Non-linear Modeling of Electricity Price: Self Exciting Threshold Auto-Regressive Approach

Miroslav Čulík, Jiří Valecký[1]

Abstract

The aim of this paper is to estimate and test non-linearities in the electricity prices of three selected regions (California, Nord Europe and Austria). To exploit non-linearity, we apply the SETAR (Self Exciting Threshold Auto-Regressive) models that imply and distinct regimes in time series dynamics with potentially different parameters (and thus dynamics properties) of each regimes. We find the most appropriate SETAR model for modelling electricity prices at selected markets, next we perform the statistical verification of each model and we also find out if our model outperform the linear AR model.

Keywords

Electricity, electricity price, regime-switching model, SETAR model, non-linear time series.

1 Introduction

In many parts of the world, sector of electricity generation is gradually converting to the competitive market structure replacing traditional monopolistic environment and therefore there arises the needs to model the time series of electricity price.

An accurate modelling and forecasting of electricity prices including analyzing of factors affecting them has become a very important tool both for generators and consumers. In a short time period, the generating company needs to forecast electricity prices to set its generating strategy and to optimally schedule energy resources. This is important for the reason of necessity of profit planning and forecasting and that is why accurate electricity prices modelling and forecasting is crucial information for any decision-making. Customers needs short-time forecast of electricity prices for the same or similar reasons as producers.

It is also necessary to point out, that electricity prices gather characteristics which reflects in the time series evolution: high frequency, non constant mean, autocorrelation, non-normal distribution, heteroskedasticity, seasonality, high volatility and high frequency of occurrence of unusual prices etc. This can incur the occurrence of outages, blackouts, and price peaks, which happen seldom in the regulated environment.

There is wide range of papers concentration on modelling, forecasting electricity prices. A group authors have tried to develop models for electricity prices at European electricity market. Results from Čulík, Valecký (2007), Čulík, Valecký (2008), etc. confirm the fact, that in the short time, daily electricity prices evolve randomly, but in the long-run period have the tendency to revert to the long-run level. Similarly, Ecsribano et al (2002) analyze the factors affecting electricity prices and the presence of mean-reversion process at five different electricity markets. They conclude that in five markets analyzed using daily data; Argentina,

[1] Ing. Miroslav Čulík, Ph.D., Ing. Jiří Valecký, VŠB – Technical University of Ostrava, Faculty of Economics, Department of Finance, Sokolska třída 33, 70800 Ostrava, Czech Republic, email: miroslav.culik@vsb.cz, jiri.valecky@vsb.cz.
This paper is solved within the project support by the Grant Agency of the Czech Republic (GAČR) Nr. 402/07/P121.

Australia (Victoria), New Zealand (Hayward), Nordpool, PJM and Spain, equilibrium elektricity prices are mean-reverting. Similar results can be found in Deng, Jiang (2005), Kian, Keyhani (2001), Garcia, Contreras, Akkern (2003), Guirguis, Felder (2004), Catalao et al (2007) etc. Next, Bunn, Karakatsani (2003) focused on explanation of the extraordinary stochastic properties of electricity price time series for example how prices react to temporal market irregulation, reaction conditional volatility to past volatility and shocks etc.

We employ the Self Exciting Threshold Auto-Regressive model and verify if application of these models gives better results than linear. We aim at the comparison of linear and non-linear models if it is possible improving data fitting and diagnostic checks of model residuals. On the one hand, we show that they are more appropriate for modelling the financial time series than linear models, but on the other hand, we also conclude that using of non-linear SETAR models does not improve the diagnostic checks of residuals in the sense of heteroscedasticity and non-normality testing.

The paper is organized as follows: SETAR model is described in Section 2 including the description of estimation method and constructing confidence interval of estimated parameters. Section 3 is devoted to statistical verification of the model in the sense of residual testing and particular estimated model and their comparison with linear model are presented in Section 4. Section 5 concludes this paper.

