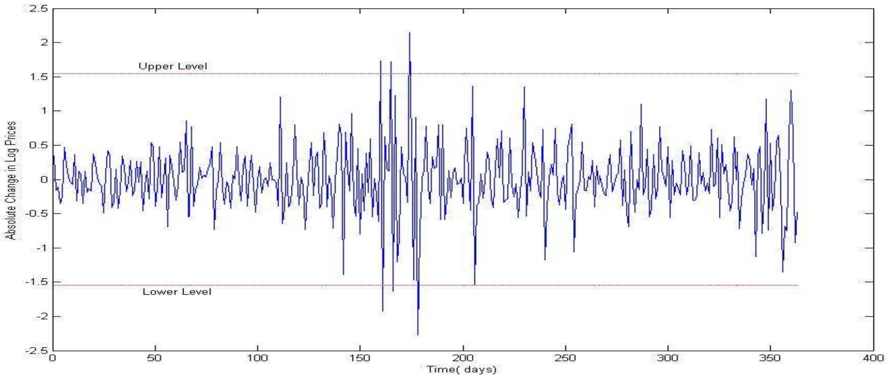
$$L_i^{AC} = \mu_{AC} - 3\sigma_{AC} \tag{2.38}$$

Figure 2.6: Scatterplot of Difference in Log Prices vs Log Prices



In order to visualize the truncation as explained above for a single iteration [7], we observe Fig(2.6) and Fig(2.7) shown below corresponding to the residual and absolute deviation methods respectively.

Figure 2.7: Absolute Change in Log Price with Upper and Lower Bounds



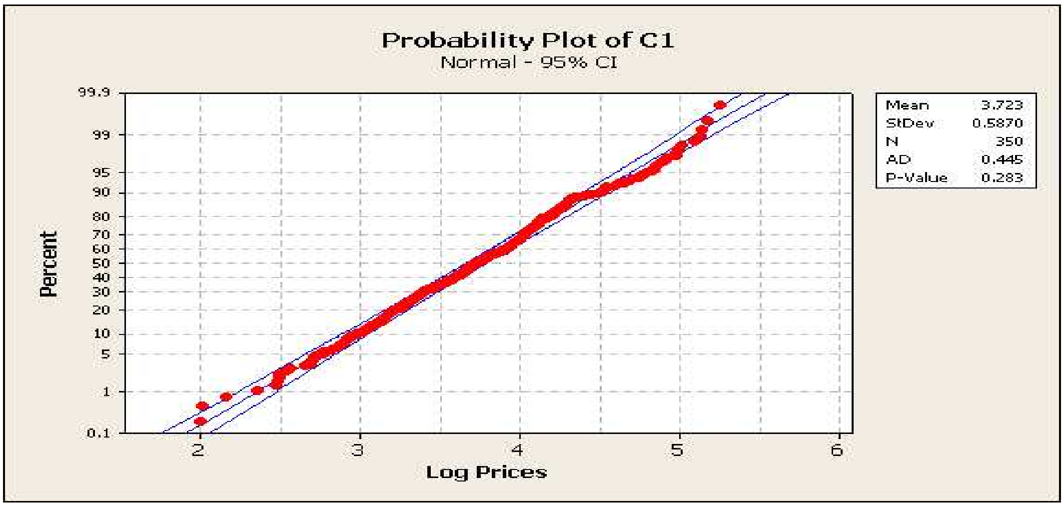
After careful examination it was observed that the **residual approach does not converge**. Therefore, we resort to the latter for further analysis. At this point, it is checked to see if the **remaining data** for the returns (excluding the jumps) satisfies the normality test.

According to Fig(2.8) there is strong evidence that, indeed, the jump-exclusive returns are Normally distributed by observing the test statistics[8] given in the probability plot.

[7] The important criteria to choose, amongst the proposed methods for calibrating the jumps enumerated above, was to check that after each iteration of removing the jumps, the next iteration, with recomputed values for the variables given by equations (2.32) through (2.38), should yield fewer number of jumps and eventually convergence is achieved.

