CHALMERS | GÖTEBORG UNIVERSITY

MASTER'S THESIS

Overview of Stochastic Models for Asset and Commodity Prices.

CHENYANG ZHANG

Department of Mathematical Statistics
CHALMERS UNIVERSITY OF TECHNOLOGY
GÖTEBORG UNIVERSITY
Göteborg, Sweden 2012

Thesis for the Degree of Master of Science (30 credits)

Overview of Stochastic Models for Asset and Commodity Prices.

Chenyang Zhang

CHALMERS | GÖTEBORG UNIVERSITY

Department of Mathematical Statistics
Chalmers University of Technology and Göteborg University
SE – 412 96 Göteborg, Sweden

Göteborg, 2012

Abstract

In this thesis we investigate and overview the pricing of different assets and commodities using stochastic models, with particular focus in finance, the electricity-market and the evolution of temperature.

Acknowledgements

First and foremost I would like to express my deepest gratitude to my supervisor Patrik Albin, for his guidance, unending support and encouragement no matter the circumstances. I would also like to thank my friends, colleagues and teachers for making the time at Chalmers memorable. Finally I would like to thank my family for their support.

Contents

1 Introduction

1.1 Brownian Motion

Definition 1. *A stochastic process* $W = \{W(t), t \geq 0\}$ *defined on a probability space* $(\omega, \mathcal{F}, \mathbb{P})$ *is a standard Brownian motion if the following properties are satisfied:*

(i) $W(0) = 0$ *almost surely.*

(ii) $W(t)$ *has independent increments. that is for any* $0 \leq s \leq t$*, we have that* $W(t) - W(s)$ *is independent of all* $W(r)$ *for* $r < s$*.*

(iii) $W(t)$ *has stationary increments; that is for any* $0 \leq s \leq t$*, we have that* $W(t) - W(s) \stackrel{d}{=} W(t-s)$*.*

(iv) $W(t+s) - W(t)$ *is Normally distributed with mean* 0 *and variance* $s > 0$*:* $W(t+s) - W(t) \sim N(0, s)$*.*

1.2 Lévy Processes

Definition 2. *A lévy process* $L = \{L(t), t \geq 0\}$ *is a stochastic process defined on a probability space* $(\omega, \mathcal{F}, \mathbb{P})$ *which satisfies the following properties:*

(i) The paths of $L(t)$ *are right continuous with left limits almost surely(Cádlág).*

(ii) $L(0) = 0$ *almost surely.*

(iii) $L(t)$ *has independent increments. that is for any* $0 \leq s \leq t$*, we have that* $L(t) - L(s)$ *is independent of all* $L(r)$ *for* $r < s$*.*

(iv) $L(t)$ *has stationary increments; that is for any* $0 \leq s \leq t$*, we have that* $L(t) - L(s) \stackrel{d}{=} L(t-s)$*.*

1.2.1 Some Useful Distributions

Here we list some useful distributions used in conjunction with Lévy processes.

Gaussian Distribution

The Gaussian or Normal distribution is by far the most well-known and important distributions, its density function is given by:

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

and the characteristic function:

$$\varphi_f(t) = \exp\{i\mu t - \frac{1}{2}\sigma^2 t^2\}$$

Poisson Distribution

The Poisson distribution is a discrete probability distribution with the density function:

$$f(k;\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$$

and the characteristic function:

$$\varphi_f(t) = \exp(\lambda(e^t - 1))$$

Gamma Distribution

The Gamma distribution is a continuous probability distribution with the density function:

$$f(x;k,\theta) = x^{k-1}\frac{e^{-x/\theta}}{\theta^k \Gamma(k)}$$

and the characteristic function:

$$\varphi_f(t) = (1 - \theta i t)^{-k}$$

Inverse Gaussian Distribution
The Inverse Gaussian distribution has the probability density:

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

and characteristic function:

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

Generalized Hyperbolic Distribution
Used in finance for its heavy tails, the Generalized Hyperbolic Distribution has the following pdf:

$$f(x;\lambda,\alpha,\beta,\delta,\mu) = \frac{(\sqrt{\alpha^2-\beta^2}/\delta)^\lambda}{\sqrt{2\pi}K_\lambda(\delta\sqrt{\alpha^2-\beta^2})}\, e^{\beta(x-\mu)} \times \frac{K_{\lambda-1/2}\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)}{\left(\sqrt{\delta^2+(x-\mu)^2}/\alpha\right)^{1/2-\lambda}}$$

where K is the modified Bessel function.

The Meixner Distribution

The Meixner distribution introduced by Schoutens and Teugels(1998) is given by:

$$f(x; a, b, d, m) = \frac{(2\cos(b/2))^{2d}}{2a\pi\Gamma(2d)} \exp\left(\frac{b(x-m)}{a}\right) \left|\Gamma\left(d + \frac{i(x-m)}{a}\right)\right|^2$$

The characteristic function of the distribution is given by:

$$\varphi_f(t) = \left(\frac{\cos(b/2)}{\cosh\frac{at-ib}{2}}\right)^{2d} \exp(imt)$$

1.3 Time series

A common practice in stochastic modeling is the use of time series forecasting. A time series is a sequence of data points, measured typically at uniform time intervals, as such, time series are treated in the discrete time frame. Time series forecasting uses a model to predict future data points based on previous known data points.

Time series operate in the discrete time framework, and in this thesis time will be denoted with a subscript X_t whereas the continuous time framework counterpart will be denoted by $X(t)$.

1.3.1 Useful time series models

Below we have a list of useful time-series in price modeling.

Autoregressive (AR) model

The autoregressive model of order p: AR(p), is defined as:

$$X_t = \mu + \sum_{i=1}^{p} \alpha_i X_{t-i} + \epsilon_t,$$

where μ is the mean, α_i's are parameters of the model and ϵ_t is white noise.

Autoregressive conditional heteroscedasticity (ARCH)

$$X_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2$$

Where ϵ_t is white noise.

Generalized autoregressive conditional heteroscedasticity (GARCH)

Introduced by Bollerslev(1986), the GARCH(p,q) process is given by:

$$X_t = \sigma_t \epsilon_t$$

where ϵ_t is a strong white noise process and:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2$$

Autoregressive fractionally integrated moving average (ARFIMA)

An ARFIMA(p,d,q) model is described by:

$$\left(1 - \sum_{i=1}^{p} \phi_i B^i\right) (1 - B)^d X_t = \left(1 + \sum_{i=1}^{q} \theta_i B^i\right) \epsilon_t \, .$$

where B^i is the differencing operator: $B^i X_t = X_{t-i}$.

1.4 Mean-Reversion

Mean reversion is a methodology commonly used when dealing with stochastic processes. It is built on the assumption that the high's and low's of a process is temporary and that the process has a tendency to gravitate towards its mean value. . This makes mean-reversion an invaluable tool in the modeling any asset or commodity whose price cannot rise indefinitely.

A common way to model mean-reversion in a price process is through the use of Ornstein-Uhlenbeck processes, given by the differential equation:

$$dX(t) = \alpha(\mu - X(t))dt + \sigma dW_t$$

where θ, σ and μ are parameters and $W(t)$ is standard Brownian motion. Conceptually we see that this process has an expected change at time t:

$$E[dX(t)] = \alpha(\mu - X(t))dt$$

we see that if $X(t) > \mu$, the expected change becomes negative, and if $X(t) < \mu$, the expected change becomes positive, so the process is expected to always approach the mean level (represented by the parameter μ) as it fluctuates above and below it.

The AR(1) process is the discrete time counterpart of the Ornstein-Uhlenbeck process, and is used for the same purpose in time series modeling.

2 Stock Market

2.1 Introduction

The strive to understand the behavior of stock markets has its roots back in 1900, when Bachelier(1900) suggested the Bachelier model for modeling the price $S(t)$ of a financial asset:

$$S(t) = S(0) + \sigma dW(t),$$

where $\sigma > 0$ is a parameter and $W(t)$ is a standard Brownian motion. The model described the change in the price as a standard Brownian Motion. The model had a major weakness in that it could allow asset prices to take negative values, which is not feasible in the real world.

