Forecasting wholesale electricity prices: A review of time series models

Weron, Rafal

2009

Appeared in: "Financial Markets: Principles of Modelling, Forecasting and Decision-Making", eds. W. Milo, P. Wdowiński, FindEcon Monograph Series, WUŁ, Łódź (2009) 71-82.

Rafał Weron*

Forecasting wholesale electricity prices: a review of time series models

Abstract. In this paper we assess the short-term forecasting power of different time series models in the electricity spot market. We calibrate autoregression (AR) models, including specifications with a fundamental (exogenous) variable – system load, to California Power Exchange (CalPX) system spot prices. Then we evaluate their point and interval forecasting performance in relatively calm and extremely volatile periods preceding the market crash in winter 2000/2001. In particular, we test which innovations distributions/processes – Gaussian, GARCH, heavy-tailed (NIG, α-stable) or non-parametric – lead to best predictions.

Keywords: Electricity price forecasting; heavy tailed distribution; autoregression model; GARCH model; non-parametric noise; system load
JEL Classification: C22, C46, C53, Q47

1. Introduction

In the last decades, with deregulation of power markets and introduction of competition, electricity price forecasts have become a fundamental input to an energy company's decision-making mechanism (Bunn 2004; Weron 2006). Extreme price volatility, which can be even two orders of magnitude higher than for other commodities or financial instruments, has forced producers and wholesale consumers to hedge not only against volume risk but also against price movements. Short-term price forecasts (STPF) are of particular interest for participants of auction-type spot electricity markets who are requested to express their bids in terms of prices and quantities. In such markets buy (sell) orders are accepted in order of increasing (decreasing) prices until total demand (supply) is met. Consequently, a market participant that is able to forecast spot prices can adjust its own production and (to some extent) consumption schedule accordingly and hence maximize its profits.

This paper is a continuation and a review of our earlier studies on STPF of California electricity prices with time series models (Misiorek et al. 2006; Weron 2006; Weron and Misiorek 2007). Consequently, as a benchmark and a starting point we choose the autoregressive time series specification that has been found to perform well for pre-crash California power market data. We compare it not only with autoregressive models allowing for heteroskedastic (GARCH) or heavy-tailed (NIG, α-stable) innovations, but also with an autoregressive model calibrated within a non-parametric framework, where the innovations density is estimated by the Parzen–Rosenblatt kernel. We call the latter model semi-parametric because it has a parametric autoregressive part and a non-parametric noise distribution. In a detailed empirical study we evaluate the point and interval forecasting

* Hugo Steinhaus Center, Institute of Mathematics and Computer Science, Wrocław University of Technology, 50-370, Wrocław, Poland.

performance of all investigated models. In particular, the novel in electricity price forecasting literature semi-parametric approach seems to be promising.

2. The data and the base model

Like in our previous studies, an assumption is made that only publicly available information is used to predict spot prices, i.e. generation constraints or line and production capacity limits are not considered. The dataset used in this analysis – *CA_hourly.dat* – is distributed with the MFE Toolbox (Weron 2006). It was constructed using data obtained from the UCEI institute (www.ucei.berkeley.edu) and the California independent system operator (CAISO; oasis.caiso.com). Apart from hourly system prices quoted in the California Power Exchange (CalPX), it includes two fundamental variables: system-wide loads and day-ahead CAISO load forecasts for the California market, see Figure 1.