[8] AD stands for the Anderson-Darling test statistic

Figure 2.8: Normal Probability test for returns of electricity prices from 06/20/2000 to 06/20/2001

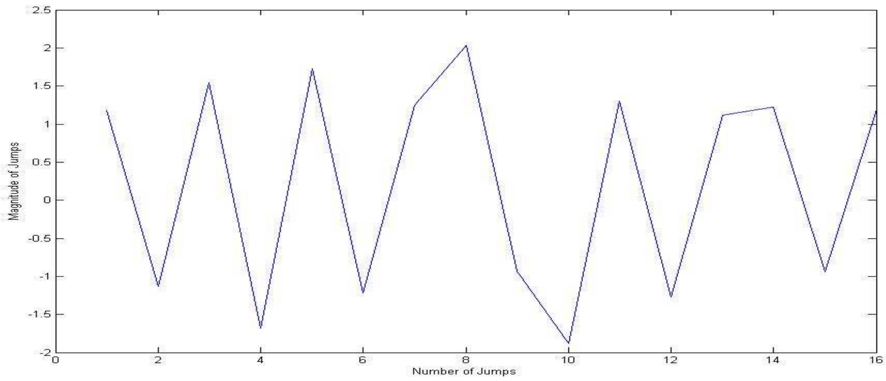


2.5 MODELING JUMPS

As seen from the Normality test, Fig(2.5), the fast tails suggest the insufficiency of a Guassian process to model the returns completely. To capture these rare but unusually high intensity returns in electricity prices a technical, data driven approach is employed. The indices at which the jumps occur are stored and the magnitudes of jumps at these indices is defined as:

$$k_i = dz_i^{data} - \alpha(\mu - \frac{\sigma^2}{2\alpha} - z_i^{data})dt - \sigma dW_i \tag{2.39}$$

Figure 2.9: Magnitudes of Jumps vs Number of Jumps



2.5.0.1 Observed Characteristics of Jumps

From Fig (2.9), it is easy to notice the following properties of the magnitude and occurrence of jumps [9].

1. Positive and negative jumps occur in pairs, that is, a positive jump is accompanied by a negative jump.
2. The absolute magnitude of the jumps, $|k_i|$, is very similar.

   Here, some reasonable assumptions are made about the jumps.

2.5.0.2 Assumptions

[9] Due to the random term , dW_i ,involved in our model, there is some variation, at each simulation, in the jump magnitudes. Therefore, I took 2000 simulations and averaged the magnitudes for each i .

- It is assumed for reasons of simplicity that every jump up is followed by a jump down the next immediate day even though there are very few occasions where this does not happen, that is, there are two consecutive jumps in the same direction.
- The mean reverting term in our model caters for the down slide after an upward jump. Hence, we need only model the positive spikes.

There are two different approaches I utilized to model the jump process.

1. Log-Normal Distribution
2. Non-parametric Distribution

The reasons for opting the enumerated methods for accommodating jumps into the model are both theoretical and empirical.

2.5.1 Log-Normal Distribution

Theoretically, the log-normal distribution is most commonly used method for modeling spikes in electricity spot prices by researchers over the years. For an empirical verification of this assumption,a normal probability plot for the logarithm of jump magnitudes is shown in Fig(2.10).

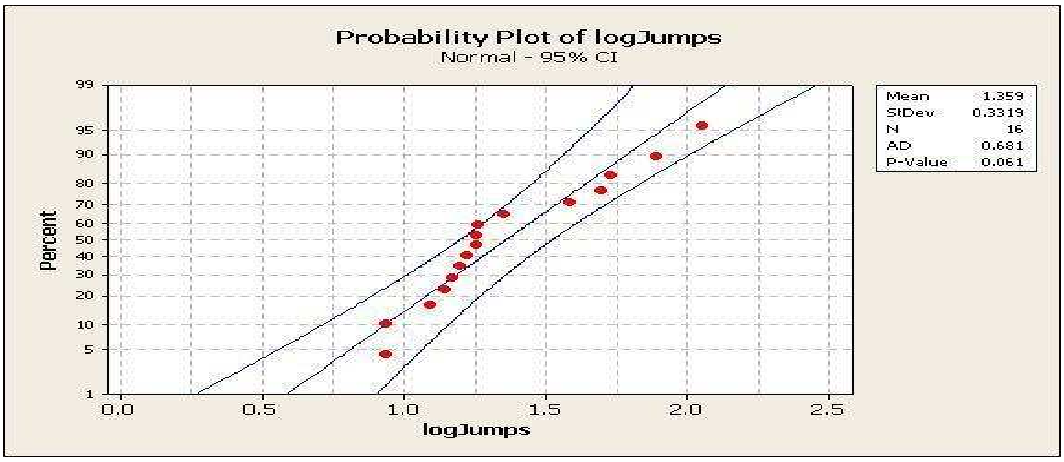
It is seen clearly by observing the statistics shown on the probability plot that indeed the jump magnitudes are log-normally distributed as the p-value is higher than α, which is 0.05. Therefore, based on the analysis, a log-normal term is introduced into the model as follows.

