Non-linear processes for electricity prices: robust estimation and forecasting

Luigi Grossi and Fany Nan

University of Verona

Emails: luigi.grossi@univr.it and fany.nan@univr.it

Abstract

In this paper we suggest the use of robust GM-SETAR (Self Exciting Threshold AutoRegressive) processes to model and forecast electricity prices observed on deregulated markets. The robustness of the model is achieved by extending to time series the generalized M-type (GM) estimator first introduced for independent multivariate data. As it has been shown in a very recent paper [5], the polynomial weighting function over-performs the classical ordinary least squares method when extreme observations are present. The main advantage of estimating robust SETAR models is the possibility to capture two very well-known stylized facts of electricity prices: nonlinearity produced by changes of regimes and the presence of sudden spikes due to inelasticity of demand. The forecasting performance of the model applied to the Italian electricity market (IPEX) is improved by the introduction of predicted demand as an exogenous regressor. The availability of this regressor is a particular feature of the Italian market. By means of prediction performance indexes and tests, it will be shown that this regressor plays a crucial role and that robust methods improve the overall forecasting performance of the model.

Keywords: Electricity prices, extreme values, GM estimator, nonlinear models

1 Introduction

A very well known stylized fact of electricity prices is the presence of isolated jumps as a consequence of sudden grid congestions which reflect immediately on prices because of lack of flexibility of the supply and demand curves [6]. This feature must be considered very carefully and robust techniques must be applied to avoid that few jumps could dramatically affect parameter estimates. Although many papers have applied quite sophisticated time series models to prices and demand time series of electricity and gas only few have considered the strong influence of jumps on estimates and the need to move to robust estimators [7], [10]. Among robust techniques for electricity prices, robust SETAR models have never been estimated. The reasons could be summarized by two main points:

1) properties of robust SETAR estimators have not been completely studied and there isn't a clear accordance on the best estimator, at least with reference to the best weighting function [4];

2) robust estimators are not implemented within the most popular statistical software platforms such as Matlab and R.

[5] have addressed the two points through a massive Monte Carlo experiment which compares the performances of classical SETAR estimator and robust estimator using different weighting functions. All the estimators (classical and robust) have been implemented in R language resulting in a set of functions which hopefully will become a library soon. The main result obtained by the authors is a quite clear prevalence of the generalized M-estimator (GM-estimator) based on the polynomial weighting function [9] with respect to the ordinary least squares (OLS) estimator when dealing with some well known features of electricity prices time series: large sample size and presence of several isolated and big spikes.

In this paper, using the results contained in [5], classical and robust GM polynomial-based estimators are applied to obtain parameters of SETAR models on Italian electricity price data (*PUN, prezzo unico nazionale*). The model is enriched by the introduction of exogenous regressors which should improve the forecasting performances. Crucial variables in predicting electricity prices are dummies for the intra-day seasonality and demanded volumes [3]. Comparisons will be made among different estimators and between pure SETAR models and nonlinear specifications with exogenous regressors.

The paper is organized as follows. In the next section, the general SETAR model is introduced and the main weighting functions are discussed to move to robust estimators. Section three contains summary and comments of the forecasting results. Conclusions and final remarks are reported in section four.

2 Robust SETAR models with exogenous regressors

Given a time series y_t, a general two-regime Self-Exciting Threshold AutoRegressive model SETAR(p,d) with exogenous regressors is specified as

$$y_t = (\mathbf{x}_t\beta_1 + \mathbf{z}_t\lambda_1)I(y_{t-d} \leq \gamma) + (\mathbf{x}_t\beta_2 + \mathbf{z}_t\lambda_2)I(y_{t-d} > \gamma) + \varepsilon_t \tag{1}$$

for $t = 1, ..., N$, where $I(\cdot)$ is an indicator function, y_{t-d} is the threshold variable with $d \geq 1$ and γ is the threshold value. The relation between y_{t-d} and γ states if y_t is observed in regime 1 or 2. $\beta_\mathbf{j}$ is the vector of auto-regressive parameters for regime $j = 1, 2$ and $\mathbf{x_t}$ is the t-th row of the $(N \times p)$ matrix $\mathbf{X}$ comprising p lagged variables of y_t (and possibly a constant). $\lambda_\mathbf{j}$ is the vector corresponding to exogenous regressors contained in the $(N \times r)$ matrix $\mathbf{Z}$ whose t-th row is $\mathbf{z_t}$. Errors ε_t are assumed to follow an iid$(0, \sigma_\varepsilon)$ distribution.