2 Model description

Let $p_1,\dots,p_k$ be an integer positive numbers representing the order of particular autoregressive models and d a delay parameter, the general SETAR model of k regimes takes the form of

$$y_t = \sum_{i=1}^{k}\left(\alpha_{i0} + \alpha_{i1} y_{t-1} + \cdots + \alpha_{i,p_i} y_{t-p_i}\right) I\left(y_{t-d} \in C_i\right) + \varepsilon_t , \tag{1}$$

where α_{ij} are slope parameters, $I(\ )$ is indicator function

$$I(\ ) = \begin{cases} 1 & \text{if} \quad y_{t-d} \in C_i \\ 0 & \text{otherwise} \end{cases}, \tag{2}$$

ε_t is error term and $\{C_i\}$ forms a partition of $(-\infty,+\infty)$ in the sense that $\sum_{i=1}^{k} C_i = (-\infty,+\infty)$ and $C_i \cap C_j = \varnothing$ for all $i \neq j$.

A general two-regime SETAR model is defined as

$$\begin{aligned} y_t &= \left(\alpha_{10} + \alpha_{11} y_{t-1} + \cdots + \alpha_{1,p_1} y_{t-p_1}\right) I\left(y_{t-d} \le r\right) \\ &\quad + \left(\alpha_{20} + \alpha_{21} y_{t-1} + \cdots + \alpha_{2,p_2} y_{t-p_2}\right) I\left(y_{t-d} > r\right) + \varepsilon_t \end{aligned}, \tag{3}$$

where r is threshold parameter. If the model takes on the form (3), the particular linear autoregressive process is partitioned by threshold value y_{t-d} with delay d and in assistance with threshold parameters r. In this paper, the process $\{\varepsilon_t\}$ we assume to be iid $\left(0,\sigma^2\right)$, although it can be also heteroskedastic.

2.1 Estimation of SETAR model

The model form of Equation (3) can be rewritten in the following representation,

$$y_t = \boldsymbol{\alpha}_1' \mathbf{Y}_{t,p_1}\left(I\left(y_{t-d} \le r\right)\right) + \boldsymbol{\alpha}_2' \mathbf{Y}_{t,p_2}\left(I\left(y_{t-d} > r\right)\right) + \varepsilon_t , \tag{4}$$

where $\mathbf{Y}_{t,p_i} = \left(1, y_{t-1},\dots, y_{t-p_i}\right)'$ and $\boldsymbol{\alpha}_i = \left(\alpha_{0,i}, \alpha_{1,i},\dots,\alpha_{p_i,i}\right)'$ for $i = 1,2$.

The unknown parameters $\boldsymbol{\alpha} = \left(\boldsymbol{\alpha}_1', \boldsymbol{\alpha}_2'\right)'$ and threshold parameter *r* must be estimated on the observed data $\mathbf{Y} = (y_1, \dots, y_T)$ and delay parameter *d* the order of p_i is needed to determined. For this purpose, the sequential conditional LS estimator is employed under the auxiliary condition that process e_t is $\text{iid}\left(0, \sigma^2\right)$. Then under this condition, the LS estimator is equivalent to maximum likelihood estimation.

LS estimation of parameters for given value of *r* is as follows,

$$\hat{\boldsymbol{\alpha}}(r) = \left(\sum_{t=1}^{T} \mathbf{Y}_t(r)\mathbf{Y}_t(r)'\right)^{-1} \left(\sum_{t=1}^{T} \mathbf{Y}_t(r) y_t\right), \tag{5}$$

with residuals $\hat{\mathbf{e}}_t(r) = y_t - \hat{\boldsymbol{\alpha}}(r)' \mathbf{Y}_t(r)$ and residual variance

$$\hat{\sigma}_e^2(r) = \frac{1}{T}\sum_{t=1}^{T} \hat{e}_t(r)^2 . \tag{6}$$

In order to estimate the parameter *r*, ordinary LS regression is run, setting $r = y_{t-1}$ for all $y_{t-1} \in R$ and for each regression compute the residual variance $\hat{\sigma}_e^2(r)$; then pick the value of *r* corresponding to the smallest variance, thus

$$\hat{r} = \min_{r \in R} \hat{\sigma}_e^2(r), \tag{7}$$

where $R = [\underline{r}, \overline{r}]$ is a set of all possible threshold parameters comprises all observed data and $\underline{r} = \min(\mathbf{Y})$, $\overline{r} = \max(\mathbf{Y})$. It is obvious that one needs to run *T* regressions in order to find parameters $\hat{\boldsymbol{\alpha}} = \hat{\boldsymbol{\alpha}}(\hat{r})$.

