Day-Ahead Electricity Price Forecasting Using the Wavelet Transform and ARIMA Models

Antonio J. Conejo, *Fellow, IEEE*, Miguel A. Plazas, Rosa Espínola, *Student Member, IEEE*, and Ana B. Molina

***Abstract*—This paper proposes a novel technique to forecast day-ahead electricity prices based on the wavelet transform and ARIMA models. The historical and usually ill-behaved price series is decomposed using the wavelet transform in a set of better-behaved constitutive series. Then, the future values of these constitutive series are forecast using properly fitted ARIMA models. In turn, the ARIMA forecasts allow, through the inverse wavelet transform, reconstructing the future behavior of the price series and therefore to forecast prices. Results from the electricity market of mainland Spain in year 2002 are reported.**

***Index Terms*—ARIMA models, electricity market, price forecasting, wavelet transform.**

NOMENCLATURE

The notation used throughout the paper is provided below for easy reference.

$\mathrm{a_h}$	First constitutive price series (Detail 1).
$\mathrm{b_h}$	Second constitutive price series (Detail 2).
$\mathrm{c_h}$	Third constitutive price series (Detail 3).
$\mathrm{d_h}$	Fourth constitutive price series (Approximation).
$\mathrm{e_{day}}$	Daily error.
$\mathrm{e_h}$	Hourly error.
$\mathrm{e_{week}}$	Weekly error.
$\mathrm{p_h}$	Price series and price value in hour h.
$\mathrm{p_h^{est}}$	Estimate of the price in hour h.
$\mathrm{p_h^{true}}$	True price in hour h.
$\mathrm{p_h^{A,est}}$	Estimate of the price in hour h using an ARIMA model.
$\mathrm{p_h^{N,est}}$	Estimate of the price in hour h using the naïve model.
$\mathrm{p_h^{W,est}}$	Estimate of the price in hour h using the proposed Wavelet-ARIMA model.
$\mathrm{\bar{p}_{day}^{true}}$	True average daily price.
$\mathrm{\bar{p}_{week}^{true}}$	True average weekly price.
T	Number of available historical price values used for prediction.
$\mathrm{W}(\cdot)$	Wavelet transform.
$\mathrm{W}^{-1}(\cdot)$	Inverse wavelet transform.
$\sigma^2_{\mathrm{e,day}}$	Daily error variance.
$\sigma^2_{\mathrm{e,week}}$	Weekly error variance.

Manuscript received June 15, 2004; revised September 21, 2004. A. J. Conejo, R. Espínola, and A. B. Molina are partly supported by the Ministry of Science and Education of Spain under CICYT Project DPI2003-01362 and by Junta de Comunidades de Castilla–La Mancha under Project GC-02-006. Paper no. TPWRS-00323-2004.

A. J. Conejo, R. Espínola, and A. B. Molina are with the Department of Electrical Engineering of the University of Castilla-La Mancha, Ciudad Real, Spain (e-mail: Antonio.Conejo@uclm.es; Rosa.Espinola@uclm.es; AnaB.Molina@alu.uclm.es).

Miguel A. Plazas is with Unión Fenosa Generación, Madrid, Spain (e-mail: maplazas@uef.es).

Digital Object Identifier 10.1109/TPWRS.2005.846054

I. INTRODUCTION

THIS paper frames itself in a pool-based electric energy market. In such a market, producers submit to the market operator selling bids consisting in energy blocks and their corresponding minimum selling prices, and consumers submit to the market operator buying bids consisting in energy blocks and their corresponding maximum buying prices. In turn, the market operator clear the market using an appropriate market clearing procedure that results in hourly energy prices and accepted selling and buying bids.

In the above framework, price forecasting is required by producers and consumers. Both producers and consumers use day-ahead price forecasts to derive their respective bidding strategies to the electricity market. Therefore, accurate price estimates are crucial for producers to maximize their profits and for consumers to maximize their utilities.

Forecasting electricity prices is difficult because unlike demand series, price series present such characteristics as nonconstant mean and variance and significant outliers.

The wavelet transform convert a price series in a set (typically three to six) of constitutive series. These series present a better behavior (more stable variance and no outliers) than the original price series, and therefore, they can be predicted more accurately. The reason for the better behavior of the constitutive series is the filtering effect of the wavelet transform. The procedure explained in this paper, denominated henceforth Wavelet-ARIMA technique, takes advantage of such behavior. To forecast the 24 hourly prices of day d, the Wavelet-ARIMA technique works as follows.

Step 1) Decompose through the wavelet transform the available historical price series (up to hour 24 of day d-1) in a set of constitutive series (typically 4).

Step 2) Use a specific ARIMA model fitted to each one of the constitutive series to forecast its 24 future values for day d.

Step 3) Use the inverse wavelet transform to forecast the hourly prices for day d using the estimates for day d of the constitutive series.

This technique is compared with an ARIMA model used directly to forecast the original price series. A naïve procedure is also used for comparison. This naïve procedure establishes just that the price estimates for a given week are the actual prices of the previous week, therefore, using a no-change criterion.

This paper shows that the use of the wavelet transform as a preprocessor of forecasting data improves the predicting behavior of any technique such as ARIMA, transfer function, neural network and others. This is way the comparison is performed with and without the wavelet transform, not across techniques. For comparison across techniques, see our previous works [1] and [2].

For the sake of simplicity and clear comparison, no explicative variables are considered in the ARIMA models used in this paper. However, we recognize that explicative variables such as demand, hydro resources, maintenance outages and temperature, do help to improve predictions. Moreover, ARIMA models are not significantly affected by explicative variables, but this is not the case with transfer function or neural network models. Nevertheless, this paper focuses on the advantages of using a wavelet preprocessor, and therefore the time series technique subsequently used is of secondary relevance.

