DAY-AHEAD ELECTRICITY PRICE FORECASTING BASED ON TIME SERIES MODELS: A COMPARISON

Rosario Espínola, Javier Contreras, Francisco J. Nogales and Antonio J. Conejo
E.T.S. de Ingenieros Industriales, Universidad de Castilla – La Mancha
13071 Ciudad Real, Spain
Rosa.Espinola@uclm.es, Javier.Contreras@uclm.es, FcoJavier.Nogales@uclm.es, Antonio.Conejo@uclm.es

Abstract – This paper presents and compares three price forecasting tools for day-ahead electricity markets: dynamic regression, transfer function and seasonal Auto Regressive Integrated Moving Average models. The three procedures are based on time series analysis and differ when modelling the relationship between prices and error terms (error measured by the difference between the actual price and the one predicted by the model). The dynamic regression model relates the current price to the values of past prices. The transfer function approach relates the current price to the values of present and past prices and demands. This relationship can include a serially correlated error that can be modelled by an Auto Regressive Moving Average process. Finally, the third model relates current prices to the values of past prices, and current error terms to previous errors. Real world case studies from mainland Spain and California electricity markets are presented to illustrate and compare the predictive behaviour of the models.

Keywords: Electricity markets, market clearing price, forecasting, time series analysis, ARIMA models

1 INTRODUCTION

The electric power industry is becoming more sophisticated after a few years of deregulation and restructuring. Electricity markets have emerged from the previous centralized operation in order to supply energy to consumers with the target of attaining high reliability and low cost.

Electricity price forecasting has become an essential tool in competitive electricity markets, both for producers and consumers. The reason is that buying and selling bidding strategies rely on next day price predictions in order to achieve benefit or utility maximization (for buying or selling agents, respectively) [1, 2]. In addition, reliable price forecasts have a definitive impact not only in electricity day-ahead markets, but also in monthly schedules and bilateral or financial contracts. For this type of portfolio decisions, it is desirable to have available predictions of price average values over a year horizon. Also, Energy Service Companies (ESCOs) buy energy from the pool or from contracts to sell it to their clients. These companies need reliable short- and long-term price forecasts.

In the past, several time series techniques have been used to predict demand in centralized markets [3]. However, there are not many applications of time series forecasting to predict day-ahead electricity prices. The most important ones are Auto Regressive Integrated Moving Average (ARIMA) methods [4] and Artificial Neural Networks (ANN) [5, 6, 7]. Fosso et al. [4] use an ARMA model to forecast prices in the Nordic market. Ramsay et al. [5] propose a hybrid fuzzy logic-neural network approach to predict prices in the England-Wales pool, with daily mean errors around 10%. Szkuta et al. [6] propose a three-layered ANN using back-propagation in the Victorian (Australia) electricity market, with daily mean errors around 15%. And Nicolaisen et al. [7] use Fourier and Hartley transforms as filters to ANNs.

This paper focuses on short-term decisions associated to the pool and presents three highly effective tools to predict day-ahead prices: dynamic regression, transfer function and seasonal Auto Regressive Integrated Moving Average (ARIMA) models. They are based on time series analysis and applied to forecast actual prices of mainland Spain [8] and California [9].

The remainder of the paper is organized as follows. In section 2, a mathematical description of the three models is provided. Section 3 presents numerical results and section 4 states some conclusions.

2 DESCRIPTION OF THE MODELS

In this section, the description of three models based on time series analysis is presented: dynamic regression, transfer function and ARIMA formulations. These three models are a class of stochastic processes used to analyze time series, and have a common methodology. The application of this methodology to the study of time series analysis is due to Box and Jenkins [10].

The analysis is based on setting up a hypothetical probability model, one for each of the proposed models, to represent the data. The models presented are selected based on a careful inspection of the main characteristics of the hourly price series. In most competitive electricity markets this series presents: high frequency; non-constant mean and variance; multiple seasonality (corresponding to a daily and weekly periodicity, respectively); calendar effect (such as weekends, holidays); high volatility; and high percentage of unusual prices (mainly in periods of high demand).

Moreover, the proposed models can include explanatory variables, for example, the demand of electricity has been included in one model because, a priori, it seems to partly explain the price behavior.

Next, the description of the general statistical methodology to build a final model is presented. The three models have been obtained through the following scheme:

Step 0. A class of models is formulated assuming certain hypotheses.