In general the value of the threshold γ is unknown, so that the parameters to estimate become $\theta = (\beta_1', \beta_2', \lambda_1', \lambda_2')$, γ and σ_ε. Parameters can be estimated by sequential conditional least squares: for a fixed threshold γ the model is linear, θ can be estimated by OLS and $\hat{\sigma}_\varepsilon = \sum_{t=1}^{N} r_t^2/N$, with $r_t = y_t - \mathbf{x}^*{}_t\hat{\theta}$, where $\mathbf{x}^*{}_t = (\mathbf{x}_t, \mathbf{z}_t)$. The least squares estimate of γ is obtained by minimizing the residual sum of squares $\gamma = \arg\min_{\gamma\in\Gamma} \sum_{t=1}^{N} r_t^2$ over a set Γ of allowable threshold values so that each regime contains at least a given fraction (ranging from 0.05 to 0.3) of all observations.

In the case of robust two-regime SETAR model, for a fixed threshold γ the GM estimate of the autoregressive parameters can be obtained by applying the iterative weighted least squares:

$$\hat{\theta}_j^{(n+1)} = \left(\mathbf{X}^{*\prime}_j \mathbf{W}^{(n)} \mathbf{X}^*{}_j\right)^{-1} \mathbf{X}^{*\prime}_j \mathbf{W}^{(n)} \mathbf{y}_j \tag{2}$$

where $\hat{\theta}_j^{(n+1)}$ is the GM estimate for the parameter vector in regime $j = 1, 2$ after the n-th iteration from an initial estimate $\hat{\theta}_j^{(0)}$, and $\mathbf{W}^{(n)}$ is a weight diagonal matrix, whose elements depend on a weighting function $w(\hat{\theta}_j^{(n)}, \hat{\sigma}_{\varepsilon,j}^{(n)})$ ranging between 0 and 1. The threshold γ can be estimate by minimizing an objective function (see 2.1) over the set Γ of allowable threshold values.

2.1 Weighting methods

Three different weighting functions have been discussed in the literature on robust estimators for nonlinear time series models. The first function is described in [1]. Weights are calculated as

$$w(\hat{\theta}_j, \hat{\sigma}_{\varepsilon,j}) = \psi\left(\frac{y_t - m_{y,j}}{C_y \hat{\sigma}_{y,j}}\right) \psi\left(\frac{y_t - \mathbf{x}^*{}_t\hat{\theta}_j}{C_\varepsilon \hat{\sigma}_{\varepsilon,j}}\right)$$

where ψ is the Tukey bisquare weighting function ([9], p.216) and $m_{y,j}$ is a robust estimate of the location parameter (sample median) in the j-th regime. $\hat{\sigma}_{y,j}$ and $\hat{\sigma}_{\varepsilon,j}$ are robust estimates of the scale parameters σ_y and σ_ε respectively, obtained by the median absolute deviation multiplied by 1.483. C_y and C_ε are tuning constants fixed at 6.0 and 3.9 respectively.

The objective function to minimize for the search of the threshold depends on Tukey bisquare weights. We use the same function described in [1].

