IWQW

Institut für Wirtschaftspolitik und Quantitative Wirtschaftsforschung

Diskussionspapier
Discussion Papers

No. 04/2009

A Two-Factor Model for Electricity Prices with Dynamic Volatility

Stephan Schlüter
University of Nuremberg

ISSN 1867-6707

A Two-Factor Model for Electricity Prices with Dynamic Volatility

Stephan Schlüter

Department of Statistics and Econometrics, University of Erlangen-Nürnberg, Lange Gasse 20, D-90403 Nürnberg

Email: stephan.schlueter@wiso.uni-erlangen.de

ABSTRACT

The wavelet transform is used to identify a biannual and an annual seasonality in the Phelix Day Peak and to separate the long-term trend from its short-term motion. The short-term/long-term model for commodity prices of Schwartz & Smith (2000) is applied but generalised to account for weekly periodicities and time-varying volatility. Eventually we find a bivariate SARMA-CCC-GARCH model to fit best. Moreover it surpasses the goodness of fit of an univariate GARCH model, which shows that the additional effort of dealing with a two-factor model is worthwile.

Keywords: Wavelets; Seasonal Filter; Relative Wavelet Energy; Multivariate GARCH; Energy Price Modelling.

JEL-Classification: C32, C51.

1 Introduction

Since the liberalisation of European energy markets is proceeding, the statistical modelling of electricity prices is becoming more and more in vogue. A reason for that is the improving data situation and the awareness that a decent understanding of the electricity price process helps to reduce the financial risk for power producers, traders and also large electricity consumers.
A simple approach to model power prices is to treat electricity as a regular commodity and use e.g. the long-term/short model of Schwartz & Smith (2000). But this proves to be not sufficient to capture the particularities of electricity: it is non-storable (at least in the short term), its price is reverting to a cost-demand equilibrium, it shows explicite intra-day patterns that vary with the season, and – due to transportation constraints – the power price can differ between regions or coutries. We neither concern with the intra-day patterns nor with the regional price differences

as we focus on the Phelix Day Peak, which is the average day-ahead energy price of the hours between 08:00 a.m. and 08:00 p.m. traded at the European Energy Exchange (EEX), Leipzig. Beyond that we disagree with the property of absolute non-storability (see Section 2), but consider the prices' reversion to a long-term equilibrium as relevant. Because of that we propose to use an adjusted version of the concept of Schwartz & Smith (2000) that contains dynamic volatility[1]. To validate our model we compare it to the original approach, various generalisations and – to test if the additional effort of a two-factor model is worthwhile – to a one-factor model with dynamic volatility. Eventually we find in Section 4 that the correlation can be modeled time-constant but the volatiliy cannot. But initially we motivate our concept and give an overview over the existing literature in Section 2. An introduction into multivariate dynamic volatility modelling is added. In Section 3 we explain how to filter seasonal patterns and extract the long-term drift. Therefore we introduce and apply the wavelet transform. A conclusion summarizes the paper.

2 A Long-Term/Short-Term Power Price Model

Electricity price models can be clustered into non-parametric, fundamental and stochastic ones. Non-parametric concepts are, for example, artificial neural networks (see Szkuta et al., 1999, or Wang & Ramsay, 1998) or the agent-based approach of Weidlich (2008). Fundamental models estimate the (log-)price based on factors like weather or power load, for example. Vehviläinen & Pyykkönen (2004) construct a price model for the Nordic market and Schindlmayr (2005) applies the same concept – however being more complex – to the spot price traded at the EEX. The major part of authors concerns with stochastic price models. The Ornstein-Uhlenbeck diffusion process suggested by Lucia & Schwartz (2002) is used in different extensions. Borovkova & Permana (2004) or Cartea & Figueroa (2005), for example, add a Poisson jump process in order to explicitely incorporate the characteristics of price spikes. Weron et al. (2003) extend the model to a regime-switching approach and Benth et al. (2007) propose a non-Gaussian Ornstein-Uhlenbeck process. However all mentioned authors assume the power price's volatility to be constant over time. Using the example of Figure 2 it can be shown that this is not the case: there are periods of weak and periods of strong oszillation (so-called volatility clusters). Serletis & Shahmoradi (2006) recognise this and design, just as Mugele et al. (2005), the volatility as an autoregressive process. We conclued this brief literature review by referring to Weron & Misiorek (2008) and Barlow (2002) for a broader overview over and discussion of existing stochastic concepts.

To motivate our approach we follow Barlow (2002) and respect that the power price is fixed where a predictable (inelastic) demand meets a supply generated by a coutry-specific mix of power plant types with varying production flexibility and marginal costs. The German 2007 energy mix, for

[1] We define volatility as standard deviation that can be interpreted as a measure for variation.

example, breaks down to 22% nuclear power, 24% brown coal, 22% hard coal, 12% natural gas, 6% oil, 7% wind and 7% of other renewable energies (see Kiesel & Herkner, 2008). In the short run there are various scenarios where a unpredictable supply drop (e.g. an outage of a nuclear power plant due to technical reasons) or supply increment (i.e. an unpredicted wind energy production) does influence the price equilibrium which is - due to the short-term non-storability of electricity - highly volatile over time (see also e.g. Benth et al., 2007). But in the long run over 80% of the energy sources, i.e. coal, gas, oil and – to a certain extend – also nuclear power, is storable. The power plant operators are able to decide whether to switch on or off the machines (or at what level to produce energy) and store or sell the respective fuel at the market. The periods in which the switching or leveling is possible vary with the type of power plant but the essence is that on the long run and to a certain extend electricity is storable in form of its fuel which is a commodity and can be therefore modeled as one.

