Lappeenranta University of Technology
Faculty of Technology
Department of Mathematics and Physics

A comparison of electricity spot prices simulation using ARMA-GARCH and mean-reverting models

The topic of this Thesis was approved
by the Department of Mathematics and Physics

...March 10, 2010

Supervisors: Prof. Ph.D. Heikki Haario and Ph.D. Tuomo Kauranne.
Examiners: Prof. Ph.D. Heikki Haario and Ph.D. Tuomo Kauranne.

Lappeenranta, March XX, 2010

Muhammad Naeem
Ruskonlahdenkatu 13-15, C8
53850 Lappeenranta
Muhammad.Naeem [at] lut.fi

Abstract

Lappeenranta University of Technology
Department of Mathematics and Physics

Muhammad Naeem

A comparison of electricity spot prices simulation using ARMA-GARCH and mean-reverting models

Master's thesis

2010

56 pages, 54 figures, 14 tables

Supervisors: Prof. Ph.D. Heikki Haario and Ph.D. Tuomo Kauranne.

Key words: time series, electricity spot price, volatility, ARMA-GARCH

The aim of this work is to compare two families of mathematical models for their respective capability to capture the statistical properties of real electricity spot market time series. The first model family is ARMA-GARCH models and the second model family is mean-reverting Ornstein-Uhlenbeck models. These two models have been applied to two price series of Nordic Nord Pool spot market for electricity namely to the System prices and to the DenmarkW prices. The parameters of both models were calibrated from the real time series. After carrying out simulation with optimal models from both families we conclude that neither ARMA-GARCH models, nor conventional mean-reverting Ornstein-Uhlenbeck models, even when calibrated optimally with real electricity spot market price or return series, capture the statistical characteristics of the real series. But in the case of less spiky behavior (System prices), the mean-reverting Ornstein-Uhlenbeck model could be seen to partially succeeded in this task.

Acknowledgements

Thanks to Allah almighty for making it possible.
First of all, I would like to express my gratitude to the Department of Mathematics at Lappeenranta University of Technology for the scholarship during my studies, otherwise my studies would not be possible. Specially thanks to Head of Department Heikki Haario for accepting my request for the scholarship, at right time.

I would like to express my gratitude for valuable supervising support from Ph.D. Tuomo Kauranne and Matylda Jabłońska, who stated directions of this work. Without them I could have not completed this work.

I would like to say thanks to my friends at departments of mathematics for cooperating with me during my studies.
My special thanks goes to my parents for their prayers and support at each step of my life.

Lappeenranta; Feb 23, 2010
Muhammad Naeem

Contents

1 Introduction

A large number of countries around the world, including the US, have recently started restructuring processes in their energy sectors. However, the pace and aim of the improvements varies across countries. With the introduction of competitive wholesale electricity markets, and power derivative contracts, both exchange-traded and over the counter (OTC), providing a number of different contract provisions to meet the needs of the electricity market participants. Electricity is strongly characterized by its very limited storability and transportability and may be considered as a flow commodity. In other words, arbitrage across time and space, which is based on transportation and storability, is seriously limited, if not completely eliminated in power markets. If the relation across time and space provided by arbitrage broke down, we would expect spot prices to be highly dependent on temporal and local supply and demand conditions. It is also expected to affect the relationship between electricity spot and derivative prices, decisively.

Keeping balance between demand and supply is very challenging due to non-storable nature of electricity which makes electricity a distinct commodity. In other words, demand of electricity at different times and on different dates becomes a cause of price fluctuations. Electricity has also transportation constraints in the form of transportation losses and capacity limits in the transmission lines, which can make non-profitable or impossible the transmission of power/electricity among certain regions. These constraints make electricity contracts and prices highly local which means that prices vary from region to region according to the weather conditions and capacities of local plants etc.

For the purpose of electricity derivative pricing several papers [12], [9] and [6] have been published which have pointed some general features of the electricity price behavior that should be considered. Especially, some papers stated that a model for electricity pricing should include a form of time varying volatility and possibility of spikes (jumps in prices). On the other hand, seasonal behavior of electricity prices, and its reversion to mean has been described in [20]. The bulk of the research on modeling electricity spot and derivative prices is still to be continued because only very limited and tentative research papers have been published to date.

In this study we will investigate the goodness of fit of two models: The first model belongs to the family of ARMA/GARCH and the second model belongs to mean reverting processes, called Ornstein-Uhlenbeck process. Our aim is to get a normal distribution of residuals and to see which method seems to fit better. We have found one thesis [17] in the literature about the investigation of goodness of fit of two model: The first one is ARMA/GARCH and the second one is two-state Markov Regime Switching model assuming that the base regime follows an Ornstein-Uhlenbeck process.

We have divided our work in several sections. First section gives an introduction to electricity market, second section is about the theoretical background on ARMA-GARCH models, practical results of ARMA-GARCH can be seen in the section three. Both theoretical and practical part about Ornstein-Uhlenbeck process have been written down in section four and results and conclusion have been given in section five.

2 Introduction to electricity markets

2.1 The Key Features of Electricity Prices

Since we want to model the spot price of electricity and we believe that in competitive electricity markets a proper representation of the dynamics of spot prices becomes a necessary tool for optimal design of supply contracts and trading purposes. As discussed in [9] the non-storability of electricity indicates the breakdown of the spot-forward connection and, in turn, the possibility of deriving the basic properties of spot prices from the analysis of forward curves. Moreover, it can be seen empirically in different markets such as Nord Pool, the U.K. and U.S. markets, electricity forward curve moves are less dramatic than spot price changes.

A first feature of electricity prices is mean reversion towards an equilibrium level and may be constant, periodic or periodic with a trend. A 127-year period for crude oil and bituminous coal and a 75-year period for natural gas has been analyzed by Pyndick [21]. According to him, prices deflated (and denoted by their natural logarithms), demonstrating mean-reversion to a stochastically fluctuating trend line. With a few years horizon in mind, in the case of electricity, we suggest to represent the diffusion part of the spot prices as mean-reverting to a non-probabilistic periodical trend driven by seasonal effects. It can be observed that the mean reversion is more or less distinct, across different markets.

The second characteristic of electricity prices is seasonality. Weather conditions and business activities are the major factors that explain seasonality of the electricity spot prices. Data exhibit various seasonal patterns such as intra-daily, weekly and monthly seasonality. We have assumed in this study that the components that generate the seasonality are non-probabilistic and we are using the daily average system prices, and monthly and weekly patterns are considered. The importance of deterministic patterns in the dynamics of electricity spot prices was studied by Lucia and Schwartz [16]. They have analyzed the Nordic spot, future and forward prices and concluded that the seasonal systematic pattern throughout the year, indeed, is of high importance in describing the shape of the future/forward curve. They have proposed that a simple sinusoidal function is adequate in order to capture the seasonal pattern of the future and forward curve directly implied by the seasonal behavior of the electricity spot prices.

Thirdly, extremely high volatility is also one of the distinct features of electricity spot prices. Moreover, volatility observed in electricity spot prices is exceptional and not comparable with other commodities. Volatility means the standard deviation of the returns on a daily level. Weron [24] has applied a standard concept of volatility and obtained:

1. notes and treasury bills less than: 0.5%
2. stock indices: 1-1.5%
3. commodities like natural gas or crude oil: 1.5-4%
4. very volatile stocks: not more than 4%
5. and electricity up to 50%

Transmission and storage limitations result in high volatility patterns thus setting equilibrium prices in real time is one of the requirements of the market. Provisional imbalance of supply and demand cannot be corrected in short time. Therefore, electricity market exhibits extreme changes when compared to other commodities.

The fourth characteristic of the price process, expectedly, is existence of small random moves around the average trend, which exhibit the temporary supply/demand imbalances in the network. This effect is locally unpredictable and may be denoted by a white noise term affecting daily price variations.

The fifth and intrinsic characteristic of power price processes is presence of spikes, namely one (or several) upward jumps shortly followed by a steep downward move, for example, when summer is over or the generation outage appears.

2.2 The Nordic Market

Nord Pool, known as an electricity market for Nordic region, had been founded in 1992 as a result of the Norwegian energy act of 1991 that formally made the way for the deregulation of the electricity sector of Norway. At first it was only a Norwegian market, but later Sweden (1996), Finland (1998) and Denmark (2000) joined in. Nord Pool is known as pioneer in power exchange. In this market, players from outside the Nordic region can participate on equal terms with 'local' exchange members. To join the spot market, called Elspot, a grid connection enabling power to be delivered to or taken from the main grid is required. Nowadays, there are more than 330 market participants from over 10 countries active on Nord Pool. Nord Pool Spot provides to generators, suppliers/retailers, traders, large consumers, energy companies, TSOs and financial institutions a market place on which they can buy or sell physical power [25]. Nord Pool has become very successful because of several factors. Firstly, the industry structure consists of more than 350 generation companies. Secondly, the system operator, government, and regulators are cooperating well with each other, which makes it distinct from other continental European countries [24].

2.3 Price Setting at Nord Pool

The spot price is a result of a two-sided uniform price auction for hourly time intervals at Nord Pool. It is determined from the various bids presented to the market administrator up to the time when the auction is closed.

Elspot is known as the market for trading power for physical delivery. Elspot is also called a day-ahead market. One-hour long physical power contracts are traded in Elspot, and the minimum contract size is 0.1MWh. Each day, the market participants submit to the market administrator their offers at 12 p.m. for the next 24 hours starting at 1 a.m. of the next day. This offer is provided electronically via the internet (Elweb) with a resolution of one hour, *i.e.* one for each hour of the next day. Such an offer should contain both price and volume of the bids.

Bidding can be done in three ways at Elspot. One way of bidding is hourly bidding consisting of a pair of price and volume for each hour. For the second kind of bidding, the price and the volume are fixed for a number of consecutive hours. Finally, flexible hourly bidding means a fixed price and volume sales bid where the hour of the sale is flexible and determined by the highest (next day) spot price that is above the price indicated by the bid.

The fact is that power generators are also willing to buy power for their large consumers. Actually, they intend to increase their profit and this is achieved only if they buy electricity during low price periods. By 12 p.m. Nord Pool closes the bidding for the next 24 hours and for each hour continues to make cumulative supply and demand curves. The system spot price for that individual hour is decided/specified as the price where supply and demand curves cross. This is known as an equilibrium point. No transaction will take place for that particular hour if the data does not define the equilibrium point. Moreover, the system price is determined by the equilibrium point independent of potential grid congestions. And area prices will differ from system price only for those hours when transmission capacity in the central grid is limited [24].

3 Theoretical background for ARMA-GARCH Approach

Collection of observations of a variable that become available sequentially through time are called time series data, or simply, time series [4]. The order of observations is represented by a subscript t. Therefore, we write z_t as a t_{th} observation of time t and a proceeding observation is written as z_{t-1}, and succeeded observation as z_{t+1}. Time series are applicable in different fields. Examples of fields are economics, sociology, meteorology, medicine, vibrating physical systems, seismology, oceanography, geomorphology, astronomy, *etc.*

3.1 Box-Jenkins Time Series Models

George Box and Gwilyn Jenkins introduced Autoregressive Integrated Moving Average (ARIMA) [23] time series models in 1970. These models are mathematical models and used for short term forecast of 'well behaved' data and find the best fit of time series in order to get a forecast. The Box-Jenkins approach uses an iterative model-building strategy that consist of four stages in general. In the first stage we identify the model or we find the order of the model that we are using and in the second step, estimation of the model coefficients is done. Then model checking and at the end forecast is obtained with that model.

3.1.1 Autoregressive Model

An autoregressive model [23] is a model where the current value of a variable z_t depends only upon the value that the variable took in previous periods plus an error term. An autoregressive model is denoted by AR(r), where r is the order of the process and the order of the process represented number of parameters that need to be estimated. An rth order autoregressive process is written as

$$z_t = C + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \ldots + \phi_r z_{t-r} + a_t \tag{1}$$

where z_t is the series and C is constant. Also, $\phi_1, \ldots, \phi_r$ are the autoregressive parameters which describe the effect of a unit change in two consecutive time series observations (z_{t-1} on z_t) and which need to to be estimated. The a_t term is a white noise or error term assumed to be normally and independently distributed with mean zero and variance constant over time, ($a_t \sim N(0, \sigma^2)$) and no significant autocorrelation. The constant mean condition does not suggest that any restriction should be imposed on ϕ_1 but for the time series to be stationary it is necessary that roots of its characteristic equation lie outside the unit circle. For instance, in case of AR(1) model that would mean $|\phi_1| \leq 1$. If the the observation at time $t - k$, z_{t-k}, has some effect on the observation at time t, z_t, then autocorrelation would be significant.

3.1.2 Moving Average Model

An extension of the AR(1) model would be to include past shocks to see if they can improve on the time series representation of the data. Now we can modify the AR(1) model to obtain an Autoregressive Moving Average Model of order (1,1) [23]

$$z_t = \phi_1 z_{t-1} + a_t - \theta_1 a_{t-1}$$

where a_{t-1} is the error at period $t - 1$, and θ_1 is called the *moving average parameter*, which describes the effect of the past error on at, and which needs to be estimated. A special model is obtained after removing term z_{t-1} from the model is known as Moving Average Model of order 1 and shows the current value of the series z_t as a linear function of the current and previous shocks a_t and a_{t-1} and mathematically, we express a first-order Moving Average Model

$$z_t = a_t - \theta_1 a_{t-1}$$

We observe that the autocorrelations for the AR(1) process die out gradually but for MA(1) process they die out abruptly. For MA(1) process the error at time t, a_t, will influence the observation at time t and $t+1$ but will not have any effect beyond period $t+1$.

A higher order moving average model can be written as

$$z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_p a_{t-q}$$

3.1.3 ARMA-GARCH Models

A model with a combination of autoregressive terms and moving average terms is known as mixed autoregressive moving average model [10]. We use notation ARMA(p,q) to represent these models for our convenience, where p is the order of the autoregressive part and q is the order of the moving average part. The use of mixed ARMA(p,q) model achieves great parsimony in model specification

$$z_t = C + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \ldots + \phi_p z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q}.$$

The behavior of stationary a time series can be widely described by the ARMA(p,q). A forecast generated by an ARMA(p,q) model will depend on current and past values of the response as well as current and past values of the errors (residuals). The order of autoregressive and moving average terms in an ARMA model are determined from the pattern of sample autocorrelation and partial autocorrelations.

The information recovery process in time series analysis uses past observations to derive estimates of current or future values of the dependent variable. The ARMA models are used in different kind of applied problems including time series analysis. The basic assumptions on the error terms include zero mean and constant variance for the the traditional ARMA estimation. But for some time series homoskedastic assumption of constant variance does not hold. Time series models for which the constant variance assumption does not hold are called hetroskedastic. We have two models which deal with heteroskedasticity, the first one is autoregressive conditional heteroskedasticity (ARCH) [8] introduced by Engle and the second one is its extension, namely generalized autoregressive conditional heteroskedasticity (GARCH) [2] introduced by Bollerslev. Conditional variance σ_t^2 is considered as time dependent. Therefore, if $a_t \sim N(0,1)$ then $u_t = a_t \sigma_t$, with

$$\sigma_t^2 = C + \sum_{i=1}^{r} \alpha_i u_{t-i}^2$$

As a_t is white noise which is assumed to be normally distributed ($a_t = N(0, 1)$), so that u_t will also be normally distributed with zero mean and variance σ_t^2. In practical applications it turns out that the order of the calibrated model is rather large $u_t = a_t\sigma_t$ and the current variance appears to be sometimes dependent not only on past squared disturbances, but the past variance of the errors as well. This comes as a GARCH model

$$\sigma_t^2 = C1 + \sum_{i=1}^{r} \alpha_i u_{t-i}^2 + \sum_{i=1}^{m} \beta_i \sigma_{t-i}^2$$

3.2 Autocorrelation and Partial Autocorrelation

In practice, it is difficult to know the values of theoretical autocorrelation [10] and partial autocorrelation [23] of the underlying stochastic process. Therefore, in identifying a tentative model one must use the sample autocorrelation and partial autocorrelation functions to check if they are similar to those of typical models for which the parameters are known. Since autocorrelation and partial autocorrelation are only estimates, they are subject to sampling error, and as such will never correspond exactly to the underlying true autocorrelations and partial autocorrelations. Sample autocorrelation is denoted by γ_k and sample partial autocorrelation is denoted by $\hat{\phi_{kk}}$.