Reported techniques to forecast day-ahead prices include ARIMA models [1] and [3], dynamic regression models [2], other time series techniques [4] and [5], neural network procedures [6]–[13], wavelet transform models [14] and [15], heuristics [16], Bayesian techniques [17], and simulations and others [18]–[20]. We believe that the hybrid technique (wavelet transform plus ARIMA models) proposed in this paper is both novel and effective.

The fundamental and novel contribution of the paper is to use the wavelet transform to decompose an ill-behaved price series into a set of better-behaved constitutive series. The future behavior of all the constitutive series is then predicted and the wavelet transform reversed to generate price prediction.

This paper is organized as follows. Section II specifies the market time-framework used to forecast day-ahead prices. Section III provides the details of the proposed forecasting technique based on the wavelet transform and ARIMA models. Section IV contains the error and error variance definitions used to assess the behavior of the technique proposed. Section V is a case study based on the day-ahead electric energy market of mainland Spain. Finally, Section VI provides some relevant conclusions. Two appendices provide background information.

II. Forecasting Framework

The time framework to forecast the day-ahead market prices in most markets is explained below and illustrated in Fig. 1.

The market price forecasts for day d are required on day d-1, typically at hour h_b (around 10 am). On the other hand, data concerning results for day d-1 are available on day d-2 at hour h_c (around noon). Therefore, the actual forecasting of market prices for day d can take place between hour h_c of day d-2 and hour h_b of day d-1. Therefore, to forecast prices for day d, price data up to hour 24 of day d-1 are considered known.

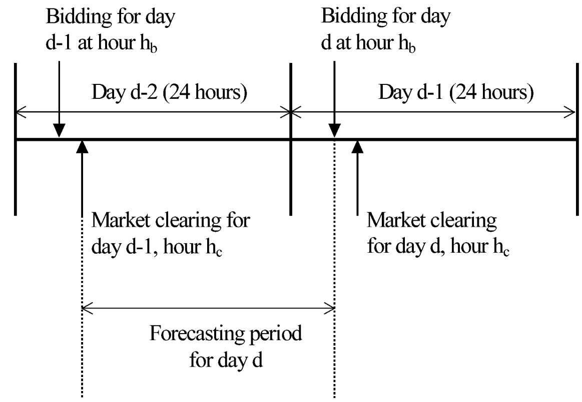

Fig. 1. Time framework to forecast market prices for day d.

III. Forecasting Using the Wavelet Transform and ARIMA Models

The available historical price series data to forecast the 24 hourly prices of day d is denoted by

$$p_h; \quad h = 1, \ldots, T. \tag{1}$$

This series includes historical data up to hour 24 of day d-1. The value of T ranges usually between 168 (1 week) to 1344 (2 months).

The Wavelet-ARIMA technique works as follows.

Step 1) Decompose through the wavelet transform the available historical price series (up to hour 24 of day d-1) in a set of four constitutive series.

A wavelet function of type Daubechies of order 5 and decomposition level 3 is used in this case study. This wavelet offers an appropriate tradeoff between wave-length and smoothness. This results in an appropriate behavior for price prediction.

The wavelet transform applied to price series $p_h \ (h = 1, \ldots, T)$ result in 4 series denoted by a_h, b_h, c_h, and $d_h; h = 1, \ldots, T$. Series a_h, b_h, and $c_h; h = 1, \ldots, T$ are denominated Detail (1, 2, and 3) series, while $d_h; h = 1, \ldots, T$ is denominated the Approximation series. This Approximation series constitutes the main component of the transform, while the three details series provides "small" adjustments. Thus, applying the wavelet transform to the original prices series results in

$$W(p_h; h = 1, \ldots, T) = \{a_h, b_h, c_h, d_h; h = 1, \ldots, T\}. \tag{2}$$

For reader's convenience, the wavelet transform [21] is briefly described in Appendix A.

Step 2) Use a specific ARIMA model of each one of the constitutive series to forecast its 24 future values for day d.

An ARIMA model is then used to forecast hours $T+1$ to $T+24$ for each one of the constitutive series a_h, b_h, c_h, and $d_h; h = 1, \ldots, T$; resulting is estimated series $a_h^{est}, b_h^{est}, c_h^{est}$, and $d_h^{est}; h = T+1, \ldots, T+24$.

For reader's convenience, the ARIMA technique [22] is briefly explained in Appendix B.

Step 3) Use the inverse wavelet transform to estimate the hourly prices for day d using the estimates for day d of the constitutive series.

The inverse wavelet transform is used in turn to reconstruct the estimate series for prices, i.e.,

$$W^{-1}\left(\{a_h^{est}, b_h^{est}, c_h^{est}, d_h^{est}; h = T+1, \ldots, T+24\}\right) = p_h^{W,est} \quad h = T+1, \ldots, T+24. \tag{3}$$

It is crucial to note that the approximation series $d_h(h = 1, \ldots, T)$ consistently presents a more stable variance than the

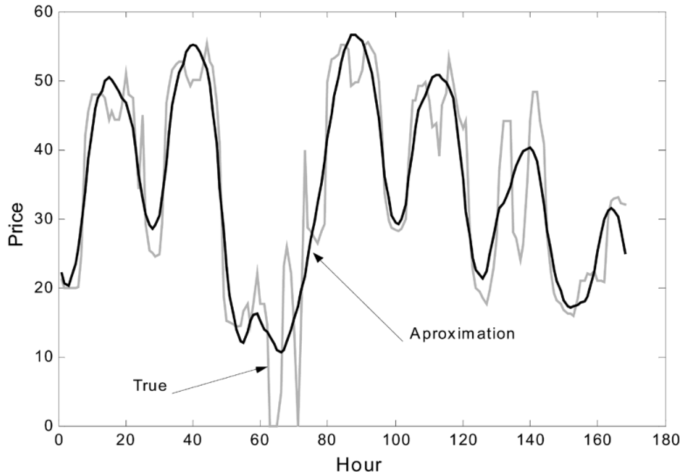

Fig. 2. Part of the original and approximation series (November 4–10, 2002) used to forecast the winter week.

original price series and no outliers. Fig. 2 shows one of the original price series used in the case study (part of the data used to forecast the winter week) and its corresponding approximation series. The beneficial filtering effect of the wavelet transform is apparent observing Fig. 2. The rationale above has motivated the development of the Wavelet-ARIMA technique.

