

# Pricing of Hourly Exercisable Electricity Swing Options Using Different Price Processes

Guido Hirsch

EnBW Trading GmbH  
Durlacher Allee 93, 76131 Karlsruhe, Germany  
gu.hirsch@enbw.com

## Abstract

In this paper fair values of hourly exercisable swing options written on the EEX spot price as an underlying are calculated using three different two factor models (regime-switching AR-process, jump-diffusion process with Bernoulli jump-terms and a normal inverse gaussian process). Therefore, an efficient Least Squares Monte Carlo algorithm (LSM) is introduced and applied to swing options with up to 5000 exercise rights. Finally, the three models are compared with a focus on their ability to reproduce the characteristics of the EEX spot prices and the swing option values resulting for different numbers of exercise rights.

A Bermudan option gives the buyer the right to exercise the option at a specified number of times. As Bermuda is geographically situated between America and Europe, the name of this option is a pun — it reflects that a Bermudan option is intermediate between a European option (exercise only possible at expiry) and an American option (exercise possible at any time until expiry). Each of these plain vanilla options may be specified as either a call or a put option. By contrast, a swing option gives the buyer the right to exercise one and only one call, which is called up-swing right in this context, or put, named down-swing right in case of a swing option, at any one of a number of specified exercise dates. Therefore, a swing option is in turn a generalization of a Bermudan option with multiple exercise rights, combining call and put properties. Typically, there are  $J$  predefined swing opportunities —  $J$  could for example be the number of hours in a year, whereas the number of up-swings is restricted to  $U$  and the number of down-swings to  $D$ . While analytical pricing formulas exist for European options, only numerical calculations are possible in case of swing options.

In this paper we concentrate on power markets and as power plants can be seen as swing options containing only up-swing rights, we deal only with swing options that solely contain up-swing rights. Nevertheless, the algorithm can also be applied to down-swing rights and an extension to combinations of both types of rights is possible. This is important for gas storage, which can be seen as a swing option with up- and down-swing rights.

If there is only one up-swing right contained in the option, i.e.  $U = 1$ , the swing option reduces to a Bermudan call option and if the number of up-swing rights equals the number

of swing opportunities, i.e.  $U = J$ , the swing option is nothing else but a strip of European call options. In all other cases the exercise strategy for the option is non-trivial, as the choice to exercise is always a decision between receiving an immediate payoff, which is certain, and skipping this opportunity in favor of an uncertain, but hopefully larger future payoff. In case of an option with multiple exercise rights, an immediate exercise can be seen as a decision for an immediate payoff plus an option with one right less left for later exercises, i.e. a swing option with less optionality and therefore usually less value.

Over all, the paper is organized as follows: In the first section the different price processes as well as the associated parameter estimation procedures are described. In the second, independent section an efficient Least Squares Monte Carlo algorithm (LSM) is introduced. Finally, the obtained results are discussed in a third section.

## 1 Price Processes

Electricity is a very unique flow commodity, with rather limited storability that requires immediate delivery. On the one hand, the demand shows high variability and dependence on weather conditions, that can only be reliably forecasted a few days ahead. On the other hand, demand is very inelastic to price variations in the short term, as only a few large industrial consumers have the flexibility to change their power consumption in response to market prices. As far as the supply side is concerned, many power plants are only able to vary their power generation with a significant time lag and the overall available generation capacity changes due to maintenance. Finally, power plant outages as well as wind power generation add randomness and complexity as a reliable wind prognosis is only available a few days ahead. Consequently, electricity spot prices exhibit a very high volatility and abrupt, extreme price changes called spikes occur, followed by a reversion back to normal price level within hours or days.

Because of these characteristics, special models for electricity spot prices have been developed. So far, there are roughly three different model classes existing, which are able to reproduce the fat tails associated with jumps and spikes, namely jump-diffusion (Clewlow and Strickland 2000, Geman and Roncoroni 2006, Cartea and Figueroa 2005), regime-switching (Huisman and Mahieu 2003, Schindelmayer 2005, de Jong 2006) and processes based on generalized hyperbolic distributions like the normal inverse gaussian distribution (Benth and Saltyte-Benth 2004, Benth 2007, Weron and Misiorek 2007). The three models for the German spot market EEX used in this paper belong to these three model classes and combine as well as extend the ideas that can be found in the cited literature — especially as an hourly model is necessary for the valuation of hourly swing options.

The EEX spot market is a day-ahead-market, i.e. every day  $d$  hourly power contracts for the 24 hours  $h$  of the next day are traded. Therefore two different views on the time series of spot prices can be taken: It can be regarded as

1. a scalar price process with an hourly granularity that contains a strong auto-correlation for a 24-hour time lag or
2. a vector process with a daily granularity that consists of 24 correlated components.

In this paper as well as in Schindelmayer (2005), the second point of view is taken and the spot market price is described by a discrete stochastic two-factor 24-dimensional vector process  $S_{d,h}$  ( $d = 1, \dots, T$ ,  $h = 1, \dots, 24$ ) with a daily granularity, including short and long term price

variations as stochastic factors:

$$S_{d,h} = \exp(f(d,h) + X_{d,h} + Y_{d,h}) . \quad (1)$$

The stochastic short term factor  $X_{d,h}$  in combination with the deterministic function  $f(d,h)$  generates the short term variations of the spot price, while the factor  $Y_{d,h}$  is responsible for the long term variations. Both factors  $X_{d,h}$  and  $Y_{d,h}$  are – as common (Burger, Klar, Müller and Schindlmayr 2004) – assumed to be stochastically independent.

The short term factor  $X_{d,h}$  is in turn split up into a scalar daily component  $\bar{X}_d$  describing the dynamics of the daily average spot price and an additional vector process consisting of the hourly deviations from the daily average:

$$X_{d,h} = \bar{X}_d + \Delta X_{d,h} . \quad (2)$$

In the same way the deterministic function  $f(d,h)$  is split up in a daily and an hourly part:

$$f(d,h) = \bar{f}(d) + \Delta f(d,h) . \quad (3)$$

In the following subsections, models for the short and long term factor as well as for the deterministic component have to be specified and algorithms for their parameter estimation have to be introduced. Altogether, three different models for the short-term factor  $X_{d,h}$  are used and compared in this paper, namely:

- Model A: combines a regime-switching approach for the daily average and ARMA-processes for the hourly deviations.
- Model B: consists of a jump-diffusion model with Bernoulli jumps instead of regime-switching. The hourly deviations are treated as in model A.
- Model C: makes — in contrast to the first two models — use of the normal inverse gaussian process. Nevertheless, the hourly deviations are again treated as in model A and B.

All three models are based on the same deterministic function and the same long term factor.

### 1.1 Long Term Factor

As we assume a deterministic interest rate framework, it is not necessary to distinguish between forward and futures prices. Therefore, single hour futures prices at time  $t$  for delivery at time  $T = (d,h)$  are conditional expectations under the equivalent martingale measure  $Q$

$$F_{t,T} = \mathbb{E}^Q[S_T | \mathcal{F}_t] , \quad (4)$$

where  $\mathcal{F}_t = \sigma(S_s : s \leq t)$  is the natural filtration generated by the price process. The futures price for a futures contract with a set  $H$  of delivery hours is then given by

$$F_{t,H} = \mathbb{E}^Q \left[ \frac{1}{|H|} \sum_{T \in H} S_T | \mathcal{F}_t \right] = \frac{1}{|H|} \sum_{T \in H} F_{t,T} . \quad (5)$$

In this equation  $|H|$  denotes the number of delivery hours of the futures contracts. In complete markets the  $Q$ -martingale measure is unique, ensuring that only one arbitrage free price for the

futures contract exists. However, there are no derivatives with an hourly granularity traded on the electricity market. Therefore it is an incomplete market, the equivalent martingale measure  $Q$  is not unique and we are left to choose an appropriate measure for the examined market (Cartea and Figueroa 2005).

In our model we assume that the long term process  $\tilde{Y}_t$  follows a random walk with drift given by

$$\tilde{Y}_{t+1} = \tilde{Y}_t + (\mu_t - \frac{1}{2}\sigma_Y^2) + \sigma_Y \varepsilon_t^Y, \quad (6)$$

as also done in Burger et al. (2004). We switch to an equivalent martingale measure  $Q$  assuming a zero market price of risk for the non hedgeable short-term process  $X_t$ . The equation for the long-term process under  $Q$  becomes

$$\tilde{Y}_{t+1} = \tilde{Y}_t + (\mu_t^* - \frac{1}{2}\sigma_Y^2) + \sigma_Y \varepsilon_t^Y, \quad (7)$$

where the new  $\mu_t^* = \mu_t - \lambda_t$  includes the market price of risk  $\lambda_t$ . The conditional expectation of the deterministic function is simply

$$\mathbb{E}^Q[\exp(f(t))|\mathcal{F}_t] = \exp(f(t)). \quad (8)$$

If  $T - t$  is large enough, the conditional distribution can be approximated by the stationary distribution

$$\mathbb{E}^Q[\exp(X_T)|\mathcal{F}_t] \approx \mathbb{E}[\exp(X_T)] \approx \exp(\mathbb{E}[X_T] + \text{Var}[X_T]/2). \quad (9)$$

Due to the fact that we have assumed a zero price of risk for the short term process  $X_t$ , we have been able to use the statistical expectation instead of the expectation under  $Q$ . Neglecting the previous approximation error, the futures prices can be written as

$$F_{t,T} = \hat{S}_T \cdot \mathbb{E}^Q[\exp(Y_T)|\mathcal{F}_t] = \hat{S}_T \exp(\tilde{Y}_t + \sum_{s=t}^{T-1} \mu_s^*) \quad (10a)$$

with

$$\hat{S}_T = \exp(\mathbb{E}[X_T] + \text{Var}[X_T]/2 + f(T)). \quad (10b)$$

The hourly futures price  $F_{t,T}$ , which is also called hourly price forward curve (HPFC) in the electricity sector, plays a central role for the pricing of hourly structures in the electricity OTC and retail markets. As it reflects both – the actual futures prices as well as today's expectation for the spot prices at a future time  $T$  – it is for example often used for the valuation of fixed delivery schedules. But, as we will see later, the HPFC can also be used to calculate a lower boundary for the price of a swing option. A typical approach, that market participants use to construct such an HPFC, is to extract daily and weekly patterns from historical spot prices (e.g. using weighted averages) and use these patterns as weights to break down the quoted forward or futures prices to an hourly granularity. Such an approach is not necessarily consistent with price processes like the ones used in this paper. As equation (10b) shows, the run of the price curve is influenced by the short term factor  $X_T$  (and thus the price process used) and not only by the deterministic component  $f(T)$ .

Starting with the price forward curve, the futures prices for delivery during a set  $H$  of delivery hours results to

$$\begin{aligned} F_{t,H} &= \frac{1}{|H|} \sum_{T \in H} F_{t,T} \\ &= \frac{1}{|H|} \sum_{T \in H} \hat{S}_T \cdot \exp \left( \tilde{Y}_t + \sum_{s=t}^{T-1} \mu_s^* \right). \end{aligned} \quad (11)$$

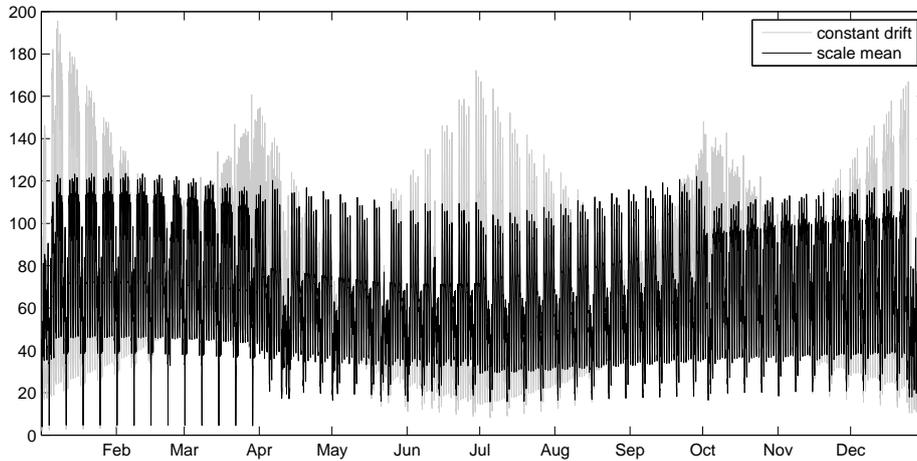
We are still free to choose an appropriate functional form for  $\mu_s^*$  (and thus select an equivalent martingale measure  $Q$ ). If the time series  $\hat{S}_T$  is given,  $\tilde{Y}_t = 0$  is chosen and  $\mu_s^*$  is assumed to be constant for all delivery hours  $H$ , i.e.  $\mu_s^* = \mu_H^* \forall s \in H$ , this nonlinear equation (11) can be solved using a trust region dogleg algorithm (The MathWorks 2008). So far, no futures with overlapping delivery periods have been considered. Unfortunately the matter is more complicated, as base- and peak-futures have to be treated in parallel. This problem can be solved introducing independent  $\mu_{P,s}^*$  and  $\mu_{OP,s}^*$  for peak and offpeak which are zero during offpeak- or rather peak-hours and constant for the delivery hours of the offpeak or rather peak futures contract:

$$\mu_{P,s}^* = \begin{cases} 0, & \forall s \in H_{OP} \\ \mu_{P,H_P}, & \forall s \in H_P \end{cases} \quad (12a)$$

and

$$\mu_{OP,s}^* = \begin{cases} 0, & \forall s \in H_P \\ \mu_{OP,H_{OP}}, & \forall s \in H_{OP}. \end{cases} \quad (12b)$$

This is equivalent to using one time series  $\mu_s^*$  that consists of the sum of  $\mu_{P,s}^*$  and  $\mu_{OP,s}^*$



**Figure 1:** Comparison of the different scaling methods for the HPFC namely constant drift (i.e. using equation (12)) versus scaling the mean (i.e. using equation (14)) for EEX futures prices settled on 28.04.2008

plus correction terms for the marginal peak hours at the crossover from offpeak as well as the

marginal offpeak hour at the crossover from peak. Instead of introducing this correction terms, we prefer to think of two independent time series. As a consequence a system of two equations per base-peak-pair of futures prices

$$F_{t,H_P} = \frac{1}{|H_P|} \sum_{T \in H_P} \hat{S}_T \cdot \exp \left( Y_t + \sum_{s=t}^{T-1} \mu_{P,s}^* \right), \quad (13a)$$

$$F_{t,H_B} = \frac{1}{|H_B|} \sum_{T \in H_B} \hat{S}_T \cdot \exp \left( \tilde{Y}_t + \sum_{s=t}^{T-1} (\mu_{OP,s}^* + \mu_{P,s}^*) \right) \quad (13b)$$

results, that again can be solved using a trust region dogleg algorithm. Depending on the relation of the futures prices  $F_{t,H_P}, F_{t,H_B}$  with subsequent maturities and the level of  $\hat{S}_T$ , a more or less pronounced exponential growth or decay of the price forward curve  $F_{t,T}$  during the different delivery periods of the quoted futures results. It is even possible that strong decay- and growth-periods alternate, as shown in figure 1. Altogether, an unrealistic variation of the price forward curve in time can result. Therefore we prefer another choice of the functional form of  $\mu_s^*$ . If  $T_{P,i}$  and  $T_{OP,i}$  are the beginnings of the delivery periods of the  $i$ -th peak as well as offpeak futures, we choose

$$\mu_{P,s}^* = \begin{cases} \mu_{P,i}, & \text{if } s = T_{P,i} \\ 0, & \text{else} \end{cases} \quad (14a)$$

and

$$\mu_{OP,s}^* = \begin{cases} \mu_{OP,i}, & \text{if } s = T_{OP,i} \\ 0, & \text{else} . \end{cases} \quad (14b)$$