3.2.1 Autocorrelation

Autocorrelation is the measure of correlation of two consecutive observations of time series. These are the statistical measures that show how a time series is related to itself over time. It shows a degree of similarity between a given time series and a lagged version of itself over successive time intervals. It is the same as calculating the correlation between two different time series, except the same time series is used twice, once in its original form and once lagged by one or more time periods. Let
γ_k : autocorrelation for k period lag
Y_t : value of time series at time t
Y_{t-k} : value of time series at time t-k
μ : mean of the time series
then

$$\gamma_k = \frac{E[(y_{t-1} - \mu)(y_t - \mu)]}{\sigma_y^2}.$$

If the time series is stationary, γ_k decreases rapidly to zero. If the time has trend, γ_k declines toward zero slowly. If seasonal pattern exists in the time series, the value of γ_k will be significantly different from zero at $k = n * 4$ for quarterly data, k = n*12 for monthly data, where $n = 1, 2, \ldots, N$. A plot of autocorrelation is called an Autocorrelation Function (ACF) or a autocorrelogram.

3.2.2 Partial Autocorrelation

The partial autocorrelation function (PACF) measures the correlation between current observation and an observation k periods ago, after controlling for observations at intermediate lags ($lags < k$), *i.e.* the correlation between y_t and y_{t-k}, after removing the effects of $y_{t-k+1}, y_{t-k+2}, \ldots, y_{t-1}$. For example the PACF for lag 4 would measure the correlation between y_t and y_{t-4} after controlling for the effects of y_{t-1}, y_{t-2}, and y_{t-3}. Autocorrelation and partial autocorrelation coefficients are equal at lag 1, since there are no intermediate lag effects to eliminate. The partial autocorrelation at lag k is given by

$$\hat{\phi_{kk}} = \frac{y_k - y_{k-1}^2}{1 - y_{k-1}^2}$$

3.3 Stationarity

If the process has the mean, variance and autocorrelation structure constant over time then process is known as stationary process. It can be assessed by run sequence plot. Run sequence plot is the graph that displays observed data in a time sequence. Thus a time series is known as stationary if there is no systematic change in mean for instance series has no trend, and if there is no systematic change in variance or series has no periodic variations. Stationary time series are considered best in modern theory of time series, and for this reason time series analysis often requires one to transform a non-stationary series to a stationarity one [23].

3.3.1 Making data series stationary

The analysis of the time series demands that the series is stationary and, in particular, that the variance (the volatility of the series) is constant over time. There are some possible transformation that could be used to induce a constant variance. The basic idea is to transform the data such that an originally curved plot straighter, and at the same time make variance constant over the whole series. Two of the frequently used transformations are logarithmic transformations and the square root transformations. We can make the data stationary by *differencing* also. Usually, the method of stationarizing data is decided on based on their graprical representation or on the plot of ACF and PACF [23].

3.3.2 Testing Stationarity

Dickey and Fuller (1979) developed the basic test for testing the stationarity. The test is called Dickey-Fuller (DF) [7] test for stationarity. The aim is to test the null hypothesis. Thus the DF test checks the null hypothesis that $\phi = 1$ (the process contains a unit root, *i.e.* its current realization appears to be an infinite sum of past disturbances with some

starting value y_0) versus the one-side alternative $\phi < 1$ (the process is stationary). The test statistics look as follows
H_0: series contains a unit root
H_1: series is stationary

$$DF = \frac{1 - \hat{\phi}}{\hat{SE}(1 - \hat{\phi})}$$

where H_0, H_1 and SE are the null hypothesis, alternative hypothesis and Standard Error respectively.

The DF test statistic does not follow the t-distribution under the null hypothesis, because of non-stationarity, but rather DF test follows a non-standard distribution. Experimental simulations were used to derive critical values for comparison.

A similar test to DF [7] test is known as the Phillips-Perron(PP) [19] test. The PP test is similar to DF test. It incorporates an automatic correction to the DF test procedure for autocorrelated residuals to be used. However, this test relaxes assumptions about lack of autocorrelation in the error term. Its critical values used for comparison are the same as for Dickey-Fuller test.

Weakness of Dickey-Fuller and Phillips-Perron-type tests: main limitation is that the tests have low power if the process is stationary but with a root close to the non-stationary boundary. A problem arises when the process has the ϕ value close to the non-stationarity boundary, *i.e.* $\phi = 0.95$. This kind of process is, by definition, still stationary for DF and PP tests. These tests often fail to distinguish for the values $\phi = 1$ or $\phi = 0.95$, if the size of the sample is small. Therefore, to avoid this failure of DF and PP-type tests, there is another test called KPSS [14] test (Kwiatkowski, Phillips, Schmidt and Shin, 1992) with the opposite null hypothesis *i.e.* stationarity as a null hypothesis:
H_0: series is a stationary
H_1: series is not stationary

$$KPSS = \frac{\sum_{i=1}^{n} \hat{S}_i^2}{N^2 \hat{s^2}}$$

where $\hat{S_i^2} = \sum_{j=1}^{i} e_j$, e are residuals given as $e = [e_1, e_2, \ldots, e_T]'$, and $\hat{s^2}$ represents an estimate of the long run variance of the residuals. We reject the null hypothesis when KPSS is large, since that is evidence that the series wanders from its mean. If the AR model is known, we can test the stationarity by evaluating the roots of the characteristic equation and, if all of roots lie outside the unit circle, the given model is stationary. For instance, $x_t = 3x_{t-1} - 2.75x_{t-2} + 0.75x_{t-3} + u_t$ does not meet stationarity requirement because out of its roots 1, $\frac{2}{3}$ and 2 only one lies outside the unit circle.

3.4 Box-jenkins Model Stages

There are four stages in Box-Jenkins [23] model building:

1. identification of the preliminary specification of the model,
2. estimation of the parameter of the model,
3. diagnostic checking of model adequacy,
4. forecasting future realizations.

3.4.1 Identification

In the identification stage the first task is to obtain the order of the ARMA first and then order of the GARCH, if ARCH effect is there. In the the identification stage we use the autocorrelation and partial autocorrelation functions to identify the order of the model. This step may give really different results dependent on subjective look of the researcher, and requires a great deal of judgment. Model identification is a stage where statistically inefficient methods are used since there is no precise formulation of the problem. We use the graphical methods where the judgments are exercised. The first task is to identify what is the appropriate model from general ARMA family. This is done when the data are stationary. Therefore, once stationarity, seasonality and trend have been addressed, one needs to identify the order of ARMA(p,q). Here the plots of sample autocorrelation and the sample partial autocorrelation are compared to the theoretical behavior of these plots when the order is known. The ARMA model identification is based on autocorrelation and partial autocorrelation function values. Naturally, the model whose values are closest to calculated ones is chosen. Understanding the concept of autocorrelation can be tested by trying to conclude the mentioned theoretical values; here they are presented in Table 1.

Table 1: Theoretical characteristics of ACF and PACF for basic ARMA models

Model	Theoretical r_k	Theoretical r_{kk}
AR(0)	All zero	All zero
AR(1)	Vanish toward zero	Zero after 1st lag
AR(2)	Vanish toward zero	Zero after 2nd lag
MA(1)	Zero after 1st lag	Vanish toward zero
MA(2)	Zero after 2nd lag	Vanish toward zero
ARMA(1,1)	Vanish toward zero	Vanish toward zero

3.4.2 ARMA-GARCH Model Identification using SLEIC

This section illustrates a way to find a good ARMA-GARCH model for the Nord Pool data. It also describes a criteria function build on Schwarz's Bayesian information criteria (SBIC) [22]. The output of Engel's and Ljung-Box tests are given in a binary form, 1 or

0. Here, 0 indicates lack of GARCH/ARMA effect in the series, while on the other hand 1 indicates its presence. The SBIC is formulated as follows:

$$SBIC = log(\sigma^2_{res} + \frac{k}{L}.log(L)$$

where:

Term σ^2_{res} is the variance of residuals between returns and its fitted model k number of parameters of GARCH model L length of tested time series

A new information criteria function, called SLEIC [22] is being suggested here as follows.

$$SLEIC = [SBIC \cdot (1 + \frac{\alpha}{2N} \sum_{i=1}^{N} (H_{1,i} + H_{2,i}))]$$

SLEIC: information criteria based on Schwartz-Bayesian information criteria,Ljung-Box test and Engel's test

$H_{1,i}$: vector of logical outputs for Ljung-Box test, $i = 1, 2, ..., 2L$

$H_{2,i}$: vector of logical outputs for Engel's test, $i = 1, 2, ..., 2L$

α: importance coefficient of Ljung-Box and Engel's tests

N: Number of lags analyzed by Engel's/Ljung-Box test

To find an appropriate model for System and DenmarkW price series, we maximize SLEIC function while varying orders p, q, r and m of GARCH(r, m) and ARMA(p, q) models.

$max_{P,Q}\, SLEIC(res, k, H_1, H_2)$

3.4.3 Model Estimation

Parameters are estimated, after selecting a particular model from the general class of models. Then by applying various diagnostic checks, one can determine whether or not the model adequately represents the data. If any inadequacies are found, a new model must be identified and the cycle of the identification, estimation and diagnostic checking are repeated. One can make logical modifications to come up with a formulation which more adequately depicts the behavior of the series, just by studying the residual patterns of an inappropriate model.

4 Modeling Electricity Spot Prices using ARMA and GARCH

4.1 Statistical Analysis of Nord Pool Electricity prices Series Data

In this section we investigate the general statistical features of the given time series: System and DenmarkW electricity prices .

4.1.1 Data description and basic statistics

The original data set consist of 3712 daily observations of the Nord Pool electricity prices (7 days a week) from 01 Jan 1999 to 28 Feb 2009. Moreover, about six month data is missing from the DenmarkW series. To avoid the data equal zero we have added 0.1 to the price series data. This operation does not change the overall characteristics of the data. We first plot the original price series of the Nord Pool region. Figure 1 is the graph of the electricity prices of the Nord Pool system and area prices. Figure 2 shows the system and the area prices separately. Now it is clear from Figure 2 that all the region are showing seasonal increase in prices. From here we are going to select price series of System and DenmarkW for further analysis.

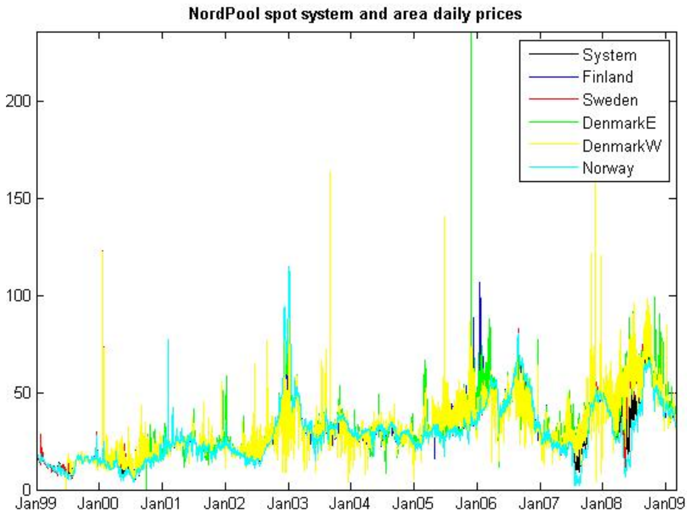

Figure 1: Nord Pool spot system and area prices.

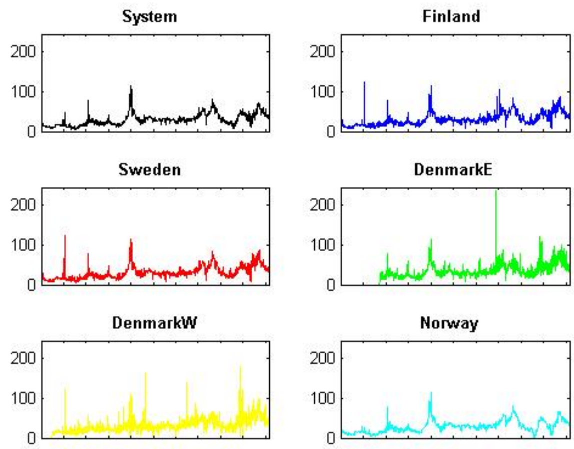

Figure 2: Nord Pool spot system and area price separately.

Usually, first information about time series data comes from the following graphical representation. Now we plot prices for System and DenmarkW in Figure 3 and returns for system and DenmarkW in Figure 4.

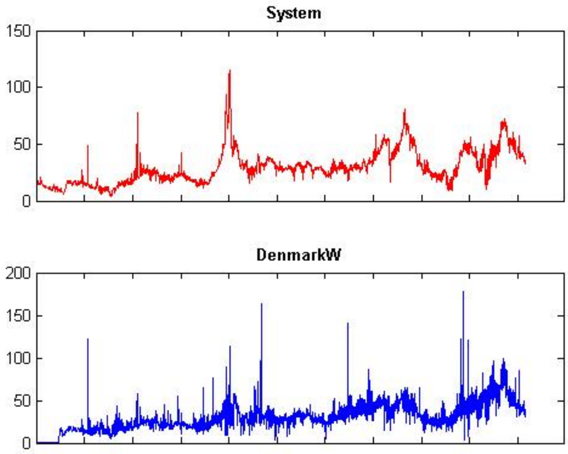

Figure 3: Plot of System and DenmarkW prices.

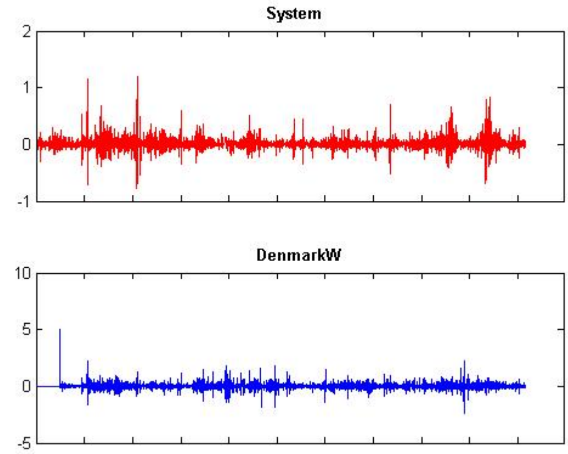

Figure 4: Plot of System and DenmarkW returns.

The aim of using the return series is to get stationarity. The logarithmic return series is based upon the following formula.

$$r_t = ln\frac{Y_t}{Y_{t-1}}$$

where

1. r_t is return for any time t
2. Y_t is the price of asset at moment t

3. Y_{t-1} is the price at moment $t-1$

High volatility can be seen from Table 2 for both system and DenmarkW series.

Table 2: Basic statistics for System and DenmarkW prices and price log returns.

case	System prices	DenmarkW prices	System return	Denmark return
count	3712	3531	3712	3531
mean	29.5141	30.8579	$2.0423.10^{-4}$	0.0016
std	14.7107	16.8975	0.1012	0.2715
max	114.7137	178.3012	1.1860	5.0484
min	3.9867	0.1	-0.7708	-2.4102

4.2 Normality

The next step is to verify the type of distributions that both System and DenmarkW prices and return series have. In financial time series often we get log-normal distribution for prices and normal one for price returns. But for the case of electricity prices neither prices nor returns follow the theoretical distributions. One reason is that electricity cannot be stored in warehouses. Therefore, let us investigate the behavior of System and DenmarkW series. We have already obtained the plots for both prices and return series, now we plot normalized histograms for both series against theoretical normal probability density functions (PDF).

Following Figure 5 and Figure 6 show the normalized histograms for both series.

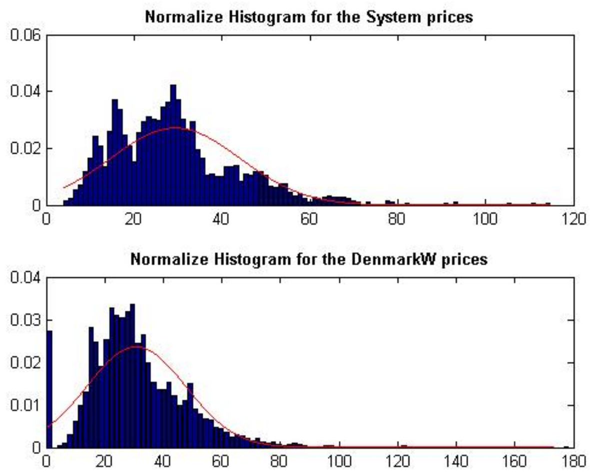

Figure 5: Normalized histogram for System and DenmarkW prices.

