WORKING PAPERS

Jesús Crespo Cuaresma, Jaroslava Hlouskova
Stephan Kossmeier, Michael Obersteiner

Forecasting electricity spot prices using linear univariate time series models

Juni 2002

Working Paper No: 0208

DEPARTMENT OF ECONOMICS

UNIVERSITY OF VIENNA

Forecasting electricity spot prices using linear univariate time series models*

Jesús Crespo Cuaresma,† Jaroslava Hlouskova‡
Stephan Kossmeier,§ Michael Obersteiner¶

*The authors would like to thank Raimund Alt, Peter Gal, Alfred Hoffmann, Alexander Schnabl and Christoph Weber for helpful comments and research assistance. This work was partly supported by the European Community, 5th Framework Programme (1998-2002), Contract No. ENK5-CT-2000-00094, by the Swiss Federal Agency for Education and Science, Contract No. BBW 00.0627 and by Termoelekrarna toplarna Ljubljana, Contract No. CEU-BS-1/2001, as part of the OSCOGEN EU Project.

†University of Vienna; Department of Economics; Brünnerstrasse 72, 1210 Vienna, Austria, e-mail: jesus.crespo-cuaresma@univie.ac.at

‡Institute for Advanced Studies; Department of Economics and Finance; Stumpergasse 56, 1060 Vienna, Austria, e-mail: hlouskov@ihs.ac.at

§Vienna University of Technology; Department of Business Administration; Favoritenstrasse 9-11, 1040 Vienna, Austria, e-mail: stephan.kossmeier@tuwien.ac.at

¶International Institute for Applied Systems Analysis; Schlossplatz 1, 2361 Laxenburg, Austria, e-mail:oberstei@iiasa.ac.at

Abstract

This paper studies the forecasting abilities of a battery of univariate models on hourly electricity spot prices, using data from the Leipzig Power Exchange. The specifications studied include autoregressive models, autoregressive-moving average models and unobserved components models. The results show that specifications where each hour of the day is modelled separately present uniformly better forecasting properties than specifications for the whole time series, and that the inclusion of simple probabilistic processes for the arrival of extreme price events can lead to improvements in the forecasting abilities of univariate models for electricity spot prices.

Keywords: Electricity spot prices, ARMA models,
Structural time series, Forecasting.

1 Introduction

The ongoing worldwide deregulation of electricity markets has created an increasing interest for building econometric models of electricity prices with the aim of both understanding the dynamics of electricity price formation and obtaining reliable forecasts. Due to the nonstorable nature of the underlying good (and therefore to the impossibility of inventories to be used as price arbitrage devices linking expectations and spot prices), the dynamics of electricity prices present certain stylized facts that cannot be adequately captured by models whose focus is the price behaviour of storable commodities and financial securities. In particular, time series of electricity spot prices exhibit more structure which can be used for forecasting compared to time series of financial securities, which are usually quite well described by Markovian processes. Due to the youth of deregulated electricity markets, few studies have approached the subject of time series modelling of electricity prices directly. De Vany and Walls (1999) use cointegration methods in order to assess the issue of integration of regional electricity markets and convergence of electricity prices in the US. Knittel and Roberts (2001) present an empirical analysis of electricity prices in California and comment on the forecasting properties of several simple time series models, while Escribano, Peña and Villaplana (2002) present a relatively general modelling strategy for electricity prices, and apply it to four different markets.

The aim of this paper is to compare the performance of univariate time series models in forecasting electricity spot prices. Electricity spot prices present several types of superposed seasonal cycles, mean reversion and price spikes. Including such stylized facts in time series models for electricity prices implies that in some cases we may want to drift away from the assumption of linearity in the modelling strategy. While nonlinear time series analysis provides an interesting framework in order to approach modelling prices of nonstorable goods [see, e.g., Robinson (2000) for an example in the electricity market], this paper concentrates exclusively on linear univariate models, and it focuses on electricity spot price predictions.[1] The performance of a battery of univariate models, including various types of ARMA models, models with unobserved components and with jumps, is studied in forecasting electric-

[1] As it will appear clear below, some of the models used in this paper could also be viewed as piecewise linear models.

ity spot prices using hourly data from the German Leipzig Power Exchange (LPX) from its opening on June 16th, 2000 at 1:00 to October 15th, 2001 at 24:00. The set of models includes both models based on the complete hourly time series of electricity prices and models where the dynamic behaviour of each hour is modelled using different (potentially interrelated) dynamic processes, and the forecasting superiority will be assesed using predictive accuracy tests.

The paper is organized as follows. Section two presents a description of the market and the data used in the study. The models that will be used in order to obtain forecasts for electricity spot prices are presented and motivated in section three, while section four presents the results of the forecasting exercise and section five concludes and indicates potential future paths of research.

2 The LPX market and the price data

The Leipzig Power Exchange (LPX) commenced trading in June 2000. Since then it has established itself as the most liquid spot market for electricity in Germany. In June 2002, LPX will be merged with its former rival, the European Energy Exchange (EEX) in Frankfurt, to form potentially the most important continental European electricity exchange. Today about 90 trading participants, German as well as international, are active at LPX. Monthly trading volumes regularly exceed 1.5 million MWh.

The LPX is organized as a day ahead market. Participants submit buying and selling bid curves for each of the next 24 hours. Every day at noon the exchange aggregates bids for each hour and determines market clearing prices and volumes for each hour of the following day. Afterwards LPX transmits schedules to the Transmission System Operators and informs the bidders about prices and amounts to be delivered and received. Contracts can be concluded for 365 days of a year.

A total of 11,688 observations of hourly LPX electricity spot prices in Euro per megawatt (Euro/MWh) are available and they are plotted in Figure 1. The sample period begins on June 16th, 2000 (the opening of the market) and ends on October 15th, 2001. Figure 2 presents the average price for each hour for weekdays (Monday-Friday) - solid line - and weekends - dotted line - over the whole sample. Prices are significantly higher during weekdays, and both for weekends and weekdays a relatively similar intraday pattern emerges: the price begins to increase at around 5:00 during the workday (7:00 for weekends) and continues to increase until 12:00 when there is the first and biggest peak of the day. Then the price begins to fall until 17:00 and after reaching its locally lowest point it starts to increase again until 19:00-20:00 when it reaches the second peak of the day. Prices begin to fall thereafter, until the 5:00 (7:00) turning point appears.