For the sake of comparison, ARIMA and naïve models are used to forecast directly the prices of hours 1 to 24 of day d, resulting respectively in estimate price series

$$p_h^{A,est}; \quad h = T+1, \ldots, T+24 \tag{4}$$

and

$$p_h^{N,est}; \quad h = T+1, \ldots, T+24. \tag{5}$$

The technique reported in the paper is restricted to predict the general price trend including no spikes. If the market under consideration suffers from cyclical occurrence of spikes, specific procedures to estimate such spikes are required.

IV. Prediction Accuracy

To assess the prediction capacity of the Wavelet-ARIMA model, in addition to the hourly error, two type of average prediction errors are computed: the one corresponding to the 24 hours of each day and the one corresponding to the 168 hours of each week.

The per unit hourly error is computed as

$$e_h = \frac{|p_h^{true} - p_h^{est}|}{p_h^{true}}. \tag{6}$$

The per unit daily error is computed as

$$e_{day} = \frac{1}{24} \sum_{h=1}^{24} \frac{|p_h^{true} - p_h^{est}|}{\bar{p}_{day}^{true}} \tag{7}$$

where

$$\bar{p}_{day}^{true} = \frac{1}{24} \sum_{h=1}^{24} p_h^{true}. \tag{8}$$

The denominator of the right hand side of (7) is the average daily price to avoid the adverse effect of hourly prices close to zero.

Analogously to the daily error, the per unit weekly error, e_{week}, is computed as

$$e_{week} = \frac{1}{168} \sum_{h=1}^{168} \frac{|p_h^{true} - p_h^{est}|}{\bar{p}_{week}^{true}} \tag{9}$$

where

$$\bar{p}_{week}^{true} = \frac{1}{168} \sum_{h=1}^{168} p_h^{true}. \tag{10}$$

Again, the denominator of the right hand side of (9) is the average weekly price to avoid the adverse effect of prices close to zero.

A measure of the uncertainty of a model is the variability of what is still unexplained after fitting the model, which can be measured through the estimation of the variance of the error. The smaller this variance, the more precise is the prediction of prices.

Consistent with definitions (7) and (9), daily and weekly error variances can be estimated as

$$\sigma_{e,day}^2 = \frac{1}{24} \sum_{h=1}^{24} \left(\frac{|p_h^{true} - p_h^{est}|}{\bar{p}_{24}^{true}} - (e_{day}) \right)^2 \tag{11}$$

$$\sigma_{e,week}^2 = \frac{1}{168} \sum_{h=1}^{168} \left(\frac{|p_h^{true} - p_h^{est}|}{\bar{p}_{168}^{true}} - (e_{week}) \right)^2. \tag{12}$$

The errors and error variances above are used in the case study below.

V. Case Study: Market of Mainland Spain 2002

The day-ahead electricity market of mainland Spain [23] is considered in this real-world case study. Price forecasting is performed using data of year 2002. Moreover, models similar to the one reported in this paper (not including the wavelet preprocessor) are routinely used by the power industry in Spain for price forecasting. It should be noted that the electricity market of mainland Spain is a duopoly with a dominant player. This results in price changes related to the strategic behavior of the dominant player, which are hard to predict.

To illustrate the behavior of the proposed technique, results comprising four weeks corresponding to the four seasons of year 2002 are presented. In this manner representative results for the whole year are provided.

For the sake of a fair comparison the fourth week of February, May, August, and September (months 2, 5, 8, and 11) are selected, i.e., weeks with particularly good price behavior are purposely not sought. This results in an uneven accuracy distribution throughout the year that reflects reality. It should be noted that accuracy for the winter and spring weeks is around 5% while for summer and fall weeks is around 10%. We believe those results are reasonably accurate for a study spanning one whole year. However, for shorter time horizons accuracy distribution can be improved.

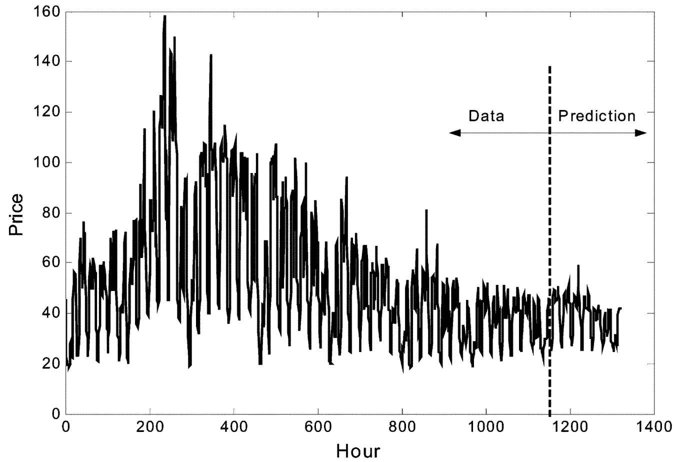

Fig. 3. Forecasting data for the winter season.