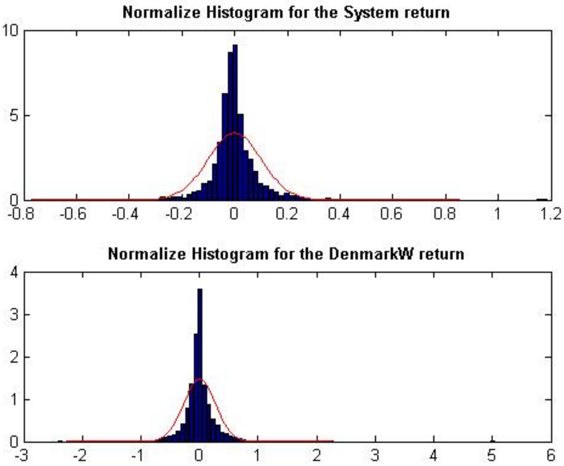

Figure 6: Normalized histogram for System and DenmarkW returns.

Now we compute the two most common parameters used for comparing a given probability distribution with the normal one, that is kurtosis and skewness. The results can be seen in Table 3. Now, skewness should be 0 and kurtosis should be 3 in order to see normal distribution in our logarithmic return series. We can easily see that neither prices nor price returns follow normal distribution. The final step is to perform a formal statistical test for verifying normality of a given distribution. Here we select the Lilliefors test with statistic calculated as follows:

$$L = max_x |scdf(x) - cdf(x)|$$

where scdf is called empirical cumulative density function estimated from the sample and cdf is known as normal CDF with mean and standard deviation equal to the mean and standard deviation of the sample. Results can be found in Table 3 - the null hypothesis was rejected for both series of prices and return with 5 percent significance level.

Table 3: Basic statistics for System and DenmarkW prices and price log returns.

case	System prices	DenmarkW prices	System return	Denmark return
skewness	1.2176	1.1643	1.5781	2.1171
kurtosis	5.6114	7.3818	24.0999	44.7962
Lillifors test H_0	rejected	rejected	rejected	rejected

We have seen that the values of skewness and kurtosis for our series are different from the theoretical values. And Lilliefors test has also rejected normality. It means that neither System and DenmarkW prices nor their returns follow the normal distribution.

4.3 Stationarity Test

In this section different test have been conducted to check stationarity in our data. For checking the stationarity here we have used Dickey-Fuller (DF) test, Phillips-Perron (PP) test and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test for checking whether our data are stationarity or not. We have performed all these tests on both price and return series. As can be seen from Table 4, for all System prices, System return and DenmarkW return null hypothesis have been rejected by using DF test and PP test and accepted by using KPSS test, which clearly indicate that our data is stationary. Moreover, for the DenmarkW price series, null hypothesis was rejected by using DF, PP and KPSS test, which mean that our DenmarkW series raises some doubts about stationarity. However, we are going to use return series in order to get a tentative model.

Table 4: Results of DF , PP and KPSS tests for System and DenmarkW prices and price returns.

case	DF test	PP test	KPSS test
System prices	rejected	rejected	accepted
DenmarkW prices	rejected	rejected	rejected
System return	rejected	rejected	accepted
DenmarkW return	rejected	rejected	accepted

4.4 Identification of the model

After obtaining stationarity the next stage is to identify a suitable model for our data. Identification is the key step in time series model building. The two most useful tools for time series model identification are the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF). The procedure in the identification stage are inexact and require great deal of judgement [23]. The sample ACF and PACF never match the theoretical autocorrelations and partial autocorrelation. There is no exact deterministic approach for identifying an ARMA model.

In the first step we have found the ACF and PACF of the original price series. Following Figure 7 and Figure 8 are showing the result of the ACF and PACF of System and DenmarkW prices respectively. Here we can see ACF of the System price series is not dying out quickly which is a clear sign of non-stationarity. So at this moment it is difficult to choose appropriate model for our data. In the next step we are going to apply logarithmic transformation to our series in order to get stationarity.

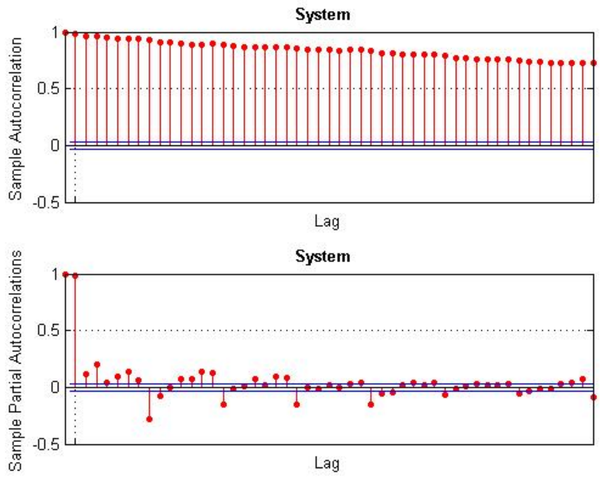

Figure 7: ACF and PACF of System price series.

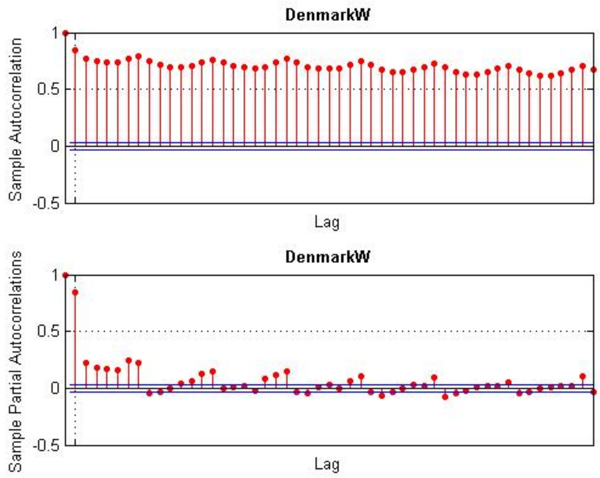

Figure 8: ACF and PACF of DenmarkW price series.

After applying the logarithmic transformation, again ACF and PACF of the transformed series are found as shown in Figure 9 and Figure 9. Now, seasonal pattern can be seen from the plot of ACF and PACF after every lag 7. So still data is not fully stationary. Thus in the next step we are going to apply differencing in order to get the stationary series.

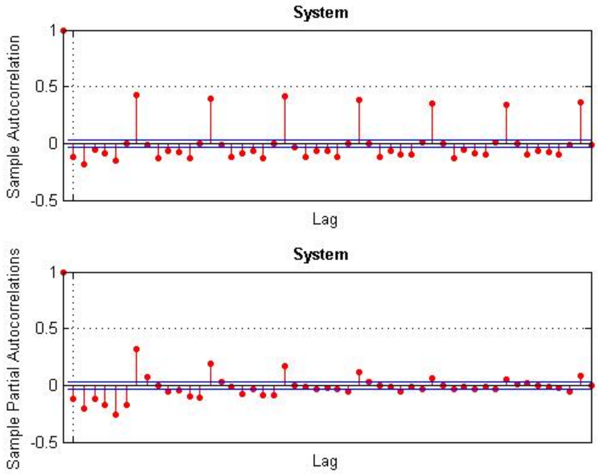

Figure 9: ACf and PACF of logarithm transformed of System price series.

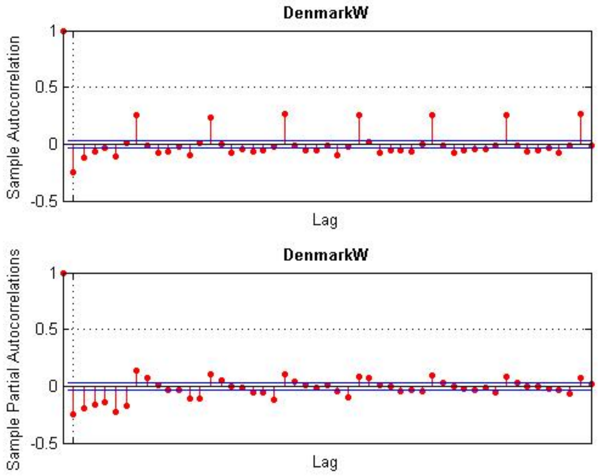

Figure 10: ACF and PACf of logarithm transformed of DenmarkW price series.

We have applied differencing in order to get stationarity. Now we again find the ACF and PACF of the differenced series. The results of ACF and PACF of the differenced series can be seen in Figure 11 and Figure 12. The plots of ACF and PACF of our series are partially similar with theoretical ACF and PACF but after lag 7.

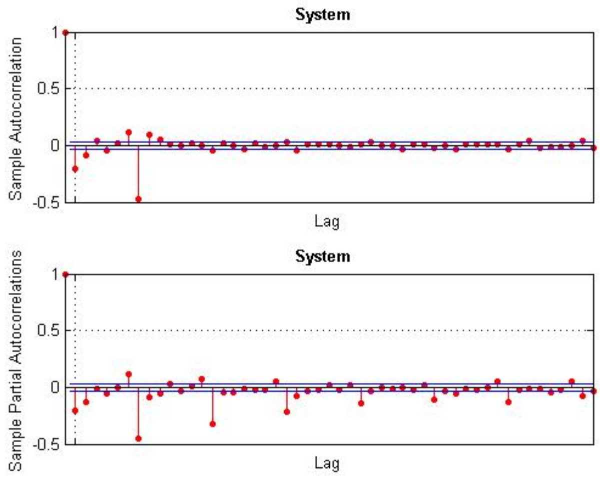

Figure 11: ACf and PACF of differenced System price series.

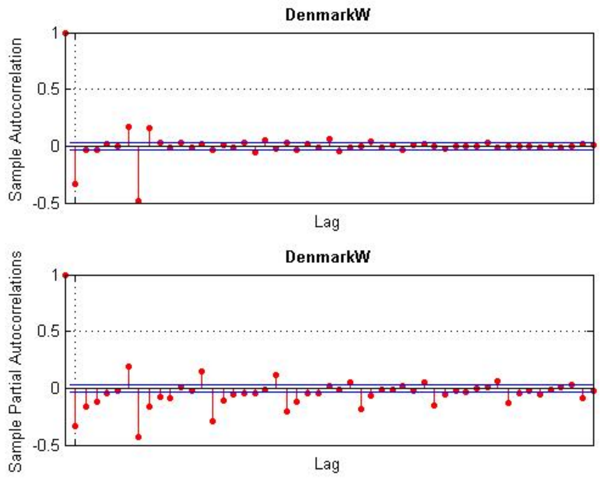

Figure 12: ACf and PACf of differenced DenmarkW price series.

Once the stationarity has been obtained, the next stage is to identify the model.

4.4.1 Order of ARMA-GARCH

Here in the model selection we are going to use the SLEIC function which does not only find order of ARMA but GARCH also. This function by itself also selects the model for GARCH, if there is any hetroskedasticity present in the series. As our results of ACF and PACF of the differenced series are not exactly similar with the theoretical ACF and PACF but after lag 7 both ACF and PACF appear to die out quickly. We can see that ACF is very significant about 0.5 at lag 7. Also PACF is dying out quickly from lag 7. But if ACF would be significant at lag 1 and PACF would be dying out from lag 1 then it would be very easy to select ARMA(0,1). But in this case it would not be appropriate to select ARMA(0,1). Now for the GARCH effect we have already seen that our differenced series for both System and DenmarkW have some correlation present specially at lag 7 the correlation is significant up to 0.5. Figure 13 shows the ACF of the squared differenced series.

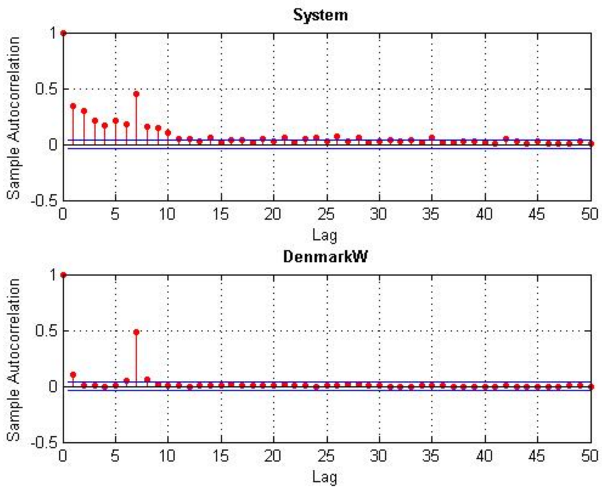

Figure 13: ACf of squared differenced series.

Now we can see in Figure 13 that ACF of squared differenced series indicate significant correlation. Here we have performed two tests on our data. One is known as Engle's test for presence of ARCH/GARCH effects and other is called Ljung-Box-Pierce Q-test [3] lack-of-fit hypothesis test. These tests also play some role selecting an appropriate model.

Ljung-Box test verifies if there is a significant serial correlation in differenced series for System and DenmarkW tested for 1 to 50 lags of the ACF at the 5 percent level of significance. The same test for squared differenced series indicates that both System and DenmarkW contain significant serial correlation. Engel's test for the differenced series of System and DenmarkW rejects hypothesis that both series do not contain ARCH effect at the 5 percent level of significance. Squared differenced series of System and DenmarkW both have ARCH effect. Therefore, the presence of heteroscedasticity for both System and DenmarkW indicates that GARCH modeling is appropriate.

4.5 SLEIC Results

Figure 14 and Figure 15 show the information criteria level (SLEIC) with respect to model complexity. Chosen models for System differenced series and DenmarkW differenced series are ARMA(7,7) GARCH(3,1) and GARCH(1,4), respectively.

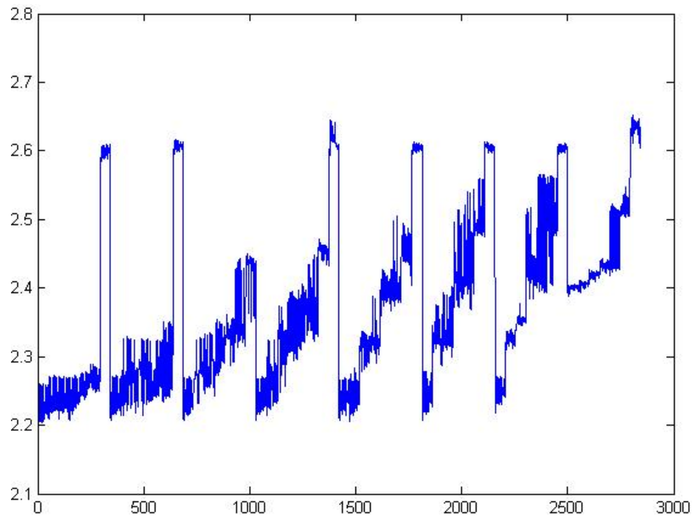

Figure 14: SLEIC results for System differenced series with subject to realizations of different ARMA-GARCH models.

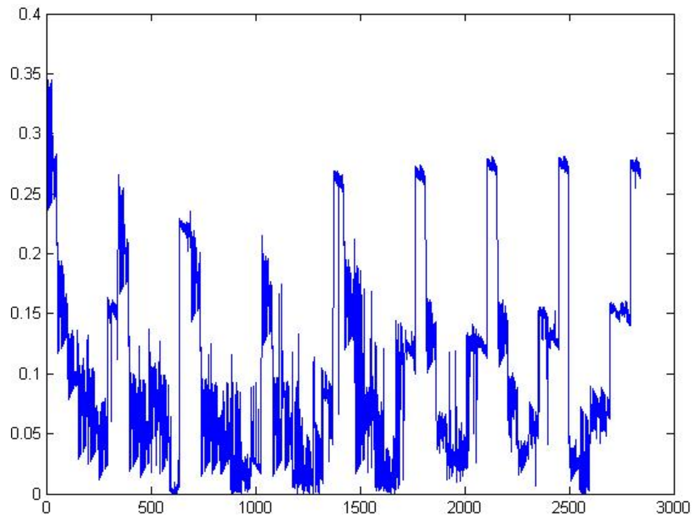

Figure 15: SLEIC results for DenmarkW differenced series with subject to realizations of different ARMA-GARCH models.