Table 1 presents summary statistics for the whole sample of hourly electricity prices and for each one of the 24 hours of the day. The null hypothesis of a normal distribution tested by the Jarque-Bera test statistic is not rejected at a 5% significance level for hours 1:00−5:00, 8:00, 24:00. Intuitively, this suggests that models based on the normality assumption applied on the electricity prices of hours 1:00−5:00, 8:00 and 24:00 have bigger chances of accurately representing the data generating process than for hours 6:00, 7:00, 9:00−23:00. Note that the electricity prices for the first peak hour, 12:00, have the highest mean and standard deviation and the electricity prices for the second peak hour, 18:00, have the highest skewness and kurtosis.

A remarkable characteristic of energy commodity prices is the presence of price spikes (see Figure 1). Such a phenomenon is usually explained by either supply-sided (unplanned outage of a large power plant) or demand-sided shocks (heat wave in summer). On the other hand, also market mechanism failure and capacity constraints of the network can cause spikes, because they lead to temporary deviations from perfect competition in the market and therefore to price spikes when temporary monopolists or oligopolists make use of their market power. Table 2 shows the mean and variance of the jumps which are identified using a recursive filter in the spirit of Clewlow and Stickland (2000) for both the whole sample and each hour separately. The underlying probability of a jump, λ, assumed to be constant for a given hour of the day, is also presented in Table 2. Note that (positive) jump probabilities for each hour vary during the day from 0.004 to 0.014. The recursive filter approach used to extract the price spikes begins by identifying as jumps those observations which are higher than the ν-quantile of the normal distribution centered around the empirical mean of the electricity spot price time series, and with the standard deviation of the complete sample (as long as this proportion of observations is higher than $(1-\nu)$). In our case, ν was set to be 0.997. The identified jumps are removed from the sample, the standard deviation is computed again and the same procedure is repeated until the proportion of observations with a higher value than the ν-quantile is smaller than ν. Notice as well that the existence of jumps is associated with a leptokurtic and long right-tailed distribution of electricity prices (see the histogram of the time series in Figure 3).

3 Time series models for electricity prices

This section presents the battery of univariate models that will be used to obtain predictions of electricity spot prices. All models will be estimated using both the complete in-sample dataset as a single time series and twenty four time series, each one corresponding to an hour of the day.[2]

3.1 AR(1) process

It has been well documented that an important property of energy spot prices is mean-reversion [see, e.g., Gibson and Schwartz (1990), Brennan (1991)]. The benchmark model for mean reverting processes used will be a simple first order autoregressive process [AR(1)], which could be thought of as an exact discrete time version of an Ohrstein-Uhlenbeck process [see, e.g., Knittel and Roberts (2001)]. The stochastic process for electricity prices (p_t) is, thus,

$$p_t = \alpha + \beta p_{t-1} + \eta_t, \qquad \eta_t \sim \mathrm{NID}(0, \sigma_\eta^2) \tag{3.1}$$

where the error term, η_t, is assumed to be white noise with constant variance σ_η^2. The conditional mean of p_t given information up to and including period $t-1$ is, thus,

$$\hat{p}_t = E(p_t | \mathcal{I}_{t-1}) = \alpha + \beta p_{t-1},$$

where $\mathcal{I}_{t-1}$ is the σ-algebra generated by $\{p_0, p_1, \ldots, p_{t-2}, p_{t-1}\}$, and the conditional variance of $\hat{p}_t$ is σ_η^2.

[2]The approach based on disaggregating high frequency data into different time series corresponding to a time unit has been used by, e.g., Ramanathan *et. al* (1997) for electricity loads and Bauer, Deistler and Scherrer (2001) for ozone data.

This class of models is able to reproduce mean reversion and therefore to capture some of the autocorrelation present in the price series, but it ignores certain other cycles present in the series (intraday, weekend/weekday and seasonal patterns), and it assumes that the error structure is independent across time. Furthermore, linear Gaussian models such as (3.1) cannot accomodate the price spikes found in the data. The AR(1) model will, thus, act as a benchmark model in the forecasting exercise.

For the case of the disaggregated daily time series for hour $\{z\}$, $z \in \{1, 2, \ldots, 23, 24\}$, the analogous model can be specified as

$$p_t^{\{z\}} = \alpha^{\{z\}} + \beta^{\{z\}} p_{t-1}^{\{z\}} + \eta_t^{\{z\}}, \qquad \eta_t^{\{z\}} \sim \mathrm{NID}(0, \sigma^2_{\eta^{\{z\}}})$$

where $\eta_t^{\{z\}}$ is assumed to be white noise, uncorrelated with $\eta^{\{s\}} \forall s \neq z$. Notice that $p_{t-1}^{\{z\}}$ corresponds to p_{t-24} in the data with hourly frequency. The model using disaggregated time series accounts for intraday seasonality by estimating different parameters for each different hour of the day, and could be viewed as a global piecewise-linear model in which the parameters depend on the hour corresponding to period t. That is,

$$p_t = \sum_{z=1}^{24} (\alpha^{\{z\}} + \beta^{\{z\}} p_{t-24} + \eta_t^{\{z\}}) I(t \text{ is hour } z),$$

where $I(\cdot)$ is an indicator function taking value one if the argument is true and zero otherwise, and $\eta_t^{\{z\}}$ is white noise, assumed to be uncorrelated with $\eta_t^{\{s\}}$ $\forall s \neq z$.

3.2 AR (1) process with time-varying intercept

The second model addresses the systematic seasonal variation found in electricity prices by allowing the intercept in the simple AR(1) process described above to change depending on the hour of the day, day of the week and month of the year corresponding to the period being modelled. The seasonal patterns will be assumed to remain constant through time and will all be modelled using dummies. The resulting AR(1) process with time-varying-intercept that will actually be used is

$$p_t = \alpha_t + \beta p_{t-1} + \eta_t, \qquad \eta_t \sim \mathrm{NID}(0, \sigma^2_\eta) \tag{3.2}$$

with

$$\alpha_t = \alpha_0 t + \sum_{i=1}^{24} \alpha_{1,i} I(t \text{ is hour } i) + \sum_{j=1}^{7} \alpha_{2,j} I(t \text{ is in day } j) + \sum_{k=1}^{12} \alpha_{3,k} I(t \text{ is in month } k). \tag{3.3}$$

Note that this specification allows as well for a global trend, independent of the seasonal pattern, as implied by the first summand of the expression of α_t above.