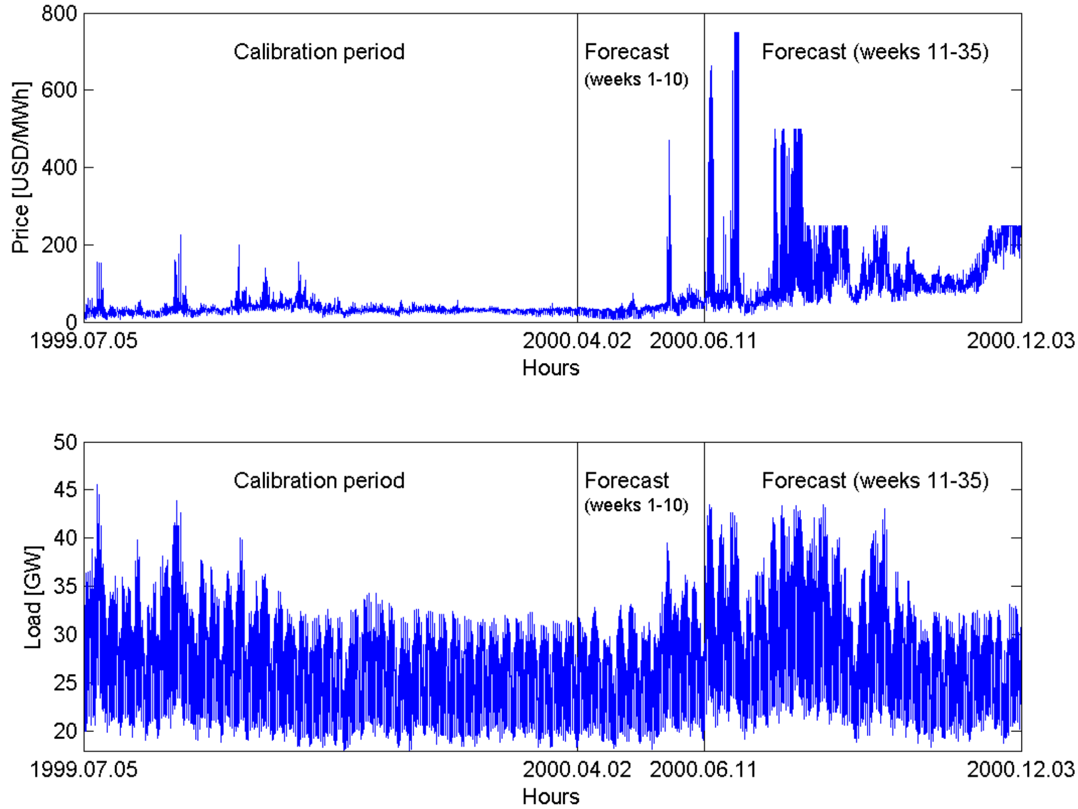

Figure 1: Hourly system prices (*top panel*) and hourly system loads (*bottom panel*) in California for the period July 5, 1999 – December 3, 2000. The changing price cap (750 → 500 → 250 USD/MWh) is clearly visible in the top panel. The day-ahead load forecasts of the system operator CAISO are indistinguishable from the actual loads at this resolution; only the latter have been plotted.

The logarithmic transformation was applied to price, $p_t = \log(P_t)$, and load, $z_t = \log(Z_t)$, data to attain a more stable variance. Furthermore, the mean price and the median load were removed to center the data around zero. Removing the mean load resulted in worse forecasts, perhaps, due to the very distinct and regular asymmetric weekly structure with the five weekday values lying in the high-load region and the two weekend values in the low-load region. The data from the period July 5, 1999 – April 2, 2000 was used only for calibration. Such a relatively long period of data was needed to achieve high accuracy. For example, limiting the calibration period to data coming only from the year 2000, like in Contreras et al. (2003), led to a decrease in forecasting performance by up to 70%. Consequently, the period April 3 – December 3, 2000 was used for out-of-sample testing. Since in practice the market-clearing price forecasts for a given day are required on the day before, we used the following (adaptive) testing scheme. To calibrate the models and compute price forecasts for hour 1 to 24 of a given day, data available to all procedures included price and system load historical data up to hour 24 of the previous day plus day-ahead load predictions for the 24 hours of that day.