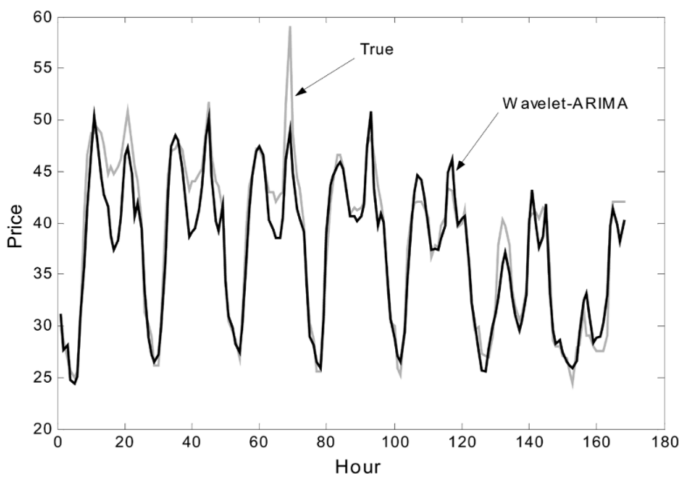

Fig. 4. Winter week: Actual prices and Wavelet-ARIMA estimates in € per megawatt hour.

To build the forecasting model for each one of the considered weeks, the information available includes hourly price historical data of the 48 days previous to the day of the week whose prices are to be predicted.

The winter week is February 18 to February 24; historical data available includes hourly prices from January 1 to February 17. The spring week is May 20 to May 26; historical data includes prices from April 2 to May 19. The summer week is August 19 to August 25; historical data includes prices from July 2 to August 18. The fall week is November 18 to November 24; historical data include prices from October 1 to November 17.

Fig. 3 provides the forecasting data for the winter season. Observe the unstable mean and variance presented by this series. Note that this unstable behavior makes forecasting hard.

For the winter week, Fig. 4 provides actual prices and price forecasts using the proposed technique. For the weeks of the other seasons, the corresponding plots are provided in Figs. 5, 6, and 7.

The prediction behavior of the wavelet-ARIMA technique for the winter week is very appropriate with a weekly error below 4.8%. Weekly error for the ARIMA technique is 6.3%. Only the mild spike (20% over the usual peak) of Wednesday evening

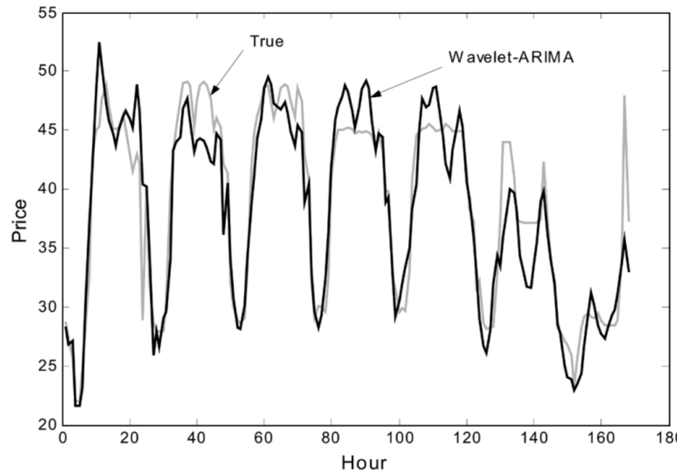

Fig. 5. Spring week: Actual prices and Wavelet-ARIMA estimates in € per megawatt hour.

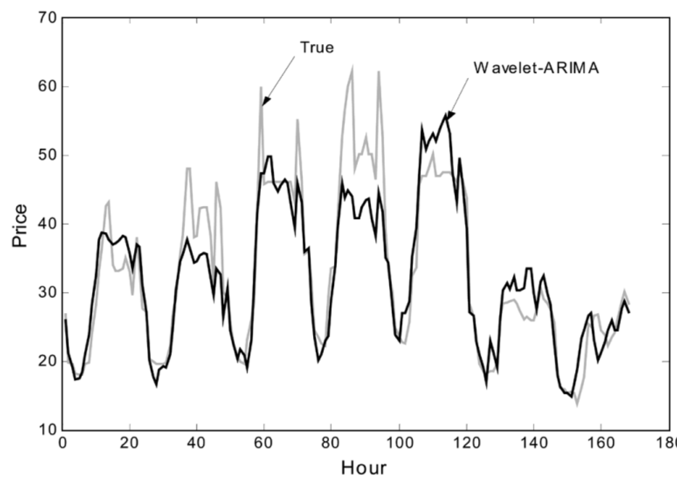

Fig. 6. Summer week: Actual prices and Wavelet-ARIMA estimates in € per megawatt hour.

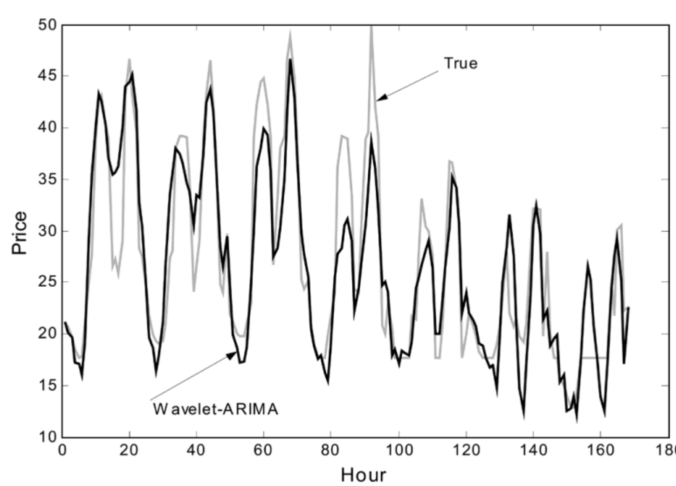

Fig. 7. Fall week: Actual prices and Wavelet-ARIMA estimates in € per megawatt hour.

is not properly reproduced. This spike is properly due to the strategic behavior of the dominant player of the market.

The performance of the wavelet-ARIMA technique for the spring week is accurate, with a weekly error below 5.7%.