When dealing with the disaggregated time series, that is, with twenty four time series each one corresponding to an hour, the specified model is

$$p_t^{\{z\}} = \alpha_t^{\{z\}} + \beta^{\{z\}} p_{t-1}^{\{z\}} + \eta_t^{\{z\}}, \qquad \eta_t^{\{z\}} \sim \text{NID}(0, \sigma^2_{\eta^{\{z\}}})$$

for hour $z \in \{1, 2, \ldots, 23, 24\}$ where

$$\alpha_t^{\{z\}} = \alpha_0^{\{z\}} t + \sum_{j=1}^{7} \alpha_{2,j}^{\{z\}} \mathcal{I}(t \text{ is in day } j) + \sum_{k=1}^{12} \alpha_{3,k}^{\{z\}} \mathcal{I}(t \text{ is in month } k). \tag{3.4}$$

3.3 ARMA process with time-varying intercept

Working in a discrete time framework, price dynamics can be specified as generalizations of the mean reverting processes presented above by specifying a complete autoregressive-moving average [ARMA (p, q)] model with time varying intercept such as

$$p_t = \alpha_t + \sum_{i=1}^{p} \phi_i p_{t-i} + \epsilon_t + \sum_{j=1}^{q} \theta_j \epsilon_{t-j}, \qquad \epsilon_t \sim \text{NID}(0, \sigma^2_\epsilon) \tag{3.5}$$

where α_t is specified as in (3.3) and $\{\epsilon_t\}$ is a white noise process with constant variance σ^2_ϵ. For hourly disaggregated data the specified ARMA model for hour $\{z\}$ is

$$p_t^{\{z\}} = \alpha_t^{\{z\}} + \sum_{i=1}^{p} \phi_i^{\{z\}} p_{t-i}^{\{z\}} + \epsilon_t^{\{z\}} + \sum_{j=1}^{q} \theta_j^{\{z\}} \epsilon_{t-j}^{\{z\}}, \qquad \epsilon_t^{\{z\}} \sim \text{NID}(0, \sigma^2_{\epsilon^{\{z\}}}) \tag{3.6}$$

where the hour-specific time-varying intercept $\alpha_t^{\{z\}}$ is defined as in (3.4), and $\epsilon_t^{\{z\}}$ is assumed to be white noise.

3.4 Crossed ARMA process with time-varying intercept

More flexibility could be achieved by allowing the electricity spot price in hour z to depend upon price realizations in hour $s \neq z$, under the assumption that all hour-specific shocks are uncorrelated with each other. The model specification is, thus,

$$p_t^{\{z\}} = \alpha_t^{\{z\}} + \sum_{i=1}^{p} \phi_i^{\{z\}} p_{t-i}^{\{z\}} + \sum_{j=1}^{24-z} \pi_j^{\{z\}} p_{t-1}^{\{z+j\}} + \sum_{k=1}^{z-1} \varsigma_k^{\{z\}} p_t^{\{z-k\}} + \epsilon_t^{\{z\}} + \sum_{l=1}^{q} \theta_l^{\{z\}} \epsilon_{t-l}^{\{z\}}, \tag{3.7}$$

where $\epsilon_t^{\{z\}}$ is assumed to be white noise, uncorrelated with $\epsilon_t^{\{s\}}$ $\forall s \neq z$.

3.5 ARMA processes with jumps

A remarkable characteristic of electricity spot prices is the presence of price spikes. In a continuous-time framework, jump-diffusion models consider the possibility of large short-lived variations of the underlying variable and thus might be appropriate for modeling electricity spot prices. Jump-diffusion models link price changes to the arrival of information, and considers the existence of two types of news: normal news, which produce continuous price dynamics and abnormal news, which cause discrete price jumps and whose arrival is modelled using a probabilistic discrete time process. Jump-diffusion models can account for conditional density functions with fat tails and non-zero skewness, whose sign depends on the mean jump size. Taking into account the inclusion of jumps, the more general price process specified by (3.5) can be augmented by appending an additional term that represents the arrival of abnormal shocks, yielding

$$p_t = \alpha_t + \sum_{i=1}^{p} \phi_i p_{t-i} + \epsilon_t + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \Xi_t j_t, \tag{3.8}$$

where j_t is a discrete time probability process governing the arrival of price jumps and Ξ_t is the jump size. This generalization can be applied to any of the models exposed above. In a continuous-time setting the arrival of shocks is usually modelled by a Poisson process [see e.g. Knittel and Roberts (2001), Huisman and Mahieu (2001)], but for simplicity the forecasting experiment will be carried out using a binomial process, that is,

$$j_t = \begin{cases} 0 & \text{with probability} \quad 1-\lambda, \\ 1 & \text{with probability} \quad \lambda. \end{cases}$$

This implies that the shock arrival process is constant over time (over days at a given hour for the models treating with 24 time series): in a given period prices are drawn from a process such as (3.5) with probability $1-\lambda$ and with probability λ the electricity price is drawn at time t from the same process augmented additively with Ξ_t. The jump size, Ξ_t will be modelled as the realization of a normally distributed random variable with expectation ξ and variance σ^2_{Ξ}. The jumps, its mean and variance as well as the probability of a jump ocurring are identified using the recursive filter presented in section 2. Given the simple process governing the arrival of jumps, they can be easily accomodated to the conditional expectation of the models described above in order to compute forecasts just by adding $\hat{\Xi}_t\hat{\lambda}$ to the forecasts computed from the data without jumps.[3] When the electricity spot price in the LPX market is modelled in an hour-by-hour fashion, we will allow for different λs for each daily time series of prices at a given hour, as the uneven distribution of price spikes across hours of the day seems to be a stylized fact of electricity spot prices (see Table 2).

3.6 Unobserved components model

Structural time series models are based on the identification of unobserved components wich are directly interpretable out of the data, and have been widely used in economics [see Harvey (1985, 1989) or Harvey and Jaeger (1993) for the most relevant examples]. Assume that the process of interest (p_t) can be decomposed in an additive fashion into a trend component (μ_t), a cycle component (ψ_t), a seasonal component (γ_t) and an irregular component (ϵ_t), where the trend captures long-term movements of the series, the cyclical component is a sine-cosine wave with constant frequency, ι, the seasonal component will be modelled using seasonal dummies and the irregular component is assumed to be white noise. That is,

$$p_t = \mu_t + \psi_t + \gamma_t + \epsilon_t, \qquad \epsilon_t \sim \text{NID}(0, \sigma^2_\epsilon). \tag{3.9}$$

[3] Actually, if jumps are removed, the models refer to the time series without jumps (say $\tilde{p}_t$), so that forecasts can be computed as $\hat{p}_t = \hat{\tilde{p}}_t + \hat{\Xi}_t\hat{\lambda}$.