$$dz_t = \alpha(\mu - \frac{\sigma^2}{2\alpha} - z_t) + \sigma dW_t + \ln J dN_t \tag{2.40}$$

where, dN_t is a Poisson process with frequency λ such that

$$dN_t = \begin{cases} 1 & \text{with probability } \lambda dt \\ 0 & \text{with probability } 1 - \lambda dt \end{cases} \tag{2.41}$$

Figure 2.10: Normality test for Logarithm of jump data



The parameters for the log-normal distribution are found through the distribution fitting tool in Matlab. Given the parameters by the fit[10] a **Monte-Carlo simulation** is formulated to carry out the simulation results for energy spot prices as described by the modified model given by equation (2.40).

[10]The log-normal distribution fit tool in Matlab is applied 3 times on the averaged jump magnitudes for 2000 simulations and it is observed that each time the distribution gives almost a replica.

2.5.2 Non-parametric Distribution

In this section I introduce some non-parametric statistics to analyze the jump process due to the following, known, theoretical reasons.

- The jump data was observed to be highly non-normal.
- The data set is very small in size.
- The data belongs to an unknown distribution.

The idea involved in this approach is that it sets out to estimate the unknown probability density function of the random variable by using a kernel density approximation. In simple words, rather than grouping observations in bins, in case of the Histogram density estimator, the kernel density estimator can be thought to place small "bumps" at each observation, determined by the kernel function and the estimator consists of a "sum of bumps" and is clearly smoother.

In particular, if $x_1, x_2, ..., x_N$ f is an **independently identically distributed** sample of a random variable, then the kernel density approximation of its probability distribution function is given by,

$$\hat{f}_h(x) = \frac{1}{Nh} K \sum_{j=1}^{N} \left(\frac{x - x_j}{h} \right) \tag{2.42}$$

where, K is usually taken to be a Gaussian Kernel with mean 0 and variance σ^2 and h is the bandwidth (smoothing parameter).

$$K = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \tag{2.43}$$

In order to simulate these random numbers, I use the standard method of sampling through a uniform distribution where we find the cumulative distribution function for the random variables to be generated. In this case, it is calculated explicitly as given below.

$$F(y) = \frac{1}{Nh\sigma\sqrt{2\pi}} \int_{-\infty}^{y} \sum_{j=1}^{N} e^{-\frac{(\frac{x-x_j}{h})^2}{2\sigma^2}} dx \tag{2.44}$$

Using the distribution fitting tool, it is known that for this particular data set the kernel is the standard normal distribution with mean 0 and variance 1 with a certain value for the smoothing parameter h[11]. Accordingly, the integral is simplified and solved by making a suitable substitution as shown below.

$$F(y) = \frac{1}{Nh\sqrt{2\pi}} \sum_{j=1}^{N} \int_{-\infty}^{y} e^{-\frac{(x-x_j)^2}{2h^2}} dx$$
$$\text{Let } \theta = \frac{x - x_j}{h}, \text{ then } d\theta = \frac{1}{h} dx \Rightarrow$$
$$F(y) = \frac{1}{N} \sum_{j=1}^{N} \int_{-\infty}^{\frac{y-x_i}{h}} e^{\frac{-\theta^2}{2}} d\theta$$

$\Rightarrow$

$$F(y) = \frac{1}{N} \sum_{j=1}^{N} \phi \left(\frac{y - x_j}{h} \right) \tag{2.45}$$

where, ϕ is the **standard normal** probability distribution function.