The time of the year, the day of the week and the hour of the day influence price patterns. Price forecasting models should take these time factors into account. However, since we are focused on short-term (day-ahead) forecasting the annual seasonal behavior does not play a major role (also the used adaptive testing scheme allows the analyzed autoregressive models to quickly adapt to the changing conditions throughout the year).

The hourly and weekly seasonal patterns were handled in two different ways. Since each hour displays a rather distinct price profile reflecting the daily variation of demand, costs and operational constraints the modeling was implemented separately across the hours, leading to 24 sets of parameters. This approach was also inspired by the extensive research on demand forecasting, which has generally favored the multi-model specification for short-term predictions (Bunn 2000; Weron 2006). On the other hand, the weekly seasonal behavior – which is mostly due to variable intensity of business activities throughout the week – was captured by a combination of (i) the autoregressive structure of the models and (ii) daily dummy variables. The log-price was made dependent on the log-prices for the same hour on the previous days, and the previous weeks, as well as the minimum of all prices on the previous day (the latter created the desired link between bidding and price signals from the entire day). Furthermore, three dummy variables (for Monday, Saturday and Sunday) were considered to differentiate between the weekends, the first working day of the week and the remaining business days (this particular choice of the dummies is a consequence of the significance of the dummy coefficients for particular days).

The electricity spot price is not only dependent on the weekly and daily business cycles but also on other fundamental variables that can significantly alter this deterministic seasonal behavior. Recall, that the equilibrium between demand and supply defines the spot price. Both – demand and supply – are influenced by weather conditions, most notably air temperatures. In the short-term horizon, the variable cost of power generation is essentially just the cost of the fuel, consequently, the fuel price is another influential exogenous factor. Other factors like power plant availability (capacity) or grid traffic (for zonal and modal pricing) could also be considered. However, including all these factors would make the model not only cumbersome but also sensitive to the quality of the inputs and conditional on their availability at a given time. Instead we have decided to use only publicly available, high-frequency (hourly) information. In the California market of the late 1990s this includes system-wide loads and day-ahead CAISO load forecasts. In particular, the latter are important as they include the system operator's (and to some extent – the market's) expectations regarding weather, demand, generation and power grid conditions prevailing at the hour of delivery. The knowledge of these forecasts allows, in general, for more accurate spot price predictions. In the studied period (with some deviations in the volatile weeks 11-35), the logarithms of loads (or load forecasts) and the log-prices were approximately linearly dependent (the Pearson correlation was positive, $\rho > 0.6$, and highly significant with a p-value of approximately 0; null of no correlation).

At lag 0 the CAISO day-ahead load forecast for a given hour was used, while for larger lags the actual system load was used. Interestingly, the best models turned out to be the ones with only lag 0 dependence. Using the actual load at lag 0, in general, did not improve the forecasts either. This phenomenon can be attributed to the fact that the prices are an outcome of the bids, which in turn are placed with the knowledge of the CAISO load forecasts but not actual future loads.

Extensive studies performed by Weron (2006) led to the conclusion that the best autoregressive model structure for the (log-)price p_t, in terms of forecasting performance for the first week of the test period (April 3-9, 2000), was given by:

$$\phi(B)p_t = \psi_1 z_t + d_1 D_{Mon} + d_2 D_{Sat} + d_3 D_{Sun} + \varepsilon_t, \tag{1}$$

where the autoregressive part $\phi(B)p_t = p_t - a_1 p_{t-24} - a_2 p_{t-48} - a_3 p_{t-168} - a_4 mp_t$, mp_t was the minimum of the 24 hourly (log-)prices on the previous day, z_t was the (log-)load forecast and $D_{Mon}, D_{Sat}, D_{Sun}$ were the dummy variables (for Monday, Saturday and Sunday). In this base model, denoted in the text as **ARX**, the noise term ε_t is i.i.d. Gaussian. Recall that the model, as well as its extensions described in the following Section, were estimated using an adaptive scheme, i.e. instead of using a single model for the whole sample, for every day (and hour) in the test period we calibrated the model (given its structure) to the previous values of prices and loads and obtained a forecasted value for that day (and hour).

