

# Electricity Pool Prices: Long-Term Uncertainty Characterization for Futures-Market Trading and Risk Management \*

Antonio J. Conejo

Department of Electrical Engineering

University of Castilla – La Mancha

Ciudad Real, Spain

<http://www.uclm.es/area/gsee>

[Antonio.Conejo@uclm.es](mailto:Antonio.Conejo@uclm.es)

Francisco J. Nogales

Department of Statistics

Universidad Carlos III de Madrid

Madrid, Spain

<http://www.est.uc3m.es/Nogales>

[FcoJavier.Nogales@uc3m.es](mailto:FcoJavier.Nogales@uc3m.es)

Miguel Carrión

Department of Electrical Engineering

University of Castilla – La Mancha

Toledo, Spain

<http://www.uclm.es/area/gsee>

[Miguel.Carrion@uclm.es](mailto:Miguel.Carrion@uclm.es)

Juan M. Morales

Department of Electrical Engineering

University of Castilla – La Mancha

Ciudad Real, Spain

<http://www.uclm.es/area/gsee>

[JuanMiguel.Morales@uclm.es](mailto:JuanMiguel.Morales@uclm.es)

---

\*We are grateful to the Statistical Department of EEX for kindly and efficiently providing us with all the required data. A. J. Conejo, M. Carrión and J. M. Morales are partly supported by Junta de Comunidades de Castilla - La Mancha through project PCI-08-0102 and by the Ministry of Education of Spain through CICYT Project DPI2006-08001. F. J. Nogales is partly supported by the Comunidad de Madrid through project CCG07-UC3M/ESP-3389 and by the Ministry of Education of Spain through CICYT Project MTM2007-63140.

**Abstract.** Organized trading for electricity includes both the pool and the futures market. Pool prices are volatile while the prices of the futures-market products are comparatively more stable. Thus, futures-market products constitute hedging instruments to reduce the risk suffered by any market agent. Electricity market agents engage in both pool and futures market transactions seeking to maximize their respective profits/utilities for a given risk level on profit variability. To make informed decisions, the market agent must gather as much accurate information as possible on the pool prices covering the whole time horizon spanned by the futures-market product. This paper provides a novel technique to represent conveniently the uncertainty associated with pool prices during long- or medium-term horizons through a set of scenarios, i.e., pool price realizations. The proposed technique uses the prices of the futures-market products as long-term explanatory variables and exploits the short-term structure of the pool prices.

**Keywords:** Electricity pool prices, year-ahead forecasting, forward trading, futures prices, scenarios.

## 1 Introduction

Organized trading for electricity includes both the pool and the futures market. The pool allows short-term transactions, e.g., buying/selling electric energy on an hourly basis one-day ahead. On the other hand, the futures market allows energy transactions spanning from one week to several years and involving a specified amount of energy to be delivered somewhere in the future, e.g., two-hundred megawatts to be supplied throughout the next quarter.

During the last two decades pool-based electricity markets have flourished in many places all over the world. Nowadays, they constitute generally mature trading floors. The pioneering experience in Europe corresponds to the England and Wales electricity pool created in the early 1990s. In North America, California established in 1998 the first pool-based electricity market. The electricity markets in New Zealand and Australia started to operate in 1996 and 1998, respectively. Currently, China is considering different market alternatives to manage efficiently its vast electric energy systems.

On the other hand, futures markets were established generally later and they have not

been yet fully developed. Futures markets in Europe include the pioneering Nordpool market in Scandinavia, EEX in Germany, PowerNext in France, and the recently established OMIP in the Iberian Peninsula. Nordpool, EEX, PowerNext and OMIP were launched in 1993, 2001, 2004 and 2006, respectively. Among others, two major electricity markets in North America are New England ISO mostly covering New England, and PJM Interconnection mostly covering Pennsylvania, New Jersey and Maryland. The financial electricity products of these markets are traded at the New York Mercantile Exchange, NYMEX. Trading for PJM and New England ISO in NYMEX started in 2003 and 2004, respectively.

Pool prices are generally highly volatile while the prices of the futures-market products are comparatively more stable. Thus, futures-market products constitute hedging instruments to reduce the risk on profit/utility variability suffered by any market agent. Needless to say, the prices of the futures-market products are usually less competitive than pool prices. Hence, a risky agent opts mostly for the pool while a conservative one opts mostly for the futures market.

Electricity market agents engage in both pool and futures market transactions seeking to maximize their respective profits/utilities for a given risk level on profit/utility variability. To make informed decisions on the futures market pertaining to any product (e.g., a monthly forward contract), the market agent must gather as much accurate information as possible on the pool prices covering the whole time horizon spanned by the futures-market product. Thus, characterizing the future behavior of electricity pool prices several months or years in advance is important for electricity trading and risk management (see, for instance, Carrión et al. (2007a), Carrión et al. (2007b) and Conejo et al. (2008)).

Specifically, this paper provides a novel technique to represent conveniently the uncertainty associated with pool prices during long- or medium-term horizons (e.g., one year) through a set of scenarios (pool price realizations). The proposed technique uses the prices of the futures-market products as long-term explanatory variables and exploits the short-term structure of the pool prices. Prices of futures-market products and pool prices are clearly dependent, as a result of the unique fact that electricity is a non-storable commodity. The technique proposed in this paper materializes in a set of time series models, which are easy to understand and simple to use. To the best of our knowledge, no pre-

vious attempt to generate long-term electricity-price scenarios based on futures markets information has been reported in the technical literature.

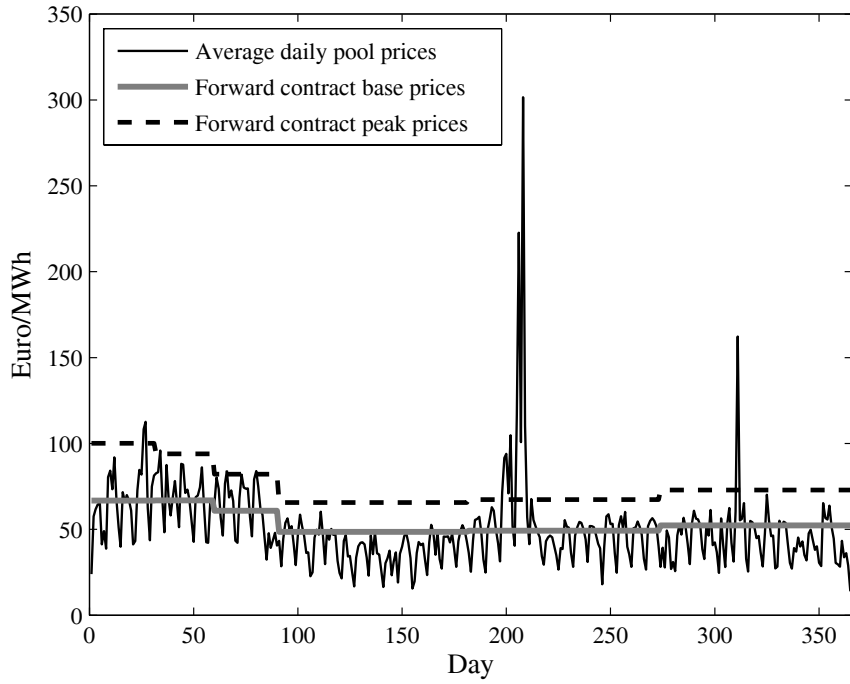
Besides electricity non-storability, electricity markets are incomplete and not sufficiently liquid, hence, pool and futures prices present a complex dynamic structure that other commodities such as oil, gas, etc. do not present in general. A careful analysis of such structure provides the relevant information needed to build appropriate models to generate accurate year-ahead pool price scenarios. If electricity markets were liquid and complete, then, futures and pool prices would have a simple dynamic structure similar to that of other financial assets. But these circumstances do not hold in electricity markets, implying the existence of specific patterns (over short- and long-term horizons) inexistent in markets for other commodities. These patterns can be exploited to generate scenarios that efficiently describe future pool price realizations.

