Pekka Malo

MULTIFRACTALITY IN NORDIC ELECTRICITY MARKETS

HELSINKI SCHOOL OF ECONOMICS
WORKING PAPERS
W-402

Pekka Malo

MULTIFRACTALITY IN NORDIC ELECTRICITY MARKETS

Quantitative Methods in Economics and
Management Science

June
2006

HELSINGIN KAUPPAKORKEAKOULU
HELSINKI SCHOOL OF ECONOMICS
WORKING PAPERS
W-402

HELSINGIN KAUPPAKORKEAKOULU
HELSINKI SCHOOL OF ECONOMICS
PL 1210
FI-00101 HELSINKI
FINLAND

ISSN 1235-5674
(Electronic working paper)
ISBN-10: 952-488-048-2
ISBN-13: 978-952-488-048-0

Helsinki School of Economics -
HSE Print 2006

MULTIFRACTALITY IN NORDIC ELECTRICITY MARKETS

Pekka Malo
Helsinki School of Economics, P.O. Box 1210, FIN-00101
pekka.malo@hse.fi

Abstract. The recent research of turbulent cascades in hydrodynamics has inspired the newly emerged field of econophysics to develop multifractal processes as a competitive alternative to the standard models of continuous time finance. The essential new features of these models are their scale invariance, multiscaling, and clustering in volatility. In this paper we perform multifractal analysis of the NordPool electricity spot prices and compare the results with GARCH line of research. The approach is based on the wavelet transformation of the scaling process. We find that electricity prices are consistent with the multifractal framework and it appears that a simulated Infinitely Divisible Cascading (IDC) process is a reasonably good choice for replicating the empirical scaling properties and the wavelet scalogram of the original signal.

1. Introduction

During the last decade electricity market has experienced a rapid deregulation from a vertically integrated industry into a commodity market, where trading and risk management play key roles. As a result several power exchanges have emerged, where transmission capacity and energy are coupled and traded simultaneously to ensure allocation of capacity according to the bids and offers. The first international power exchange is Nord Pool, which was established 1993 with an intention to improve cost efficiency and competition. Currently Nord Pool spans Norway, Sweden, Finland, and Denmark and over 30 percent of the total Nordic electricity consumption is now exchange traded. At Nord Pool the system price is determined by the intersection of aggregated supply and demand curves, which are constructed from the bids and offers for each hour of the following day. The market is divided into separate bidding areas and Nord Pool manages transmission capacity to conduct power out of low price areas and into high price areas.

However, in addition to improved efficiency, the opening of markets has exposed energy producers and consumers to new forms of market and price risks thus calling for models that are able to capture the dynamics in electricity prices. Given that electricity is expensive and difficult to store at a reasonable cost, this task is particularly challenging and has inspired a variety of different lines of models. The prices are not only seasonal but exhibit levels of volatility and spikes unparalleled by the traditional commodity markets. The fact that production and consumption of electricity must always balance makes the grid vulnerable to transmission failures or generation outages. Combined with inelastic supply and demand, also weather conditions can cause tenfold changes within a single hour. Since Nord Pool is essentially a hydroelectric exchange, the weather is an important factor affecting both demand load and supply. Hydrounits are heavily dependent on availability

of water reservoirs, which vary on a seasonal basis according to precipitation and factors such as snow melting.

As a consequence of the complex characteristics of electricity markets, a considerable number of models ranging from GARCH and jump diffusions to stochastic volatility have been proposed to tackle the problem of modelling the price dynamics. So far, there does not seem to be much concensus on which line of modelling to prefer as they tend to focus on different aspects electricity prices. In the last few years also a completely new family of models has emerged, generally referred to as multifractals, and are currently widely applied in a number of fields including nonlinear physics, geophysics, biology, and computer networking. In economics the concept of multifractal processes was first introduced by Mandelbrot, Fisher and Calvet [28] as an alternative to GARCH and its variants with ability to incorporate long-tailed asset returns and long memory in volatility. Since then multifractal information processes have started to gain popularity as a convenient framework to analyze signals and process that exhibit scaling properties in terms of sample moments. By the term, scaling, we refer to a specific relationship between data samples of different time scales (intra-day, daily, weekly, or monthly).

The idea of multifractal modelling has typically been sold with the following arguments, which have been well recognized in empirical finance [27]: 1) clustering in volatility, 2) compatibility with the martingale property of returns, 3) scale-consistency, and 4) multiscaling. Recently especially scale-consistency and multiscaling have received increasing attention, since quite often the traditional models are not able to provide consistent representations at different time-scales. For example Drost and Nijman [17] have studied this problem for GARCH by aggregating log returns of various processes. The property, that aggregated processes belong to the same class as the original processes, has become known as scale-consistency as it implies that the models have equivalent representations at different time scales. Generally GARCH family models are not always scale consistent. The other property, multiscaling, then again refers to the fact that a price process can not be always described by a single characteristic scale exponent or Hurst exponent. This is a central issue in defining a multifractal process and separates them from Fractional Brownian Motion. Especially when no a priori scale preference is given, the multifractal framework makes a competitive alternative.

The purpose of this paper is to consider an application of the multifractal framework to the case of Nordic electricity markets. First of all we are interested in analysing the scaling behaviour associated with electricity spot prices, that is the strength of multifractality as given by the curvature of the empirical scaling function. The second step is to propose a simple multifractal model for electricity, where the seasonal component is combined with an infinitely divisible cascade which is consistent with the observed scaling function. The third, and perhaps the most interesting issue, is to benchmark the multifractal approach against GARCH family models by comparing their abilities to capture the underlying scaling behaviour of electricity spot prices. Since no well established statistical tests exist, the comparison is done using a simulation based analysis, where we study the properties of synthesized time paths of each model and try to discriminate between different models by considering the obtained scaling exponents and wavelet scalograms.

The main finding of this paper is that electricity prices exhibit multiscaling properties and can be quite successfully modelled using an Infinitely Divisible Scaling

(IDC) process. It appeared that on average the IDC-process was able to replicate the empirical scaling function of spot prices quite closely, whereas GARCH-models had increading deviations in scaling exponents for higher moments. When wavelet scalograms computed for the simulated processes were considered, it seemed that the IDC-process managed to produce similar magnitude patterns as the original signal. Also Skew-t-GARCH did §rather well in this respect. The general impression was that IDC-model can be considered as an interesting alternative for modelling spot price dynamics. However, also this method has its drawbacks. One fundamental problem of multifractal models is that there are no established methods of comparison against other models. Furthermore the estimation of parameters for these models is a delicate issue, since we often find it difficult to tell anything about the signicance of the obtained parameters except in some special cases. Another important point to remember is that multifractal models are mostly simulation tools rather than forecasting models. This of course limits their range of applicability in traditional econometric analysis, yet they can be powerful tools in analysing and predicting risks. Thus these processes become interesting in cases such as electricity price modelling, where no previously established models have found to give satisfactory results.

The paper is organized as follows. In the next section, we define multifractality and explain the concept of infinitely divisible cascading (IDC) random walk as a building block for various multifractal processes [13]. The third section proposes a simple multifractal model for electricity spot prices and considers GARCH-family as benchmark models. The fourth section discusses the estimation tools for the IDC random walk. Since the approach is based on wavelet techniques, we provide a short general introduction to the multiresolution analysis with wavelets. The next section continues by presenting empirical results on estimation of the different models and discusses their ability to capture the scaling behaviour in electricity spot prices using simulation. The sixth section concludes the paper.

2. Multiscaling of asset returns

In order to make the idea of multifractality and scaling behaviour more precise and accessible, we begin this paper by defining the concepts of power law scaling which we then extend to the framework of infinitely divisible scaling. As a next step, we introduce infinitely divisible cascading noise with its extensions to associated cascading motion and finally into random walk. Throughout this introduction we refer to Chainais et al. [10, 12, 13], who have done considerable effort to make these processes accessible for modelling.

2.1. Power law and multiscaling.

Multifractality is a form of generalized scaling, which emphasizes both extreme variations and long-memory. The concept was first introduced by Mandelbrot [25, 26] in the context of turbulent dissipations, which he later extended from measures to stochastic processes [27] by defining the framework of power law scaling.

Definition 2.1 (Power law)**.** We say that a stochastic process $\{X(t)\}$ has power law scaling if it has stationary increments $\delta_\tau X(t) = X(t+\tau) - X(t)$ and satisfies

$$\mathbf{E}(|\delta_\tau X(t)|^q) = C_q(\tau)\tau^{\zeta(q)} \quad as\ \tau \to 0 \tag{2.1}$$

where $C_q(\tau)$ is either assumed to be bounded between positive constants, or to be a more general function depending on the context such that $\liminf_{\tau\to 0} \log_\tau C_q(\tau) = 0$.