The weakness was addressed later by Samuelson(1965), who proposed the Bachelier-Samuelson model which states that:

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dW(t),$$

which has the analytic solution:

$$S(t) = S(0)\exp\left((\mu - \frac{\sigma^2}{2})t + \sigma W(t)\right).$$

The Bacherlier-Samuelson model instead describes the log-asset-price as a Brownian motion. This process was coined geometric Brownian motion.

2.2 Geometric Brownian motion

Geometric Brownian motion has turned out to be a central model in finance. Amongst its triumphs are the success of the Black-Scholes option pricing formulas presented by Black and Scholes(1973) in their seminal paper(for which the authors received the Nobel prize in economics), which assume the stock price process to obey Geometric Brownian Motion. Geometric Brownian motion has henceforth been called the Black-Scholes model in the context of mathematical finance and stock price models.

2.2.1 Weaknesses of the Black-Scholes model

The Black-Scholes model has shown to have many imperfections. The most prominent ones pointed out by Schoutens(2003) are:

- We see that log returns, which are assumed to follow the Normal distribution, in fact, do not. This can be, and has been verified empirically.
- Squared ACF from empirical data show that the assumption that the log-returns are time-independent, in fact may not be.

- Empirical data show negative skewness and excess kurtosis, compared to that of the Normal distribution.
- The volatility parameter which is assumed to be constant in the Black-Scholes model, appears to be invalid, especially with the common occurrence of volatility clusters.

2.2.2 Some possible solutions

Below we provide some of the innovations aimed to remedy the weaknesses of the Black-Scholes model.

- The Lévy market model:

  The Lévy market model has gained greatly in popularity recently, it uses Lévy processes to model the randomness in log-returns instead of standard Brownian motion shocks. This is to address the inadequate normality-assumption in the Black-Scholes model.
- Time-series models:

  Time series models are used mainly to address the issues of time-dependency of log-returns as well as incorporating dynamic evolution of the volatility into the model. The most commonly employed time-series in the context of stock prices are of the family ARCH.
- Stochastic volatility models:

  Stochastic volatility models aim to address the volatility trends that stock prices exhibit which the Black-Scholes model doesn't capture. This includes describing the volatility itself as a stochastic process.

2.3 Lévy market models

As empirical log-returns of stock prices has proven not to be normally distributed and has shown to exhibit skewness and excess kurtosis, one natural approach is to model the log-returns with some other distribution that fits the empirical properties better. The Lévy market model assumes the stock price to behave like an exponential Lévy process:

$$S(t) = S(0)\exp(L(t))$$

where $L(t)$ is a Lévy process. The log returns of the stock price under this model will have the same distribution as the increments of the chosen Lévy process. The aim is then to choose an appropriate Lévy process so that the model fits empirical data better. The primary aim of the Lévy market model is to address the fallacy of the normality-assumption of the Black-Scholes model, as well as the finding distributions with better fits to the empirical moments.

2.3.1 Interesting Lévy processes in finance

It is quite obvious that the distributions that exhibit the same excess kurtosis and skewness are of particular interest when it comes to using Lévy processes in finance. Jump processes are also sometimes incorporated into models to capture the sudden large movement that occurs in the financial market.

Brownian Motion

Brownian motion is a Lévy process with normally distributed increments. Geometric Brownian motion is a Lévy market model driven by Brownian motion with drift.

Poisson process

The Poisson process is a pure jump process which has stationary increments that are distributed according to the Poisson distribution. This processes is typically incorporated into models to capture jump behavior, much like the way it is used for the same purpose in modeling electricity prices.

Generalized Hyperbolic Process

The use of the Generalized Hyperbolic Distribution to model log-returns has been met with success as of recently, this is mainly due to the distribution having semi-heavy tails; a very desirable quality judging from empirical data.

The Mexiner Process

Introduced by Schoutens and also suggested as a viable choice for modeling stock prices. The Meixner process has increments which are distributed according to the Meixner distribution specified in section 1.3.6.

The Meixner distribution has the following properties:

	Meixner(a, b, d)	Meixner$(a, 0, d)$
mean	$ad\tan(b/2)$	0
variance	$\frac{1}{2}a^2d(cos^{-2}(b/2)$	$\frac{1}{2}a^2d$
skewness	$\sin(b/2)\sqrt{2/d}$	0
kurtosis	$3 + (2 - cos(b))/d$	$3 + 1/d$

As empirical data has consistently shown to have a kurtosis greater than that of the Gaussian distribution, this makes the Meixner distribution a desirable choice for the Lévy market model.

2.3.2 Addition of Drift Term

As Schoutens(2003) points out, the Meixner an General Hyperbolic process has the added advantage that a drift term can be added without complicating the model too much. Much like the need to find a risk-neutral measure in the case of the Black Scholes model, the Meixner and General Hyperbolic processes require a similar transformation when a drift term is added.

2.3.3 Calibration and Simulation

The calibration of Lévy market models conceptually is as simple as fitting the empirical log-returns of the process to the distribution of the driving Lévy process. Maximum likelihood estimation is typically used in fitting the log-returns to the distribution of interest, in fact all Lévy processes presented can be fitted using maximum likelihood. For example, Geometric Brownian Motion, which is a Lévy market model driven by Brownian motion with drift, can be calibrated by simply fitting the empirical log-returns to the normal distribution, whose maximum likelihood estimates are the sample mean and variance.

Simulation of Lévy processes typically involves generating random variables from the underlying distribution, and then a simple Euler-scheme to construct the process from the generated random variables. The Lévy market model can then be constructed from the Lévy process. For an overview of the simulation of some Lévy processes, we refer to Schoutens(2003).

2.4 Modeling volatility

The constant volatility assumption made by the Black-Scholes model has proven to be unrealistic, therefore there have been models created to address this. One can model the volatility with a deterministic function, creating a so-called local volatility model. Another approach is to assume that the volatility itself is a stochastic process, resulting in a so-called stochastic volatility model.

2.4.1 Heston model

Perhaps the most well-known stochastic volatility model is the Heston model, named after Steven Heston(1993). It assumes that the volatility of the asset follows a random process rather than being constant. The Heston model assumes that the asset price $S(t)$ follows the following stochastic process:

$$dS(t) = \mu S(t)dt + \sqrt{\sigma(t)}S(t)dW_1(t)$$

where $\sigma(t)$ is determined by the stochastic process:

$$d\sigma(t) = \alpha(\theta - \sigma(t))dt + \xi\sqrt{\sigma(t)}dW_2(t)$$

$$W_1(t) \cdot W_2(t) = \rho$$

where μ is the mean rate of return, $\sigma(t)$ the stochastic volatility, θ the mean volatility, α the mean-reversion factor of the volatility, and ξ the variance of the volatility.

The Heston model has closed form solutions for call prices, and therefore can be calibrated to market price of standard calls using least-squares fit.

2.5 Time-Series

Time Series are also extensively researched for their applications within finance. The main advantage is that time series can be used to great effect in capturing time dependencies particularly for the volatility. Volatility clusters that empirical data tend to exhibit can be modeled to great effect with the use of time series.

Typical time series model involve ARMA models for the price, and GARCH models for volatility. There are many also variations of the GARCH model being implemented to capture specific behaviors.

2.6 Concluding Remarks

While innovation is at an all time high in finding ever better models for finance, Black-Scholes remains the most popular model, mostly due to the extensive and clear as well as tried and proven results in option pricing and other applications it has had over the past decades, it is for a lack of a better description the most practical model still to date.

In essence, practicality remains the biggest issue when it comes to finding a good applicable model to the financial market. Besides the methods discussed in trying to improve upon the Black Scholes model, combinations have also been made, such as Lévy models with stochastic volatility. Efforts have also been made in combining Lévy processes into Ornstein-Uhlenbeck processes.

3 Interest Rates

3.1 Introduction

Interest rates are a fundamental asset to finance.

Interest rates exhibit mean reversion, an essential characteristic of the interest rate that sets it apart from other financial prices. Thus, as opposed to stock prices for instance, interest rates cannot rise indefinitely. This is because at very high levels they would hamper economic activity, prompting a decrease in interest rates. Similarly, interest rates can not decrease indefinitely. As a result, interest rates move in a limited range, showing a tendency to revert to a long run value.

3.2 Short-rate models

The most well-known short-rate models can be divided into two categories:

- Equilibrium Models: These models are derived as a result of certain market assumptions for achieving economic equilibrium. Here the model itself is the output.
- No Arbitrage models: In contrast to the equilibrium models, no arbitrage models takes the short rate model as input and calibrates it to existing market data.