The same problem arises in determination the delay variable $d \in \left[1, \overline{d}\right]$, where *d* hat is the maximum considered delay. It follows that the amount of *T* regressions is not final. The minimization problem of Equation (7) is augmented to include a search over *d*, so instead of *T* regressions, the search method requires the amount of $T\overline{d}$ regressions and the parameters are used for estimating the slope parameters that satisfy following function

$$\left(\hat{r}, \hat{d}\right) = \min_{r \in R, d} \hat{\sigma}_e^2(r, d). \tag{8}$$

Finally, we add some remarks concerning the practical implementation of this framework. In practical use of non-linear models, practitioner can find several appropriate models for fitting data. Therefore, some goodness of fit measures of an estimated model was developed. We present Akaike's information criterion for *k* regimes in the form of

$$AIC = \sum_{i=1}^{k} T_i \ln \hat{\sigma}_i^2 + 2\left(p_i + 1\right). \tag{9}$$

It is also necessary to further noted that for the reliable model estimation the set of threshold parameters *R* must be selected so that each regime contains the sufficient observations. Therefore, the set of threshold parameter *R* is not bounded by the observed data, but by the technique ensuring sufficient number of observation in each regime. For instance, the 15th and 85th quartile are used to determinate the boundary of set *R*.

2.2 Confidence intervals

To test the statistical significance of estimated parameters, the confidence intervals are necessary to construct. Firstly, we explain the difficulties occurring in constructing the confidence interval of threshold parameter and afterwards we present the same for the slope parameters.

The confidence interval of threshold parameter is given by

$$\hat{\Gamma}(r)=\left\{r: LR_T(r)\le z_\beta\right\}, \tag{10}$$

where LR_T is likelihood ratio for the null hypothesis $H_0: \hat{r}=r_0$ in the form of

$$LR_T(r)=T\left(\frac{\hat{\sigma}_\varepsilon^2(r_0)-\hat{\sigma}_\varepsilon^2(\hat{r})}{\hat{\sigma}_\varepsilon^2(\hat{r})}\right) \tag{11}$$

and z_β is β-level critical value that is available in Hansen (1997, 2000). The graphical method of finding the confidence interval relies on plotting values of $LR_T(r)$ against r and drawing the horizontal line at value of z_β. Needless to say, there might arise a problem in practice because the region can be disjoint and for that reason difficult applicable. Therefore, the convexified region is constructed and used $\hat{\Gamma}(r)=[\underline{r},\overline{r}]$, where $\underline{r}=\left\{r:\min\left(\hat{\Gamma}(r)\right)\right\}$ and $\overline{r}=\left\{r:\max\left(\hat{\Gamma}(r)\right)\right\}$.

The confidence interval of the estimated slope parameters can be constructed in the standard way as they are in linear models. Let $\hat{r}$ be a estimated threshold value, the α-level confidence interval $\hat{\Theta}_{\hat{\boldsymbol{\alpha}}}(\hat{r})$ of the slope parameters $\tilde{\boldsymbol{\alpha}}$ is given by

$$\hat{\Theta}_{\hat{\boldsymbol{\alpha}}}(\hat{r})=\hat{\boldsymbol{\alpha}}(\hat{r})\pm z_\eta \hat{s}(\hat{r}), \tag{12}$$

where z_α is α-level critical value for the normal distribution and $\hat{s}(\hat{r})$ denotes a standard error. Hansen (1996) pointed out that such constructed confidence intervals are not reliable in case of finite sample because the threshold parameter can be estimated not very precisely and can contamine the estimate of $\tilde{\boldsymbol{\alpha}}$. Therefore, he proposed to take the union of all constructed confidence interval of $\hat{\boldsymbol{\alpha}}(r)$ for all $r\in\hat{\Gamma}(r)$, thus

$$\hat{\Theta}_{\hat{\boldsymbol{\alpha}}}=\bigcup_{r\in\hat{\Gamma}(r)}\hat{\Theta}(r). \tag{13}$$

