MEAN-REVERSION PROCESSES TO MODEL HOURLY ELECTRICITY SPOT PRICES

Fernando L. Aiube
Department of Industrial Engineering, PUC-Rio
Rua Marquês de São Vicente, 225, Predio Leme sala 960 –22453-900, Rio de Janeiro, Brazil
faiube@puc-rio.br

Tara K. N. Baidya
Department of Industrial Engineering, PUC-Rio
Rua Marquês de São Vicente, 225, Predio Leme sala 960 –22453-900, Rio de Janeiro, Brazil
baidya@puc-rio.br

Frances F. Blank
Department of Industrial Engineering, PUC-Rio
Rua Marquês de São Vicente, 225, Predio Leme sala 960 –22453-900, Rio de Janeiro, Brazil
frances.blank@gmail.com

Wagner Saboia
Department of Industrial Engineering, PUC-Rio
Rua Marquês de São Vicente, 225, Predio Leme sala 960 –22453-900, Rio de Janeiro, Brazil
wagnersaboia@hotmail.com

Ana Beatriz Mattos
Department of Eletrical Engineering, PUC-Rio
Rua Marquês de São Vicente, 225, Predio Leme 4º andar –22453-900, Rio de Janeiro, Brazil
anabeatrizmattos@gmail.com

ABSTRACT

The process of liberalization and deregulation has resulted in great changes in European electricity markets. Competitive markets have been introduced in the last decades and the interest and the need for electricity price modeling became an important aspect of risk management in industry. The aim of this work is to compare characteristics of mean-reversion processes when modeling hourly spot electricity prices. We propose two approaches, modeling 24 different series, one for each hour of the day, or one entire hourly series in sequence, and we applied the models to Austria market. This paper is an extension of the work we developed for Stockholm University as a part of Energy Efficiency and Risk Management in Public Buildings project (EnRiMa).

KEYWORDS. Electricity. Prices. Spot market. Mean-reversion process. Austria.

EN – OR in Energy

1. Introduction

In the last few decades, the process of liberalization and deregulation has resulted in great changes in European electricity markets, reflecting directly in pricing behavior. The electricity industry used to be a regulated monopoly in most countries until early 1990s, when structural reforms started to take place worldwide. Competitive markets have been introduced and, consequently, the interest and the need for electricity price modeling became an important aspect of risk management in this industry. Wilson (2002) examines the economic issues that emerge in the context of liberalized power markets.

The Energy Efficiency and Risk Management in Public Buildings project (EnRiMa) aims to develop a decision-support system to enable operators to control energy flows in energy-efficient buildings and areas of public use. It is important to obtain and compare electricity prices forecasting models in Europe in order to facilitate the making of robust decisions under uncertainty. This paper is an extension work we developed for Stockholm University as part of EnRiMa project.

When working with hourly prices, there are two different approaches in the literature: (i) models based on one entire series with all available hourly data in sequence and (ii) models with 24 separate series, one for each hour of the day. According to Huisman et al. (2007), there is an important difference between modeling daily average prices and hourly prices. Hourly prices for next day delivery are determined at the same time and could not be treated as a pure time series process. On the other hand, there are some models in the literature using an entire 24-hourly series, where observations are taken in sequence. Our objective is to analyze mean-reverting processes to model hourly electricity prices, estimate its parameters, obtaining the results for both alternative approaches and compare their forecasting performance. The second approach will generate 24 models with different parameters, one for each hour of the day, compared to the first approach, which will produce only one model. The data used is from Austrian market.

Electricity is a special commodity because of its limited storability and transportability. These characteristics are reflected in electricity price time series, which present stylized facts that cannot be completely described by models based on other storable commodities or financial securities. Serati *et al.* (2007) and Härdle and Trück (2010) also describe the observed characteristics. The first feature is seasonality, which can be detected as three different types: annual, related to the seasons during the year and to the economic and social activities during different months; weekly, related to working days and weekends; and intraday cycle, related to variations among different hours of the day. Electricity prices also present mean-reverting characteristic. Many studies have been made to analyze reversion in different markets, which happens because weather is a dominant factor and influences equilibrium prices (Koopman *et al.*, 2007). Another observed fact is that the series present a very high volatility in both hourly and daily data, much higher than that for any other traditional commodities. The non-storable characteristic influences this behavior since electricity is an instantaneous consumption good. Price jumps and spikes are also observed and can be caused both by supply-side and demand-side shocks, as generation outages and occurrence of extreme loads. Market mechanism failure and network capacity constraints can also cause spikes in prices (Cuaresma *et al.*, 2004).