Throughout this paper, the term base refers to all the 24 hours in a day and the term peak refers to the hours in a day with the highest pool-price values, i.e., the 12 hours from 8 am to 8 pm. Off-peak hours are the remaining 12 hours.

We consider below the so-called forward base/peak curve that allows relating prices of futures products and pool prices. A forward base/peak curve is constructed at forecasting time (day  $d$ ) spanning the study horizon (e.g., one year) and including the following prices:

1. day  $d$  prices for all available one-month futures-market products (e.g., three months),
2. day  $d$  prices for three-month futures-market products for the remaining periods to complete the planning horizon (e.g., three quarter futures-market products to complete one year).

As an example, Figure 1 shows for the EEX market the forward base/peak curves computed on December 30, 2005 and spanning the whole 2006. This figure shows also the average daily pool prices throughout 2006. Besides some outliers, it can be observed that the forward curves provide appropriate indicators of the long-term behavior of electricity pool prices. This observation constitutes the rationale of the approach used in this paper: to exploit the patterns observed in the differences between the forward curves and the associated pool prices. Note that these differences are known as forward premia.



**Figure 1: EEX market: Forward base and peak curves (computed at December 30, 2005) and average daily pool prices throughout 2006.**

Hence, the proposed methodology is based on combining the information provided by the long-term properties of the forward premium and the information provided by the short-term properties of the pool price. However, note that the objective of this paper is not to evaluate the financial properties of the forward premium. These properties are studied in Longstaff and Ashley (2004) and Karakatsani and Bunn (2003), where forward premium is found to vary systematically in the short-term. On the contrary, in this paper we exploit existing systematic patterns in the year-ahead forward premium to generate long-term price scenarios.

References on long-term electricity price forecasting include Bunn (2006), which provides background on the price formation in electricity markets, and Davison et al. (2002), which proposes a stochastic electricity price model that is adequate for pricing options in financial markets. References on long-term commodity prices are unfortunately scarce. Relevant ones include Schwartz (1997) and Schwartz & Smith (2000). In these papers the authors propose several models for the stochastic behavior of commodity prices, with emphasis on the valuation of financial and real assets.

In conclusion, the main contribution of this paper is to provide a technique to generate base/peak year-ahead pool-price scenarios with the help of information pertaining to base/peak futures-market products. Off-peak pool-price scenarios are subsequently obtained. Results of a real-world case study pertaining to the Germany-based EEX electricity market (a sufficiently liquid market), which illustrate the relevance and accuracy of the technique proposed, are provided. Needless to say, the proposed methodology can be adapted to markets other than EEX.

The rest of this paper is organized as follows. Section 2 describes in detail the models proposed and the scenario-generation procedure. Section 3 provides, analyzes and discusses results pertaining to a real-world case study based on the EEX electricity market. Section 4 draws some relevant conclusions. Finally, and for the sake of clarity, an appendix provides the notation used throughout the paper.

## **2 Long-Term Uncertainty Characterization of Pool Prices**

### **2.1 Time Discretization**

We describe the pool electricity prices throughout the year using 48 base values, 48 peak values and 48 off-peak values. Independent models for base and peak prices, respectively, are developed. Off-peak prices are then obtained from base and peak forecasts.

The 48 base price values correspond to 12 months, each one including 1 Monday base period, 1 working day (other than Monday) base period, 1 Saturday base period and 1 Sunday base period. The Monday base period is the average value of all Monday base hourly prices throughout the month. The Saturday/Sunday base period is defined similarly as the Monday base period. The working (other than Monday) base period includes all base hourly prices pertaining to working days (other than Mondays) throughout the month. Peak and off-peak periods are defined similarly as base periods.

It should be noted that the above discretization captures pool price variations during the whole year and leads to yearly price scenarios of a tractable dimension, for both uncertainty characterization of pool prices and long-term decision making through stochastic

programming models, Birge & Louveaux (1997). Indeed, the discretization above constitutes an appropriate tradeoff since it provides sufficient accuracy for moderate computational effort. Moreover, the above discretization is based on actual practice of producers in the Iberian Electricity Market, OMIP.

Needless to say, the considered discretization should be tailored to the specific circumstances of the electricity market under consideration and the 48-value discretization changed as considered appropriate.

## 2.2 Forward Curves

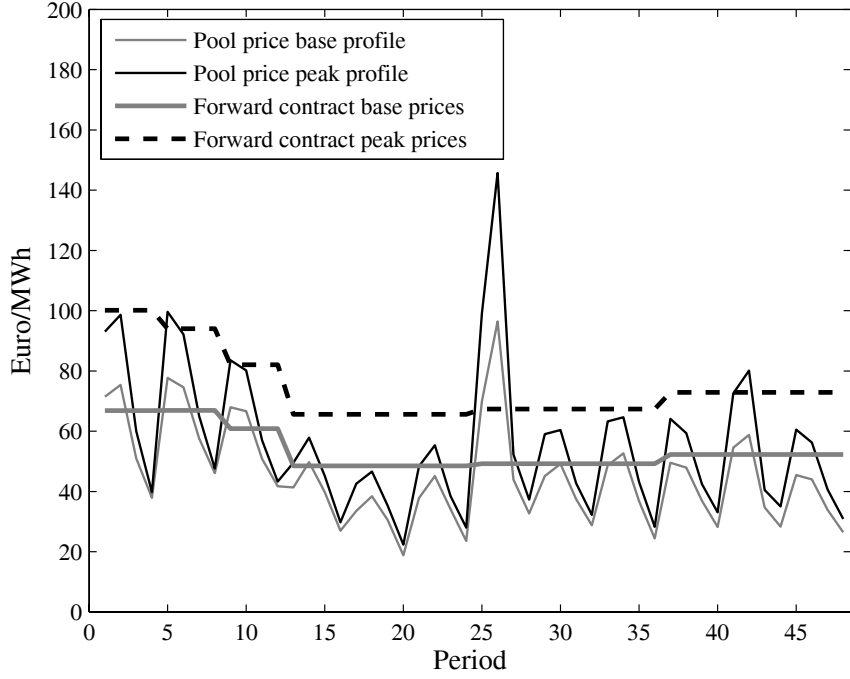
Forward base and peak curves spanning the considered yearly horizon are built as follows:

1. The forward base/peak curve spans one year and includes 48 values according with the time discretization established in Section 2.1 above. The forward base/peak curve is computed at the decision-making (forecasting) day, day  $d$ .
2. Prices for the first three months (months 1, 2 and 3) are the futures prices at day  $d$  of the corresponding three monthly futures products.
3. Prices for months 4-6, 7-9 and 10-12 are the futures prices at day  $d$  of the three futures three-month (quarters) products that correspond to months 4-6, 7-9 and 10-12, respectively.

