

Michael Eichler, Dennis Tuerk

**Fitting semiparametric Markov  
regime-switching models to  
electricity spot prices**

RM/12/035

**METEOR**

Maastricht University School of Business and Economics  
Maastricht Research School of Economics  
of Technology and Organization

P.O. Box 616  
NL - 6200 MD Maastricht  
The Netherlands

# Fitting semiparametric Markov regime-switching models to electricity spot prices

M. Eichler, D. Tuerk\*

*Department of Quantitative Economics, Maastricht University, The Netherlands*

---

## Abstract

Recently regime-switching models have become the standard tool for modeling electricity prices. These models capture the main properties of electricity spot prices well but estimation of the model parameters requires computer intensive methods. Moreover, the distribution of the price spikes must be assumed given although the high volatility of the spikes makes it difficult to check this assumption. Consequently, there are a number of competing proposals. Alternatively we propose the use of a semiparametric Markov regime-switching model that does not specify the distribution under the spike regime. To estimate the model we use robust estimation techniques as an alternative to commonly applied estimation approaches. The model in combination with the estimation framework is easier to estimate, needs less computation time and distributional assumptions. To show its advantages we compare the proposed model with a well established Markov-switching model in a simulation-study. Further we apply the model to Australian log-prices. The results are in accordance with the results from the simulation-study, indicating that the proposed model might be advantageous whenever the distribution of the spike process is not sufficiently known. The results are thus encouraging and suggest the use of our approach when modeling electricity prices and pricing derivatives.

*Keywords:* electricity spot prices, mean-reversion, Markov regime-switching, robust estimation, semiparametric estimation, simulation-study

---

## 1. Introduction

In the past it has been frequently shown that electricity prices are mean-reverting (see e.g. Lucia and Schwartz, 2002, Huisman and Mahieu, 2003 and Weron, 2006). Unfortunately the mean-reverting model is known to be hampered severely through spikes when estimated using standard estimation techniques as quasi maximum likelihood (QML) estimation. For that reason model extensions (which nest the mean-reverting model) were introduced into the literature on electricity price modeling. A first extension was proposed by Deng (2000) using jump-diffusion models to model electricity spot prices. Although jump-diffusion models are still used it is commonly known that they are not well suited to cope with price spikes, regularly observed on electricity markets.

For this reason regime-switching models, which distinguish between the base and the spike process, were introduced as a next step (see e.g. Robinson, 2000, Huisman and Mahieu, 2003, Huisman and De Jong, 2003 and Rambharat et al., 2005).

It is important to note that the specifications of regime-switching models differ in the way the regime-switching mechanism is implemented and that in this context two main classes can be

---

<sup>\*</sup>Financial support from Electrabel, a subsidiary of the GDF Suez Group is gratefully acknowledged. We thank Marco Avarucci, Francisco Blasques, Kris Boudt, Cyriel De Jong, Sébastien Laurent, Franz Palm and Francesco Violante for helpful comments. The usual disclaimer applies.

\*Corresponding author. Address: Department of Quantitative Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands. Tel: +31 43 388 3239 Fax: +31 43 388 2000

*Email addresses:* m.eichler@maastrichtuniversity.nl (M. Eichler), d.tuerk@maastrichtuniversity.nl (D. Tuerk)

distinguished when modeling electricity prices. For the first class, called threshold autoregressive (TAR) models, the regime at time  $t$  can be determined by an observable variable at time  $t - d$  (with  $0 < d < t$ ) and thus the regimes that have occurred in the past and present are known with certainty. For the second class, called Markov-switching (MS) models, the regime is determined by an unobservable, latent variable. In this case one cannot be certain that a particular regime has occurred at a particular point in time but can only assign or estimate probabilities of their occurrences. MS models further use information up to time  $t$  to assign probabilities concerning the state of the process at this time.

As shown, for example, by De Jong (2006) and Misiorek et al. (2006) these models seem to be better suited to account for the distinct nature of electricity prices. In this context it is not surprising that a lot of research has been focused on the extension and evaluation of different regime-switching models. Rambharat et al. (2005) and Misiorek et al. (2006) extended the TAR model by an exogenous variable (temperature recorded at the same time as the maximum price of the day and load). For MS models Weron et al. (2004) and Bierbrauer et al. (2007) compared the model introduced by Huisman and De Jong (2003) using different distributional assumptions for the spike process, while Misiorek et al. (2006) compared the regime-switching model with other model classes evaluating their short-term forecasting power. Extensions were for example given by Kosater and Mosler (2006) who introduced a third regime to account for the occurrence of negative spikes and by Mount et al. (2006) who allowed the transition probabilities to depend on the load and/or the implicit reserve margin. A further extension where given by Janczura and Weron (2009), who extended the base regime to account for time-varying volatility, while restricting the spike process to follow a shifted distribution. Haldrup and Nielsen (2006) proposed a Markov-switching type long memory model. In difference to the standard MS approach which uses a latent state variable, they used congestion as an observable state variable. An extension for VAR models followed in Haldrup et al. (2010). An overview concerning the different extensions can be found in Janczura and Weron (2010).

Although regime-switching models might be better suited to capture the characteristics of electricity prices than their predecessors, their implementation is often quite complicated, computationally demanding and individual parameters might be unidentified.<sup>1</sup> Rambharat et al. (2005) used a block Metropolis hastings algorithm, whereas Huisman and De Jong (2003) applied a Kalman filter based approach with Bayesian updating. With reference to gains in computational intensity Janczura and Weron (2012) propose to adapt the estimation technique used by Huisman and De Jong (2003) and replace the latent variables from the base regime with their expectations.

In the face of the given problem the contribution of the paper at hand is two-fold. First we propose a semiparametric Markov regime-switching model that does not specify the distribution under the spike regime. The extension of the original model by Huisman and De Jong (2003) can thus be seen as a direct response to several earlier publications, which treated the question of the correct choice for the distribution of spikes (see e.g. Bierbrauer et al., 2007 and Janczura and Weron, 2009).

Further to estimate the model we use robust estimation techniques as an alternative to commonly applied estimation approaches. To our knowledge the only attempt to introduce robust statistics into the literature on electricity price modeling has been made by Trueck et al. (2007), analyzing the effects of different prefiltering techniques on model estimation and outlier detection. The idea is based on Clewlow and Strickland (2000) who proposed recursive prefiltering using non-robust scale estimates. Our approach is different insofar that we use estimation techniques which by construction in a first step neglect the influence of spikes when estimating the parameters of the base process. The advantage of our multi-step based approach, compared to the approaches mentioned before to estimate MS models, is that neither complicated numerical optimization nor any distributional assumption concerning the spike process are needed. Therefor the estimation is extremely time efficient and flexible. The procedure further allows to account for negative outliers without making it necessary to model them explicitly thus providing a valid alternative to the

---

<sup>1</sup>This is for example the case when assuming a compound poisson process for the spikes.

three-regime models which were e.g. used in Lindstroem and Regland (2012). This robustified estimation is advantageous when modeling data that exhibits very scarce occurrence of negative outliers.

The paper is organized as follows. In Section 2 we present the theoretical framework of the semiparametric model while the proposed estimation technique is described in Section 3. In Section 4 we evaluate the properties of the proposed approach in a simulation-study and compare it to a standard fully parametric Markov regime-switching model. In Section 5, we discuss an application to electricity spot prices from the Australian market. Section 6 concludes.