Another specific phenomenon in electricity markets may happen: negative and zero prices. Generally, negative prices occur only for a short period and mainly at night. It can happen, for instance, by the installation of combined-cycle facilities with non-flexible generators, causing an imbalance during night hours (Sewalt and de Jong, 2003). This situation becomes a problem when modeling prices and should receive proper treatment, e.g., just simply excluding these observations, or shifting prices to zero level or working with a transformation that deals with negative prices (Schneider, 2011). The presence of renewable sources and cogeneration also reduce the prices because of the effects of public support policies. This information would be important as explanatory variables to improve forecasting models (Gelabert *et al.*, 2011).

In this work, a mean-reverting diffusion process is used to model hourly electricity prices. The logarithm of prices is decomposed in a sum of two factors: a deterministic seasonal function and a stochastic part. The deterministic part is related to seasonality and modeled using dummy variables. To model the stochastic part, a mean-reverting process is proposed. Section 2 presents some models in the literature. Section 3 describes the methodology used. Sections 4 present electricity prices data and results for Austrian market. Section 5 provides the conclusions.

2. Models in literature

The literature on electricity price modeling is very rich. Different models have been proposed based on the observed stylized facts. According to Serati *et al.* (2007), considering the methodologies used in this field of research, three main classes can be built: autoregressive models; jump-diffusion and regime-switching models; and volatility models. The authors mention that, although many papers on this subject have been published, there is not one specific model supported by empirical evidence. The models differ in used data frequency, usually daily or hourly; in time-series transformation, usually logarithm of prices or log-returns; in the treatment of seasonality, that could be deterministic or stochastic; and in statistical approaches.

Financial models using continuous stochastic processes have also been used for this purpose, i.e., capturing the main characteristics of electricity prices such as mean reversion and spikes. Lucia and Schwartz (2002) express daily spot price and log spot price as a sum of two components: a deterministic part to model the seasonality and a stochastic part, for which they propose either a one-factor mean reverting stochastic process or a two-factor stochastic process combining a mean-reverting process and geometric Brownian motion to model the correlation between spot and future prices. Heydari and Siddiqui (2010) also propose the decomposition of log prices, modeling the stochastic part as different linear stochastic processes and non-linear stochastic models to account for the spikes. Huisman et al. (2007) work with separate hourly prices series, one for each hour of the day, and propose the same decomposition for all hours (in deterministic seasonal function and stochastic part, which is modeled as a mean-reverting process). They identify mean reversion parameters for each hour of the day.

Time-series traditional models are also frequent in literature, based on Box-Jenkins models. Bisaglia *et al.* (2010) decompose hourly prices into a deterministic part, to model the seasonality, and a stochastic part, described by autoregressive-GARCH and Markov-switching models. Weron and Misiorek (2005) study simple ARMA and ARMAX (including exogenous variable) models using log transformation for hourly prices. They also tried to introduce weekly seasonal dummies for a specific period to compare the models. Contreras *et al.* (2003) use SARIMA models and Garcia *et al.* (2005) use the AR-GARCH.

The interest in our work is in high-frequency data, i.e., hourly electricity prices. In the European Energy Exchange (EEX), for example, the spot price is an hourly contract with physical delivery, and every day is divided into 24 hourly contracts (Härdle and Trück, 2010). In the day-ahead markets, prices for all hours of the next day are determined at the same time.

Some authors model each hour time-series separately, while others treat as an entire time series in sequence. There are some arguments that hourly prices cannot be treated as a pure time series process because of the specific structure of day-ahead markets (Härdle and Trück, 2010; Huisman *et al.*, 2007). Huisman *et al.* (2007) propose a panel model, thereby resulting in one model for each hour of the day and a cross-sectional correlation matrix. On the other hand, many authors present models using an entire 24-hourly series, where observations are taken in sequence (Contreras *et al.*, 2003; Weron and Misiorek, 2005; Thomas and Mitchell, 2005). An advantage in this case is that it is possible to work with only one model for every hour in a day, and the correlation between the hours can be treated within the same model. Cuaresma *et al.* (2004) present models based on both approaches. Using Leipzig Power Exchange data, they obtained better forecasting properties when modeling each hour separately. Sewalt and de Jong (2003) mention that when modeling hourly prices, it is important to capture the interdependencies in different hours during the same day and between equivalent hours on different days.

Summarizing the applications presented in several papers, which analyse markets in US and Europe, the time span used to both in-sample period (used to estimate unknown parameters) and out-of-sample period (used to forecasting) varies a lot. There are some cases that less than a year is taken to estimate the parameters, because of different characteristics of crisis periods, for example. As for the forecast period, up to one week forecasts are normally chosen.

The contribution of this paper is to compare mean-reverting properties of electricity hourly prices under the two aforementioned different approaches, i.e., modeling the data as one entire series or 24 different series. The application will be using Austrian data.