The trend component is specified in its most general form as follows,

$$\begin{aligned}\mu_t &= \mu_{t-1} + \beta_{t-1} + \upsilon_t, & \upsilon_t &\sim \text{NID}(0, \sigma_\upsilon^2),\\ \beta_t &= \beta_{t-1} + \delta_t, & \delta_t &\sim \text{NID}(0, \sigma_\delta^2),\end{aligned}$$

where υ_t and δ_t are disturbances uncorrelated mutually and with the irregular component, ϵ_t. It can be easily noticed that such a specification of the trend nests as special cases the linear time trend model (if σ_δ^2=0 and σ_υ^2=0), the random walk with drift (if σ_δ^2=0 and σ_υ^2 >0) and the smooth trend model (if σ_δ^2 >0 and σ_υ^2=0).

The cyclical component is specified as

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos\iota & \sin\iota \\ -\sin\iota & \cos\iota \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \omega_t \\ \omega_t^* \end{bmatrix},$$

where $\iota \in [0, \pi]$ is the cyclical frequency, $\rho \in (0, 1)$ is the damping factor of the cycle, ψ_t^* appears by construction and ω_t and ω_t^* are iid normally distributed disturbances, mutually uncorrelated and with equal, fixed variance σ_ω^2. Notice that the cycle is an ARMA(2,1) process where the autoregressive roots are constrained to be complex [see Harvey (1985)], leading to pseudo-cyclical behaviour.

The seasonal component is defined in a similar fashion like for the ARMA models presented above, that is,

$$\gamma_t = \sum_{i=1}^{24} \alpha_{1,i} I(t \text{ is hour } i) + \sum_{j=1}^{7} \alpha_{2,j}^{\{z\}} I(t \text{ is in day } j),$$

where the monthly seasonal pattern will be captured by the cyclical component and is therefore excluded from the specification of the seasonal component.

4 Forecasting evaluation

4.1 The forecasting exercise: setting and results

The dataset is divided into an in-sample period comprising the first 10,607 observations (June 16$^{\text{th}}$, 1:00, 2000 to August 31$^{\text{st}}$, 24:00, 2001) and an out-of-sample period composed by the remaining 1,080 observations (from September, 1st, 1:00, 2001 to October, 15th, 24:00, 2001), which will be used to assess the forecasting abilities of the different models. The out-of-sample period is plotted in Figure 4, and includes several price spikes. The forecasting exercise is designed as follows. The in-sample data is used to estimate the model of interest, and up to 168 hours (one week)-ahead forecasts are computed from the estimated model.[4] Two measures of the forecast error, root mean square error (RMSE) and mean absolute error (MAE), are then calculated for each model as

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{1}{168}\sum_{t=S+1}^{S+168}(\hat{p}_t - p_t)^2}, \\ \text{MAE} &= \frac{1}{168}\sum_{t=S+1}^{S+168}|\hat{p}_t - p_t|, \end{aligned}$$

where $\hat{p}_t$ refers to the forecasted price for period t, p_t is the actually realized electricity price in period t and S is the period corresponding to the last in-sample observation.[5] The in-sample period is then enlarged by one observation and again forecasts up to 168 steps-ahead are computed, together with the corresponding RMSE and MAE. This procedure is repeated 913 times and the average RMSE and MAE are computed for each model. The results are presented in Table 3 for the whole collection of models.

For the case of the benchmark AR(1) model, four different models were estimated and used for forecasting using the complete dataset without hourly disaggregation (global models): AR(1), AR(1) in logs, AR(1) with jumps and AR(1) in logs with jumps. The AR(1) model using logged data avoids negative forecasts of prices, but proves to produce worse forecasts than the AR(1) model using raw data.[6] Two extra models are considered for the AR(1) class

---

[4]The results of the model estimation are not presented, and are available from the authors upon request.

[5]Note that, although the loss function differs, the same weight is given to each step-ahead forecast in computing the root mean square error and the mean absolute error. Although a different weighting strategy could have been carried out if the forecaster has reasons to believe that observations at some forecasting horizon are more valuable than others, we decided to use a more unconstrained setting.

[6]Although our dataset does not contain negative prices on electricity, such a phe-

with time-varying mean. The unrestricted AR(1) model with time-varying intercept includes all estimated parameters from the specification (3.2), while the restricted AR(1) model with time-varying intercept elliminates the parameters which appear statistically insignificant using sequential t-tests for the significance of the parameters of the seasonal dummies, starting with the most insignificant parameter until all remaining parameters are individually significant at least at a 5% significance level. In the model with only significant parameters, F-tests were performed to assess equality of seasonal effects across all possible combinations of (significant) hours of the day, days of the week and months of the year. If the F-test did not reject equality of parameters at 5% significance level and the proportion of the variance explained by the model did not decrease, the model was further restricted to contain equal parameters in those seasonal dummies which were being tested. Using the global dataset for the AR(1) class with time-varying intercept, models with restricted parameters present better forecasting properties than their unrestricted analogs. While the inclusion of jumps improves the forecasting performance of the AR(1) model with constant intercept, it worsens the forecasting properties of the AR(1) model with time-varying mean.

Eight different ARMA models are estimated and used for forecasting for the complete hourly time series. The AR and MA orders were chosen after examining the correlogram and for most models include lag 1, 23, 24 and 25. Inside the ARMA class of models for the hourly time series the unrestricted model in logs performs best according to the RMSE, and the unrestricted ARMA model in levels according to MAE. The inclusion of jumps worsens the forecasting performance, and the best models of this class perform relatively worse that the best models of the AR(1) class with time-varying intercept.

The unobserved components model is estimated by maximum likelihood after using Kalman filtering.[7] After trying different trend specifications, the restriction $\sigma_{\delta}^2 = \sigma_{v}^2$=0 was imposed, as allowing for more flexibility in the trend component lead to worse forecasts. The relatively bad forecasting properties of the structural time series model in the global setting can be

nomenon appears in other time series of electricity prices [see e.g. Knittel and Roberts (2001), using data for the electricity market in Northern California].

[7]For technical details on the estimation methodology, see Harvey (1989).

partly explained by the fact that part of the seasonal dynamics are captured by the cyclical error, leading to predictions that tend to understate intraday seasonal variations in the price of electricity.