Due to the argumentation above we propose to apply the model of Schwartz & Smith (2000) in order to respect the difference between the long-term and the short-term power price. Their approach breaks down the process of log-prices $\ln S_t$ into a sum consisting of an Ornstein-Uhlenbeck process χ with mean reversion $\kappa \in \mathbb{R}$ for the short-term dynamics, and a Brownian motion ξ_t with long-term mean μ_ξ for the long-term price development. We now additionally add a seasonal adjustment component g_t to the model, incorporate both an autoregressive and a moving average factor in ξ_t and a weekly pattern in χ_t. For the volatilities $\sigma_{\chi,t}, \sigma_{\xi,t}$ we use the dynamic model defined below. Moreover let $\theta_1, \theta_2 \in \mathbb{R}$ be the autoregression parameters and $\psi \in \mathbb{R}$ the moving average coefficient; the error vector $(z_{\chi,t}, z_{\xi,t})$ is bivariately Gaussian distributed with zero mean and variance-covariance matrix Σ_t. Using these definitions our model reads as

$$\begin{aligned} \ln(S_t - g_t) &= \chi_t + \xi_t, \\ \chi_t &= (1-\kappa)\chi_{t-1} + \theta_1 \chi_{t-7} + \sigma_{\chi,t} z_{\chi,t}, \\ \xi_t &= \theta_2 \xi_{t-1} + \psi(\sigma_{\xi,t-1} z_{\xi,t-1}) + \sigma_{\xi,t} z_{\xi,t}. \end{aligned} \tag{2.1}$$

For a dynamic volatility model we refer to the concept of generalised autoregressive conditional heteroscedasticity (GARCH),which was designed by Engle (1982) and Bollerslev (1986) to model volatility clustering. For a detailed introduction into univariate GARCH models we refer to McNeil et al. (2005). Here we merely define the multivariate extension, whereby the simplest one is the constant conditional correlation GARCH model (CCC-GARCH). It assumes the volatilities to be time-dependent, but the correlation-matrix to be constant. Let $\Sigma_t = \Delta_t P_C \Delta_t$,with P_C being the positive-definite correlation matrix. Then Δ_t is a diagonal 2×2 scaling matrix with univariate

volatility processes $\sigma_{\chi,t}$ and $\sigma_{\xi,t}$ as elements. These are calculated as

$$\begin{aligned}\sigma^2_{\chi,t}|\mathcal{F}^{\chi}_{t-1} &= \omega_{\chi,t} + \sum_{i=1}^{p_\chi} \alpha_{\chi,i}(\chi_{t-i} - (1-\kappa)\chi_{t-1-i})^2 + \sum_{i=1}^{q_\chi} \beta_{\chi,i}\sigma^2_{\chi,t-i},\\ \sigma^2_{\xi,t}|\mathcal{F}^{\xi}_{t-1} &= \omega_{\xi,t} + \sum_{i=1}^{p_\xi} \alpha_{\xi,i}(\xi_{t-i} - \xi_{t-1-i})^2 + \sum_{i=1}^{q_\xi} \beta_{\xi,i}\sigma^2_{\xi,t-i},\end{aligned}$$

with $\mathcal{F}_t$ being the filtration generated by χ_t resp. ξ_t (i.e. the information up till t). Let $k \in \{\chi, \xi\}$, then holds: $p_k, p_k \in \mathbb{N}$, $\omega_{k,t} > 0, \alpha_{k,i} \geq 0, i = 1, \ldots, p_k$ and $\beta_{k,j} \geq 0, j = 1, \ldots, q_k$. The unconditional variance exists, if $\sum_{p_k} \alpha_{k,i} + \sum_{q_k} \beta_{k,j} < 1$. As this definition depends on the choice of lags $p = (p_\chi, p_\xi)$ resp. $q = (q_\chi, q_\xi)$ we speak of a GARCH(p,q) process (see McNeil et al., 2005). In case of the correlation is time-varying the concept is called dynamic conditional correlation GARCH model (DCC-GARCH). Tse & Tsui (2002) propose an autoregressive moving average (ARMA) process for the correlation, Engle & Sheppart (2001) use a GARCH(1,1) process.

3 Wavelets as a Preprocessing Tool

In (2.1) the variables χ_t and ξ_t are designed as hidden variables; they have to be estimated using the observations of S_t. Schwartz & Smith (2000) apply therefore the Kalman-Filter which is a recursive algorithm to estimate unobserved state variables that do influence observable ones (see Kalman, 1960). They consider it as an efficient tool, but it has the drawback that it requires a prior distribution for the unobserved variables. We therefore prefer a wavelet-based approach which does not need such an assumption.

The wavelet transform is a generalisation of the Fourier transform. It is a mapping from time space into time-frequency space that is – in contrary to the latter – capable of identifying regular oszillations with time-varying intensity and frequency. We can additionally use it to extract a long-term trend from a time series. Stationarity of the process is not required but square-integrability. For a detailed introduction into continuous wavelet transform we refer to Mallat (2003). Kaiser (1994) and Jensen & Cour-Harbo (2001) focus on the discrete version. Within this Section we give definitions that are relevant for our analysis in Section 4.