Figure 2 includes plots of the peak and base forward curves for EEX (computed at December 30, 2005), the base prices (48 values) and the peak prices (48 values) for year 2006.

The above three-step procedure illustrates the way in which forward curves have to be built. It should be emphasized that this procedure should be tailored to the specific financial market under study.

Moreover, note that unlike traditional financial markets in which the forward price represents the discounted expected future value of the pool price, pool and forward electricity prices do not comply with this relationship because electricity markets are not sufficiently liquid and complete.



**Figure 2: EEEX market: Forward base/peak curves, computed at December 30, 2005 (48 values), and base/peak pool prices (48 values) throughout 2006.**

## 2.3 Basic Models

Based on the considerations provided in the previous Sections 2.1 and 2.2, the following basic models for the pool price are initially considered for base and peak periods, respectively:

$$\log p_t^{\text{Base}} = \log f_{t_0}^{\text{Base}}(t) + \varepsilon_{t,t_0}^{\text{Base}} , \quad (1)$$

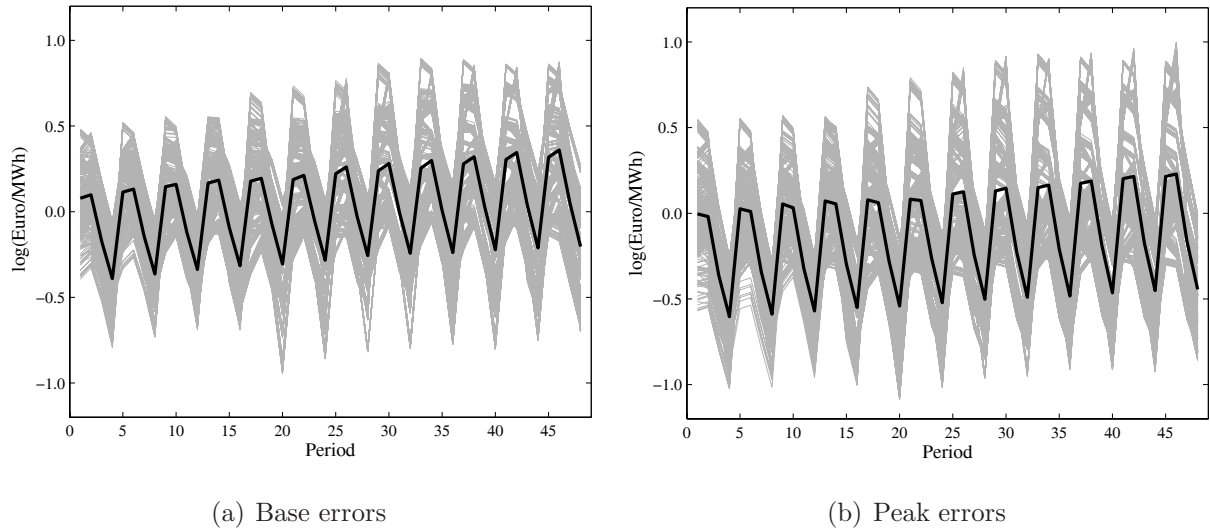
$$\log p_t^{\text{Peak}} = \log f_{t_0}^{\text{Peak}}(t) + \varepsilon_{t,t_0}^{\text{Peak}} . \quad (2)$$

Note that equations (1)/(2) relate base/peak prices ( $p_t^{\text{Base}}/p_t^{\text{Peak}}$ ) with base/peak forward curves ( $f_{t_0}^{\text{Base}}(t)/f_{t_0}^{\text{Peak}}(t)$ ) and base/peak forecasting errors ( $\varepsilon_{t,t_0}^{\text{Base}}/\varepsilon_{t,t_0}^{\text{Peak}}$ ). The motivation to take logarithms is to stabilize the variance. Because of this transformation, the direct un-transformation of the forecasts is biased. Thus, we use the method provided in Guerrero (1993) to adjust for bias when un-transforming forecasts and confidence intervals. Further details are given in Section 2.7.

For each trading day in the EEEX market, throughout 2004-2005, errors for models (1) and (2) are depicted in Figure 3. In order to clearly visualize error patterns, the base/peak



48-error values of each trading day are plotted in a superimposed manner, together with the error means of each of the time periods.



**Figure 3: EEX Market: Base and peak errors throughout the trading days of 2004-2005 for basic models (1) and (2), respectively.**

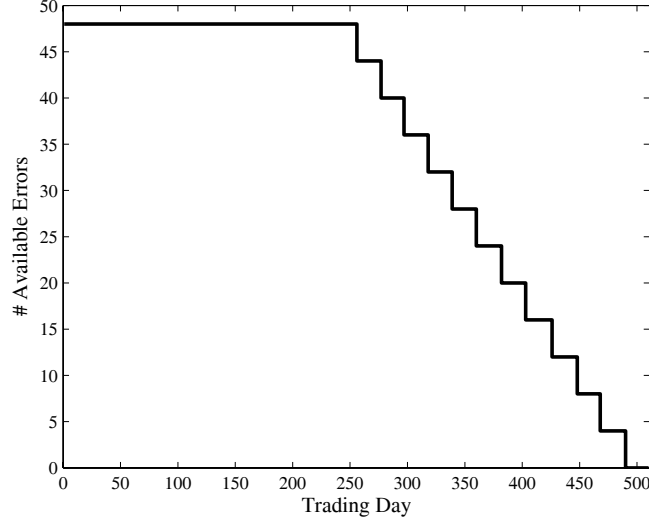
Figure 3 reveals a significant pattern in the time evolution of the forward premium within a long term horizon, i.e., this premium increases with time.

## 2.4 Features of the Basic Models

First of all, note that not all the errors in (1) and (2) are known at the forecasting day. For instance, if the forecasting day is December 30, 2005, since pool prices for 2005 are available but pool prices for 2006 are not, then all the errors terms in (1) and (2) are available for each trading day in 2004. However, for the first trading day in 2005, the first 44 errors are available (the first 44 pool prices in equations in (1) and (2) belong to 2005) but the last 4 errors are not available (the last 4 pool prices in equations in (1) and (2) belong to 2006). Thus, the latter the trading day in 2005, the higher the number of non-available errors.

As an example, Figure 4 shows the pattern of available errors at forecasting day December 30, 2005. Specifically, each column represents the available errors (up to 48) for the base/peak models obtained for each trading day from the first trading day in 2004 (January 2, 2004) to the last trading day in 2005 (December 30, 2005).

These non-available errors need to be treated as missing data to be estimated through the



**Figure 4: EEX Market: Pattern of available errors (computed at December 30, 2005) corresponding to each trading day from January 2, 2004 to December 30, 2005.**

proposed time series models. In Section 2.6, we propose a recursive procedure to estimate these non-available errors.

Finally, to reproduce the special patterns shown in Figure 3, the cyclic terms in the errors of (1) and (2) are treated with seasonal differences. That is, the first 48 error terms in (1) and (2) corresponds to the 48 observed error terms at the first trading day in 2004. Then, the second 48 error terms in (1) and (2) corresponds to the 48 observed error terms at the second trading day in 2004, and so on. Then, we differentiate these data with seasonality differences of orders 4 and 48, respectively. Note that the 4-order difference is motivated by the 4 blocks considered in a month: Monday, working day (other than Monday), Saturday and Sunday. The 48-order difference is motivated by the forecasting horizon: 48-period ahead.