The second part of Table 3 presents the results for the models based on 24 time series (separable models), each one corresponding to an hour of the day. It should be noted that every single model in the separable setting outperforms its corresponding global model independently of the forecast error statistic (RMSE or MAE) used to compare the forecasting abilities. Two different classes of ARMA models are considered in the separable setting, depending on whether the autoregressive lags refer exclusively to past days of the hour to be modelled (hourly ARMA models) or if cross-correlations among hour t and hour $t-v$, for $v = 1, 2, \ldots, 23$ are also included in the model (crossed ARMA models). Notice that this last class of models assumes that the error corresponding to a given hour $z \in \{1, 2, \ldots, 24\}$ is uncorrelated with contemporaneous, past and future errors of hour $s \in \{1, 2, \ldots, 24\}, s \neq z$.

The inclusion of jumps, even when it is in the simple form described above, improves the forecasting abilities in all classes of separable models independently of the loss function taken into account when computing the forecast error. The model that outperforms all others and, therefore, presents the best forecasting capabilities among the fifty estimated models is the separable restricted crossed ARMA model with jumps, with a RMSE of 3.99 and a MAE of 2.57. A graphical illustration of the different forecasting abilities of the models considered can be found in Figures 5 and 6. Figure 5 presents the point forecasts generated by the model with worst forecasting abilities among all models studied [global AR(1) with constant intercept] for the last out-of-sample week. Figure 6 presents the forecasts in the last out-of-sample week for the model with the smallest forecast error (separable restricted crossed ARMA model with jumps).

4.2 Testing for equal predictive accuracy

In order to assess whether the observed differences in forecasting power across models are actually significant, the Diebold-Mariano (DM) test for predictive accuracy [Diebold and Mariano (1995)] was performed among the models which present better forecasting power inside each class. The DM test approach aims to test the null hypothesis of equality of expected forecast accuracy against the alternative of different forecasting ability across models. Assume that we are trying to compare the forecasts produced by two models, A and B up to h-steps ahead. The null hypothesis of the test can be, thus, written as

$$d_t = E[g(e_t^A) - g(e_t^B)] = 0, \tag{4.10}$$

where e_t^i refers to the difference between the forecasted price and the actual price for model i ($i = A, B$), and $g(\cdot)$ is the corresponding loss function.

The Diebold-Mariano test uses the autocorrelation-corrected sample mean of d_t in order to test for (4.10). If n observations and forecasts are available, the test statistic is, therefore,

$$S = [\hat{V}(\bar{d})]^{-1/2}\bar{d},$$

where

$$\hat{V}(\bar{d}) = \frac{1}{n}(\gamma_0 + 2\sum_{k=1}^{h-1}\hat{\gamma}_k),$$

and

$$\hat{\gamma}_k = \frac{1}{n}\sum_{t=k+1}^{n}(d_t - \bar{d})(d_{t-k} - \bar{d}).$$

Under the null hypothesis of equal forecast accuracy, S is asymptotically normally distributed. The model that presents best forecasting properties inside a given class is chosen as the representative of that class, and a test of equal forecasting accuracy is performed against the model with best overall forecasting properties (the separable crossed ARMA model with jumps and restricted coefficients). The results are presented in Table 4 for both of the loss functions used in the forecasting experiment (RMSE and MAE). The separable crossed ARMA model with jumps and restricted coefficients presents highly significant improvements in forecasting accuracy when compared to the rest of the representatives of the model classes. Given the large number of out of sample observations in the forecasting exercise, the results give very strong evidence of better predictive abilities of this model against all others.

5 Conclusions and paths for further research

The current liberalization process taking place in electricity markets worldwide has increased the interest for econometric models with good forecasting properties for electricity spot prices. Using data for the LPX market, the results presented in this paper indicate that an hour-by-hour modelling strategy for electricity spot prices improves significantly the forecasting abilities of linear univariate time series models, and that assessing the process of arrival of price spikes, even if it is in a simple manner, can also lead to better forecasts. The result is not trivial, as it is not always the case that models that are able to reproduce in-sample stylized facts produce better out-of-sample forecasts. We have abstracted from modelling volatility clustering and potential nonlinearities of the data generating process (DGP), and the process leading to the existence of price spikes has been kept as simple as possible. A straightforward path of further research would involve accounting for more sophistication in the DGP by studying the forecasting properties of electricity spot price models with time-varying volatility and potential nonlinearities in the conditional mean process. The computational burden, however, is greatly enlarged by allowing for nonlinearity, as in general there exists no closed-form analytic expressions for multi-step forecasts in nonlinear models.

Table 1: Descriptive statistics for LPX electricity spot prices over period June 16^{th}, 2000 - October 15^{th}, 2001. Hours marked with * do not reject at 5% significance level the null hypothesis of a normal distribution tested by Jarque-Bera test statistic.

Hour	Mean	St. Dev.	Skewness	Kurtosis
1:00*	15.01	4.04	0.077	2.960
2:00*	13.25	4.03	-0.038	2.574
3:00*	12.29	4.04	-0.004	2.650
4:00*	11.88	4.08	0.010	2.570
5:00*	12.18	4.23	-0.147	2.620
6:00	13.19	4.38	-0.434	2.785
7:00	15.62	5.43	-0.601	2.694
8:00*	20.43	8.11	-0.028	2.734
9:00	23.52	8.89	0.348	3.500
10:00	25.75	9.26	0.861	5.359
11:00	28.30	10.01	1.105	5.623
12:00	**34.87**	**16.59**	2.758	19.950
13:00	28.34	10.07	3.526	36.815
14:00	25.84	9.14	1.201	8.464
15:00	23.31	8.19	0.783	4.687
16:00	21.44	7.08	0.567	4.435
17:00	20.37	6.51	0.603	4.247
18:00	21.52	9.54	**6.412**	**88.920**
19:00	22.25	7.69	1.329	6.311
20:00	22.04	7.03	1.683	12.333
21:00	20.90	5.53	1.360	9.290
22:00	19.60	4.25	1.340	13.200
23:00	19.29	3.65	0.641	7.140
24:00*	16.86	4.05	-0.135	3.388
Whole sample	20.33	9.50	2.370	23.500

Table 2: Mean of jumps (μ_j), jump variances (σ_j^2) and jump probabilities (λ) for specific hours and the whole sample