## 2. Model specification

In absence of spikes electricity spot prices are generally assumed to be log-normally distributed. For this reason it is a common convention in the literature to model log-spot prices instead. When applied to log-spot prices,  $P_t$ , the price process can be assumed to be separable into a deterministic part,  $D_t$ , and a stochastic part,  $X_t$ :

$$P_t = D_t + X_t. \quad (2.1)$$

While  $D_t$  accounts for a possibly time-varying deterministic component including mean, trend and seasonal effects,  $X_t$  represents the stochastic component, which is generally assumed to be mean-reverting. For this reason one of the first models proposed by Lucia and Schwartz (2002) to examine dynamics on electricity markets was based the assumption that electricity prices follow an Ornstein-Uhlenbeck process. In this framework changes in  $P_t$  can be written as:

$$\begin{aligned} dP_t &= dD_t + \alpha(D_t - P_t) + \sigma dW_t \\ &= dD_t - \alpha X_t dt + \sigma dW_t. \end{aligned} \quad (2.2)$$

Here  $\alpha$  measures the speed with which  $P_t$  reverts to  $D_t$  (or  $X_t$  to 0) and  $\sigma$  is a positive real constant, scaling movements in the Wiener process,  $W_t$ . Following the literature we concentrate on modeling the stochastic component of the log-price process after applying an appropriate procedure to eliminate  $D_t$ . With respect to equation (2.2) this leaves:

$$dX_t = -\alpha X_t dt + \sigma dW_t. \quad (2.3)$$

As the paper at hand is concerned with modeling the stochastic part using regime-switching models, we will leave the explanation of the demeaning procedure for Section 5.

The disadvantage of model 2.3 is the fact that it abstracts from the existence of price spikes. As already pointed out to overcome this problem, different model specifications have been proposed in the past (for overviews see e.g. Misiorek et al., 2006, Bierbrauer et al., 2007 and De Jong, 2006). The possibly most successful one is the MS model of Huisman and De Jong (2003) on which we are going to extend by loosening its distributional assumption for the spike process.

The proposed semiparametric MS (SMS) model is in spirit close to the MS model of Huisman and De Jong (2003). To incorporate the important feature of extreme price movements, we assume that the (discretized) stochastic component,  $X_t$ , follows a regime-switching process with two independent states. The base regime,  $R_t = B$ , models the ordinary price dynamics from equation (2.3) and the spike regime,  $R_t = S$ , models times of extreme prices. In this context  $R_t \in \{B, S\}$  represents a time-homogenous (hidden) two-state Markov-chain with transition matrix,  $\Pi$ , which gives the conditional probabilities,  $\pi_{ij}$ , of switching from regime  $i$  (in  $t$ ) to  $j$  (in  $t + 1$ ):

$$\Pi = (\mathbb{P}(R_{t+1} = j | R_t = i))_{i,j \in \{B,S\}} = \begin{pmatrix} \pi_{BB} & \pi_{BS} \\ \pi_{SB} & \pi_{SS} \end{pmatrix}. \quad (2.4)$$

Thus  $X_t = X_{t,R_t}$  with:

$$X_{t,B} = \gamma X_{t-1,B} + \varepsilon_t, \quad (2.5)$$

$$X_{t,S} \stackrel{iid}{\sim} F, \quad (2.6)$$

As equation (2.5) is the discrete time equivalent to (2.3),  $X_{t,B}$  converges with speed  $(1 - \gamma)$  to zero and the innovations,  $\varepsilon_t$ , are assumed to be independent and identically normal distributed with mean zero and variance  $\sigma_\varepsilon^2$ . In the parametric model the spike regime (see equation 2.6) is assumed to follow an *iid* process, represented by a known distribution  $F$ . Our model is insofar different that we do not make any distributional assumptions for the spike process and require (for its simplest specification) only that  $F$  has support  $[a, \infty)$  for some threshold  $a$  corresponding to a pre-specified quantile of the time-series. The assumption that observations from the spike regime should not be smaller than  $a$  seems reasonable as the spike regime is used to model extreme (positive) prices and was already made earlier in the literature (see Janczura and Weron, 2009 and Janczura and Weron, 2010).

So far the estimation methods proposed in the literature (see e.g. Huisman and De Jong, 2003, Janczura and Weron, 2012 and references therein) do need a distributional assumption for the spike process in order to be applicable. To allow for the proposed generalization concerning the spike process, we will thus introduce an alternative estimation approach which will be presented in the next section.

### 3. Estimation

We present robust estimation techniques to estimate the model. The logic behind our proposal to use a robust estimation approach is to treat the base regime as the underlying dynamic core process. Observations belonging to the spike regime are in contrast treated as extreme observations with respect to the base regime.

This reasoning suggests to first estimate the parameters of the base regime by use of robust estimation techniques. In a second step, the parameters for the spike regime are estimated. The latter is also needed to compute the Markov-transition matrix.

#### 3.1. Robust estimation of the base regime

To properly model the base regime we have to adequately estimate  $\gamma$  and the variance of residuals,  $\sigma_\varepsilon^2$ , from equation (2.5). The needed procedures will be presented in the following.

##### 3.1.1. Estimation of $\gamma$

The estimator for the autocorrelation coefficient  $\gamma$  proposed by Ma and Genton (2000), is based on the following identity:

$$\gamma = \frac{\text{var}(X_t + X_{t-1}) - \text{var}(X_t - X_{t-1})}{\text{var}(X_t + X_{t-1}) + \text{var}(X_t - X_{t-1})}. \quad (3.1)$$

Thus estimation of  $\gamma$  can be based on a robust estimate for  $\text{var}(X_t + X_{t-1})$  and  $\text{var}(X_t - X_{t-1})$ .

To compute a robust scale estimator Ma and Genton (2000) used the  $Q_n$  estimator, introduced by Rousseeuw and Croux (1992, 1993):

$$\hat{Q}_n = d_{Q_n} \{ |X_i - X_j|; i < j \}_{(l)}, \quad (3.2)$$

which is defined as the  $l$ th order statistic of  $\binom{n}{2}$  interpoint distances between the observations of sample  $X = (X_1, \dots, X_T)'$  with  $l = \lceil \binom{n}{2} / 4 \rceil + 1$ . The correction factor  $d_{Q_n} = 1 / (\Phi^{-1}(0.625)\sqrt{2})$  is needed to ensure consistency in the case that  $X$  is normally distributed as in this case  $Q_n$  corresponds with the 62.5% quantile of  $(X_i - X_j) \sim N(0, 2\sigma_x^2)$ . The estimator has the highest breakdown point possible (which in context of (3.1) corresponds to 25%). Furthermore it has a high efficiency of 82% under normality.