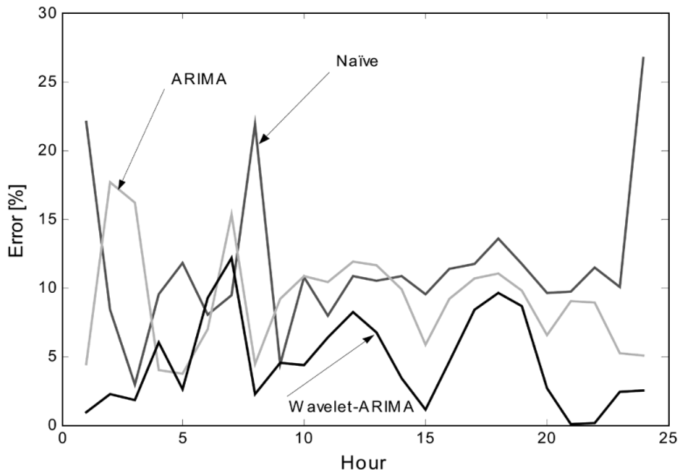

Fig. 8. Thursday of the spring week: Hourly percentage errors for (i) Naïve, (ii) ARIMA, and (iii) Wavelet-ARIMA forecasts, respectively.

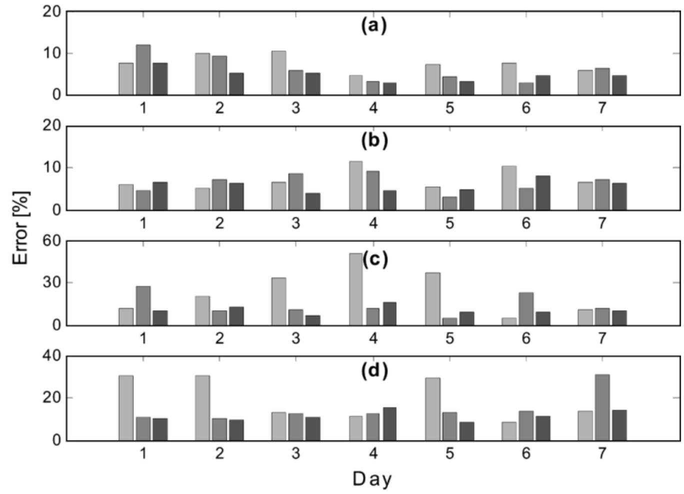

Fig. 9. Daily errors corresponding to Naïve, ARIMA, and Wavelet-ARIMA models for the weeks analyzed: (a) winter, (b) spring, (c) summer, and (d) fall.

Weekly error for the ARIMA technique is 6.4%. The prediction is more accurate during weekdays (where the bulk of profits lay) than during the weekend.

For the summer week, the prediction behavior of the wavelet-ARIMA technique is less accurate than for the winter and spring weeks. However, it is reasonable accurate with a weekly error below 10.7%. Weekly error for the ARIMA technique is 13.4%. Tuesday and Thursday peak are not accurately reproduced due to significant changes in prices between two consecutive hours within a day, two consecutive days within the considered week and with respect to the previous week. Note the significant prediction accuracy for the weekend.

As for the summer week, the performance of the wavelet-ARIMA technique for the fall week is not as good as for the winter and spring weeks. Nevertheless, accuracy is reasonable enough with a weekly error below 11.3%. Weekly error for the ARIMA technique is 13.8%. Observe that this week is particularly unstable in respect to price behavior, probably due to the strategic behavior of the dominant player in the market. The prediction is particular inaccurate for the morning and evening peaks of Thursday, and for Sunday morning.

In summary, the forecasting behavior of the proposed wavelet-ARIMA technique, as illustrated in Figs. 4–7, is appropriate. This technique outperforms ARIMA models in all considered weeks, which shows the usefulness and practical interest of the wavelet preprocessor proposed in this paper.

The forecasting behavior of the proposed technique, as illustrated in Figs. 4–7, is appropriate. Note that it mildly deteriorates in the summer week. Nevertheless, the proposed technique outperforms ARIMA models in all weeks.

For the Thursday of the spring week, Fig. 8 provides hourly percentage errors for the proposed Wavelet-ARIMA technique, the ARIMA model and the Naïve procedure. Error plots for the other days of this week and for the weeks of the other seasons are similar, and they are not plotted for the sake of conciseness. However, it should be noted that hourly errors for both the Wavelet-ARIMA and ARIMA techniques increase during weekends.

The four plots of Fig. 9 provide daily errors for the considered four weeks, using the Naïve, ARIMA and the proposed Wavelet-ARIMA technique, respectively.

Note that the performance of the Wavelet-ARIMA technique is generally better than the performance of the ARIMA and Naïve techniques. However, the superiority of the Wavelet-ARIMA technique is more apparent observing Tables I and II below.

Finally, Tables I and II provide weekly errors and error variances for the four weeks analyzed, which correspond to the four seasons of the year. For the sake of comparison, these errors are computed for the Naïve procedure, the ARIMA model and the proposed Wavelet-ARIMA technique.

Note that weekly errors (see Table I) and weekly error variances (see Table II) are smaller for the Wavelet-ARIMA technique than for the other techniques in all scenarios.

The CPU time required to forecast the 24 prices of the day-ahead market is below 5 minutes using a Dell PowerEdge 6600 with 2 processors at 1.60 GHz and 2 Gb of RAM memory. Although this is a significant amount of time, it is reasonable within a day-ahead decision-making framework.

Analyzing the different plots represented in Figs. 4–8, and Tables I and II, it can be concluded that the behavior of the proposed technique is superior to the behaviors of the ARIMA and the Naïve procedures. These experimental results confirm the intuition that the wavelet transform produces constitutive series that can be predicted more accurately than the original price series. The superior predicting behavior of the proposed technique is apparent in all weeks, as can be deducted observing Tables I and II.