Hour	μ_j	σ_j^2	λ
1:00	7.19	172.69	0.008
2:00	no jumps are identified		
3:00	no jumps are identified		
4:00	no jumps are identified		
5:00	no jumps are identified		
6:00	no jumps are identified		
7:00	no jumps are identified		
8:00	no jumps are identified		
9:00	29.60	19.72	0.010
10:00	36.88	119.85	0.008
11:00	39.74	54.90	0.012
12:00	75.85	1258.28	0.014
13:00	68.45	2128.27	0.006
14:00	41.72	495.72	0.006
15:00	29.04	40.44	0.014
16:00	28.79	13.19	0.008
17:00	23.79	7.45	0.010
18:00	82.87	5859.03	0.004
19:00	31.50	77.43	0.010
20:00	29.03	200.77	0.012
21:00	25.52	65.78	0.010
22:00	17.92	75.60	0.012
23:00	19.17	67.63	0.004
24:00	7.93	182.66	0.008
Whole sample	36.41	17.30	0.018

Table 3: Forecast performance. Out of sample period: September 1st, 1:00, 2001 to October 15th, 24:00, 2001 (1080 observations) Forecast horizon: one week (168 hours)

Global Models (single time series)		
Model	RMSE	MAE
AR(1)	9.464	6.697
AR(1), in logs	10.048	7.131
AR(1), jumps	9.396	6.676
AR(1), jumps, in logs	9.833	6.967
AR(1), varying intercept	5.306	3.852
AR(1), varying intercept, significant	5.461	4.067
AR(1), varying intercept, in logs	4.917	3.259
AR(1), varying intercept, in logs, significant	4.893	3.249
AR(1), varying intercept, jumps	5.547	3.999
AR(1), varying intercept, jumps, significant	5.719	4.267
AR(1), varying intercept, jumps, in logs	5.028	3.347
AR(1), varying intercept, jumps, in logs, significant	4.983	3.325
ARMA, varying intercept	5.354	3.542
ARMA, varying intercept, significant	5.252	3.428
ARMA, varying intercept, in logs	5.164	3.465
ARMA, varying intercept, in logs, significant	5.228	3.505
ARMA, varying intercept, jumps	5.585	3.712
ARMA, varying intercept, jumps, significant	5.659	3.809
ARMA, varying intercept, jumps, in logs	5.259	3.476
ARMA, varying intercept, jumps, in logs, significant	5.375	3.605
Unobserved components model	7.923	5.942
Separable Models (24 time series)		
AR(1)	6.536	4.641
AR(1), in logs	6.995	5.007
AR(1), jumps	6.459	4.583
AR(1), jumps, in logs	6.866	4.922
AR(1), varying intercept	4.212	2.769
AR(1), varying intercept, significant	4.347	2.886

Table 3: (continued)

AR(1), varying intercept, in logs	4.319	2.854
AR(1), varying intercept, in logs, significant	4.476	2.979
AR(1), varying intercept, jumps	4.130	2.702
AR(1), varying intercept, jumps, significant	4.197	2.752
AR(1), varying intercept, jumps, in logs	4.211	2.775
AR(1), varying intercept, jumps, in logs, significant	4.266	2.813
ARMA, varying intercept	4.527	2.840
ARMA, varying intercept, significant	4.447	2.784
ARMA, varying intercept, in logs	4.288	2.733
ARMA, varying intercept, in logs, significant	4.266	2.702
ARMA, varying intercept, jumps	4.210	2.731
ARMA, varying intercept, jumps, significant	4.139	2.671
ARMA, varying intercept, jumps, in logs	4.194	2.703
ARMA, varying intercept, jumps, in logs, significant	4.185	2.688
Crossed ARMA, varying intercept	4.231	2.697
Crossed ARMA, varying intercept, significant	4.138	2.641
Crossed ARMA, varying intercept, in logs	4.049	2.617
Crossed ARMA, varying intercept, in logs, significant	4.158	2.670
Crossed ARMA, varying intercept, jumps	4.065	2.601
Crossed ARMA, varying intercept, jumps, significant	**3.993**	**2.568**
Crossed ARMA, varying intercept, jumps, in logs	4.019	2.619
Crossed ARMA, varying intercept, jumps, in logs, significant	4.037	2.627
Unobserved components model	5.837	3.734

Table 4: Diebold-Mariano test statistics for equal forecasting accuracy. Restricted crossed ARMA with varying intercept and jumps against best forecasting models of all other classes. * (**) [***] indicates rejection of the null hypothesis of equal forecasting accuracy at 10% (5%) [1%] significance level.

Model class	**DM (RMSE)**	**DM (MAE)**
Global Models (single time series)		
AR(1)	-5.403***	-4.108***
AR(1), varying intercept	-0.900***	-0.682***
ARMA, varying intercept	-1.235***	-0.860***
Unobserved components model	-3.930***	-3.374***
Separable Models (24 time series)		
AR(1)	-2.560***	-2.080***
AR(1), varying intercept	-0.211***	-0.183***
Hourly ARMA, varying intercept	-0.157***	-0.086***
Unobserved components model	-1.843***	-1.167***

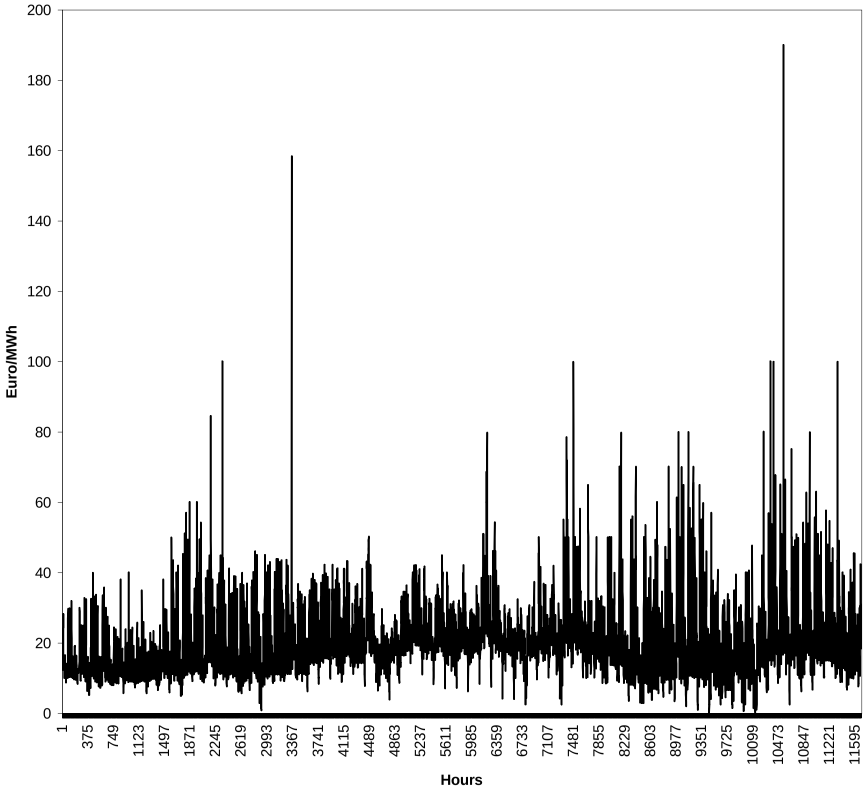

Figure 1: LPX electricity spot price over the period June 16$^{\mathrm{th}}$, 2000 - October 15$^{\mathrm{th}}$, 2001