3.3 Equilibriums Models

3.3.1 Vasicek model

One of the most well-known models for describing the evolution of interest rates is the Vasicek model introduced by Vasicek(1977), the model states that the instantaneous interest rate follows the stochastic differential equation:

$$dr(t) = \alpha(\mu - r(t))dt + \sigma dW(t)$$

where $a > 0$,b and σ are constant parameters and $W(t)$ is standard Brownian motion, this is a mean-reverting process also known as the Ornstein-Uhlenbeck process

One notable weakness in the Vasicek model is the fact that the model can undertake negative values, which is not desirable. This weakness was addressed by the Cox-Ingersoll-Ross model.

3.3.2 Cox-Ingersoll-Ross model

Cox-Ingersoll-Ross(1985) (hereby referred to as CIR) provides the following model, given by the following stochastic differential equation which is also known as the CIR process:

$$dr(t) = \alpha(\mu - r(t))dt + \sigma\sqrt{r(t)}dW(t)$$

The model contains a mean-reverting drift factor just like in the Vasicek model, and both models provide analytical formulae for the zero-coupon bond prices, the main difference lies in the standard deviation factor. The factor $\sigma\sqrt{r(t)}$ in the Cox-Ingersoll-Ross model approaches zero as the interest rate approaches zero, this ensures that the interest rate process will never take negative values, thus addressing the main weakness in the Vasicek model.

3.4 Multi-factor models

The equilibrium modeling approach can be easily extended to several factors. We give some examples of the best-known models in this category

3.4.1 Longstaff-Schwartz model

Longstaff and Schwartz extends general framework of the CIR model by assuming that the short-rate dynamics are given by the two equations:

$$dX(t) = \alpha_x(\mu_x(t) - X(t))dt + \sqrt{X(t)}\sigma_x(t)dW_1(t)$$

$$dY(t) = \alpha_y(\mu_y(t) - Y(t))dt + \sqrt{Y(t)}\sigma_y(t)dW_2(t)$$

where both factors are assumed to affect the mean; r is assumed to be a linear combination of X and Y, but only one factor is assumed to affect the instantaneous variance.

3.4.2 Chen Model

The Chen(1996) model is a three factor model which assumes that the mean and volatility of the short rate are stochastic, it has the following specification:

$$dr(t) = (\mu(t) - r(t))dt + \sqrt{r(t)}\sigma(t)dW_1(t)$$

$$d\mu(t) = (\alpha(t) - \mu(t))dt + \sqrt{\mu(t)}\sigma(t)dW_2(t)$$

$$d\sigma(t) = (\beta(t) - \sigma(t))dt + \sqrt{\sigma(t)}\eta(t)dW_3(t)$$

This model provides more flexibility than one-factor models, but at the expense of higher computing demands.

3.5 No arbitrage models

3.5.1 Ho-Lee model

The following model proposed by Ho and Lee(1986) typically given in the discrete time framework, is given by the SDE:

$$dr_t = \theta_t dt + \sigma dW_t$$

The model's main advantage is its simplicity, and the ease at which it can be calibrated to market data, where the deterministic function θ_t is chosen so that observed and theoretical prices for the zero coupon bond match. This model is a no-arbitrage model, because its calibration involves an initial calibration to observed prices there is no arbitrage opportunity at the outset. It is also one of the few interest rate models that does not allow for mean reversion.

3.5.2 Hull-White model

A possible extension of the Vasicek model to being a time-inhomogeneous no-arbitrage model, or an extension of the Ho-Lee model that allows for mean reversion. We get the following model:

$$dr(t) = (b(t) - a(t)r(t))dt + \sigma^{\beta}(t)dW(t)$$

where $b(t)$ and $a(t)$ and $\sigma(t)$ are deterministic functions chosen to match the theoretical and observed zero coupon bond prices or possibly other derivative prices much like in the Ho-Lee model. Choosing $a(t)$ and $\sigma(t)$ to be constant gives us the Hull-White model. This model has the advantage of analytic tractability and like the Vasicek, CIR, and Ho-Lee models, are examples of affine term-structure models, which means that the zero coupon bond prices $P(t, T)$ can be written in the form:

$$P(t, T) = \exp(A(T - t) + B(T - t)r(t))$$

Worthy to note is that the SDE presented above, depending on the choice of parameters can nest several other known models and/or extensions of them.

3.5.3 Black-Karasinski model

The Black-Karasinski model follows the SDE:

$$dr(t) = r(t)(a(t) - b(t)\ln r(t))dt + \sigma(t)r(t)dW(t)$$

where $a(t)$,$b(t)$ and $\sigma(t)$ are calibrated to initial market observations much like in the Hull-White model. This model does not have any analytical solutions, other than the fact that it can be shown that $r(t)$ is log-normally distributed, however it has uses in the development of fast numerical methods for market calibration and price calculation.

3.6 Nesting Models

Chan, Karolyi, Longstaff and Sanders(1992) (CKLS) proposed a generalized stochastic differential equation which nests many different models depending on parameter choices:

$$dr(t) = (\alpha + \beta r(t))dt + \sigma r(t)^{\gamma} dW(t)$$

this is a good way of providing an overview over many different models and comparing them, in similar spirit Ait-Sahalia(1996) proposed the following SDE:

$$dr = (\alpha_0 + \alpha_1 r + \alpha_2 r^2 + \alpha_3 / r)dt + \sqrt{\beta_0 + \beta_1 r + \beta_2 r^{\beta_3}} dW(t).$$

It nests most of the well known interest rate models (and some previously unmentioned) as well as other known diffusion models as shown in the table below:

Model	α_0	α_1	α_2	α_3	β_0	β_1	β_2	β_3
Merton	α	-	-	-	σ^2	-	-	-
Geometric Brownian Motion	-	β	-	-	-	-	σ^2	2
Vasicek	α	β	-	-	σ^2	-	-	-
CIR	α	β	-	-	-	σ^2	-	-
CIR VR	-	-	-	-	-	-	σ^2	3
Brennan Scwartz	α	β	-	-	-	-	σ^2	2
Dothan	-	-	-	-	-	-	σ^2	2
CEV	-	β	-	-	-	-	σ^2	2γ
CKLS	α	β	-	-	-	-	σ^2	2γ

Durham(2002) used this approach to test a selection of interest rate models:

Model	α_0	α_1	α_2	α_3	β_0	β_1	β_2	β_3
Affine	✓	✓-	-	-	✓	✓	-	-
CEV1	✓	-	-	-	-	-	✓	✓
CEV2	✓	✓	-	-	-	-	✓	✓
CEV4	✓	✓	✓	✓	-	-	✓	✓
GEN1	✓	-	-	-	✓	✓	✓	✓
GEN2	✓	✓	-	-	✓	✓	✓	✓
GEN4	✓	✓	✓	✓	✓	✓	✓	✓

3.7 Calibration and Simulation

Estimating parameters from real data is typically done with maximum likelihood estimation. The Vasicek and CIR models can be calibrated via least-squares regression, but also since they have well-known closed-form transition densities they can be easily calibrated through maximum likelihood estimation. The Vasicek model can be simulated from its analytical solution, while the CIR can be simulated via exact simulation in lieu of its known transition density. Typically Euler-schemes can be used to simulate the models, though it must be noted that for the CKLS model, the

Euler-scheme will not work for a high parameter γ, therefore an alternate method, such as simulation by transformation has to be used.

However, not all models have such nice properties or transition densities in closed form. Typically approximated maximum likelihood estimations are used in those cases. Approximated maximum likelihood involves approximating a transition density after discretizing the process. This may also involve the need to approximate unobserved process values. One such method is the simulated maximum likelihood method by Pedersen(1995).

4 Electricity prices

4.1 Introduction

Deregulation of electricity markets is a recent and still ongoing process. Previously, price variations were often minimal due to control from the regulators, and customers had little to no influence over the prices, deregulation introduces competition into the market.

Deregulation results in a market where there can be any number of distributors, prices variations have increased as a result of a competitive market because customers the choice to purchase electricity from the distributor of their choice. The result of deregulation has made the trading of electricity a very special commodity market in itself, and many power exchanges were established around the world to facilitate the trading of this commodity, the market itself provides a very interesting object of study.