VI. Conclusion

If the wavelet transform is applied to an ill-behaved price series (non constant mean and variance, outliers, and seasonal and calendar effects), the resulting constitutive series behave in general better than the original price series. The approximation series (main component of the transform) consistently presents a more stable variance than the original price series and no outliers.

TABLE I
WEEKLY FORECASTING ERRORS FOR THE FOUR WEEKS ANALYZED

Week	Weekly error [%]		
	Naïve	ARIMA	Wavelet-ARIMA
Winter	7.68	6.32	4.78
Spring	7.27	6.36	5.69
Summer	27.30	13.39	10.70
Fall	19.98	13.78	11.27

TABLE II
WEEKLY FORECASTING ERROR VARIANCES FOR THE FOUR WEEKS ANALYZED

Week	Error variances		
	Naïve	ARIMA	Wavelet-ARIMA
Winter	0.0034	0.0034	0.0019
Spring	0.0043	0.0020	0.0025
Summer	0.0738	0.0158	0.0108
Fall	0.0338	0.0157	0.0103

Therefore, the future values of the constitutive series can be forecast accurately, even using simple procedures such as ARIMA models. The application of the inverse wavelet transform to the predictions of the constitutive series allows producing accurate forecast of the original price series. The above has been illustrated using data of the day-ahead electric energy market of mainland Spain for year 2002. The functioning of the proposed technique, which is based on the wavelet transform and ARIMA models, outperforms the direct use of ARIMA models.

APPENDIX A
WAVELET TRANSFORM

This appendix describes the wavelet transform that works basically as follows.

1. Select a typical wavelet function and the origin of the series to be analyzed. This function is used to project the price series under analysis.
2. Project the series on the wavelet function and obtain the decomposition coefficient that provides the degree of similitude of the price series and the wavelet function.
3. Shift forward the wavelet function and repeat 1 and 2 until the original series has been fully analyzed.
4. Scale appropriately the wavelet function, and repeat steps 1 to 3. This allows generating a new set of decomposition coefficients for a lower resolution level.
5. Repeat Step 4 to consider all required levels of resolution.

The wavelet transform procedure is carried out considering a finite number of positions and resolution levels (discrete wavelet transform). Using this technique, the decomposition coefficients of the wavelet transform of the hourly price series are given by

$$\begin{aligned} \mathrm{p}_{\mathrm{mn}}^{\mathrm{W}} &= 2^{-(\mathrm{m}/2)} \sum_{\mathrm{t}=0}^{\mathrm{T}-1} \mathrm{p}_{\mathrm{t}} \mathrm{w}\left(\frac{\mathrm{t}-\mathrm{n}\cdot 2^{\mathrm{m}}}{2^{\mathrm{m}}}\right) \\ &= 2^{-(\mathrm{m}/2)} \sum_{\mathrm{t}=0}^{\mathrm{T}-1} \mathrm{p}_{\mathrm{t}} \mathrm{w}_{\mathrm{mn}}(\mathrm{t}) \end{aligned} \quad (13)$$

where $\mathrm{w}(\cdot)$ is the selected wavelet function, $\mathrm{p_t}$ is the value of the price at hour t, T is the length of the series, $\mathrm{p}_{\mathrm{mn}}^{\mathrm{W}}$ is the decomposition coefficient corresponding to position n and resolution level m. Note that if the number of observations, T, is divisible by 2^{m}, then the number of coefficients at each resolution level is $\mathrm{T}/2^{\mathrm{m}}$. To speed up calculations, expression (13) can be treated as a convolution, and the efficient Fast Fourier Transform used, as stated in [21].

An efficacious manner to select the wavelet functions is the multi-resolution technique based on using a father wavelet function and its complementary, a mother wavelet function. The father function allows deriving the low frequency components of the series, while the mother one allows deriving components of high frequency. Additionally, it is convenient to choose orthogonal wavelet functions due to their appropriate mathematical properties.

The so-called "approximation series", $\mathrm{A_m}$ $(\mathrm{m}=1,\ldots,\mathrm{M})$, and the "detail series", $\mathrm{D_m}$ $(\mathrm{m}=1,\ldots,\mathrm{M})$, are defined as

$$\mathrm{A_m} = \sum_{\mathrm{n}} \mathrm{p}_{\mathrm{mn}}^{\Phi} \varphi_{\mathrm{mn}}(\mathrm{t}); \quad \mathrm{m}=1,\ldots,\mathrm{M} \quad (14)$$

and

$$\mathrm{D_m} = \sum_{\mathrm{n}} \mathrm{p}_{\mathrm{mn}}^{\Psi} \psi_{\mathrm{mn}}(\mathrm{t}); \quad \mathrm{m}=1,\ldots,\mathrm{M} \quad (15)$$

where $\varphi_{\mathrm{mn}}(\mathrm{h})$ and $\psi_{\mathrm{mn}}(\mathrm{h})$ are the father and mother wavelet functions, and p_{mn}^{Φ} and $\mathrm{p}_{\mathrm{mn}}^{\Psi}$ are the coefficients obtained through (13). Note that $\mathrm{A_m}$ $(\mathrm{m}=1,\ldots,\mathrm{M})$ and $\mathrm{D_m}$ $(\mathrm{m}=1,\ldots,\mathrm{M})$ are series, i.e., $\mathrm{A_m}=\{\mathrm{A_{m1}},\ldots,\mathrm{A_{mT}}\}$ and $\mathrm{D_m}=\{\mathrm{D_{m1}},\ldots,\mathrm{D_{mT}}\}$.