This choice is equivalent to scaling the mean of the forward curve  $F_{t,T}$  over the delivery period of each futures contract to the futures price via one factor for each futures contract (peak as well as base (scaling is done via the offpeak hours)). Using this scaling method, the resulting price forward curve, which is also shown in figure 1, is much more smooth and realistic. If  $F_{H_P,i}^S$  denotes the scaling factor during the delivery hours  $H_{P,i}$  of the  $i$ -th peak futures contract and  $F_{H_{OP,i}}^S$  during the associated offpeak hours  $H_{OP,i}$  (for  $i = 1, \dots, N_F$  quoted futures), the scaling factor can be written as

$$F^S(T) = \sum_{i=1}^{N_F} \left( F_{H_P,i}^S 1_{H_P,i}(T) + F_{H_{OP,i}}^S 1_{H_{OP,i}}(T) \right). \quad (15)$$

Here  $1_{H_P,i}(T)$  and  $1_{H_{OP,i}}(T)$  denote indicator functions, i.e. for example

$$1_{H_{OP,i}}(T) = \begin{cases} 1, & \forall T \in H_{OP,i} \\ 0, & \text{else} . \end{cases} \quad (16)$$

We have implicitly assumed, that the futures prices are totally arbitrage-free and the redundant futures prices (e.g. the yearly contract if all four quarterly contracts are also quoted) have been removed. Using this notation, the price forward curve can be written as

$$F_{0,T} = F^S(T) \cdot \exp(\mathbb{E}[X_T] + \text{Var}[X_T]/2 + f(T)), \quad (17)$$

the long term process becomes

$$Y_{t+1} = Y_t - \frac{1}{2}\sigma_Y^2 + \sigma_Y \varepsilon_t^Y \quad (18)$$

and is assumed to start with  $Y_0 = 0$ . Finally, the price process can be rewritten based on the price forward curve and the modified long term process  $Y_{d,h}$ :

$$S_{d,h} = F_{0,(d,h)} \cdot \exp(-\mathbb{E}[X_{d,h}] - \text{Var}[X_{d,h}]/2 + X_{d,h} + Y_{d,h}) . \quad (19)$$

Altogether, the introduced one factor dynamic of  $Y_t$  has enough degrees of freedom to explain the dynamics of one base futures contract  $F_{t,H}$  as well as all the observed base and peak futures prices on a particular day  $t$  via the choice of the drift function  $\mu_s^*$ .

Burger et al. (2004) show, that only a small error occurs if the parameters of  $X_t$  are estimated assuming  $Y_t = 0$  – instead of using a Kalman filter. Independently from this estimation procedure, the volatility of the long term process can be obtained as intrinsic volatility from quoted option prices. In this paper we use a yearly volatility of  $\sigma_Y = 0.18$ .

In addition to the long term factor, all price processes used in this paper also have the same deterministic component in common.

## 1.2 Deterministic Component

An important input – but not the only one – for the hourly price forward curve (c.f. equations (10a) and (10b)) is the deterministic component. Starting with the historical spot prices for a time interval  $H_S$ , e.g. the last 5 years, a first problem occurs as every now and then zero prices are observed, e.g. on holidays like Christmas. After a replacement of these zero prices by small values like 0.001€/MWh, the logarithm of the spot price time series  $S_{d,h}$  can be taken. The resulting time series  $s_{d,h} = \ln(S_{d,h})$  is split up in a daily average and hourly deviations from this average:

$$s_{d,h} = \bar{s}_d + \Delta s_{d,h}, \quad \forall (d, h) \in H_S . \quad (20)$$

Now, two regression models are introduced to determine the deterministic component:

1. The daily part  $\bar{f}(d)$  of the deterministic function is assumed to consist of  $N_h$  harmonic functions (first, second and possibly third harmonics) and dummy variables for the different day types (i.e. each weekday, holidays, vacation periods):

$$\begin{aligned} \bar{f}(d) = & c_0 + c_1 \cdot (d - d_0) + \sum_{i=2}^7 c_i \delta_{W(d),i} (1 - w_d) + c_8 \cdot w_d + c_9 \delta_{d,\text{CN}} \\ & + \sum_{j=1}^{N_h} (c_{2j+8} \sin(2\pi j \cdot (d - d_0)/365) + c_{2j+9} \cos(2\pi j \cdot (d - d_0)/365)) . \end{aligned} \quad (21)$$

Here  $W(d)$  denotes the number of the weekday of day  $d$  (starting with 1 for Sunday), the origin  $d_0$  is chosen to be the beginning of the first year of the historical spot period  $H_S$ .  $w_d$  is the holiday weight of day  $d$ :

$$w_d = \begin{cases} 0 & \text{if } d \text{ is neither a holiday nor a bridge day} \\ p & \text{if } d \text{ is a holiday (} p \text{ is the part of the population that has a holiday)} \\ g \cdot p & \text{if } d \text{ is a bridge day and } p \text{ the population weight of the appropriate holiday} \end{cases} \quad (22)$$

and

$$\delta_{d,\text{CN}} = \begin{cases} 1 & \text{if } d \text{ is a business day between Christmas and New Year's Eve} \\ 0 & \text{else .} \end{cases} \quad (23)$$

In case of  $N_S = 3$ , 16 coefficients  $c_0, \dots, c_{15}$  have to be fitted. Comparing the  $R^2$  of the resulting models,  $g = 0.4$  is a good choice for the EEX, i.e. a possible and plausible explanation is that about and around 40% of the population who have got a holiday are also on holiday on the appropriate bridge day.

2. The hourly part  $\Delta f(d, h)$  is subdivided into four time series for the quarters of a year and for each quarter modeled using dummy variables for each hour of the day and each weekday (Sundays and Holidays are collected in a single dummy variable):

$$\Delta f(d, h) = c_1^{Q(d)} \cdot (1 - \delta_{W(d),1} \cdot \delta_{h,1}) + \sum_{i=1}^7 \sum_{j=1}^{24} c_{24(i-1)+j}^{Q(d)} \cdot \delta_{W(d),i} \cdot \delta_{h,j} , \quad (24)$$

where  $Q(d)$  is the number of the quarter that day  $d$  belongs to ( $Q = 1, \dots, 4$ ). Altogether this hourly model contains 168 coefficients  $c_1, \dots, c_{168}$ .

The coefficients of these two models are fitted using robust regression, which in MATLAB is nothing else but an iteratively reweighted least squares algorithm that can use different weighting functions — in this paper the Talwar weighting function has been chosen. The advantage in comparison to normal regression (i.e. ordinary least squares) is that due to the iterative reweighting scheme of this algorithm in combination with the Talwar weighting function, outliers are assigned zero weights and therefore these outliers do not disturb the estimation of the deterministic component but end up in the short term component  $X_{d,h}$ . As we have seen, they nevertheless find their way into the price forward curve via the expected value and the variance of the short term component  $X_{d,h}$  (c.f. equation (17)).

After fitting the deterministic function, daily residuals  $r_d$  as well as hourly residuals  $\Delta r_{d,h}$  result which are input for the parameter estimation of the daily and hourly price processes for the short term factor.

### 1.3 Short Term Factor

Business and non-business days show a completely different behavior as far as volatility and occurrence of spikes are concerned. For example a closer look at the spot price history of the last five years reveals, that no positive spikes have been observed on non-business days, whereas on business days both, positive and negative spikes, can be found. This observation is also the reason, why in contrast to Schindelmayer (2005) three regimes are used: one is the normal price regime, one represents the positive spikes and the third one the negative spikes. Of course, in reality the prices on business and non-business days are correlated and not independent, but this effect is neglected in this paper. Nevertheless, this topic might be worth spending some time and effort in the future.

Fitting the price process involves the following two steps:

1. The daily residuals resulting from regression are the input for the estimation of the daily part of the short term price processes. The time series of the daily residuals  $r_d$  is divided

into business (B) and non-business (NB) days:

$$r_d = 1_B(d) \cdot r_d^B + (1 - 1_B(t)) \cdot r_d^{NB} \quad (25)$$

Here  $1_B(d)$  denotes the indicator function

$$1_B(d) = \begin{cases} 1, & d \in B \\ 0, & \text{otherwise.} \end{cases} \quad (26)$$

For either of the processes  $r_d^B$  and  $r_d^{NB}$  the modeling and estimation procedure is identical. Thus the superscript is left out in the following for ease of notation and we work with a time series  $x_t$  ( $t = 1, \dots, N_d$ ) observed on days  $t = d_t$ .

2. Finally, the hourly part of the short term price process is fitted to the hourly residuals. Once again business and non-business days are approximated as being independent:

$$\Delta r_{d,h} = 1_B(d) \cdot \Delta r_{d,h}^B + (1 - 1_B(t)) \cdot \Delta r_{d,h}^{NB} . \quad (27)$$

Again the superscript is left out and we work with a time series  $\Delta x_{t,h}$  ( $t = 1, \dots, N_h$ ,  $h = 1, \dots, 24$ ) observed on days  $t = d_t$ .

In this case the orthogonal transformation of a principal component analysis is used to decompose the 24-dimensional vector process into 24 factor loads that can in turn be modeled as independent ARMA(1,1) processes.

In the following subsections the daily price processes for model A to C are described.

### 1.3.1 Model A: Regime-switching for the Daily Average

In model A the daily average of the spot price is modeled by a regime-switching AR(1)-process with three regimes (normal price mean-reversion regime M, positive S+ and negative S- spike regime). We have chosen three regimes, as the resulting likelihood is larger than the one obtained in case of two regimes. Characteristic for every regime-switching model is, that the transition between the regimes is described by a Markov chain, i.e. there are always all the regimes existing in parallel, but at a particular time only one of the regimes is observable:

$$\text{Regime M with mean-reversion: } x_t = \alpha_1 \cdot x_{t-1} + \mu_1 + \sigma_1 \cdot \varepsilon_t , \quad (28a)$$

$$\text{negative Spike-Regime S-: } x_t = \alpha_2 \cdot x_{t-1} + \mu_2 + \sigma_2 \cdot \varepsilon_t , \quad (28b)$$

$$\text{positive Spike-Regime S+: } x_t = \alpha_3 \cdot x_{t-1} + \mu_3 + \sigma_3 \cdot \varepsilon_t , \quad (28c)$$

$$\text{Markov transition probability matrix: } \Pi = \begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{13} \\ \pi_{21} & \pi_{22} & \pi_{23} \\ \pi_{31} & \pi_{32} & \pi_{33} \end{bmatrix} . \quad (28d)$$

In each case  $\varepsilon_t \sim N(0, 1)$  is normally distributed.

Of course the next step after the definition of the price process is to estimate the parameters from historical spot prices.

### Parameter Estimation for Model A: Hamilton Filter

As only one of the regimes is observable at a certain point in time, the Hamilton filter (Burger, Graeber and Schindelmayer 2007, Hamilton 1994) is suitable for parameter estimation. All the parameters that have to be estimated are collected in a vector  $\theta$ :

$$\theta = (\mu_1, \mu_2, \mu_3, \sigma_1, \sigma_2, \sigma_3, \pi_{11}, \pi_{12}, \pi_{13}, \pi_{21}, \pi_{22}, \pi_{23}, \alpha_1, \alpha_2, \alpha_3). \quad (29)$$

An initial point  $\theta_0$  to start the estimation procedure has to be chosen. If the observed values  $\bar{r}_t$  are given and the values of the parameter vector  $\theta$  are fixed, the likelihood is obtained applying the Hamilton filter. Via a maximization of the log-likelihood  $\mathcal{L}(\theta)$  or equivalent a minimization of the negative log-likelihood eventually the optimal estimated values  $\hat{\theta}$  for the parameters of the price process result, which are shown in table 1 for the regime-switching model.

Parameter	business days regime			non-business days regime		
	1	2	3	1	2	3
$\mu$	-0.000	-0.086	0.078	0.007	-0.109	-0.826
$\alpha$	0.917	1.164	0.570	0.819	0.094	1.420
$\sigma$	0.101	0.245	0.154	0.173	0.318	0.833
$\Pi$	0.924	0.000	0.370	0.929	0.262	0.000
	0.076	0.186	0.516	0.056	0.592	0.926
	0.000	0.814	0.114	0.015	0.146	0.074

**Table 1:** Parameters for the regime-switching model (28.04.2008)

As already mentioned before, the different character of business and non-business days is reflected by the parameters. While on business days regime 3 is more responsible for positive spikes and regime 2 for negative spike as far as the parameter  $\mu$  is concerned, on non-business days two negative spike regimes result that show a different mean-reversion as well as volatility. In any case, the volatility is much higher in the spike-regimes in comparison to the normal price regime 1. As far as the transition probability matrix is concerned, it is interesting that on business days the transition from the normal price regime to the positive spike can only occur via the negative price regime, while a direct jump from the negative spike regime back to the normal price regime is impossible. On non-business days a direct transition into both spike-regimes is possible, but the more probable regime 2 shows a very strong mean reversion as  $\alpha$  is quite small for this regime. In each case the probability for staying in the normal price regime is almost identical.

#### 1.3.2 Model B: Jump-diffusion for the Daily Average

Instead of using different regimes, the up- and down-jumps can also be modeled via special jump-terms added to a mean-reverting price process. For these jump-terms there are different possibilities existing (Cartea and Figueroa 2005, Geman and Roncoroni 2006). Usually the jump-height is assumed to be normally distributed while the arrival of the jumps is described by a poisson process. Due to the fact, that positive as well as negative spikes are quite rare events, we use a Bernoulli process in this paper instead of the Poisson process:

$$x_t = \alpha_{JD} \cdot x_{t-1} + \mu_{JD} + \sigma_{JD} \cdot \varepsilon_t + \kappa^+ \cdot \nu_t^+ + \kappa^- \cdot \nu_t^- \quad (30)$$

with

$$\nu_t^+ \sim \text{Bernoulli}(p_{\text{JD}}^+), \quad \kappa^+ \sim N(\mu_{\text{JD}}^+, (\sigma_{\text{JD}}^+)^2), \quad (31)$$

$$\nu_t^- \sim \text{Bernoulli}(p_{\text{JD}}^-), \quad \kappa^- \sim N(\mu_{\text{JD}}^-, (\sigma_{\text{JD}}^-)^2). \quad (32)$$

The parameters  $p_{\text{JD}}^+$  and  $p_{\text{JD}}^-$  of the Bernoulli distributions are the probabilities for an occurrence of a positive or negative spike. In comparison to the regime-switching model with three regimes as introduced in equations (28), the most significant difference is, that only one mean-reversion parameter  $\alpha_{\text{JD}}$  exists in the jump-diffusion model. Whenever a spike occurs, i.e.  $\nu_t^+ = 1$  or  $\nu_t^- = 1$ , the reversion back to the normal price level can only be driven by this single parameter. By contrast, the regime-switching process inhibits three different mean-reversion rates (one for each regime) and – even more important – three different means. Whenever a jump from a spike regime back to the normal price regime occurs, the mean of the process changes – back to a normal level. This is nothing else but an additional contribution to "mean reversion".