Here the scaling function $\zeta(q)$ contains all information about the rate of growth of the multifractal process. If the scaling function is linear and fully determined by its slope, the corresponding process is self-affine or monofractal. This means that the shape of the return distribution should be the same when the time scale is changed. Typical example of such a process is the fractional Brownian motion. If the process has finite second moments with stationary increments, we can use the value of H to describe the autocorrelation structure of the increment sequence. For $H \in (1/2, 1)$, the increments have slowly decaying serial correlations i.e. it exhibits long-range dependence. For $H \in (0, 1/2)$ the increments decay rapidly and sum to zero. In the special case of Wiener process, $H = 1/2$, the increments are serially uncorrelated.

However, in modelling, we can rarely assume that the scaling function is completely characterized by a single exponent. For this purpose, we need multifractal models to allow for more flexible scaling structures. In order to appreciate the geometric intuition one defines the local Hölder exponent as $\alpha(t) = \sup\{\gamma \geq 0 : |X(t + \delta t) - X(t)| = O(|\delta t|^{\gamma}) \; as \; \delta t \to 0\}$, which provides a measure of local regularity of the price path at each time instant t. Whereas the standard Itô diffusions have always local variation proportional to $(\delta t)^{1/2}$, the multifractal processes can generate a variety of local scales filling the gaps between $(\delta t)^{1/2}$ and jump process behaviour $(\delta t)^0$. The higher the Hölder exponent, the smoother the process. For example, the fractional Brownian motion with self-affinity index H is characterized by a unique Hölder exponent $\alpha(t) = H$. Such a process has a modification with Hölder continuous sample paths of order $\gamma \in [0, H)$.

Recently, the multifractal framework has been extended to infinitely divisible cascades (IDC). The concept was originally used to analyse fluid dynamics with a limited range of scales. The idea is to integrate the contribution of all scales in a range of interest:

Definition 2.2 (Infinitely divisible scaling)**.** We say that a stochastic process $\{X(t)\}$ has infinitely divisible scaling if it has stationary increments $\delta_\tau X(t) = X(t + \tau) - X(t)$ and satisfies

$$\mathbf{E}(|\delta_\tau X(t)|^q) = C_q(\tau) \exp\left[-\zeta(q) n(\tau)\right] \quad for \; \tau_{min} \leq \tau \leq \tau_{max}, \tag{2.2}$$

where the function $n(\tau)$ is assumed monotonous and can be interpreted as the depth of the cascade.

This is consistent with power law scaling when we set $n(\tau) = -\ln(\tau)$. An essential difference to the power law scaling definition is, however, to limit the range of scales considered to the ones we can actually observe from the data. This is quite useful, since in practice we cannot observe infinitely small scales as required by 2.1. Another good point to choose this definition is the supply of algorithms for synthesis of infinitely divisible cascades [13]. Therefore, throughout this paper we prefer to work with definition 2.2 when referring to multifractal processes. By the term multiscaling we emphasize the fact that the scaling laws are meant to hold within the limited scales obtained from actual data.

2.2. **Infinitely divisible cascading noise.** Inspired by the ideas of Barral and Mandelbrot [3, 4] infinitely divisible cascading processes were introduced in [12, 11, 39] as an attempt to design process with controllable scaling with their associated random walks. The best known example of a cascading process, which illustrates the

underlying multiplicative construction, is the canonical binomial cascade proposed by Mandelbrot [26]

$$Q_r(t) = \prod_{\{(j,k):1\leq j\leq n, k2^{-j}\leq t<(k+1)2^{-j}\}} W_{j,k},$$

where $W_{j,k}$ denote i.i.d. positive random variables with mean one. Thus from a time-scale view, the binomial cascade is obtained by dividing the time-scale plane $\mathbb{R}\times(0,1]$ using a dyadic grid, which gives the process as the limit of densities $Q_r(t)$ when $r\to 0$. Later on binomial cascades have been improved by replacing the dyadic grid based arrangement of multipliers with compound Poisson cascades (CPC) in order to get rid of non-stationarity and time-shift problems with binomial framework. Now, as discussed by Chainais et al. [13], the idea of infinitely divisible cascades is to generalize the CPC framework of Barral and Mandelbrot [4] one step further.

Let G be an infinitely divisible distribution with moment generating function

$$\tilde{G}(q) = \exp\left[-\rho(q)\right] = \int \exp[qx]dG(x).$$

Typical choices for G range from Normal and Poisson distributions to Stable laws. In this paper we will use lognormal distribution as it leads to a process with a simple parabolic scaling function, which corresponds to our empirical observations.

Let $dm(t,r) = g(r)dtdr$ define a positive measure on the half-plane $S^+ = \{(t,r) : t\in\mathbb{R}, r\in\mathbb{R}^+\}$. Let M denote an infinitely divisible, independently scattered random measure distributed by G in S^+ such that for any Borel set ξ we have $\mathbf{E}[\exp[qM(\xi)]] = \exp[-\rho(q)m(\xi)]$. If subsets ξ_1 and ξ_2 are disjoint their measures $M(\xi_1)$ and $M(\xi_2)$ are independent random variables and $M(\xi_1\cup\xi_2) = M(\xi_1) + M(\xi_2)$. We also define a cone of influence for each time instant as

$$C_r(t) = \{(t',r') : r\leq r'\leq 1, t - r'/2\leq t' < t + r'/2)\}, \tag{2.3}$$

where the choice of large scale equal to 1 is arbitrary as it amounts to choice of time and scale units. Thus by using cones rather than dyadic grids we obtain a considerably more flexible framework, which still allows us to recover the classical cascades by an appropriate choice of control measure m (see for example Muzy and Bacry [32]). It is also interesting to notice that while distribution G controls the structure function through ρ, the control measure m and the shape of the cone $C_r(t)$ determine the speed of the cascade. By choosing $g(r) = 1/r^2$ for $0 < r\leq 1$ and $g(r) = 0$ for $1\leq r$, that is $m(C_r(t)) = -\ln(r)$, we have exact power law behaviour within the given range of scales.

As proposed by Chainais et al. [13] we now define an infinitely divisible cascading noise (IDC-noise) as follows

Definition 2.3 (IDC-noise)**.** An infinitely divisible cascading noise is a family of processes $Q_r(t)$ parametrized by r of the form

$$Q_r(t) = \frac{\exp[M(C_r(t))]}{\mathbf{E}[\exp[M(C_r(t))]]} \tag{2.4}$$

where $C_r(t)$ is the cone of influence given by (2.3).

Assuming time-invariance of the control measure and the cone of influence, the process $\{Q_r(t)\}_{r>0}$ is a positive, left-continuous martingale with log-infinitely divisible distribution. Thus the process converges almost surely as $r\to 0$. Furthermore,

the limit degenerates to zero for almost all t as a consequence of the Law of Large Numbers. This is an essential feature to motivate the definition of IDC-motion in the next section. Another interesting property of the IDC-noise is that the correlation structure of the process has a nice interpretation through the intersections of cones in S^+

$$\mathbf{E}[Q_r(t)Q_r(s)] = \exp[-\varphi(2)m(C_r(t) \cap C_r(s))]$$

where $\varphi(q) = \rho(q) - q\rho(1)$ for all q for which $\rho(q) = -\ln(\tilde{G}(q))$.

2.3. **IDC random walk.** Having defined the fundamental building block, IDC-noise, we are now ready to extend it to an IDC motion and IDC random walk. The problem with using IDC-noise directly is that it is always positive, which makes it necessary to take a couple of additional steps to overcome this restriction. Following Chainais et al. [13] we first define an infinitely divisible cascading motion (IDC-motion) as the limiting integral of an IDC-noise $Q_r(t)$

$$A(t) = \lim_{r \to 0} \int_0^t Q_r(s)ds. \tag{2.5}$$

Denoting $A_r(t) = \int_0^t Q_r(t)dt$, we have an increment process

$$\delta_\tau A_r(t) = A_r(t+\tau) - A_r(t) = \int_t^{t+\tau} Q_r(s)ds,$$

which inherits its stationarity directly from from Q_r as insured by time invariance of both the control measure m and the shape of the cone C_r. The obtained cascading motion A is a non-decreasing cadlag-process with stationary increments.

This gives us the following definition of a multifractal random walk in IDC time [13]

Definition 2.4 (IDC Random Walk)**.** Let A be an infinitely divisible cascading motion, and B_H the fractional Brownian motion with Hölder exponent H. The compound process

$$V_H(t) = B_H(A(t)), \quad t \in \mathbb{R}^+ \tag{2.6}$$

is an Infinitely Divisible Cascading Random Walk.