Although, depending on the relative proportion and size of contamination, the  $Q_n$  estimator can have a non-negligible bias, the resulting autocorrelation estimate,  $\hat{\gamma}$ , is highly robust. In other words, despite the fact that the  $Q_n$  estimator does exhibit a non-negligible maximal bias when both location and scale of the contaminating outlier distribution tend to infinity, the  $Q_n$  based autocorrelation estimator stays unbiased in this case. If on the other hand outliers (even of high

absolute value) do exhibit a standard deviation that is not significantly higher than the one of the base process, the  $Q_n$  based autocorrelation estimator might be biased. As will be seen later, for electricity prices the absolute values of outliers as well as the corresponding standard deviation appear to be high enough such that the autocorrelation coefficient stays unbiased.<sup>2</sup>

### 3.1.2. Estimation of $\mu_B$ and $\sigma_B^2$

As stated before, despite its good efficiency under normality, the  $Q_n$  estimator has a non-neglectable maximum bias caused by outliers of large absolute size and variance. This makes it less suited to estimate the variance  $\sigma_B^2$  of the base process since the observations in the spike regime consist of extreme prices with high variance. However, under our assumption that observations in the spike regime are above some threshold  $a$ , a consistent estimator of  $\sigma_B^2$  can be obtained by the truncated maximum likelihood approach (see e.g. Boudt et al., 2011 and Marazzi and Yohai, 2004). Such if  $x^B \sim N(\mu_B, \sigma_B^2)$  and letting  $y$  be distributed as a truncated normal,  $y \sim TN(\mu_B, \sigma_B^2, a)$ , with threshold  $a$ , the truncated likelihood function can be written as:

$$\frac{1}{n} \ell(\sigma_B^2) = \frac{-1}{2n} \sum_{t=1}^T \frac{(y_t - \mu_B)^2}{\sigma_B^2} \mathbb{1}_{\{y_t \leq a\}} - \frac{1}{2} \log \sigma_B^2 - \log \Phi \left( \frac{a - \mu_B}{\sigma_B} \right), \quad (3.3)$$

with  $\mathbb{1}_{\{y_t \leq a\}} = 1$  if  $y_t \leq a$  and zero otherwise while  $n = \sum_{t=1}^T \mathbb{1}_{\{y_t \leq a\}}$ . Consistent estimates for  $\mu_B$  and  $\sigma_B^2$  can then be achieved by maximizing the truncated likelihood function as long as spikes are greater than the chosen cut-off quantile of the time-series. The variance of residuals,  $\sigma_\varepsilon^2$ , from equation (2.5) can be calculated as a byproduct using  $\sigma_\varepsilon^2 = \sigma_B^2 (1 - \gamma^2)$ .

We opted to include  $\mu_B$  in the truncated likelihood function despite the fact, that a proper detrending would make its estimation obsolete. Reasoning is that even robust detrending methods can exhibit a small bias, for which the estimation of  $\mu_B$  can (partly) correct. The needed numerical optimization is non-iterative and thus advantageous against the Kalman filter based approach with Bayesian updating by Huisman and De Jong (2003), which is known to be slow in convergence (see also Janczura and Weron, 2012). Further in case that one opts to neglect  $\mu_B$  solely one parameter needs to be estimated and the procedure could even be run using a spreadsheet software.

### 3.2. Estimation of the spike process and the transition matrix

After having identified the correct parameters for the base process we still have to calculate the probabilities of being in one regime or the other and to calculate the Markov-transition matrix from equation (2.4). Further the mean and standard deviation of the spike process will be derived.

#### 3.2.1. Spike probabilities and transition matrix

In order to be able to calculate the conditional probabilities which form the transition matrix we first need to derive the unconditional probabilities for each of the two regimes. In this context to estimate the unconditional probability of being in the base process,  $\pi_B$ , the conditions implied by the SMS model have to be fulfilled and  $\hat{\mu}_B$  and  $\hat{\sigma}_B$  have to be consistent estimates for the mean and the variance of the distribution that corresponds to the base regime. Using  $\hat{\mu}_B$  and  $\hat{\sigma}_B$  and having decided on the support of the spike process,  $b$  one can calculate  $\mathbb{P}(X_t^B \in (-\infty, b])$ . In difference to  $a$ , which gives a quantile of the whole time-series,  $b$  does resemble a quantile of the base process. We decided to implement this further assumption of spikes being greater than a

---

<sup>2</sup>It has to be noted that during a simulation study we found the  $Q_n$  based AR(1) coefficients, which are unequal zero, to be consistently (although by no means significantly) smaller than their theoretical value. Reason for this observation seems to be the fact, that for an AR process  $Q_n$  might tend to be (very slightly) biased by construction. To correct for this problem we propose to solely use differences of observations that are far enough away from each other, to not be influenced through the autocorrelation of the process anymore:

$$\hat{Q}_n^* = d_{Q_n} \{|X_i - X_j|; (i + 100) < j\}_{(l)}.$$

We choose 100 as an arbitrary lag and found that the change in the calculation of  $Q_n$  mitigates the biasing effect when applied to an AR process.

certain quantile of the base process in order to yield higher efficiency. This is guaranteed whenever  $b > a$ . Thus we can write:

$$\mathbb{P}(X_t^B \leq b) = \mathbb{P}(X_t^B \leq b | R_t = B) = \mathbb{P}(X_t \leq b | R_t = B) = \frac{\mathbb{P}(X_t \leq b)}{\mathbb{P}(R_t = B)},$$

whereas the last equality relies on Bayes' theorem and the fact that  $\mathbb{P}(X_t \leq b \cap R_t = B) = \mathbb{P}(X_t \leq b)$  as  $\mathbb{P}(X_t \leq b) \subseteq \mathbb{P}(R_t = B)$ . Rearranging then yields the probability for an observation of process  $X$  to be in the base regime:

$$\pi_B = \mathbb{P}(R_t = B) = \frac{\mathbb{P}(X_t \leq b)}{\mathbb{P}(X_t^B \leq b)}$$

The numerator can be estimated by the proportion of observations that do not exceed the threshold  $b$ ,  $\widehat{\mathbb{P}}(X_t \leq b) = \frac{1}{T} \sum_{t=1}^T \mathbb{1}_{\{X_t \leq b\}}$ , and the denominator is given by  $\mathbb{P}(X_t^B \leq b) = \Phi((b - \mu_B)/\sigma_B)$ , which can be estimated by  $\widehat{\mathbb{P}}(X_t^B \leq b) = \Phi((b - \hat{\mu}_B)/\hat{\sigma}_B)$ .

After having correctly identified  $\pi_B$  and thus  $\pi_S$  one still has to calculate the transition probabilities from equation (2.4). For this purpose we use  $\hat{\mu}_B$  and  $\hat{\sigma}_B^2$ , assuming (as done before) that  $X_t^B \sim N(\hat{\mu}_B, \hat{\sigma}_B^2)$ . Further we state that for  $c$  being a sufficiently large quantile of  $X_t^B$ , it is appropriate to write  $\mathbb{P}(X_t^B > c) \approx 0$ . We note that under these assumptions  $\mathbb{P}(X_t > c)$  is equivalent to  $\mathbb{P}(X_t^S > c \cap R_t = S)$ . Thus we have

$$\mathbb{P}(X_t > c | X_{t-1} > c) = \mathbb{P}(X_t^S > c \cap R_t = S | X_{t-1}^S > c \cap R_{t-1} = S)$$

and further by independence of the spike process  $X_t^S$  and the regime process  $R_t$

$$= \mathbb{P}(X_t^S > c) \mathbb{P}(R_t = S | R_{t-1} = S).$$

Since  $\mathbb{P}(X_t^S > c) = \mathbb{P}(X_t^S > c | R_t = S) = \mathbb{P}(X_t > c | R_t = S)$  we finally obtain for  $\pi_{SS}$

$$\pi_{SS} = \frac{\mathbb{P}(X_t > c | X_{t-1} > c) \pi_S}{\mathbb{P}(X_t > c)}. \quad (3.4)$$

Using stationarity of the regime process, we further obtain for  $\pi_{BS}$

$$\pi_{BS} = \frac{\pi_S (1 - \pi_{SS})}{1 - \pi_S}. \quad (3.5)$$

The first part of the nominator from equation (3.4) can be calculated by counting all consecutive pairs of observations greater than  $c$  and dividing them by all observations greater than  $c$ ,  $\hat{\pi}_S = 1 - \hat{\pi}_B$  and  $\widehat{\mathbb{P}}(X_t > c)$  can be estimated by counting the number of observations greater  $c$ . The estimator  $\hat{\pi}_{BS}$  is then calculated using  $\hat{\pi}_S$  and  $\hat{\pi}_{SS}$ .