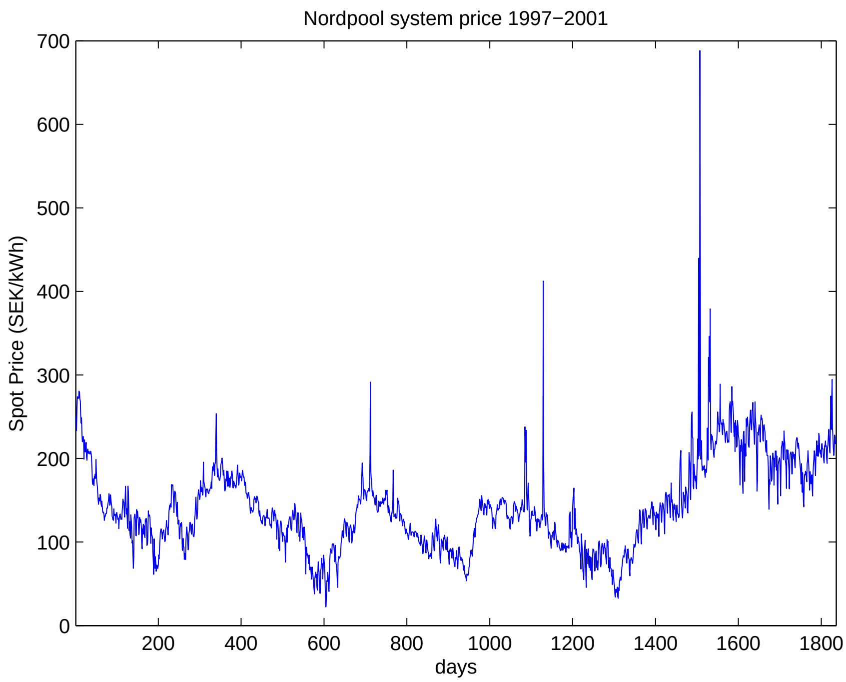

Figure 1: The daily spot price at Nordpool from 1997-01-01 to 2001-12-31

In figure 1 we have the daily spot price at nordic power exchange Nordpool spanning 1997 to 2001, this is a very typical example of a electricity price trajectory. We see large jumps and spikes, small sporadic variations around a mean level, and periodic behavior.

A well-known feature of electricity prices stems from the fact that electricity is very expensive and difficult to store, but yet it is an essential and vital commodity for everyday life. As a result, with supply and demand and healthy competition in modern day electricity market, electricity prices have tendencies of big fluctuations with strong daily, weekly and yearly periodicity, and markets have to be kept up to date on a second to second basis.

Another common feature of electricity prices is the existence of large seemingly ineligible jumps in the price. These jumps are mainly due to the fact that demand is a fairly inelastic function of price, because the majority of customers must be serviced at all times. Since electricity is difficult to store, there is no buffer in supply in the case of supply changes due to outages or failures, hence big price spikes are often observed in markets worldwide. Perhaps the most well-known case is in Chicago 1998, where electricity prices increased 100-fold, this event was influential in showing the need for ways to accurately model electricity prices.

Finally, like many other commodities, electricity prices exhibit mean-reversion, unlike the stock-market, where prices are allowed to evolve freely, electricity prices over longer periods of time gravitate towards the levels of the production costs.

To sum it up, there are some key features of electricity prices that most authors recognize when attempting to model it. These features of the price process include:

- Seasonality and periodicity: there is higher consumption of electricity during winters, and also during summers in countries with hot summers(not as much so in Europe), in countries where hydropower is prominent, seasonality is exhibited through the conditions of water reservoirs and their effect on production. There is also seasonality on much smaller times frames, in the form of prices during on-peak and off-peak hours.

- Mean reversion: the tendency of electricity prices to gravitate towards the levels of the production costs.

- Small variations around average: there are always small variations around the average caused by temporary supply/demand imbalances in the network.

- Jumps and Spikes: extreme price variations characterized by sharp jumps, most of which are relatively short-lived. Spikes occur in the form a sharp jump in one direction followed immediately by another jump in the opposite direction.

Any good model for electricity prices should take these features into account, and most authors incorporate most if not all features into their models, depending on which electricity market they attempt to model.

4.2 The Overview:

The study of electricity spot price models is an ongoing process, often consisting of extensions of existing models in an attempt to achieve a model that is better than the previous, many ideas are recurring. We shall hence start by looking at the most basic models, and various recurring themes and extensions, to get a wider picture over how many models are related.

4.3 Seasonality and Mean-reversion

We begin with the one of the most basic model for electricity spot price, presented by Lucia and Schwartz. It gives the spot price at time t as:

$$E(t) = \alpha(\mu(t) - E(t))dt + \sigma dW(t)$$

where $\mu(t)$ is a deterministic function accounting for seasonality.

This model, proves to be inadequate for modeling electricity due to its failure to capture jumps and spikes; albeit not surprising as the author's did not consider jump components in the first place. Despite this however, this model provides the basic building block of almost all power-market models, as it incorporates seasonality and mean-reversion.

4.4 Capturing Spikes

The main challenge of modeling electricity prices is to effectively and accurately capture jumps and spikes that electricity prices tend to exhibit. Models that attempt to capture spike behavior typically come in one of two categories: regime switching models, which treat spikes separately from the normal process, and non-regime switching models which treat spikes as a part of the price process.

4.5 Non-regime switching models

Models that fall under this category are typically jump-diffusion models, that is, diffusions models with jump components. The main difficulty in this approach is finding an accurate way to bring prices down to "normal" levels after a jump, and reproducing spike behavior.

4.5.1 Jump-diffusion models

The most basic jump-diffusion models for electricity spot prices builds upon the Merton equation, where mean reversion has been introduced into the drift term. A jump-diffusion process takes the general form:

$$dE(t) = \mu(t, E(t))dt + \sigma dW(t) + \phi dN(t)$$

- μ is the drift term, in the context of modeling electricity prices this is chosen to account for seasonality and mean-reversion.

- $W(t)$ is standard Wiener process, σ the volatility.

- $N(t)$ is a counting process describing jump arrival.

- ϕ is a random variable describing jump magnitude.

Many different specifications to the particular parameters exist. The volatility may be chosen to be constant, deterministically varying, or stochastic. The choice of counting process is typically a Poisson process with constant intensity, however, there are many options, and many options also exist for the choice of random variable ϕ. One of the main weaknesses of the basic jump diffusion process is its reliance on a high mean-reverting parameter to return prices to normal levels after a jump.

A typical example of a jump-diffusion model include the following jump-diffusion model with stochastic jump size considered by Escribano, Peña and Villaplana(2002) which describes the log spot price $E(t)$ as:

$$dE(t) = \alpha(\mu - E(t))dt + \sigma dW(t) + dJ(t)$$

where $J(t)$ is a compound Poisson process with jump size J_i which are independently distributed random variables, that is to say:

$$J(t) = \sum_{i=1}^{N(t)} J_i$$

where $N(t)$ is a poisson process with constant intensity. This particular model includes the pretty standard mean-reverting component along with Brownian motion fluctuations with constant volatility, however, the jump component only captures upward jumps, and hence requires a usually unrealistically high mean-reversion component to pull prices back after a jump. The constant jump intensity is also unrealistic. In a similar vein, Cartea and Figueroa (2005) proposes a similar model with time-varying volatility.

Another notable model that fall under the category of non-regime switching models is a model involving two sources of risk proposed by Villaplana(2004). It considers two different variables X and Y representing short-term and long-term price variations respectively. The model describes the spot price $S(t)$ as:

$$
\begin{aligned}
&\ln S(t) = f(t) + X(t) + Y(t), \\
&dX(t) = \alpha_1[\mu_1 - X(t)]dt + \sigma_1 dW_1(t) + J(t)dN(t), \\
&dY(t) = \alpha_2[\mu_2 - Y(t)]dt + \sigma_2 dW_2(t), \\
&dW_1(t) \cdot dW_2(t) = \rho dt,
\end{aligned}
$$

where $f(t)$ is the deterministic component of the log spot price. This model likewise suffers from relying solely on the mean reversion to bring high prices to a standard level.