The idea behind this definition is essentially the same as in the MMAR model proposed by Mandelbrot, Calvet and Fisher [28], where they combine the concept of trading time with a multifractal measure. In this case we could interpret the IDC-motion A as a stochastic trading time process, which adjusts time according to different levels of trading activity. Thus the trading time will be highly variable and potentially of long memory. Since the fractional Brownian motion is a self-affine process with a constant Hölder exponent, we find that the scaling properties of the IDC-motion are passed on to the compound process. From $A(0) = 0$, $V_H(0) = 0$, and the fact that A is nondecreasing with $\mathbf{E}[\delta_\tau A^q] = C_q(\tau)\tau^q \exp[-\varphi(q)m(C_\tau)]$, we find that V_H must satisfy

$$\mathbf{E}|\delta_\tau V_H|^q = C'_q(\tau)\tau^{qH} \exp[-\varphi(q)m(C_\tau)] \tag{2.7}$$

where $C'_q(\tau)$ is bounded and close to constant for $\tau << 1$.

The significance of compounding is simply to allow direct modelling of process variablity without affecting the direction of increments or their correlations. We also assume that the trading time is independent from the fractional Brownian motion.

3. Models

Once we have introduced the concept of infinitely divisible scaling and multifractality, we are now ready to propose a simple multifractal model for electricity spot prices along with benchmark models. However, before doing so, we begin this section with a brief summary of the stylized facts of Nordic power markets. For a more detailed discussion see for example Weron [41].

3.1. **Stylized facts.** In order to have an empirically plausible model, we need to take into account the following key features of electricity prices:

(1) Price spikes: One of the most salient features of the power markets are the unanticipated extreme changes in spot prices caused by severe weather conditions, generation outages, and transmission failures. Although the spikes are usually short lived they account for a large part of the observed total price variation and as such are a key motivator for designing hedging strategies. The spikes are particularly powerful during high demand periods such as winter in Scandinavia, when electricity consumption is heavy due to excess heating.
(2) Seasonality: Electricity is highly dependent on weather conditions, consumption pattern and economic activities. In Scandinavia the weather effects are a major factor, because a considerable amount of energy is generated by hydrounits which are strongly dependent on fluctuations in water reservoirs. Combined with inelasticities in demand and supply, the weather effects lead to highly volatile electricity spot prices. For more discussion on deterministic patterns see for example Lucia and Schwartz [24] and Bhanot [5].
(3) Mean-reversion: Generally electricity prices have been found to exhibit considerable mean-reversion. As such electricity is perhaps one of the best known examples of anti-persistent data. Various explanations for mean reversion have been proposed. For example Knittel and Roberts [23], who consider this as a consequence of weather cycles and the tendency of weather to revert to its mean level (possibly time-varying).

3.2. **IDC model for electricity spot price.** Based on the above criteria, we formulate the following simple multifractal as starting point.

$$X(t) = f(t) + V_H(t), \quad f(t) = \sum_{i=1}^{6} \beta_i x_i(t) \tag{3.1}$$

Here $X(t) = \ln P(t)$ denotes the logarithm of the spot price and $V_H(t)$ is an IDC random walk with Hölder exponent H. To define a simple seasonal component $f(t)$ we can choose for example $x_1(t) = 1$, $x_2(t) = t$, $x_3(t) = \sin(\lambda t)$, $x_4(t) = \cos(\lambda t)$, $x_5(t) = \sin(2\lambda t)$, and $x_6(t) = \cos(2\lambda t)$(see e.g. Carnero [9]).

When proposing this model our main interest is to capture the wildly random component of the electricity price. By incorporating $V_H(t)$ we can generate continuous anti-persistent price paths, which are spiky enough to resemble a jump diffusion. The specification of the seasonal part is of course important for forecasting purposes, but it is not that much of an issue as various well established techniques for modelling seasonal behaviour exist. Furthermore, the estimation of the IDC model is not sensitive to different specifications of the seasonal component,

since the wavelet techniques are quite robust against various trends. By choosing the number of vanishing moments high enough, when applying the wavelets, we can eliminate the effect of complex sinusoidal trends. The issues concerning the estimation methodology are discussed more closely in the next section.

3.3. **Traditional modelling.** In order to get a few comparisons with traditional models we consider standard ARMAX-GARCH models, which have become common tools in analyzing financial time series. In addition to the standard GARCH model introduced by Bollerslev [6], we include GJR-GARCH of Glosten et al. [20], and Skew-t-GARCH with Hansen [21] density function. The GARCH models are estimated using maximum likelihood.

The estimation of the models is carried out in two steps. Given the seasonally detrended time series $X(t) - f(t)$, we first use ARMA-filter to remove the observed autocorrelation structure and then calibrate the GARCH models to the residuals. Let $r(t) = d(X(t) - f(t))$ denote differences of the deseasonalized time series, the general model reads as follows

$$r(t) = \mu + \sum_{i=1}^{m} \phi_i r(t-i) + \sum_{j=1}^{n} \theta_j \epsilon(t-j) + \epsilon(t) \tag{3.2}$$

$$\epsilon(t) = z(t)\sigma(t), \quad z(t) \sim iid(0,1) \tag{3.3}$$

The different specifications for $\sigma(t)$ are listed below:

$$GARCH: \quad \sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

$$GJR-GARCH: \quad \sigma_t^2 = \omega + \sum_{i=1}^{p} (\alpha_i + \gamma_i I_{\{\epsilon_{t-i}<0\}}) \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

For Skew-t-GARCH the innovation density function, f, is the skewed t distribution defined as [21]

$$f(z|\nu,\lambda) = \begin{cases} bc\left(1 + \frac{1}{\nu-2}\left(\frac{bz+a}{1-\lambda}^2\right)\right)^{-\frac{\nu+1}{2}}, & z < -\frac{a}{b} \\ bc\left(1 + \frac{1}{\nu-2}\left(\frac{bz+a}{1+\lambda}^2\right)\right)^{-\frac{\nu+1}{2}}, & z \geq -\frac{a}{b} \end{cases} \tag{3.4}$$

where $2 < \nu < \infty$ is the degrees of freedom and $-1 < \lambda < 1$ is the skewness parameter. The coefficients a, b and c are given by

$$a = 4\lambda c\left(\frac{\nu-2}{\nu-1}\right), \quad b^2 = 1 + 3\lambda^2 - a^2,$$

$$c = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi(\nu-2)}\Gamma(\frac{\nu}{2})}$$

4. Wavelets and estimation

Given the preceeding definition of an IDC random walk, an interesting question is how to calibrate models involving such components to data. For this purpose, we need to estimate the Hölder exponent driving the fractional Brownian motion and parametrize a random multifractal measure such that it is consistent with the empirical scaling function estimated from the data.

So far, there is no unified estimation framework, but wavelets have become an increasingly popular choice due to the structural affinity which exists between the mathematical framework they offer and the physical nature of fractal processes. Wavelets are also quite robust against various trends and seasonalities found in data, which is particularly useful in our case as the Nord pool electricity spots are well known for strong seasonal patterns. In this section, we introduce the theory of multiresolution analysis and wavelet representation as a tool for efficient analysis of scaling processes. Then we consider the practical issues in detection of scaling, interpreting the results and estimation of the scaling parameters.

4.1. **Review on wavelets.** Although wavelet theory was originally established for deterministic finite energy processes, it has proven to be a useful tool for stochastic processes as well [8, 29]. The term wavelets refers to a set of basis functions with a very special structure, which allows a convenient series expansion of a function in terms of the generating wavelets. The idea is related to Fourier analysis, but with a notable distinction: in Fourier analysis the time evolution of the frequencies is not directly reflected, whereas wavelets are well designed for time-dependent frequency analysis.

In order to understand the idea of wavelet representation we introduce the concept of multiresolution analysis (MRA), which can be used to generate an orthonormal basis for $L^2(\mathbf{R})$. A multiresolution analysis consists of a collection of nested closed linear subspaces $\{V_j : j \in \mathbf{Z}\} \subset L^2(\mathbf{R})$ and a scaling function ϕ_0 such that the following conditions are satisfied

(1) $V_{j-1} \supset V_j$, $\cap V_j = \{0\}$, and $\cup V_j$ is dense in $L^2(\mathbf{R})$,
(2) $X(t) \in V_j \Leftrightarrow X(2^j t) \in V_0$
(3) A function $\phi \in V_0$ such that $\{\phi(t-k) : k \in \mathbf{Z}\}$ is an orthonormal basis for V_0.