For the second method, we follow [2]. The GM weights are presented in Schweppe's form $w(\hat{\theta}_j, \hat{\sigma}_{\varepsilon,j}) = \psi(r_t)/r_t$ with standardized residuals $r_t = (y_t - \mathbf{x}^*{}_t\hat{\theta}_j)/(\hat{\sigma}_{\varepsilon,j} w(\mathbf{x}^*{}_t))$ and $w(\mathbf{x}^*{}_t) = \psi(\mathrm{d}(\mathbf{x}^*{}_t)^\alpha)/\mathrm{d}(\mathbf{x}^*{}_t)^\alpha$. $\mathrm{d}(\mathbf{x}^*{}_t) = |\mathbf{x}^*{}_t - m_{y,j}|/\hat{\sigma}_{y,j}$ is the Mahalanobis distance and α is a constant usually set equal to 2 to obtain robustness of standard errors. The chosen weighting function is the polynomial ψ function. The threshold γ is estimated by minimizing the objective function $\sum_{t=1}^{N} w(\hat{\theta}, \hat{\sigma}_\varepsilon)(y_t - \mathbf{x}^*{}_t\hat{\theta})^2$ over the set Γ of allowable threshold values.

The third method is based on the same methodologies of the second but with ψ being the Huber weighting function which is a convex non-negative function. The consistency of GM estimators of autoregressive parameters in each regime of SETAR models when the threshold is unknown and using the Huber weights has been proven by [13].

The estimation performances of the three methods have been studied through an extensive Monte Carlo simulation experiment in [5]. From the experiment a prevalence of the polynomial (POL from now on) has been pointed out, which will be then used throughout the present paper.

3 Application: Italian electricity price

In this section, we apply LS and the robust POL weighting functions, presented in the previous section, to estimate parameters of SETAR models on the Italian electricity price data (*PUN, prezzo unico nazionale*). Moreover, a comparison of prediction accuracy among the methods is conducted.

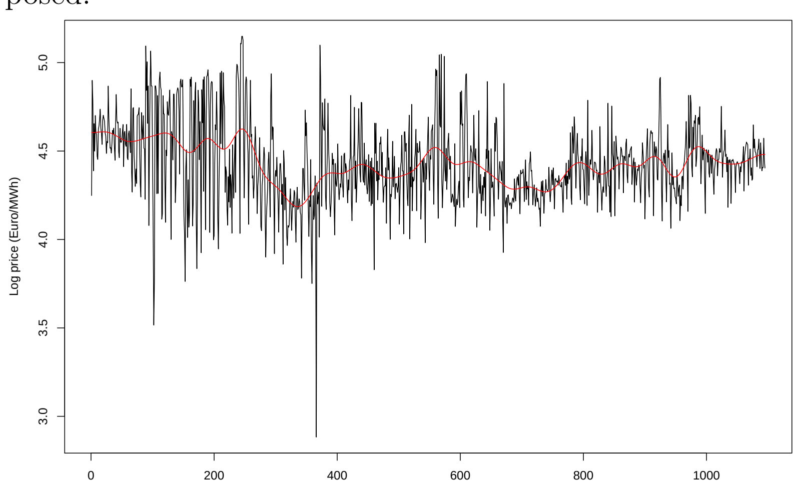
The time series of prices used in the present work covers the period from January 1st, 2010 to December 31th, 2012 (26,304 data points, for $N = 1{,}096$ days): year 2012 has been left for out-of-sample forecasting. The data have an hourly frequency, therefore each day consists of 24 load periods with 00:00–01:00am defined as period 1. Spot price is denoted as P_{tj}, where t specifies the day and j the load period ($t = 1, 2, ..., N$; $j = 1, 2, ..., 24$).

In this study, following a widespread practice in literature, each hourly time series is modeled separately.

Differences in load periods can cause significant variations in price time series. A first inspection, based on graphs, spectra and ACFs (Figures are not reported for lack of space but can be obtained from the authors) for different hours, shows that the series have long-run behaviour and annual dynamics, which change according to the load period. A common characteristic of price time series is the weekly periodic component (of period 7), suggested by the spectra that show three peaks at the frequencies 1/7, 2/7 and 3/7, and a very persistent autocorrelation function.

We assume that the dynamics of log prices can be represented by a nonstationary level component L_{tj}, accounting for level changes and/or long-term behaviour, and a residual stationary component p_{tj}, formally, $\log P_{tj} = L_{tj} + p_{tj}$.

Figure 1: $\log P_{tj}$ for hour 11, with the estimated nonstationary level component superimposed.