The continous wavelet transform is defined as the convolution of a function $f(t) \in \mathcal{L}^2(\mathbb{R})$ with a complex-valued function $\Psi(t) \in \mathcal{L}^1(\mathbb{R}) \cap \mathcal{L}^2(\mathbb{R})$ which – to be called wavelet – has to fulfill the admissible condition

$$C_\Psi = \int_{-\infty}^{\infty} \frac{\left|\hat{\Psi}(\omega)\right|^2}{|\omega|} d\omega < \infty, \tag{3.1}$$

whereby the hat denotes the Fourier transform (see Fabert, 2004). From (3.1) follows that Ψ tends to zero for $|\omega| \to \infty$ and is located in a finite mean. In order to cover the whole time-axis with Ψ we introduce a translation parameter $b \in \mathbb{R}$. A scaling parameter $a > 0$ can be used to dilute the shape of Ψ. To indicate this modification we write

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}} \Psi \left(\frac{t-b}{a} \right),$$

and define the continuous wavelet transform of a function $f(t)$ as

$$WT(a,b) = \langle f, \Psi_{a,b} \rangle = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{a}} \Psi^* \left(\frac{t-b}{a} \right) dt. \tag{3.2}$$

The $a^{-1/2}$ normalizes the energy and $*$ denotes the conjugate complex. The wavelet transform given a pair (a,b) can be interpreted as an orthogonal mapping onto the time-scale plane that quantifies the proportion of $f(t)$ explained by $\Psi_{a,b}$. In (3.2) we also see that the values of $f(t)$ influence the wavelet coefficients within a certain interval (around $b_0 = t$) that depends on the choice of Ψ. This region is called cone of influence (COI) and varies with the respective scale. The COI is a half-plane (b,a) shaped by $|b_0 - b| < \gamma \cdot a$ (see Lau & Weng, 1995).

Figure 1: The Real Part of the Morlet Wavelet at Different Scales

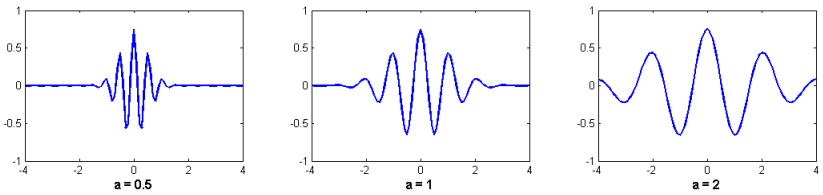


Figure 1 shows the Morlet wavelet at three different scales. We can clearly see the character of a local osziallation, i.e. it is diminishing outside a certain scale-dependent time-window. The formula of Morlet's wavelet reads as

$$\begin{aligned} \Psi_M(t) &= c_{\omega_0} \pi^{-1/4} e^{-t^2/2\sigma^2} \left(e^{i\omega_0 t} - e^{-\frac{1}{2}\omega_0^2} \right), \\ c_{\omega_0} &= \left(1 - e^{-\omega_0^2} - 2e^{-\frac{3}{4}\omega_0^2} \right)^{-\frac{1}{2}}, \end{aligned} \tag{3.3}$$

whereby ω_0 denotes a certain basis frequency and $\sigma > 0$ (see Daubechies, 1992). For $\omega_0 > 5$ holds $c_{\omega_0} \approx 1, e^{-\omega_0^2/2} \approx 0$. The COI parameter γ of the Morlet wavelet is $\sqrt{2}$ (see Lau & Weng, 1995).

The inverse operation to (3.2) does also exist (see e.g. Mallat, 2003),

$$f(t) \quad = \quad \frac{1}{C_\Psi} \int_0^\infty \int_{-\infty}^\infty WT(a,b) \frac{1}{a^2\sqrt{a}} \Psi\left(\frac{t-b}{a}\right) dbda. \tag{3.4}$$

Shrinking the domain of a in (3.4) to a scale window $[a_l, a^u]$ gives the portion of f explained by these scales. However, although reducing this scale window to one point is mathematically possible, the result cannot be interpreted as the influence of the chosen scale on f at time t. The reason for that is the "uncertainty principle" of time-frequency analysis (see Fabert, 2004) which states that not both scale and location of a signal can be exactly specified simultaneously; i.e. a good time resolution reduces the frequency resolution and vice versa (see Lau & Weng, 1995). We merely can derive statements about a certain window of scales and time. The minimal width of this time-scale window varies with the wavelet function. Applying the Morlet wavelet yields among all admissible functions the smallest possible window size (see Ahuja et al., 2005), which we give – based on Fabert (2004) – as a set-valued function ζ of a and b:

$$\zeta(a,b) = \left[b - \frac{a\sigma}{\sqrt{2}}, b + \frac{a\sigma}{\sqrt{2}}\right] \times \left[a \cdot \frac{2\sqrt{2}\pi\sigma}{\omega_0\sqrt{2}\sigma + 1}, a \cdot \frac{2\sqrt{2}\pi\sigma}{\omega_0\sqrt{2}\sigma - 1}\right]. \tag{3.5}$$

This window size is only one possible criterion for choosing a specific wavelet; depending on the situation there may be other requirements. For a broad list of wavelet features we refer to Farge (1992), Meyers et al. (1993) or Ahuja et al. (2005).

The uncertainty principle makes it harder to identify the intensity of a relevant scale a^* via wavelet transform, as also the wavelet coefficients within a certain proximity of a^* are influenced and contain therefore information about it. In the wavelet coefficients' contour plot this fact appears as oval contour lines centered around a^* with alternating positive and negative height (see e.g. Figure 3 or Torrence & Compo, 1998).

We can additionally compress the information contained in the matrix of wavelet coefficients by computing the (relative) wavelet energy E_a (E_a^{rel}) which is a function of a and defined as (see e.g. Salwani & Jasmy, 2005)

$$E_a = \|WT(a,b)\|_2^2 = \int_{-\infty}^\infty WT(a,b) \cdot WT^*(a,b) db, \quad E_a^{rel} = \frac{E_a}{\|f\|^2}. \tag{3.6}$$

The function E_a^{rel} does exist due to (3.1) and gives the fraction of variation in f explained by the scale a (see e.g. Rosso et al., 2001). The bigger E_a^{rel} the more important is a. Thereby – because of the uncertainty principle – E_a^{rel} is increasing for $a \to a^*$ within a certain interval around a^*.