4.5.1 System price series

Here, results of ACF and PACF of standardized innovations are presented for model chosen as optimal by the SLEIC function. Figure 16 shows ACF and PACF of standardized innovations obtained after applying ARMA(7,7) and GARCH(3,1) model for System differenced series.

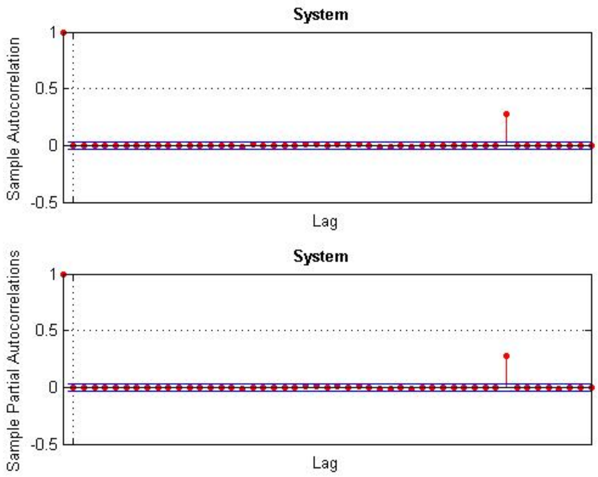

Figure 16: ACF and PACf of standardized innovations using ARMA(7,7)/GARCH(3,1) of system.

4.5.2 DenmarkW price series

Here results of ACF and PACF of standardized innovations are presented for model chosen as optimal by the SLEIC function. Figure 17 shows ACF and PACF of standardized innovation by applying ARMA(0,0) and GARCH(1,4) model in the prices of the DenmarkW.

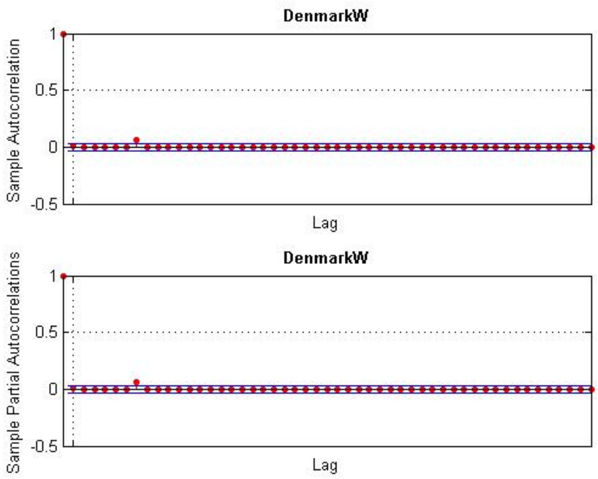

Figure 17: ACF and PACf of standardized innovation using ARMA(0,0)/GARCH(1,4) of DenmarkW.

In the standardized residuals plots for the System we can notice significant ACF estimates at 42nd lag which seems to be associated with weekly and half-season periodicity, *i.e.* weather seasons are considered to last on average for 3 months and thus 6 weeks would mean half a season. Then 6 weeks times 7 week days form the 42nd significant lag.

Also, for DenmarkW we can see a slight significance ACF value at lag 7, which would stand for the weekly price pattern. Therefore, to verify this hypothesis we form weekly averaged data series and perform the whole previous analysis on the weekly series.

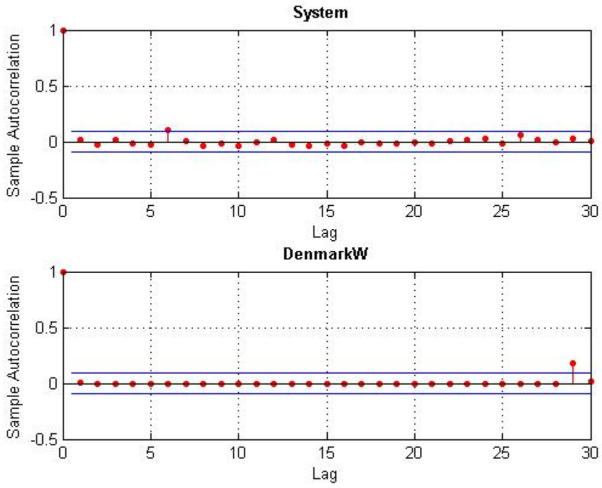

Figure 18: ACF of standardized innovation using ARMA(0,1)/GARCH(1,1) for both System and DenmarkW.

As anticipated, we can see from figure 18 that the ACF for System has very small significance (0.1) at lag 6 which confirms that our System series is showing weekly and half-season periodicity. By averaging the prices over weeks we do remove the weekly pattern, but a sign of 6-weekly seasonality remains. And we can also see from the same Figure 18 that ACF for DenmarkW is not significant in the beginning which means that our DenmarkW has some weekly patterns. Further ACF is significant at the end for DenmarkW which does not violate our assumption.

4.6 Simple ARMA-GARCH Model

Here we have consider theoretical view point in order to select models for both System and DenmarkW prices. By analyzing figure 11 and figure 12 ARMA(0,1)/GARCH(1,1) have been chosen for both System and DenmarkW prices. ACf and PACF for both series can be seen in the figures 19 and 20. Here we can see that standardized innovations are similar with the innovations of the models ARMA(7,7)/GARCH(3,1) and ARMA(0,0)/GARCH(1,4), that have been selected with the help of SLEIC function.

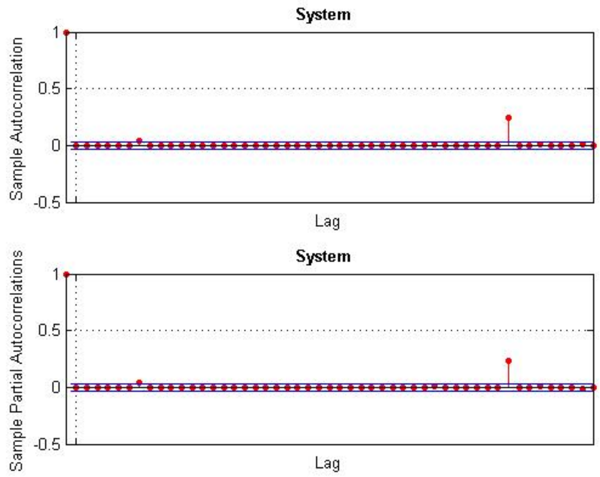

Figure 19: ACF and PACf of standardized innovation using ARMA(0,1)/GARCH(1,1) for System.

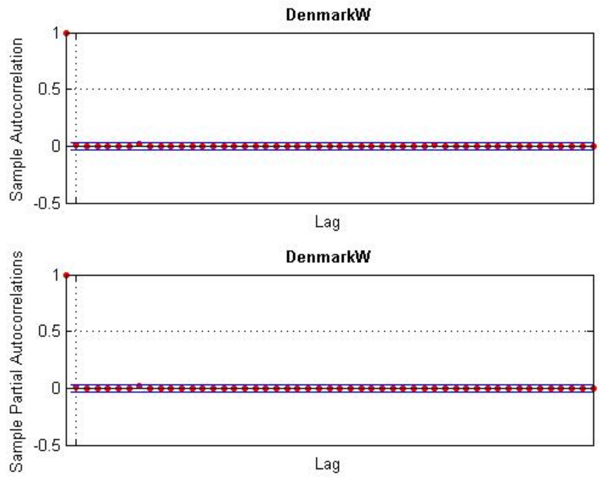

Figure 20: ACF and PACf of standardized innovation using ARMA(0,1)/GARCH(1,1) for DenmarkW.

4.7 Parameter Estimation for System

As we have selected model ARMA(7,7) and GARCH(3,1) for System differenced series therefore we have estimated the parameters of that model using Matlab command "garchdisp". These model parameters can be seen in Table 5.

Table 5: Parameter Estimation for System differenced series.

Parameter	Value	Standard Error
AR(1)	-0.16768	0.01856
AR(2)	-0.17685	0.01985
AR(3)	-0.077395	0.01809
AR(4)	-0.040064	0.01097
AR(5)	0.056526	0.0163
AR(6)	0.05271	0.013447
AR(7)	0.14082	0.014505
MA(1)	0.061311	0.0068507
MA(2)	0.015765	0.0074546
MA(3)	-0.028546	0.0073832
MA(4)	-0.036277	0.0074584
MA(5)	-0.0010396	0.0072419
MA(6)	0.47602	0.0064032
MA(7)	-0.92138	0.0060216
K	0.00042038	2.5828e005
GARCH(1)	0.15876	0.022498
GARCH(2)	0.28927	0.027541
GARCH(3)	0.087204	0.021683
ARCH(1)	0.46476	0.018219

4.8 Parameter Estimation for DenmarkW

We have also estimated the parameters of the model ARMA(0,0) and GARCH(1,4) for DenmarkW differenced series, which can be seen in Table 6. Parameter have been estimated with help Matlab command "garchdisp".

Table 6: Parameter Estimation for DenmarkW differenced series.

Parameter	Value	Standard Error
C	0.0012329	0.0046622
K	0.013464	0.0017974
GARCH(1)	0.68397	0.0422
ARCH(1)	0.15796	0.012376
ARCH(2)	0	0.023113
ARCH(3)	0	0.031152
ARCH(4)	0.037709	0.027048

4.9 Fitting of Original Prices

Simulated prices for System and DenmarkW prices with the help of models ARMA(7,7)/GARCH(3,1) and ARMA(0,0)/GARCH(1,4) can be seen in the figures 21 and 22 and simulated prices for both System and DenmarkW price series with the help of ARMA(0,1)/GARCH(1,1) can be seen in figure 23 and figure 24. Here batter results can also be seen in case of simple ARMA-GARCH model i.e ARMA(0,1)/GARCH(1,1).

Here we can see that our fit is not good and it is not capturing the data. We have tried simulation for a shorter horizon (200 days), the conclusion have been drawn is that the main problem lies in the fact that the simulation tends to generate a few consecutive positive or negative returns which makes the reconstructed price explode or get down to 0. Then it can drop or raise from the zero neighborhood only when having a few consecutive returns of the opposite sign. Whereas we have to remember that in real returns always after a high positive return there will come a high negative one within 1 to 3 days, which brings prices back down.

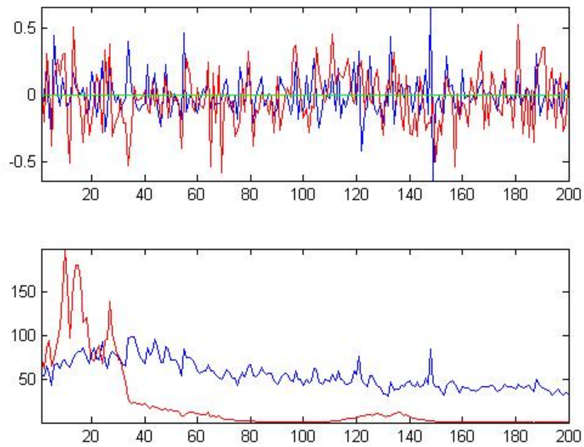

Figure 21: ARMA(7,7)/GARCH(3,1):Original (blue) and simulated (red) prices for System.

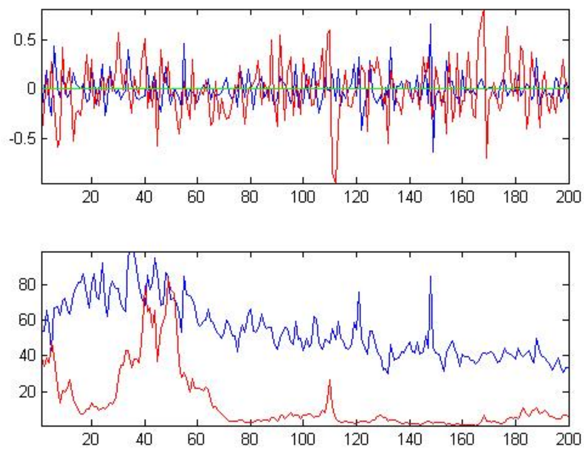

Figure 22: ARMA(0,0)/GARCH(1,4):Original (blue) and simulated (red) prices for DenmarkW.

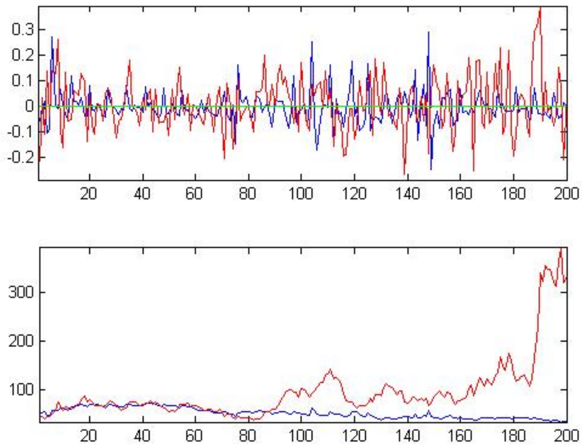

Figure 23: ARMA(0,1)/GARCH(1,1):Original (blue) and simulated (red) prices for System.

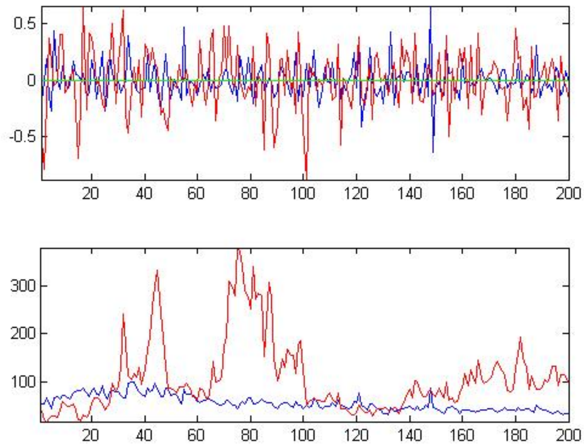

Figure 24: ARMA(0,1)/GARCH(1,1):Original (blue) and simulated (red) prices for DenmarkW.

One could think that since price reconstruction from the simulated returns tends to explode or drop to neighborhood of zero, a better approach would be to use ARMA-GARCH modeling on prices themselves. However, the price series are not stationary, even if it is possible to find theoretically optimal models, the simulations will always by stationarity definition oscillates around some constants level, whereas the real prices have local trends which can seen in the figure 25

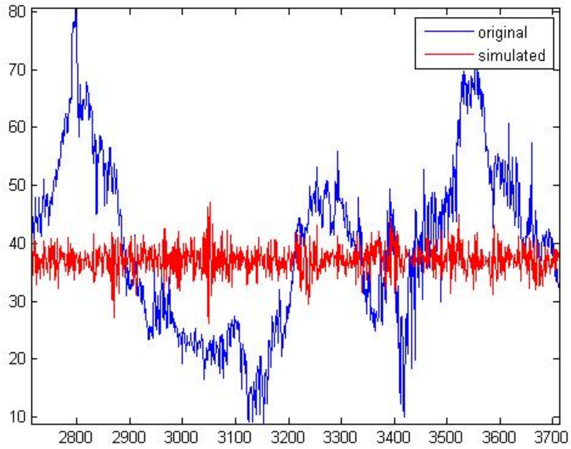

Figure 25: Electricity price behavior (blue) and simulated ARMA-GARCH behavior (red).

5 Modeling Electricity using Ornstein-Uhlenbeck Process

5.1 Probability Models

Uncertainty can be modeled by using probability models [5]. The main object in a probability model is a probability space, which is a triple (Ω, F, P) consisting of a set Ω, usually denoted as the sample space, a σ-field F of subsets of Ω and a probability is defined on F. All possible scenarios that can occur are considered in Ω. If A for example is an event which is subset of Ω then F is collection of all events in Ω. Mathematically it is important to consider only collections of events that have the form of σ field.

5.1.1 Stochastic Process

A stochastic process [5] with state space S is a family (collection) of random variables $X_t, t \in T$ defined on the same probability space (Ω, F, P). The set T is known as its parameter set. The process is called a discrete parameter process if $T \in N$. The process is said to have a continuous parameter if T is not countable. The time is represented as index t, and then one thinks of X_t as the "position" or the "state" of the process at time t. The state space is R in most cases, and then the process is real-valued. There will be also examples where S is a finite set, N, or the set of all integers. The mapping for every fixed $\omega \in \Omega$ on the parameter set T is called a realization, trajectory, sample path or sample function of the process is given by.

$$t \to X_t(\omega)$$

5.1.2 Brownian Motion

The complex and erratic motion of grains of pollen suspended in a liquid had been observed by Robert Brown in 1827. It was later found that such irregular motion comes from the extremely large number of collisions of the suspended pollen grains with the molecules of the liquid. Norbert Wiener presented a mathematical model for this motion based on the theory of stochastic processes in the 20's. The location of a particle at each time $t \geq 0$ is a three dimensional random vector B_t. The Brownian Motion [13] is defined mathematically.