Thus by definition $\{\phi_{j,k}(t) = 2^{-j/2}\phi(2^{-j}t - k) : k \in \mathbf{Z}\}$ forms an orthonormal basis for V_j. Since $V_j \subset V_{j-1}$, we can now define $W_j = V_j \ominus V_{j-1}$. Then assuming that the scaling function ϕ_0 (father wavelet), is chosen well enough, there exists a function $\psi \in W_0$ (mother wavelet) such that $\{\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k) : k \in \mathbf{Z}\}$ is an orthonormal basis for W_j. Here, the scaling coefficients $2^{-j/2}$ are needed for normalization of $L^2(\mathbf{R})$ norm $\|\psi_{j,k}\|_2 = 1$. The mother wavelet should also satisfy $\int \psi(t)dt = 0$ to guarantee regularity and have N vanishing moments $\int t^k \psi(t)dt = 0$ for $k = 1, .., N$.

Now the idea of performing multiresolution analysis is to successively project $X(t)$ onto each of the subspaces V_j defined by $Proj_{V_j}(X(t)) = \sum_k a_X(j,k)\phi_{j,k}(t)$. Since the projection at level j is coarser than projection at level $j-1$, we have a loss of information at each stage equal to $detail_j(t) = Proj_{V_{j-1}}(X(t)) - Proj_{V_j}(X(t))$. These are equivalent to considering projections onto W_j in terms of mother wavelets: $detail_j(t) = Proj_{W_j}(X(t)) = \sum_k d_X(j,k)\psi_{j,k}(t)$. This means that we can rewrite the fine scale approximation for X(t) as a collection of details at different resolutions together with a final low-resolution approximation which belongs to V_J

$$X(t) = \sum_{k\in\mathbf{Z}} a_X(J,k)\phi_{J,k}(t) + \sum_{j=1}^{J}\sum_{k\in\mathbf{Z}} d_X(j,k)\psi_{j,k}(t) \tag{4.1}$$

where the coefficients are given by $L^2(\mathbf{R})$ inner products $a_x(J,k) = \langle x, \phi_{J,k}\rangle$ and $d_x(j,k) = \langle x, \psi_{j,k}\rangle$. The coefficients $d_x(j,k)$ are often referred to as a subsample

of the continuous wavelet transform $\{T_X(a,t) = \langle X, \psi_{a,t}\rangle, a \in \mathbf{R}^+, t \in \mathbf{R}\}$ located on a dyadic grid. In this paper we use a fast pyramidal algorithm to compute the coefficients, which has lower computational cost than the FFT [15]. For more details on this approach see e.g [8, 29]. Although the above presentation is essentially designed for functions in L^2, it has been well documented that the approach can be applied to stochastic processes when the wavelets satisfy some mild regularity conditions [8, 29].

4.2. **Estimation of scaling exponents and detection of multifractality.** In this part we explain how wavelets can be used to detect potential multifractality. The actual multifractal model and its estimation procedure is presented in next subsection.

The connection between fractal processes and wavelets arises from the fact that the increments involved in the study of the local regularity of a sample path can be seen as simple examples of wavelet coefficients [30]. In order to describe the idea, we consider first a few fundamental properties of the second order processes and then discuss how the approach can be extended to analyze the multifractal processes (qth order processes). One major reason for this is that we will use these techniques in preliminary analysis before considering multifractal models.

Now, let X be either a self-similar process with stationary increments with scaling parameter H (H-sssi for short), or a long-range dependence (LRD) process, or a second-order stationary 1/f-type process or a fractal process, then the wavelet coefficients will, exhibit following properties which are fundamental for construction of the estimators and analysis of scaling [2].

P1: The details $\{d_X(j,k), k \in \mathbf{Z}\}$ is a stationary process if $N \geq (\alpha-1)/2$ and the variance of the $d_X(j,k)$ reproduces the underlying scaling behavior of data within a given range of octaves $j_1 \leq j \leq j_2$:

$$\mathbf{E}\left[d_X(j,k)^2\right] = 2^{j\alpha} c_f C(\alpha,\psi)$$

Here N is the number of vanishing moments and the definitions of α, c_f and $C(\alpha,\psi)$ depend on the underlying process. For example in the case of a H-sssi process, we have $\alpha = 2H+1$, $C(\alpha,\psi)$ is to be identified from $\mathbf{E}\left[d_X(j,k)^2\right] = 2^{j(2H+1)} \int |t|^{2H} (\int \psi(u)\psi(u-t)du)dt$, and $j_1 = -\infty$ and $j_2 = +\infty$. Here it is worthwhile to note that the results extend for the multifractal processes as well [2].

P2: The details $\{d_X(j,k), k \in \mathbf{Z}\}$ are short-range dependent on condition that $N \geq \alpha/2$. Furthermore the number of vanishing moments can be used to control the correlation: $\mathbf{E}\left[d_X(j,k)d_X(j',k'\right] \approx |k-k'|^{\alpha-1-2N}$, $|k-k'| \to +\infty$.

Given these properties we find that the effectiveness of the wavelet approach stems from the fact that the wavelet basis of functions possesses a scaling property and thereby constitutes an optimal co-ordinate system to study such phenomena [34]. The starting point for analysis is obtained from (P1) as it implies a regression approach for estimating α and c_f (Logscale diagram): $\log_2(\mathbf{E}d_x(j,\cdot)^2) = j\alpha + \log_2(c_f C)$. The approach is feasible, since any kind of linear regression constitutes an unbiased estimator due to Property 1. The lack of bias does not require knowledge of their variances or distriubtions [2]. However, it can be shown that the weighted regression is preferable as it is the minimum variance unbiased estimator. This semiparametric wavelet estimator reads, see e.g. [1, 38, 2, 40]

$$\hat{\alpha} = \frac{\sum y_j(jS - S_j)/\sigma_j^2}{SS_{jj} - S_j^2} = \sum w_j y_j \tag{4.2}$$

where $y_j = \log_2(1/n_j \sum_k d_X(j,k)^2) - g_j$ with correction term $g_j = \Psi(n_j/2)[\ln(2)]^{-1} - \log_2(n_j/2)$ to take into account the change of order in taking expectations and logarithms. Here Ψ is the digamma function and n_j is the number of coefficients at octave $j \in [j_1, j_2]$. The other terms are defined by

$$S = \sum_{j=j_1}^{j_2} \sigma_j^{-2}, \quad S_j = \sum_{j=j_1}^{j_2} j\sigma_j^{-2}, \quad S_{jj} = \sum_{j=j_1}^{j_2} j^2\sigma_j^{-2}, \quad \sigma_j^2 = \frac{\zeta(2, n_j/2)}{\ln^2(2)}$$

where $\zeta(2,z) = \sum_{n=0}^{\infty} 1/(z+n)^2$ is Riemann's zeta function.

As discussed by Chainais et al. [10, 2], the above procedure can be generalized for the multifractal formalism, where $\mathbf{E}|d_X(j,\cdot)|^q \sim 2^{j\zeta(q)}$. That is, for the case where relevant information for the analysis of scaling is beyond the reach of second order statistics. The estimation of the $\zeta(q)$ is achieved via

$$S_q(j) = \frac{1}{n_j} \sum_{k=1}^{n_j} |d_X(j,k)|^q \tag{4.3}$$

where n_j is the number of coefficients at octave j. Here we can use the generalized logscale diagrams ($\log_2(S_q(j))$ vs. $\log_2(2^j) = j$) to check for straight lines and estimate $\zeta(q)$ through linear regression as the slope in those diagrams. Confidence intervals are obtained for each $\log_2(S_q(j))$. The traditional Hurst or Hölder exponent is recovered by considering the special second order case. Finally, an estimate of the multifractal spectrum is given by the Legendre transform of the estimated $\{\zeta(q)\}$:

$$f(\alpha) = \inf_q \{q\alpha - \zeta(q) + 1\}. \tag{4.4}$$

An interesting question when studying the forms of $\zeta(q)$ is to check whether it takes a simple linear form $\zeta(q) = qH$ or not. If $\zeta(q)$ is linear the process is essentially a monofractal with a degenerate multifractal spectrum. A convenient tool for this analysis has been designed by Abry et al. [2] referred to as multiscale diagram, which will be discussed more in the results section.