When estimating (3.6) from a finite data set $f(t), t = 1, \ldots, T$, we have to be aware of certain edge effects. These occur because (3.2) is based on an infinite (continuous) function f which contradicts

the finiteness of the data set that is to be used for estimating the coefficients. The COI of the data's end points shapes the region where these effects show. They can be reduced by padding the ends of the data sample with zeros, but thereby we cause frictions which do also skew the wavelet coefficients. Meyers et al. (1993) discuss other methods but in all cases it cannot be quantified how much of the skewness is reduced. A satisfying solution has - to our knowledge - not yet been found. We therefore focus on the coefficients outside the COI for estimating the wavelet energy. Let $c_l(a)$ ($c_r(a)$) be a function with values in $\{1,\ldots,T\}$ that maps the scale on the left (right) margin of the endpoints' COI. Then we can define an unskewed estimate of the (relative) wavelet energy by

$$\hat{E}_a = \frac{1}{c_r(a) - c_l(a) - 1} \sum_{c_l(a)+1}^{c_r(a)-1} |WT(a,b)|^2, \quad \hat{E}_a^{rel} = \frac{E_a}{\sum_a E_a}. \tag{3.7}$$

With (3.7) and the COI's general formula we see that the number of addends decreases with the scale. This reduces the estimate's quality with increasing size of a. However (3.7) is consistent, if the process of wavelet coefficients at scale a is erdgodic which again follows directly from the ergodicity of the underlying process if the latter is of second-order (see Wu & Su, 1996).
For a pattern identification procedure it is required to combine the visual with the analytical analysis. The contour plot gives information about the permanence of a regular pattern contained in f and about possible shifts in its period. The relative wavelet energy, again, is used for specifying the exact scales.

Now, having identified a relevant pattern with scale a^* in f, we can quantify it in case of using Morlet wavelet by reducing (3.4) to $f_a^*(t)$ with

$$f_{a^*}(t) = \frac{1}{C_\Psi} \int_{\left[\frac{a^* 2\pi\sqrt{2}\sigma}{\omega_0\sqrt{2\sigma+1}}, \frac{a^* 2\pi\sqrt{2}\sigma}{\omega_0\sqrt{2\sigma-1}}\right]} \int_{-\infty}^{\infty} WT(u,b) \frac{1}{u^2\sqrt{u}} \Psi\left(\frac{t-b}{u}\right) dbdu. \tag{3.8}$$

The Morlet wavelet has the smallest-possible time-scale window, but it lacks an important feature: it has no corresponding scaling function $\phi(a_0)$ that aggregates the influence of all scales larger than a_0 on f. This ϕ can be interpreted as the long-term trend of f and Mallat (2003) shows

$$\begin{aligned} f(t) = & \frac{1}{C_\Psi} \int_0^{a_0} \int_{-\infty}^{\infty} WT(a,b) \frac{1}{a^2\sqrt{a}} \Psi\left(\frac{t-b}{a}\right) dbda \\ & + \frac{1}{C_\Psi a_0} \int_{-\infty}^{\infty} \left\langle f(s), \frac{1}{\sqrt{a_0}} \phi\left(\frac{s-b}{a_0}\right) \right\rangle \frac{1}{\sqrt{a_0}} \phi\left(\frac{t-b}{a_0}\right) db. \end{aligned} \tag{3.9}$$

Formula (3.9) can be simplified according to Shannon's Sampling Theorem which states that a continuous signal can be exactly discretely reconstructed if it is band-limited, i.e. its Fourier transform equals zero above a certain finite threshold (see Blatter, 2002). As this is by definition

the case for

$$f^*(t) = f(t) - \frac{1}{C_\Psi a_0} \int_{-\infty}^{\infty} \left\langle f(s), \frac{1}{\sqrt{a_0}} \phi\left(\frac{s-b}{a_0}\right) \right\rangle \frac{1}{\sqrt{a_0}} \phi\left(\frac{t-b}{a_0}\right) db, \tag{3.10}$$

we can replace the integration over all scales between zero and a_0 in (3.9) by a finite sum. For the specification of ϕ we follow Ahuja et al. (2005) and Unser (1999) who suggest to use B-splines, i.e. the function

$$\phi(x) = \frac{1}{(L-1)!} \sum_{k=0}^{L} (-1)^k \begin{pmatrix} L \\ k \end{pmatrix} (x-k)_+^{L-1} \quad L > 1. \tag{3.11}$$

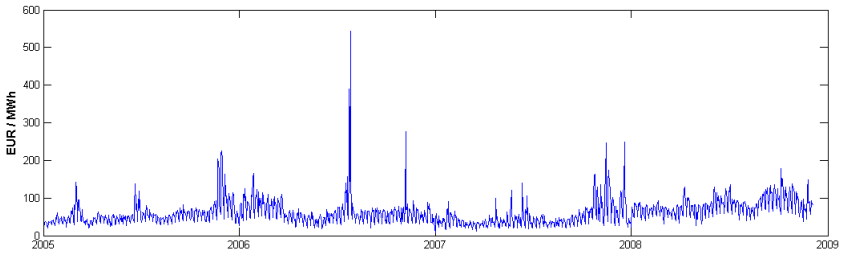
So, applying the concepts defined in this section we can identify resp. quantify relevant patterns contained in a function f. Moreover, by using a scaling function, it is possible to extract the long-term trend of f, whereby the optimal choice of a_0 is still to be discussed.

4 Benchmarking the Model

We use observations from the Phelix Day Peak to benchmark our model. This index has been traded since 2002 but we leave out the first three years as the data show a structural break between 2004 and 2005. The time series plotted in Picture 2 includes weekends, it starts on January 1st, 2005 and ends on December 4th, 2008. We see the typical properties of a short-term energy price: extreme price spikes do occure and the volatility is occasionally high, i.e. clustered.