3. Methodology

In order to obtain the parameters of the models the data set is split into two periods: an in-sample period, which is used to estimate the unknown parameters; and an out-of-sample period, which is used to assess the forecasts.

According to Heydari and Siddiqui (2010), the natural logarithms of spot prices are decomposed into two factors:

$$\ln(S_t) = f_t + X_t \qquad (1)$$

where
S_t – is the energy spot price observed, at time t;
X_t – is the stochastic part of log prices, at time t; and
f_t – is a deterministic seasonal function, at time t.

The right-hand side of Eq. (1) is composed of two terms. The first one, f_t, is related to seasonality observed in the series, which will be modeled using hourly, daily and monthly dummy variables to take into account intraday, weekly and annual seasonality. The coefficients are calculated using multiple regression. The model is represented by the following equation:

$$\begin{aligned} \ln(S_t) = (\alpha + \varepsilon_t) + \{(I_W D_W + I_S D_S) \\ + (I_1 D_1 + I_2 D_2 + I_4 D_4 + I_5 D_5 + \cdots + I_{12} D_{12}) \\ + (A_1 H_1 + A_2 H_2 + A_3 H_3 + A_5 H_5 + \cdots + A_{24} H_{24})\} \end{aligned} \qquad (2)$$

where
$\ln(S_t)$ – is the natural logarithm of the energy price;
α – is the independent coefficient of the regression;
D_i – are the daily dummy variables. Three types of days are defined for this purpose (i = W for Monday to Friday and S for Saturday; Sunday is used as reference);
I_i – are the linear coefficients of the regression for weekdays;
D_j – are the monthly dummy variables (j = 1, 2, 4 to 12; March is used as reference);
I_j – are the linear coefficients of the regression for the months of the year;
A_l – are the hourly dummy variables (l = 1, 2, 3, 5 to 24; hour 4 is used as reference);
H_l – are the linear coefficients of the regression for the hours of the day.

We execute the regression of Eq. (2) to estimate the coefficients. Afterwards, the seasonal effects are eliminated as follows:

$$\begin{aligned} X_t = \ln(S_t) - (\hat{I}_1 D_1 + \hat{I}_2 D_2 + \hat{I}_4 D_4 + \hat{I}_5 D_5 + \cdots + \hat{I}_{12} D_{12} + \hat{I}_W D_W + \hat{I}_S D_S \\ + \hat{A}_1 H_1 + \hat{A}_2 H_2 + \hat{A}_3 H_3 + \hat{A}_5 H_5 + \cdots + \hat{A}_{24} H_{24}) \end{aligned} \qquad (3)$$

When working with 24 separate series, one for each hour of the day, the decomposition presented in Eq.(1) can be written as

$$\ln(S_{h_t}) = X_{h_t} + f_{h_t} \tag{4}$$

where $h = 1 \text{ to } 24$ corresponds to each hour of the day
t corresponds to each observed date.

In this case, since each hour is modeled separately, two types of seasonality should be treated, using daily and monthly dummy variables to take into account weekly and annual seasonality. The coefficients are calculated using multiple regression and the model is represented by the following equation:

$$\ln(S_{h_t}) = (\alpha_h + \varepsilon_{h_t}) + \{(I_{Wh}D_W + I_{Sh}D_S) + \\ + \{(I_{1h}D_1 + I_{2h}D_2 + I_{4h}D_4 + I_{5h}D_5 + \cdots + I_{12h}D_{12})\} \tag{5}$$

The description of each variable and coefficient is analogous to the Eq. (2).

The second term in the right-hand side of Eq. (1) and Eq. (2) is the stochastic part of log electricity prices. Schwartz (1997) modeled log prices of commodities as an Ornstein-Uhlenbeck stochastic process as follows:

$$dX_t = \kappa(\alpha - X_t)dt + \sigma dz_t \tag{6}$$

where
X_t – is the natural logarithm of the commodity's price, at time t;
κ – is the magnitude of the speed of adjustment, which measures the degree of mean reversion to the long term log price;
α – is the long-term mean natural logarithm of the price;
σ – is the term of volatility of the process;
dz_t – is the increment of a standard Brownian motion.

It is worth mentioning that when modeling the 24 hours separately, there are different stochastic processes, one for each hour of the day. In this case, the diffusion processes are:

$$dX_{h_t} = \kappa_h(\alpha_h - X_{h_t})dt + \sigma_h dz_t \tag{7}$$

where $h = 1 \text{ to } 24$ and t corresponds to each observed date.

In this case, there are different parameters for different hours of the day, i.e., for each hour the long-term mean price and the speed of adjustment are different.