3. Model extensions

The residuals obtained from the fitted ARX model seemed to exhibit a non-constant variance. Indeed, when tested with the Lagrange multiplier 'ARCH' test statistics (Engle 1982) the heteroskedastic effects were significant at the 5% level. Following Weron (2006) we calibrate an **ARX-G** model, where '**G**' stands for GARCH(1,1). It differs from the ARX model in that the noise term ε_t in eqn. (1) is not just $iid(0,\sigma^2)$ but is given by $h_t = \varepsilon_t \sigma_t$ with $\sigma_t^2 = \alpha_0 + \alpha_1 h_{t-1}^2 + \beta_1 \sigma_{t-1}^2$.

It has been long known that financial asset returns are not normally distributed. Rather, the empirical observations exhibit excess kurtosis (Carr et al. 2002; Rachev and Mittnik 2000). Bottazzi et al. (2005) and Weron (2006) have shown that electricity prices are also heavy-tailed. In particular, normal inverse Gaussian (NIG) and α-stable probability distributions provide a very good fit. The pertinent question is whether models with heavy-tailed innovations perform better in terms of forecasting accuracy than their Gaussian counterparts.

Following Weron and Misiorek (2007), we extend the basic model by allowing for a noise term ε_t that is governed by a heavy-tailed distribution: NIG or α-stable. The resulting models are denoted by **ARX-N** and **ARX-S**, respectively. Recall, that the NIG distribution is defined as a normal variance-mean mixture where the mixing distribution is the generalized inverse Gaussian law with parameter λ=–0.5, i.e. it is conditionally Gaussian. The probability density function of the NIG(α, β, δ, μ) distribution is given by:

$$f_{NIG}(x) = \frac{\alpha\delta K_1(\alpha\sqrt{\delta^2 + (x-\mu)^2})}{\pi\sqrt{\delta^2 + (x-\mu)^2}} e^{\delta\sqrt{\alpha^2-\beta^2}+\beta(x-\mu)}, \tag{2}$$

where $\delta > 0$ and $\mu \in R$ are the usual scale and location parameters, while α and β determine the shape, with α being responsible for the steepness and β, $|\beta| < \alpha$, for the skewness. The normalizing constant $K_1(t)$ is the modified Bessel function of the third kind with index 1. The tail behavior is often classified as 'semi-heavy', i.e. the tails are lighter than those of non-Gaussian stable laws, but much heavier than Gaussian (Weron 2004).

Stable laws – also called α-stable, stable Paretian or Lévy stable – require four parameters for complete description: the tail exponent $\alpha \in (0,2]$, which determines the tail thickness, the skewness parameter $\beta \in [-1, 1]$ and the usual scale, $\sigma > 0$, and location, $\mu \in R$, parameters. When $\alpha = 2$, the Gaussian distribution results. When $\alpha < 2$, the variance is infinite and the tails are asymptotically equivalent to a Pareto law, i.e. they exhibit a power-law decay of order $x^{-\alpha}$. In contrast, for $\alpha = 2$ the decay is exponential. From a practitioner's point of view the crucial drawback of the stable distribution is that, with the exception of three special cases (α = 2, 1, 0.5), its density and distribution function do not have closed form expressions. They have to be evaluated numerically, either by approximating complicated integral formulas or by taking the Fourier transform of the characteristic function (Weron 2004).