Definition A stochastic process $\{B_t, t \geq 0\}$ is known as Brownian Motion if it satisfies the following conditions:

1. $B_0 = 0$
2. For all $0 \leq t_1 < \ldots < t_n$ the increments $B_{t_n} - B_{t_{n-1}}, \ldots, B_{t_2} - B_{t_1}$ are independent random variables.
3. If $0 \leq s < t$ the increments $B_t - B_s$ has the normal distribution with mean 0 and standard deviation $t - s$.
4. The process $\{B_t\}$ has continuous trajectories or paths.

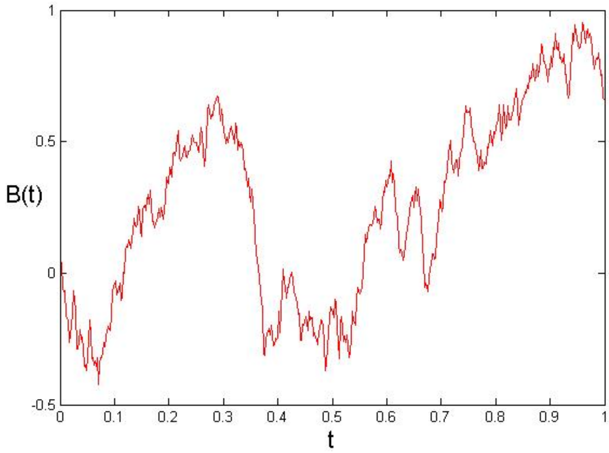

Figure 26: Geometric Brownian Motion.

5.2 Stochastic Differential Equations

Suppose a Brownian Motion $\{B_t, t \geq 0\}$ defined on a probability space (Ω, F, P). Consider that $\{F_t, t \geq 0\}$ is a filtration such that B_t is F_t-adapted and for any $0 \leq s < t$, the increment $B_t - B_s$ is independent of F_s. We aim to solve stochastic differential equations [18] of the form

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t$$

X_0 is an initial condition, which is a random variable independent from the Brownian Motion B_t. The coefficients $b(t, x)$ and $\sigma(t, x)$ are called drift and diffusion coefficient respectively.

5.3 Ornstein-Uhlenbeck process

Traditional financial models start with the Black-Scholes assumption of the Geometric Brownian Motion or log-normal prices. This assumption does not make sense in the context of the electricity prices for many reasons including the non-predictability of the electricity prices. A model which has been used in practice is known as Ornstein-Uhlenbeck process. This continuous time model permit for autocorrelation in the series and is written as

$$dX_t = \kappa[\mu - X_t]dt + \sigma dB_t$$

where κ,$\sigma > 0$ andμ is a real number.

5.4 Modeling of Elecricity using Mean Reversion Process

It is necessary to incorporate mean reversion when modeling electricity prices, because some time we observe that electricity prices jump from 30 €/MWh to 150 €/MWh due to an unexpected event (*e.g.* transmission constraints, plant outages heat wave, *etc.*). Most market participants would agree that it is highly probable that prices will eventually return to their average level once the cause of the jump goes away. For similar reasons if the price of a barrel of WTI falls to US. This simple example illustrates the limitations of Geometric Brownian Motion (GBM) when applied to electricity prices. In the example above GBM would accept the 150 €/MWh price as a normal event and would proceed randomly from there (via a continuous diffusion process) with no higher probability of returning to the average price level and no consideration of prior price levels (no memory). This result is clearly strange in market reality, and provides evidence to select mean reverting price process.

5.5 Geometric Brownian Motion With Mean Reversion

Geometric Brownian Motion is a random walk process which is used to model prices based on the assumption that price changes are independent of one another. Which means, the historical path of the price following to achieve its current price is irrelevant for predicting the future price path. Modification of the random walk is known as the mean reversion, where price changes are not completely independent of one another but rather are related of one another.

5.6 Geometric Brownian Motion With Mean Reversion: Mathematical Representation

Mathematically, we can write the phenomena of mean reversion with a modification to the Geometric Brownian Motion assumption.

$$dX_t = \kappa[\mu - X_t]dt + \sigma dB_t$$

Where

1. μ is long run equilibrium price or mean reversion level
2. X_t is the spot price
3. κ is the mean reversion rate
4. σ is the volatility
5. B_t is standard Brownian Motion

we can notice from this equation that the drift term or mean reversion component is governed by the distance between the spot price and the mean reversion level as well as by the mean reversion rate. If the spot price is above the mean reversion level, the mean reversion component will be negative, resulting in a decrease on the spot price. On the other hand, if the spot price is below the mean reversion level, the mean reversion component will be positive, thus exerting a upward influence on the spot price. Thus in the long run, this results in a price path pulls towards the mean reversion level or equilibrium level, at a speed determined by the mean reversion rate. Moreover mean reversion rate is high when spot price is high and low when spot price is low. The intuition behind all mean reverting processes is captured by this simple case. This model is suggested for the spot price of energy commodity in [20].

5.7 Numerical Methods

There are number of numerical methods which can solve our model. One way is to discretize the stochastic differential equation (SDE) giving X_t and to use simulations to regenerate the distribution of returns at the given time T. The whole information needed for both option pricing and risk analysis, at each instance of time is taken from the distribution of returns. Finite difference method is also used for option pricing, in which a partial differential equation (PDE) is discretized. Binomial tree is the third one which can be considered either as a finite difference discretization of the option-pricing PDE or as a discretization of the underlying SDE. In reality algorithms for both binomial and finite difference are model dependent and will have to be reconstructed for other models, while Monte Carlo routine would require only minor modification. From [11], we know that there are many methods to solve such problems.

5.8 Asset Pricing by Monte Carlo Simulation

There are two ways of measuring accuracy in numerical methods, namely strong convergence and week convergence. For strong convergence, an instance of the stochastic process is required to match the exact solution of the process which is driven by the same random function as closely as possible. Any numerical scheme for the approximation of an Stochastic Differential Equation (SDE) modeling X_t starts with the discretization of that SDE and leads to stochastic difference equations. The maximum step size is the main parameter, which characterizes any numerical scheme. The interval $[0, T]$ can be divided into equal subintervals ($\Delta t = \frac{T}{n}$) or unequal subintervals.

The parameters for our model can be calibrated from historical prices by considering the Euler discretization of SDE that we are using. This method has been described in [15].

$$X_{i+1} - X_i = \kappa \Delta t(\mu - X_i) + \sigma \sqrt{\Delta t} Bi$$

where X_i stands for $X_{i\Delta t}$, and the B_i are normally distributed random variables. In order to simplify our notation we write $\overline{\kappa} = \kappa \Delta t$ and $\overline{\sigma} = \sigma \sqrt{\Delta t}$. We will use also these notations.

$$\overline{X} = \frac{1}{N} \sum_{i=0}^{N-1} X_i$$

$$\overline{X_2} = \sum_{i=0}^{N-1} X_i^2$$

$$\overline{X_{12}} = \frac{1}{N} \sum_{i=0}^{N-1} X_i X_{i+1}$$

$$\delta = \frac{X_N - X_0}{N}$$

and the finally

$$\Delta = \frac{1}{N} \sum_{i=0}^{N-1} (X_i - X_{i+1})^2$$

All quantities are observable. First, summing given equation from $i = 0$ to $i = N \ldots 1$ and dividing by N gives

$$\delta = \overline{\kappa}\mu - \overline{\kappa}\overline{X} + \overline{\sigma} \frac{1}{N} \sum_{i=0}^{N-1} B_i \approx \overline{\kappa}\mu - \overline{\kappa}\overline{X}$$

$$\overline{X_{12}} = \overline{X_2} + \overline{\kappa}\mu\overline{X} - \overline{\kappa}\overline{X_2} + \overline{\sigma} \sum_{i=0}^{N-1} X_i B_i \approx \overline{X_2} + \overline{\kappa}\mu\overline{X} - \overline{\kappa}\overline{X_2}$$

$$\Delta = \overline{\kappa}^2\mu^2 - 2\overline{\kappa}^2\mu\overline{X} + \overline{\kappa}^2\overline{X_2} + \overline{\sigma}^2 \frac{1}{N} \sum_{i=0}^{N-1} B_i{}^2 \approx \overline{\kappa}^2\mu^2 - 2\overline{\kappa}^2\mu\overline{X} + \overline{\kappa}^2\overline{X_2} + \overline{\sigma}^2$$

If we suppose that the approximations made in the above equations are in fact exact, we have three equations for three unknowns which may be written, after solving, as

$$\overline{\kappa} = \frac{\overline{X_{12}} - \overline{X}\delta - \overline{X_2}}{\overline{X}^2 - \overline{X_2}}$$

$$\mu = \frac{\overline{XX_{12}} - \overline{X_2}\delta - \overline{XX_2}}{\overline{X_{12}} - \overline{X}\delta - \overline{X_2}}$$

$$\overline{\sigma} = \sqrt{\Delta - (\overline{\kappa}^2\mu^2 - 2\overline{\kappa}^2\mu\overline{X} + \overline{\kappa}^2\overline{X_2})}$$

5.9 Least Square Method

The stochastic differential equation (SDE) for the Ornstein-Uhlenbeck process constitutes our model which is given as.

$$dX_t = \kappa[\mu - X_t]dt + \sigma dB_t$$

We have used the following equation for generating paths that are sampled with fixed time steps of $dt = 1$. This equation is an exact solution of the stochastic differential equation.

$$X_{i+1} = X_i e^{-\kappa dt} + \mu(1 - e^{-\kappa dt}) + \sigma\sqrt{\frac{1 - e^{-2\kappa dt}}{2\kappa}}N_{1,0}$$

Calibration is done by using the following relationship between consecutive observations X_i, X_{i+1} is linear with i.i.d. normal random term ϵ.

$$X_{i+1} = aX_i + b + \epsilon$$

The relationship between the linear fit and the model parameter is given by

$$a = e^{-\kappa dt}$$

$$b = \mu(1 - e^{-\kappa dt})$$

$$sd(\epsilon) = \sigma\sqrt{\frac{1 - e^{-2\kappa dt}}{2\kappa}}$$

Rewriting above three equation gives

$$\kappa = -\frac{lna}{dt}$$

$$\mu = \frac{b}{1 - a}$$

$$\sigma = sd(\epsilon)\sqrt{\frac{-2lna}{dt(1 - a^2)}}$$

5.10 Maximum Likelihood Estimation

Let us consider data vector (X) which is a random sample from an unknown population. The aim of data analysis is to recognize the population that is most likely to have generated the sample. In statistics, each population is recognized by a corresponding probability distribution. A unique value of the model parameter is associated with each probability distribution. Different probability distributions are generated as the parameter changes in value. Formally, a model is written as the family of probability distributions indexed by the model parameters. For a given set of parameter values, the corresponding probability distribution function will show that some data are more probable than other data. However, we have already observed the data. Therefore, we are faced with an inverse problem: specified the observed data and a model of interest, find the one PDF, among all the PDF's that the model prescribes, that is most likely to have generated the data. Likelihood function is defined by interchanging the roles of the data vector X and the parameter vector w in $f(X|w)$, to solve this inverse problem, *i.e.*

$$L(w|X) = f(X|w).$$

Hence $L(w|X)$ represents the likelihood of the parameter w given the observed data X. R.A. Fisher [1] in the 1920s basically developed the principle of maximum likelihood estimation (MLE), which states that the desired probability distribution is the one that makes the noticed data *most likely*, which means that one must seek the value of the parameter vector that maximizes the likelihood function $L(w|X)$.

5.10.1 Calibration using Maximum Likelihood Estimates

The equation of the conditional probability density of an observation X_{i+1}, given a previous observation X_i, (with a dt time step between them) is given by

$$f(X_{i+1}|X_i; \mu, \kappa, \hat{\sigma}) = \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} \exp\left[-\frac{\left(X_i - X_{i-1}e^{-\kappa dt} - \mu(1 - e^{-\kappa dt})\right)^2}{2\hat{\sigma}^2}\right].$$

In order to simplify the notation of this equation we have introduced

$$\hat{\sigma}^2 = \sigma^2 \frac{1 - e^{-2\kappa dt}}{2\kappa}.$$

5.10.2 Log-likelihood function

The log-likelihood function of a set of observation $(X_0, X_1, \ldots, X_n)$, can be derived from the conditional density function

$$L(\mu, \kappa, \hat{\sigma}) = \sum_{i=1}^{n} lnf(X_i|X_{i-1}; \mu, \kappa, \hat{\sigma})$$

$$= -\frac{n}{2}ln(2\pi) - nln(\hat{\sigma}) - \frac{1}{2\hat{\sigma}}\sum_{i=1}^{n}\left[X_i - X_{i-1}e^{-\kappa dt} - \mu(1 - e^{-\kappa dt})\right]^2$$

5.10.3 Maximum Likelihood Conditions

The maximum of this log-likelihood surface can be seen at the place where all the partial derivatives are zero. This leads to the following set of equation.

$$\frac{\partial L(\mu, \kappa, \hat{\sigma})}{\partial \mu} = 0 = \frac{1}{\hat{\sigma}}\sum_{i=1}^{n}\left[X_i - X_{i-1}e^{-\kappa dt} - \mu(1 - e^{-\kappa dt})\right]$$

$$\mu = \frac{\sum_{i=1}^{n}\left[X_i - X_{i-1}e^{-\kappa dt}\right]}{n(1 - e^{\kappa dt})}$$

$$\frac{\partial L(\mu, \kappa\hat{\sigma})}{\partial \kappa} = 0 = -\frac{dte^{-\kappa dt}}{\hat{\sigma}^2}\sum_{i=1}^{n}\left[(X_i - \mu)(X_{i-1} - \mu) - e^{\kappa dt}(X_{i-1} - \mu)^2\right]$$

$$\kappa = \frac{1}{dt}ln\frac{\sum_{i=1}^{n}(X_i - \mu)(X_{i-1} - \mu)}{\sum_{i=1}^{n}(X_{i-1} - \mu)^2}$$

$$\frac{\partial L(\mu, \kappa\hat{\sigma})}{\partial \hat{\sigma}} = 0 = \frac{n}{\hat{\sigma}} - \frac{1}{\hat{\sigma}^3}\sum_{i=1}^{n}\left[X_i - \mu - e^{-\kappa dt}(X_i - \mu)\right]^2$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left[X_i - \mu - e^{-\kappa dt}(X_i - \mu)\right]^2$$

Further simplification and Matlab code for unknown parameters can be found from [26].

5.11 Model Calibration Using Monte Carlo Simulation

Model has been calibrated using all above methods and found the values of unknown parameters. The following Table 7 shows all the parameters values using Euler method. We have three unknown parameters values κ, μ and σ.

Table 7: Parameter estimates for System and DenmarkW differenced prices with and without spikes by using Ornstein-Uhlenbeck process.

case	Sys-Spiky	Sys-NoSpikes	Den-Spiky	Den-NoSpikes
κ	1.2076	1.2487	1.3793	1.3582
μ	-2.2415e-005	-2.2914e-005	-1.1461e-005	-1.0693e-005
σ	0.1314	0.1264	0.4598	0.4055

Here we can see that the parameters are not much different for both spiky and non spiky behavior of System and DenmarkW differenced series. It means that removing spikes does not change much the parameter values, as we have got the similar values for both spiky and non spiky behavior. Now we are moving on to plots of the estimated differenced series and residual for each case. The following Figure 27 shows that the estimated differenced series got from the Euler method captures well the System differenced series for spiky behavior.