According to Dixit and Pindyck (1994), Eq. (6) is the continuous-time version of the first-order autoregressive process in discrete as $\Delta t \rightarrow 0$.

$$x_t - x_{t-1} = \alpha(1 - e^{-\kappa}) + (e^{-\kappa} - 1)x_{t-1} + \epsilon_t \tag{8}$$

where ϵ_t is normally distributed with mean zero and standard deviation σ_ϵ, and

$$\sigma_\epsilon^2 = \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa}) \tag{9}$$

Again, the parameters of Eq. (6) are estimated using the discrete time data available to run the regression of Eq. (10) for the entire series model and for 24 hours models:

$$x_t - x_{t-1} = a + bx_{t-1} + \epsilon_t \tag{10}$$

The parameters are then given by:

$$\hat{\alpha} = -\hat{a}/\hat{b} \tag{11}$$

$$\hat{\kappa} = -\ ln(1+\hat{b}) \tag{12}$$

$$\hat{\sigma} = \hat{\sigma}_{\varepsilon}\sqrt{\frac{2\ ln(1+\hat{b})}{(1+\hat{b})^2-1}} \tag{13}$$

where $\hat{\sigma}_{\varepsilon}$ is the standard error of the regression of Eq. (10).

When working with an entire series with hourly observations in sequence, only one model is estimated, i.e., only one $\hat{\alpha}$, one $\hat{\kappa}$ and one $\hat{\sigma}$. But when working with the hour series separately, 24 models are estimated, producing 24 $\hat{\alpha}_h$, 24 $\hat{\kappa}_h$ and 24 $\hat{\sigma}_h$.

From the fitted models, the price forecasts can be obtained for the out-of-sample period. The value of the stochastic process in Eq. (6) for a future date t, $X(t)$, conditional on the initial value $X(0)$, may be written as the stochastic integral (Bjerksund and Ekern, 1995):

$$X(t) = e^{-\kappa t}X(0) + (1-e^{-\kappa t})\alpha + \sigma e^{-\kappa t}\int_0^t e^{-\kappa u}dZ(u) \tag{14}$$

$X(t)$ is normally distributed, and its expected value and its variance are given by:

$$\mathbb{E}_0[X(t)] = e^{-\kappa t}X(0) + (1-e^{-\kappa t})\alpha \tag{15}$$

$$\mathbb{V}ar_0[X(t)] = \frac{\sigma^2}{2\kappa}(1-e^{-2\kappa t}) \tag{16}$$

As $X(t)$ is normally distributed, with mean and variance given by Eqs. (15) and (16), respectively, we can find forecasted price values using the following expression:

$$\mathbb{E}_0[S_t] = \mathbb{E}\left[e^{f_t}e^{X_t}\right] = e^{\{f_t+\mathbb{E}_0[X(t)]+\frac{1}{2}\mathbb{V}ar_0[X(t)]\}} \tag{17}$$

where e^{f_t} is deterministic and e^{X_t} is stochastic.

To verify the forecast performance, aggregate error measures are used: Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) for the out-of-sample period, defined by:

$$MAE = \frac{\sum_{t=1}^{N} |e_t|}{N} \tag{18}$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{N} e_t^2}{N}} \tag{19}$$

where

$e_t = S_t - \hat{S}_t$ is the forecast error, i.e., the difference between the actual price S_t and the forecasted price $\hat{S}_t$ at time *t*

N is the length of the out-of-sample period

The forecasted price $\hat{S}_t$ is a conditional expectation. We used two methodologies to analyze the forecast performance of the models. The first methodology, called here in-sample information, forecasts prices given all the data until the last in-sample observation. Given the last in-sample observation at time *t*, the forecasted price at time *t+s* is given by

$$\hat{S}_{t+s} = E[S_{t+s}\ |S_t\ ] \tag{20}$$

The second methodology, called here one-step ahead, calculates the forecasted price given the last most updated observation, i.e., each forecast within the out-of-sample period is given one step-ahead. The forecasted price at time t+s is given by

$$\hat{S}_{t+s} = E[S_{t+s} \,|S_{t+s-1}\,] \tag{21}$$

4. Austrian Market results

4.1. Electricity prices data

A total of 43,080 hourly observations over five years of electricity spot prices in €/MWh from Austrian market are available, provided by EXAA - Energy Exchange Austria. The sample period begins on January 1st, 2007 and ends on November 30th, 2011. The data set is split into two periods: in-sample (January 1st, 2007 to December 31st, 2010), which has 35,064 hourly observations; and out-of-sample (January 1st, 2011 to November 30th, 2011), which has 8,016 hourly observations. Considering the data set provided by EXAA, there are: (i) one entire series, with 43,080 hourly observations, for which the descriptive statistics are presented in Table 1; (ii) and 24 separate series for each hour of the day, with 1,795 daily observations, for which the descriptive statistics are presented in Table 2.