It can be shown, see [21], that the expression of the original price series $\mathrm{p_h}$ $(\mathrm{h}=1,\ldots,\mathrm{T})$ can now be reconstructed by

$$\mathrm{p_h} = \mathrm{D_1} + \cdots + \mathrm{D}_M + \mathrm{A_M} \quad (16)$$

which is the denominated multiresolution decomposition of the price series $\mathrm{p_h}$ $(\mathrm{h}=1,\ldots,\mathrm{T})$.

In this paper we consider 3 detail series and the resulting approximation series. That is, $\mathrm{a_h}=\mathrm{D_1}, \mathrm{b_h}=\mathrm{D_2}, \mathrm{c_h}=\mathrm{D_3}$, and $\mathrm{d_h}=\mathrm{A_3}$.

Some comments are appropriate. First, Daubechies wavelets are the most commonly used in applications. For these families of wavelets, the smoothness increases as the order of the functions do; nevertheless, the support intervals also increase, which may deteriorate the prediction. Therefore, low order wavelet functions are generally advisable. Second, a long estimation period may originate inaccuracies because market conditions evolve with time. On the other hand, a short estimation period may originate volatile estimations. Therefore, an appropriate estimation period should be selected based on trial and error. Finally, it should be noted that efficient numerical implementations of the procedure above are readily available (see, e.g., [24]).

APPENDIX B
ARIMA FORECASTING

The standard statistical methodology to construct an ARIMA model includes the following.

Step 0) A class of models is formulated assuming certain hypotheses.
Step 1) A model is identified for the series considered.
Step 2) The parameters of the model are estimated.
Step 3) If the hypotheses of the model are validated, the procedure continues in Step 4; otherwise, the procedure continues in Step 1 to refine the model.
Step 4) The model is used to forecast.

These steps are briefly explained below.

Model Selection (Step 0). The proposed general ARIMA model has the form

$$\phi(\mathrm{B})\mathrm{p_h} = \mathrm{c} + \theta(\mathrm{B})\varepsilon_\mathrm{h} \quad (17)$$

where $\mathrm{p_h}$ is the price at hour h and ε_h is the error term. Polynomials $\phi(\mathrm{B})$ and $\theta(\mathrm{B})$ are functions of the back-shift operator B (observe that $\mathrm{B^s p_h} = \mathrm{p_{h-s}}$). That is, $\phi(\mathrm{B})\mathrm{p_h} = \mathrm{p_h} - \phi_1\mathrm{p_{h-1}} - \phi_2\mathrm{p_{h-2}} - \cdots - \phi_{\mathrm{n_F}}\mathrm{p_{h-n_F}}$, and ϕ_k $(\mathrm{k} = 1, \ldots, \mathrm{n_F})$ are polynomial coefficients, and $\theta(\mathrm{B})\varepsilon_\mathrm{h} = \varepsilon_\mathrm{h} - \theta_1\varepsilon_\mathrm{h-1} - \theta_2\varepsilon_\mathrm{h-2} - \cdots - \theta_{\mathrm{n_T}}\varepsilon_{\mathrm{h-n_T}}$, and θ_k $(\mathrm{k} = 1, \ldots, \mathrm{n_T})$ are polynomial coefficients.

The number of terms of the polynomial functions $\phi(\mathrm{B})$ and $\theta(\mathrm{B})$, $\mathrm{n_F}$ and $\mathrm{n_T}$, respectively, depends on the time series under analysis.

Note that including in the left hand side of expression (17) factors of the form $(1 - \mathrm{B^s})$ allows taking into account appropriately the seasonality effects. Finally, certain hypotheses on the error terms, ε_h, are needed to ensure the effectiveness of the predictions.

Model identification (Step 1). The target of this step is to identify which polynomial parameters should be estimated. The initial selection is based on the observation of the autocorrelation and partial autocorrelation plots [22]. Further refinement of the selection is based on physical knowledge and on engineering judgment.

Polynomial parameter estimation (Step 2). Once the parameters of the polynomials different from 0 have been identified (through plot observation, physical knowledge and engineering judgment), these parameters should be estimated. The estimation procedure is based on available historical data. Good estimators are usually found assuming that the data constitute observations of a stationary time series and maximizing the likelihood function with respect to the polynomial parameters. Good estimations can be obtained using commercial software, such as [25].

Validation of model hypotheses (Step 3). In this step, a diagnosis check is used to validate the model assumptions. If the estimated model is appropriate, then, the residuals (actual prices minus predicted prices) should behave in a manner consistent with the model. Residuals must satisfy the requirements of a white noise process: zero mean, constant variance, uncorrelation and normal distribution. If the hypotheses on the residuals are validated, then the corresponding model can be used to forecast prices and this step concludes successfully. Otherwise, the residuals contain a certain structure that should be analyzed to refine the model, and the procedure continues in Step 1. To refine the model a careful inspection of the autocorrelation and partial autocorrelation plots of the residuals is advisable.

Actual prediction (Step 4). In this step, the corresponding model from Step 2 is used to predict future values of prices, typically 24 hours ahead. It should be noted that prediction quality deteriorates as the predicted hour increases, i.e., the error of the estimate of hour 24 is typically greater than the error of the estimate of hour 1.