### 3.2.2. Estimating mean, $\mu_S$ , and variance, $\sigma_S^2$ , of the spike process

To calculate the mean and variance of the spike process we make use of the fact, that  $X$  is a finite mixture process. Following Frühwirth-Schnatter (2006) the mean of the spike process can be calculated as  $\mu_X = \pi_B \mu_B + \pi_S \mu_S$ . Rearranging gives:

$$\mu_S = \frac{\mu_X - \pi_B \mu_B}{\pi_S} \quad (3.6)$$

Further the variance of the whole process  $X$  can be written as  $\sigma_X^2 = E(X^2) - \mu_X^2$ . Making use of the fact that  $E(X^2) = \pi_B(\mu_B^2 + \sigma_B^2) + \pi_S(\mu_S^2 + \sigma_S^2)$  we can rewrite the variance of  $X$  as  $\sigma_X^2 = \pi_B(\mu_B^2 + \sigma_B^2) + \pi_S(\mu_S^2 + \sigma_S^2) - \mu_X^2$ . Then solving for  $\sigma_S^2$  yields:

$$\sigma_S^2 = \frac{\sigma_X^2 - \pi_B(\mu_B^2 + \sigma_B^2) + \mu_X^2}{\pi_S} - \mu_S^2. \quad (3.7)$$

Thus the mean and the variance of the spike process can be easily derived using the before calculated parameters.

To get further information concerning the empirical distribution of the spike process nonparametrically Gaussian kernels can be fitted. The basic idea is to first fit a kernel to the complete data-set. Next its theoretic counterpart, which is based on the assumption that the base process is normally distributed with  $\hat{\mu}_B$  and  $\hat{\sigma}_B^2$ , and further scaled down using  $\hat{\pi}_B$ , can be calculated. In a last step the difference between the two kernels (with support  $[max(a, b), \infty)$ ) is multiplied by  $\hat{\pi}_B/\hat{\pi}_S$ , scaling the spike process upwards. We thus get an empirical approximation of the probability distribution function for the spike process with a scale that corresponds to the one of the theoretical distribution of the base process. A graphical illustration is given in Figure 4.

### 3.3. Extension to include the possibility of negative spikes

Electricity prices sometimes exhibit extreme negative log-prices that are not in accordance with the assumption of normality.

The  $Q_n$  based autocorrelation estimator is robust against negative outliers and no further adaption is needed when estimating  $\hat{\gamma}$ .

To robustify the estimation of the standard deviation and the mean of the base process against such rare events apart from an upper truncation we also can consider a lower truncation. In this case we have to define an upper and a lower threshold,  $a_1$  and  $a_2$  with  $y \sim TN(0, \sigma_B^2, a_1, a_2)$  and thus the truncated likelihood function is:

$$\frac{1}{n} \ell(\sigma_B^2) = \frac{-1}{2n} \sum_{t=1}^T \frac{(y_t - \mu_B)^2}{\sigma_B^2} \mathbb{1}_{\{a_2 \leq y_t \leq a_1\}} - \frac{1}{2} \log \sigma_B^2 - \log \left( \Phi \left( \frac{a_1 - \mu_B}{\sigma_B} \right) - \Phi \left( \frac{a_2 - \mu_B}{\sigma_B} \right) \right), \quad (3.8)$$

with  $\mathbb{1}_{\{a_2 \leq y_t \leq a_1\}} = 1$  if  $a_2 \leq y_t \leq a_1$  and zero otherwise while  $n = \sum_{t=1}^T \mathbb{1}_{\{a_2 \leq y_t \leq a_1\}}$ . The resulting estimation is still consistent as long as spikes are greater in absolute value than the corresponding chosen cut-off quantiles of the time-series. Consistent estimates for  $\mu_B$  and  $\sigma_B^2$  can then be again achieved by maximizing the truncated likelihood.

In case that negative outliers are present some slight adjustments have to be further made for the estimation of the spike probabilities and the transition matrix. Define  $b_1$  and  $b_2$  as an upper and a lower threshold corresponding to certain quantiles of the base process that give support for the spike process. Then one can write:

$$\mathbb{P}(X_t \leq b_1 | X_t \geq b_2) = \mathbb{P}(X_t^B \leq b_1 | X_t^B \geq b_2) \mathbb{P}(R_t = B).$$

Alike before, rearranging allows to express  $\pi_B$  in terms of a ratio:

$$\pi_B = \mathbb{P}(R_t = B) = \frac{\mathbb{P}(X_t \leq b_1 | X_t \geq b_2)}{\mathbb{P}(X_t^B \leq b_1 | X_t^B \geq b_2)}.$$

Here the nominator can be estimated by  $\hat{\mathbb{P}}(X_t \leq b_1 | X_t \geq b_2) = \sum_{t=1}^T \mathbb{1}_{\{b_2 \leq X_t \leq b_1\}} / \sum_{t=1}^T \mathbb{1}_{\{X_t > b_2\}}$ . The denominator is given by  $\mathbb{P}(X_t^B \leq b_1 | X_t^B \geq b_2) = (\Phi(\frac{b_1 - \mu}{\sigma}) - \Phi(\frac{b_2 - \mu}{\sigma})) / (1 - \Phi(\frac{b_2 - \mu}{\sigma}))$  and thus can be estimated by  $\hat{\mathbb{P}}(X_t^B \leq b_1 | X_t^B \geq b_2) = (\Phi(\frac{b_1 - \hat{\mu}_B}{\hat{\sigma}_B}) - \Phi(\frac{b_2 - \hat{\mu}_B}{\hat{\sigma}_B})) / (1 - \Phi(\frac{b_2 - \hat{\mu}_B}{\hat{\sigma}_B}))$ .

In case of frequent negative spikes an additional second spike regime can be introduced to include these observations in the model (see e.g. Kosater and Mosler, 2006, De Jong and Schneider, 2009 and Janczura and Weron, 2010). The extension of the estimation method to this case is straightforward.

## 4. Simulation-study

In order to check the performance of the proposed techniques we conduct a simulation-study. For this purpose we will run simulations of the Markov-switching process presented in Section 2 with

uniformly distributed spikes. For each specification we simulate 1000 sample trajectories for 1000, 2000, 5000 and 10 000 observations. Here 1000 observations can be interpreted as a lower bound of the number of observations that are normally used in applications to data while 10 000 observations will give a good idea of the underlying asymptotics.

As already mentioned, we first make use of the fact that the discrete version of the standard mean-reverting model is a simple AR(1)-process. Further we do not account for seasonality nor use a constant for the simulation-study. The resulting process of the base regime is then

$$X_t^B = \gamma X_{t-1}^B + u_t, \quad u_t \sim N(0, \sigma_u^2). \quad (4.1)$$

Next, the independent spike process,  $X_t^S$  is simulated from a uniform distribution,

$$X_t^S \sim U(d, e). \quad (4.2)$$

The regime process  $r_t$  takes value one for a spike and zero otherwise. It is simulated as a series of Bernoulli distributed random variables with probabilities depending on the previous state, that is,

$$r_t | r_{t-1} \sim \text{Ber}(r_{t-1} \pi_{SS} + (1 - r_{t-1}) \pi_{BS}).$$

The resulting Markov regime-switching process  $X_t$  is given by

$$X_t = (1 - r_t) X_t^B + r_t X_t^S.$$

After having set the framework, we choose parameter values, based on results from former publications using regime-switching models for electricity markets (see Janczura and Weron, 2010 and references therein) to conduct the simulation-studies. At the same time we restrict the spike process not to yield values that are smaller than the 90% quantile of the base process:

Table 1: *parameters used for simulations*

$\gamma$	$\sigma_u$	$\mu_S$	$\sigma_S$	$\pi_{BS}$	$\pi_{SS}$	$\pi_S$
0.7	0.1414	2.2538	1.1547	0.06	0.4	0.0909

To check if the simulation framework can be interpreted as realistic we give a graphical example. Figure 1 shows how the characteristics of a representative simulation, closely resemble the characteristics of NSW.