To estimate L_{tj} we used the wavelets approach ([11]). Wavelets have been used in many studies, including [8] and [12]. We considered the Daubechies least asymmetric wavelet family, LA(8), and the coefficients were estimated *via* the maximal overlap discrete wavelet transform (MODWT) method (for details, see [11]). As an example of the time series of prices and corresponding estimated long-term component, Figure 1 shows $\log P_{tj}$ for hour 11, with the estimated nonstationary level component superimposed. After removing the long-term component, we estimated on the stationary time series p_{tj} the SETAR(p,d) model with exogenous regressors, as reported in eq. (1). According to the empirical ACFs, a SETAR(7,1) model has been estimated over all the price series to highlight differences in the estimation given by different dynamics characterizing each load period. Matrix $\mathbf{Z}$ of exogenous regressors could contain day-of-the-week dummies, D_k, with $k = 1, ..., 6$ and the day-ahead predicted demand of electricity made available by GME (Gestore Mercato Elettrico). Next step of the analysis will be to compare the forecasting performances of the robust methods with the forecasting performance of the LS estimator. The base model contains only autoregressive components, excluding in eq. (1) matrix $\mathbf{Z}$. The second model we consider is a SETAR(7,1) with day-of-the-week dummies, D_k, with $k = 1, ..., 6$. In the third model, matrix $\mathbf{Z}$ contains the detrended day-ahead predicted demand of electricity made available by GME (Gestore Mercato Elettrico). As for the price series, the level component of the predicted demand has been estimated using the wavelets approach.

For comparing our robust/non robust SETAR models, we reproduced 366 one day-ahead forecasts $\hat{p}_{t+1}$ for each model estimated on a rolling window of 2 years. Comparisons are based on the predictions of the original spot prices that are given by $\hat{P}_{t+1} = \exp(\hat{L}_{t+1} + \hat{p}_{t+1})$, where $\hat{L}_{t+1}$ and $\hat{p}_{t+1}$ are predictions of the components, which are based only on the information available in t. In particular, we set $\hat{L}_{t+1} = \hat{L}_t$, that is, we used the estimated value in t as a forecast for $t+1$. Besides its simplicity, the motivation to use this equation comes from the fact that the long-term component, by definition, should be basically the same for two contiguous days. Forecasts have been compared in terms of MSE (Mean Square Error) and MAE (Mean Absolute Error), and of the Diebold and Mariano test, that are based on the forecasting errors $e_{tj} = P_{tj} - \hat{P}_{tj}$, $(t = 1, 2, ..., M;\ j = 1, 2, ..., 24)$ for each method. We used the one-tailed Diebold and Mariano test (DM), whose null hypothesis is that the prediction accuracy of procedure (say) A is equal or lower than that of procedure B. The test has been performed both with MSE and MAE using data observed in three hours which can be considered as representative of the main patterns during the day: 11, 18 and 21. Table 1 shows MSE and MAE values on the whole year 2012 and on the four quarters. A SETAR model with dummies for days of the week has been estimated with LS and POL methods. In the last two columns we reported an indicator variable which assumes two values: 1 when the robust (POL) forecasts are statistically better than the non robust (LS) forecasts using the Diebold and Mariano test at 5% significance level, 2 when the opposite event happens (LS forecasts are better than POL).

For example, for hour 11 LS are statistically better in the fourth quarter, both considering MSE and MAE, while, for hour 21, POL is statistically better than LS in the second quarter, considering MSE, and in the second, third and whole period, considering the MAE loss function.

In order to summarize the results contained in Table 1, we have reported the number of hours (out of the three we have analyzed) the LS forecasts are better than the POL (Table 2) and the number of hours the POL forecasts are better than LS (Table 3). To help interpreting the tables, the first column of each table shows the number of hours the MSE value of an estimator is the lowest, the second column gives the same information but with reference to MAE, the third and fourth column show the number of hours in which the prevalence of an estimator is statistically significant looking at the MSE and the MAE, respectively.