After computing the cumulative probability distribution function, a numerical scheme is formulated to find the inverse mapping , $y = F^{-1} \left[\frac{1}{N} \sum_{j=1}^{N} \phi \left(\frac{y-x_j}{h} \right) \right]$. This mapping yields the random numbers, namely y's, which belong to the kernel density approximated probability distribution function for the log jumps.

2.5.3 Simulation Results with Jumps

The simulated electricity spot prices for log-normal and non-parametric approaches are presented in Fig(2.13) and Fig(2.14) below.

[11] There is minimal variation in the value of the smoothing parameter over each simulation since the data set is averaged over 2000 simulations. So, the simulation with non-parametric jumps is run using one particular value for h.

Figure 2.11: Simulation of Energy Spot Prices with Log-normally distributed Jumps

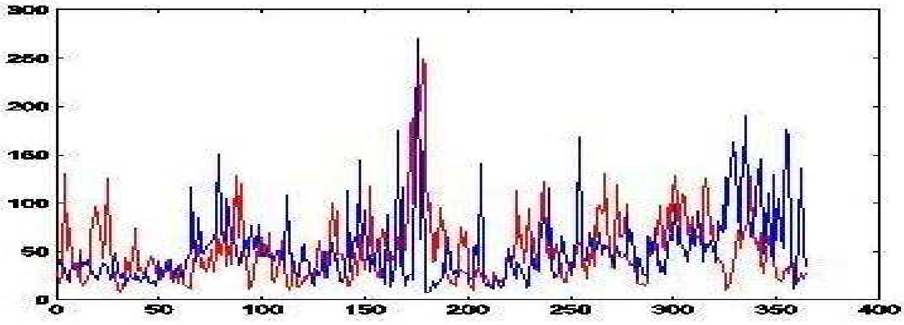
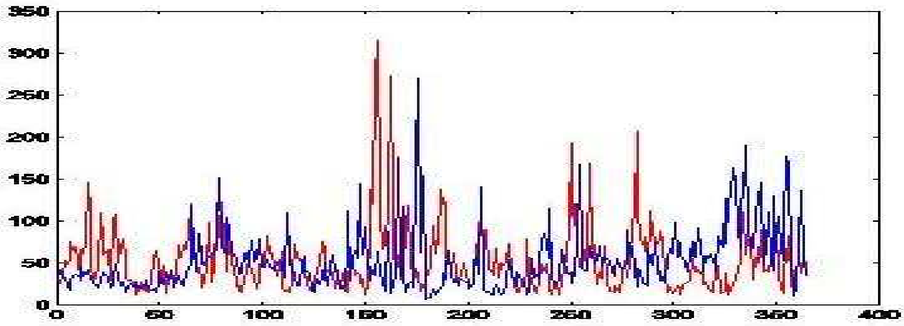


Figure 2.12: Simulation of electricity spot prices with a non-parametric of Jumps



3.0 FORWARD SPOT PRICES

As discussed earlier in this paper, electricity is a very expensive commodity to keep in storage. Once purchased, it must be consumed almost immediately. Therefore, the usual hedging strategies adopted in case of other financial assets, like holding certain amounts of the underlying, which in this case is electricity, is not a pragmatic solution. Consequently, in the electricity market forwards on the spot prices are typically used as a hedging strategy.

3.1 CLOSED-FORM REPRESENTATION OF THE FORWARD PRICE

In this chapter, I adopt the same approach as in [7] who find the price at time t of a forward contract with expiry T, by taking the expectation of the spot price at expiry under an equivalent Q -martingale measure conditional on the information available uptill time 't'; precisely

$$F(t,T) = \mathbb{E}_t^Q[P_T|B_t{}^1] \tag{3.1}$$

At first, in order to get a closed-form expression for the spot prices at expiry, P_T, the solution, z_T to Equation (2.40) is found using a similar method (Vasicek Model approach) as adopted earlier in the paper. The price at expiry, P_T is then computed by calculating the