For the estimation we use the one-sided estimation techniques, applicable when negative outliers are not present. We choose  $a_1$  to equal the 80% quantile of the time-series whereas  $b_1$  and  $c$  are chosen to correspond to the 90% and 99.9% quantile of the base process. Table 2 shows in the first and second line the parameters used for the simulation of the process. Lines 3,5,7 and 9 give the corresponding estimates for an increasing amount of observations. The standard deviations of the estimates over the 1000 trajectories are given in parenthesis. It can be seen that the estimates do not appear to be biased. This conclusion still holds when increasing the number of observations to 10 000 for which the standard deviations are substantially smaller than for 1000 observations.

To show the relative strength of our approach we also calculate the MS model of De Jong (2006) with the Kalman filter based approach and assuming compound poisson distributed spikes, while using the same simulation framework as presented before. The corresponding results are given in Table 3.

Although the parametric approach appears to yield more efficient estimates, it can be stated (see underlined estimates) that the parameters, describing the distribution of the spike process are significantly larger than the simulated ones at the corresponding 1% significance level. The fact that the parameters are significantly biased can lead to serious problems when valuing derivatives. The same holds when forecasting electricity prices. It thus can be concluded that the semiparametric approach is advantageous whenever the distribution of spikes is not sufficiently known. Another important point when considering to use this kind of MS type models in the context of high-frequency data, is the immense speed difference between the algorithm proposed by Huisman and De Jong (2003) and our approach (studies using Australian spot prices in the context of

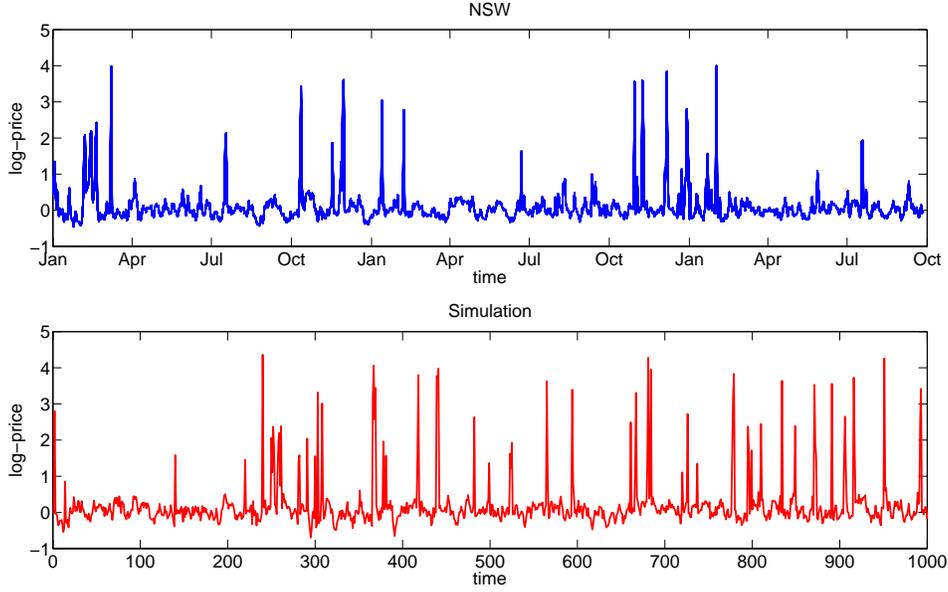


Figure 1: Simulated time-series with uniformly distributed spikes and robustly deseasonalized NSW starting in January 2004.

high-frequency data are amongst others conducted by Chan et al., 2008, Thomas et al., 2011, Christensen et al., 2012 and Eichler et al., 2012). The last column in Table 3 gives the needed computation time in minutes to run the simulation-study. It can be seen, that our algorithm is around 2000 times faster than the Kalman filter based one.

## 5. Data and empirical evaluation

To evaluate our approach we will use electricity prices from four regions in Australia, namely Victoria (VIC), New South Wales (NSW), Queensland (QLD) and South Australia (SA) for the period from 1st of January 2004 till 30st of June 2010. The data is publicly available at <http://www.aemo.com.au/>. In accordance with the literature we choose to analyze daily log-prices instead of prices. Reason is the idea that under abstraction from spikes, electricity prices are

Table 2: Parameters estimated with the SMS model for simulated model with uniformly distributed spikes. Average parameter estimates for sample sizes of  $T = 1000, 2000, 5000, 10000$  were obtained over 1000 repetitions, the empirical standard deviation of the estimates is given in parentheses. Line two gives the parameters used for the simulations.

	$\gamma$	$\sigma_u$	$\mu_s$	$\sigma_s$	$\pi_{BS}$	$\pi_{SS}$	$\pi_S$
Theory	0.7	0.1414	2.2538	1.1547	0.06	0.4	0.0909
1000	0.6896 (0.0254)	0.1431 (0.0068)	2.3701 (0.6534)	0.9682 (0.4118)	0.0590 (0.0122)	0.3951 (0.1139)	0.0924 (0.0269)
2000	0.6916 (0.0187)	0.1430 (0.0050)	2.2965 (0.3717)	1.0527 (0.2963)	0.0603 (0.0080)	0.3943 (0.0800)	0.0923 (0.0187)
5000	0.6928 (0.0114)	0.1427 (0.0031)	2.2652 (0.2165)	1.1225 (0.1327)	0.0605 (0.0050)	0.3957 (0.0500)	0.0917 (0.0120)
10 000	0.6931 (0.0082)	0.1427 (0.0022)	2.2574 (0.1505)	1.1404 (0.0765)	0.0606 (0.0036)	0.3957 (0.0355)	0.0915 (0.0086)

Table 3: Parameters estimated with the SMS and the MS model for simulated model with uniformly distributed spikes. Average parameter estimates for sample size of  $T = 10000$  were obtained over 1000 repetitions, the empirical standard deviation of the estimates is given in parentheses. Line two gives the parameters used for the simulations.

	$\gamma$	$\sigma_\epsilon$	$\mu_s$	$\sigma_s$	$\pi_{BS}$	$\pi_{SS}$	$\pi_S$	time
Theory	0.7	0.1414	2.2538	1.1547	0.06	0.4	0.0909	-
SMS	0.6931 (0.0082)	0.1427 (0.0022)	2.2574 (0.1505)	1.1404 (0.0765)	0.0606 (0.0036)	0.3957 (0.0355)	0.0915 (0.0086)	2 -
MS	0.6889 (0.0085)	0.1418 (0.0012)	<u>2.9553</u> (0.0684)	<u>1.4001</u> (0.0338)	0.0597 (0.0026)	0.3993 (0.0182)	0.0904 (0.0043)	4058 -

Table 4: Descriptive statistics for detrended log-prices

	VIC	NSW	QLD	SA
Mean	0.0511	0.1004	0.0889	0.0903
Median	0	0	0	0
Std dev	0.4006	0.4704	0.4447	0.5216
Skewness	3.6320	4.2087	3.5392	3.7000
Kurtosis	30.9734	26.6940	25.7021	26.9188

approximately log-normally distributed. The calculation of daily prices is done by taking the arithmetic mean of the 48 half-hourly prices that are recorded during a day.