Initially we apply the wavelet transform to identify regular patterns in the observations. As moti-

Figure 2: The Phelix Day Peak



vated in Section 3 we use the Morlet wavelet with $\omega_0 = 6$ (so the scale approximately equals the

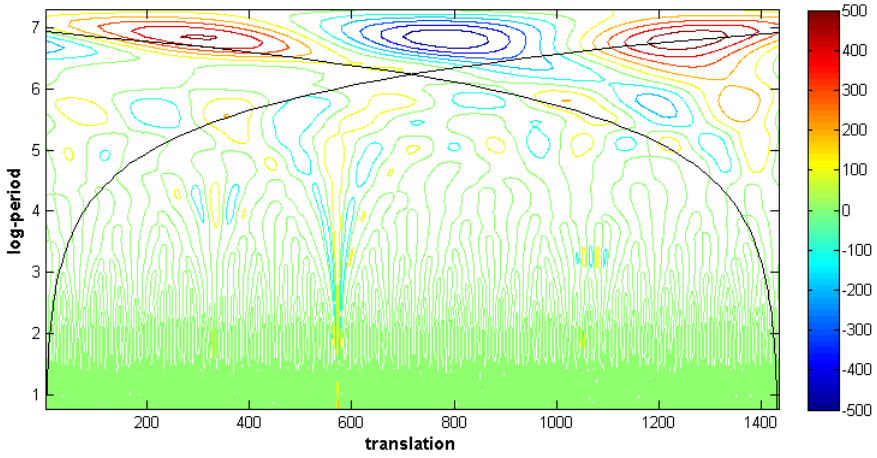
inverse frequency) and $\sigma = 1$. The long-term trend is extracted later but we keep (3.10) in mind and assume that this yet to be trend-adjusted signal is band-limited. A discrete range of scales is therefore sufficient and we apply the dyadic scheme given e.g. by Torrence & Compo (1998). It is based on a time series which length can be written as $2^\gamma, \gamma \in \mathbb{N}$, so we set $T' = 2^{\lfloor \log_2 T \rfloor}$ and construct our grid as follows:

$$\begin{aligned} a_0 &= 2, \Delta j = 0.125, J = \left\lfloor \Delta j^{-1} \log_2 \left(T'/a_0 \right) \right\rfloor + 1 \\ a_j &= a_0 2^{j \delta j}, \quad j = 0, 1 \ldots, J. \end{aligned} \tag{4.1}$$

Translation will be performed along a time grid of daily granulation and therefore b can be interpreted as time index. Figure 3 shows the real part of the wavelet coefficients. Regular patterns are indicated by a series of relatively high contour lines with alternating sign. In black lines the end points' COI is additionally plotted. The long-term trend appears as a high-scale (i.e. low-frequency) pattern and spikes like the one in Summer 2006 are indicated by relatively big wavelet coefficients at a low scale.

Besides that we see some relevant patterns between the log-scales 5 and 6 and compute the relative

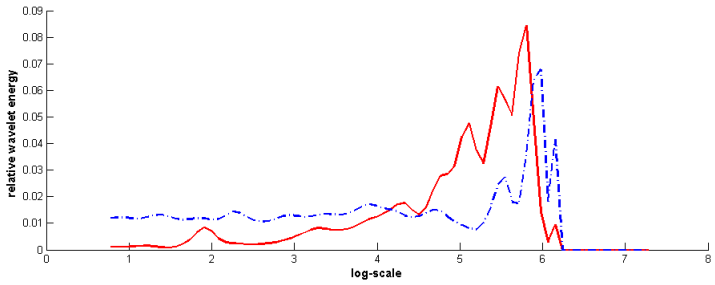
Figure 3: The Wavelet Transform of the Phelix Day Peak



wavelet energy, i.e. (3.7), in order to retrieve more information. In Figure 4 it is represented by the red line. The additional blue dashed line is the relative wavelet energy estimated from an equally sized sample of random numbers generated using the method of Marsaglia & Zaman (1991). It serves as a significance line as it indicates the relative wavelet energy of a process without any

regular pattern. Both lines increase with the scale but abruptly return to zero around a log-scale of 6. This is due to the construction of (3.7) and the proximity of the log-scale 6 has to be interpreted carefully as only a few wavelet coefficients can be used for estimating the energy.
In Figure 4 we can distinctly measure peaks in the energy plot at a log-scale of 5.1 (i.e. 160 days resp. about half a year), 5.5 (about 235 days) and 5.8 (332 days resp. about a year). This goes along with the findings from Figure 3 where we see that the two latter scales describe one relevant pattern that changes – more precisely: increases – its period over the considered time. This fact again shows the advantage of wavelet transform over other seasonal filters as it can identify changes in the pattern's period over time.

Figure 4: The Estimated Relative Wavelet Energy



The red line denotes the estimated relative wavelet energy, the blue dashed one the significance line generated by a random sample.

In order to quantify the identified seasonal patterns we discretize (3.8) according to (4.1). Additionally we sum up the two latter scales (as they represent one pattern) and eventually obtain two log-scale windows w_1,w_2 to be inverted:

$$w_1 = \left[\left(\frac{5.1 \cdot 2\pi\sqrt{2}}{6\sqrt{2}+1}\right), \left(\frac{5.1 \cdot 2\pi\sqrt{2}}{6\sqrt{2}-1}\right)\right] \quad , \quad w_2 = \left[\left(\frac{5.5 \cdot 2\pi\sqrt{2}}{6\sqrt{2}+1}\right), \left(\frac{5.8 \cdot 2\pi\sqrt{2}}{6\sqrt{2}-1}\right)\right]$$

For the discretization of (3.4) we use that while dealing with the continuous wavelet transform any wavelet function can be applied for the retransformation and opt for the δ function (see Torrence & Compo, 1998). This simplifies (3.8) and determining

$$j_{low} = max\left\{j \mid a_j \leq \left(\frac{5.1 \cdot 2\pi\sqrt{2}}{6\sqrt{2}+1}\right)\right\}, \quad j^{up} = min\left\{j \mid a_j \geq \left(\frac{5.1 \cdot 2\pi\sqrt{2}}{6\sqrt{2}-1}\right)\right\}$$