## 2.5 Detailed Models for Base and Peak Errors

Considering the features exhibited by the error plots of Figure 3, and the standard autocorrelation and partial autocorrelation functions of the final time series involved, the two models below are suggested.

The proposed base-error model has the form:

$$\begin{aligned} & (1 - \phi_1^{\text{Base}} B - \phi_2^{\text{Base}} B^2) (1 - \phi_4^{\text{Base}} B^4) (1 - \phi_{48}^{\text{Base}} B^{48}) (1 - B^4) (1 - B^{48}) \varepsilon_{t,t_0}^{\text{Base}} \\ &= (1 - \theta_1^{\text{Base}} B) (1 - \theta_4^{\text{Base}} B^4) (1 - \theta_{48}^{\text{Base}} B^{48}) a_{t,t_0}^{\text{Base}}, \end{aligned} \quad (3)$$

where  $a_{t,t_0}^{\text{Base}}$  is the innovation term, and, as usual, it is assumed to be white noise.

The proposed peak-error model has the alternative and simpler form:

$$\begin{aligned} & (1 - \phi_1^{\text{Peak}} B) (1 - \phi_4^{\text{Peak}} B^4) (1 - B^4) (1 - B^{48}) \varepsilon_{t,t_0}^{\text{Peak}} \\ &= (1 - \theta_1^{\text{Peak}} B - \theta_2^{\text{Peak}} B^2) (1 - \theta_4^{\text{Peak}} B^4) (1 - \theta_{48}^{\text{Peak}} B^{48}) a_{t,t_0}^{\text{Peak}}, \end{aligned} \quad (4)$$

where  $a_{t,t_0}^{\text{Peak}}$  is the innovation term.

Note that in models (3) and (4) the short-term properties of the pool prices are modelled through components with lags 1, 2 and 4, and the long-term properties of the forward curve are modelled through components with lag 48. Note also that the proposed models can be adapted to markets other than EEX by changing the corresponding ARIMA specifications.

The parameter estimation for models (3) and (4) is based on maximizing a conditional likelihood function for the available data, as described in Hillmer & Tiao (1979). Since the number of observations used to estimate the parameters is large enough (more than 12,000), instead of the exact likelihood function, a conditional function is used because both functions produce similar results, while the conditional one has a lower computational burden.

## 2.6 Estimation of Non-Available Errors

We characterize a non-available error as an additive outlier, as in Chen & Liu (1993). This characterization allows estimating the missing error values in an optimal manner. Nevertheless, given the large amount of non-available errors present in the data, specially at the end of the estimation window, we use a recursive procedure to estimate the non-available errors by blocks (using blocks of 4 periods at a time).

We illustrate this estimation considering that the forecasting day is December 30, 2005. At that time, we cannot use all available error data from January 2, 2004 to December 30, 2005 (24,480 periods) because there are missing error values. The first 4 missing error

values appear in the first trading day in January 2005. That is, in positions 12,285 to 12,288. To estimate these 4 missing error values we use the procedure in Chen & Liu (1993) considering error data from positions 1 to 12,284. Once these 4 missing values are estimated, the next 4 missing values appear in the second trading day of January 2005. That is, in positions 12,333 to 12,336. To estimate these 4 missing values we use again the procedure in Chen & Liu (1993) but now considering error data from positions 1 to 12,332. This procedure is repeated until the last 4 missing error values appear in the last trading day of January 2005. That is, we estimate the missing error values in positions 13,245 to 13,248 using data from positions 1 to 13,244.

Then, for each trading day in February 2005, there are 8 missing error values (instead of 4) to be estimated. For simplicity, we estimate these 8-error values using two blocks of 4 values. That is, for the first trading day in February 2005, the first block of 4 missing values appear in positions 13,289 to 13,292. We estimate these 4 missing values with data from positions 1,009 to 13,288. Once these values are estimated, we proceed to estimate the second block of 4 missing values in positions 13,293 to 13,296 but using data from positions 1,009 to 13,292. Once these 8 missing values are estimated, we use the same procedure to estimate the next 8 missing values appearing in the second trading day in February 2005.

We repeat the same procedure explained above for the remaining months in 2005, taking into account that the later the month in 2005, the higher the number of non-available errors and, hence, the higher the number of blocks of 4-error values to be estimated.

It should be noted that although the length of the estimation window in this recursive procedure is not equal in all iterations, it remains above 12,000 values.

After the missing error estimation, 48-period ahead forecasts (from position 24,432) can be obtained as indicated in the next section. The required information to generate a scenario tree of appropriate size is also readily available.

## 2.7 Forecasting Base and Peak Prices

Before detailing the forecasting algorithm, the following intuitive explanation of the approach is provided.

Consider that we intend to forecast the future spot price at time  $t$  based on data up to and including time  $t_0$ . For this we need to collect all error terms up to time  $t_0$ . However, as explained in Section 2.6, not all *past* “errors” are available up to time  $t_0$ , thus, we need to “fill” these *gaps* in a efficient way. To do so, we propose a recursive estimation procedure by blocks to compute the missing errors. Each step of this procedure is based on Chen & Liu approach, which considers each missing value as an additive outlier, and hence estimates the non-available errors in an optimal way.

Once we have replaced all non-available past errors by their corresponding estimates, we can estimate all parameters in equations (3) and (4), and finally we can forecast future errors from the same equations.

In summary, we have a collection of past errors with several missing values up to time  $t_0$ . We propose a recursive method to compute these missing values, which results in a complete collection of past errors. Then, we propose time series models to fit these past errors. Finally, we estimate model parameters and compute error forecasts to obtain price forecasts.

With this explanation, the forecasting algorithm for base prices is described below. For peak prices, an analogous algorithm can be built.

1. Parameter fitting and error characterization:

- (a) Based on historical data up to period  $t_0$ , obtain  $\phi_i^{\text{Base}}, \forall i; \theta_j^{\text{Base}}, \forall j$ .
- (b) Obtain also an estimation of the standard error of the white noise term,  $\sigma_{a_{t,t_0}^{\text{Base}}}$ .

2. Error forecasting:

Use model (3) to forecast errors  $\hat{\varepsilon}_{t_1,t_0}^{\text{Base}}, \dots, \hat{\varepsilon}_{t_{n_B},t_0}^{\text{Base}}$ .

3. Price forecasting:

- (a) Use model (1) and the un-transformation proposed by Guerrero (1993) to obtain the  $n_B$  price forecasts:

$$\hat{p}_{t_j}^{\text{Base}} = f_{t_0}^{\text{Base}}(t_j) \exp \left( \hat{\varepsilon}_{t_j,t_0}^{\text{Base}} + \frac{1}{2}(\sigma_{a_{t_j,t_0}^{\text{Base}}}^2) \right), \quad j = 1, \dots, n_B. \quad (5)$$

- (b)  $n_B$  confidence intervals at 95% level are computed below.