Table 1 – Summary of descriptive statistics for Austrian electricity prices (entire series)

Statistic	*Price*
Mean	48.20
Standard Deviation	23.92
Variance	572.32
Skewness	2.40
Kurtosis	25.67
Number of observations	43080

Table 2 – Summary of descriptive statistics for Austrian electricity prices (24 series)

Statistic	*Electricity price*											
	hr 1	*hr 2*	*hr 3*	*hr 4*	*hr 5*	*hr 6*	*hr 7*	*hr 8*	*hr 9*	*hr 10*	*hr 11*	*hr 12*
Mean	37.23	32.72	29.32	26.90	27.24	31.68	38.51	50.55	55.44	58.81	61.17	64.79
Std. Dev.	12.37	12.34	12.41	12.31	12.47	13.40	17.50	24.16	23.75	23.74	24.29	27.30

Statistic	*Electricity price*											
	hr 13	*hr 14*	*hr 15*	*hr 16*	*hr 17*	*hr 18*	*hr 19*	*hr 20*	*hr 21*	*hr 22*	*hr 23*	*hr 24*
Mean	60.41	56.93	53.72	51.17	51.39	57.86	61.72	59.02	54.09	48.56	47.18	40.43
Std. Dev.	22.45	21.86	21.38	20.47	21.63	34.35	33.82	23.64	18.25	14.72	13.52	12.10

Because of the outliers, a database treatment is performed in order to obtain reliable forecasts. There are 5 missing values, besides very low prices in the sample. For example, there are 184 values that are equal or less than 1€/MWh, which provide log prices less than zero. Those observations (in a total of 189, less than 0.5% of the database) are removed in order to proceed with the analysis. By doing this, we have an in-sample period with 34,890 valid hourly observations, in a total of 35,064 and an out-of-sample period with 8,001 valid hourly observations, in a total of 8,016. Figures 1(a) and 1(b) present the price and log-price time series.

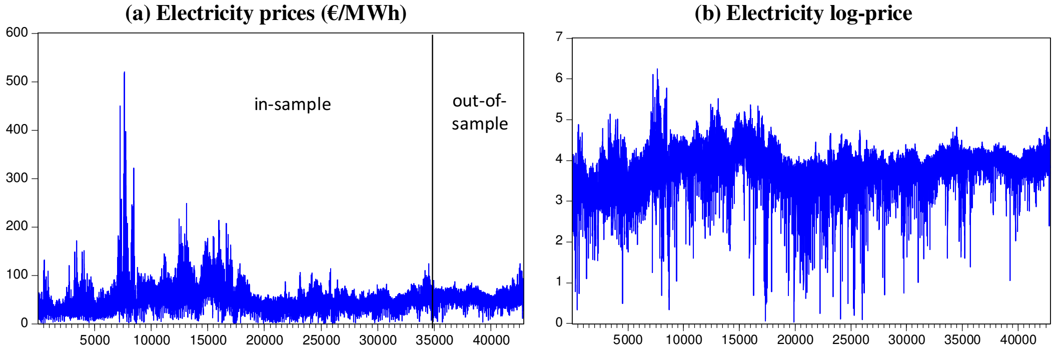

Fig. 1 - Austrian prices and log-prices series

Figure 2(a) presents the sample autocorrelation function for the electricity prices. We can observe the existence of spikes at lags equal to 24 and multiples revealing intraday seasonality. When taking a first and a 24-periodic differentiation for log prices, a weekly seasonality in lag 168 is also observed. Figure 2(b) presents the correspondent autocorrelation function after differentiations. The annual seasonality would be given in very high lags since the data are in high frequency (hourly data).

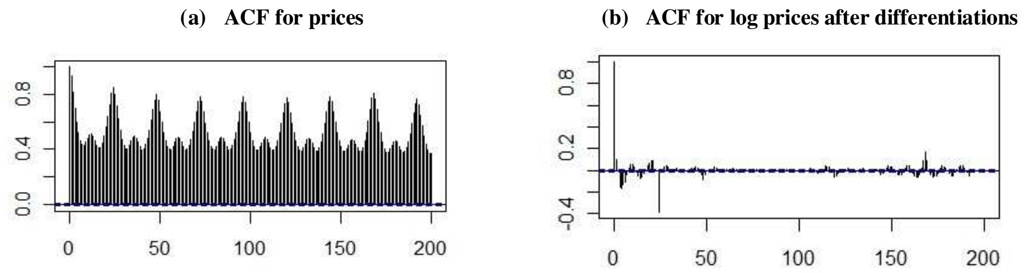

Fig. 2 – ACF for prices and log-prices after differentiation

4.2. Model for one entire hourly price series

In this case, the three types of seasonality should be modeled. Considering the dummy coefficients, the estimated parameters are shown in Table 3.