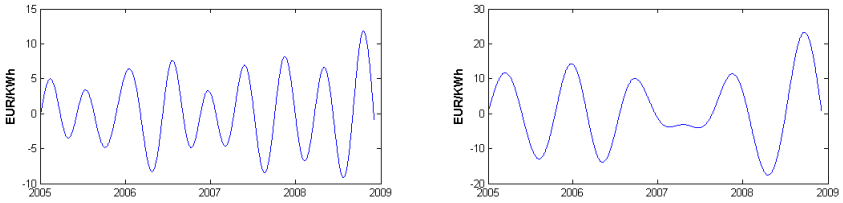
we can calculate according to Torrence & Compo (1998)

$$f_{[\log a=5.1]}(t) = \frac{\Delta j}{C_\delta \Psi(0)} \sum_{j=j_{low}}^{j^{up}} \frac{Re\left[WT(a_j,t)\right]}{\sqrt{a_j}} = \frac{0.125}{0.776 * \pi^{1/4}} \sum_{j=j_{low}}^{j^{up}} \frac{Re\left[WT(a_j,t)\right]}{\sqrt{a_j}}, \tag{4.2}$$

whereby C_δ denotes a wavelet-specific constant and $Re\left[\cdot\right]$ the real part of the respective number. The influence of the second log-scale window is computed analogously and both results are plotted in Figure 5. The shift of the annual pattern towards a longer period in 2007 is clearly visible as in this transition period the oszillations of the two frequencies compensate each other.

We respect the edge effects while computing the relative wavelet energy but – as we cannot quan-

Figure 5: The Biannual and Annual Oszillation



The left graph shows the biannual pattern, the right one the annual pendant.

tify them – not in (4.2). Because the existing damping methods aren't a solution either, we omit the first and last 150 observations after finishing the data preprocession stage. Having performed the seasonal-adjustment (and thus estimated the influence of the (bi)-annual pattern) we apply the logarithm on the data to damp volatility and extract the trend by computing the second addend of (3.9) using (3.11) for ϕ. But that procedure requires to specify a scale a_0 which seperates "long-term" from "short-term". We filtered already the influence of the biannual and annual pattern so it is reasonable to draw the border in this intervall in order to make sure that no relevant frequency window is cut in two. Respecting the scale's dyadic discretization scheme we choose $a_0 = 2^8$.

Subtracting the computed trend from the original seasonal-adjusted log-data yields the short-term oszillation. Both time series are of equal length and can be aggregated in a bivariate data vector which is the basis for our further proceeding. We apply different models of increasing complexity to this vector in order to benchmark the approach described in (2.1). For the estimated parameter values, see Appendix a. As goodness of fit (GOF) measures we choose the log-likelihood value (LLH), Akaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC)[2]. As not all models are nested we additionally compute the mean squared approximation error (MSE)

[2]For formulae see McNeil et al. (2005).

and the mean absolute error (MAE) of approximation. The higher LLH and the lower AIC, BIC, MSE and MAE the better the fit.

First we apply the basic long-term/short-term model of Schwartz & Smith (2005), which we denote by SCSM. Evaluating the partial autocorrelation of the short-term oszillation yields a weekly pattern, so we add a seasonal component and call the model S-SCSM. Comparing both models in Table 1 we can see that including the seasonal pattern improves the GOF. Moreover we perform the LM-Test for presence of (G)ARCH-effects (see Engle, 1982) and find time-varying volatility in both SCSM and S-SCSM. However, the test on dynamic correlation from Engle & Sheppard (2001) yields that the correlation doesn't need to be modeled dynamic so we extend the S-SCSM to a multivariate CCC-GARCH model, incorporate the weekly pattern and add moving average in the trend (we call it SARMA-CCC-GARCH). This proceeding is justified by an increasing LLH and decreasing AIC, BIC, MSE resp. MAE. Because we found the correlation in the CCC-GARCH model to be significant but quite small we additionally fit various univariate error distributions to both the trend and the short-term oszillation. We clearly see that the skewed Student-t distribution proposed by Hansen (1994) reflects more the short-term price behaviour than the Gaussian, which is adequate for the trend's error term. This approach (which we call SARMA-GARCH2) and the other models are not nested so we use MSE and MAE to compare them. We observe an inferior GOF of SARMA-GARCH2 which means that incorporating dependence does make sense.

Table 1: Measuring the Goodness-of-fit of various models

Model	LLH	BIC	AIC	MSE	MAE
SCSM	2306.586	-4585.042	-4605.172	0.118	0.253
SCSM-I	2585.637	-5138.111	-5163.274	0.075	0.181
SARMA-CCC-GARCH	4391.132	-8697.873	-8758.264	0.011	0.055
SARMA-GARCH2	4445.180	-8798.936	-8864.360	0.073	0.180
UGARCH	372.257	-695.279	-730.513	0.155	0.282

LLH = log-likelihood function, BIC = Bayesian Information Criterion, AIC = Akaike's Information Criterion, number of observations = 1134.

Eventually we check if the additional effort of a multivariate model is worthwhile by comparing the derived concepts to an univariate autoregressive GARCH model with Student-t distributed innovations (denoted by UGARCH) as Mugele et al. (2005) propose. But both the MSE and the MAE are distinctly higher than the corresponding values of SARMA-CCC-GARCH.