Before comparing the SMS model with the MS model we detrend the data sets. To do so we follow the idea of Rambharat et al. (2005) and subtract a time-varying mean, which we calculate using robust locally weighted regression (see Cleveland, 1979) with a window of 100 observations.<sup>3</sup> To account for weekly seasonality we subtract the median for weekdays, Saturdays and Sundays.

The characteristics of the detrended daily Australian log-prices are reported in Table 4. As can be seen, the time-series exhibit high kurtosis and skewness with values that indicate non-normally distributed data sets. This observation is supported by the positive difference between the mean and the median.

The resulting fit for the different time-series is given in Figure 2. We observe that very few extreme negative log-prices do occur for three of the four time-series. As those observations are not in line with the assumption of a normally distributed base process, we do use the estimation framework which is robust to negative outliers when calibrating the SMS model. The quantiles for  $a_1$ ,  $b_1$  and  $c$  are chosen to be the same as already for the simulation-study. The quantiles for  $a_2$  and  $b_2$  are chosen to be the 1% quantile for the whole time-series and the 0.1% quantile for the base process. To allow for comparable conditions we delete negative observations which under the assumption of a normally distributed base process would have a probability smaller than 0.1% before fitting the MS model. We thus delete between 3 and 23 observations depending on the time-series.

Next we control the stochastic part of log-prices (to which we will further refer as log-prices) for the order of autocorrelation. To do so we calculate the partial autocorrelations for lags one to ten for each of the time-series. To get results that are robust to outliers we replace the needed autocorrelations by their robust counterparts using the  $Q_n$  estimator. The results are shown in Figure 3. The autocorrelation-function (AC) appears to decrease exponentially. This is in line with the assumption that the log-prices follow an AR(1) process. Further support for this assumption is given by the partial autocorrelation functions (PAC) and the corresponding 5% significance levels. For the first lag, partial autocorrelations are far larger than the corresponding 5% significance levels and thus can be considered significantly different from zero for all four time-series. For larger lags partial autocorrelations are wether not significantly different from zero or do slightly pass the corresponding level.

<sup>3</sup>The needed procedure is implemented in the Curve Fitting Toolbox of MATLAB.

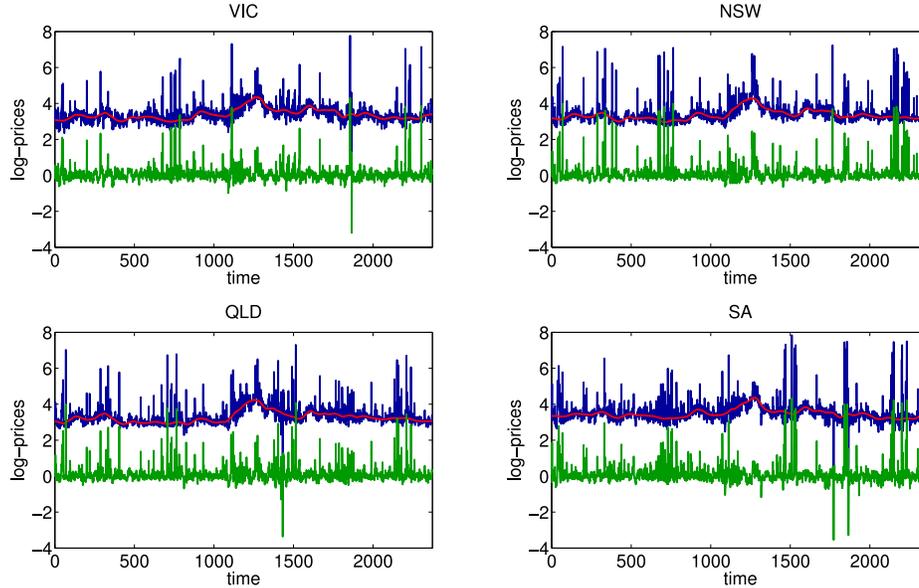


Figure 2: Time-varying mean using robust locally weighted regression, log-prices and detrended log-prices: The red lines show the time-varying means, while the blue lines exhibit the log-prices of the different time-series. The green lines finally show the difference between log-prices and time-varying mean. All data sets are plotted from 1st of January 2004 till 30st of June 2010.

Table 5: Parameters estimated with the MS and the SMS model for log-prices from January 2004 to June 2010

		$\gamma$	$\mu_B$	$\sigma_\varepsilon$	$\mu_S$	$\sigma_S$	$\pi_S$	$\pi_{BS}$	$\pi_{SS}$
VIC	MS	0.6818	-0.0215	0.1438	0.9210	0.5984	0.0819	0.0325	0.6197
	SMS	0.6986	-0.0285	0.1342	0.7796	0.8846	0.0986	0.0226	0.7934
NSW	MS	0.7065	-0.0161	0.1232	0.9619	0.7161	0.1233	0.0453	0.6764
	SMS	0.7083	-0.0263	0.1134	0.8827	0.8449	0.1394	0.0212	0.8688
QLD	MS	0.6939	-0.0177	0.1294	0.8531	0.6613	0.1291	0.0547	0.6309
	SMS	0.6782	-0.0270	0.1220	0.8268	0.7934	0.1358	0.0442	0.7185
SA	MS	0.6237	-0.0143	0.1518	1.1695	0.8127	0.0985	0.0516	0.5273
	SMS	0.6308	-0.0140	0.1483	1.2126	1.3235	0.0850	0.0483	0.4803

After having analyzed and cleaned the time-series, we now turn our focus on the empirical application. For this purpose we first compare the parameters yielded by the different models and check if the results are in line with the theoretical expectations. As the robust locally weighted regression and the median, which were used to account for the seasonal trend, can still be expected to exhibit a small bias we extend the base process of the MS and the SMS model by a constant, in order to model the mean of the base process,  $\mu_B$ .

Columns 3 to 5 of Table 5 show that the parameters concerning the base process are close to each other for the MS and SMS model, although the standard deviation of the error terms is consistently smaller for the SMS model. As already seen in the simulation-study, concerning the spike process the SMS model yields parameters that differ from the ones, that are based on the MS model. For the first moments no clear pattern can be established meanwhile for second moments the SMS appears to yield higher values for the data-sets at hand. Further the estimated overall probability of spikes is for three out of four data-sets higher when using the SMS model. This is in accordance with the fact, that the SMS model yielded smaller values for the standard

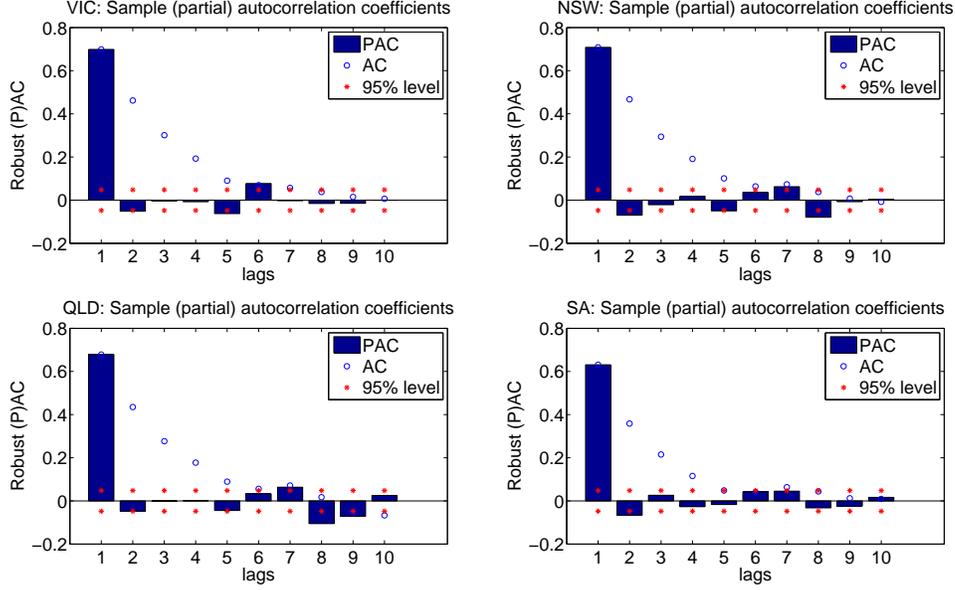


Figure 3: (P)ACFs for log-prices of the four time series under consideration from January 2004 till June 2010: The blue bars show the autocorrelation functions, while the blue circles exhibit the partial autocorrelation function. The red asterisks give the corresponding 95% significance levels for the partial autocorrelation functions.