Table 1: *MSE and MAE of forecasts obtained with LS and POL models with dummies on the four estimation periods and on the whole year. Last two columns: 1 indicates that forecasts obtained with POL are statistically better than predictions of LS, 2 indicates that forecasts obtained with LS are statistically better than predictions of POL, 0 indicates that forecasts are not significantly different (1-tailed Diebold and Mariano test at 5% significance level, MSE and MAE loss functions).*

Period	**MSE**		**MAE**		**D-M (MSE)**	**D-M (MAE)**
	LS	POL	LS	POL		
				Hour 11		
Jan-Mar	406.51	451.88	12.21	12.1	0	0
Apr-Jun	213.99	216.83	11.3	11.49	0	0
Jul-Sep	75	71.47	6.86	6.6	0	0
Oct-Dec	65.97	73.49	5.85	6.28	2	2
Year	189.71	202.7	9.04	9.1	0	0
				Hour 18		
Jan-Mar	422.54	357.16	16.25	14.49	1	1
Apr-Jun	159.95	160.66	9.51	9.33	0	0
Jul-Sep	97.72	108.53	6.52	6.96	2	0
Oct-Dec	245.43	250.92	11.13	11.24	0	0
Year	231.08	219.1	10.84	10.5	0	0
				Hour 21		
Jan-Mar	209.89	217.7	9.59	9.6	0	0
Apr-Jun	171.39	160.18	10.43	9.97	1	1
Jul-Sep	723.46	711.44	14.57	13.57	0	1
Oct-Dec	61.1	59.6	5.65	5.55	0	0
Year	292.01	287.77	10.06	9.67	0	1

Table 2: *Number of cases LS model with dummies gives better results than POL model with dummies, considering the three analyzed hours.*

Period	**MSE**	**MAE**	**D-M (MSE)**	**D-M (MAE)**
Jan-Mar	2	1	0	0
Apr-Jun	2	1	0	0
Jul-Sep	1	1	1	0
Oct-Dec	2	2	1	1
Year	1	1	0	0

Table 3: *Number of cases POL model with dummies gives better results than LS model with dummies, considering the three analyzed hours.*

Period	**MSE**	**MAE**	**D-M (MSE)**	**D-M (MAE)**
Jan-Mar	1	2	1	1
Apr-Jun	1	2	1	1
Jul-Sep	2	2	0	1
Oct-Dec	1	1	0	0
Year	2	2	0	1

Table 4: *Ratios of MSE and MAE of forecasts obtained with models with dummies to MSE and MAE of forecasts obtained with models with a pure SETAR (four estimation periods and whole year). Last two columns: 1 indicates that forecasts obtained with dummies are statistically better than predictions with pure SETAR model, 2 indicates that forecasts obtained with AR are statistically better than predictions with dummies, 0 indicates that forecasts are not significantly different (1-tailed Diebold and Mariano test at* 5% *significance level, MSE and MAE loss functions).*

Period	**MSE**		**MAE**		**D-M (MSE)**		**D-M (MAE)**	
	LS	POL	LS	POL	LS	POL	LS	POL
				Hour 11				
Jan-Mar	0.934	1.104	0.886	0.915	0	0	1	0
Apr-Jun	0.988	0.958	1.007	0.985	0	0	0	0
Jul-Sep	0.877	0.771	1.017	0.95	0	1	0	0
Oct-Dec	0.8	0.868	0.871	0.951	1	0	1	0
Year	0.929	1	0.942	0.949	0	0	1	0
				Hour 18				
Jan-Mar	0.94	0.904	0.97	0.932	0	0	0	0
Apr-Jun	0.896	0.928	0.984	0.953	1	0	0	0
Jul-Sep	0.715	0.807	0.847	0.893	1	1	1	1
Oct-Dec	0.919	0.903	0.924	0.932	0	0	1	1
Year	0.897	0.894	0.94	0.93	1	1	1	1
				Hour 21				
Jan-Mar	1.046	1.015	0.953	0.967	0	0	0	0
Apr-Jun	1.077	1.081	1.085	1.051	0	0	2	0
Jul-Sep	1.051	1.021	1	1.006	0	0	0	0
Oct-Dec	0.908	0.914	0.937	0.973	1	0	0	0
Year	1.045	1.021	0.999	1.002	0	0	0	0

Table 5: *Number of cases models with a pure SETAR give better results than models with dummies, considering the three analyzed hours.*