4.5.2 Geman Roncoroni threshold Model

Geman and Roncoroni(2006) proposes a jump diffusion model with a unique feature called a threshold parameter. This model is particularly promising because it includes a mechanic to bring prices from high levels to standard levels without having to use a high mean-reversion parameter. The authors coined this type of models jump-reversion:

$$
\begin{aligned}
&E(t) = \ln S(t), \\
&dE(t) = d\mu(t) + \theta_1[\mu(t) - E(t^-)]dt + \sigma dW(t) + h(t^-)dJ(t),
\end{aligned}
$$

where:

- S(t) is the spot price, E(t) is therefor the log spot price.
- $\mu(t)$ is the mean trend, which is a deterministic function accounting for seasonality, Geman and Roncoroni specifies it as of the form: $\mu(t) = \alpha + \beta t + \gamma \cos(\epsilon + 2\pi t) + \delta \cos(\zeta + 4\pi t)$
- $\theta_1[\mu(t) - E(t^-)]dt$ is the mean reversion term pulling the price process towards the seasonal mean level, θ_1 is mean reversion parameter pulling prices to the mean trend $\mu(t)$
- $W(t)$ is standard Brownian motion.
- σ is the volatility.
- The discontinuous part of the process which is a compound Poisson process aims to reproduce spike occurrence.

Jumps are characterized by time of occurrence, size and direction:

$$J(t) = \sum_{i=1}^{N(t)} J_i$$

where:

- $N(t)$ is a Poisson process with time-varying intensity. More specifically the intensity function $i(t)$ is defined as:

  $$i(t) = \theta_2 s(t)$$

  where θ_2 is the maximum number of jumps per unit of time, and $s(t)$ is the normalized jump intensity shape:

  $$s(t) = \left[\frac{2}{1 + |\sin(\pi(t-\tau))/k|} - 1\right]^d$$

  The shape function is in place to make jump occurrences occur more often around peak dates and disperse away around it. The parameter k gives a periodicity of jump occurrence peaks of k years. The parameter τ sets the peak times, and d is the dispersion term that adjusts the speed of dispersion. To illustrate this we look at a plot of the shape intensity function for $k = 1$(peaking occurs every 1 year), $\tau = 1/2$(peak date at half way through the k years), for different values on the dispersion parameter d over a period of 3 years.

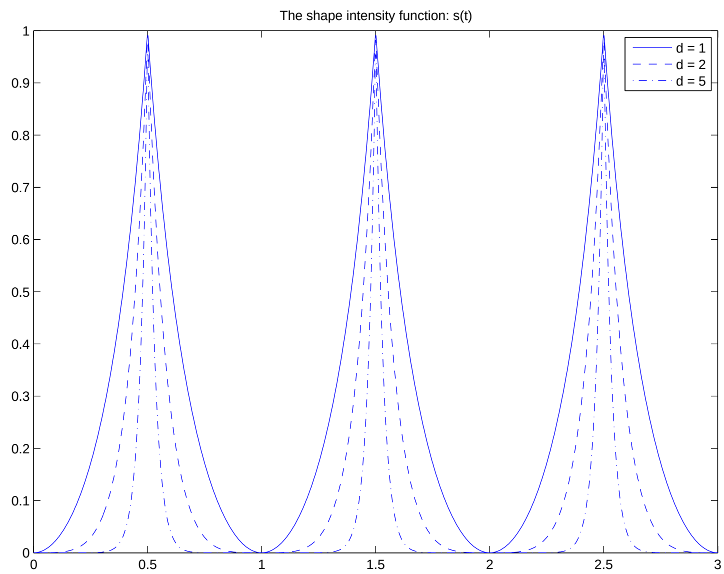

Figure 2: The shape intensity function for $k = 1$, $\tau = 1/2$, and $d = 1, 2, 5$

We see that the purpose of the shape intensity function is to make spikes more likely to occur during peak times, in the above case, spikes are most likely to occur June every year.

- J_i are independent and identically distributed random variables. The authors have chosen a truncated exponential density according to:

$$p(x; \theta_3, \psi) = c \times \exp(\theta_3 f(x)), \quad 0 \leq x \leq \psi$$

where c is a constant to ensure p is a probability distribution. The truncated density within the exponential family is meant to be chosen to reproduce the observed higher order moments.

- h is a threshold parameter that determines the direction of jumps:

$$h(E(t)) = \begin{cases} +1 & \text{if} \quad E(t) < \tau \\ -1 & \text{if} \quad E(t) \geq \tau \end{cases}$$

where the threshold τ is calibrated to market data.

The threshold parameter acts as a regime parameter of sorts. When the price process is under the threshold, the price is considered to be in a normal state, and all jumps that occur will be upward directed. When the price process is above the threshold the price process is considered to be in an abnormal state, and all jumps that occur will be downward directed.

The use of this threshold parameter is what makes this particular model unique, and it is also what makes this particular model very promising as it addresses in a relatively good way the issue of a signed jump without having to revert of high mean-reversion to bring prices back to a normal state.

This model was implemented by the authors on three American markets, and it has been revealed that this model captures the trajectorial and statistical properties of Electricity prices very well. The most promising aspect of this model is the fact that it does not rely on a strong mean-reversion parameter to bring prices back to normal, and the fact that its design all aims at capturing all the unique features of electricity. One possible weakness is that the intensity at which jumps tend to appear is very periodic, and this may not be appropriate for every market out there.

4.5.3 Calibration and Simulation

The principal parts involved in calibrating jump diffusion models consists of the following:

- Determining parameters of seasonality function
- Filtering the process into its continuous and discontinuous parts

- Determining the parameters of the jump component
- Determining mean reversion parameter and volatility of the diffusion component

All aforementioned models requires fitting the data to a deterministic seasonality function, this is usually done through ordinary least squares.

Finding the jump components from the data is typically done with a form of iterative filtering process. One example of such a filter was used by Cartea and Figueroa. It involves calculating the standard deviation of the returns of the data, and filtering out any returns 3 times greater than the standard deviation. This procedure is then iterated on the remaining returns, until no more outliers are filtered. The standard deviation of the jumps and the jump frequency can then be determined from the returns that were filtered out.

The mean reversion parameter of the remaining diffusion process is typically estimated through linear regression by regressing the returns of the log-spot price against the log-spot price itself. The volatility parameter, if assumed to be constant, can be estimated via the quadratic variation. If a time-dependent volatility is assumed, typically rolling historic volatility is used, see Eydeland and Wolyniec(2003).

Geman and Roncoroni proposes a different calibration procedure as their model is a bit more complex than the typical jump diffusion model. The determination of seasonality function and the disentangling of the data are done in similar fashion, although the filtration can be done with help from the threshold parameter instead of a recursive filter. The threshold parameter is typically chosen such that model matches the first four moments of the daily log-price return distribution.

Determination of the parameters for mean reversion, jump frequency and jump size are then done via a approximated likelihood function for the unknown process, which they derived from a prior process of the same class. Maximizing the likelihood function then yields a maximum likelihood estimate of the parameters. For the full details of the derivation of the log-likelihood function, we refer to Geman and Roncoroni(2006).

Simulation of the aforementioned models can all be done via Euler-discretization.

4.6 Regime Switching Models

The basic idea behind regime-switching models is the assumption that the price behaves differently depending on the state of the world for which they are in. Originally introduced by Hamilton in 1989, a regime-switching model is defined by how the process behaves in different regimes, and how the process alternates between those regimes. In the case of modeling electricity prices, the regimes typically represent "normal" price conditions and "extreme" price conditions.

4.6.1 Two-Regime model with stochastic jumps

Amongst the models that De Jong(2006) discusses, we have the following two regime model for the log spot price E_t with stochastic Poisson jumps.

An example of a typical regime switching model in discrete time framework, the price process has two regimes it can undertake: The Mean Reverting Regime:

$$dE_t = \alpha(\mu - E_{t-1}) + \sigma\epsilon_t, \quad \epsilon_t \sim N(0,1)$$

The Spike Regime:

$$dE_t = \alpha(\mu - E_{t-1}) + \sum_{i=1}^{n_t+1} Z_{t,i}, \quad \begin{cases} Z_{t,i} \sim N(\mu^S, \sigma^S) \\ n_t \sim \text{Poisson}(\lambda) \end{cases}$$

With the following Markov transition matrix:

$$\Pi = \begin{pmatrix} 1-\pi^S & \pi^S \\ \pi^M & 1-\pi^M \end{pmatrix}$$

This equations presented above are discrete versions of a standard mean-reverting process for the normal regime, for the spike regime the diffusion term has been replaced by a compound poisson process, to produce jump behavior.