[1] B_t is defined as the filtration on F, where a filtration is a sequence of σ-algebras on a measureable space

expectation of e^{z_T} as discussed by (3.1). Formally, Let

$$Y = e^{\alpha t} z_t$$

Then, $$dY = d(e^{\alpha t} z_t) = \alpha e^{\alpha t} z_t + e^{\alpha t} dz_t$$

Substituting for dz_t from Equation (2.40) and simplifying yields,

$$d(e^{\alpha t} z_t) = \alpha e^{\alpha t}(\mu - \frac{\sigma^2}{2\alpha})dt + \sigma e^{\alpha t} dW + e^{\alpha t} \ln J dN_t$$

Integrating both sides from t$\longrightarrow$ T gives,

$$e^{\alpha T} z_T - e^{\alpha t} z_t = \alpha e^{\alpha t}(\mu - \frac{\sigma^2}{2\alpha})dt + \sigma e^{\alpha t} dW + e^{\alpha t} \ln J dN_t$$

$\Rightarrow$

$$z_T = (\mu - \frac{\sigma^2}{2\alpha}) + \left(z_t - \mu + \frac{\sigma^2}{2\alpha}\right) e^{-\alpha(T-t)} + \sigma \int_t^T e^{-\alpha(T-s)} dW_s + \int_t^T e^{-\alpha(T-s)} \ln J dN_s \quad (3.2)$$

Before evaluating the expectation, I assume that even though the market is incomplete[2] the *drift term* $(\mu - \frac{\sigma^2}{2\alpha})$ in the solution includes the market price of risk, often referred in literature as λ and hence we can compute the expectation directly without adjusting for finding an appropriate equivalent Q-martingale measure.

Moreover, it is important to recall that the random processes W_t, J_t and N_t are all independent.

So,

$$\begin{aligned} F(t,T) &= \mathbb{E}_t[e^{z_T}|B_t] \\ &= \mathbb{E}_t\left[e^{(\mu - \frac{\sigma^2}{2\alpha})}.e^{\left(z_t - \mu + \frac{\sigma^2}{2\alpha}\right)e^{-\alpha(T-t)}}.e^{\sigma \int_t^T e^{-\alpha(T-s)} dW_s}.e^{\int_t^T e^{-\alpha(T-s)} \ln J dN_s}|B_t\right] \\ F(t,T) &= e^C \left(\frac{P(t)}{e^C}\right)^{e^{-\alpha(T-t)}} \mathbb{E}_t\left[e^{\sigma \int_t^T e^{-\alpha(T-s)dW_s}}|B_t\right].\mathbb{E}_t\left[e^{\int_t^T e^{-\alpha(T-s)} \ln J dN_s}|B_t\right] \quad (3.3) \end{aligned}$$

[2]Incomplete market refers to the fact that we have 1 tradable asset and 3 random processes, W_t, J_t and N_t to model it.

where, $C = (\mu - \frac{\sigma^2}{2\alpha})$.

The first expectation in the above equation is calculated explicitly with the help of probability theory and **Theorem 1** . In particular,

$$\mathbb{E}_t\left[e^{\sigma\int_t^T e^{-\alpha(T-s)dW_s}}|B_t\right] = e^{\frac{\sigma^2}{2}\int_t^T e^{-2\alpha(T-s)}ds}$$

$\Rightarrow$

$$\mathbb{E}_t\left[e^{\sigma\int_t^T e^{-\alpha(T-s)dW_s}}|B_t\right] = e^{\frac{\sigma^2}{4\alpha}[1-e^{-2\alpha(T-t)}]} \tag{3.4}$$

The remaining unknown quantity in the forward price expression is the expectation of the log-normal process. In the subsequent part of the paper, I present a closed form expression for this expectation from [0,t] and then extend this result from [t,T] as done by Alvaro and Marcelo in [7].