3 Model verification

After the model estimation, its verification is necessary. Firstly and foremost, the estimated model of Equation (4) has to be statistically significant relative to linear AR model. Consequently, obtained residual $\hat{\mathbf{e}}_t(r)$ have to meet the assumption of the white noise and the slope parameters are needed to be statistically significant.

Firstly, we show the linearity test according to Hansen (1996, 1997) under the conditions that the parameter r is known and ε_t is assumed to be iid. The relevant null hypothesis $H_0: \hat{\boldsymbol{\alpha}}_1=\hat{\boldsymbol{\alpha}}_2$ is tested against hypothesis $H_1: \hat{\boldsymbol{\alpha}}_1\ne\hat{\boldsymbol{\alpha}}_2$.

The relevant *F*-statistics $F_T(\hat{r})$ is equivalent to the supremum over the set *R* of the point-wise test-statistic $F_T(r)$,

$$F_T(\hat{r}) = \sup_{r \in R} F_T(r), \tag{14}$$

where

$$F_T(r) = T\left(\frac{\tilde{\sigma}_\varepsilon^2 - \hat{\sigma}_\varepsilon^2(r)}{\hat{\sigma}_\varepsilon^2(r)}\right) \tag{15}$$

and

$$\tilde{\sigma}_\varepsilon^2 = \frac{1}{T}\sum_{t=1}^{T}\left(Y_t - \mathbf{Y}_t'\tilde{\boldsymbol{\alpha}}_{AR}\right)^2, \tag{16}$$

$$\tilde{\boldsymbol{\alpha}}_{AR} = \left(\sum_{t=1}^{T}\mathbf{Y}_t\mathbf{Y}_t'\right)^{-1}\left(\sum_{t=1}^{T}\mathbf{Y}_t Y_t\right) \tag{17}$$

is OLS estimate of parameters $\tilde{\boldsymbol{\alpha}}_{AR}$ from linear autoregressive model of order *p*.

For slope parameters testing, one is allowed to use standard *T*-test by using the united confidence interval from the Equation (13), see Hannsen (1996, 2000),.

The last step of the statistical verification poses (consists in) the necessity to perform diagnostic checks of residuals $\hat{\mathbf{e}}_t(r)$. Some tests that are used in the traditional linear framework can be applied also to testing of the non-linear models. For instance, Jarque-Bera test can be used for normality testing in both frameworks. On the other hand, common Ljung-Box test does not remain valid; see Eitrheim and Terasvirta (1996).

Here, generalized LM test is employed for serial correlation in an AR(*p*) model of Breusch and Pagan (1979), which is based on the auxiliary regression,

$$\hat{\varepsilon}_t = \beta_1 y_{t-1} + \cdots + \beta_p y_{t-p} + \delta_1\hat{\varepsilon}_{t-1} + \cdots + \delta_q\hat{\varepsilon}_{t-q} + v_t. \tag{18}$$

The LM test for *q*-th order serial dependence in ε_t is obtained as TR^2, where R^2 is coefficient determination from the regression $\hat{\varepsilon}_t$ on $\hat{z}_t$, which $\hat{z}_t$ are relevant partial derivation of non-linear model, thus

$$\hat{\varepsilon}_t = \beta_1\hat{z}_t + \cdots + \beta_p\hat{z}_\theta + \delta_1\hat{\varepsilon}_{t-1} + \cdots + \delta_q\hat{\varepsilon}_{t-q} + v_t, \tag{19}$$

where

$$\hat{z}_t = \partial F\left(y_t;\hat{\boldsymbol{\theta}}\right)\big/\partial\boldsymbol{\theta} \tag{20}$$

and $F\left(y_t;\hat{\boldsymbol{\theta}}\right)$ is non-linear SETAR model of Equation (3) and $\boldsymbol{\theta} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, r, d)$ are estimated parameters.