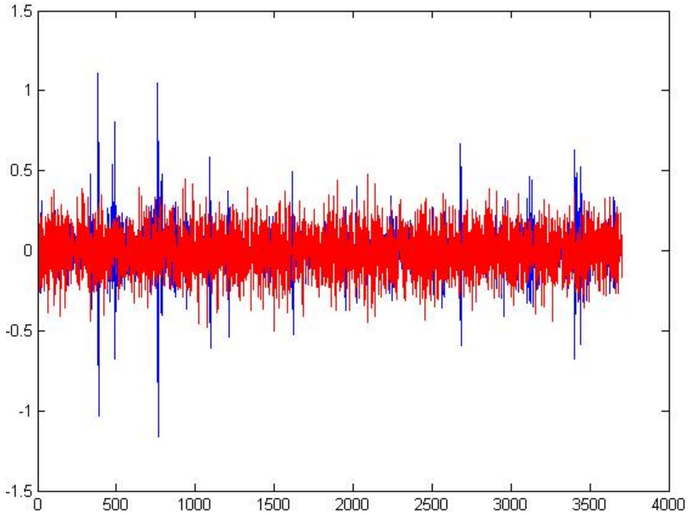

Figure 27: Estimated (red) and original (blue) diffrenced series for System with spikes.

and we have got the plot for the residual of the the System differenced series which can be seen in Figure 28.

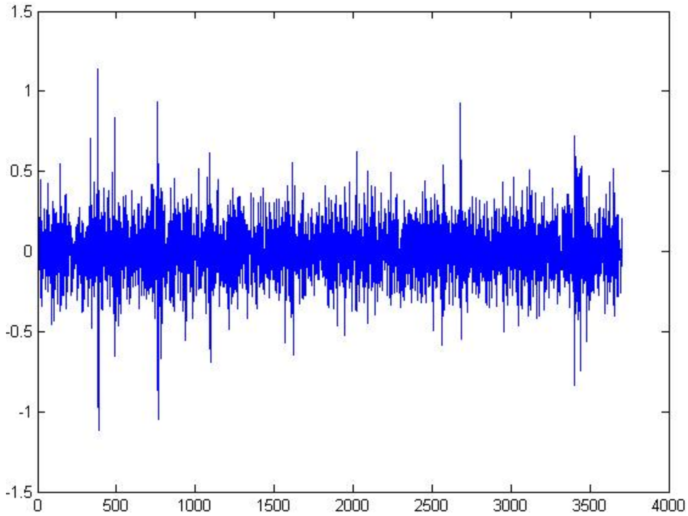

Figure 28: Plot of the residual series for System with spikes.

Now we move on to find the plot of the estimated and original differenced series for System but without spikes. We can see from Figure 29 that the behavior is quite similar with system with spikes. It mean means that when we are dealing with the differenced series it does not matter whether we have removed spikes or not.

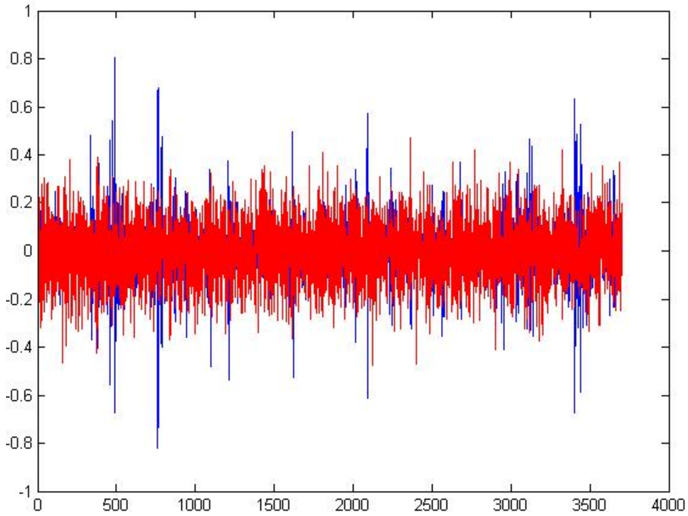

Figure 29: Estimated (red) and original (blue) differenced series for System without spikes.

We can also see the plot of the residual series in Figure 30.

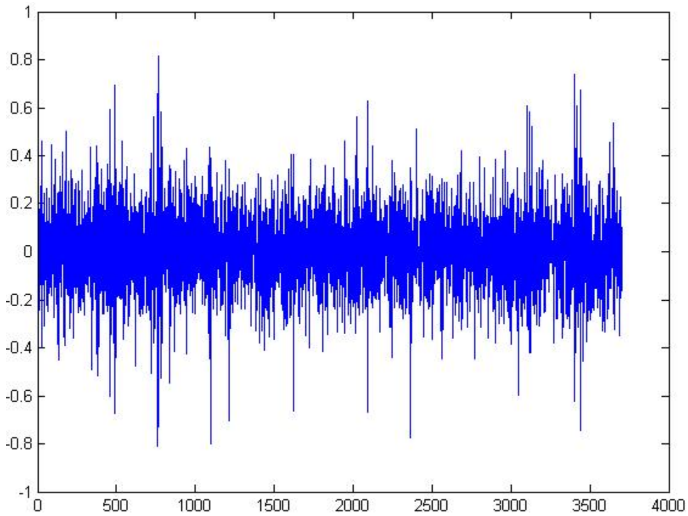

Figure 30: plot of residual for System differenced series without spikes.

Furthermore, we have also got estimated and original series for DenmarkW differenced series with spikes. The following figure 31 shows the the plot of the both estimated and original series.

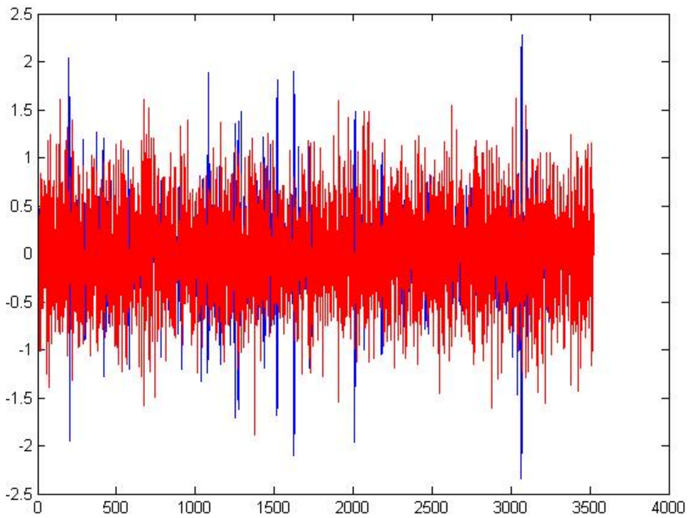

Figure 31: Estimated (red) and original (blue) diffrenced series for DenmarkW with spikes.

We can also see the plot of the residual for DenmarkW differenced series with spikes (see Figure 32).

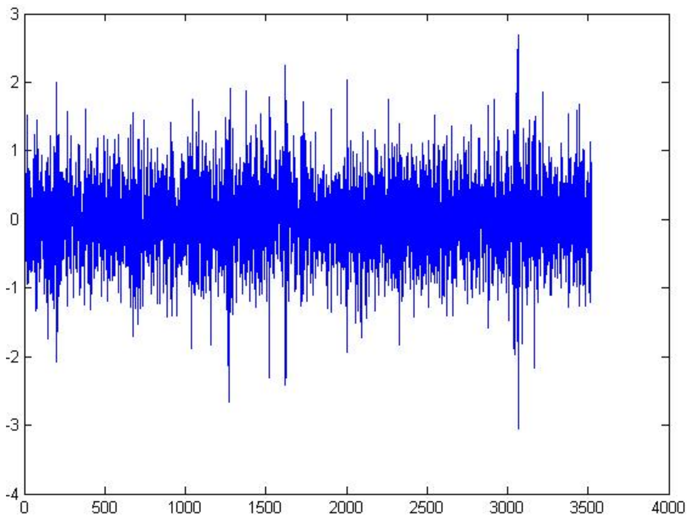

Figure 32: Plot of the residual series for DenmarkW differenced with spikes.

Now we plot the estimated and original differenced series without spikes. Here we can see from the following Figure 33 that the estimated differenced series is capturing well the original differenced series. As it can be seen from Table 7, for both spiky and non spiky behavior the parameters are not much different.

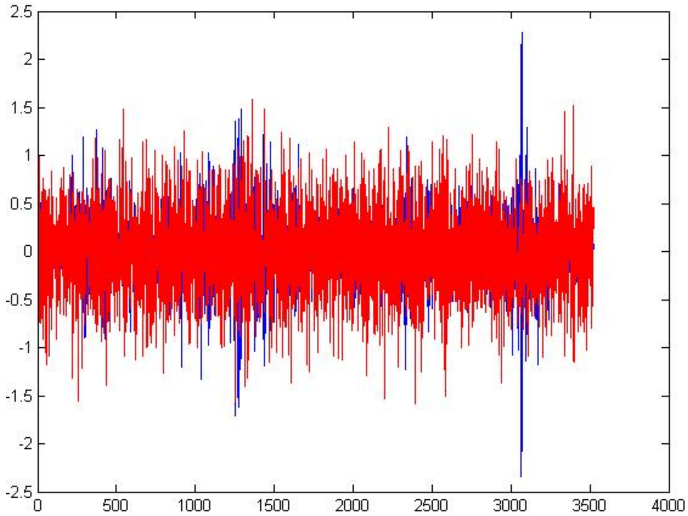

Figure 33: Estimated (red) and original (blue) differenced series for DenmarkW without spikes.

Now, residuals are plotted for DenmarkW differenced series and we have also removed spikes in this case. Figure 34 shows the residual for DenmarkW differenced series.

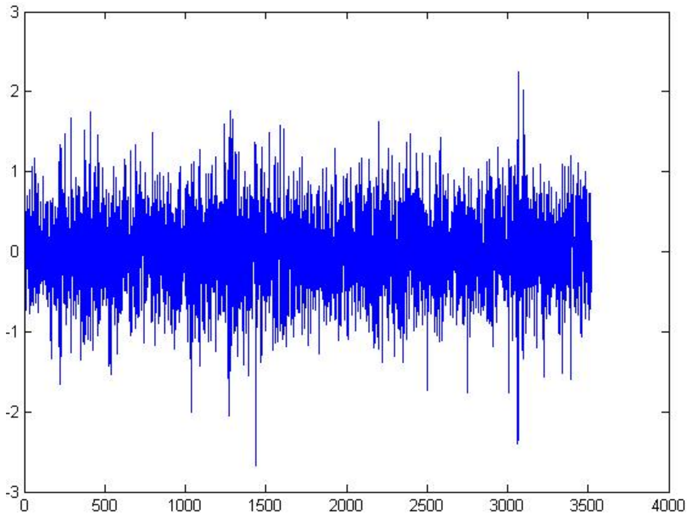

Figure 34: Plot of the residual series for DenmarkW differenced seires without spikes.

5.12 Parameter Estimation using Least Square Method

Parameters of our model have been obtained using least square method, which can be seen in Table 8. Here we can see that the value of κ and σ are complex numbers.

Table 8: Model Parameter Estimation for System and DenmarkW differenced prices with and without spikes by using Least Square Method.

case	Sys-Spiky	Sys-NoSpikes	Den-Spiky	Den-NoSpikes
κ	1.5723-3.1416i	1.2236-3.1416i	0.9696-3.1416i	1.0081-3.1416i
μ	-3.1813e-005	-3.2828e-005	-1.9307e-005	-1.8787e-005
σ	0.2455-0.1516i	0.2224-0.1521i	0.6634-0.4895i	0.6015-0.4387i

5.13 Parameter Estimation of Model by using Maximum Likelihood Function

The method of maximum likelihood has also been used to estimate the parameters of the model and again we have obtained complex numbers for κ and σ. Model parameters values can be seen in Table 9.

Table 9: Model Parameter Estimation for System and DenmarkW differenced prices with and without spikes by using Maximum Likelihood Function.

case	Sys-Spiky	Sys-NoSpikes	Den-Spiky	Den-NoSpikes
κ	1.5723-3.1416i	1.5549-3.1416i	0.9696-3.1416i	0.9691-3.1416i
μ	-3.1813e-005	-3.1858e-005	-1.9307e-005	-1.9313e-005
σ	0.2454-0.1516i	0.2219-0.1378i	0.6632-0.4894i	0.5971-0.4407i

5.14 Fitting of Original Data

For fitting the original data we have estimated the parameters which can be seen in Table 10. Two parameters κ and μ are calculated by the method that has been described in the Monte Carlo simulation but the third parameter has been estimated by maximum likelihood method. We can see parameters for both System and DenmarkW price series with spikes and without spikes.

Table 10: Model Parameters Estimation for System and DenmarkW prices with and without spikes by using Ornstein-Uhlenbeck process.

case	Sys-Spiky	Sys-NoSpikes	Den-Spiky	Den-NoSpikes
κ	0.0207	0.0166	0.1590	0.1304
μ	29.7427	29.6724	30.9140	32.0604
σ	2.8965	2.5453	9.8924	7.9671

First of all we have got estimation for the system prices with spikes in Figure 35.

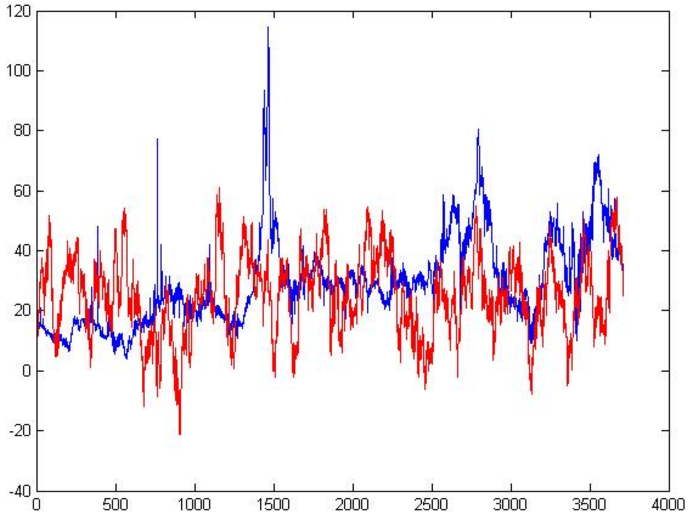

Figure 35: Estimated(red) and original(blue) price series for System with spikes.

We have got the residual series for the System prices in Figure 36

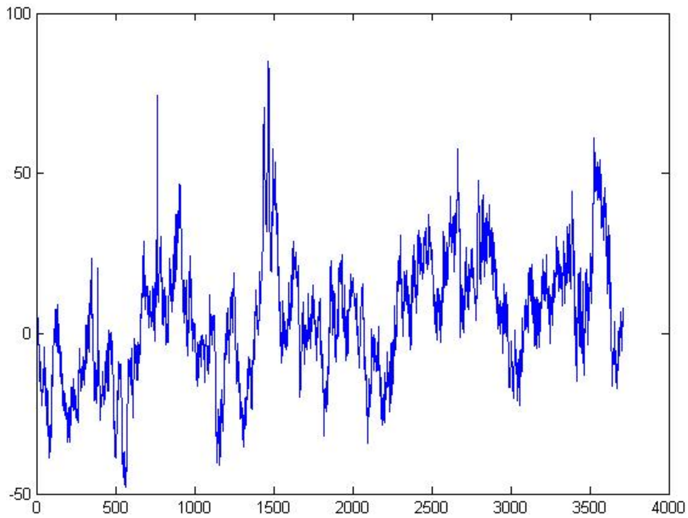

Figure 36: Residual series for System with spikes.

Now we have got the estimation of the System prices without spikes, which can be seen in Figure 37.

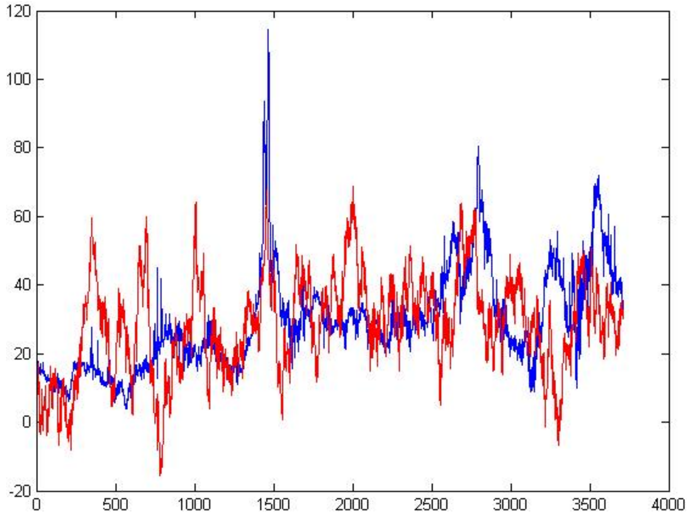

Figure 37: Estimated(red) and original(blue) price series for System without spikes.

and in the same way we have got the residual for the system prices without spikes, which can be seen in Figure 38.