4.3. **Parametrizing IDC-noise.** In the case of electricity markets, we find that the empirical scaling function obtained using the above procedure is well approximated by a second order polynomial (see section 5). Thus, the choice of a random measure with log-normal distribution $N(\mu, \sigma^2)$ appears natural since it leads to simple polynomial forms for φ. Using notation presented in section 2.2 we have $\rho(q) = -\mu q - \sigma^2 q/2$. Then condition $\varphi(1) = 0$ requires that $\mu = -\sigma^2/2$ which gives us

$$\varphi(q) = \frac{\sigma^2}{2} q(1-q). \tag{4.5}$$

By setting $m(C_\tau) = -\ln(\tau)$ in (2.7), we get $\mathbf{E}|\delta_\tau V_H|^q = C_q'(\tau)\tau^{qH(1+\sigma^2/2) - \sigma^2 q^2 H^2/2}$, which implies that the theoretical scaling function of V_H should the polynomial given by

$$\zeta_{V_H}(q) = qH(1 + \sigma^2/2) - \sigma^2 q^2 H^2/2 \tag{4.6}$$

Thus we can recover the parameter σ^2 by matching the second derivatives of the empirical and theoretical scaling functions. The interpretation of σ^2 is then associated with the curvature of the scaling function or the strength of multifractality.

5. Experiments

The Nordpool data is extracted from Datastream International and consists of system reference spot prices covering the period between 10-Mar-1998 and 12-Jan-2006. A visual inspection of the price path (Figure 1) confirms the extremal behaviour with jumps and strong seasonal fluctuations. We find that several large price spikes have occured during the sample range, and most of them have been in January or February. Typically the spikes are short-lived and the overall price path is driven by a sinusoidal pattern. In the nordic region demand for electricity is high in winter due to heating and light, while prices decline in summer. Also seasonal changes in water reservoirs add their effect on electricity generation. These factors are particularly important when designing models for prediction purposes. However, given that our interest is more on the spiky behaviour and extreme volatility, we have decided to take a rather crude approach to remove the seasonal component by approximating it with a sinusoidal trend as defined in 3.1. The preprocessed data is shown in Figure 2 along with the trend component. Descriptive statistics for both original and preprocessed data are reported in Table 1.

5.1. **Detecting multifractality.** Given the peculiar nature of electricity as a commodity it is not too surprising to find that it appears to exhibit strong multifractal characteristics. The evaluation was done using multiscale diagrams proposed by Abry et al. [1]. The idea in this approach is to plot the empirical scaling function $\hat{\zeta}(q) = \hat{\alpha}_q - q/2$ against q, together with confidence intervals, in order to see whether it deviates significantly from a simple linear function over a range of moments. Thus a lack of alignment in the multiscale diagram strongly suggests scaling behaviour, which cannot be explained by a linear scaling function $\zeta(q) = qH$. For curiosity we present also the linear multiscale diagram with $h_q = \hat{\alpha}_q/q - 1/2$, where linear forms appear as horizontal alignments. This is essentially the same as the actual multiscale diagram, but visually more convenient.

Figure 3. shows the multiscale diagram for detrended electricity prices, where it is confirmed that they do not obey any linear scaling function. Since no horizontal regions are found in linear multiscale diagram, the results indicate that a multifractal model might be appropriate. However, the test does not tell anything about the model formulation, and therefore we should be careful when drawing these conclusions. In this paper we have considered an IDC-model, but there is a variety of alternative scaling models which are left as an issue for future research. When running the estimates, q ranged from -1 to 7, the number of vanishing moments was set to 4, and L^2 normalisation was used.

5.2. **IDC-model calibration.** Now we can take the first steps to calibrate an IDC-model. Using the quadratic approximation $\hat{\alpha}_q = -0.052q^2 + 0.92q - 0.11$ and $\hat{\zeta}(q) = \hat{\alpha}_q - q/2$, we can match the second derivative of empirical scaling function $\hat{\zeta}''(q) = -0.104$ with the theoretical curvature of (4.6) to obtain $\sigma^2 = 0.104/H^2$. The Hölder exponent can be computed from the second order statistic α_2 by $H = (\alpha_2 - 1)/2 = 0.3380$, which gives $\sigma^2 = 0.9103$. One sample path generated with these parameters is shown in Figure 4. The finding that Hölder or Hurst exponent is

well below 0.5 is consistent with the stylized fact of fast mean-reversion in electricity prices.

5.3. **Comparisons.** As benchmark models we have considered ARMA-GARCH(1,1), ARMA-GJR-GARCH(1,1), and ARMA-Skew-t-GARCH(1,1) with Hansen [21] skew-t density. The parameter estimates for these models are presented in Table 3. Although all of the parameters were not significant, we decided to leave them in the models as they appeared to improve the stability of the models.

Given the characteristics of the IDC-model and the lack of tests designed to compare multifractals with the classical models, we decided to take a simulation approach and produced 1000 series with 2000 observations for each process in order to compare their empirical scaling properties. Table 2 presents the identified scaling exponents α_q for a range of moments between 0 and 7 including also some of the fractional moments. The mean scaling exponents are calculated by averaging those scaling exponents of the simulated series for each model.

Based on these results we make three interesting notions. First of all, we find that IDC-process appears to outperform the rest in its ability to trace out the empirical scaling function as it follows the observed parabolic shape most convincingly. This becomes most visible when considering the range of moments above 4. The second observation concerns the differences between the GARCH models. On average GJR-GARCH yields better results than the standard GARCH, although they are quite close. However, switching the error density from Gaussian to Hansen's Skew-t appears to have a considerable effect on scaling properties. The fact that the scaling exponent estimates for higher moments have surprisingly large deviations from the empirical exponents is perhaps best explained by having only 4.1 as the optimal degrees of freedom for the density function. The third observation concerns the large variation in the simulated scaling exponents as can be observed from the min-max ranges given in Table 2. The variation was largest for the Skew-t-GARCH and IDC-process, whereas GARCH and GJR-GARCH appeared to have considerably smaller bounds. In all of the cases variation increased as we moved toward higher moments. These are important points to keep in mind when trying to interpret the results. Although we can discriminate between different models on average and thereby claim that an IDC-process seems to the most appropriate choice for capturing the empirical scaling properties, it remains a delicate issue to be able to make a difference between genuine and apparent multiscaling. This is quite problematic when trying to find reasonable criteria to separate multifractals from other processes, but for the moment there are no clear solutions for this issue.

In addition to the comparision of scaling exponents, we have considered the ability of the different models to replicate the empirical scalogram of detrended spot prices. The results are presented in Figure 5. This idea is based on the continuous wavelet multiresolution analysis (MRA), which gives us a graph with time on the horizontal axis and different scales on the vertical axis. In a scalogram, the absolute values of the localized wavelet resonance coefficients generated by the transform at different scales and time are represented by a colour bar ranging from blue (minimum) to red (maximum). Here the number of scales is 1-64. The interpretation is that the scalogram traces the second moment, variance, or energy of the spot price series over the given time period. The connection is easily verified by considering the variation of price paths in Figures 1 and 2 with 5(a).

The graphs (b)-(e) in Figure 5 show the scalograms for typical simulated path computed for each model. The main conclusion is that the IDC-process and Skew-t-GARCH generated the most compatible scalograms with the original signal, whereas GARCH and GJR-GARCH fail to generate the green-reddish patterns observed between scales 40-60. It could be that Skew-t is perhaps too strong when considering the colour patterns at scales 15-30, where we find more light blue-green stripes than in the original signal. But in general the results are quite expected, and it seems that IDC-model makes a reasonable alternative for modelling spot price variations.

6. Conclusions

Given the stylized facts of electricity price spikes, seasonality, and anti-persistence combined with empirically observed scaling behaviour, the task of modelling electricity price dynamics has posed a considerable challenge. In this paper we have considered a novel approach for capturing these properties using Infinitely Divisible Cascading (IDC) processes. The origin of these processes resides mostly in nonlinear physics and hydrodynamics, but they have proven to be useful in various applications. Recently the multiscaling models have emerged as new risk management tools in economics and motivation for their usage stems from the ability to capture clustering effects and long memory in volatility, yet still remain scale-consistent and compatible with the martingale property of returns.

However, the approach does not come without additional costs. The main disadvantage of the IDC-models is the lack statistical methods to test the robustness of the specification and compare it with classical lines of modelling. Therefore we considered a set of simulation based tests to compare the IDC-model with different GARCH-models in terms of their ability to capture the empirically observed scaling properties along with their ability to replicate the scalogram of electricity spot prices. The overall impression based on these simple tests was that an IDC-model appears to be a reasonably consistent alternative for modelling electricity spot prices as it did rather good job in reproducing the empirical scaling behaviour and variance patterns. However, we recognize that the IDC-process does not exhaust the set of possible scaling models and acknowledge the limitations of simulation analysis when no robust statistical methods are available.