Table 3 - Results for entire hourly series regression

Parameter	*Estimated Value*	*Parameter*	*Estimated Value*
α	2.580	A_7	0.353
I_W	0.475	A_8	0.646
I_S	0.275	A_9	0.787
I_1	0.147	A_{10}	0.879
I_2	0.150	A_{11}	0.932
I_4	0.053	A_{12}	0.997
I_5	-0.002	A_{13}	0.929
I_6	0.107	A_{14}	0.861
I_7	0.114	A_{15}	0.792
I_8	0.037	A_{16}	0.737
I_9	0.239	A_{17}	0.735

I_{10}	0.407	A_{18}	0.832
I_{11}	0.285	A_{19}	0.905
I_{12}	0.180	A_{20}	0.877
A_1	0.399	A_{21}	0.799
A_2	0.243	A_{22}	0.697
A_3	0.119	A_{23}	0.673
A_5	0.018	A_{24}	0.500
A_6	0.184		

Higher values are detected for the months of September, October, November and December and lower values for the months of March, April and May. The hourly coefficients present an expected behavior based on the prices levels for peak hours (hours 9 to 20), which are higher, and off-peak hours.

After removing seasonal effects, a mean-reversion process should be estimated to capture the stochastic behavior. Table 4 displays the results of the regression of Eq. (10). Both parameters are significant at 5% confidence level. As the null hypothesis of $b = 0$ is rejected, the hypothesis of unit root (random walk) is also rejected, what reinforces the idea of mean-reversion in log hourly prices. The estimated values for mean-reversion parameters are given in Table 5.

Table 4 - Results of the regression of Eq. (10)

	Coefficients	*Std Error*	*t Stat*	*P Value*
$\hat{a}$	0.221	0.006	39.066	0.000
$\hat{b}$	-0.086	0.002	-39.534	0.000

Table 5 - Estimated mean-reversion parameters for hourly price series

$\hat{\alpha} = 2.580$	$\hat{\kappa} = 0.090$	$\hat{\sigma} = 0.170$

4.2. Model for 24 separate price series

In this case, since we have separate model for each hour of the day, only weekly and annual seasonality should be modeled. The coefficients for seasonality were obtained for the 24 different models. The main conclusions observed are:

- The parameter of the weekdays (I_W) is statistically different from that of Sundays for every hour. The parameter of Saturdays (I_s) is statistically different from that of Sundays in hours 1 to 19, including almost all peak hours, except hour 20. The parameter of the Saturdays (I_s) is not significantly different from that of Sundays only in hours 20 to 24. We decided to keep the same weekly seasonality treatment for all the hours, using the classification of weekdays, Saturdays and Sundays.
- Regarding the annual seasonality, the absolute and relative values of the monthly coefficients of the dummy variables vary over the hours. All monthly parameters are statistically different from zero for hours 11 to 15 and 20. In general, for peak hours, we observe parameters significantly different from zero. On the other hand, for off-peak hours, there is a large number of parameters that are statistically equal to zero. We kept the same treatment with all dummy variables for all the hours.

After removing seasonal effects, we propose a mean-reversion process to capture the stochastic behavior. Again, the hypothesis of unit root (random walk) is also rejected for the 24 models, reinforcing the idea of mean-reversion in log-hourly prices. The mean-reversion parameters are different for each one of the 24 models. The results differ a lot form hour to hour and are summarized in the following Table 6.

Table 6 – Estimated mean-reversion parameters

Parameters							
hour	α_h	κ_h	σ_h	*hour*	α_h	κ_h	σ_h
1	3.264	0.282	0.290	**13**	3.484	0.163	0.179
2	2.721	0.291	0.393	**14**	3.359	0.170	0.185
3	2.895	0.365	0.394	**15**	3.226	0.225	0.222
4	2.773	0.382	0.437	**16**	3.149	0.189	0.202
5	2.697	0.349	0.408	**17**	3.137	0.195	0.208
6	2.602	0.443	0.429	**18**	3.309	0.185	0.212
7	2.325	0.546	0.476	**19**	3.621	0.158	0.192
8	2.627	0.451	0.417	**20**	3.755	0.132	0.116
9	3.041	0.336	0.302	**21**	3.677	0.099	0.136
10	3.305	0.203	0.210	**22**	3.583	0.108	0.138
11	3.424	0.173	0.187	**23**	3.628	0.071	0.111
12	3.482	0.176	0.198	**24**	3.420	0.106	0.145

Electricity prices revert more rapidly to its long term average for hours in the morning, followed by afternoon hours and night hours. The different long term average for each hour indicates the different levels of prices over the day. Huisman, Huurman and Mahieu (2007) also observe that mean-reversion is not stable over the day and super-peak hours exhibit less mean-reversion. We observe this fact when comparing peak hours to off-peak hours in the morning, but not when comparing to off-peak hours late at night. Actually, when considering the classification of peak hours (9-20) and off-peak hours, we observe that off-peak hours early in the morning are pretty much different from the rest of the day. Off-peak hours late at night seem to behave similarly to peak hours.