### Parameter Estimation for Model B: Recursive Filtering

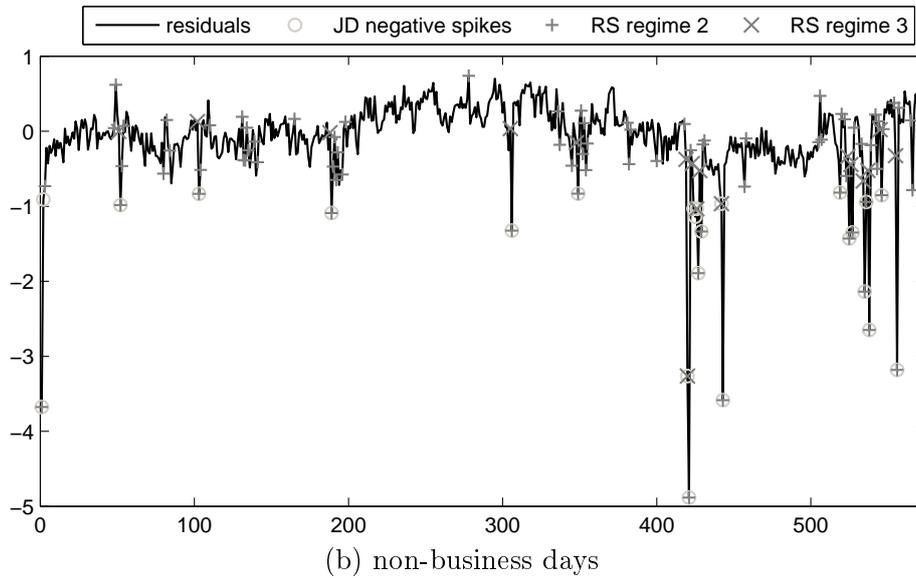
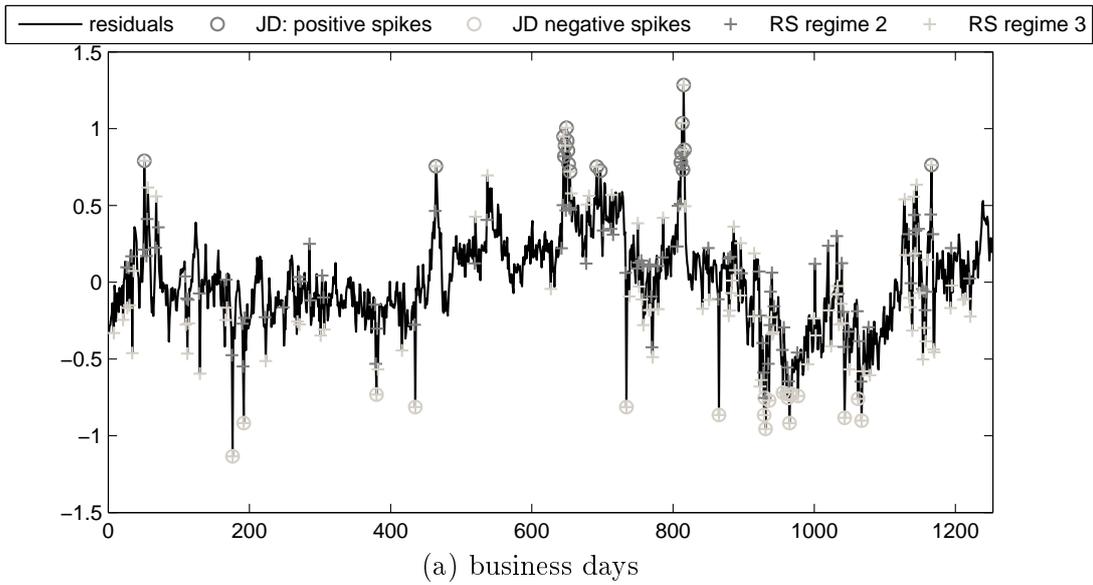
Because of the Bernoulli distribution used in the model, the parameter estimation is rather easily done in this case. In a first step, recursive filtering can be used to detect the spikes (Weron 2006). In this filtering procedure the following steps are recursively taken

1. Calculate the standard deviation  $\sigma_i$  of the time series  $x_t$ .
2. Mark all values greater than  $2.5 \cdot \sigma_i$  as positive spikes, all values smaller than  $-2.5 \cdot \sigma_i$  as negative spikes — but only, if they have not been marked in an earlier recursion step.
3. If no new spikes are found, finish the filtering procedure.
4. Otherwise, extract or replace all marked values. In contrast to what is usually done in the literature, in this paper we have not simply extracted these outliers out of the time series or replaced them by a value of  $2.5 \cdot \sigma_i$ . Instead they are replaced by values from a random walk using the  $\sigma_i$  estimated in the actual recursion step. If for example  $x_t$  is detected to be a spike, it is replaced by  $\hat{x}_t = x_{t-1} + \varepsilon_t$ , with  $\varepsilon_t$  randomly drawn from a normal distribution  $\varepsilon_t \sim N(0, \sigma_i^2)$ . The difference  $x_t - \hat{x}_t$  is saved as the jump-height of the replaced spike.
5. Let  $i \rightarrow i+1$  and goto step 1 with the time series obtained after replacement of the spikes.

An AR(1) model can easily be fitted to the time series resulting from this filtering procedure using simple linear regression. Thus, the parameters  $\mu_{\text{JD}}$ ,  $\alpha_{\text{JD}}$  and  $\sigma_{\text{JD}}$  are obtained.

Finally, the extracted positive and negative spikes are used to estimate the probabilities  $p^+$  and  $p^-$ , which are approximated by the empirical frequencies of the detected outliers. Moreover, the mean and standard deviation of the heights of the extracted outliers are used as estimates for the parameters  $\mu^+$  and  $\sigma^+$  for the positive outliers as well as for  $\mu^-$  and  $\sigma^-$  in case of negative outliers.

Table 2 shows the parameters found for business and non-business days. For a spot history that consists of the last five years no positive spikes are detected on non-business days and the negative spikes show a large absolute value of the mean  $\mu_{\text{JD}}^-$ , while on business days the probability for positive and negative spikes are comparable and the absolute values of the



day type	jump-diffusion		regime-switching	
	positive	negative	regime 2	regime 3
business days	19	21	105	115
non-business days	0	24	78	17

(c) number of spikes detected

**Figure 2:** Comparison of the spike identification via Hamilton filter and recursive filtering

	$p_{JD}^+$	$\mu_{JD}^+$	$\sigma_{JD}^+$	$p_{JD}^-$	$\mu_{JD}^-$	$\sigma_{JD}^-$	$\mu_{JD}$	$\alpha_{JD}$	$\sigma_{JD}$
business days	0.015	0.292	0.337	0.017	-0.248	0.416	0.001	0.857	0.146
non-business days	0.000	0.000	0.000	0.035	-1.409	1.311	-0.008	0.648	0.255

**Table 2:** Parameter estimated for the jump-diffusion model using recursive filtering (28.04.2008)

means  $\mu_{JD}^+, \mu_{JD}^-$  are much smaller than on non-business days. Therefore it makes very much sense to use two jump-terms on business days (whereas one term would be enough on non-business days). Finally, the mean-reversion  $\alpha_{JD}$  is comparable to the average one found for the regime-switching price process and therefore somehow at the same time right and wrong.

In figure 2 the identification of the spikes via recursive filtering (jump-diffusion) and via the Hamilton filter (regime-switching) are compared. Obviously, the Hamilton filter identifies many more residuals as spikes than recursive filtering with  $2.5 \cdot \sigma$ . While the recursive filtering procedure only finds the largest outliers, and even misses some that seem to be spikes, the Hamilton filter is able to identify all possible "spikes" – maybe even more than obvious. From these observations it can be concluded, that both models are based on a different implicit definition of a spike. While in model B a spike is an outlier and therefore defined through its mean and rareness, in model A a spike results from another regime and is characterized by another volatility, mean as well as mean reversion rate. After all, this is rather different.

### 1.3.3 Model C: Normal Inverse Gaussian Process for the Daily Average

While the first two models are based on the normal distribution and the heavy tails of the residuals caused by spikes and jumps are reproduced via regime-switching or jump-terms, the third model uses the normal inverse gaussian distribution – a generalized hyperbolic distribution that is by itself flexible enough to reproduce heavy tails. In model C we assume, that the short term factor is NIG distributed:

$$\bar{x}_d \sim \text{NIG}(\alpha_{\text{nig}}, \beta_{\text{nig}}, \delta_{\text{nig}}, \mu_{\text{nig}}) \quad (33)$$

The four parameters  $\alpha_{\text{nig}}, \beta_{\text{nig}}, \delta_{\text{nig}}$ , and  $\mu_{\text{nig}}$  can easily be estimated in MATLAB using the routine `nigest`, which uses maximum likelihood estimation and is part of the MFE toolbox (Weron 2006, Weron 2007).

	$\alpha_{\text{nig}}$	$\beta_{\text{nig}}$	$\delta_{\text{nig}}$	$\mu_{\text{nig}}$
business days	5.617	0.197	0.504	-0.017
non-business days	2.752	-1.372	0.351	0.128

**Table 3:** Parameters resulting for the NIG-process using maximum likelihood estimation (28.04.2008)

In table 3 the results of the parameter estimation are shown. Again, completely different parameters are obtained for business and non-business days. On business days the distribution is slightly skewed to the right, while on non-business days strongly to the left. The tail heaviness is larger on business than on non-business days. Finally, the volatility is comparable to what is found in case of the other price processes.

### 1.3.4 The Hourly Profile Process

All three models described so far are daily models. All of them have to be combined with an hourly model. In each case ARMA(1,1) processes can be used as basis for the hourly model, in case of model A and B supplemented by sampling historical spike profiles.

#### Sampling of Spike Profiles

An analysis of historical profiles for spike days reveals, that the daily average on spike days is not simply higher while the hourly profile is the same as on days with normal price levels. In fact, the profiles of the hourly deviations from the average can be completely different in comparison to normal days. Furthermore, seasonal differences in the spike behavior can be observed and therefore the historical profiles have to be classified as summer and winter profiles according to the season during the occurrence of the spike (summer is defined as April to September, winter the rest of the year). Altogether,  $N_{sum}^+$  positive and  $N_{sum}^-$  negative summer profiles ( $p_{i,h}^{sum,+}, p_{i,h}^{sum,-}$  with  $h = 1, \dots, 24$ ) result from this sampling as well as  $N_{win}^+$  positive and  $N_{win}^-$  negative winter profiles ( $p_{i,h}^{win,+}, p_{i,h}^{win,-}$ ).

On days where model A is in the positive spike regime, a positive historical winter or summer spike profile is randomly drawn and on days with a negative spike regime a negative historical profile. In case of model B the same is done, if the corresponding Bernoulli term equals 1. Of course nothing comparable can or has to be done for model C. The historical spike profiles result during the parameter estimation of the daily process. Within model A, for every day of the price history the Hamilton filter assigns a probability to each regime. Accordingly, the corresponding hourly profile is assumed to be a positive spike if the positive spike regime is the most probable one. The same is done for negative spikes. In case of model B, the recursive filtering procedure automatically classifies the outliers as positive or negative spikes and the same can be done for the corresponding hourly profiles.

#### ARMA process

For all other days, which have not been classified as spike days, the hourly residuals are collected in a 24 dimensional time series. The components of this vector process reflect the volatilities of each single hour as well as the correlations between the time series' for the different hours. As already done for the daily average the process is again separated in business and non-business days:

$$\Delta \mathbf{r}_d = 1_B(d) \cdot \Delta \mathbf{r}_d^B + (1 - 1_B(t)) \cdot \Delta \mathbf{r}_d^{NB}. \quad (34)$$

In order to simplify the notation, only one of these two processes is examined and denoted  $\Delta \mathbf{r}_k$ ,  $k = 1, \dots, N_k$ . As done in a principal component analysis (pca) (Burger et al. 2007) for each column  $i$ , with  $i = 1, \dots, 24$ , the mean  $\hat{\mu}_{pca,i}$  is subtracted and afterwards the row is divided by its standard deviation  $\hat{\sigma}_{pca,i}$ :

$$\Delta \tilde{r}_k^i = \frac{\Delta r_k^i - \hat{\mu}_{PCA,i}}{\hat{\sigma}_{PCA,i}} \quad (35)$$

with:

$$\hat{\mu}_{PCA,i} = 1/N_k \sum_{k=1}^{N_k} \Delta r_k^i, \quad (36)$$

$$\hat{\sigma}_{PCA,i} = \sqrt{\frac{1}{N_k} \sum_{k=1}^{N_k} (\Delta r_k^i - \hat{\mu}_{PCA,i})^2}. \quad (37)$$

For the calculation of the orthogonal transformation matrix the components of the resulting normalized vector process are collected in the  $(N_k \times 24)$ -matrix

$$\Delta \tilde{\mathbf{r}} = \begin{pmatrix} \Delta \tilde{\mathbf{r}}_1^T \\ \vdots \\ \Delta \tilde{\mathbf{r}}_{N_k}^T \end{pmatrix}. \quad (38)$$

The transformation matrix  $\mathbf{Q}$  is nothing else but the matrix of the eigenvectors of the eigenvalue problem

$$((\Delta \tilde{\mathbf{r}})^T \cdot \Delta \tilde{\mathbf{r}}/N_k) \cdot \mathbf{Q} = \Lambda \cdot \mathbf{Q}. \quad (39)$$

After an application of this orthogonal transformation 24 independent processes result which are collected in the matrix  $\mathbf{w}$

$$\mathbf{w} = \Delta \tilde{\mathbf{r}} \cdot \mathbf{Q}. \quad (40)$$

For each time series  $w_k^i$  with  $k = 1, \dots, N_k (i = 1, \dots, 24)$ , that can be found in the  $i$ -th column of the matrix  $\mathbf{w}$  an ARMA(1,1)-process

$$y_t^i = C_{ARMA}^i + \phi_{AR}^i y_{t-1}^i + \varepsilon_t^i + \theta_{MA}^i \varepsilon_{t-1}^i \quad (41)$$

with  $\varepsilon_t^i \sim N(0, \sigma_{ARMA,i}^2)$  is fitted using the GARCH-toolbox in MATLAB. For the scenario generation the single steps of this transformation have to be run through in reverse order to obtain correlated scenarios.

Thus, all the models and their parameter estimations are specified – only the variance corrections are still missing.

### 1.3.5 Variance Corrections for the HPFC

For each model the expected value and variance corrections for the price forward curve have to be calculated, which are an additional input for the HPFC (c.f. equation (17)).