In general, when analysing empirical scaling functions, it is a rather delicate issue to say anything about the reasons for departure from linearity as it appears to be notoriously difficult to claim whether it is genuine multifractality or other nonstationarities which lead to apparent multiscaling. For example Bouchaud et al. [7] have shown that it is possible to design stochastic volatility models with slow crossover effects which lead to multiscaling, although the processes themselves are asymptotically monofractal. Thus although we will leave apparent multiscaling as an issue for further research, it is good to keep in mind that the multiscaling analysis is more likely a tool for pointing out behaviour which is important to take into account without saying much about the applicable models. In our case the IDC-model was quite successful in producing price paths with close resemblance to the original signal, yet we feel that there is a considerable need for more accurate tests and criteria to be able to distinguish between apparent and genuine scaling behaviour. Another issue, which calls for development is the tools to evaluate and test the significance of parameter estimates in multifractal models.

References

[1] P. Abry, D. Veitch and P. Flandrin, Long-range dependence: revisiting aggregation with wavelets, *Journal of Time Series Analysis*, Vol. 19, No. 3, 1998, pp. 253-266.

[2] P. Abry, P. Flandrin, M. S. Taqqu and D. Veitch, *Wavelets for the Analysis, Estimation, and Synthesis of Scaling Data Self-Similar Network Traffic and Performance Evaluation*, K. Park and W. Willinger eds., New York: John Wiley & Sons, 2000.

[3] J. Barral, and B. Mandelbrot, Multiplicative products of cylindrical pulses, Cowles Foundation, discussion paper 1287, 1999.

[4] J. Barral, and B. Mandelbrot, Multiplicative products of cylindrical pulses, *Probability Theory Relat. Fields*, Vol. 124, 2002, pp. 409-430.

[5] K. Bhanot, Behavior of Power Prices: Implications for the Valuation and Hedging of Financial Contracts, *Journal of Risk*, Vol. 2, 2000, pp. 43-62.

[6] T. Bollerslev, Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, Vol. 31, 1986, pp. 307-327.

[7] J. Bouchoud, M. Potters, and M. Meyer, Apparent multifractality in financial time series, *The European Physical Journal B*, Vol. 13, 2000, pp. 595-599.

[8] S. Cambanis and C. Houdré, On the continuous wavelet transform of second-order random processes. In *IEEE Trans. on Info. Theory*, Vol. 41, No. 3, 1995, pp. 628-642.

[9] M. Carnero, S. Koopman, and M. Ooms, Periodic Heteroskedastic RegARFIMA Models for Daily Electricity Spot Prices, Tinbergen Institute Discussion Paper, 2003.

[10] P. Chainais, P. Abry, and D. Veitch, Multifractal analysis and α-stable processes: a methodological contribution, *ICASSP-2000*, Istanbul, 2000.

[11] P. Chainais, R. Riedi, and P. Abry, Compound poisson cascades, *Colloque Autosimilarité et Applications*, Clermont-Ferrand, France, 2002.

[12] P. Chainais, R. Riedi, and P. Abry, Scale invariant infinitely divisible cascades, *Int. Symp. on Physics in Signal and Image Processing.* Grenoble, France, 2003.

[13] P. Chainais, R. Riedi, and P. Abry, On non scale invariant infinitely divisible cascades, *IEEE Transactions on Information Theory*, Vol. 51, No. 3, 2005.

[14] P. Chainais, R. Riedi, and P. Abry, Warped infinitely divisible cascades: beyond power laws, *Traitement du Signal*, Vol. 22, No. 1, 2005.

[15] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, 1992.

[16] Z. Ding, C. Granger, R. Engle, A long memory property of stock market returns and a new model, *Journal of Empirical Finance*, Vol. 1, 1993, pp. 83-106.

[17] F. Drost, and T. Nijman, Temporal Aggregation of GARCH Processes, *Econometrica*, Vol. 61, 1993, pp. 83-106.

[18] R. Engle, and V. Ng, Measuring and testing the impact of news on volatility, *Journal of Finance*, Vol. 48, 1993, pp. 1749-1778.

[19] A. Fisher, L. Calvet, and B. Mandelbrot, Multifractality of Deutschemark / US Dollar Exchange rates, working paper, 1997.

[20] L. Glosten, R. Jagannathan, and D. Runkle, On the relation between the expected value and the volatility of the nominal excess returns on stocks, *Journal of Finance*, Vol. 48, 1993, pp. 1779-1801.

[21] B. Hansen, Autoregressive conditional density estimation, *International Economic Review*, Vol. 35, pp. 705-730.

[22] M. Higgins, and A. Bera, A class of nonlinear ARCH models, *International Economic Review*, Vol. 33, 1992, pp. 137-158.

[23] C. Knittel, and M. Roberts, An Empirical Examination of Deregulated Electricity Prices, working paper, University of California Energy Institute, 2001.

[24] J. Lucia, and E. Schwartz, Electricity Prices and Power Derivatives: Evidence from the Nordic Power Exchange, working paper, 2000.

[25] B. B. Mandelbrot, Possible Refinements of the Lognormal Hypothesis Concerning the Distribution of Energy dissipation in Intermitent Turbulence, In M. Rosenblatt and C. Van Atta eds., *Statistical Models and Turbulence*, New York: Springer Verlag, 1972.

[26] B. B. Mandelbrot, Intermittent Turbulence in Self Similar Cascades; Divergence of High Moments and Dimension of the Carrier, *Journal of Fluid Mechanics*, Vol. 62, 1974.

[27] B. B. Mandelbrot, *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk*, New York: Springer Verlag, 1997.

[28] B. B. Mandelbrot, L. Calvet, A. Fisher, Large Deviation Theory and the Multifractal Spectrum of Financial Prices, working paper, 1997.
[29] E. Masry, The Wavelet transform of stochastic processes with stationary increments and its application to fractional Brownian motion, In *IEEE Trans. on Info. Theory*, Vol. 39, No. 1, 1993, pp. 260-264.
[30] J. F. Muzy, E. Bacry, and A. Arnéodo, The multifractal formalism revisited with wavelets, *Int. J. of Bifurc. and Chaos*, Vol. 4, No. 2, 1994, pp. 245-301.
[31] J. F. Muzy, J. Delour, and E. Bacry, Modelling fluctuations of financial time series: from cascade process to stochastic volatility model, *The European Physical Journal B*, Vol. 17, 2000, pp. 437-548.
[32] J. F. Muzy, and E. Bacry, Multifractal stationary random measures and multifractal random walks with log-infinitely divisible scaling laws, working paper, 2002.
[33] D. Nelson, Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica*, Vol. 59, No. 2, 1991, pp. 347-370.
[34] M. Roughan, D. Veitch, P. Abry, On-line estimation of the parameters of long-range dependence, *Proceeding Globecom*, Vol. 6, Sydney 1998, pp.3716-3721.
[35] G. Schwert, Why does stock market volatility change over time?, Journal of Finance, Vol. 44, 1989, pp. 1115-1153.
[36] P. van der Sluis, EmmPack 1.01: C/C++ Code for Use with Ox for Estimation of Univariate Stochastic Volatility Models with the Efficient Method of Moments, Tinbergen Institute Discussion Papers, Tinbergen Institute, 1997.
[37] S. Taylor, *Modelling Financial Time Series*, John Wiley and Sons, 1986.
[38] D. Veitch, P. Abry, A wavelet based joint estimator for the parameters of LRD, *Special issue on Multiscale Statistical Signal Analysis and its Applications, IEEE Trans. Info. Th.* Vol. 45, No. 3, 1999.
[39] D. Veitch, P. Abry, P. Flandrin, P. Chainais, Infinitely divisble cascade analysis of network traffic data, Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing - ICASSP, 2000.
[40] D. Veitch, P. Abry, M. S. Taqqu, On the Automatic Selection of the Onset of Scaling, working paper, 2002.
[41] R. Weron, Heavy tails and electricity prices, conference paper, The Deutsche Bundesbanks 2005 Annual Fall Conference, 2005.
[42] J. Zakoian, Threshold heteroskedastic models, *Journal of Economic Dynamics and Control*, Vol. 18, 1994, pp. 931-955.