The following definitions are needed to simplify the arithmetic calculations. Let

$$\alpha_s \equiv e^{-\alpha(T-s)}\ln J_s \tag{3.5}$$

$$m_t = \int_0^t \alpha_s dN_s \tag{3.6}$$

$$L_t \equiv e^{m_t} \tag{3.7}$$

Equation (3.6) is rewritten equivalently, as

$$dm_t = \alpha_t dN_t \tag{3.8}$$

The process described above incorporates jumps and consequently, in order to write the SDE followed by L_t, I use the generalized Ito's Lemma for Jump processes as given in [8][3] according to which,

$$dL_t = \frac{\partial L_t(m_{t^-})}{\partial m_t}dm_t - \frac{\partial L_t(m_{t^-})}{\partial m_t}(m_t - m_{t^-})dN + (L_t - L_{t^-})dN \tag{3.9}$$

[3]Please refer to pages 176-177 on the reference for more details

Notice here that there is no double derivative term in the above expression because our underlying process given by equation(3.8) is a pure jump process without the drift and Wiener process terms as explained in [8] and we know that $dN \sim dt$.

For the purpose of clarity, the notation m_{t^-} refers to the time index immediately before a jump occurs so that if there is a jump in $\{m_t\}_{t>0}$, it is of magnitude α_t. Mathematically,

$$m_t = m_{t^-} + \alpha_t \tag{3.10}$$

Using the result above and the fact that $\frac{\partial L_t(m_{t^-})}{\partial m_t} = L_{t^-}$ Equation(3.7) is rewritten as,

$$\begin{aligned} L_t &= e^{m_{t^-}+\alpha_t} \\ &= L_{t^-}e^{\alpha t} \end{aligned}$$

and back substituting the transformed equations into Equation (3.9), we get

$$dL_t = L_{t^-}\alpha_t dN_t - L_{t^-}(\alpha_t)dN_t + (L_{t^-}e^{\alpha t} - L_{t^-})dN_t \tag{3.11}$$

$$dL_t = L_{t^-}(e^{\alpha t} - 1)dN_t \tag{3.12}$$

As stated earlier, we integrate the SDE above from [0,t] and then extend the result to the interval [t,T].

$$\begin{aligned} \int_0^t (dL_t) &= \int_0^t L_{\tau^-}(e^{\alpha\tau} - 1)dN_\tau \\ L_t - L_0 &= \int_0^t L_{\tau^-}(e^{\alpha\tau} - 1)dN_\tau \end{aligned}$$

Since $L_0 = 1$ by definition, we get

$$L_t = 1 + \int_0^t L_\tau(e^{\alpha\tau} - 1)dN_\tau$$

Taking expectation of the above and using linearity of the expectation operator with the fact that $\mathbb{E}_0[dN_t] = \lambda dt$, gives,

$$\mathbb{E}_0[L_t] = 1 + \int_0^t \mathbb{E}_0[L_\tau](\mathbb{E}_0[(e^{\alpha\tau}] - 1)dN_\tau \lambda d\tau \tag{3.13}$$

Letting $\mathbb{E}_0[L_t] = \eta_t$ and using separation of variables, the above equation is rewritten in the differential form as,

$$\frac{d\eta_t}{dt} = \eta_t(\mathbb{E}_0[e^{\alpha_t}])\lambda$$

$$\frac{1}{\eta_t} d\eta_t = (\mathbb{E}_0[e^{\alpha_t}])\lambda dt \tag{3.14}$$

Integrating both sides from 0→t and using the result that $\eta_0 = 1$ by definition of η_t, it is found that

$$\ln \eta_t - \ln \eta_0 = \int_0^t \mathbb{E}_0[e^{\alpha_\tau}])\lambda d\tau$$
$$\eta_t = e^{\int_0^t (\mathbb{E}_0[e^{\alpha_\tau}]-1)\lambda d\tau}$$

By combining equations (3.6) and (3.7), the above equation can be written in terms of the jump process dN_t as,

$$\mathbb{E}_0\left[e^{\int_0^t \alpha_\tau dN_\tau}\right] = e^{\int_0^t (\mathbb{E}_0[e^{\alpha_\tau}]-1)\lambda d\tau} \tag{3.15}$$

As stated earlier, the result above can be easily extended for the interval [t,T] by using the fact that for any integrable function on the interval [0,T],

$$\int_t^T f(x)dx = \int_0^T f(x)dx - \int_0^t f(x)dx$$

where, $t \in (0, T)$.