Lower:

$$f_{t_0}^{\text{Base}}(t_j) \exp \left( \hat{\varepsilon}_{t_j, t_0}^{\text{Base}} + \frac{1}{2} (\sigma_{a_{t_j, t_0}^{\text{Base}}})^2 - 1.96 \sigma_{a_{t_j, t_0}^{\text{Base}}} \right), j = 1, \dots, n_B \quad (6)$$

Upper:

$$f_{t_0}^{\text{Base}}(t_j) \exp \left( \hat{\varepsilon}_{t_j, t_0}^{\text{Base}} + \frac{1}{2} (\sigma_{a_{t_j, t_0}^{\text{Base}}})^2 + 1.96 \sigma_{a_{t_j, t_0}^{\text{Base}}} \right), j = 1, \dots, n_B \quad (7)$$

Note that these intervals are not symmetric.

## 2.8 Off-Peak Prices

Once base and peak forecasts are available, off-peak forecasts are obtained as explained below.

Base, peak and off-peak prices are related as:

$$p_d^{\text{Base}} = \frac{1}{24} \sum_{h=1}^{24} p_{d,h}^{\text{Base}} \quad (8)$$

$$= \frac{1}{24} \left( \sum_{h \in \Omega_{\text{Peak}}} p_{d,h}^{\text{Peak}} + \sum_{h \in \Omega_{\text{Off-peak}}} p_{d,h}^{\text{Off-peak}} \right) \quad (9)$$

$$= \frac{1}{24} (|\Omega_{\text{Peak}}| p_d^{\text{Peak}} + |\Omega_{\text{Off-peak}}| p_d^{\text{Off-peak}}), \quad (10)$$

thus,

$$p_d^{\text{Off-peak}} = \frac{1}{|\Omega_{\text{Off-peak}}|} (24 p_d^{\text{Base}} - |\Omega_{\text{Peak}}| p_d^{\text{Peak}}). \quad (11)$$

Hence, the off-peak price forecasts depend on base price forecasts and peak price ones as

$$\hat{p}_t^{\text{Off-peak}} = \frac{1}{|\Omega_{\text{Off-peak}}|} (24 \hat{p}_t^{\text{Base}} - |\Omega_{\text{Peak}}| \hat{p}_t^{\text{Peak}}). \quad (12)$$

Since base and peak prices are correlated, caution should be exercised in applying formula (12) above. Particularly, we propose to use the same seed to generate both base and peak prices.

## 2.9 Scenario Generation Algorithm

The scenario generation algorithm for base prices is described below. For peak prices, an analogous algorithm can be built.

1. Single scenario generation:

(a) Randomly generate the  $n_B$  white noise terms  $a_{t_j, t_0}^{\text{Base}} \sim N(0, \sigma_{a_{t_0}^{\text{Base}}})$ ;  $j = 1, \dots, n_B$ .

Note that  $\sigma_{a_{t_0}^{\text{Base}}}$  is the standard deviation of the estimation error using data up to  $t_0$ .

(b) Then, use equation (3) to obtain a realization of errors  $\varepsilon_{t_1, t_0}^{\text{Base}}, \dots, \varepsilon_{t_{n_B}, t_0}^{\text{Base}}$ .

(c) Finally, use equation (5) to obtain a realization of prices  $p_{t_1}^{\text{Base}}, \dots, p_{t_{n_B}}^{\text{Base}}$ .

2. Scenario tree generation:

Repeat step (1) above as many times as needed to generate a scenario tree of appropriate size, e.g., 1,000 times.

Once base and peak price scenarios are generated, off-peak price scenarios are derived using (12). Finally, the peak and off-peak prices of each scenario are merged to produce a resulting scenario covering the whole year and following the actual time sequence. Note that the resulting scenario tree comprises  $2 \times n_B$  periods ( $n_B$  peak values and  $n_B$  off-peak values).

## 3 Case study

### 3.1 Introduction and Data

A real-world case study is comprehensively analyzed in this section and the obtained results reported. This case study is based on the European Energy Exchange (EEX), an exchange under public law located in Leipzig, Germany.

EEX operates both the pool trading for electric energy with physical delivery (one-day ahead) and the futures market. The EEX futures market is well developed and has sufficient financial liquidity, i.e., it has a sufficient volume of forward trade.

We have collected all EEX (month and quarter) futures prices (for base and peak hours) for each trading day in 2004 and 2005, that is, for 510 trading days. Additionally, we have collected the 24 hourly pool prices for each day in 2004-2006. From this data and using the 48-value yearly discretization proposed in Section 2.1, we calculate the  $510 \times 48 = 24,480$

error values for the base model (including missing errors) and the corresponding 24,480 error values for the peak model (including missing errors as well).

The models presented in Section 2 are then used for out-of-sample forecasting as explained below.

### 3.2 Forecasting

To evaluate the out-of-sample performance of the proposed models, we calculate 48 periods-ahead forecasts for each Friday in October, November and December, 2005 (that is, for 13 trading days). Hence, in the out-of-sample study, we have selected 13 trading days and, for each of them, electricity prices are forecasted for 48 periods ahead (one whole year).

Note that no previous trading day in 2005 is considered due to the short history of the EEX market, which makes the proportion of non-available errors over the available ones unreasonable for dates previous to October 2005.

The out-of-sample evaluation is summarized as follows. First and for each of these 13 days, we estimate the missing error values by the recursive procedure stated in Section 2.6. Second, the parameters of the models are estimated and 48-period ahead forecasts (confidence intervals included) are computed. Third, the differentiation and transformation (logarithm) for the forecasts are reversed and diverse performance metrics regarding prediction errors are computed.

Finally, note that in order to appropriately consider the strong periodic pattern exhibited by the data, we re-estimate the model parameters each time we change the trading day. Particularly, we re-estimate models (3) and (4), and the corresponding forecasting equation (5), thirteen times. Doing so, we ensure that the proposed models capture the corresponding seasonal patterns, no matter the day of the week in which we initiate the forecasting horizon.

Results are reported below.



### 3.3 Performance Metrics and Benchmark Models

To appraise the accuracy of the proposed models, we use the standard performance metrics below:

$$\text{MAPE} = \frac{100}{n_F} \times \sum_{t=1}^{n_F} \frac{|\hat{p}_t - p_t^{\text{true}}|}{p_t^{\text{true}}} \quad (13)$$

$$\text{EMax} = 100 \times \max \left\{ \frac{|\hat{p}_1 - p_1^{\text{true}}|}{p_1^{\text{true}}}, \frac{|\hat{p}_2 - p_2^{\text{true}}|}{p_2^{\text{true}}}, \dots, \frac{|\hat{p}_{n_F} - p_{n_F}^{\text{true}}|}{p_{n_F}^{\text{true}}} \right\} \quad (14)$$

$$\text{RMSE} = \sqrt{\frac{1}{n_F} \sum_{t=1}^{n_F} (\hat{p}_t - p_t^{\text{true}})^2} \quad (15)$$

Additionally, two benchmark models are used to assess the behavior of the proposed technique. Note that long-term forecasting methods for electricity prices have not been considered previously in the forecasting literature.

The first benchmark model for the pool price is:

$$\log p_t^{\text{Base}} = \log f_{t_0}^{\text{Base}}(t) + a_{t,t_0}^{\text{Base}} \quad (16)$$

$$\log p_t^{\text{Peak}} = \log f_{t_0}^{\text{Peak}}(t) + a_{t,t_0}^{\text{Peak}}, \quad (17)$$

where  $a_{t,t_0}^{\text{Base}}$  and  $a_{t,t_0}^{\text{Peak}}$  are the innovation terms, and, as usual, they are assumed to be white noise. Note that this model assumes no specific patterns between expected pool prices and futures-product prices. Therefore, there is no need to estimate non-available errors as the standard deviation characterizing the innovation term can be directly computed with the available data.