4 Empirical results

In this part, we present the empirical results that we gathered in daily electricity price modelling by applying non-linear SETAR models. We employed this approach on electricity price data series form California (prices obtained from Energy Information Administration), Nord Europe (prices obtained from Nord Pool) and Austria (prices obtained from Energy Exchange Austria). Data set were obtained consists of eight annual time series from 2006 – 2008 and contains the daily electricity prices. To obtain our data sample, we worked with discrete daily returns.

The accuracy of fitting the time series by non-linear model was compared with accuracy of linear AR model. As criterions for the comparison, the residual variance and results of diagnostic checks were used. Firstly, we estimated all SETAR models and constructed confidence intervals of all parameters. Then we conducted the diagnostic checks and compared gathered empirical results with linear models. Thus, it was necessary to verify if nonlinear model gives better results than linear for the purpose of modelling times series.

4.1 Model estimation

The crucial problem in SETAR estimation poses a determination of delay parameter and order of particular autoregressive process. The main slope parameters estimation follows immediately. One can explicitly define the delay parameter and order of AR processes, but this approach is susceptible to be misspecified. Therefore, we employed a special algorithm representing the complete enumeration of all possible models combining our conditions.

We form a set of possible delay parameter $d \in \{1, 2, \ldots, 5\}$, a set representing the order of autoregressive model $p_1, p_2 \in \{1, 2, \ldots, 6\}$ and the set of threshold parameter $r \in \{p_{0.15}, P_{0.85}\}$, where $p_{0.15}$ and $P_{0.85}$ is 15[th] and 85[th] percentile. For all combination, we estimate the slope parameters. We choose order of autoregressive part corresponding to the minimal AIC criterion of Equation (9) and select the delay and threshold parameter corresponding to the minimal residual variance. Thus, we need to perform 172,970 of estimations for each time series.

The estimation results are summarized in the following tables. Table 1 reports determined delay parameter, estimated threshold parameters and order of autoregressive parts for each time series.

Table 1: Threshold parameters and orders of delay and autoregressive parts

	Nord Europe	California	Austria
delay parameter	1	1	1
threshold parameter	5.160	-1.3876	-27.0497

Next, the construction of confidence intervals follows. For this purpose, we perform a Monte-Carlo experiment. We generate 3,500 values of threshold parameter and for each of them we computed LR statistics according to the Equation (11). Then we plot them against generated threshold parameter and draw the line z_α at the critical value of 7.35, see Hansen (1997, 2000). The Figure 1 depicts results of our experiment.