Period	**MSE**		**MAE**		**D-M (MSE)**		**D-M (MAE)**	
	LS	POL	LS	POL	LS	POL	LS	POL
Jan-Mar	1	2	0	0	0	0	0	0
Apr-Jun	1	1	2	1	0	0	1	0
Jul-Sep	1	1	1	1	0	0	0	0
Oct-Dec	0	0	0	0	0	0	0	0
Year	1	1	0	1	0	0	0	0

Table 6: *Number of cases models with dummies give better results than models with a pure SETAR, considering the three analyzed hours.*

Period	**MSE**		**MAE**		**D-M (MSE)**		**D-M (MAE)**	
	LS	POL	LS	POL	LS	POL	LS	POL
Jan-Mar	2	1	3	3	0	0	1	0
Apr-Jun	2	2	1	2	1	0	0	0
Jul-Sep	2	2	1	2	1	2	1	1
Oct-Dec	3	3	3	3	2	0	2	1
Year	2	1	3	2	1	1	2	1

A further point we have explored is the effectiveness of dummies in the improvement of the model forecasting performance with respect to a pure SETAR model. Results are reported in Table 4. Values shown in the Table 4 are ratios of MSE (MAE) of the SETAR with dummies and pure SETAR estimated both with LS and POL. Values less than 1 mean that the model with dummies performs better than the pure autoregressive model. For example, the first value of the table at the top left corner is 0.934 and means that SETAR model with dummies estimated by LS is 6.66% better than the pure SETAR estimated with LS. The last four columns of the table contain, as before, an indicator variable which is 1 when the model with dummies is statistically better, using the Diebold-Mariano test, than the pure SETAR model, while is 2 in the opposite case. As can be seen, the indicator variable is 1 in many cases while is 2 only in one case, showing that the dummies improve the forecasting performance of the model. Summaries contained, as done earlier, in Table 5 and 6 confirm the prevalence of the model with dummies.

Finally, the forecasted demand has been introduced in the model to see whether the robust POL estimator is still better than the LS in presence of spikes and if this exogenous variable can help improving the forecasting performance of the model. The comparison between robust and non-robust estimator is reported in Table 7 and results are summarized, as usual, in 8 and 9. Last two columns of Table 7 contains some 1 and are never equal 2, showing the prevalence of the robust estimator even in complex models when dummies and exogenous variables are introduced.

4 Conclusions

In this paper the forecasting performances of the robust GM estimator based on the polynomial weighting function and of the classical Least Squares estimator applied to electricity prices have been studied. Dummy variables have been included in the model

Table 7: *MSE and MAE of forecasts obtained with LS and POL models on the four estimation periods and on the whole year. Last two columns: 1 indicates that forecasts obtained with POL are statistically better than predictions of LS, 2 indicates that forecasts obtained with LS are statistically better than predictions of POL, 0 indicates that forecasts are not significantly different (1-tailed Diebold and Mariano test at 5% significance level, MSE and MAE loss functions).*

Period	**MSE**		**MAE**		**D-M (MSE)**	**D-M (MAE)**
	LS	POL	LS	POL		
				Hour 11		
Jan-Mar	453.85	456.64	12.37	12.28	0	0
Apr-Jun	158.61	171.46	9.72	10.09	0	0
Jul-Sep	61.47	65.58	6.1	6.05	0	0
Oct-Dec	57.97	61.64	5.28	5.43	0	0
Year	196.28	202.37	8.7	8.79	0	0
				Hour 18		
Jan-Mar	397.22	374.12	15.41	14.6	0	1
Apr-Jun	122.95	122.13	8.05	8.12	0	0
Jul-Sep	110.58	115.72	7.4	7.28	0	0
Oct-Dec	166.25	157.19	8.67	8.57	0	0
Year	202.59	195.9	10.01	9.75	0	0
				Hour 21		
Jan-Mar	210.59	209.32	9.82	9.61	0	0
Apr-Jun	159.42	155.74	9.89	9.55	0	0
Jul-Sep	716.16	691.46	13.8	12.89	1	1
Oct-Dec	63.36	59.18	5.83	5.58	0	1
Year	313.2	304.22	10.29	9.84	1	1