For the second benchmark model, we have adapted the standard exponential weighted moving average (frequently used in practice without seasonality). We have incorporated additional moving average terms to appropriately capture the monthly and yearly seasonality patterns of the forecast error series. The resulting model for this benchmark is:

$$(1 - B) \log p_t^{\text{Base}} = (1 - \theta_1^{\text{Base}} B) (1 - \theta_4^{\text{Base}} B^4) (1 - \theta_{48}^{\text{Base}} B^{48}) a_{t,t_0}^{\text{Base}} \quad (18)$$

$$(1 - B) \log p_t^{\text{Peak}} = (1 - \theta_1^{\text{Peak}} B - \theta_2^{\text{Peak}} B^2) (1 - \theta_4^{\text{Peak}} B^4) (1 - \theta_{48}^{\text{Peak}} B^{48}) a_{t,t_0}^{\text{Peak}}, \quad (19)$$

where  $a_{t,t_0}^{\text{Base}}$  and  $a_{t,t_0}^{\text{Peak}}$  are the innovation terms. We have selected this benchmark for its good performance in forecasting short-term pool prices, but note this second benchmark

does not include relevant long-term information (futures-product prices). It expresses past predictions as weighted combinations of past prices and past predictions of orders 1, 4, and 48.

### 3.4 Results

All models have been implemented in the SCA System, Liu and Hudak (1994). Using this tool, we can estimate the missing error values and the parameters in the models and obtain 48-period ahead forecasts.

The cases have been run on a Dell PowerEdge 6600 with 2 processors at 1.60 GHz and 2 GB of RAM memory. For each one of the 13 forecasting studies (including parameter estimation and forecasting of prices), running times pertaining to peak and base prices are under 5 and 7 CPU minutes, respectively.

Tables 1 and 2 provide the estimates of the parameters characterizing the proposed models (3) and (4), respectively, for a specific date (December 30, 2005). Particularly, each table contains the estimation for each of the parameters of the corresponding models, the corresponding standard error, and the corresponding  $t$ -value to measure parameter significativeness. Finally, each table contains the residual standard error, which is an estimation of the standard error of the white noise term:  $\sigma_{a_{t,t_0}^{\text{Base}}}$  for the base model and  $\sigma_{a_{t,t_0}^{\text{Peak}}}$  for the peak model.

From Tables 1 and 2, we can see the significant effect of the pool price properties over the short-term together with the significant effect of the forward curve (or forward premium) over the long-term. Moreover, diagnostic checks reveal that the proposed methodology performs adequately in this case study. Finally, note that the information provided in Tables 1 and 2 need to be updated for each trading day.

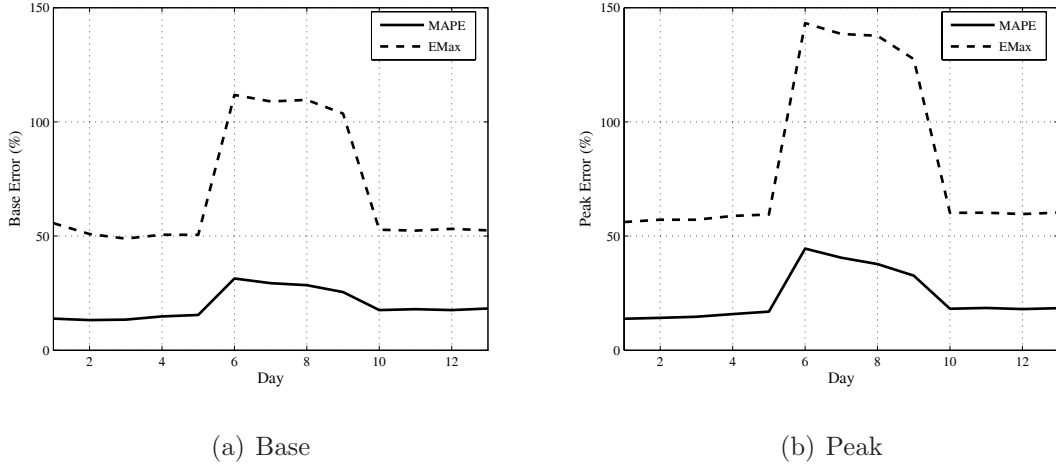
Figure 5 shows the evolution of MAPE and EMax for the 13 forecasting days and the base and peak models, respectively.

Figure 6 shows the base and peak true pool prices, respectively, for 2006 and the corresponding confidence intervals computed December 30, 2005.

Figure 7 depicts 1,000 scenarios of EEX pool prices covering peak and off-peak periods for 2006, computed December 30, 2005. Peak and off-peak prices are generated separately

**Table 1: Base-error model parameters computed at December 30, 2005.**

Base			
Parameter	Estimation	Std. Error	<i>t</i> -value
$\phi_1^{\text{Base}}$	1.4905	0.0224	66.3
$\phi_2^{\text{Base}}$	-0.5143	0.0197	-26.0
$\phi_4^{\text{Base}}$	-0.6693	0.0066	-101.2
$\phi_{48}^{\text{Base}}$	-0.9032	0.0663	-13.6
$\theta_1^{\text{Base}}$	0.7978	0.0186	42.9
$\theta_4^{\text{Base}}$	0.9974	0.0410	24.3
$\theta_{48}^{\text{Base}}$	-0.9022	0.0636	-14.1
Residual standard error = 0.020			



**Figure 5: EEX market: MAPE and EMax evolution for base/peak prices and the 13 forecasting days.**

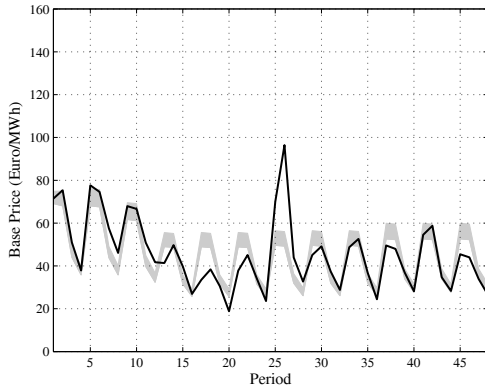
and then merged in the right time sequence to produce the scenarios plotted in Figure 7.

Finally, a comparison between the proposed methodology and the two benchmark models presented in Section 3.3 is carried out.

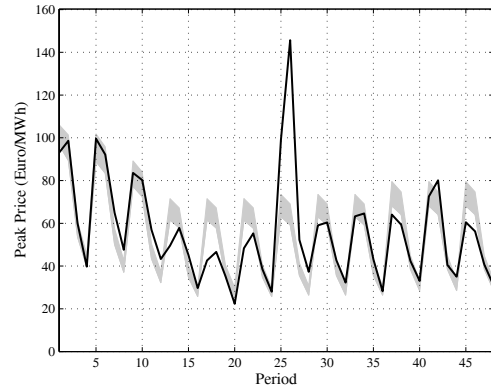
First of all, Table 3 presents the performance metrics (MAPE, EMax and RMSE) of the proposed methodology for the selected 13 forecasting days (and the 48 predictions). Note that these metrics imply an accurate performance of the proposed forecasting technique for the considered forecasting horizon that spans one whole year.