REFERENCES

[1] J. Contreras, R. Espínola, F. J. Nogales, and A. J. Conejo, "ARIMA models to predict next-day electricity prices," *IEEE Trans. Power Syst.*, vol. 18, no. 3, pp. 1014–1020, Aug. 2003.
[2] F. J. Nogales, J. Contreras, A. J. Conejo, and R. Espínola, "Forecasting next-day electricity prices by time series models," *IEEE Trans. Power Syst.*, vol. 17, no. 2, pp. 342–348, May 2002.
[3] O. B. Fosso, A. Gjelsvik, A. Haugstad, M. Birger, and I. Wangensteen *et al.*, "Generation scheduling in a deregulated system," *IEEE Trans. Power Syst.*, vol. 14, no. Feb., pp. 75–81, 1999.
[4] Z. Obradovic and K. Tomsovic, "Time series methods for forecasting electricity market pricing," in *Proc. IEEE Power Eng. Soc. Summer Meeting*, vol. 2, Edmonton, AB, Canada, 1999, pp. 1264–1265.
[5] J. Crespo, J. Hlouskova, S. Kossmeier, and M. Obersteiner, "Forecasting electricity spot prices using linear univariate time series models," *App. Energy*, vol. 77, no. 1, pp. 87–106, 2002.
[6] B. Ramsay and A. J. Wang, "A neural network based estimator for electricity spot-pricing with particular reference to weekend and public holidays," *Neurocomput.*, vol. 23, pp. 47–57, 1998.
[7] B. R. Szkuta, L. A. Sanabria, and T. S. Dillon, "Electricity price short-term forecasting using artificial neural networks," *IEEE Trans. Power Syst.*, vol. 14, no. 3, pp. 851–857, Aug. 1999.
[8] F. Gao, X. Guan, X.-R. Cao, and A. Papalexopoulos, "Forecasting power market clearing price and quantity using a neural network method," in *Proc. IEEE Power Eng. Soc. Summer Meeting*, vol. 4, Seattle, WA, 2000, pp. 2183–2188.
[9] P. Doulai and W. Cahill, "Short-term price forecasting in electric energy market," in *Proc. Int. Power Eng. Conf.*, Singapore, 2001, pp. 749–754.
[10] Y.-Y. Hong and C.-Y. Hsiao, "Locational marginal price forecasting in deregulated electricity markets using artificial intelligence," *Proc. Inst. Elect. Eng., Gen., Transm., Distrib.*, vol. 149, no. 5, pp. 621–626, Sep. 2002.
[11] L. Zhang and P. B. Luh, "Power market clearing price prediction and confidence interval estimation with fast neural network learning," in *Proc. IEEE Power Eng. Soc. Winter Meeting*, vol. 1, New York, 2002, pp. 268–273.
[12] L. Zhang, P. B. Luh, and K. Kasiviswanathan, "Energy clearing price prediction and confidence interval estimation with cascaded neural networks," *IEEE Trans. Power Syst.*, vol. 18, no. 1, pp. 99–105, Feb. 2003.
[13] C. P. Rodríguez and G. J. Anders, "Energy price forecasting in the Ontario competitive power system market," *IEEE Trans. Power Syst.*, vol. 19, no. 1, pp. 366–374, Feb. 2004.
[14] S. J. Yao and Y. H. Song, "Prediction of system marginal prices by wavelet transform and neural network," *Elect. Mach. Power Syst.*, vol. 28, no. 10, pp. 983–993, 2000.
[15] C.-I. Kim, I.-K. Yu, and Y. H. Song, "Prediction of system marginal price of electricity using wavelet transform analysis," *Energy Convers. Manage.*, vol. 43, pp. 1839–1851, 2002.
[16] G. Jau-Jia and P. B. Luh, "Market clearing price prediction using a committee machine with adaptive weighting coefficients," in *Proc. IEEE Power Eng. Soc. Winter Meeting*, vol. 1, New York, 2002, pp. 77–82.

[17] E. Ni and P. B. Luh, "Forecasting power market clearing price and its discrete PDF using a Bayesian-based classification method," in *Proc. IEEE Power Eng. Soc. Winter Meeting*, vol. 3, Columbus, OH, 2001, pp. 1518–1523.
[18] D. W. Bunn, "Forecasting loads and prices in competitive power markets," *Proc. IEEE*, vol. 88, no. 2, pp. 163–169, Feb. 2000.
[19] A. Angelus, "Electricity price forecasting in deregulated markets," *Elect. J.*, vol. 14, no. 3, pp. 32–41, 2001.
[20] A. M. Breipohl, "Electricity price forecasting models," in *Proc. IEEE Power Eng. Soc. Winter Meeting*, vol. 2, New York, 2002, pp. 963–966.
[21] Y. Nievergelt, *Wavelets Made Easy*. Cambridge, MA: Birkhäuser, 1999.
[22] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting & Control*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1994.
[23] Market operator of the electricity market of mainland Spain, OMEL. (2004). [Online]. Available: http://www.omel.es/
[24] The Math Works, MATLAB. (2004). [Online]. Available: http://www.mathworks.com/
[25] Forecasting and time series software, Scientific Computing Associates (SCA). (2004). [Online]. Available: http://www.scausa.com/

Antonio J. Conejo (F'04) received the M.S. degree from Massachusetts Institute of Technology, Cambridge, MA, in 1987 and the Ph.D. degree from the Royal Institute of Technology, Stockholm, Sweden, in 1990.

He is currently Professor of Electrical Engineering at the Universidad de Castilla–La Mancha, Ciudad Real, Spain. His research interests include control, operations, planning and economics of electric energy systems, as well as statistics and optimization theory and its applications.

Miguel A. Plazas received the Industrial Engineering degree from the Universidad Politécnica de Valencia, Valencia, Spain, in 1995.

He is currently working in electrical markets as a trader for Unión Fenosa Generación.

His main research interest is optimization under uncertainty applied to electricity markets.

Rosa Espínola (S'02) received the B.S. degree in statistics from Universidad de Granada, Granada, Spain, in 1999. She is currently working toward the Ph.D. degree at Universidad de Castilla–La Mancha, Ciudad Real, Spain.

Her research interests include planning and economics of power systems, forecasting, and time series analysis.

Ana B. Molina received the Ingeniero Industrial degree from the Universidad de Castilla–La Mancha, Ciudad Real, Spain, in 2004.

Her research interests include control, operations, planning, and economics of electric energy systems.