5 Conclusion

Based on an analysis of the demand-supply equilibrium on the German power market we designed a two-factor model for the Phelix Day Peak. In a preprocessing stage we identified a biannual as well as an annual pattern. After extracting the trend from the seasonal-adjusted log-data (and thus splitting up the time series) we found another (weekly) pattern in the short-term oszillation. Eventually we fitted models of various complexity to the bivariate time series. Measured by LLH, BIC, AIC and MSE resp. MAE we identified as the best approach a bivariate model with constant correlation but dynamic volatility including a weekly pattern in the short-term oszillation and an additional moving average in the trend.
Within this paper we have shown some advantages of wavelet transform and how to use it as a data preprocessing tool. Moreover we see that the additional effort of using a two-factor model for the German power price is justified. However we have to admit that the univariate Student-t distribution definitely fits better to the short-term oszillation than the Gaussian, which is adequate for the trend. Yet the SARMA-CCC-GARCH model, which includes a positive correlation, shows a better GOF – despite of this discrepancy. Therefore, to respect dependence and individual distribution properties, we suggest as a next step to model the margins separately from the dependence structure. For the latter a copula (which is a multivariate distribution function with uniform margins) is an adequate concept.

References

[1] Ahuja N, Lertrattanapanich S, Bose N K. Properties determining choice of mother wavelet. IEEE Proceedings - Vision, Image & Signal Processing 2005; 152(5); 659–664.

[2] Barlow M T. A Diffusion Model for Electricity Prices. Mathematical Finance 2002; 12(4); 287–298.

[3] Blatter C. Wavelets, a primer. A K Peters Ltd: Wellesley; 2002.

[4] Benth F E, Kallsen J, Meyer-Brandis T. A Non-Gaussian Ornstein-Uhlenbeck- Process for Electricity Spot Price Modeling and Derivatives Pricing. Applied Mathematical Finance 2007: 14(2); 153–169.

[5] Bollerslev T. Generalized Autoregressive Conditional Heteroscedasticity. Journal of Econometrics 1986; 31; 307–327.

[6] Borovkova S, Permana F J. Modeling electricity prices by the potential jump-diffusion. Springer: New York; 2004.

[7] Cartea A, Figueroa M. Pricing in Electricity Markets – a mean reverting jump diffusion model with seasonality. Applied Mathematical Finance 2005; 12(4); 313–335.

[8] Daubechies I. Tel Lectures on Wavelets. Society for Industrial and Applied Mathematics: Philadelphia; 1992.

[9] Engle R F. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica 1982; 50(4); 987–1007.

[10] Engle R F, Sheppard K. Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. Working Paper 2001; URL: http://EconPapers.repec.org/RePEc:cdl:ucsdec:2001-15.

[11] Fabert O. Effiziente Wavelet Filterung mit hoher Zeit-Frequenz-Auflösung. Verlag der Bayerischen Akademie der Wissenschaften: Munich; 2004.

[12] Farge M. Wavelet transforms and their applications to turbulence. Annual Review Fluid Mechanics 1992; 24; 395–457.

[13] Hansen B E. Autoregressive conditional density estimation. International Economic Review 1994; 3; 705–730.

[14] Jensen A, Cour-Harbo. Ripples in Mathematics, The Discrete Wavelet Transformm. Springer: Berlin; 2001.

[15] Kaiser G. A Friendly Guide to Wavelets. Birkhäuser: Boston; 1994.

[16] Kalman R E. A New Approach to Linear Filtering and Prediction Problems. Transactions of the ASMEJournal of Basic Engineering 1960; 82 D; 35–45.

[17] Kiesel F, Herkner T: Energiemarkt Deutschland, Zahlen und Fakten zur Gas- und Stromversorgung. VWEW Energieverlag GmbH: Frankfurt am Main; 2008.

[18] Lau K-M, Weng H. Climate Signal Detection Using Wavelet Transform: How to Make a Time Series Sing. Bulletin of the American Meteorological Society 1995; 76(12); 2391–2402.

[19] Lucia J J, Schwartz E S. Electricity prices and power derivatives: Evidence from the Nordic Power Exchange. Review of Derivatives Research 2002; 5; 5–50.

[20] Mallat S. A wavelet tour of signal processing, 2nd edition. Academic Press: Manchester; 2003.

[21] Marsaglia G, Zaman A. A New Class of Random Number Generators. Annals of Applied Probability 1991; 3; 462–480.

[22] McNeil A J, Frey R, Embrechts P. Quantitative Risk Management: Concepts, Techniques and Tools. Princeton University Press: New Jersey; 1991.

[23] Meyers S D, Kelly B G, O'Brien J J. An Introduction to Wavelet Analysis in Oceanography and Meteorology: With Application to the Dispersion of Yanai Waves. Monthly Weather Review 1993; 121(10); 2858-2866.

[24] Mugele C, Rachev S T, Trück S. Stable Modeling of different European Power Markets. Investment Management and Financial Innovations 2005; 2(3); 65–85.

[25] Rosso O, Blanco S, Yordanova J, Kolev V, Figliola A, Schürmann M, Basar E. Wavelet entropy: a new tool for analysis of short duration brain electrical signals. Journal of Neuroscience Methods 2001; 105; 65–75.

[26] Salwani M D, Jasmy Y. Relative Wavelet Energy as a Tool to Select Suitable Wavelet for Artifact Removal in EEG. Proceeding of 1st Conference on Computers, Communication, and Signal Processing, 14–15 November 2005: Kuala Lumpur; 2005.

[27] Schwartz E, Smith J E. Short-Term Variations and Long-Term Dynmaics in Commodity Prices. Management Science 2000; 46(7); 893–911.

[28] Schindlmayr G A. A Regime-Switching Model for Electricity Spot Prices. 10th Symposium on Banking, Finance and Insurance, University of Karlsruhe; 2005.

[29] Serletis A, Shamoradi A. Measuring and Testing Natural Gas and Electricity Markets Volatility: Evidence from Alberta's Deregulated Markets. Studies in Nonliear Dynamics & Econometrics 2006; 10(3); 1341–1341.