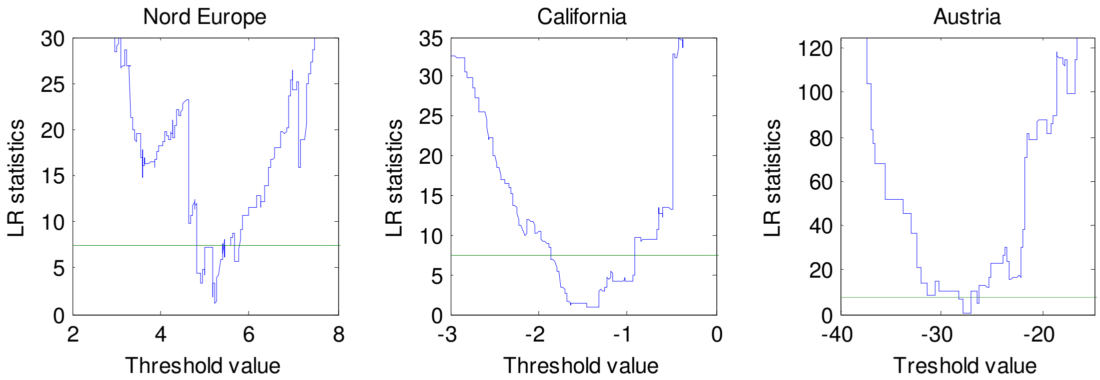

Figure 1: Confidence intervals for threshold value

We can see that the 95 % confidence intervals are not really tight in some cases (especially for California) and are disjoint (Austria). Therefore, we have to convexify the obtained region in accordance with above-mentioned technique in Section 2.2. The Table 2 records the results.

Table 2: Convexifed 95 % confidence intervals of threshold parameters

	Nord Europe	California	Austria
min	4.832	-1.850	-28.291
threshold parameter	5.160	-1.388	-27.050
max	5.788	-0.918	-26.366

For precision assessing of the estimated threshold parameters, we split our sample data into regime 1 for $X_{t-d} < \underline{r}$ (left column in Table 3) and into regime 2 for $X_{t-d} > \overline{r}$ (right column). We also extract the data belonging to the gray zone, thus $X_{t-d} \in \{\underline{r}, \overline{r}\}$, see next table.

Table 3: Regime splitting of data

	Regime 1		**Gray Zone**		**Regime 2**	
Nord Europe	800	79.60%	36	3.58%	169	16.82%
California	403	40.10%	50	4.98%	552	54.93%
Austria	153	15.25%	5	0.50%	845	84.25%

The threshold parameters are estimated precisely for all cases, especially for Austria. For this time series, more than 99.5 percent of all observations fall to one of the two regimes with certainty and only 0.5 percent of all data are in gray zone.

The next Table 4 presents estimated slope parameters of particular regimes and for all time series. Statistically significant coefficients are in bold. In the last row of the table, we show the number of observations used for slope parameters estimation in each regime.

Table 4: Slope parameters of non-linear SETAR model

Parameter	Nord Europe		California		Austria	
	regime		regime		regime	
	1	2	1	2	1	2
p0	**10.2184**	**-3.3102**	**5.0907**	**-6.9629**	**9.5961**	**-44.8087**
p1	-0.5154	0.0050	**0.2546**	**-0.3558**	**-0.2373**	-0.0881
p2	-0.4791	**-0.0079**	**-0.0813**	**-0.3625**	-0.3497	**-0.2859**
p3	**-0.3716**	**-0.0647**	0.0015	**-0.2388**	**-0.1858**	**-0.2247**
p4	-0.4976	-0.0789	-0.0402	**-0.3234**	**-0.1678**	-0.1458
p5	**0.0000**	**-0.1991**	0.0364	0.0205	**-0.0972**	**-0.3974**
p6			-0.0261	-0.0620	0.0171	
No. of obs.	894	201	427	578	155	848

The last step poses a diagnostic checking of residuals. This is presented in the next Section 4.2, where we compare non-linear models with linear autoregressive models. Nevertheless, we can say that estimated models face the same problem as many other models (heteroscedasticity and autocorrelation presence and non-normality of residuals). Our models are not exception and they suffer the same imperfections as linear AR models, see next section.

4.2 Comparison with linear AR model

In this section, we provide a comparison of linear and non-linear AR models. Using the same data sets, we estimated linear AR model by employing LS estimator and determine the order process with similar principle described in Section 4.1. Thus, for particular value from predetermined set of orders, $p \in \{1, 2, \ldots, 6\}$, we run LS estimate and choose the one that is corresponding to the minimal residual variance. From obtained residuals, we also perform diagnostic checks and compare all results with non-linear SETAR models.

Next Table 5 records the estimated slope parameters of AR models. Bold coefficients are statistically significant.