deviations of the error terms from the base process. Finally the estimated conditional probabilities to move from the base to the spike process are lower for the SMS model whereas the conditional probabilities to stay in the spike process are higher when using SMS with exception of  $\pi_{SS}$  for SA. A possible explanation for this fact could be that because of computational intensity, the algorithm of Huisman and De Jong (2003) can just handle a certain amount of consecutive spikes. We followed the recommendation of the authors and set the maximum number of consecutive spikes, that the algorithm accounts for equal to 10 (see also Janczura and Weron, 2012). The algorithm should thus only be used if the probability of more than 10 consecutive observations from the spike regime is negligible, whereas our approach does not exhibit any restrictions concerning the possible number of consecutive spikes.

The quality of our model is assessed through a graphical illustration of nonparametric density estimates using gaussian kernels with a bandwidth of 0.05.<sup>4</sup> The first row of Figure 4 shows estimation results that are based on the SMS model. The theoretic density of the base process (represented by the red, dotted curve) relies on the assumption that the base process is normally distributed with  $\hat{\mu}_B$  and  $\hat{\sigma}_B^2$ . It is scaled down by multiplying it with  $\hat{\pi}_B$  and fits the kernel (given by the blue curve) of the corresponding data set quite well up to a certain upper threshold. The green curve gives the difference between the theoretical and the empirical density (multiplied by the factor  $\hat{\pi}_B/\hat{\pi}_S$  in order to have the same scale for the spike distribution as for the theoretical distribution of the base process) and has support  $[max(a_1, b_1), \infty)$ . The fact that it starts at zero with respect to the y-axis for all four time-series, indicates that the assumption of spikes not being smaller than the 80% quantile of the whole data set and 90% of the base process is sensible for the data sets at hand. We thus get an approximation of the pdf for the spike process. It can be seen that there is no clear pattern that could be associated with a specific theoretical distribution. The lower row of Figure 4 shows the corresponding results for the MS model. From comparing the center of the empirical kernel of the data and the theoretical one for the base process it can

<sup>4</sup>The bandwidth was chosen with the bandwidth selection algorithm proposed by Botev et al. (2010).

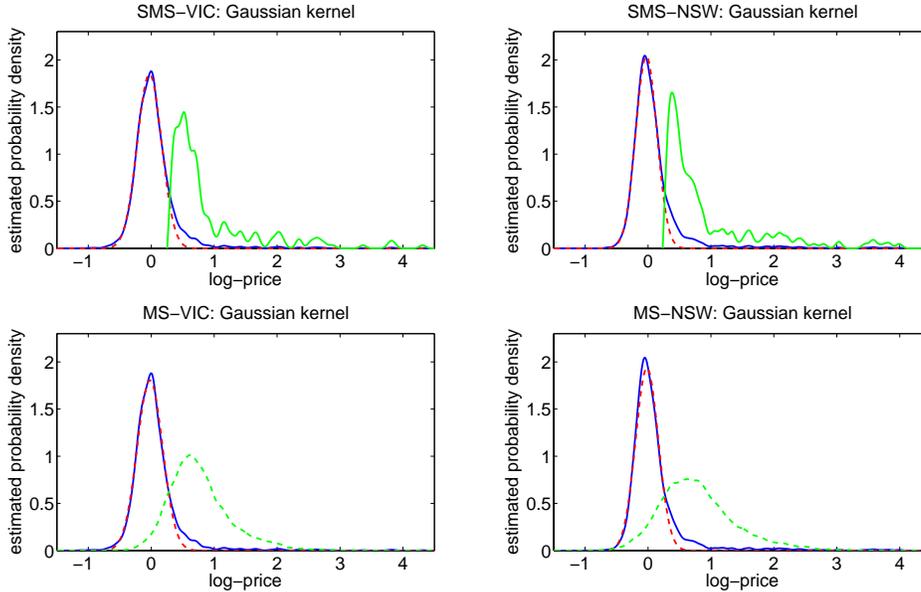


Figure 4: Nonparametric density estimates via Gaussian kernel fits: The red, dotted curve shows the theoretical kernel of the base process. It is calculated using  $\hat{\mu}_B$  and  $\hat{\sigma}_B$ . The resulting kernel density is then scaled down by multiplying it with  $\hat{\pi}_B$ . The blue curve shows the empirical density, which emerges when using log-prices. The green line shows the difference between the theoretical (scaled) density and the empirical one with support  $[\max(a_1, b_1), \infty]$ . To show how the resulting density of the spike process itself can be expected to look we divide the resulting density by  $\hat{\pi}_S$ . The same is done for the dotted green line, which shows the theoretical density for the spikes, which is used by the MS model. The time series under consideration are VIC and NSW and comprise again observations from January 2004 till June 2010.

be intuited that the fit is slightly worse than for the SMS model. Further it can be seen, that the theoretical distribution of the spike process (given in form of the green dotted line) is not sufficiently capable of modeling an empirical spike distribution as the one given in the upper row of Figure 4. These results indicate that the proposed semiparametric approach is recommendable in order not to risk producing biased estimates.

## 6. Conclusion

Although MS models do demand a certain amount of prior assumptions, they are said to be exceptionally well suited to model electricity price characteristics. Reasoning is their ability to calculate transition probabilities from one state to another. We propose a semiparametric extension, using robust estimation techniques. The advantage of our approach is that neither complex forms of numerical optimization nor any distributional assumption concerning the spike process are needed. The approach offers a clear advantage in computational complexity and (under realistic conditions) assures against biased estimates when the distribution of the spike process is unknown. As a side product, the procedure allows to account for negative outliers without making it necessary to model them specifically. In this context extensions to account for a third regime with negative spikes as proposed in De Jong and Schneider (2009) and Janczura and Weron (2010) would be straight forward. However the data at hand exhibited very scarce occurrence of negative outliers, making it advantageous to robustify the estimation against these occurrences without modeling them explicitly. Our approach is further in line with contemporaneous literature in which several different distributions for the spike process (see Bierbrauer et al., 2007 and Janczura and Weron, 2009) were proposed. We take this idea consequently to a next step by leaving the spike distribution apriori unspecified. Further Janczura and Weron (2012) recently proposed an algorithm to

reduce the computational burden, that is faced by the algorithm used by Huisman and De Jong (2003). Our proposed algorithm does not just lower the computational burden (measured in computational speed) by a factor of 2000. For the simplest model specification without a constant for the base model it even can be calculated using spreadsheet software.