This particular specification suffers from one problem, because the price process at a given time depend on previous levels, one still faces the problem of pulling prices from extreme price conditions back to normal. This model also has another problem in reproducing multiple consecutive spikes. The two models that follow are improvements upon this one.

4.6.2 Three-regime model

Huisman and Mahieu(2001) proposes the following three regime model for the log spot price E_t, as an improvement upon the previous two regime model: The Mean Reverting Regime M:

$$dE_t = \alpha(\mu - E_{t-1}) + \sigma\epsilon_t, \quad \epsilon_t \sim N(0,1)$$

The spike regimes take the form:

$$dE_t = \alpha(\mu - E_{t-1}) + \sum_{i=1}^{n_t+1} Z_{t,i}, \quad n_t \sim \text{Poisson}(\lambda)$$

where for the "up regime" U we have:

$$Z_{t,i} \sim N(\mu, \sigma)$$

and for the "down regime" D we have:

$$Z_{t,i} \sim N(-\mu, \sigma)$$

The process has the following Markov transition matrix:

$$\Pi = \begin{pmatrix} 1-\pi^S & \pi^S & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

This model ensures that a upward jump is directly followed by a downward jump, the intention is to allow the occurrence of spikes in the price process without needing excessively large mean reversion parameters to bring prices back to normal levels, this however, does not come without drawbacks.

The main drawback of this model is that multiple consecutive spikes are still not possible due to the restrictions of allowed transitions. Multiple up-ward jumps are not possible either, and it is also unrealistic to assume that in an electricity market every up jump is directly followed by a down jump.

4.6.3 Independent regimes model

De Jong and Huisman(2003) proposes the following model for the logspotprice E_t:
The Mean Reverting Regime:

$$dE_{M,t} = \alpha(\mu_M - E_{M,t}) + \sigma_M \epsilon_{M,t}, \quad \epsilon_{M,t} \sim N(0,1)$$

The Spike Regime:

$$E_{S,t} = \mu_S + \sigma_S \epsilon_{S,t}, \quad \epsilon_{S,t} \sim N(0,1)$$

With the following Markov transition matrix:

$$\Pi = \begin{pmatrix} 1-\pi^S & \pi^S \\ \pi^M & 1-\pi^M \end{pmatrix}$$

This regime switching model differs from the previous two in the way that here, the price processes for the mean-reverting regime and the spike regime are totally independent. Instead of having different descriptions for the evolution of the price process for the different regimes, where prices depend on previous price levels, this approach considers two processes running completely independently in parallel, with the price process assuming the values of one of the regimes at a time, while the other regime still runs, but is dormant and not observed.

This separation of the process into two parts also solves the issue of reproducing jump and spike behavior without having to rely on high mean-reverting parameter. De Jong and Huisman found this model to capture up to 50% more spikes than the Huisman-Mahieu three regime model while having similar mean-reverting parameters, which is not surprising seeing the three regime model is more restrictive on jump behavior as discussed earlier.

In De Jong's(2006) comparison between several models (mean reverting, jump-diffusion, the three aforementioned regime-switching models, and the Geman-Roncoroni Threshold model) on eight different(2 North-Eastern U.S and 6 European) electricity markets, De Jong concluded that while the independent and 3-regime model identified spike behavior better, the 2-regime jump diffusion regime switching model provided the best overall fit, suggesting some kind of tradeoff between overall model fit and spike identification. It is also worthy to note that in that particular study, the regime-switching models outperformed Geman-Roncoroni's threshold model.

This model was applied by the authors for pricing of European options. This model has the advantage that since the regimes are independent, option prices can be calculated for each component of the model, thus the final price is simply the sum of the two component prices. The option prices can be derived analytically using Black's formula.

4.6.4 Calibration and Simulation

The Regime Switching models presented above can all be calibrated using maximum likelihood estimation conditioned on each regime. The main challenge when calibrating therefore lies in determining which regime is active and dormant.

To determine the Markov transition probabilities, typically some form of filtering method is employed to disentangle the data into different regimes. Both Huisman and Mahieu as well as De Jong and Huisman proposes a Kalman filter and refer to Harvey(1989). Once the data has been disentangled, standard maximum likelihood estimation can then be used to determine the model parameters.

4.7 Other approaches

There are other methods that do not fit into the two major categories above. Barlow(2002) from an econometric angle proposes a Non-linear Ornstein Uhlenbeck process that is able to exhibit spikes behavior. A small yet significant number of models deal with modeling with time-series, the main motivation is that some markets(in particular the Nordic market) have shown to exhibit long memory properties, typically ARIMA models are considered in this context, such as by Tórro(2007). Benth et al.(2005) proposes a non-Gaussian Ornstein Uhlenbeck process which is a linear combination of pure jump Lévy processes, this approach manages to fit the stylized facts about electricity prices well, the use of non-Gaussian OU processes could also remedy problems involving the returns not being normally distributed.

4.8 Concluding remarks

To accurately model electricity spot prices the aspects of mean-reversion, seasonality, jumps and spike behavior must be considered. We see that most existing models account for seasonality and mean-reversion, but their treatise of spike behavior typically fall into one of two categories, jump-models and regime-switching models.

The most promising models appear to be those that are able bring prices back to normal levels after a jump without resorting to large mean-reverting parameters. Also worthy to note is that electricity spot price behaves differently from market to market, and markets change with time, making flexibility also a possible factor in determining what makes a good model.

5 Weather

5.1 Introduction

As of recently, weather has becomes a very hot area of research. Weather plays a key role in almost all commodity markets, being the prime factor behind yield, quality of agricultural commodities. Another example is the energy market, where temperature has a profound effect on demand, thus affecting prices. The strong correlation between energy market and weather is one of the main reasons why weather has become such an active field of research, often in tandem with different energy markets.

Weather is a prominent risk factor in many industries and businesses as well: Frost and heat-waves can ruin crops, as well as the droughts of over-precipitation. A myriad of companies may suffer when product sales drop from either a mild summer or warm winter. One main concern these companies face is to be able to protect themselves from the financial risks incurred by weather, and this has led to the rise in prominence in the weather derivatives market, as well as making it a very active area of research.

5.1.1 Weather derivatives

In short, weather derivatives are financial contracts much like existing derivatives such as options that provide a payoff based some measurable weather factor, thereby allowing companies to offset/hedge the financial risk caused by weather. Examples of how weather derivatives could be used would be:

- A farmer could purchase an option with payoff based on rainfall, to hedge against the risk of a drought.
- An ice cream vendor could safeguard against a mild summer by investing in options based on temperature.
- Energy providers can hedge against sudden fluctuations in demand caused by unexpectedly warm winters or cool summers.
- Weather derivatives can even be used in a portfolio in conjunction with other investments to reduce risk as weather is an asset class that is almost uncorrelated with any other class of investments.

Currently, the most common weather derivative traded are those based on temperature, the most well known market doing so is the Chicago Mercantile Exchange(CME). The index being used is expressed in "degree days", where a degree day is the difference between a reference temperature and the mean temperature on a given day. The typical reference temperature is $18°C(65°F)$, and the mean temperature is given be an arithmetic mean of the daily maximum and minimum temperatures. Which gives the following indices:

- Heating Degree Days: The number of degrees that a day's average temperature T is below the reference temperature: $\text{HDD}_j = \max(0, 18°C - T)$
- Cooling Degree Days: The number of degrees that a day's average temperature T is above the reference temperature: $\text{HDD}_j = \max(0, T - 18°C)$

Options are then written on "cumulative degree days" over a set period of time:

$$\text{CumHDD} = \sum_j \text{HDD}_j, \quad \text{CumCDD} = \sum_j \text{CDD}_j$$

Examples of how this index can be used is for example, if a company wishes to hedge its revenues against a warm winter, they could invest in put options whose pay-off at maturity T as:

$$P(T) = A \cdot \max(0, k - \text{CumHDD}(T))$$

where the number k is the total HDD over winter below which the company will suffer revenue damage, and A is the loss in revenue per missing HDD. Likewise, a company whose revenues would be damaged in case of a very cold winter can invest in a call option providing:

$$C(T) = A \cdot \max(0, \text{CumHDD(T)} - k)$$

at maturity to hedge against a cold winter.