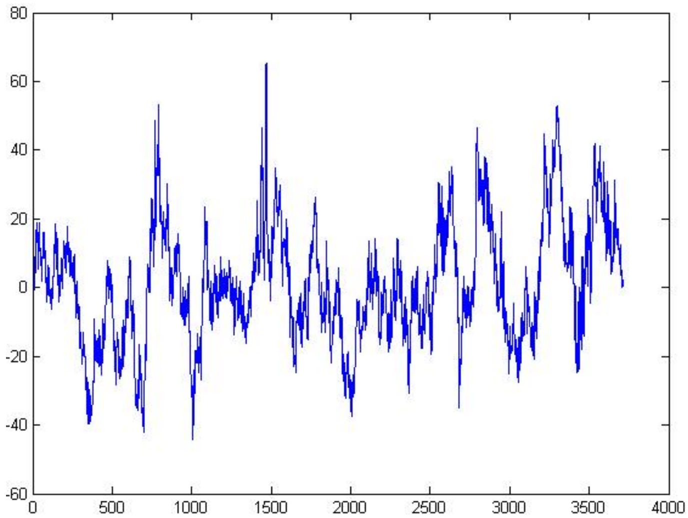

Figure 38: Residual series for System without spikes.

We have estimated parameters for DenmarkW price series with and without spikes, which can also be seen in Table 10. By using these parameters we have estimated the new prices which can be in Figure 39. Furthermore, two parameters mean reversion rate and equilibrium level have been estimated by Monte Carlo Simulation method and the standard deviation has been obtained by the maximum likelihood method.

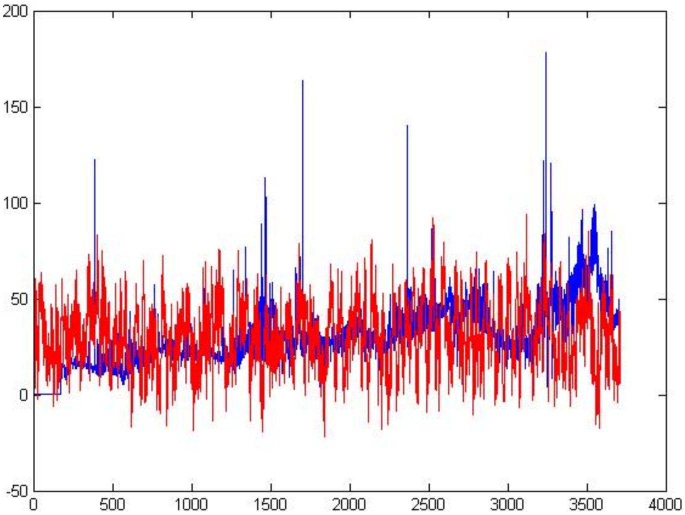

Figure 39: Estimated (red) and Original (blue) series for DenmarkW with spikes.

and we have got the residual for DenmarkW prices with spikes which can be seen in Figure 40.

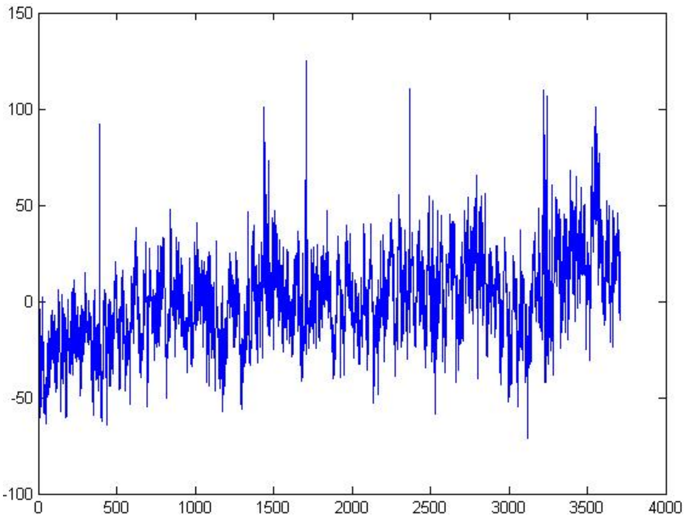

Figure 40: Residual series for DenmarkW with spikes.

We have got the estimated price series for DenmarkW without spikes also, which can be seen the in Figure 41.

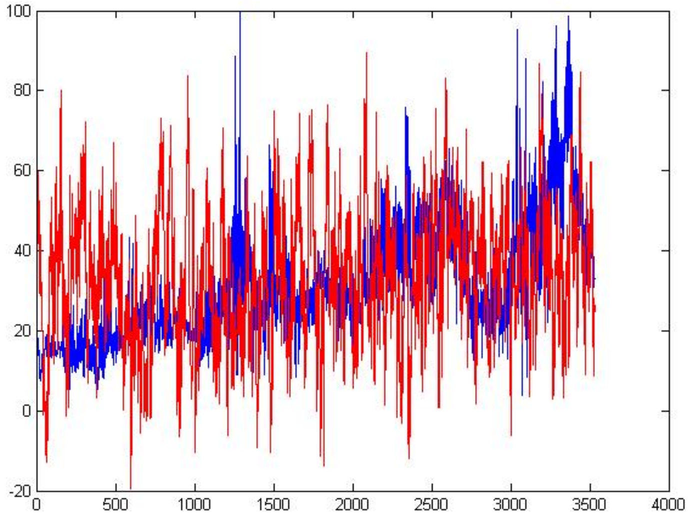

Figure 41: Estimated (red) and Original (blue) series for DenmarkW without spikes.

Residual series have also been obtained which can be seen in Figure 42. This residual series have been obtained by taking the difference between the original and estimated prices for DenmarkW.

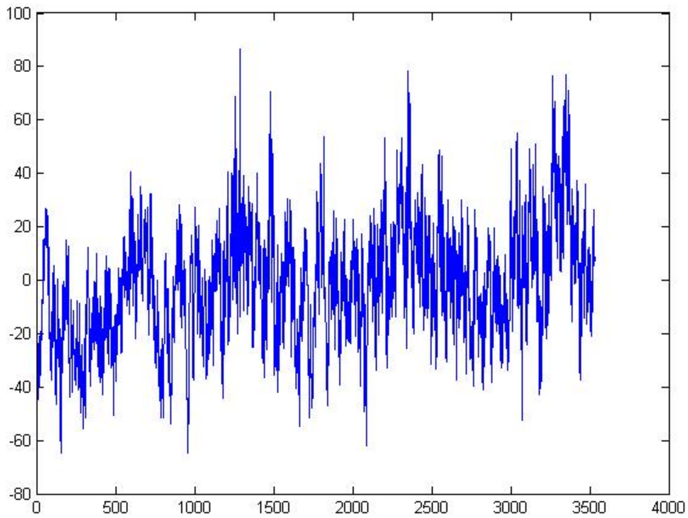

Figure 42: Residual series for DenmarkW without spikes.

6 Final Results

6.1 Residual from ARMA-GARCH and Mean Reversion

In this section we discuss behavior of residuals for both System and DenmarkW differenced series obtained after fitting ARMA-GARCH and mean reverting models. In case of ARMA-GARCH, we have got residual of both spiky and non-spiky series by using Matlab built-in function garchfit. Further, in the case of mean reversion, we have used

Euler approach to approximate both differenced series and got residual for both spiky and non spiky series also. Results are presented at the following Tables 11 and 12 . The normalized histograms of residuals for both cases are presented in Figures 43-44, 45-46,47-48 and 49-50.

Table 11: Statistics for System and DenmarkW differenced prices with and without spikes by using ARMA-GARCH.

case	Sys-Spiky	Sys-NoSpikes	Den-Spiky	Den-NoSpikes
skewness	1.2310	0.4587	0.0195	-0.1527
kurtosis	30.3489	16.6132	36.2776	11.2213
Lillifors test H_0	rejected	rejected	rejected	rejected

Table 12: Statistics for System and DenmarkW differenced prices with and without spikes by using Ornstein-Uhlenbeck process.

case	Sys-Spiky	Sys-NoSpikes	Den-Spiky	Den-NoSpikes
skewness	-0.0279	-0.0083	0.00017	0.0317
kurtosis	6.3515	4.1178	3.9541	3.4048
Lillifors test H_0	rejected	rejected	rejected	accepted

Here we can clearly see that the results obtained by Ornstein-Uhlenbeck are much better than the ARMA-GARCH for both spiky and non spiky series. Now the following Figures 44-50 show the results of normalized histograms

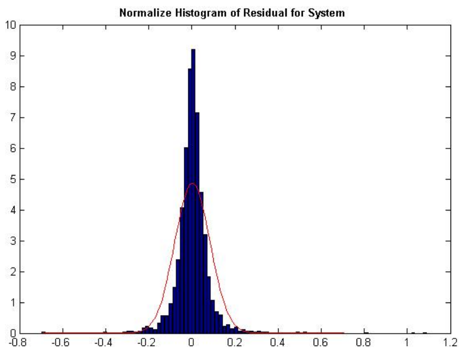

Figure 43: Histogram of Residual for Spiky System Differenced series using ARMA-GARCH.

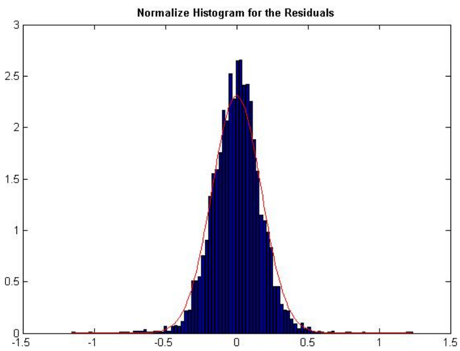

Figure 44: Histogram of Residual for Spiky System Differenced series using Ornstein-Uhlenbeck Process.

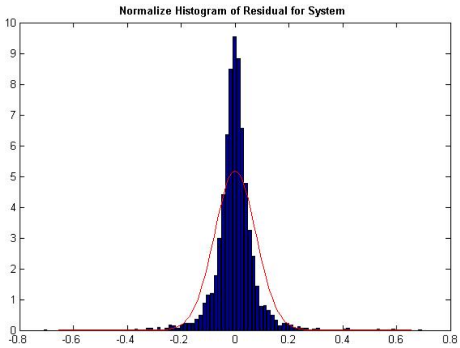

Figure 45: Histogram of Residual for non Spiky System Differenced series using ARMA-GARCH.

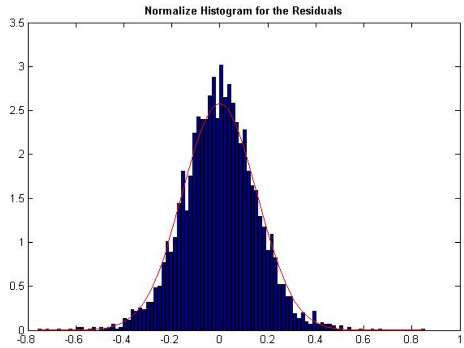

Figure 46: Histogram of Residual for non Spiky System Differenced series using Ornstein-Uhlenbeck Process.

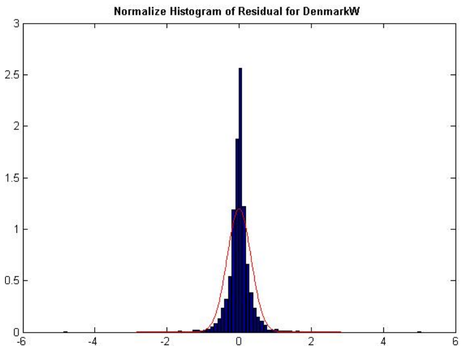

Figure 47: Histogram of Residual for Spiky DenmarkW Differenced series using ARMA-GARCH.

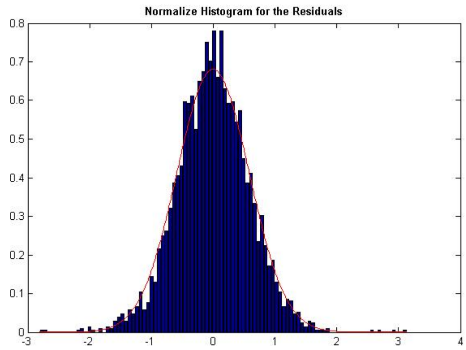

Figure 48: Histogram of Residual for Spiky DenmarkW Differenced series using Ornstein-Uhlenbeck Process.

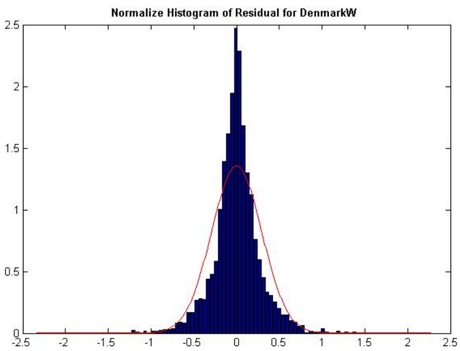

Figure 49: Histogram of Residual for non Spiky DenmarkW Differenced series using ARMA-GARCH.

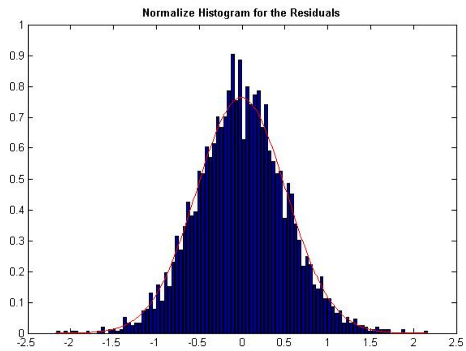

Figure 50: Histogram of Residual for non Spiky DenmarkW Differenced series using Ornstein-Uhlenbeck Process.

As far as the differenced series is concerned, we have seen the results of residuals approximated by ARMA-GARCH and Ornstein-Uhlenbeck process. We have also seen that residuals obtained by Ornstein-Uhlenbeck process are close to normal distribution.

Moreover as we move on to approximation of the original prices by using both methods ARMA-GARCH and mean reversion Ornstein-Uhlenbeck process we have seen that the results obtained by the ARMA-GARCH are not good as compared with the results obtained by mean reverting Ornstein-Uhlenbeck process. For ARMA-GARCH we have used Matlab command "ret2price" for getting the simulated prices.

6.2 Original and Differenced Series Comparison

On the other hand for mean reversion we have used the original prices for both System and DenmarkW. We got the simulated prices for both System and DenmarkW with and without spikes. Further, we have calculated the residual for each case and plotted the normalized histograms of each case which can be seen in Figures 51, 52, 53 and 54.

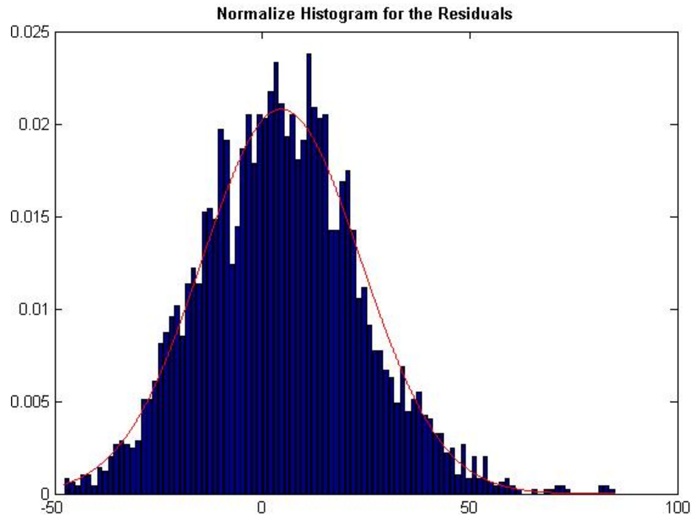

Figure 51: Normalized Histogram of the Residual for System with spikes.