Table 8: *Exo1: number of cases LS model gives better results than POL model.*

Period	**MSE**	**MAE**	**D-M (MSE)**	**D-M (MAE)**
Jan-Mar	1	0	0	0
Apr-Jun	1	2	0	0
Jul-Sep	2	0	0	0
Oct-Dec	1	1	0	0
Year	1	1	0	0

Table 9: *Exo1: number of cases POL model gives better results than LS model.*

Period	MSE	MAE	D-M (MSE)	D-M (MAE)
Jan-Mar	2	3	0	1
Apr-Jun	2	1	0	0
Jul-Sep	1	3	1	1
Oct-Dec	2	2	0	1
Year	2	2	1	1

in order to account for the presence of seasonality in the data and day-ahead predicted demand has been considered to test its relevance in predicting day-ahead prices. Summarizing the main results we could say that robust estimators over-perform the classical Least Squares estimator both when simple SETAR models are considered and when exogenous variables are included in the model. The use of regressors seems to increase the forecasting power of models when dummies and forecasted demand are included as external regressors compared to pure SETAR models. Other comparisons have been carried out but results have not been reported for lack of space. Summarizing briefly we could say that the model with forecasted demand as exogenous variable has better performances not only than the pure SETAR model, but also than the SETAR model with dummies. Moreover we have also tried to model directly the original series of prices including the trend estimated by wavelets as regressor, but the corresponding forecasting performance is worse than the model estimated directly on the de-trended prices.

Acknowledgment

The authors would like to thank the Department of Economics and the University of Verona for financing the research with a FAR/2013 (ex 60%) grant. They also thank the financial support of the project MIUR PRIN MISURA - Multivariate models for risk

assessment.

References

[1] Chan, W. S. and Cheung, S. H. (1994). On robust estimation of threshold autoregressions, *Journal of Forecasting* 13: 37–49.

[2] Franses, P. H. and van Dijk, D. (2000). *Non-linear time series models in empirical finance*, Cambridge University Press.

[3] Gianfreda, A. and Grossi, L. (2012). Forecasting italian electricity zonal prices with exogenous variables, *Energy Economics* 34: 2228–2239.

[4] Giordani, P., Kohn, R. and van Dijk, D. (2007). A unified approach to nonlinearity, structural change, and outliers, *Journal of Econometrics* 137(1): 112–133.

[5] Grossi, L. and Nan, F. (2013). Robust estimation of regime switching models, *CLADAG13, Modena, Italy, September 2013*, pp. 77–80.

[6] Huisman, R. and Mahieu, R. (2003). Regime jumps in electricity prices, *Energy Economics* 25: 425–34.

[7] Janczura, J., Trueck, S., Weron, R. and Wolff, R. C. (2013). Identifying spikes and seasonal components in electricity spot price data: A guide to robust modeling, *Energy Economics* 38: 96–110.

[8] Janczura, J. and Weron, R. (2010). An empirical comparison of alternate regimeswitching models for electricity spot prices, *Energy Economics* 32: 1059–1073.

[9] Maronna R.A., R.D. Martin and V.J. Yohai (2006), *Robust Statistics: Theory and Methods*, John Wiley and Sons, Chichester.

[10] Nowotarski, J., Tomczyk, J. and Weron, R. (2013). Robust estimation and forecasting of the long-term seasonal component of electricity spot prices, *Energy Economics* 39(0): 13–27.

[11] Percival, D. and Walden, A. (2000). *Wavelet Methods for Time Series Analysis*, Cambridge University Press.

[12] Treuck, S., Weron, R. and Wolff, R. (2007). Outlier treatment and robust approaches for modeling electricity spot prices, *MPRA Paper No. 4711*, Hugo Steinhaus Center, Wroclaw University of Technology.

[13] Zhang, L. X., Chan, W. S., Cheung, S. H. and Hung, K. C. (2009). A note on the consistency of a robust estimator for threshold autoregressive processes, *Statistics and Probability Letters* 79: 807-813.