Model A: In this case the corrections consist of three components, resulting from the mean and variance of the daily model, the historical spike profiles and the ARMA processes. Altogether the following term results if  $T = (d, h)$  lies in the summer half year (superscript "sum")

$$\mathbb{E}[x_{d,h}] + \text{Var}[x_{d,h}]/2 = \sum_{i=1}^3 (p_i \cdot \mu_i + p_i^2 \cdot \sigma_i^2 / (1 - \alpha_i^2) / 2) \quad (42)$$

$$+ p_2 \cdot \mu_h^{sum,+} + p_2^2 \cdot (\sigma_h^{sum,+})^2 / 2 \quad (43)$$

$$+ p_3 \cdot \mu_h^{sum,-} + p_3^2 \cdot (\sigma_h^{sum,-})^2 / 2 + \mu_h \quad (44)$$

Where  $\mathbf{p} = (p_1, p_2, p_3)^T$  denotes the ergodic probabilities of the three different regimes (1: mean-reversion, 2,3: spike-regimes) and the profile mean and variance are calculated as follows:

$$\mu_h^{sum,+} = \frac{1}{N^{sum,+}} \sum_{i=1}^{N^{sum,+}} p_{i,h}^{sum,+} ; \quad (\sigma_h^{sum,+})^2 = \sum_{i=1}^{N^{sum,+}} \frac{(p_{i,h}^{sum,+} - \mu_h^{sum,+})^2}{N^{sum,+}} \quad (45)$$

$$\mu_h^{sum,-} = \frac{1}{N^{sum,-}} \sum_{i=1}^{N^{sum,-}} p_{i,h}^{sum,-} ; \quad (\sigma_h^{sum,-})^2 = \sum_{i=1}^{N^{sum,-}} \frac{(p_{i,h}^{sum,-} - \mu_h^{sum,-}(p_1))^2}{N^{sum,-}} \quad (46)$$

and the hourly process produces the correction term

$$\begin{aligned} \mu_h(p) = p & \left( \hat{\mu}_{PCA,h} + \hat{\sigma}_{PCA,h} \cdot \sum_{l=1}^{24} q_{h,l} \cdot C_{ARMA}^l \right) \\ & + p^2 \cdot \hat{\sigma}_{PCA,h}^2 \cdot \sum_{l=1}^{24} q_{h,l}^2 \cdot \frac{1 + (\theta_{MA}^l)^2 + 2\phi_{AR}^l \theta_{MA}^l}{1 - (\phi_{AR}^l)^2} \cdot \sigma_{ARMA,l}^2 \end{aligned} \quad (47)$$

In this equation  $q_{h,l}$  is an Element (row h, column l) of the transformation matrix  $Q$ . Replacing the mean and variance resulting from the historical spike profiles by the ones calculated from winter profiles, an analogue result can be obtained for the winter half year.

Model B: Again three components result and the third one is the same as shown in equation (47). Altogether, the following expression results

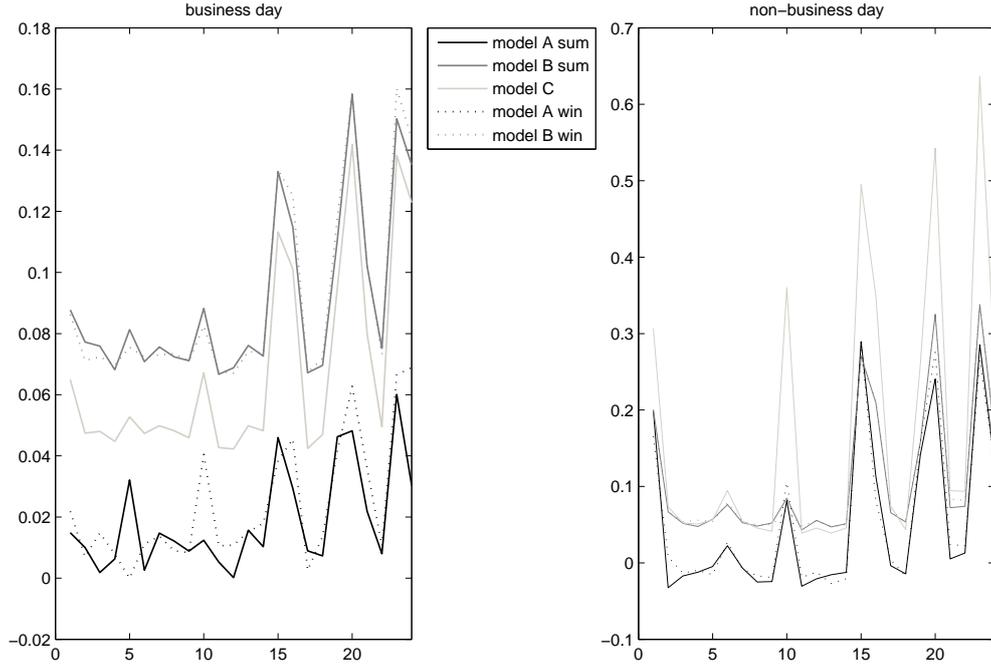
$$\begin{aligned} \mathbb{E}[x_{d,h}] + \text{Var}[x_{d,h}]/2 = & (\mu_{JD} + p_{JD}^+ \cdot \mu_{JD}^+ + p_{JD}^- \cdot \mu_{JD}^-) \\ & + \frac{\sigma_{JD}^2 + (p_{JD}^+)^2 \cdot (\sigma_{JD}^+)^2 + (p_{JD}^-)^2 \cdot (\sigma_{JD}^-)^2}{2(1 - \alpha_{JD}^2)} \\ & + p_{JD}^+ \cdot \mu_h^{sum,+} + (p_{JD}^+)^2 \cdot (\sigma_h^{sum,+})^2 / 2 \\ & + p_{JD}^- \cdot \mu_h^{sum,-} + (p_{JD}^-)^2 \cdot (\sigma_h^{sum,-})^2 / 2 \\ & + \mu_h(1 - p_{JD}^+ - p_{JD}^-) . \end{aligned} \quad (48)$$

As the sampling of historical profiles is comparable to regime-switching, the expression (45) for the mean and variance still holds – of course calculated using different spike profiles (e.g. the number of spikes detected differs).

Model C: In the third model the resulting expression is more simple, as no historical sampling is used:

$$\mathbb{E}[x_{d,h}] + \text{Var}[x_{d,h}]/2 = \mu_{nig} + \frac{\delta_{nig} \cdot \beta_{nig}}{\sqrt{\alpha_{nig}^2 - \beta_{nig}^2}} + \frac{\delta_{nig} \cdot \alpha_{nig}^2}{2 \cdot \sqrt{(\alpha_{nig}^2 - \beta_{nig}^2)^3}} + \mu_h(1) \quad (49)$$

Again, the corrections  $\mu_h$  resulting from the hourly price process can be calculated using equation (47), but now with  $p = 1$ .

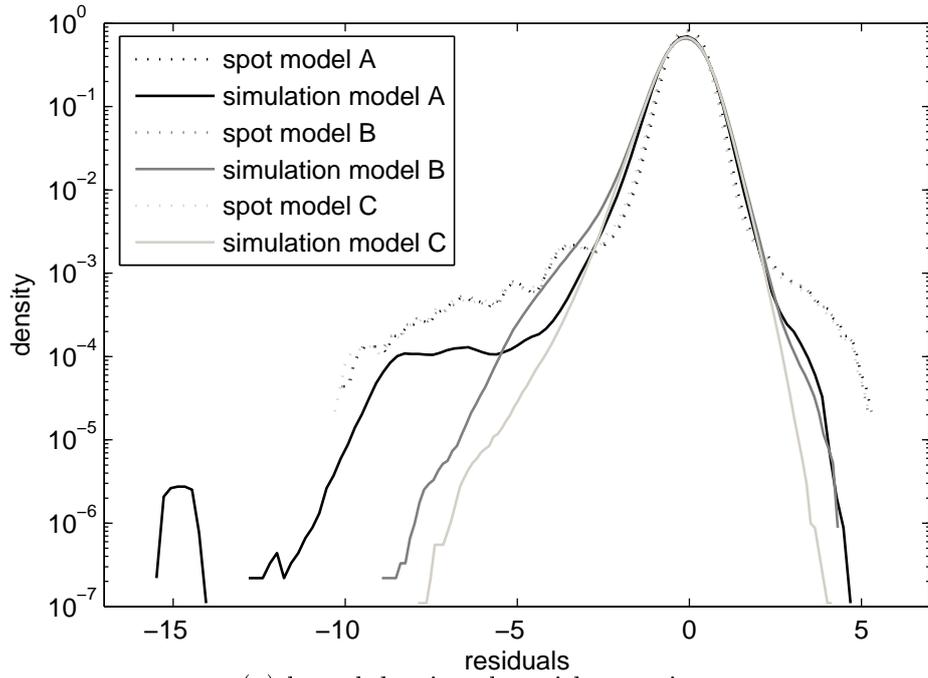


**Figure 3:** Corrections for the HPFC on business and non-business days

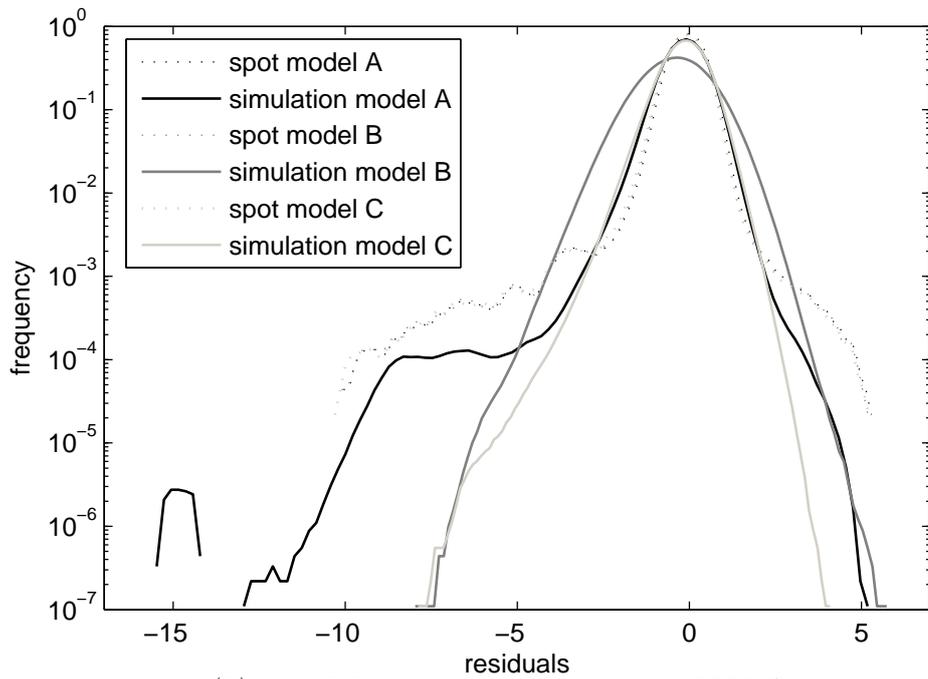
In each model different corrections for business and non-business days result, which can be calculated according to the same equations using the parameters estimated for the different day types. The resulting corrections are shown in figure 3. Obviously, the corrections in winter and summer do not differ very much.

#### 1.4 Generation of price scenarios

Using models A to C as well as the estimated parameters for these models, any given number of price scenarios can be generated for a delivery period in the future. For example 1000 scenarios have been generated for the year 2009 using the different models, the residuals  $r_{d,h} = \ln(S_{d,h}/F_{0,(d,h)})$  (starting with eq. (19)) have been calculated and the density has been estimated using a kernel density with a box kernel and a window width of 0.3. The results are compared in figure 4 and they are contrasted to the spot residuals obtained during the fit of the deterministic component including the variance corrections. Of course the spot residuals and the ones resulting from the simulation differ – as they represent periods with different price levels they should do so. But nevertheless the shape of the density should be the same. From this criterion model A is the best spot price model – it fits the quadratic functional form around zero quite good and it is able to reproduce the fat tails and the asymmetry caused by negative as well as positive spikes. Model B and C show comparable results - while model C fits the quadratic functional form better, model B reproduces the positive spikes better. But they do not perform as good as model A. In all the models very high spike prices, reaching 10.000 €/MWh and more, result that are extremely unrealistic — no trader is ever willing to pay such high prices. Therefore, the price generation has been modified by introducing a price



(a) kernel density plot without price cap



(b) kernel density plot with price cap 3000 €

**Figure 4:** Scenario generation using the three specified models with and without a price cap. The spot residuals are taken from the last five years while the scenarios are generated for 2009.

cap, e.g. 3000 €/MWh. In figure 4 (b) the resulting densities obtained after an introduction of this price cap are shown. Nevertheless, this analysis is strictly speaking nothing else but an in-sample analysis. Beyond it, model A has to prove its performance also in an out-of-sample analysis.

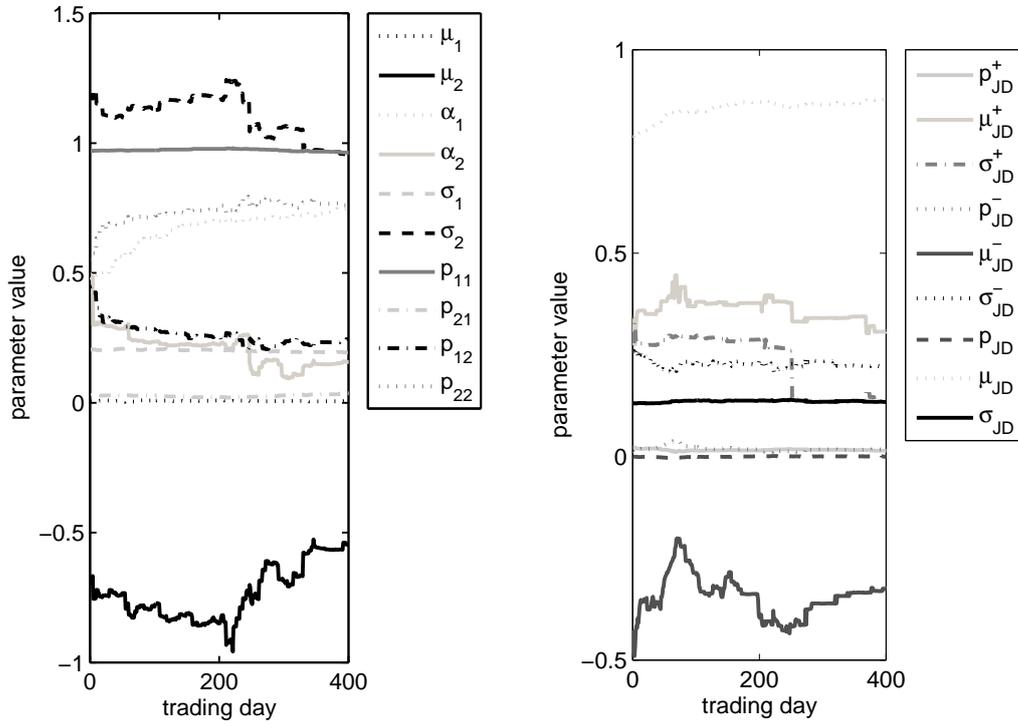
## 1.5 Out-of-sample Analysis

One important question is the stability of the parameters estimated with the algorithms described above. Thus, the parameters for the three models have been estimated for a period of 400 trading days starting with 2.1.2007. In figure 5 (a)-(c) the resulting parameter values for the three models are plotted against the trading days. For the purpose of clearness, the parameters for the non-business days are shown exemplarily for model A. Only two regimes are chosen in the case of non-business days and therefore the transition matrix contains only 4 elements, the whole model includes a total of 10 parameters (and not 18 as on business days). Nevertheless the results for business days have also been examined and are comparable as far as parameter stability is concerned. Over all, the parameter stability of the three models is comparable and the parameters fluctuate not too much over time – except for the parameter  $\alpha_{NIG}$  in model C describing the tail heaviness, which varies over a wide range during the examined 400 trading days. From this point of view none of the three models is really outstanding, only model C falls slightly back.

In figure 6 the resulting distribution of price deviations from realized spot prices are shown for the three models. For this purpose 1000 scenarios have been generated for the period 1.1.2007-31.12.2008 as resulting from a parameter-estimation based on spot and futures prices up to 28.12.2006 as well as for the period 1.1.2008-31.12.2008 based on spot and futures prices up to 28.12.2007. For each scenario the observed spot prices (which have not been included in the parameter estimation) have been subtracted and the density of the price deviations has been estimated using a kernel density with a box kernel and a window width of 0.3. An ideal model would show no price deviations from the spot price at all and the density plot would be a flat line. The closer the models are to the flat line, the better they perform in predicting the future spot prices. Thus, model A and C perform comparably good, only model B clearly overemphasizes the high price tail.