Other weather indices such as precipitation, wind speed, humidity are also traded albeit to a much lesser extent.

The pricing of weather derivatives is still an active area of research. In fact, its a matter of speculation what pricing approaches are actually being used in the market. Many authors agree however that obtaining a good model for the underlying asset, which in this case is the weather parameter of interest, is an important step.

5.2 Modeling temperature: mean reversion

We shall here look at some notable models used to model temperature. One of the most common approaches is the use of mean-reversion models, Many existing models for temperature fall into this category. The choice of a mean-reversion model is intuitive in the sense that temperature does not rise or drop indefinitely and tend to gravitate towards a trend.

5.2.1 Mean-Reversion model with seasonality

Ornstein Uhlenbeck dynamics with seasonality was considered for example by Alaton et al.(2005):

$$dT(t) = ds(t) + \alpha(T(t) - s(t))dt + \sigma(t)dW(t)$$

where $s(t)$ is a deterministic function that accounts for seasonality of the form $s(t) = A + Bt + C\sin(\omega t - \phi)$, the addition of the linear term is due to the observed trend that mean temperature has a tendency to increase with time. The choice of the volatility parameter $\sigma(t)$ may vary. Alaton et al. chose to model it as a piecewise constant function, under the assumption that temperature volatilities vary on a monthly basis, but are constant for each month.

The analytical tractability of the OU process allows for explicit analytical expressions for pricing of weather derivatives such as an HDD option. Monte Carlo simulations is also an appropriate method of pricing such derivatives.

The results from Alaton et al.'s model show that the use of a mean-reverting model with seasonality fit temperature data of their region of interest very well. They also found that quadratic variation $\sigma^2(t)$ was nearly constant for each month n their data set, which meant the choice of volatility function was a good one. The only potential weakness in this model approach as with many other fields is that temperature differences may not be normally distributed, but the author's found the normality-assumption adequate, but they do make note that a better choice of background noise could contribute to improving the model.

5.2.2 Choice of background noise

Some variations to the above mean-reversion model can be found by changing the noise which drives the OU process. Brody et al(2002). after having observed considerable fractional behavior in temperatures from central England proposed a fractional Ornstein Uhlenbeck process:

$$dT(t) = \alpha(T(t) - s(t))dt + \sigma(t)dB^H(t)$$

where B^H is fractional Brownian motion, the authors however, did not attempt to fit data to the model at all. Benth and Saltyte-Benth(2004) however found out that fractional Brownian motion is a poor choice for modeling Norwegian temperatures, so the effectiveness of fractional Brownian motion is definitely region dependant.

Instead Benth and Saltyte-Benth proses a generalization of the Ornstein Uhlenbeck process above:

$$dT(t) = ds(t) + \alpha(T(t) - s(t))dt + \sigma(t)dL(t)$$

where $L(t)$ is a Lévy process, this approach is much in line with the recent and popular Lévy market model for stock prices, aimed in particular to model skewness and heavier tails than those exhibited in the normal distribution. The choice made by Benth and Saltyte-Benth based on flexibility and analytic tractability is the use of the Generalized Hyperbolic distribution to model the residuals. The SDE has the explicit solution:

$$T(t) = s(t) + (T(0) - s(0))e^{\kappa t} + \int_0^t \sigma(u)e^{\kappa(u-t)}dL(u)$$

5.2.3 Choice of volatility parameter

This is also a factor that varies a lot from model to model, mostly depending on the properties of the market under research. Therefore, the best choice of volatility varies depending on the data set.

Alaton et al. found that a piecewise constant function representing constant volatilities on a monthly basis was sufficient for their data. Doernier and Querel used constant volatility in their study of 20 years of temperature data for Chicago. Torró et al.(2001) found that a volatility parameter exhibiting GARCH(1,1) behavior suited their data set from the Spanish Temperature Index very well. Finally some models attempt to incorporate seasonality into the volatility parameter, which is a fair assumption.

5.2.4 Calibration and Simulation

Much like the electricity market, temperature models incorporate mean-reversion and seasonality. Calibration therefore tend to be very similar to that of the models presented in section 4.5.3 without the jump component calibration, that is the calibration of a seasonal component, along with a mean-reverting diffusion process. One has to note that when calibrating one could and probably should take into accounts the information provided by meteorological forecasts.

5.3 Modeling Temperature using Time Series

Another large category of models for temperature uses time series. In time series jargon, the aforementioned mean-reversion models all fall into the category of time series called AR(1)(or some extension of it), which is widely considered as the discrete time counterpart of Ornstein-Uhlenbeck processes. The main motivation is that temperature fluctuations may exhibit more time-dependencies than the markovian nature diffusion models, which in turn may require more complex models to capture, the use of time series is one way to tackle this problem.

5.3.1 An auto-regressive model

Under the motivation that daily temperatures should strong auto-regressive behavior; for example, a warm day is likely to be followed by another warm day, the following model was proposed by Cao and Wei(2000), who chose an auto-regressive model with 3 lags over a diffusion model.

Let $U_{yr,t}$ denote the temperature residuals, that is:

$$U_{yr,t} = T_{yr,t} - \overline{T_t}, \qquad yr = 1, ..., 20, \qquad t = 1, ..., 365$$

where the trend $\overline{T_t}$ is mean temperature of date t over the past years, in the authors' case 20 years:

$$\overline{T_t} = \sum_{yr=1}^{20} T_{yr,t}, \qquad t = 1, ..., 365$$

the temperature residuals $U_{yr,t}$ is then assumed to follow the following k-lag autocorrelation system:

$$U_{yr,t} = \textstyle\sum_{i=1}^{k} \rho_i U_{yr,t-i} + \sigma_{yr,t} \cdot \epsilon_{yr,t}$$

$$\sigma_{yr,t} = \sigma - \sigma_1 |(\sin(\pi t/365 + \psi))|$$

$$\epsilon_{yr,t} \sim \text{iid } N(0,1)$$

$$yr = 1, 2, \ldots, 20. \quad t = 1, 2, \ldots, 365$$

where ρ_i is the autocorrelation coefficient for the i:th lag.

5.3.2 The use of GARCH models to model volatility

Another common practice is the use of GARCH models to model volatility, much like what Torró et al.(2001) has done. This particular choice has turned out to be very popular, as temperature data tends to demonstrate volatility clusters, and as such, GARCH are used by some to model volatility in temperature residuals for the same reason why they are used in finance to account for volatility clusters.

Campbell and Diebold(2002) proposes the following ARMA model with ARCH dynamics in volatility:

$$T(t) = \text{Trend}_t + \text{Seasonal}_t + \sum_{l=1}^{L} \rho_{t-l} T_{t-l} + \sigma_t \epsilon_t$$

where

$$\text{Trend}_t = \beta_0 + \beta_1 t$$

$$\text{Seasonal}_t = \textstyle\sum_{p=1}^{P} \left[\delta_{c,p} \cos(2\pi p \frac{d_t}{365}) + \delta_{s,p} \sin(2\pi p \frac{d_t}{365})\right]$$

$$\sigma_t^2 = \textstyle\sum_{q=1}^{Q} \left[\gamma_{c,q} \cos(2\pi p \frac{d_t}{365}) + \gamma_{s,q} \sin(2\pi p \frac{d_t}{365})\right] + \sum_{r=1}^{R} \alpha_r \epsilon_{t-r}^2$$

$$\epsilon \sim N(0,1)$$

where d_t is a repeating step function taking the values 1 to 365, one value for each day in a year. We see the use of a linear trend, seasonality in both the temperature as well as its variance.

Benth and Saltyte-Benth(2007) propose a similar framework with an AR(3) process and seasonal volatility, the authors show that despite the relatively much lower lag count compared to Campbell and Diebold's model, the model is no less effective at capturing desirable properties.