Table 5: Slope parameters of linear AR model

Parameter	Nord Europe	California	Austria
p0	0.0247	0.6259	0.1731
p1	**-0.3255**	**-0.1620**	**-0.5301**
p2	**-0.3667**	**-0.3141**	**-0.6444**
p3	**-0.3012**	**-0.1526**	**-0.5448**
p4	**-0.3276**	**-0.2227**	**-0.5664**
p5	**-0.3465**	**-0.1212**	**-0.5956**
p6	**-0.2331**	**-0.1198**	**-0.4859**
No. of obs.	1089	1089	1089

From the Table above, it is obvious, that all parameters of time series are statistically significant with the exception of constant.

Hereafter, the main comparison of models is following. Next Table 6 records the residual variances of linear and nonlinear models for each return of electricity price at selected market.

Table 6: Residual variance of linear and non-linear models

Model	Nord Europe	California	Austria
Linear AR	100.317	150.304	450.888
Non-linear SETAR	55.222	68.935	191.181

According to the results, we can see that the non-linear models fit the observed data doubly better.

Next comparison consists in potential improvement of the diagnostic checks if they give better results. Whereas we test autocorrelation in linear models by Portemanteau test with order $\lg k = 20$, we are made to employ the modified Breusch-Pagan test (described in Section 3) in order to test the same lag order of autocorrelation in non-linear model. To detect heteroscedasticity, we employ ARCH effect test and for testing of non-normality we used Jaque-Bera test. Results in the form of *p*-values are summarised in next Table.

Table 7: Diagnostic checks – p-values of particular tests

	Autocorr.		Heterosc.		Skewness		Kurtosis		Normality	
	lin	(non)	lin	(non)	lin	(non)	lin	(non)	lin	(non)
Nord Europe	0	0	0	0	0.006	0	0	0	0	0
California	0	0.674	0	0	0	0	0	0	0	0
Austria	0	0	0	0	0.417	0	0	0	0	0

It is apparent from the results, that using non-linear SETAR models do not improve the diagnostic checks results significantly. The autocorrelation of residuals was getting rid off in non-linear model of California time series, but all the others results are almost the same. Heteroscedasticity is present for all time series and it is not dependent on the model employed. In every case, the ARCH effect test indicates the no-constant conditional residual variance. Thereafter, using non-linear model is related to the following consequence: the normal distribution of residuals is more skewed in two cases (Nord Pool and Austria) and kurtosis is non-normal for both types of models. In the end, the Jarque Bera test of normality indicates non-normality of residuals for all estimated econometric models.

5 Conclusion

The aim of the paper was to propose non-linear SETAR models of daily electricity prices at selected regions (Nord Europe, California and Austria) and comparing estimated non-linear models with linear auto-regressive. We compared the non-linear models with linear AR on the basis of residual variance and also results of diagnostic checks.

On the basis of the obtained results, we can conclude that non-linear models fit the data better than linear AR model. The residual variance of non-linear models was half in comparison to residual variance of linear model. However, using non-linear models did not improve the diagnostic checks and in our cases we obtained the same or very similar results.

Generally, we can say that the non-linear SETAR models are more appropriate to model the electricity prices than linear AR model, but none of them capture the time-varying conditional variance and non-normality of probability distributions. This crucial problem is in the fact that the electricity prices are affected by many factors (i.e. seasonality, occurrence of price peaks) and the process of time series is characteristic with very high volatility and with high frequency of spikes resulting in the fact that time series seems to be rather mean-reverting or even non-linearly mean-reverting. Furthermore, the dynamics of time series can be so tangled that the process comprises a lot of process.

References

[1] BOX, G. E. P.; PIERCE, D. A. (1970): *Distribution of the Autocorrelations in Autoregressive Moving Average Time Series Models*, Journal of American Statistical Association 65, p. 1509-1526.

[2] BREUSCH, T.S.; PAGAN, A.R. (1979): *A simple test for heteroscedasticity and random coefficient variation*. Econometrica 47, 1287-94.

[3] BUNN, D.W., KARAKATSANI, N. (2003): *Forecasting Electricity Prices*. London Business School, London.