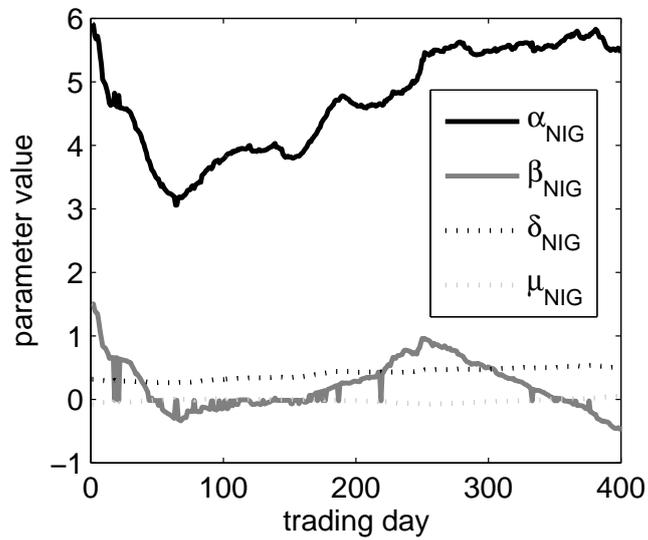
## 1.6 Correlation Structure and Volatilities of Forward Prices

As the models consider the spot price dynamics, another important question is, how well they explain the volatilities of the forward prices and their correlation structure. Figure 7 shows the volatilities resulting from model A to C in comparison to option implied volatilities calculated from option prices quoted at the EEX. In this figure only the implied volatilities of at-the-money options are shown, as the models are not able to reproduce volatility smiles (see also (Kiesel, Schindlmayr and Boerger 2007)). It can be seen that the implied volatilities are quite well reproduced by model A and even a bit better by model B - in contrast to model C that underestimates the volatilities. Nevertheless, the correlation structure between forward prices is not reproduced by any of the three models, as can be seen in table 4. For model B the correlation between different forwards is mostly close to zero and for model A only for forwards with delivery far in the future the correlation becomes larger and is closer to the empirical values. The results for model C are not shown in the table, but they are similar to those for model A.



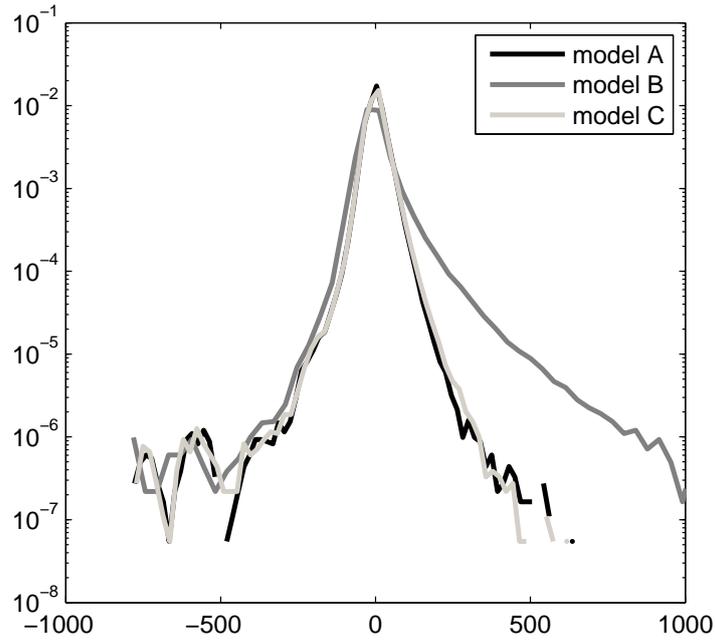
(a) Model A (regime-switching)

(b) Model B (jump-diffusion)

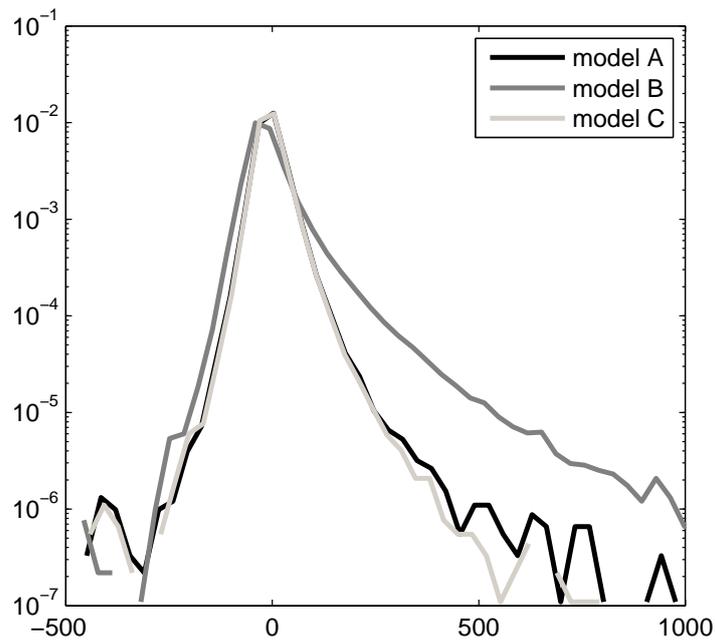


(c) Model C (NIG)

**Figure 5:** Parameters estimated for the different models over a period of 400 trading days

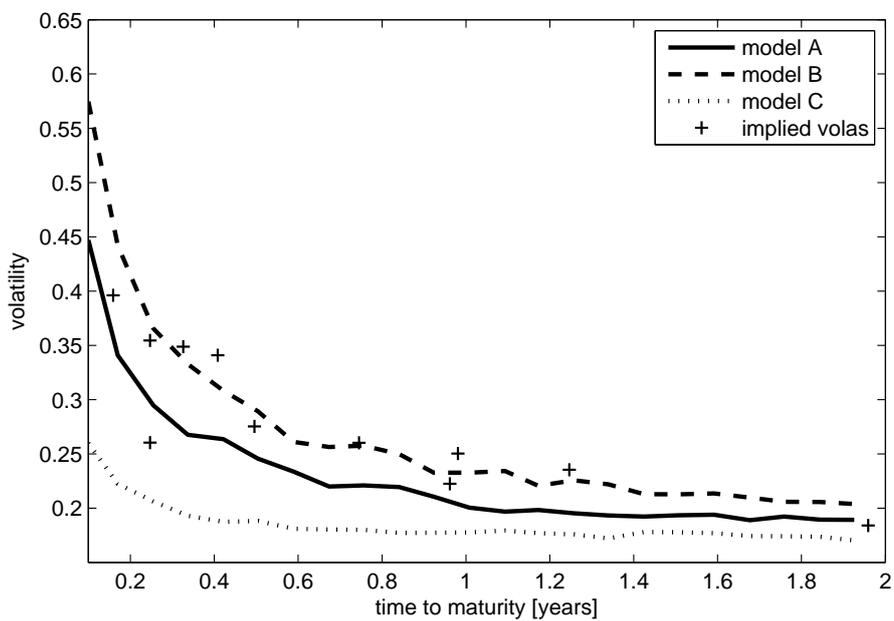


(a) Parameter estimation for business day 28.12.2006, price deviations for time period 1.1.2007-31.12.2008

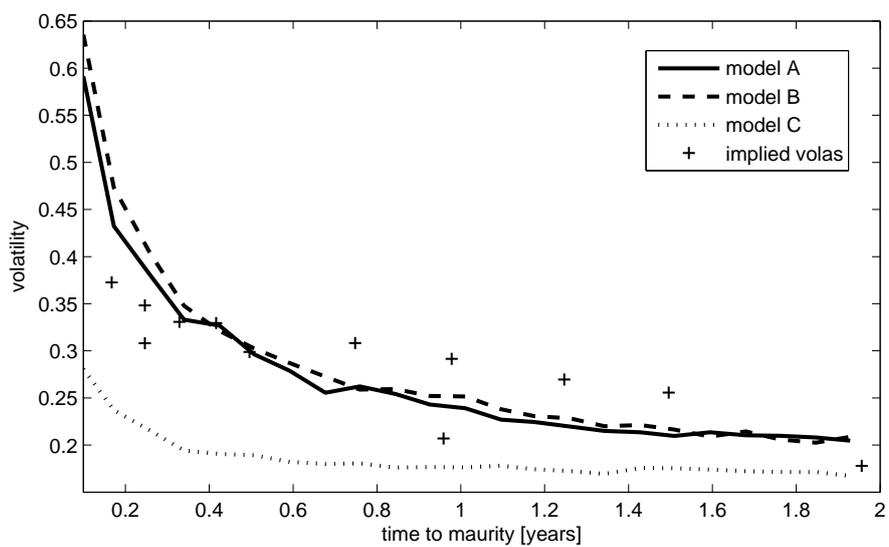


(b) Parameter estimation for business day 28.12.2007, price deviations for time period 1.1.-31.12.2008

**Figure 6:** Distribution of the price deviations of price scenario generated for a future period from observed spot prices for the three models



(a) Volatilities of futures prices compared to implied volatilities resulting for business day 28.12.2006



(b) The same as in (a) for business day 28.12.2007

**Figure 7:** Volatilities of the futures prices resulting from the three specified models in comparison to implied volatilities calculated from option prices.

(a) model correlation matrix (model A: regime switching)

	Jan07	Feb07	Mar07	Q207	Q307	Q407	Cal08
Jan07	1.00	0.26	0.07	0.13	0.12	0.12	0.12
Feb07	0.26	1.00	0.29	0.27	0.26	0.21	0.18
Mar07	0.07	0.29	1.00	0.45	0.34	0.30	0.26
Q207	0.13	0.27	0.45	1.00	0.63	0.58	0.48
Q307	0.12	0.26	0.34	0.63	1.00	0.71	0.61
Q407	0.12	0.21	0.30	0.58	0.71	1.00	0.75
Cal08	0.12	0.18	0.26	0.48	0.61	0.75	1.00

(b) model correlation matrix (model B: jump-diffusion)

	Jan07	Feb07	Mar07	Q207	Q307	Q407	Cal08
Jan07	1.00	0.08	0.13	0.05	0.03	0.04	0.01
Feb07	0.08	1.00	0.07	0.17	0.01	0.10	0.07
Mar07	0.13	0.07	1.00	0.24	0.06	0.22	0.18
Q207	0.05	0.17	0.24	1.00	-0.00	0.42	0.41
Q307	0.03	0.01	0.06	-0.00	1.00	-0.04	-0.04
Q407	0.04	0.10	0.22	0.42	-0.04	1.00	0.69
Cal08	0.01	0.07	0.18	0.41	-0.04	0.69	1.00

(c) empirical correlation matrix

	Jan07	Feb07	Mar07	Q207	Q307	Q407	Cal08
Jan07	1.00	0.98	0.94	0.93	0.84	0.87	-0.18
Feb07	0.98	1.00	0.94	0.95	0.90	0.88	-0.12
Mar07	0.94	0.94	1.00	0.96	0.94	0.81	0.04
Q207	0.93	0.95	0.96	1.00	0.91	0.83	-0.08
Q307	0.84	0.90	0.94	0.91	1.00	0.80	0.24
Q407	0.87	0.88	0.81	0.83	0.80	1.00	-0.01
Cal08	-0.18	-0.12	0.04	-0.08	0.24	-0.01	1.00

**Table 4:** Comparison of empirical and model forward correlation matrix (28.12.2006)

This wrong correlation structure results from the fact that the short and long term factors are assumed to be uncorrelated. Thus, there is still potential for improvements of the models as far as this topic is concerned. For example additional factors (as done in a forward model using a PCA by Koekebakker and Ollmar (2005)) and a correlation between short and long term factors could be introduced. In this case a parameter estimation is much more complicated and use of a Kalman filter would be inevitable.

Over all, model A performs best in-sample as well as out-of-sample and therefore generates the most realistic price paths among the three examined models. After we are able to generate realistic price paths, we can turn our attention to the pricing of swing options.

## 2 Least Squares Monte Carlo

As already mentioned in the introduction, the main issue of pricing American or Bermudan options is the early exercise decision, which is non-trivial. The optimal exercise decision relies on the estimation of the continuation value. Longstaff and Schwartz (2001) have proposed to estimate this continuation value using least squares regression in combination with a Monte Carlo simulation of the possible price paths. Their method fits the future cashflows using for example simple polynomial basis functions. The resulting Least Squares Monte Carlo (LSM) algorithm is able to value Bermudan options with a single exercise right, which are an approximation of an American option and can be seen as a special case of a swing option with one exercise right. This fundamental algorithm has been extended for the valuation of swing options with multiple exercise rights (Dörr 2003, d-fine 2004) and has recently been applied to gas storage (Boogert and de Jong 2008). Other methods to value swing options are for example based on finite element methods (Dahlgren 2005), which are fast for simple price processes but difficult or even impossible to implement for complicated price processes like the ones used in this paper. Keppo (2004) shows that swing options can be replicated with regular electricity forwards and call options in case of a complete market; in this paper also lower boundaries for swing option values have been calculated for selected swing options.

In the following section the fundamental LSM algorithm as well as the extensions for multiple exercise rights are described.

The possible exercise times of the Bermudan option are denoted

$$0 < t_1 < t_2 < \dots < t_J = T . \quad (50)$$

In each case the option is exercised if the resulting immediate payoff is larger than the continuation value, which is the value of the remaining option with a shorter lifetime and the remaining number of exercise rights left. This continuation value can be written as the conditional expectation under the risk-neutral measure:

$$F(\omega, t_k) = \mathbb{E}^Q \left[ \sum_{j=k+1}^J D(t_k, t_j) \cdot C_k(\omega, t_j) \middle| \mathcal{F}_{t_k} \right] ; \quad (51)$$

where  $D(t_k, t_j) = \exp(-r \cdot (t_k - t_j))$  denotes the discount factor from time  $t_k$  to  $t_j$  with the risk-free rate  $r$  and  $C_k(\omega, t_j)$  the cashflows generated by an application of the optimal exercise strategy if the option is not exercised at time  $t_k$ . For each price scenario  $\omega$ , at most one index  $j$  with  $C_k(\omega, t_j) > 0$  can exist, as the Bermudan option contains only one exercise right. Thus, most elements of the  $N \times J$ -matrix  $C_k(\omega, t_j)$  are zeros and a much more compact

storage is possible if only the discounted cashflows are saved in each price path. In this case the information, which exercise opportunities are used (this information could even be stored in an additional index-vector), is lost, but as a big advantage only an  $N$ -dimensional vector is necessary to keep track of the discounted cashflows. For Bermudan options with one exercise right computing time and memory usage is not an issue, but this changes later on if we try to value swing options with a large number of exercise rights.

It makes very much sense to start with  $t_J$ , and iteratively step back in time until  $t_1$  is reached, because at time  $t_J$  no future decisions exist and the immediate cashflow for an exercise can be calculated using

$$C_J(\omega, t_J) = P(S_J(\omega)) , \quad (52)$$

where  $S_J(\omega)$  is the spot price at time  $t_J$  in the price path  $\omega$  and  $P$  the payoff function. For a call this payoff function  $P$  is simply

$$P(S_J(\omega)) = \max(S_J(\omega) - K, 0) , \quad (53)$$

where  $K$  denotes the strike price of the option.