Hence, this simple extension yields the required result,

$$\mathbb{E}_t\left[e^{\int_t^T \alpha_\tau dN_\tau}\right] = e^{\int_t^T (\mathbb{E}_0[e^{\alpha_\tau}]-1)\lambda d\tau} \tag{3.16}$$

Therefore, the log-normal expectation in the forward price equation (3.3) simplifies to evaluating the integral given by equation (3.16).

In order to evaluate the integral, the following definition is given for simplification purposes.

$$g(s) = e^{-\alpha(T-s)} \tag{3.17}$$

So then

$$\begin{aligned}
\mathbb{E}_0[e^{\alpha_s}] &= \mathbb{E}_0[e^{g(s)\ln J_s}] \\
&= \mathbb{E}_0[e^{g(s)\phi_s}] \\
&= e^{\mu_J h + \sigma_J^2 h^2}
\end{aligned}$$

Recall that $\mu_J = 1.35$ and $\sigma_J^2 = 0.3162$ are known parameters given by the distribution fitting tool. For notation purposes, I denote them simply by μ and σ in what follows.

Plugging back the above expectation into the integral term in equation (3.15), I obtain

$$\mathbb{E}_t[e^{\int_t^T \alpha_s dN_s}] = e^{\int_t^T (e^{\mu g + \frac{\sigma^2}{2} g^2} - 1)\lambda ds}$$

$$\mathbb{E}_t[e^{\int_t^T \alpha_s dN_s}] = \frac{e^{\int_t^T e^{(\mu g + \frac{\sigma^2}{2} g^2)} \lambda ds}}{e^{\lambda(T-t)}} \tag{3.18}$$

Finally, combining the above result with equation (3.3) the explicit forward spot price equation is derived to be,

$$F(t,T) = e^C \left(\frac{P(t)}{e^C}\right)^{e^{-\alpha(T-t)}} e^{\frac{\sigma^2}{4\alpha}[1-e^{-2\alpha(T-t)}] + \int_t^T e^{(\mu g + \frac{\sigma^2}{2} g^2) - \lambda(T-t)}} \tag{3.19}$$

Table 3.1: Current Forward Prices with different Expiries

T (in years)	F(0,T)
5	44.78
2	48.23
1	49.43
1/2	50.05
1/4	50.36
1/12	50.57

3.2 SIMULATIONS FOR FORWARD PRICES

In this last section, some forward prices with varying dates of expiry are calculated by integrating equation (3.19) numerically and the results are tabulated below.

The forward prices observed make perfect intuitive sense and turn out to be as expected. This is because for larger times to expiry the amount of risk undertaken by the investor who holds a long or short[4] position in the forward, is higher. Consequently, he/she would only be motivated to take a higher amount of risk if the initial investment, which in this case refers to the forward price, is lower.

[4] The term 'long' in the financial jargon refers to buying and 'short' refers to selling.

BIBLIOGRAPHY

[1] A. Lari-Lavassani, A. A. Sadeghi, and A. Ware, “Mean reverting models for energy option pricing,” Tech. Rep., University of Calgary, 2001.

[2] M. Davison, C.L Marcus, and B.Anderson, “Development of a hybrid model for electrical power spot prices,” in *IEEE Transactions on Power Systems*, May 2002, vol. 17, pp. 257–264.

[3] Robert S.Pindyck and Daniel L.Rubinfeld, *Econometric Models and Economic Forecasts*, McGraw-Hill, Ohio, USA, 1998.

[4] J. C. Hull, *Options, Futures and Other Derivatives*, Prentice Hall, Upper Saddle River, NJ, 2006.

[5] T. C. Gard, *Introduction to Stochastic Differential Equations*, Dekker, Newyork, NY, 1988.

[6] G. Skorohod, *Stochastic Differential Equations*, Springer-Verlag, Newyork, NY, 1972.

[7] A. Cartea and M. G. Figueroa, “Pricing in electricity markets: a mean reverting jump diffusion model with seasonality,” in *Applied Mathematical Finance*, Dec 2005, vol. 12, pp. 313–335.

[8] Alison Etheridge, *A course in Financial Calculus*, Cambridge University Press, Cambridge, UK, 2002.