**Table 2: Peak-error model parameters computed at December 30, 2005.**

Peak			
Parameter	Estimation	Std. Error	<i>t</i> -value
$\phi_1^{\text{Peak}}$	0.9332	0.0046	202.1
$\phi_4^{\text{Peak}}$	-0.5933	0.0072	-81.9
$\theta_1^{\text{Peak}}$	0.0178	0.0098	1.8
$\theta_2^{\text{Peak}}$	0.4536	0.0098	46.5
$\theta_4^{\text{Peak}}$	0.9926	0.2695	3.6
$\theta_{48}^{\text{Peak}}$	0.0199	0.0087	2.2
Residual standard error = 0.023			



(a) Base



(b) Peak

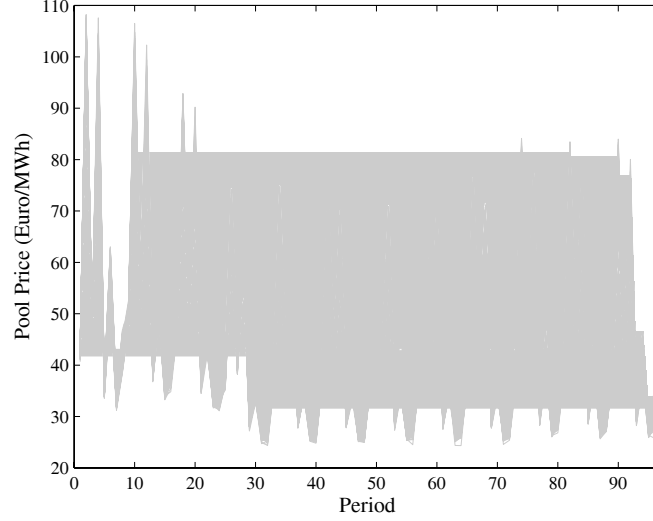
**Figure 6: EEX market: True base/peak pool prices for 2006 and confidence intervals at December 30, 2005.**

Table 4 below presents the performance metrics (MAPE, EMax and RMSE) for the first benchmark model (16)–(17) for the selected 13 forecasting days (and the 48 predictions). From this table, we conclude that the proposed methodology conveniently exploit the specific patterns observed between expected pool prices and futures-product prices. For example, the MAPE of the proposed model is approximately half of that for this first benchmark model.

Table 5 below shows the performance metrics for the second benchmark model (18)–(19). Note that forecasts at December 2 (2005) and November 4 (2005) cannot be performed as

**Table 3: Proposed methodology: Forecasting-error metrics for the last Fridays in October, November and December, 2005.**

Case	Date	Base			Peak		
		MAPE (%)	EMax (%)	RMSE (€/MWh)	MAPE (%)	EMax (%)	RMSE (€/MWh)
1	Dec. 30, 2005	13.8	55.6	9.5	13.8	56.1	15.0
2	Dec. 23, 2005	13.2	50.9	9.5	14.2	57.2	15.3
3	Dec. 16, 2005	13.4	48.9	9.8	14.7	57.2	15.4
4	Dec. 9, 2005	14.8	50.6	10.8	15.9	58.8	16.2
5	Dec. 2, 2005	15.4	50.5	11.1	16.9	59.4	16.8
6	Nov. 25, 2005	31.4	111.8	17.0	44.5	143.3	31.7
7	Nov. 18, 2005	29.4	109.0	15.8	40.5	138.5	29.0
8	Nov. 11, 2005	28.5	109.6	15.4	37.7	137.7	27.2
9	Nov. 4, 2005	25.5	103.7	14.1	32.7	127.5	24.3
10	Oct. 28, 2005	17.6	52.8	14.3	18.2	60.2	22.0
11	Oct. 21, 2005	18.0	52.4	14.6	18.5	60.2	22.3
12	Oct. 14, 2005	17.6	53.2	14.3	18.0	59.6	21.7
13	Oct. 07, 2005	18.3	52.5	14.8	18.4	60.3	22.3



**Figure 7: EEX market: 1,000 pool-price scenarios for 2006 computed at December 30, 2005.**

the data available up to those dates are not sufficient to conduct the time discretization proposed in Section 2.1 for the months of December and November 2005, respectively. We can observe from this table that the performance of this second benchmark is clearly worse than that of the proposed methodology.

### 3.5 Discussion

Observing Table 3 and analyzing Figures 5-7, the conclusions below are drawn:

1. MAPE, EMax and RMSE of the proposed technique remain below 32%, 112% and 17 €/MWh, respectively, for base prices, and below 45%, 144% and 32 €/MWh, respectively, for peak prices.
2. Results are surprisingly accurate for the considered forecasting horizon that spans one whole year. Moreover, the price forecasts by the proposed method are substantially more accurate than those provided by other methods.
3. The proposed model is therefore very appropriate for the EEX market. Other markets should be tested to further validate the proposed methodology.

**Table 4: First benchmark: Forecasting-error metrics for the last Fridays in October, November and December, 2005.**

Case	Date	Base			Peak		
		MAPE (%)	EMax (%)	RMSE (€/MWh)	MAPE (%)	EMax (%)	RMSE (€/MWh)
1	Dec. 30, 2005	40.6	169.2	17.5	63.9	212.6	32.0
2	Dec. 23, 2005	38.0	161.1	16.9	61.2	209.6	31.0
3	Dec. 16, 2005	36.8	157.7	16.5	60.9	210.7	30.9
4	Dec. 9, 2005	33.1	147.6	15.4	54.9	199.1	28.2
5	Dec. 2, 2005	32.5	147.8	15.2	52.3	193.4	27.0
6	Nov. 25, 2005	31.4	143.3	15.3	49.8	188.0	27.2
7	Nov. 18, 2005	30.5	140.0	15.1	47.4	182.3	26.0
8	Nov. 11, 2005	30.4	140.7	15.2	46.4	181.3	25.3
9	Nov. 4, 2005	29.0	133.8	15.3	42.5	169.1	24.3
10	Oct. 28, 2005	28.5	133.6	15.9	42.2	169.5	25.4
11	Oct. 21, 2005	28.1	130.8	16.0	42.0	169.0	25.4
12	Oct. 14, 2005	28.7	134.3	15.9	43.4	173.5	25.6
13	Oct. 07, 2005	27.8	129.8	16.0	42.0	169.0	25.3

## 4 Conclusions

Since forecasting errors greatly increase as the forecasting horizon expands, the year-ahead uncertainty characterization of pool prices is a hard and complex endeavor. However, year-ahead pool price scenarios covering most realizations are crucial to make informed decision in financial markets, thus achieving appropriate risk hedging.

Based on the prices of diverse financial products traded at financial markets, this paper provides an efficient yet sufficiently accurate procedure to characterize pool price uncertainty one-year ahead. Moreover, scenario trees to be used within stochastic programming decision making tools are also generated.