There are some advantages and disadvantages to work with 24 different models. First, since the mean-reverting parameters are different, separate models for each hour are expected to better describe the price behavior. On the other hand, it becomes much more complicated to manage 24 different models instead of only one if using the same model for the entire hourly series. Thus, the models can be compared under forecasting performance.

4.3. Forecasting results

The results of the aggregate error measures for the forecasted series, calculated based on both methodologies proposed (in-sample information and one-step ahead) are displayed in Table 7. In order to compare the performance of the entire series model and the 24 separate models, the results are presented for both cases.

Table 7 – Aggregate error measures for forecasted prices conditional to in-sample information

	In-sample information				*One-step ahead*			
	1 entire series		*24 separate series*		*1 entire series*		*24 separate series*	
hour	MAE	RMSE	*MAE*	*RMSE*	MAE	RMSE	*MAE*	*RMSE*
1	10.56	12.71	9.96	11.47	8.68	13.39	4.18	5.55
2	11.85	13.80	13.12	15.23	1.64	2.23	7.21	9.75
3	12.59	14.40	12.08	13.81	1.58	2.31	5.63	7.22
4	13.77	15.52	12.72	14.52	1.29	1.80	6.26	8.00
5	13.57	15.34	12.76	14.56	3.51	3.88	6.23	8.08
6	12.11	13.86	11.23	13.07	7.49	8.19	7.12	9.76
7	13.41	15.16	11.51	13.46	5.99	7.23	8.85	12.97

8	11.26	13.42	10.48	12.85	12.66	13.51	10.21	15.58
9	8.98	11.29	9.52	11.82	10.96	11.77	7.89	11.19
10	7.60	9.78	8.35	10.57	10.44	11.23	6.55	9.14
11	8.00	10.10	8.24	10.34	8.16	8.70	6.32	8.56
12	9.68	12.28	9.63	12.15	8.73	9.16	7.17	9.64
13	8.45	10.82	7.89	9.77	2.36	2.73	6.17	8.17
14	7.57	9.87	7.67	9.63	3.47	4.01	6.08	8.05
15	7.35	9.55	7.78	9.69	3.05	3.42	6.03	8.02
16	7.20	9.27	7.89	9.86	3.04	3.32	5.98	8.07
17	7.27	9.31	8.19	10.20	4.55	5.09	6.21	8.32
18	8.25	10.74	9.40	11.65	6.69	7.56	6.30	8.33
19	9.55	12.17	9.84	11.92	5.32	6.14	5.80	7.64
20	8.68	11.09	10.29	12.45	5.01	8.11	5.43	7.11
21	9.16	11.76	10.26	12.07	4.61	6.16	4.68	6.24
22	9.98	12.58	9.46	11.10	4.11	5.55	3.59	4.65
23	10.24	12.86	9.05	10.58	4.61	5.39	3.13	4.16
24	10.22	12.44	9.92	11.37	1.56	1.98	2.85	3.84

The aggregate measures show that for some hours one method can be preferred compared to the other, considering both forecast methodologies (in-sample information forecast and the one-step ahead forecast, as defined previously). Although the mean-reversion parameters vary from one hour to another, the use of 24 different models for each hour of the day did not provide real gains in terms of forecasting performance in this case. Using only one model for the entire hourly series is easier and less expensive to implement.

5. Conclusions

Some different approaches in the literature are proposed to model electricity prices. In fact, the suitability of models depends on the nature of the markets as well as the scope of the underlying decision making problem. High frequency data, as hourly prices used in this paper, are subject to be modeled under two methodologies, the first based on one entire series with all available data in sequence and the second on 24 separate series, one for each hour of the day.

Our objective is to analyze mean-reverting processes to model hourly electricity prices, estimate its parameters, obtaining the results for both alternative approaches and compare their forecasting performance. The methodology was applied to Austrian market data.