At time  $t_{J-1}$  the exercise decision is done comparing the future cashflow  $F(\omega, t_{J-1})$  and the immediate payoff  $P(S_{J-1}(\omega))$ ; while the payoff is well known, the functional form of the continuation value  $F(\omega, t_{J-1})$  is unknown. For this reason the conditional expectation  $F$  is expanded in a set of basis functions  $B_l$ :

$$F(\omega, t_k) = \sum_{l=0}^{\infty} a_l(t_k) \cdot B_l(S_k(\omega)) , \quad (54)$$

with  $B_l = x^l$  .

In practice the infinite sum is cut off after a small number of  $M$  basis functions, e.g.  $M = 5$  is used in this paper (see below for an examination of the influence of the choice of  $M$  on the option value), which is larger than the  $M = 3$  found in the literature (Dörr 2003, d-fine 2004):

$$\hat{F}(\omega, t_k) = \sum_{l=0}^M a_l(t_k) \cdot B_l(S_k(\omega)) . \quad (55)$$

As a memoryless Markov process is assumed for the state variable  $S_k$ , only realizations at time  $t_k$  (and not times  $t_l < t_k$ ) occur as arguments of the basis functions. For each of the three price processes used in this paper this assumption is valid. The coefficients  $a_l$  can be calculated using linear regression:

$$\min \sum_{\omega=1}^N \left( \sum_{j=k+1}^J D(t_J, t_k) \cdot C_k(\omega, t_j) - \sum_{l=0}^M a_l B_l(S_{J-1}(\omega)) \right)^2 . \quad (56)$$

In time step  $t_{J-1}$  this reduces to:

$$\min \sum_{\omega=1}^N \left( D(t_J, t_{J-1}) \cdot C_{J-1}(\omega, t_J) - \sum_{l=0}^M a_l B_l(S_{J-1}(\omega)) \right)^2 . \quad (57)$$

With known coefficients  $a_l$  the continuation value  $F$  is given by

$$\hat{F}(\omega, t_k) = \sum_{l=0}^M a_l B_l(S_k(\omega)) , \quad (58)$$

or specifically for time step  $t_{J-1}$ :

$$\hat{F}(\omega, t_{J-1}) = \sum_{l=0}^M a_l B_l(S_{J-1}(\omega)) . \quad (59)$$

Using this approximation for the continuation value, the option can be exercised optimally under uncertainty: For every price path  $\omega$ , where the option is in the money, the cashflow for the appropriate time step  $t_k$  can be calculated comparing the continuation value  $\hat{F}(\omega, t_k)$  and the immediate payoff  $P(S_k(\omega))$  resulting by an immediate exercise:

$$C_k(\omega, t_k) = \begin{cases} P(S_k(\omega)) , & \text{if } P(S_k(\omega)) > \hat{F}(\omega, t_k) \\ 0 , & \text{else .} \end{cases} \quad (60)$$

For all price paths  $\omega$ , where the option is exercised (i.e.  $C_k(\omega, t_k) > 0$ ) the resulting cashflows have to be placed in the cashflow vector, while otherwise the discounted future values are kept.

After the last iteration step (at time  $t_1$ ) the option value eventually results as the average

$$V = \frac{1}{N} \sum_{\omega=1}^N \sum_{j=k+1}^J D(t_j, t_k) \cdot C_1(\omega, t_j) . \quad (61)$$

Altogether, the algorithm can be summed up as follows: For an initialization of the iteration at time step  $t_J$  the cashflows  $C_J(\omega, t_J)$  have to be calculated for all price paths  $\omega$  and stored in the cashflow vector  $\mathbf{C}(t_J)$ . Subsequently, for time steps  $t_k = t_{J-1}, \dots, t_1$  the following steps are taken:

1. The cashflows in the vector  $\mathbf{C}(t_{k+1})$  are discounted to the time step  $t_k$  by multiplication with the discount factor  $D(t_{k+1}, t_k)$ :

$$\mathbf{C}(t_k) = D(t_{k+1}, t_k) \cdot \mathbf{C}(t_{k+1}) .$$

2. The cashflow vector  $P(\mathbf{S}_k)$  for an immediate exercise is calculated. For a plain vanilla call option the payoff function is simply  $P(S_k(\omega)) = \max(S_k(\omega) - K, 0)$ .
3. For all price paths  $\omega$ , where the option is in the money ( $P(S_k(\omega)) > 0$ ), the continuation value  $\hat{\mathbf{F}}(t_k)$  is calculated using linear regression:

$$\hat{\mathbf{F}}(t_k) = \sum_{l=0}^M a_l B_l(\mathbf{S}_k) .$$

4. The cashflow at time  $t_k$  results from a comparison of the continuation value  $\hat{\mathbf{F}}(t_k)$  and the cashflow vector  $P(\mathbf{S}_k)$  for immediate exercise: If an immediate exercise occurs in price path  $\omega$  (in this case  $P(S_k(\omega)) > \hat{F}(\omega, t_k)$  holds), the resulting cashflow  $P(S_k(\omega))$  is placed in the cashflow vector  $\mathbf{C}(t_k)$ , otherwise the discounted future cashflows in this vector (calculated as described in step 1) are kept.

5. Progress with the next time step  $t_k \rightarrow t_{k-1}$  if  $t_1$  has not yet been reached.

Using the entries in the cashflow vector at time  $t_1$  the option value can easily be calculated as a simple average:

$$V = \frac{1}{N} \sum_{\omega=1}^N C_{\omega}(t_1). \quad (62)$$

The algorithm described so far has to be extended for swing options with  $u_{\max}$  exercise rights (up-swing rights).

## 2.1 Extensions for Swing Options

Let  $u$  denote the number of upswing rights already exercised. Instead of a cashflow matrix a tensor with dimensions  $(N, J, u_{\max})$  would be necessary to keep account of the cashflows in the different price paths. Due to the compact storage, the cashflow vector with length  $N$  turns into a cashflow matrix with dimensions  $(N, u_{\max})$ . The algorithm for the Least Squares Monte Carlo calculation results as follows: In time step  $J$  all  $u_{\max}$  columns of the cashflow matrix are initialized with the cashflow vector  $\mathbf{C}(t_J)$  calculated as in eq. (52). After this initialization the following iteration results:

1. The cashflows in the cashflow matrix  $\mathbf{C}(t_{k+1})$  have to be discounted to time  $t_k$  by multiplication with the discount factor  $D(t_{k+1}, t_k)$ :

$$\mathbf{C}(t_k) = D(t_{k+1}, t_k) \cdot \mathbf{C}(t_{k+1}).$$

2. The cashflow vector  $P(\mathbf{S}_k)$  for an immediate exercise has to be calculated. For a plain vanilla call option the payoff in price path  $\omega$  is  $P(S_k(\omega)) = \max(S_k(\omega) - K, 0)$ .
3. For all price paths  $\omega$ , where the option is in the money (i.e.  $P(S_k(\omega)) > 0$  holds), the continuation value  $\hat{\mathbf{F}}_u(t_k)$  for the case, that  $u$  exercise rights have already been used, is calculated via linear regression:

$$\hat{\mathbf{F}}_u(t_k) = \sum_{l=0}^M a_l^u B_l(\mathbf{S}_k). \quad (63)$$

In contrast to a Bermudan option,  $u_{\max}$  regression problems need to be solved. This can be done simultaneously, as the regression matrix with elements  $B_l(\mathbf{S}_k)$  is the same for all  $u_{\max}$  problems. Altogether, these continuation values can be collected in an  $(N, u_{\max})$ -matrix  $\hat{\mathbf{F}}(t_k)$ .

4. The cashflow at time  $t_k$  results from a comparison of the continuation value  $\hat{\mathbf{F}}_u(t_k)$  and the cashflow vector  $P(\mathbf{S}_k)$  for immediate exercise: The option is exercised in price path  $\omega$  if  $P(S_k(\omega)) + \hat{F}_{u+1}(\omega, t_k) > \hat{F}_u(\omega, t_k)$  holds. In this case, the resulting cashflow  $P(S_k(\omega)) + C_{\omega, u+1}(t_k)$  has to be placed in row number  $\omega$  of the  $u$ -th column of the cashflow matrix  $\mathbf{C}(t_k)$ . Otherwise, the discounted future cashflows contained in this element (calculated as described in step 1) are left unchanged. Because of these modifications the calculations have to start with  $u = 0$  and end up with  $u = u_{\max} - 1$  and not vice versa! (In this case the next column of the cashflow matrix has to contain the (discounted) cashflows from the previous iteration step, as this ensures that after iteration step  $t_k$  only the remaining exercise rights are used (in addition to the one exercised at time  $t_k$ )).

5. Proceed to the next time step  $t_k \rightarrow t_{k-1}$ , if  $t_1$  has not yet been reached.

Calculating the average over the entries in the first column of the cashflow matrix after time step  $t_1$  the option value results:

$$V = \frac{1}{N} \sum_{\omega=1}^N C_{\omega,0}(t_1). \quad (64)$$

As a nice side effect of the algorithm, the values of all swing options with  $u = 1, \dots, u_{\max} - 1$  exercise rights can easily be calculated simultaneously as average over the appropriate remaining columns.

The values of the swing options are (independent of the number of exercise rights) discounted to time  $t_1$  (present value for this time). If the value has to be calculated for another (earlier) time, a corresponding discount factor has to be introduced.

Finally, it should be mentioned, that a few ideas concerning the implementation are crucial for the speed and accuracy of the resulting algorithm:

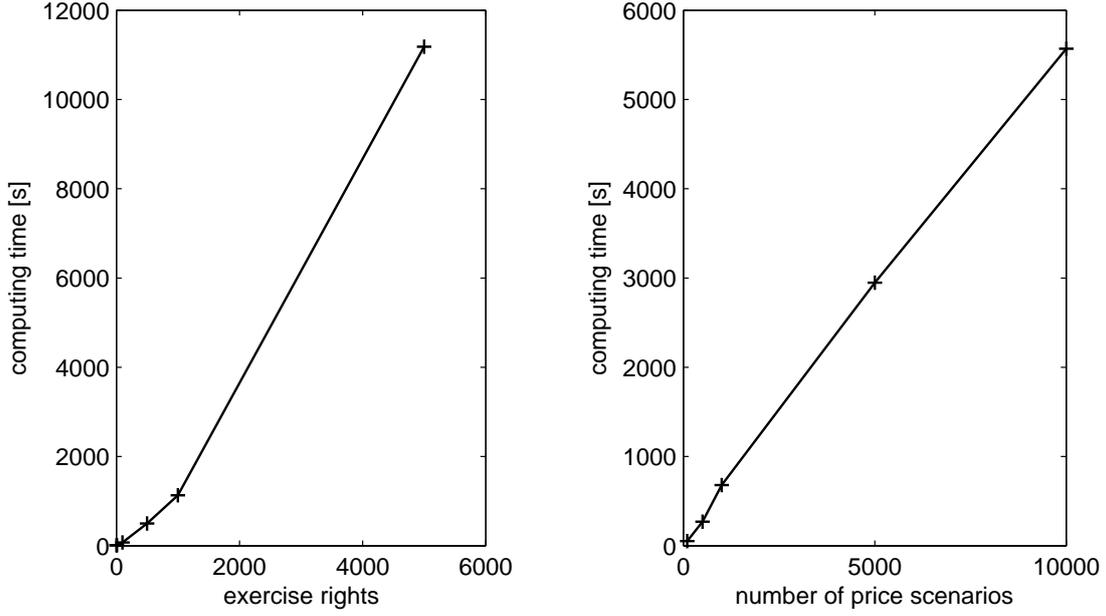
- **Efficient memory usage:** A cashflow matrix has to be used instead of a tensor. Otherwise, a valuation of swing options with 5000 upswing-right is impossible on a normal laptop computer (Intel T2080 Pentium Dual Core 1.73 GHz with 2 GB RAM). Furthermore, swapping, which is very time consuming, has to be prevented. Therefore, the generated scenarios should not be kept in memory for the whole examined lifetime but only for one day of the option's lifetime.
- **Economy size decomposition:** The commonly used regression algorithm is based on a QR decomposition of the regression matrix. In MATLAB an economy size decomposition algorithm (also called skinny QR decomposition) can be used, which improves the speed of the algorithm by round about a factor of 10 (depending of course on the number of exercise rights). Furthermore, the skinny QR decomposition should be reused to solve the  $u_{\max}$  regression problems (c.f. equation 63) simultaneously.
- **Scaling:** Due to the polynomials used as basis functions, it has turned out to be a very good idea to scale the spot scenarios and replace the regression (63) by

$$\hat{\mathbf{F}}_u(t_k) = \sum_{l=0}^M a_l^u B_l\left(\frac{\mathbf{S}_k}{\|\mathbf{S}_k\|_{\infty}}\right). \quad (65)$$

This scaling does not change the option value, but without it only calculations with up to  $M = 3$  basis functions are possible (at least in MATLAB on a laptop equipped with 2 GB RAM) as the regression matrix becomes numerically rank deficient. With this scaling up to  $M = 8$  basis functions can be used without rank deficiencies.

As figure 8 shows, the computing time of the efficient LSM algorithm presented in this paper scales linearly with the number of scenarios as well as with the number of exercise rights. For an option with 5000 exercise rights a calculation based on 1000 scenarios takes round about 1.5 hours, which is a very reasonable time.

Beneath the real option value obtained from the LSM algorithm, lower and upper boundaries for the option value can be calculated.



(a) Variation of the number of exercise rights (1000 scenarios)      (b) Variation off the number of scenarios (500 exercise rights)

**Figure 8:** Computing time for the efficient LSM algorithm as a function of the number of scenarios as well as the number of scenarios (strike 60 €/MWh,  $M = 5$  basis functions, regime-switching)

## 2.2 Intrinsic and Deterministic Value

On the one hand, the intrinsic value, i.e. the value against the price forward curve

$$V_{intrinsic} = \max_{\phi_i} \sum_{i=1}^N \exp(-r(T_i - t)) \phi_i P(F_{t,T_i}) \quad (66)$$

forms a lower boundary for the swing option value.

On the other hand, an upper boundary is given by the deterministic option value, i.e. the expected payoff that would result if the spot prices were deterministic and known in advance. In this case a number of price scenarios  $k = 1, \dots, N_{Scen}$  are generated and for each price scenario  $S_{T_i}^k$  the linear optimization problem

$$V_{deterministic}^k = \max_{\phi_i^k} \sum_{i=1}^N \exp(-r(T_i - t)) \phi_i^k P(S_{T_i}^k) \quad (67a)$$

with the energy constraint

$$\sum_{i=1}^N \phi_i^k \leq E_{max} \quad (67b)$$

is solved. Since the whole price path is assumed to be known in advance, the solution of this optimization problem can easily be obtained by

1. calculating the discounted payoffs  $\exp(-r(T_i - t))P(S_{T_i}^k)$  for each scenario,
2. sorting them in descending order for each scenario and
3. taking only the first  $U$  payoffs.

Finally, the deterministic option value is calculated as mean over all scenarios:

$$V_{deterministic} = \frac{1}{N_{scen}} \sum_{k=1}^{N_{scen}} V_{deterministic}^k . \quad (68)$$

The deterministic option value as well as the intrinsic value can be compared to the one resulting from the LSM algorithm.

### 3 Results for Selected Swing Options

All the calculations presented in the following subsections have been done in MATLAB and assume a yearly interest rate of 5% as well as an energy constraint  $E_{max} = U$  MWh, where  $U$  denotes the number of upswing rights. The set  $J$  of possible exercise times always contains all the hours of the year 2009 (i.e. 8760 hours).