Detailed studies pertaining to a real-world electricity market demonstrate the theoretical

**Table 5: Second benchmark: Forecasting-error metrics for the last Fridays in October, November and December, 2005.**

Case	Date	Base			Peak		
		MAPE (%)	EMax (%)	RMSE (€/MWh)	MAPE (%)	EMax (%)	RMSE (€/MWh)
1	Dec. 30, 2005	41.1	179.2	18.0	31.8	114.6	22.9
2	Dec. 23, 2005	47.9	198.0	19.7	34.3	120.5	23.3
3	Dec. 16, 2005	52.0	209.0	21.0	36.0	123.6	24.0
4	Dec. 9, 2005	51.2	206.9	20.7	37.5	127.4	24.6
5	Dec. 2, 2005	—	—	—	—	—	—
6	Nov. 25, 2005	31.8	148.9	16.9	59.1	206.2	30.2
7	Nov. 18, 2005	29.7	133.8	17.2	48.5	178.5	26.2
8	Nov. 11, 2005	28.7	123.3	17.7	41.1	157.7	24.6
9	Nov. 4, 2005	—	—	—	—	—	—
10	Oct. 28, 2005	29.0	129.1	18.7	38.9	150.0	26.9
11	Oct. 21, 2005	28.9	128.6	18.7	38.8	149.5	27.0
12	Oct. 14, 2005	28.8	127.5	18.8	38.6	148.9	27.0
13	Oct. 07, 2005	29.1	130.5	18.6	38.6	148.6	27.2

and practical interest of the proposed methodology.

## Notation

The notation used throughout the paper is stated below for quick reference. A price forecast is indicated with a “ $\hat{\phantom{x}}$ ”, and superscripts “Base”, “Peak” and “Off-peak” indicate base, peak and off-peak values, respectively.

### Pool Prices:

$p_{d,h}^{\text{Base}}$  Pool price at hour  $h$  of day  $d$ .



$p_{d,h}^{\text{Peak}}$  Pool price at peak hour  $h$  of day  $d$ .

$p_{d,h}^{\text{Off-peak}}$  Pool price at off-peak hour  $h$  of day  $d$ .

$p_d^{\text{Base}}$  Average pool price of day  $d$ .

$p_d^{\text{Peak}}$  Average pool price of peak hours of day  $d$ .

$p_d^{\text{Off-peak}}$  Average pool price of off-peak hours of day  $d$ .

$p_t^{\text{Base}}$  Pool price of base period  $t$ .

$p_t$  Pool price of period  $t$ .

$p_t^{\text{Peak}}$  Pool price of peak period  $t$ .

$p_t^{\text{Off-peak}}$  Pool price of off-peak period  $t$ .

$p_t^{\text{true}}$  Pool price true value in period  $t$ .

## Forward Functions:

$f_{t_0}^{\text{Base}}(t)$  Forward base curve providing at period  $t_0$  available futures prices for all relevant future periods  $t > t_0$ .

$f_{t_0}^{\text{Peak}}(t)$  Forward peak curve providing at period  $t_0$  available futures prices for all relevant future periods  $t > t_0$ .

## Errors:

$\varepsilon_{t,t_0}^{\text{Base}}$  Base price forecast error at time period  $t > t_0$  for data available up to  $t_0$ .

$\varepsilon_{t,t_0}^{\text{Peak}}$  Peak price forecast error at time period  $t > t_0$  for data available up to  $t_0$ .

$a_{t,t_0}^{\text{Base}}$  ,  $a_{t,t_0}^{\text{Peak}}$  White noise.

$\sigma_{a_{t,t_0}^{\text{Base}}}$  Standard deviation of the base-price forecast error for data available up to  $t_0$  and time  $t > t_0$ .

$\sigma_{a_{t,t_0}^{\text{Peak}}}$  Standard deviation of the peak-price forecast error for data available up to  $t_0$  and time  $t > t_0$ .

$\sigma_{a_{t_0}^{\text{Base}}}$  Standard deviation of the base-price estimation error for data available up to  $t_0$ .

$\sigma_{a_{t_0}^{\text{Peak}}}$  Standard deviation of the peak-price estimation error for data available up to  $t_0$ .

## Parameters:

$\phi_i^{\text{Base}}, \phi_i^{\text{Peak}}$  Polynomial price constants of order  $i$ .

$\theta_i^{\text{Base}}, \theta_i^{\text{Peak}}$  Polynomial error constants of order  $i$ .

## Operators:

$B^k$  Backshift operator, i.e.,  $B^k p_t = p_{t-k}$ .

$|\cdot|$  Both cardinality and absolute value.

## Indices:

$d$  Index for days.

$h$  Index for hours.

$t$  Index for periods typically including several hours.

## Numbers:

$n_F$  Number of forecast values.

$n_B$  Number of forecast values pertaining to base periods.

## Sets:

$\Omega_{\text{Peak}}$  Set of the peak hours of the day.

$\Omega_{\text{Off-peak}}$  Set of the off-peak hours of the day.

## References

- Birge, R. & Louveaux, F. (1997). *Introduction to stochastic programming*. New York: Springer-Verlag.
- Bunn, D. W. (2006). *Modelling prices in competitive electricity markets*. New York: John Wiley & Sons, Inc.
- Carrión, M., Philpott, A., Conejo, A. J. & Arroyo, J. M. (2007a). A Stochastic Programming Approach to Electric Energy Procurement for Large Consumers. *IEEE Transactions on Power Systems*, 22, 2, 744-754.
- Carrión, M., Conejo, A. J. & Arroyo, J. M. (2007b). Forward Contracting and Selling Price Determination for a Retailer. *IEEE Transactions on Power Systems*, 22, 4, 2105-2114.
- Chen, C. & Liu, L. (1993). Joint estimation of models parameters and outliers effects in time series. *Journal of the American Statistical Association*, 88, 284-297.
- Conejo, A. J., García-Bertrand, R. & Carrión, M. (2008). Optimal Involvement in Forward Markets of a Power Producer. *IEEE Transactions on Power Systems*, 23, 2, 703-711.
- Davison, M., Anderson, L., Marcus, B. & Anderson, K. (2002). Development of a hybrid model for electrical power spot prices. *IEEE Transactions on Power Systems*, 17, 257-264.
- Guerrero, V. M. (1993). Time-series analysis supported by power transformations. *Journal of Forecasting*, 12, 37-48.
- Hillmer, S. C. & Tiao, G. C. (1979). Likelihood function of stationary multiple autoregressive moving average models. *Journal of the American Statistical Association*, 74, 652-660.
- Karakatsani, N. V. & Bunn, D. W. (2003). Forecasting Electricity Prices. *DW, EMG Working Paper*,  
([http://www.london.edu/assets/documents/PDF/2.3.4.12.1\\_Karakatsani\\_and\\_Bunn\\_2003\\_FEP\\_pdf5.pdf](http://www.london.edu/assets/documents/PDF/2.3.4.12.1_Karakatsani_and_Bunn_2003_FEP_pdf5.pdf))
- Liu, L. & Hudak, G. P. (1994). *Forecasting and time series analysis using the SCA statistical system*. Chicago: Scientific Computing Associated.

- Longstaff, F. A. & Ashley, W.W. (2004). Electricity Forward Prices: A High Frequency Empirical Analysis. *Journal of Finance*, 59, 1743-1776.
- Schwartz, E. S. (1997). The stochastic behavior of commodity prices: implications for valuation and hedging. *Journal of Finance*, 52, 923-973.
- Schwartz, E. S. & Smith, J. E. (2000). Short-term variations and long-term dynamics in commodity prices. *Management Science*, 46, 893-911.