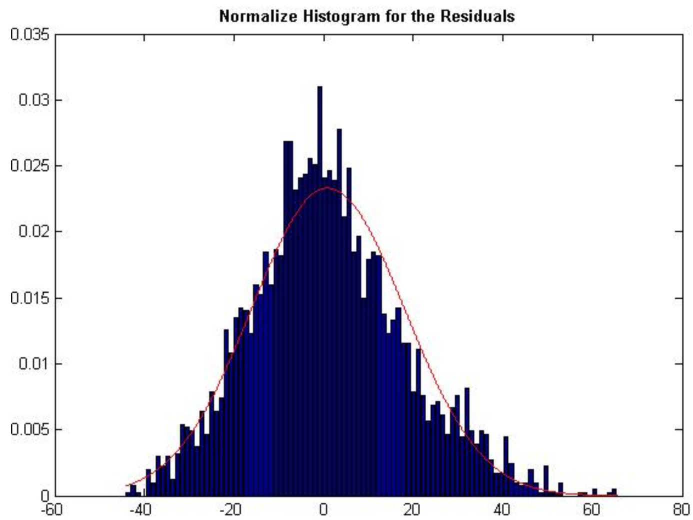

Figure 52: Normalized Histogram of the Residual for System without spikes.

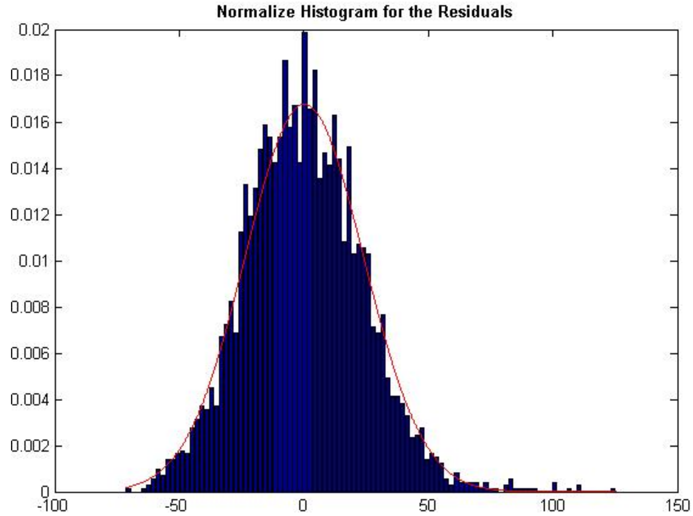

Figure 53: Normalized Histogram of the Residual for DenmarkW with spikes.

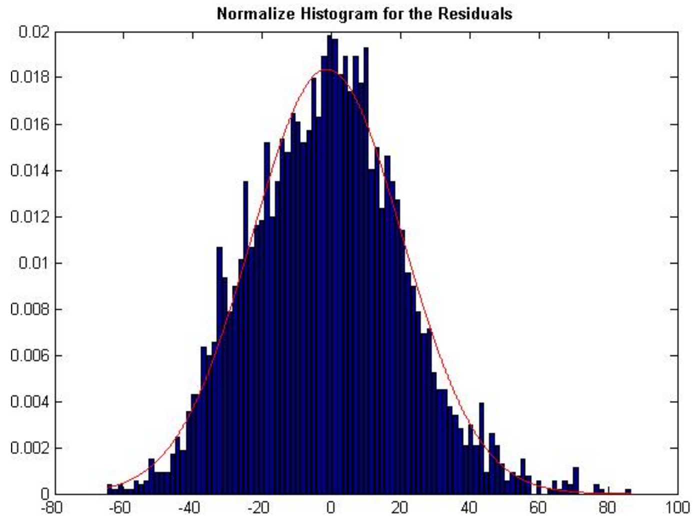

Figure 54: Normalized Histogram of the Residual for DenmarkW without spikes.

We also present statistics for the residuals for both System and DenmarkW with and without spikes, which can be found from Table 13. Here we can see that the values of skewness kurtosis and lillifors test are quite similar with the values that we have obtained for the differenced series in Table 12.

Table 13: Statistics of Residual for System and DenmarkW with and without spikes by using Ornstein-Uhlenbeck process.

case	Sys-Spiky	Sys-NoSpikes	Den-Spiky	Den-NoSpikes
skewness	0.2631	0.3477	0.3706	0.2117
kurtosis	3.2760	3.1427	3.8155	3.1399
Lillifors test H_0	rejected	rejected	rejected	accepted
standard deviation	20.3327	18.4140	24.5650	20.9555

6.3 Histogram of Simulated Prices

As we have seen that our residuals are close to normal distribution and even for DenmarkW prices without spikes the residuals are normally distributed. We have observed that even though our residuals are normally distributed, the simulated prices are not capturing the behavior of the original prices very well. We have got our simulated prices which are fluctuating around the original prices but is not following the original price path. Since we have seen that System prices are less spiky than the DenmarkW prices. But we have got normally distributed residual for DenmarkW prices without spikes. On the other hand we have seen that we have not got normally distributed residual for System prices without spikes. These kind of results compel us to investigate more about the simulated prices. In this section we are going to obtain the histograms of the simulated prices in order to see whether our simulated prices are normally distributed or not. Histogram of the simulated prices for System and DenmarkW can be seen in Figures 55, 56, 57 and 58.

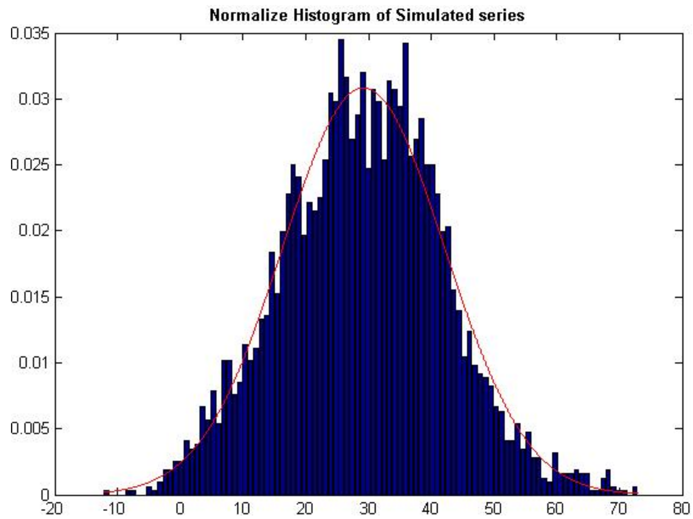

Figure 55: Histogram of Simulated Price Series for System with spikes.

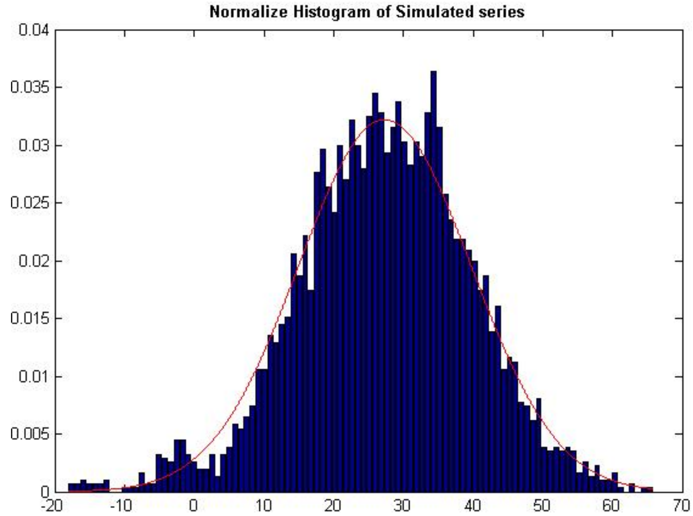

Figure 56: Histogram of Simulated Price Series for System without spikes.

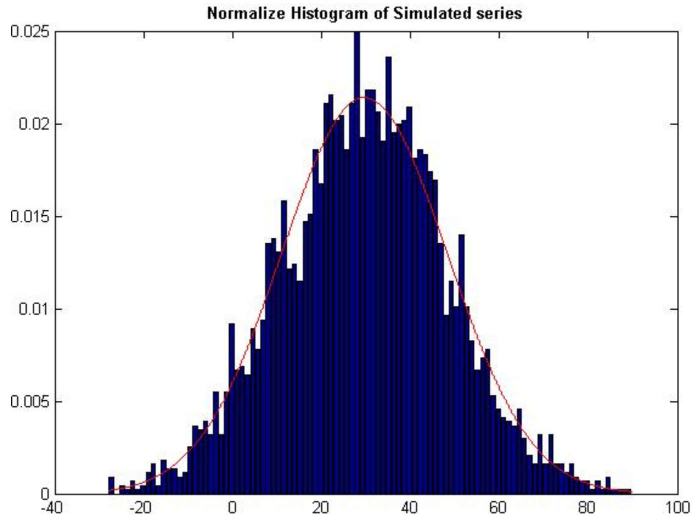

Figure 57: Histogram of Simulated Price Series for DenmarkW with Spikes.

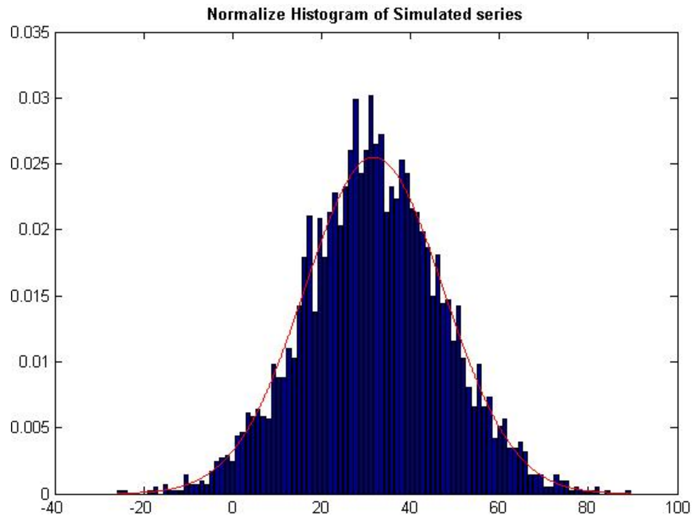

Figure 58: Histogram of Simulated Price Series for DenmarkW without Spikes.

We have present statistics for simulated prices in order to see our simulated prices are normally distributed or not, which can be seen in Table 14. We can see from the table that our simulated prices are normally distributed.

Table 14: Statistics of Simulted prices for System and DenmarkW with and without spikes by using Ornstein-Uhlenbeck process.

case	Sys-Spiky	Sys-NoSpikes	Den-Spiky	Den-NoSpikes
skewness	0.0799	-0.2264	-0.0010	0.0168
kurtosis	3.0063	3.3269	2.9812	3.1223
Lillifors test H_0	rejected	accepted	accepted	accepted
standard deviation	12.9384	12.3934	18.6417	15.6391

Moreover we have observed that for the system prices simulated prices were trying to capture the prices partially but for the DenmarkW, simulated prices were fluctuating more. From this we have concluded that for less spiky behavior Ornstein-Uhlenbeck process may capture the original prices but for more spiky behavior Ornstein-Uhlenbeck process becomes blind and start fluctuating around the mean. Residual are normally distributed not in the sense that simulated prices capturing well the original prices but in the sense that simulated prices becomes normally distributed.

7 Conclusion

In this work, we have compared two families of mathematical models for their respective capability to capture the statistical properties of real electricity spot market time series that are characterized by frequent and dramatic price spikes, and highly non-normally distributed price and even price return series. Parameters of both model families have been calibrated by using the real time series with suitable statistical criteria.

The first model family was ARMA-GARCH models. An optimal order for an ARMA model was identified with the SLEIC criterion. Subsequently the optimal coefficients for the chosen model were determined by least squares fitting. A GARCH model for volatility was then added by using the Matlab GARCH Toolbox. It was found that even an optimal ARMA-GARCH model leaves a leptokurtic residual, and hence not at all normally distributed. This implies that ARMA-GARCH models fail to capture the statistical properties of real electricity spot market time series.

The second model family was mean-reverting Ornstein-Uhlenbeck model. Optimal reversion rate and equilibrium level were approximated by Euler discretization and volatility was determined by Maximum Likelihood funtion. The residuals emerging from this optimal mean reverting model are normally distributed. But at closer inspection it becomes evident that this does not follow from a good statistical fit to the real series. Rather, this is a consequence of the excessive "jumpiness" of an optimal mean-reverting model. Since an Ornstein-Uhlenbeck model is always normally distributed by definition, this property is transferred to model residuals when there are frequent jumps in the simulated series that do not coincide with jumps in the real series.

We therefore have to conclude that neither ARMA-GARCH models, nor conventional mean-reverting Ornstein-Uhlenbeck models, even when calibrated optimally with a real electricity spot market price or return series, capture the statistical characteristics of the real series.

Future Work

We have observed that electricity price behavior is very complex for modeling purpose. We can easily see that prices do not have constant equilibrium level and volatility because volatility and equilibrium level varies from season to season. We can improve our model by introducing same model for (Ornstein-Uhlenbeck process) not only for prices but for equilibrium level and volatility also. Furthermore, we can categories electricity prices into two parts i.e summer prices and winter prices. We can also include spike or jump part for winter prices along with Ornstein-Uhlenbeck process.

References

[1] Aldrich, J. (1997). "R.A. Fisher and the making of maximum likelihood 1912–1922". *Statistical Science* 12 (3): pp. 162–176

[2] Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity, *Econometrica*, Econometric Society, USA, vol. 31, pp. 307–327.

[3] Box, G.E.P., Jenkins, G.M., Reinsel, G.C. (1994). *Time Series Analysis: Forecasting and Control*, third edition, Prentice Hall.

[4] Brockwell, P. J. and Davis, R. A. (2002). *Introduction to Time Series and Forecaseting.* 2nd ed. Springer Texts in Statistitics, Springer, New York.

[5] Cinlar, E. (1975). *Introduction to Stochastic Processes.* Prentice-Hall.

[6] Deng, S. (2000). Stochastic models of energy commodity prices and their applications: Mean-reversion with jumps and spikes. Working paper. Georgia Institute of Technology.

[7] Dickey, D. A and Fuller, W. A. (1979). Distribution of the estimates for autoregressive time series with a unit root. *Journal of Empirical Finance* **1**: 83-106

[8] Engle, R.F. (1982). "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, Econometric Society, USA, vol. 50, pp. 987-1007.

[9] Eydeland, A. and Hélyette, G. (1998). Pricing power derivatives. *Risk* (October).

[10] Hanke, J. E, Reitsch, A. G, and Wichern, D, W. (1994). *Business Forecasting*, seventh edition. Prentice Hall, Upper Saddle River, NJ.

[11] Jeffery, J. (2004). *Leader Numerical Analysis and Scientic Computation*, Addison Wesley

[12] Kaminski, V. (1997). The Challenge of pricing and risk managing electricity derivatives, Ch. 10 in: The US power market, 149-171, (*Risk Publications*, London).

[13] Karatzas, I. and Schreve, S.E. (1991). *Brownian Motion and Stochastic Calculus.* Springer-Verlag.

[14] Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. and Shin, Y. (1992). Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root. *Journal of Econometrics* 54, 159–178.

[15] Lari-Lavassani, A., Sadeghi, A.A. and Weron, A. (2001). Mean reverting models for energy option pricing. *International Energy Credit Association.*

[16] Lucia, J.J. and Schwartz, E.S., (2002). Electricity prices and power derivatives: Evidence from the Nordic Power Exchange, *Rev. Derivatives Research* 5.

[17] Mavrou, V., (2006). Modeling The Dynamics Of Electricity Prices, Master's thesis, ETH Zurich and University of Zurich.

[18] Øksendal, B. (1998): *Stochastic Differential Equations.* Springer-Verlag.

[19] Phillips, P.C.B. (1987). Time series regression with a unit root. *Econometrica* **55**: 863-898

[20] Pilipovic, D., (1998). *Energy risk: Valuing and managing energy derivatives*, McGraw-Hill, New York.

[21] Pindyck, R. (1999). The long-run evolution of energy commodity prices. *Energy Journal*, April, IAEE.

[22] Ptak, P., Jabłońska, M., Habimana, D, and Kauranne, T, (2008) Reliability of ARMA and GARCH models of electricity spot market prices. *Proceeding of European Symposium on Time Series Prediction 2008*. Porvoo, Finland

[23] Vandaele, W. (1983). *Applied time series and Box-Jenkins model.* Academic Press, INC, United States.

[24] Weron, R. (2006). *Modeling and forecasting electricity load and prices: A Statistical Approach.* John Wiley and Sons Ltd.

[25] http://www.nordpool.com

[26] http://www.sitmo.com