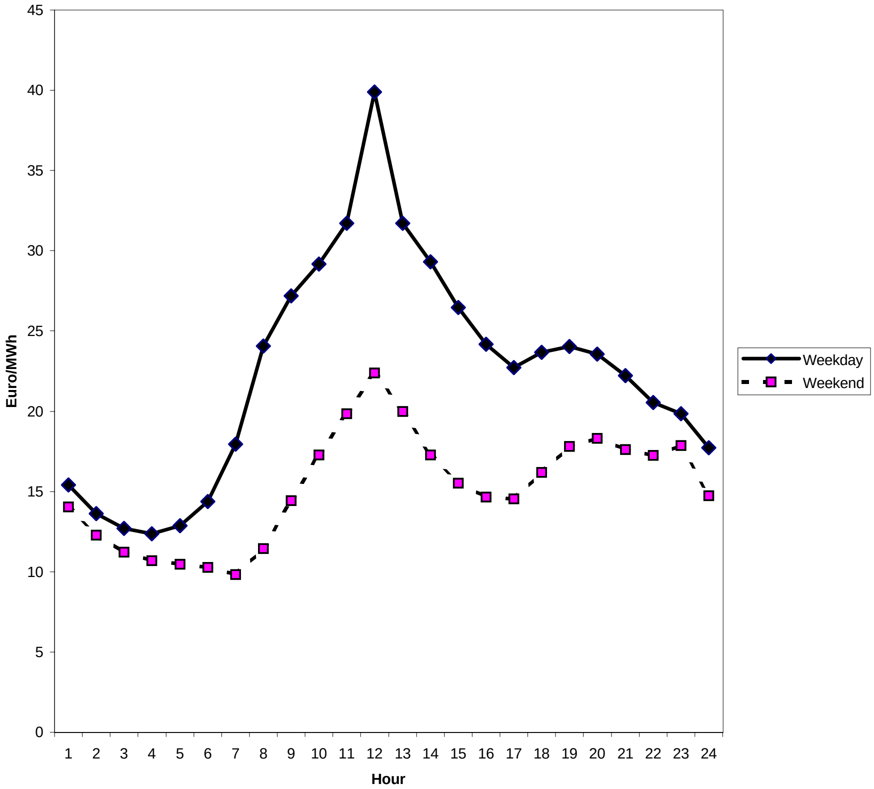

Figure 2: Average hourly LPX electricity spot prices across the entire sample

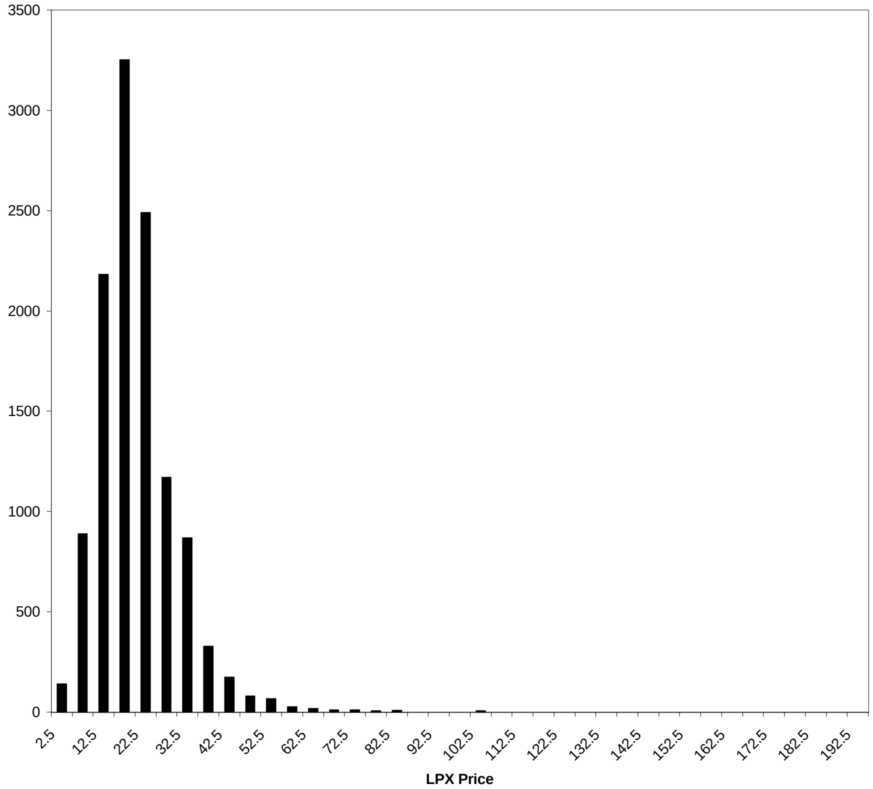

Figure 3: Histogram of LPX electricity spot price over the period June 16^{th}, 2000 - October 15^{th}, 2001 (Mean = 20.33, St. Dev. = 9.50, Skewness = 2.37, Kurtosis = 23.5)