FIGURE 1. NordPool spot price (1998-2006)

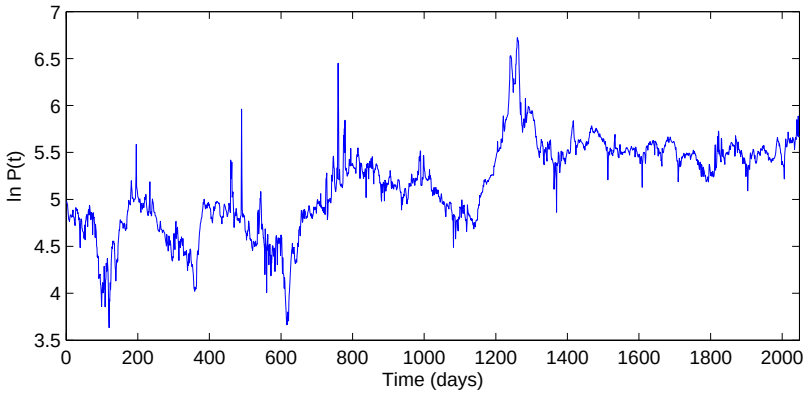


TABLE 1. Descriptive statistics

	Spot		Detrended	
	Levels	Returns	Levels	Returns
Mean	5.151	0.000	0.000	0.000
Standard deviation	0.467	0.085	0.306	0.085
Skewness	-0.238	1.145	0.227	1.141
Excess kurtosis	0.157	29.554	2.284	29.524
$Q(4)$	-	**76.949**	-	**77.685**
$Q^2(4)$	-	**280.566**	-	**283.209**
$Q(12)$	-	**97.500**	-	**99.131**
$Q^2(12)$	-	**280.864**	-	**283.516**
KPSS (with trend)	**0.676**	0.018	**0.792**	0.019

FIGURE 2. Detrended process and seasonal trend

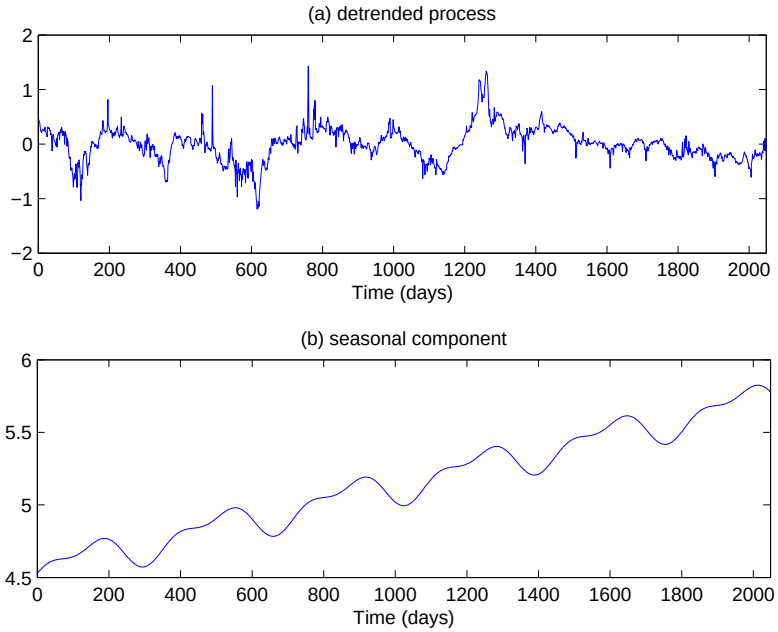


FIGURE 3. Multiscale diagram

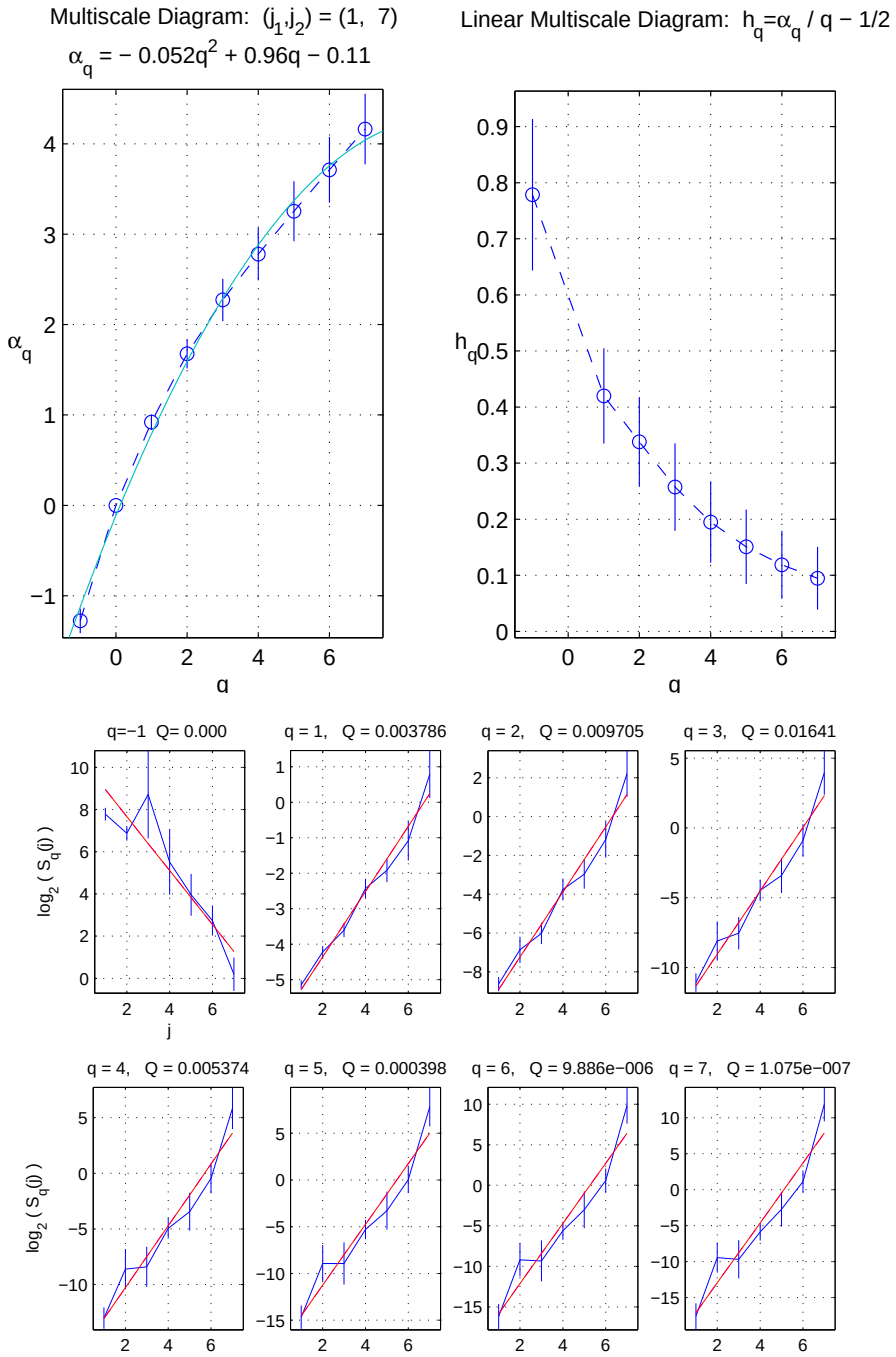


FIGURE 4. IDC-process sample path

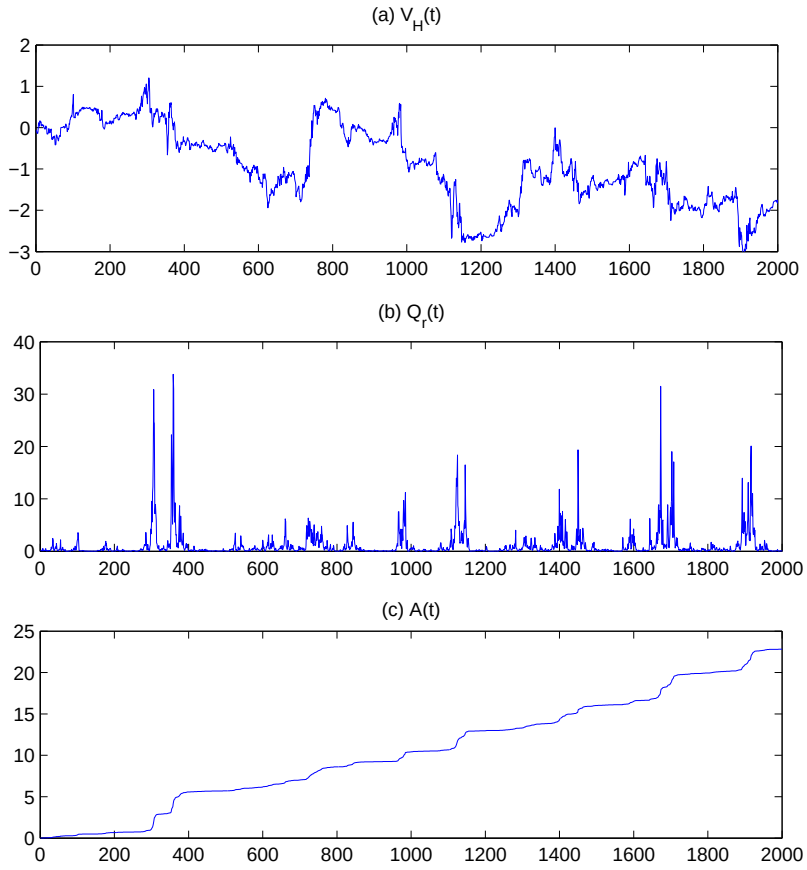
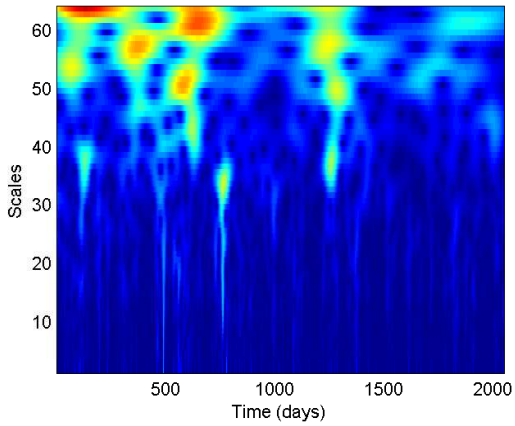