To underpin our theoretical arguments, we ran a simulation study using uniformly distributed spikes. The results show that the parametric nested MS model of Huisman and De Jong (2003) estimated by a Kalman filter based approach yields biased estimates even when using the flexible specification of a compound poisson distribution to model the spike process. The SMS model however yields unbiased estimates. This result holds for small samples as well as asymptotically.

Further we applied the different approaches to four Australian electricity spot log-price time-series. It can be stated that the only noteworthy difference in the parameters of the base process concerned the standard deviation of the residuals, which were consistently smaller when using the SMS model than when using the MS model. The estimates for the distribution of the spike process however do (in accordance with the results from the simulation study) all differ between the models. Furthermore the MS model seems to assign a lower probability to observe spikes than the SMS model whenever the conditional probability of consecutive spikes is high. Reason could be that because of the computational burden of the Kalman filter based algorithm, the probability of more than 10 consecutive spikes is neglected when estimating the parameters of the MS model. To control if the needed assumptions for the SMS model are well chosen we evaluated the empirical density of two different data-sets and the density of the spike process graphically with help of a nonparametric gaussian kernel estimation. The results indicate, that the assumptions for the SMS model are sensible and that a parametric modeling of the spike process can lead to inadequate results.

Although we were able to show the good properties through simulations and encouraging results using Australian spot data, it would still be interesting to extend the presented model further. Possible extensions could concern time-varying volatility for the base process. This was already proposed by Janczura and Weron (2009) and shown to be important for only some of the electricity spot markets by Lindstroem and Regland (2012). Extending our approach to allow the modeling of two different spike regimes would be straight forward and might be interesting for alternate data-sets in which negative prices are more frequent. Furthermore the use of a multivariate framework in order to yield further results in the line with the publications of De Jong and Schneider (2009) and Lindstroem and Regland (2012) would be of great interest. The last point is even more interesting when considering the computational ease that the proposed approach offers compared to former estimation procedures used in the literature. Implementing and testing these extensions will be of future interest for the authors.

## References

- Bierbrauer, M., Menn, C., Rachev, S. T., Trueck, S., 2007. Spot and derivative pricing in the EEX power market. *Journal of Banking & Finance* 31, 3462–3485.
- Botev, Z. I., Grotowski, J. F., Kroese, D. P., 2010. Kernel density estimation via diffusion. *Annals of Statistic* 38, 2916–2957.
- Boudt, K., Croux, C., Laurent, S., 2011. Robust estimation of intraweek periodicity in volatility and jump detection. *Journal of Empirical Finance* 18, 353–367.
- Chan, K. F., Gray, P., Van Campen, B., 2008. A new approach to characterizing and forecasting electricity price volatility. *International Journal of Forecasting* 24, 728–743.
- Christensen, T. M., Hurn, A. S., Lindsay, K. A., 2012. Forecasting spikes in electricity prices. *International Journal of Forecasting* 28, 400–411.
- Cleveland, W., 1979. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* 74, 829–836.

- Clelow, L., Strickland, C., 2000. *Energy Derivatives - Pricing and Risk Management*. Lacima Publications, London.
- De Jong, C., 2006. The nature of power spikes: A regime-switch approach. *Studies in Nonlinear Dynamics & Econometrics* 10, Article 3.
- De Jong, C., Schneider, S., 2009. Cointegration between gas and power spot prices. *The Journal of Energy Markets* 2, 27–46.
- Deng, S., 2000. Stochastic models of energy commodity prices and their applications: Mean-reversion with jumps and spikes. Working paper PWP-073, POWER, University of California Energy Institute.
- Eichler, M., Grothe, O., Manner, H., Tuerk, D., 2012. Modeling spike occurrences in electricity spot prices for forecasting. Working paper 12/029, METEOR Research Memorandum, Maastricht University.
- Frühwirth-Schnatter, S., 2006. *Finite Mixture and Markov Switching Models*. Springer, New York.
- Haldrup, N., Nielsen, F. S., Nielsen, M. O., 2010. A vector autoregressive model for electricity prices subject to long memory and regime switching. *Energy Economics* 32, 1044–1058.
- Haldrup, N., Nielsen, M. O., 2006. A regime switching long memory model for electricity prices. *Journal of Econometrics* 135, 349–376.
- Huisman, R., De Jong, C., 2003. Option pricing for power prices with spikes. *Energy Power Risk Management* 7, 12–16.
- Huisman, R., Mahieu, R., 2003. Regime jumps in electricity prices. *Energy Economics* 25, 425–434.
- Janczura, J., Weron, R., 2009. Regime-switching models for electricity spot prices: Introducing heteroskedastic base regime dynamics and shifted spike distributions. Working paper 18784, MPRA.
- Janczura, J., Weron, R., 2010. An empirical comparison of alternate regime-switching models for electricity spot prices. *Energy Economics* 32, 1059–1073.
- Janczura, J., Weron, R., 2012. Efficient estimation of markov regime-switching models: An application to electricity spot prices. *ASTA - Advances in Statistical Analysis*.  
URL <http://www.springerlink.com/content/2r1866007873514u/>
- Kosater, P., Mosler, K., 2006. Can Markov-regime switching models improve power-price forecasts? Evidence from German daily power prices. *Applied Energy* 83, 943–958.
- Lindstroem, E., Regland, F., 2012. Modeling extreme dependence between european electricity markets. *Energy Economics* 34, 899–904.
- Lucia, J. J., Schwartz, E. S., 2002. Electricity prices and power derivatives: evidence from the Nordic power exchange. *Review of Derivatives Research* 5, 5–50.
- Ma, Y., Genton, M., 2000. Highly robust estimation of the autocovariance function. *Journal of Time Series Analysis* 21, 663–684.
- Marazzi, A., Yohai, V. J., 2004. Adaptively truncated maximum likelihood regression with asymmetric errors. *Journal of Statistical Planning and Inference* 122, 271–291.
- Misiorek, A., Trueck, S., Weron, R., 2006. Point and interval forecasting of spot electricity prices: linear vs. non-linear time series models. *Studies in Nonlinear Dynamics & Econometrics* 10, Article 2.

- Mount, T. D., Ning, Y., Cai, X., 2006. Predicting price spikes in electricity markets using a regime-switching model with time-varying parameters. *Energy Economics* 28, 62–80.
- Rambharat, B. R., Brockwell, A. E., Seppi, D. J., 2005. A threshold autoregressive model for wholesale electricity prices. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 54, 287–299.
- Robinson, T. A., 2000. Electricity pool prices: a case study in nonlinear time-series modelling. *Applied Economics* 32, 527–532.
- Rousseeuw, P. J., Croux, C., 1992. Explicit scale estimators with high breakdown point. In: Dodge, Y. (Ed.), *L<sub>1</sub>-Statistical Analysis and Related Methods*. North-Holland, Amsterdam, pp. 77–92.
- Rousseeuw, P. J., Croux, C., 1993. Alternatives to the median absolute deviation. *Journal of the American Statistical Association* 88, 1273–1283.
- Thomas, S., Ramiah, V., Mitchell, H., Heaney, R., 2011. Seasonal factors and outlier effects in rate of return on electricity spot prices in Australia’s national electricity market. *Applied Economics* 43, 355–369.
- Trueck, S., Weron, R., Wolff, R., 2007. Outlier treatment and robust approaches for modeling electricity spot prices. Working paper 4711, MPRA.
- Weron, R., 2006. *Modeling and forecasting electricity loads and prices: A statistical approach*. Wiley, Chichester.
- Weron, R., Bierbrauer, M., Trueck, S., 2004. Modeling electricity prices: jump diffusion and regime switching. *Physica A* 336, 39–48.