#### Intrinsic Value

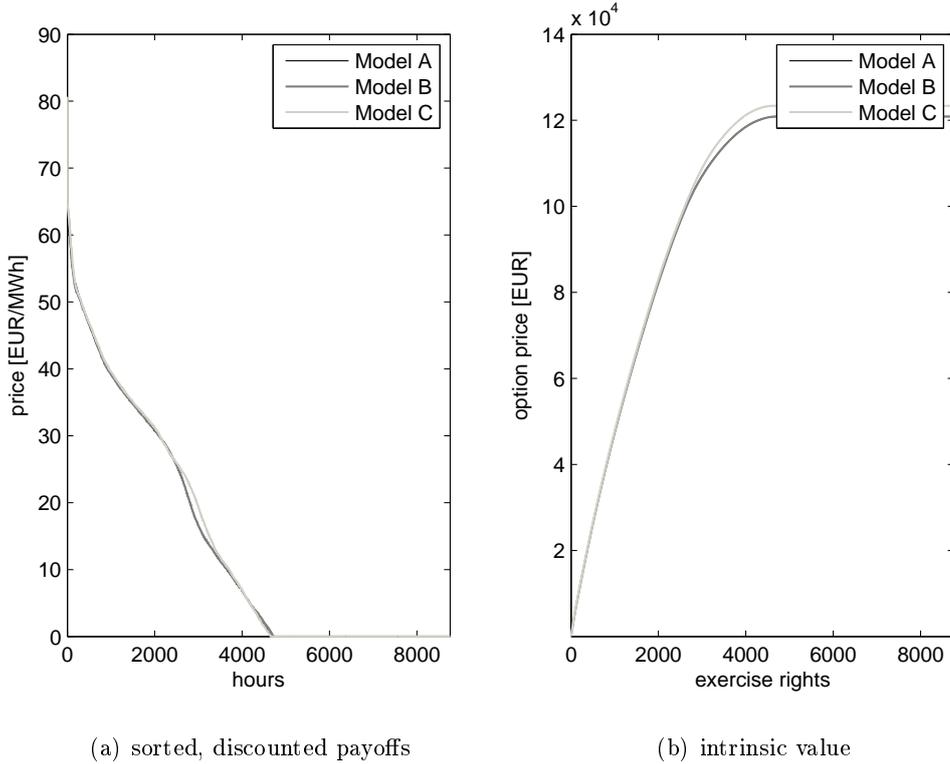
To get a first insight into typical properties of swing options we can have a closer look at the intrinsic value. In figure 9 the sorted discounted price forward curves resulting from the three models as well as the intrinsic value for the corresponding swing options with a strike price of  $K = 60$  €/MWh have been plotted. Obviously, the sorted price forward curves for model A, B and C show only little differences. Generally, the intrinsic option value grows with the number of exercise right, at first almost linearly. At round about 4700 exercise rights a maximum intrinsic value is reached – more exercise rights do not increase the option value any longer, i.e. the number of exercise rights is no longer a restriction as no more prices above the strike price exist in the price forward curve. Special interest will be spend to this special property if deterministic as well as real option values are examined later on.

	model A	model B	model C
$V_{max,intrinsic}$	120828.31	120901.30	123379.18
$U_{max}$	4693	4718	4653

**Table 5:** Comparison of maximum intrinsic value and corresponding number of exercise rights for model A to C

The maximum intrinsic value and the corresponding number of exercise rights  $U_{max}$  are given in table 5. From this table it can be seen that for  $U_{max}$  upswing rights the differences between the three models – at least as far as the intrinsic value is concerned – are rather small.

The downside of the intrinsic value is, that it is based on the HPFC containing the spikes only as an average and that it does not take the exercise decision under uncertainty (exercise today or hope for higher prices in the future) into account — in contrast to the LSM algorithm.

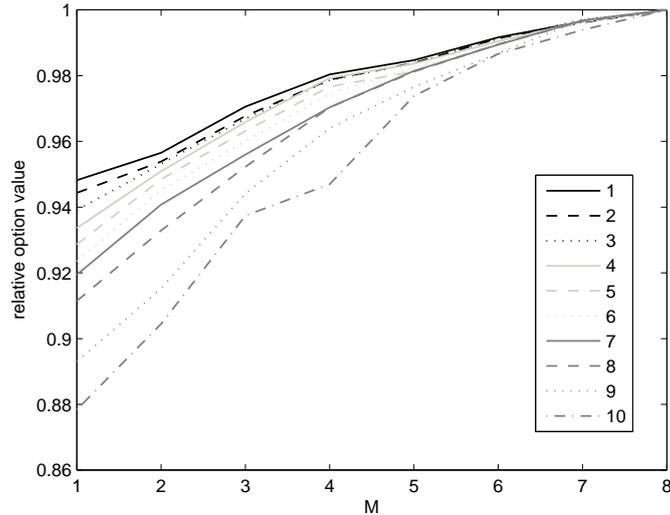


**Figure 9:** Intrinsic value for a swing option with strike  $K = 60$  €/MWh resulting from model A to D

### Required Number of Basis Functions and Scenarios

One important question concerning the LSM algorithm is, how many basis functions are necessary to get a satisfactory accuracy of the option value. To answer this question, 1000 scenarios generated with the regime-switching model A have been used and the option values of swing options with 1 to 10 exercise rights have been calculated using up to  $M = 8$  basis functions. The resulting options values have been divided by the option values obtained for  $M = 8$  basis functions and are shown in figure 10. Using  $M = 1$  instead of  $M = 8$  decreases the resulting option value by 5 % in case of ten exercise rights and up to 12 % for a single exercise right. As computation time scales almost linearly with the number of basis functions,  $M = 5$  is a good compromise between accuracy and speed.

A second major question is the convergence of the LSM algorithm as far as the number of price scenarios is concerned. For this purpose, a swing option with 500 exercise rights has been priced using different numbers of scenarios. For Monte Carlo Simulations a  $\sqrt{N}^{-1}$  law for the convergence is typical, i.e. quadrupling the number of sample paths approximately halves the error in the simulated price. This behavior can also be observed in figure 11 (a). Starting with 100 scenarios generated with model A, the error reduces nonlinearly with an increasing number of scenarios. Nevertheless, even for 100 scenarios the error in the option value (compared to a scenario number of 10000) is small if the number of exercise rights is not too small. Figure 11 (b) shows that for more than 50 exercise rights the relative error in comparison to 10000 scenarios is less than 1 % if 1000 scenarios are used. Due to this fact the further calculations in this paper



**Figure 10:** Options values for 1 to 10 exercise rights and  $M = 1, \dots, 8$  (strike 60 €/MWh)

have been based on 1000 price scenarios as well as  $M = 5$  basis functions.

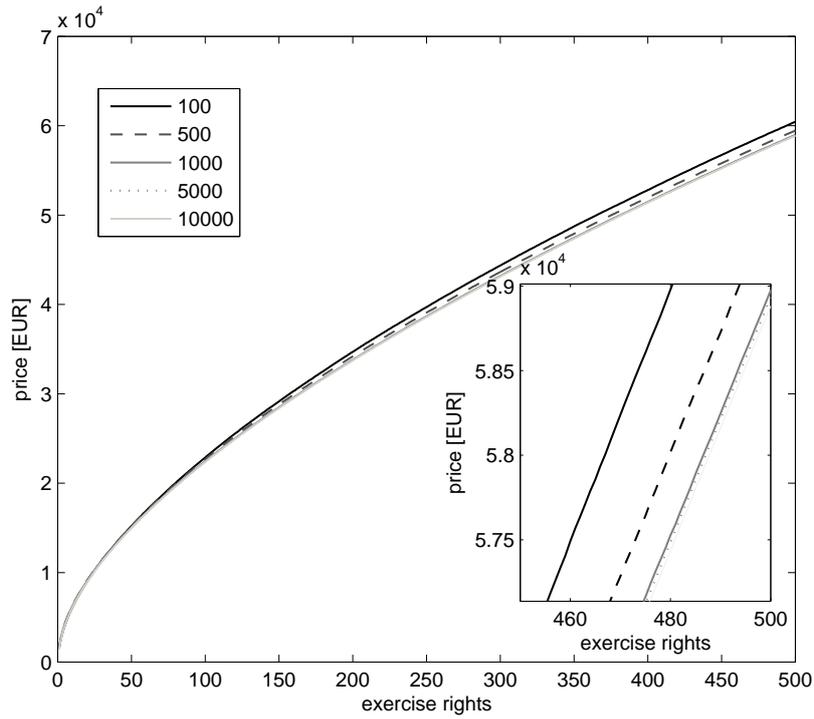
### Option Values for the Different Models

In a next step, option values resulting from the three specified models A to C are compared. In figure 12 the option values calculated with the LSM algorithm are plotted against the number of exercise rights and the corresponding intrinsic and deterministic values are also shown.

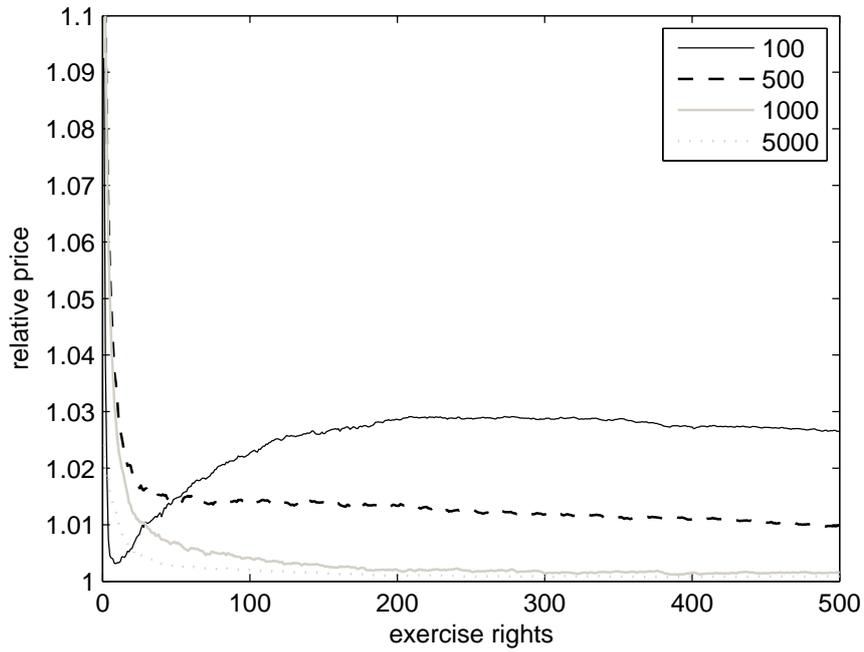
Regardless of the underlying model, the intrinsic value is very small compared to the value calculated using the LSM algorithm, which is much closer to the deterministic value. One reason for this large difference between the intrinsic and the LSM option value is, that spikes are very important for swing options with few exercise rights. Unfortunately, the spikes are only contained in the HPFC as an average and therefore have only a small influence on the HPFC and the intrinsic option price. Another reason is, of course, the volatility introduced by the two factors, which is not reflected by the HPFC. On the contrary, the deterministic value is larger as the nomination in each scenario is done without uncertainty — in difference to the LSM algorithm.

In the loglog plot shown in figure 12 (a) the option values obtained applying model B and C have been divided by the option value calculated using model A. The resulting relative prices reveal the potential model error resulting from a use of the different models. For a small number of exercise rights models B and C result in lower prices than model A – the prices differ by 32 % (model B) or even 56 % (model C) respectively. With a growing number of exercise rights the situation reverses. For 500 exercise rights the option prices obtained with model B as well as model C are 9 % higher than the ones with model A.

In figure 13 the LSM option values for models A to C as well as the relative option values are plotted for up to 5000 exercise rights. The subfigures containing the results for the three models show, that the differences between the deterministic price and the real option value calculated via LSM first become larger with a growing number of exercise rights – and shrink to zero if the number of exercise rights approaches round about 5000. This behavior reproduces

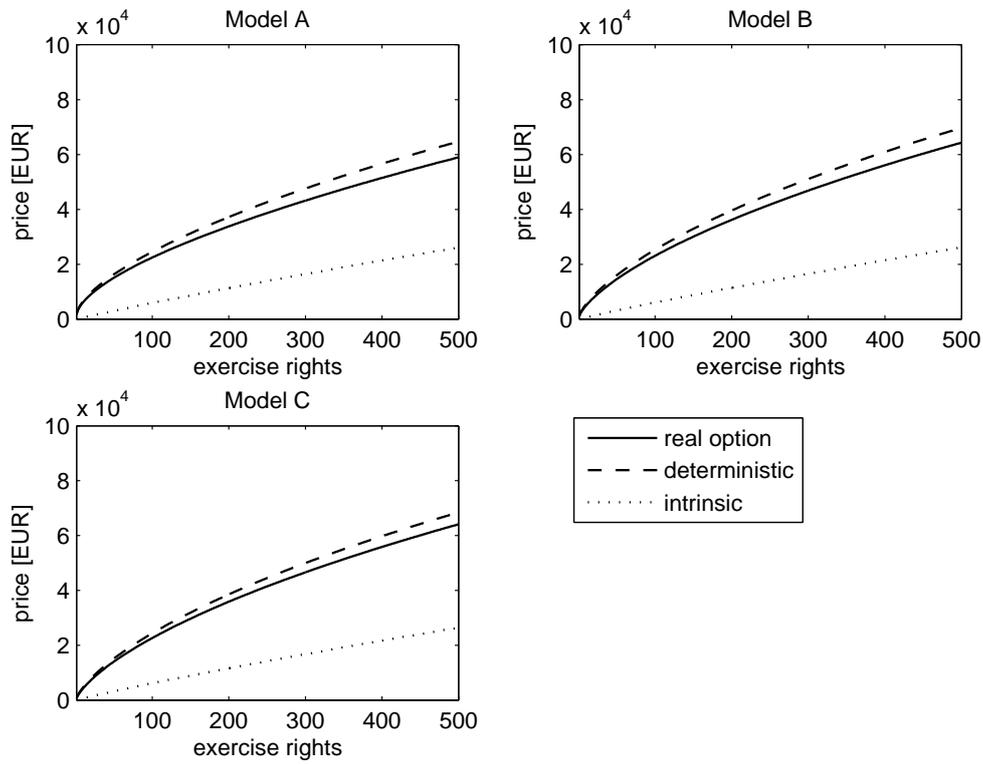


(a) Absolute change in option value due to scenario number

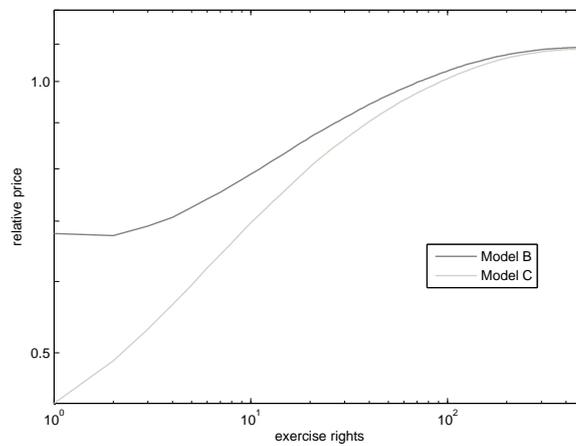


(b) Relative option value compared to 10000 scenarios

**Figure 11:** Option value calculated with LSM using model A (strike 60 €/MWh) for up to 500 exercise rights using different numbers of price scenarios