TABLE 2. Comparison between scaling exponents of empirical detrended process and mean scaling exponents of corresponding simulated processes. The first left column provides the exponents for the empirical series. The next columns provide the average scaling exponents from 1000 simulations for the IDC-model, GARCH(1,1),GJR-GARCH(1,1,1), and respectively. The minimum and maximum values are reported in parenthesis.

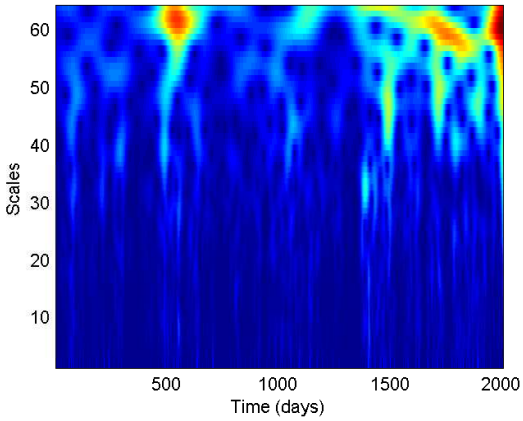
Scaling exponents $\hat{\alpha}(q)$					
Moment (q)	Emp. series	IDC-model	GARCH	GJR-GARCH	Skew-t-GARCH
0.25	0.2452	0.2314	0.1978	0.1988	0.2169
		[0.1660, 0.2855]	[0.1461, 0.2363]	[0.1482, 0.2558]	[0.1601, 0.3021]
0.50	0.4796	0.4450	0.3940	0.3957	0.4287
		[0.3289, 0.5425]	[0.2932, 0.4684]	[0.2979, 0.5035]	[0.3240, 0.6221]
1.00	0.9199	0.8497	0.7806	0.7829	0.8298
		[0.6248, 1.0349]	[0.5910, 0.9225]	[0.5836, 0.9811]	[0.6179, 1.3090]
1.50	1.3198	1.2290	1.1578	1.1592	1.1858
		[0.8794, 1.5279]	[0.8740, 1.3908]	[0.8528, 1.5144]	[0.6946, 1.8664]
2.00	1.6759	1.5832	1.5230	1.5206	1.4882
		[1.0625, 2.0258]	[1.1288, 1.8667]	[1.0845, 2.0712]	[0.5981, 2.3321]
3.00	2.2722	2.2166	2.2077	2.1847	1.9699
		[1.2064, 2.9352]	[1.5336, 2.8283]	[1.3939, 3.0893]	[0.2985, 3.2413]
4.00	2.7798	2.7713	2.8282	2.7706	2.3679
		[1.2861, 3.7509]	[1.7756 3.7694]	[1.6121, 3.9573]	[-0.0247, 4.0708]
4.50	3.0193	3.0300	3.1186	3.0417	2.5533
		[1.3187, 4.1405]	[1.8604, 4.2255]	[1.7113, 4.3633]	[-0.1855, 4.4724]
5.00	3.2536	3.2808	3.3992	3.3027	2.7339
		[1.3490, 4.5232]	[1.9357, 4.6757]	[1.8067, 4.3633]	[-0.3454, 4.8697]
5.50	3.4844	3.5261	3.6720	3.5565	2.9112
		[1.3779, 4.9024]	[2.0067, 5.1185]	[1.8993, 5.1527]	[-0.5043, 5.2643]
6.00	3.7125	3.7674	3.9390	3.8050	3.0865
		[1.4060, 5.2898]	[2.0756, 5.5544]	[1.9897, 5.5415]	[-0.6624, 5.6569]
6.50	3.9385	4.0058	4.2016	4.0496	3.2602
		[1.4335, 5.6729]	[2.1436, 5.4844]	[2.0785, 5.9279]	[-0.8201, 6.0482]
7.00	4.1628	4.2421	4.4607	4.2914	3.4328
		[1.4607, 6.0529]	[2.2113, 6.4092]	[2.1661, 6.3128]	[-0.9773, 6.4384]

TABLE 3. Estimates of benchmark model parameters. The signal was prefiltered using ARMA(1,1) with AR-parameter $\phi_1 = 0.6890$ (0.0537), MA-parameter $\theta_1 = -0.8334$ (0.0416). L denotes log-likelihood. The standard errors are reported in parenthesis. All of GARCH-models were estimated using Quasi-Maximum likelihood.

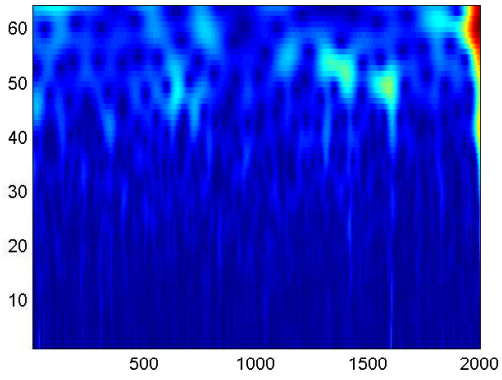
	ω	α	γ	β	ν	λ	L
GARCH	0.0021	0.3448	-	0.3833	-	-	2455.9
	(0.0015)	(0.1018)	-	(0.2096)	-	-	
GJR-GARCH	0.0028	0.2657	0.3595	0.2140	-	-	2462.8
	(0.0018)	(0.0779)	(0.2792)	(0.2324)	-	-	
Skew-t-GARCH	0.0003	0.4234	-	0.5766	4.1000	0.0688	3003.9
	(0.0002)	(0.1565)	-	(0.1323)	(1.0806)	(0.0237)	



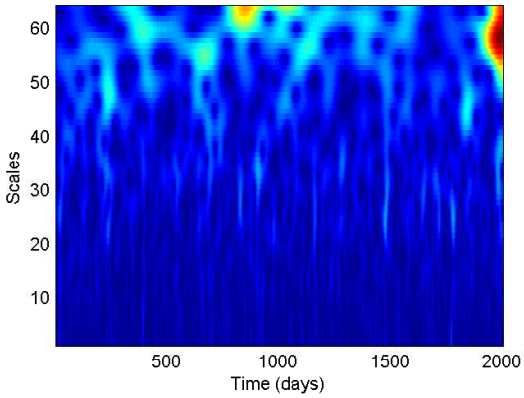
(a) Original signal



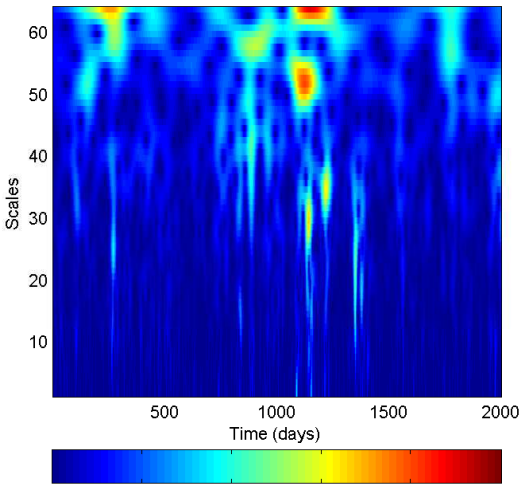
(b) IDC-process



(c) GARCH



(d) GJR-GARCH



(e) Skew-t-GARCH

FIGURE 5. Scalogram analysis