Heavy tailed laws provide a much better model for electricity price returns than the Gaussian distribution. Yet, a non-parametric kernel density estimator will generally yield a superior fit to any parametric distribution. Perhaps, time series models would lead to more accurate predictions if no specific form for the distribution of innovations was assumed. To test this conjecture we evaluate a semi-parametric model (denoted by **ARX-NP**; we call it semi-parametric because it has a parametric autoregressive part and a non-parametric noise distribution) for which no specific form for the distribution of innovations ε_t is assumed. Instead, to calibrate the parameters of the autoregression, we employ a non-parametric maximum likelihood (ML) routine. Such ML estimators can be derived by extending the ML principle to a non-parametric framework, where the innovations density is estimated by the Parzen–Rosenblatt kernel (Cao et al. 2003, Hsieh and Manski 1987). These non-parametric maximum likelihood estimators (NPMLE) generally perform well not only when the error distribution is Gaussian (or any other known parametric form), but also when only regularity conditions are assumed about the error density. On the other hand, the deficiency of these estimators with respect to ordinary ML estimators, under normality, should not be great if the non-parametric density estimator performs well.

In this study we use the smoothed NPMLE proposed by Cao et al. (2003). It (numerically) maximizes the likelihood $L_h(\Theta) = L(g_{h,\Theta}, \Theta)$, where

$$g_{h,\Theta}(x) = \frac{1}{(n-1)h} \sum_{i=2}^{n} K\left(\frac{x - \varepsilon_t(\Theta)}{h} \right) \tag{3}$$

is the non-parametric density, $K(\cdot)$ is the kernel and $\varepsilon_t(\Theta)$ are the model residuals for a given parameter vector Θ. For the sake of simplicity we use the Gaussian kernel and h = 0.105 (which roughly corresponds to the so-called 'rule of thumb' bandwidth $h = 1.06 \hat{\sigma} n^{-1/5}$, where $\hat{\sigma}$ is an estimator of the standard deviation of the error density). For more optimal bandwidth choices consult Jones et al. (1996).

Finally, let us note that nearly all computations are performed in Matlab 7.0. The ARX model is calibrated using the *armax.m* function, which minimizes the Final Prediction Error criterion (Weron 2006). The heavy-tailed and semi-parametric models are estimated by numerically maximizing the likelihood and the non-parametric likelihood function, respectively, with the ARX models' parameters as starting points of the unconstrained simplex search routine (*fminsearch.m* function). Only the ARX-G model is calibrated in SAS 9.0 (via ML), because Matlab's GARCH Toolbox yields significantly worse forecasts.

4. Empirical results

To assess the prediction performance of the models, different statistical measures can be utilized. The most widely used measures are those based on absolute errors, i.e. absolute values of differences between the actual, P_h, and predicted, $\hat{P}_h$, prices for a given hour, h. The Mean Absolute Percentage Error (MAPE) is a typical example. However, when applied to electricity prices, MAPE values could be misleading. In particular, when electricity prices drop to zero, MAPE values become very large regardless of the actual absolute differences $|P_h - \hat{P}_h|$. The reason for this is the normalization by the current (close to zero, and hence very small) price P_h. Alternative normalizations have been proposed in the literature. For instance, the absolute error $|P_h - \hat{P}_h|$ can be normalized by the average price attained during the day: $\overline{P}_{24} = \frac{1}{24}\sum_{h=1}^{24} P_h$. The resulting measure, also known as the Mean Daily Error, is given by (Conejo et al. 2005; Weron 2006):

$$\text{MDE} = \frac{1}{24}\sum_{h=1}^{24} \frac{\left|P_h - \hat{P}_h\right|}{\overline{P}_{24}}. \tag{4}$$

The Mean Weekly Error (MWE) corresponds to a situation when the number 24 is replaced by 168 in (4). Both errors are usually reported in percent, i.e. as MDE×100% or MWE×100%. The forecast accuracy was checked afterwards, once the true market prices were available. The MWE errors for the whole test period (April 3 – December 3, 2000) and all models are given in Table 1. Furthermore, to distinguish the rather calm first 10 weeks of the test period from the more volatile weeks 11-35 (see Fig. 1), in Table 2 the summary statistics are displayed separately for the two periods. These statistics are based on the 35 Mean Weekly or 245 Mean Daily Errors. In particular, the number of weeks (days) a given model was best in terms of MWE (MDE), the mean MWE (MDE) and the mean deviation from the best model in a given week (day). The latter statistics is defined as $\frac{1}{T}\sum_{t=1}^{T}(\text{E}_{i,t} - \text{E}_{\text{Best model},t})$, where i ranges over all evaluated models (i.e. i = 5), T is the number of weeks (10, 25) or days (70, 175) in the sample and E is either MWE or MDE.