[30] Szkuta B R, Sanabria L A, Dillon T S. Electricity price short-term forecasting unsign artificial neural networks. IEEE Transactions on Power Systems 1999; 14; 851–857.

[31] Torrence C, Compo G P. A Practical Guide to Wavelet Analysis. Bulletin of the American Meteorological Society 1998; 79(1); 61–78.

[32] Tse Y, Tsui A. A multivariate GARCH model with time-varying correlations. Journal of Business and Economic Statistics 2002; 20; 351–362.

[33] Unser M. Ten good reasonst for using spline wavelets. IEEE Signal Processing 1999; 16; 22–38.

[34] Vehviläinen I, Pyykkönen T. Stochastic factor model for electricity spot price - the case of the Nordic market. Energy Economics 2004; 27; 361-367.

[35] Wang A J, Ramsay B. A neural network based estimator for electricity spot-pricing withparticular reference to weekend and public holidays. Neurocomputing 1998; 23; 47–57.

[36] Weidlich A. A critical survey of agent-based wholesale electricity market models. Energy Economics 2008; 30; 1728–1759.

[37] Weron R, Bierbrauer M, Trück S. Modeling electricity prices: jump diffusion and regime switching. Physica A 2003; 336(1-2); 39–48.

[38] Weron R, Misiorek A. Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models. International Journal of Forecasting 2008; 24(4); 744-763.

[39] Wu B-F, Su Y-L. The Ergodicity and Stationarity Property Analysis of Nonstationary Stochastic Processes Using Wavelet Transform. Proceedings of the 35th IEEE Conference on Decision and Control 1996; 3(11–13); 3107 – 3112.

A The models and their fitted parameters

Table 2 : The Parameter of the UGARCH Model

μ	θ	ω	α	β	ν	λ
1.0220	0.7527	0.0018	0.1006	0.8957	9.8736	1.334
(0.1217)	(0.0297)	(0.0008)	(0.0157)	(0.0140)	(0.0780)	(3.1252)

μ = long-term mean, θ = autoregression parameter, (ω, α, β) = GARCH parameter,
ν = degrees of freedom, λ = skewing parameter. The standard errors are presented in brackets.

Table 3: Parameter of The Two-Factor-Models

	μ_ξ	κ	θ_1	θ_2	ψ	ρ	ω_χ	ω_ξ
SCSM	0.0074	0.7288	0	0.9976	0	0.030	0.3417	0.0224
	(0)	(0.0286)	(0.0019)	(0)	(0)	(0.0198)	(0)	(0)
S-SCSM	0.0074	0.7659	0.6130	0.9976	0	0.1399	0.2696	0.0224
	(0)	(0.0289)	(0.0233)	(0)	(0)	(0.0166)	(0)	(0)
SARMA-	0.0098	0.7659	0.6130	0.9976	0.9960	0.0025	0.0015	0.0000
CCC-GARCH	(0)	(0.0289)	(0.0233)	(0.0023)	(0.0023)	(0)	(0)	(0)
SARMA-	0.0098	0.7659	0.6130	0.9976	0.9960	0	0	0.0019
GARCH2	(0)	(0.0289)	(0.0233)	(0.0023)	(0.0023)	(0)	(0)	(0)

μ_ξ = mean of the trend, κ = mean-reversion, θ_1 = parameter of the weekly autoregression, θ_2 = the trend's autoregression, ψ = the trend's moving average parameter, ρ = correlation between long- and short-term motion,
ω_1, ω_2 = constant volatility.

Table 4: Parameter of The Two-Factor-Models (cont.)

	α_χ	β_χ	α_ξ	β_ξ	ν_χ	λ_χ
S-ARMA	0.1884	0.8116	0.8639	0.1361		
CCC-GARCH	(0.0010)	(0.0002)	(0.4237)	(0.3660)		
S-ARMA	0.1861	0.8011	0.8217	0.2344	5.2592	0.9872
GARCH2	(0.0333)	(0.0277)	(0.0643)	(0.0474)	(0.7975)	(0.0392)

$(\alpha_\chi, \alpha_\xi, \beta_\chi, \beta_\xi)$ = the GARCH-model's parameter, ν, λ_χ = degrees of freedom and skewness of the short-term oszillation's Student t distribution.

Diskussionspapiere 2009
Discussion Papers 2009

01/2009 **Addison, John T. and Claus Schnabel:** Worker Directors: A German Product that Didn't Export?

02/2009 **Uhde, André and Ulrich Heimeshoff:** Consolidation in banking and financial stability in Europe: Empirical evidence

03/2009 **Gu, Yiquan and Tobias Wenzel:** Product Variety, Price Elasticity of Demand and Fixed Cost in Spatial Models

Diskussionspapiere 2008
Discussion Papers 2008

01/2008 **Grimm, Veronika and Gregor Zoettl**: Strategic Capacity Choice under Uncertainty: The Impact of Market Structure on Investment and Welfare

02/2008 **Grimm, Veronika and Gregor Zoettl**: Production under Uncertainty: A Characterization of Welfare Enhancing and Optimal Price Caps

03/2008 **Engelmann, Dirk and Veronika Grimm**: Mechanisms for Efficient Voting with Private Information about Preferences

04/2008 **Schnabel, Claus and Joachim Wagner**: The Aging of the Unions in West Germany, 1980-2006

05/2008 **Wenzel, Tobias**: On the Incentives to Form Strategic Coalitions in ATM Markets

06/2008 **Herrmann, Klaus**: Models for Time-varying Moments Using Maximum Entropy Applied to a Generalized Measure of Volatility

07/2008 **Klein, Ingo and Michael Grottke:** On J.M. Keynes' "The Principal Averages and the Laws of Error which Lead to Them" - Refinement and Generalisation