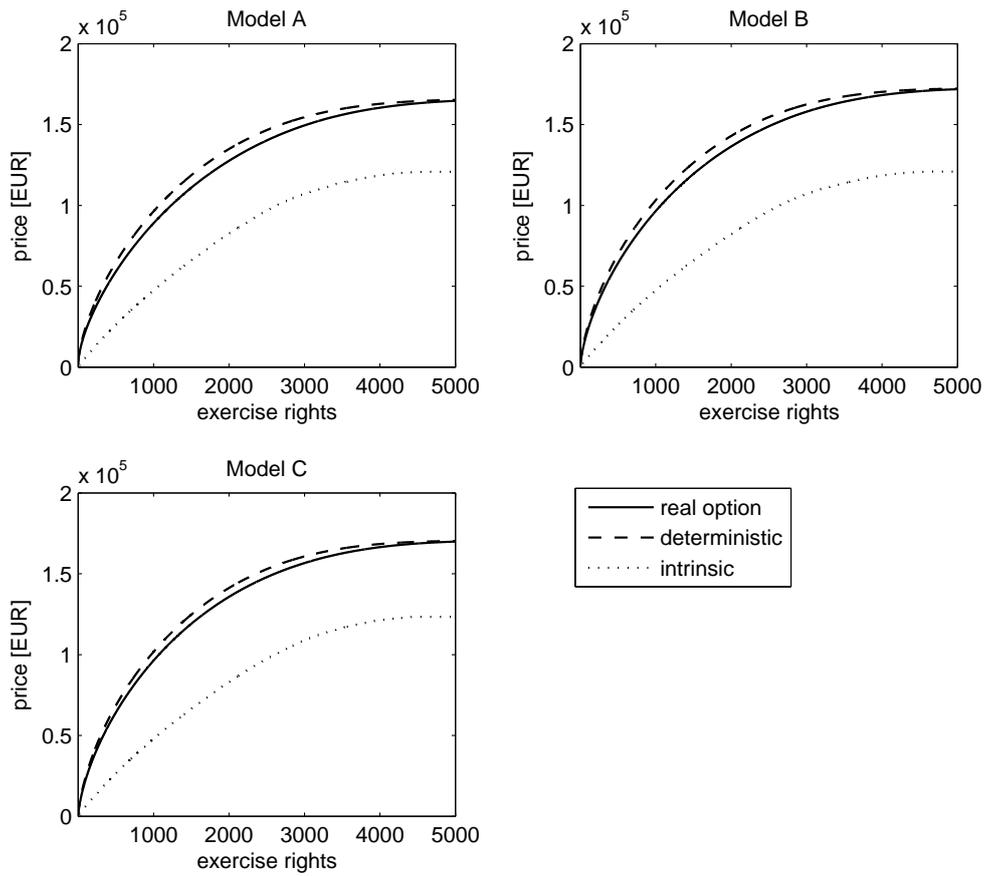


(a) absolute option values for the four models

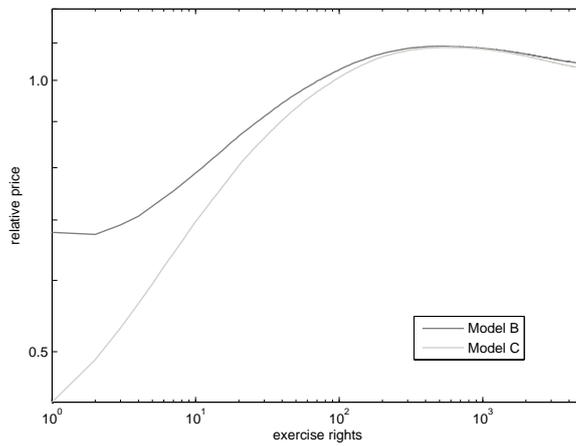


(b) option value relative to model A

**Figure 12:** Real option value calculated with LSM using models A to C in comparison to intrinsic and deterministic value (strike 60 €/MWh) for up to 500 exercise rights

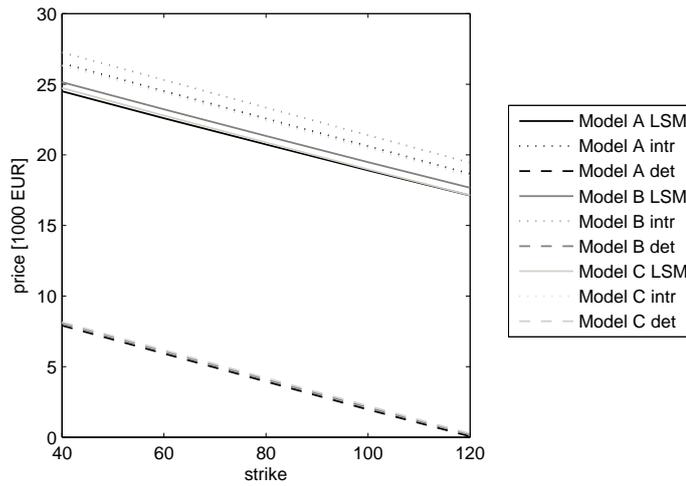


(a) absolute option values for the four models

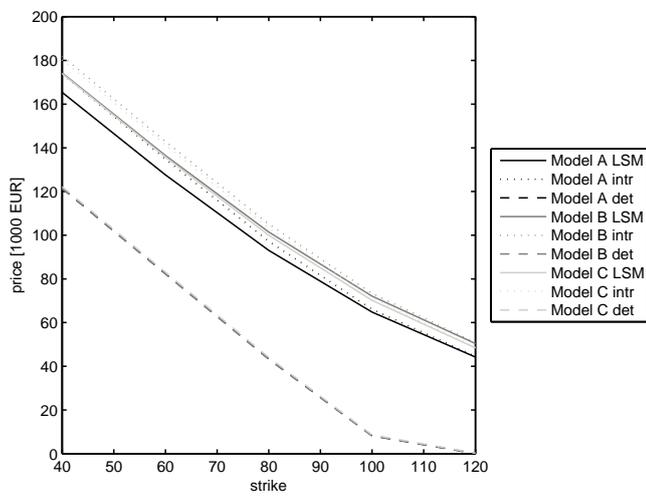


(b) LSM option value relative to model A

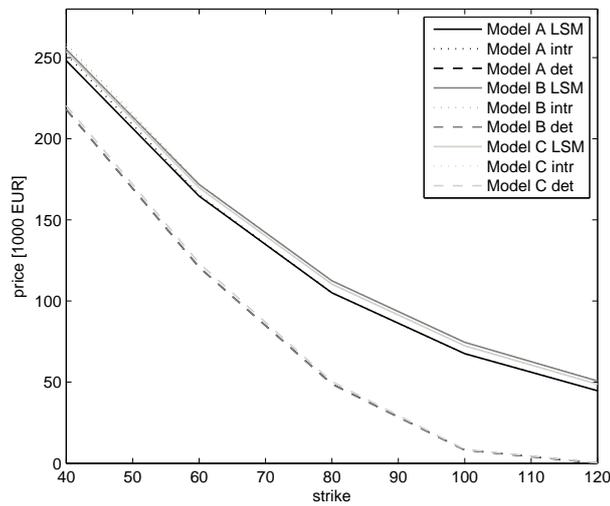
**Figure 13:** Real option value calculated with LSM using models A to C in comparison to intrinsic and deterministic value (strike 60 €/MWh) for up to 5000 exercise rights



(a) 100 exercise rights



(b) 2000 exercise rights



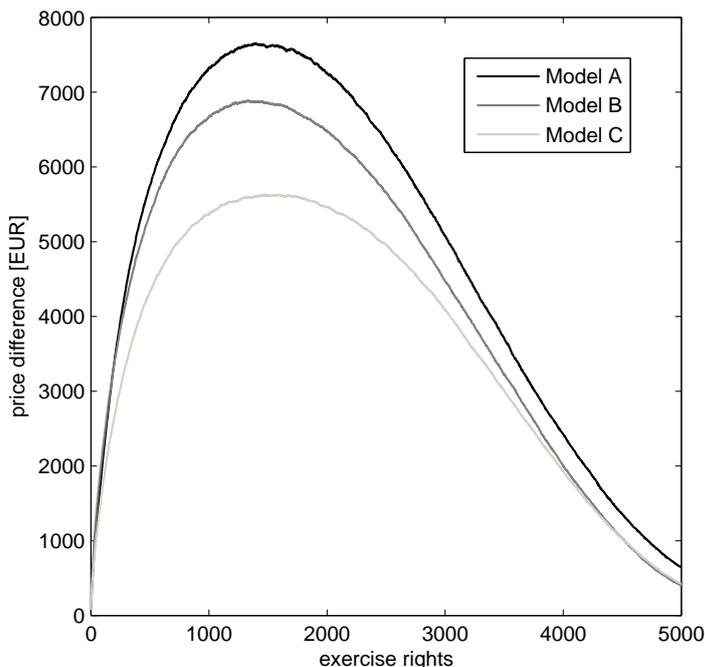
(c) 5000 exercise rights

**Figure 14:** Real option value for different strike prices using models A to C in comparison to intrinsic and deterministic value

model	B	C
relative price	1.04	1.03

**Table 6:** relative price of the different models for 5000 exercise rights (strike 60 €/MWh)

what we have seen for the intrinsic price: Even with the uncertain future spot prices, a limit of 5000 exercise rights is almost surely no longer a constraint — LSM algorithm and deterministic calculation show the same results.



**Figure 15:** Difference between the deterministic value and the real option value calculated with LSM for models A to C (strike 60 €/MWh) for up to 5000 exercise rights

Furthermore, for 5000 exercise rights the differences between the three models are smaller than 4 % in any case (c.f. table 6) and it can be concluded, that for such a large number of exercise rights the model choice is less important but nevertheless still relevant.

Finally, the option values have been calculated for strike prices varying between 40 €/MWh and 120 €/MWh and the results are shown in figure 14. In any case, for a strike price of 120 €/MWh the intrinsic value for model A to C approaches zero, as no prices higher than 120 €/MWh can be found in the HPFC. Moreover, the option value generally decreases with an increasing strike price. While this decrease is linear for 100 exercise rights (fig. 14 (a)) it becomes more and more nonlinear with a growing number of exercise rights (fig. 14 (b),(c)). Once again, the differences between the real option value and the deterministic value decrease – now independent of the number of exercise rights – with an increasing strike price. This is once again due to the fact that with a growing strike price the number of exercise rights is almost surely no longer a constraint. Therefore, the exercise decision becomes trivial and the option value converges to the one obtained in a deterministic calculation.

## 4 Summary

In this paper we have modeled the EEX spot prices using three different price processes. From a kernel density estimation it can be concluded that model A (regime-switching combined with 24 ARMA processes) reproduces the characteristics of the historical spot prices, namely fat tails and asymmetry, best. Model B (jump-diffusion combined with 24 ARMA processes) and Model C (NIG process combined with 24 ARMA processes) show comparable results but perform not as good as model A. Beyond this in-sample analysis, model A also shows the best performance in an out-of-sample analysis and the estimation algorithm delivers stable parameters as has been shown for a period of 400 trading days. The other two models either show less parameter stability or weaknesses in the out-of-sample analysis. Model A also as well as model B reproduce the implied volatilities of options quoted at the EEX, but none of the models is able to reproduce the empirical correlation structure of futures prices. Over all, model A is the most preferable one among the three examined models.

In the second part of the paper an efficient Least Squares Monte Carlo algorithm (LSM) has been introduced and applied to swing options with up to 5000 exercise rights using price scenarios generated with the three different price processes as input. Here it has become obvious that it is very important for the pricing of swing options to use simulation methods based on an appropriate price process. This is especially relevant for small numbers of exercise rights as option prices may vary by a factor of two depending on the price process. Therefore a swing option with few exercise rights is a risky bet on the price process used because of the high model risk. With a growing number of exercise rights the differences between the three models shrink and the model choice becomes less important.

In all examined cases the resulting option values obtained using the LSM algorithm as well as a deterministic calculation are much larger than the intrinsic option values. But from the calculations in this paper it can be concluded that with a growing number of exercise rights and an increasing strike price a time consuming stochastic calculation becomes superfluous – instead a simple deterministic calculation can be applied. This is an important fact for the valuation and optimal dispatch of power plants with high numbers of hours of operation, for example gas and steam plants or coal fired power plants. In a simplified model (disregarding constraints like ramps, minimum down-time, maintenance, outages etc.) these power plants can be seen as swing options with at least 5000 exercise rights. Of course the situation may be different if another price level at the EEX emerges and therefore from time to time this examination has to be repeated.

## References

- Benth, F.: 2007, *Option Theory with Stochastic Analysis. An Introduction to Mathematical Finance*, Springer.
- Benth, F. and Saltyte-Benth, J.: 2004, *International Journal of Theoretical and Applied Finance* **7**, 177–192.
- Boogert, A. and de Jong, C.: 2008, Gas storage valuation using a monte carlo method, *The Journal of Derivatives* .
- Burger, M., Graeber, B. and Schindelmayer, G.: 2007, *Managing Energy Risk*, John Wiley & Sons.

- Burger, M., Klar, B., Müller, A. and Schindlmayr, G.: 2004, A spot market model for pricing derivatives in electricity markets, *Quantitative Finance* **4**, 109–122.
- Cartea, A. and Figueroa, M.: 2005, Pricing in electricity markets: a mean reverting jump diffusion model with seasonality, *Applied Mathematical Finance* **12**(4), 313–335.
- Clewlow, L. and Strickland, C.: 2000, *Energy Derivatives - Pricing and Risk Management*, Lacima Publications.
- d-fine: 2004, Valuation of enery derivatives with monte carlo methods, Master Thesis, University of Oxford.
- Dahlgren, M.: 2005, A continuous time model to price commodity-based swing options, *Review of Derivatives Research* **8**(1), 27–47.
- de Jong, C.: 2006, The nature of power spikes: A regime-switch approach, *Studies in Nonlinear Dynamics & Econometrics* .
- Dörr, U.: 2003, Valuation of swing options and examination of exercise strategies by monte carlo techniques, Master Thesis, University of Oxford.
- Geman, H. and Roncoroni, A.: 2006, Understanding the fine structure of electricity prices, *Journal of Business* **79**(3).
- Hamilton, J.: 1994, *Time Series Analysis*, Princeton University Press.
- Huisman, R. and Mahieu, R.: 2003, Regime jumps in electricity prices, *Energy Economics* **25**, 425–434.
- Keppo, J.: 2004, Pricing of electricity swing options, *Journal of Derivatives* .
- Kiesel, R., Schindlmayr, G. and Boerger, R. H.: 2007, A two-factor model for the electricity forward market, <http://ssrn.com/abstract=982784>.
- Koekebakker, S. and Ollmar, F.: 2005, Forward curve dynamics in the nordic electricity market.
- Longstaff, F. and Schwartz, E.: 2001, Valuing american options by simulation: A simple least-squares approach.
- Schindelmayer, G.: 2005, A regime-switching model for electricity spot prices, *10th Symposium on Banking, Finance and Insurance*, University of Karlsruhe.
- The MathWorks: 2008, MATLAB, <http://www.mathworks.com>.
- Weron, R.: 2006, *Modeling and Forecasting Electricity Loads and Prices. A Statistical Approach*, John Wiley & Sons.
- Weron, R.: 2007, MFE Toolbox ver. 1.0.1, [http://www.im.pwr.wroc.pl/~rweron/MFE/MFE\\_Toolbox.html](http://www.im.pwr.wroc.pl/~rweron/MFE/MFE_Toolbox.html).
- Weron, R. and Misiorek, A.: 2007, Heavy tails and electricity prices: Do time series models with non-gaussian noise forecast better than their gaussian counterparts?, *Prace Naukowe Akademii Ekonomicznej we Wroclawiu* **1076**, 472–480.