Another example of such an approach is the model proposed by Franses et al(1998), who estimated and evaluated a univariate model for weekly mean Dutch temperature data. Their preliminary analysis revealed key features of the time series: a yearly seasonal pattern in the mean and volatility, volatility clustering, varying impact of high and low temperatures on the conditional volatility and the seasonality of that impact. As volatility clustering can be observed in finance where GARCH models are extensively used, Franses et al. used a QGARCH model on the volatility to accommodate this, the model specifications are as follows for the temperature T:

$$T_t = s(\vec{\mu}, t) + \phi T_{t-1} + Z_t$$

$$Z_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim N(0,1)$$

$$\sigma_t^2 = s(\vec{\omega}, t) + \alpha(\epsilon_{t-1} - s(\vec{\gamma}, t))^2 + \beta \sigma_{t-1}^2$$

where $\vec{\mu}, \vec{\omega}, \vec{\gamma}$ are parameter vectors, and the seasonality function is given by:

$$s(\vec{\lambda}, t) = \lambda_0 + \lambda_1 n_t + \lambda_2 n_t^2$$

where n_t is a repeating step function that goes from 1 to 52, a value for each week of the year. Note that setting the $\vec{\gamma}$ to zero would reduce the model to a standard GARCH model.

5.4 Long-memory time series

Recently, empirical evidence has been found that temperature time series exhibit long memory property, as such it has been popular to model temperature using long-memory time series, in particular time series of the class ARFIMA. We will not go into detail, we refer to Caballero and Jewson(2002), Caballero Jewson and Brix(2002) and Hamisultane(2006) for different approaches of modeling temperature using ARFIMA models.

5.5 Concluding remarks

We have seen typical examples of temperature models both in continuous and discrete time frameworks. We can see that all models include some sort of seasonality term, which is particularly useful for temperature modeling, a simple tweak in the seasonality term can ensure that the models keep up with the ever changing temperature trends in the real world. Another recurring factor are the mean-reverting or auto-regressive properties, that is temperatures cannot evolve freely, but rather must gravitate towards a trend.

Models differ mostly in the way volatility is modeled, most of these choices are based on regional differences in temperature trends, but some also based on well known imperfections and observations we've seen from finance such as normality-assumption of fluctuations and volatility clusters.

Some possible approaches that were not discussed were the use of stochastic volatility for mean-reverting models. In the time-series framework, ARFIMA models could be used on locations where temperature exhibit long-memory property.

References

[1] Ait-Sahalia Y. (1996). *Testing continuous time models of the spot interest rate.* Review of Financial Studies 9, 385-426.

[2] Alaton P., B. Djehiche and D. Stillberger (2005). *On modelling and pricing weather derivatives*

[3] Barlow M.T. (2002). *A diffusion model for electricity prices.* Mathematical Finance, 12, 287-298.

[4] Bachelier L. (1900) *Théorie de la Spéculation.* Annales Scientificques l'École Normale Supérieure 17, 21-86.

[5] Benth F.E., J. Saltyte-Benth (2004). *Stochastic modeling of temperature variations with a view towards weather derivatives.*

[6] Benth F.E., J. Saltyte-Benth (2007). *The volatility of temperature and pricing of weather derivatives.*

[7] Benth F.E., J. Kallsen and T. Meyer-Brandis (2005). *A non-Gaussian Ornstein-Uhlenbeck process for electricity spot price modeling and derivative pricing.*

[8] Black, F. and M. Scholes (1973). *The pricing of options and corporate liabilities.* J. Political Econom.,81, 637-654.

[9] Black, F., and P. Karasinski (1991). *Bond and option pricing when short rates are Lognormal.* Financial Analyst Journal, July-August, 52-59.

[10] Bollerslev T.(1986). *Generalized autoregressive conditional heteroscedasticity.* Journal of Econometrics 31, 307-327.

[11] Brody D., J.Syroka and M.Zervos (2002) *Dynamical pricing of weather derivatives.* Quantitative Finance, 2: 189-198.

[12] Caballero R. and S. Jewson (2002) *Multivariate Long-Memory Modeling of Daily Surface Air Temperatures and the Valuation of Weather Derivative Portfolios*

[13] Caballero R., S.Jewson and A.Brix (2002) *Long memory in surface air temperature: detection, modelling and application to weather derivative valuation.*

[14] Campbell S. and F.X. Diebold (2002). *Weather forecasting for weather derivatives.* Working Paper 02-046, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania.

[15] Cao M, J. Wei (2000) *Pricing the weather.* Risk Magazine May: 67-70.

[16] Cartea A, M. Figueroa (2005) *Pricing in Electricity Markets: a mean reverting jump diffusion model with seasonality*

[17] Chan K.C., G.A. Karolyi, F.A. Longstaff and A.B. Sanders (1992). *An empirical comparison of alternative models of the short-term interest rate.* The Journal of Finance 47. 1209-1227.

[18] Chen L. (1996). *Interest Rate Dynamics, Derivatives Pricing, and Risk Management.* Springer.

[19] Cox J.C., J.E. Ingersoll Jr and S.A.Ross (1985). *A Theory of the term structure of interest rates.* Econometrica 53, 385-407.

[20] Durham G.B. (2002). *Likelihood-based specification analysis on continuous-time models of the short-term interest rate.*

[21] Escribano, A., J.I. Pena and P. Villaplana (2002). *Modeling electricity prices: international evidence.* working paper, Universidad Carlos III de Madrid.

[22] Eydeland A. and H. Geman (1998). *Some fundamentals of electricity derivatives.*

[23] Franses P.H., J. Neele and D. van Dijk (1998). *Modeling asymmetric volatility in weekly dutch temperature data.* Environmental Modelling and Software 16: 131-137.

[24] Geman, H. (2005). *Commodities and Commodity Derivatives, Modeling and Pricing for Agriculturas, Metals and Energy.* Wiley.

[25] Geman H and A. Roncoroni (2006). *Understand the fine structure of electricity prices.* Journal of Business, vol. 79, no. 6.

[26] Hamisultane, H. (2006), *Pricing the weather derivatives in the presence of long memory in temperatures*

[27] Harvey, A.C. (1989). *Forecasting, Structural Time Series and the Kalman Filter* Cambridge University Press.

[28] Ho, T.S.Y. and S.-B. Lee (1986). *Term structure movements and pricing interest rate contingent claims.* Journal of Finance 41, 1011-1029

[29] Hull J. and A. White, *Pricing interest-rate derivative securities.* The Review of Financial Studies, Vol 3, No. 4 (1990) pp. 573-592

[30] Huisman R. and R. Mahieu(2001) *Regime jumps in electricity prices.* ERIM Report Series Research in Management, August.

[31] de Jong C. (2006). *The nature of power spikes: a regime switch approach.* Studies in Nonlinear Dynamics & Econometrics, vol. 10, issue 3, article 3.

[32] de Jong, C. and R. Huisman (2003). *Option formulas for mean-reverting power prices with spikes.* Energy Power Risk Management (Energy Risk), 7, p.12-16.

[33] Longstaff, F. A. and E. S. Schwartz (1992). *A two-Factor interest ratemodel and contingent claims valuation.* Journal of Fixed Income, 2, 16-23

[34] Lucia J.J. and E. Schwartz (2005). *Electricity prices and power derivatives: evidence from the nordic power exchange.* Review of Derivatives Research 5, 5-50.

[35] Pedersen, A.R. (1995a). *Consistency and asymptic normality of an approximate maximum likelihood estimator for discretely observed diffusion processes.* Bernoulli 1, 257-279.

[36] Pedersen, A.R. (1995b). *A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations.* Scandinavian Jounal of Statistics 22, 55-71.

[37] Samuelson P. (1965). *Rational theory of warrant pricing.* Industrial Management Review 6.

[38] Schoutens, W. (2003). *Lévy Process in Finance: Pricing Financial Derivatives.* Wiley, New York.

[39] Tórro H. (2007). *Forecasting weekly electricity prices at Nordpool.* Intenational Energy Markets, September.

[40] Tórro H., V. Meneu and E. Valor (2001) *Single factor stochastic models with seasonality applied to underlying weather derivatives variables.* Technical Report 60, European Financial Management Association, 2

[41] Vasicek O. (1977). *An equilibrium characterisation of the term structure.* Journal of Financial Economics 5: 177-188.

[42] Villaplana P. (2004) *A two-state variables model for electricity prices.* Third World Congress of the Bachelier Finance Society, Chicago