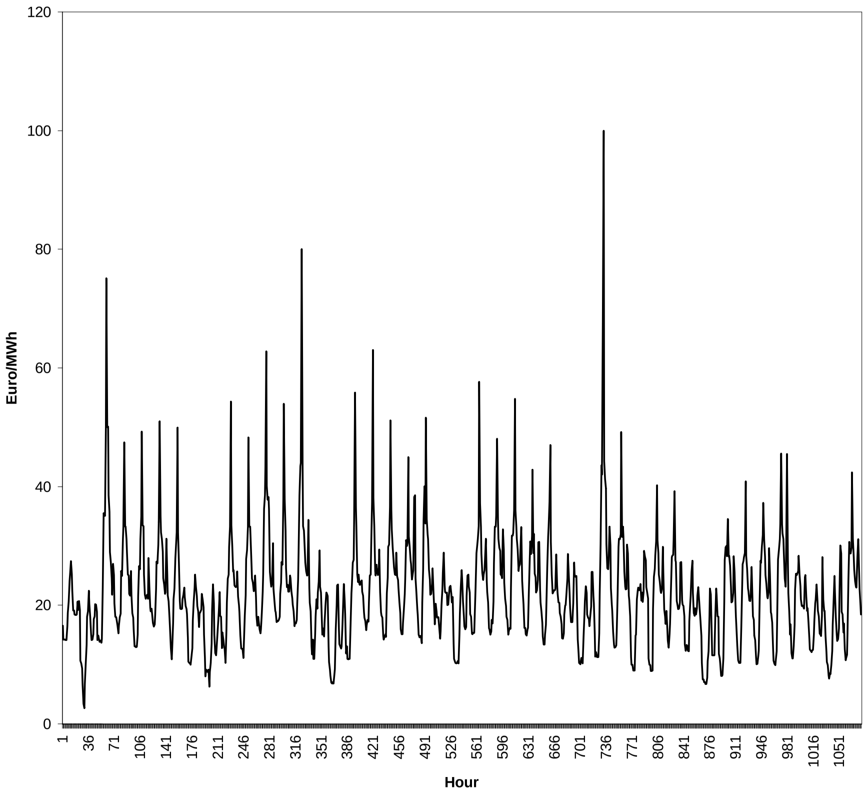

Figure 4: Out-of-sample period of LPX electricity spot price from September 1^{st}, 1:00, 2001 until October 15^{th}, 24:00 2001

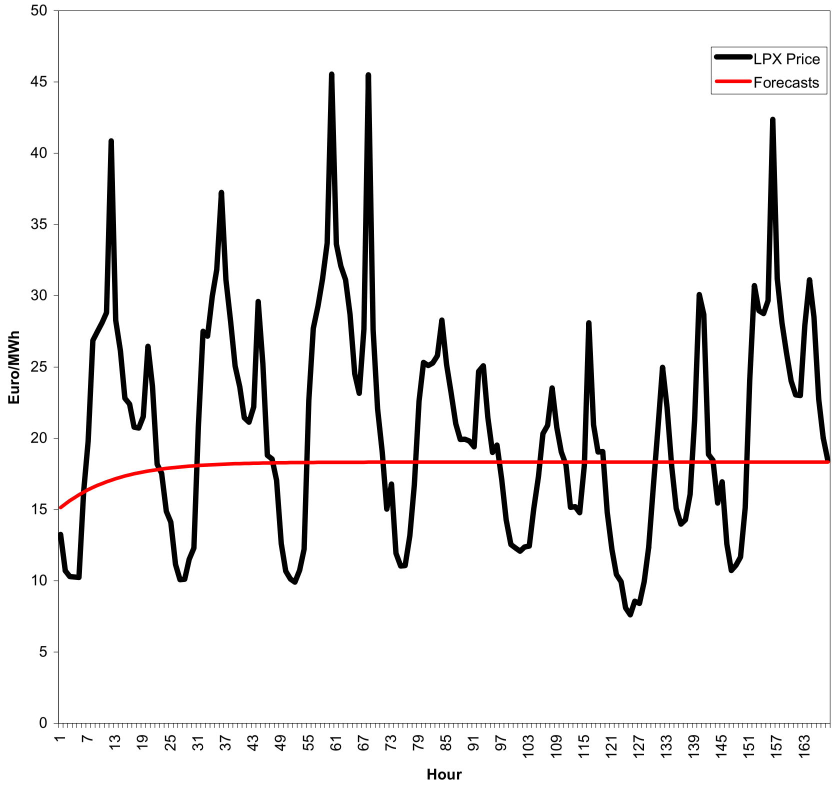

Figure 5: One week ahead forecasts for the period October 9$^{\text{th}}$, 1:00, 2001 - October 15$^{\text{th}}$, 24:00, 2001: Global AR(1) model with constant intercept

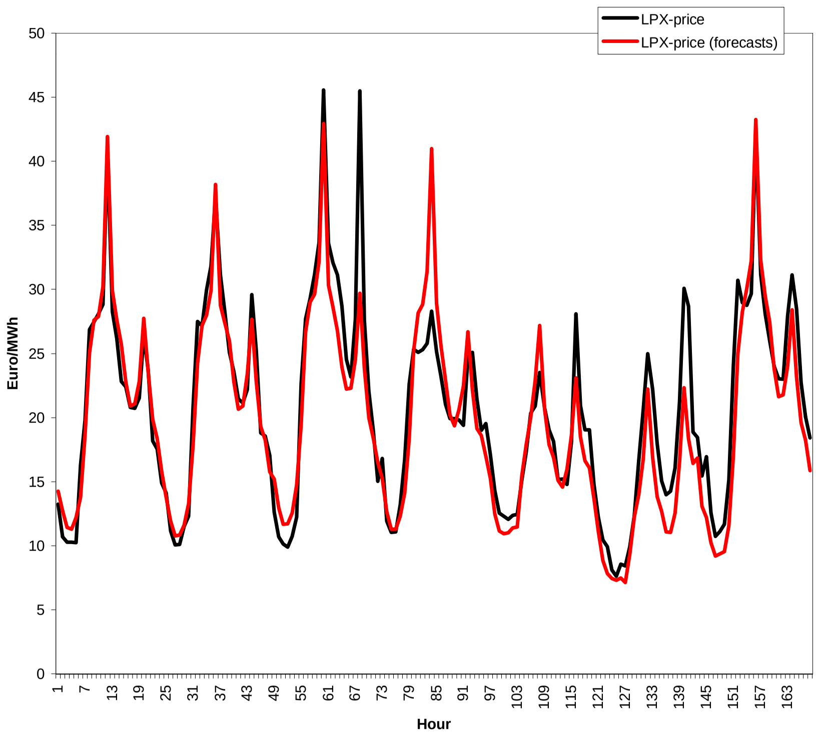

Figure 6: One week ahead forecasts for the period October 9th, 1:00, 2001 - October 15th, 24:00, 2001: Separable crossed ARMA model with jumps and restricted coefficients