[4] CATALAO, J.P.S. *et a*l (2007): *Short-term electricity prices forecasting in a competitive market: a neural network approach*. Electric Power System research, 1297-1304.

[5] CLEMENTS, M. P., SMITH, J. (1999): *A Monte Carlo Study of the Forecasting Performance of Empirical SETAR Models*. Journal of Applied Econometrics 14, 123-141

[6] EITRHEIM, Ø.; TERÄSVIRTA, T. (1996): *Testing the adequacy of smooth transition autoregressive models*. Journal of Econometrics 74, 59-76.

[7] ESCRIBANO, A. *et al* (2002): *Modeling electricity prices: international evidence*. Universidad Carlos III. de Madrid, Working Paper 02-27, Economics Series 08, 2002.

[8] FAN, J., YAO, Q. (2005): *Nonlinear Time Series: Nonparameteric and Parametric Methods*. New York : Springer Sciend+Business Media, Inc. 551 s. ISBN 0-387-26142-7.

[9] GARCIA, R.C. *et al* (2003): *A GARCH Forecasting Model to Predict Day-Ahead Electricity prices*. Workshop on Applied Infrastructure, Berlin, Germany.

[10] GUIRGUIS, H.S., FELDER, F.A. (2004): *Further Advances in Forecasting Day-Ahead Electricity prices Using Time Series Models*. KIEE International Transactions on PE, Vol. 4-A, No.3, 159-166.

[11] HANSEN, B. E. (1996): *Inference When a Nuisance Parameter is not Identified Under the Null Hypothesis*, Econometrica 64, p. 413-430.

[12] HANSEN, B.E. (2000*): Sample splitting and threshold estimation*. Econometrica 68, 575-603.

[13] JARQUE, CARLOS M.; ANIL K. BERA (1980): *Efficient tests for normality, homoscedasticity and serial independence of regression residuals*, Economics Letters 6, p. 255–259.

[14] JARQUE, CARLOS M.; ANIL K. BERA (1981): *Efficient tests for normality, homoscedasticity and serial independence of regression residuals: Monte Carlo evidence*, Economics Letters 7, p. 313–318.

[15] KANAMURA, T. AND OHASHI, K. (2007): *A Structural Model for Electricity Prices with Spikes - Measurement of Jump Risk and Optimal Policies for Hydropower Plant Operation*, Energy Economics 29, p. 1010-1032.

[16] KIAN, A., KEYHANI, A. (2001): *Stochastic Price Modeling of Electricity in Deregulated Energy Markets.* Proceedings of the 34th Hawaii International Conference on System Sciences.

[17] LJUNG, G. M.; BOX, G. E. P. (1978): *On a measure of lack of fit in time series models*, Biometrika 65, p. 297-303.

[18] MING CHIEN LO (2008): *Nonlinear PPP Deviations: A Monte Carlo Investigation of Their Unconditional Half-Life*. Studies in Nonlinear Dynamics & Econometrics 12.

[19] PHILIP HANS FRANSES, DICK VAN DIJK (2000): *Nonlinear Time Series Models in Empirical Finance.* Cambridge: Cambridge University Press. ISBN 0-521-77965-0.

[20] SHAPIRO, S. S.; WILK, M. B. (1965): *An analysis of variance test for normality (complete samples*), Biometrika 52, p. 591–611.

[21] WERON, R., MISIOREK A. (2005): *Forecasting spot electricity prices with time series models.* The European Electricity Market, 2005, Lodz, Poland.

Summary

Cílem příspěvku je provést odhad a test nelinearity cen elektrické energie ve třech vybraných regionech (Kalifornie, Severní Evropa, Rakousko). Pro analýzu nelinearity je aplikován SETAR (Self Exciting Threshold Auto-Regressive) model, který zohledňuje změnu vybraných parametrů v různých intervalech časových řad.

Je odhadnut nejvhodnější SETAR model pro modelování cen elektrické energie na vybraných trzích, model je statisticky testován a je ověřeno, zda vykazuje lepší výsledky než lineární AR model.