The first approach generates only one model and the hourly data is taken in sequence. The mean reversion parameters, i.e., the long-term log-price mean, the speed of adjustment and the volatility, reflect the behavior of all hours of the day when analyzed in sequence. The second approach produces 24 different models, one for each hour of the day. In this case, the results show that the parameters differ a lot from hour to hour. Electricity prices revert more rapidly to its long term average for hours in the morning, followed by afternoon hours and night hours. The different long term average for each hour indicates the different levels of prices over the day. Comparing peak hours to off-peak hours in the morning, super-peak hours exhibit less mean-reversion, but this does not happen when compared to off-peak hours late at night. Actually, off-peak hours early in the morning are very different from the rest of the day. Off-peak hours late at night seem to behave similarly to peak hours.

There are some advantages and disadvantages to work with 24 different models. First, since the mean-reverting parameters are different, separate models for each hour are expected to better describe the price behavior. On the other hand, it becomes much more complicated to manage 24 different models instead of only one if using the same model for the entire hourly series. Comparing the models using forecast aggregate measures to Austrian market data, given the in-sample period from 2007 to 2010 and the out-of-sample period of 2011, the results from both proposed approached are similar. The use of 24 different models for each hour of the day is not an advantage because this approach does not provide a better forecasting performance in this case. The alternative of using the entire series model approach is more recommended.

Acknowledgments

This paper is an extension work developed for Stockholm University as part the project "Energy Efficiency and Risk Management in Public Buildings" (EnRiMa). We would like to thank Stockholm University and the European Comission (via the 7th Framework Programme FP7, project number 260041) for funding the original work. This work is extended under the CNPq research grants, to whom we are grateful.

References

Bisaglia, L., Bordignon, S. and Marzovilli, M. (2010) Modelling and forecasting hourly spot electricity prices: some preliminary results. Working Paper Series, N.11. Department of Statistical Sciences, University of Parma.
Bjerksund, P. and Ekern, S. (1995) Contingent claims evaluation of mean-reverting cash flows in shipping , in Trigeorgis, L., Ed.: Real Options In Capital Investment: Models, Strategies, and Applications, Preager.
Contreras, J., Espínola, R., Nogales,F.J. and Conejo, A.J. (2003) ARIMA models to predict next-day electricity prices. IEEE Transactions on Power Systems, 18, 3, 1014-1020.
Cuaresma, J.C., Hlouskpva, J., Kossmeier, S. and Obersteiner, M. (2004) Forecasting electricity spot-prices using linear univariate time-series models. Applied Energy, 77, 1, 87-106.
Dixit, A. and Pindyck, R. (1994) Investment under uncertainty. Princeton Univ. Press, 468 p.
Garcia, R.C., Contreras, J., Van Akkeren, M. and Garcia, J.B.C. (2005) A GARCH forecasting model to predict day-ahead electricity prices. IEEE Transactions on Power Systems, 20, 2, 867-874.
Gelabert, L., Labandeira, X. and Linares, P. (2011) An ex-post analysis of the effect of renewables and cogeneration on Spanish electricity prices. Energy Economics, 33, S1, S59-S65.
Härdle, W.K. and Trück, S. (2010) The dynamics of hourly electricity prices. SFB 649 Economic Risk. Discussion Paper, 2010-13, Humboldt-Universität zu Berlin.
Heydari, S. and Siddiqui, A. (2010) Valuing a gas-fired power plant: a comparison of ordinary linear models, regime-switching approaches, and models with stochastic volatility. Energy Economics, 32, 709-725.
Huisman, R., Huurman, C. and Mahieu, R. (2007) Hourly Electricity Prices in Day-Ahead Markets, ERIM Reports Series Research in Management.
Koopman, S.J., Ooms, M. and Carnero, M.A. (2007) Periodic Seasonal Reg-ARFIMA-GARCH models for daily electricity spot prices. Journal of the American Statistical Association, 102, 477, 16-27.
Lucia J.J. and Schwartz, E.S. (2002) Electricity prices and power derivatives: evidence from the Nordic Power Exchange. Review of Derivatives Research, 5, 5-50.
Schneider, S. (2011) Power spot price models with negative prices. Journal of Energy Markets, 4, 4.
Serati, M., Manera, M. and Plotegher, M. (2007) Modelling electricity prices: from the state of the art to a draft of a new proposal. Liuc Papers n. 210, Serie Economia e Impresa, 56.
Sewalt, M. and De Jong, C. (2003) Negative prices in electricity markets. Commodities Now, June 2003, 74-77.
Schwartz, E. (1997) The stochastic behaviour of commodity prices: Implications for valuation and hedging, The Journal of Finance, 52, 3.
Thomas, S. and Mitchell, H. (2005) GARCH modeling of high-frequency volatility in Australia's National Electricity Market. System, 1-39.
Weron, R. and Misiorek, A. (2005) Forecasting spot electricity prices with time series models. Proceedings of International Conference The European Electricity Market EEM-05, 133-141.
Wilson, R. (2002) Architecture of power markets. Econometrica, 70, 4, 1299-1340.