The obtained results suggest that the semi-parametric specification ARX-NP is the best model. Of all the competitors it most often leads to the best point forecasts, both in the calm and volatile periods and both in terms of the weekly and daily measures (see rows labeled '# best(MWE)' and '# best(MDE)' in Table 2). Yet, it is not unanimously the best. While it has the lowest mean MWE in the calm period, in the latter 25 weeks, surprisingly, the ARX model beats it. Likewise, ARX-NP has the lowest mean MDE in the volatile period, but in the calm weeks both heavy-tailed specifications slightly overtake it. The mean deviations from the best model lead to the same conclusions. Nevertheless, ARX-NP can be considered the overall best model.

Table 1: Mean Weekly Errors (MWE; in percent) for all weeks of the test period. Best results in each week are emphasized in bold. Notice that the results for the ARX and ARX-G methods in this table were originally reported in Misiorek et al. (2006) and are re-produced here for comparison purpose.

Week	ARX	ARX-G	ARX-N	ARX-S	ARX-NP
1	**3.03**	3.60	3.39	3.55	3.44
2	**4.71**	5.46	5.73	5.35	5.30
3	**8.37**	8.92	8.71	8.73	8.58
4	13.51	13.48	**12.37**	12.84	13.49
5	17.82	18.22	17.70	17.64	**17.52**
6	8.04	8.26	8.00	7.89	**7.87**
7	9.43	10.72	9.93	9.82	**8.59**
8	48.15	45.55	44.44	**44.11**	45.18
9	13.11	15.19	13.38	13.30	**12.74**
10	**7.39**	8.10	7.63	7.59	8.01
11	**46.23**	53.64	52.95	52.10	48.42
12	19.23	19.18	**17.56**	17.67	18.47
13	**44.17**	56.00	56.88	56.51	49.21
14	27.99	28.22	**27.00**	27.21	28.03
15	11.11	16.99	11.76	11.94	**10.56**
16	**25.41**	33.45	30.67	29.67	26.04
17	**19.26**	32.49	26.45	25.66	19.79
18	**11.71**	26.47	16.33	15.96	11.97
19	14.47	**14.02**	14.94	14.62	15.74
20	**9.18**	15.19	10.49	10.50	9.57
21	13.91	18.51	14.10	14.58	**13.45**
22	20.28	22.40	20.55	20.27	**19.51**
23	23.27	24.64	23.78	23.72	**22.84**
24	**14.30**	17.83	14.97	14.83	14.75
25	17.28	22.92	18.29	18.20	**16.37**
26	13.97	13.30	13.43	**13.30**	13.75
27	10.65	11.13	10.67	10.56	**10.53**
28	7.93	7.57	7.46	7.50	**7.43**
29	7.36	8.41	7.65	8.25	**7.03**
30	10.22	**8.73**	9.21	9.53	10.20
31	13.35	**11.94**	12.59	12.88	13.12
32	11.43	11.29	**10.38**	11.03	11.28
33	11.09	12.92	**10.61**	11.13	11.35
34	12.40	**10.30**	10.61	11.22	12.67
35	5.07	4.74	**4.01**	4.72	4.83

Table 2: Summary statistics for the Mean Weekly Errors (MWE; presented in Table 1) and the Mean Daily Errors (MDE). The first number (before the slash) indicates performance during the first 10 weeks and the second – during the latter 25 weeks. Best results in each category are set in boldface.

Statistics	ARX	ARX-G	ARX-N	ARX-S	ARX-NP
# best(MWE)	**4**/7	0/4	1/5	1/1	**4/8**
Mean(MWE)	13.36/**16.85**	13.75/20.09	13.13/18.13	13.08/18.14	**13.07**/17.08
Mean dev. from best	0.69/**0.62**	1.08/3.86	0.46/1.90	0.41/1.91	**0.40/**0.84
# best(MDE)	17/35	9/43	15/25	5/18	**24/54**
Mean(MDE)	11.98/17.70	12.30/19.87	11.64/18.10	**11.63**/18.15	11.69/**17.64**
Mean dev. from best	1.55/2.59	1.87/4.76	1.21/2.99	**1.20**/3.04	1.26/**2.53**

Table 3: Mean percent of exceedances of the 50%, 90% and 99% two-sided day-ahead prediction intervals (PI) by the actual system price for the five considered models.

Weeks	50%	90%	99%	50%	90%	99%
		ARX			**ARX-G**	
1-10	41.96	**13.93**	5.60	41.85	13.15	4.88
11-35	46.10	**13.60**	5.52	**47.93**	15.64	6.81
		ARX-N			**ARX-S**	
1-10	60.95	14.70	3.39	58.57	15.48	**0.77**
11-35	65.02	18.40	2.05	65.38	19.31	**0.60**
		ARX-NP				
1-10	**42.92**	14.40	5.65			
11-35	46.31	14.17	5.50			

At the other end is the ARX-G model, which is inferior to the remaining competitors in most categories. It fails spectacularly in terms of the mean errors and the mean deviation from the best model. However, it can lead to the best predictions from time to time. Finally, the heavy-tailed models behave similarly. While ARX-N more often yields the best forecasts, ARX-S performs slightly better on average. Somewhat surprisingly, it is the calm period and not the volatile one that favors the heavy-tailed models relative to its competitors.

Apart from point forecasts, we investigated the ability of the models to provide interval forecasts. For all considered models interval forecasts were determined analytically; for details on calculation of conditional prediction error variance and interval forecasts we refer to Hamilton (1994) and Weron (2006). Afterwards, following Christoffersen and Diebold (2000) and Misiorek et al. (2006), we evaluated the quality of the interval forecasts by comparing the nominal coverage of the models to the true coverage. Thus, for each of the models we calculated prediction intervals (PIs) and determined the actual percentage of exceedances of the 50%, 90% and 99% two sided day-ahead PIs of the models by the actual system price, see Table 3. If the model implied interval forecasts were accurate then the percentage of exceedances should be approximately 50%, 10% and 1%, respectively. Note that in the calm period (first 10 weeks) 1680 hourly values were determined and compared to the system price for each of the models, while in the volatile period (weeks 11-35) – 4200 hourly values.

Examining the exceedances of the 50% interval we note that while the Gaussian, GARCH and semi-parametric models yield too wide PIs, the heavy-tailed alternatives behave quite the opposite. In this respect they exhibit a performance similar to the Markov regime-switching model analyzed in Misiorek et al. (2006). Looking at the exceedances of the 90% interval we see all models performing alike and yielding too narrow PIs. Yet, the ARX PIs are slightly better (wider) than those of the other models. Finally, the exceedances of the 99% interval present a different picture. The α-stable innovations lead to the widest (even a bit too wide) and closest to the optimal PIs. Next in line is the ARX-N model, the other three trail far behind. All of them yield too narrow PIs. In this category, the ARX-N model behaves comparably to the nonlinear Threshold TARX model analyzed in Misiorek et al. (2006).

Overall, the interval forecasting results are much less conclusive than the point forecasting ones. While ARX, ARX-G and ARX-NP are better in 50% and 90% intervals, they fail in 99% PIs, where the heavy-tailed models dominate. Among the three models – ARX, ARX-G and ARX-NP – the latter model could be considered the best, as it leads to more accurate point forecasts and comparable interval forecasts. In the whole group, however, the answer is not that obvious. The heavy tailed models could be preferred for risk management purposes since they yield more accurate upper quantiles of the error distribution.

Their behavior during the calm weeks is also comparable to that of the ARX-NP model. Surprisingly, only during the volatile period they perform below expectations.

5. Conclusions

In this paper we investigated the forecasting power of time series models for electricity spot prices. We expanded the standard autoregressive specification by allowing for heteroskedastic (GARCH), heavy-tailed (NIG, α-stable) and non-parametric innovations. The models were tested on a time series of hourly system prices and loads from California. We evaluated the quality of the predictions both in terms of the Mean Daily and Weekly Errors (for point forecasts) and in terms of the nominal coverage of the models to the true coverage (for interval predictions).

There is no unanimous winner of the presented competition. While in terms of point forecasts the semi-parametric ARX-NP model generally yields the best performance, when prediction intervals are considered the evidence is mixed. In particular, for risk management purposes, requiring accurate approximation of the upper quantiles, the heavy tailed models could be preferred. Although this study adds an important voice to the discussion of electricity spot price forecasting, more research – including evaluation of the models on other datasets – is needed.

Bibliography

Bottazzi, G., Sapio, S., Secchi, A. (2005) Some statistical investigations on the nature and dynamics of electricity prices, Physica A 355, 54-61.

Bunn, D.W. (2000) Forecasting loads and prices in competitive power markets, Proceedings of the IEEE 88(2), 163-169.

Bunn, D.W., ed. (2004) Modelling Prices in Competitive Electricity Markets, Wiley.

Cao, R., Hart, J.D., Saavedra, A. (2003) Nonparametric maximum likelihood estimators for AR and MA time series, Journal of Statistical Computation and Simulation 73(5), 347-360.

Carr, P., Geman, H., Madan, D.B., Yor, M. (2002) The fine structure of asset returns: An empirical investigation, Journal of Business 75, 305-332.

Christoffersen, P., Diebold, F.X. (2000) How relevant is volatility forecasting for financial risk management, Review of Economics and Statistics 82, 12-22.

Conejo, A.J., Contreras, J., Espinola, R., Plazas, M.A. (2005) Forecasting electricity prices for a day-ahead pool-based electric energy market, International Journal of Forecasting 21(3), 435-462.

Contreras, J., Espinola, R., Nogales, F.J., Conejo, A.J. (2003) ARIMA models to predict next-day electricity prices, IEEE Transactions on Power Systems 18(3), 1014-1020.

Engle, R.F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, Econometrica 50, 987-1007.

Jones, M.C., Marron, J.S., Sheather, S.J. (1996) A brief survey of bandwidth selection for density estimation, Journal of the American Statistical Association 91, 401-407.

Hamilton, J. (1994) Time Series Analysis, Princeton University Press.

Hsieh, D. A., Manski, C. F. (1987) Monte Carlo evidence on adaptive maximum likelihood estimation of a regression, Annals of Statistics 15, 541-551.

Misiorek, A., Trück, 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(3), Article 2.

Rachev, S., Mittnik, S. (2000) Stable Paretian Models in Finance, Wiley.

Weron, R. (2004) Computationally intensive Value at Risk calculations, in "Handbook of Computational Statistics: Concepts and Methods", eds. J.E. Gentle, W. Härdle, Y. Mori, Springer, 911-950.

Weron, R. (2006) Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach, Wiley. See also: http://www.im.pwr.wroc.pl/~rweron/MFE.html.

Weron, R., Misiorek, A. (2007) Heavy tails and electricity prices: Do time series models with non-Gaussian noise forecast better than their Gaussian counterparts?, Prace Naukowe Akademii Ekonomicznej we Wrocławiu Nr 1076, 472-480.