Step 1. A model is identified for the observed data.

Step 2. The model parameters are estimated.

Step 3. If the hypotheses of the model are validated go to Step 4, otherwise go to Step 1 to refine the model.

Step 4. The model can be used to forecast.

In the following subsections, each step of the above scheme is detailed.

2.1 Step 0

The building of each of the three models presented differs in this step. This step is explained below depending on the selected model.

2.1.1 Dynamic regression

The first proposed method to forecast prices is a *dynamic regression model* [10, 11]. In this model, the price at hour t is related to the values of past prices at hours $t-1$, $t-2$, …, etc. This is done to obtain a model that has uncorrelated errors.

In Step 0, the selected model used to explain the price at hour t is the following:

$$p_t = c + \omega^p(B)\, p_t + \varepsilon_t \qquad (1)$$

where p_t is the price at time t and c is a constant. Function $\omega^p(B) = \sum_{l=1}^{K} \omega_l^p B^l$ is a polynomial function of the backshift operator $B : B^l p_t = p_{t-l}$, where the total number of terms of the function, K, is subject to change in steps 1-3. Function $\omega^p(B)$ depends on parameters ω_l^p, whose values are estimated in Step 1. Finally, ε_t is the error term. In Step 0 this term is assumed to be a series drawn randomly from a normal distribution with zero mean and constant variance σ^2, that is, a white noise process.

The efficiency of this approach depends on the election of the appropriate parameters in $\omega^p(B)$ to achieve an uncorrelated set of errors. This selection is carried out through Steps 1 to 3 as it is explained below.

2.1.2 Transfer function

A second proposed method that includes a serially correlated error is called *transfer function model* [10, 11]. Specifically, it is assumed that the price and demand series are both stationary (i.e. with constant mean and variance). The general form proposed to model the (price, demand) transfer function is

$$p_t = c + \omega^d(B)\, d_t + N_t \qquad (2)$$

where p_t is the price at time t, c is a constant, d_t is the demand at time t, $\omega^d(B) = \sum_{l=0}^{K} \omega_l^d B^l$ is a polynomial function of the backshift operator, and N_t is a disturbance term that follows an ARMA model of the form

$$N_t = \frac{\theta(B)}{\phi(B)} \varepsilon_t \qquad (3)$$

with $\theta(B) = 1 - \sum_{l=1}^{\Theta} \theta_l B^l$ and $\phi(B) = 1 - \sum_{l=1}^{\Phi} \phi_l B^l$, both of which being polynomial functions of the backshift operator. The total number of terms of the functions $\theta(B)$ and $\phi(B)$, Θ and Φ, respectively, are subject to change in steps 1-3. Finally, ε_t is the error term, that is assumed to be a white noise process.

The model in (2) relates actual prices to demands through function $\omega^d(B)$ and actual prices to past prices through function $\phi(B)$.

2.1.3 ARIMA

The third proposed method is an ARIMA formulation. The proposed general ARIMA formulation in Step 0 is the following:

$$\phi(B)\, p_t = \theta(B)\, \varepsilon_t \qquad (4)$$

where $\phi(B)$ and $\theta(B)$ are functions of the backshift operator such as in (3), and ε_t is the error term. But in this case, functions $\phi(B)$ and $\theta(B)$ have special forms. They can contain factors of polynomial functions of the form

$$\phi(B) = 1 - \sum_{l=1}^{\Phi} \phi_l B^l$$

and/or $\theta(B) = 1 - \sum_{l=1}^{\Theta} \theta_l B^l$, and/or $(1 - B^s)$, where several values of ϕ_l and θ_l can be set to 0. The total number of terms of the functions $\theta(B)$ and $\phi(B)$, Θ and Φ, respectively, are subject to change in steps 1-3. For example, function $\phi(B)$ could have the following form:

$$\begin{aligned}\phi(B) &= (1 - \phi_1 B^1 - \phi_2 B^2)(1 - \phi_{24} B^{24} - \phi_{48} B^{48}) \\ &(1 - \phi_{168} B^{168})(1 - B)(1 - B^{24})\end{aligned} \qquad (5)$$

It should be noted that this example does not correspond to a standard ARIMA formulation, as presented in [10]. However, the model in (4) is sufficiently general to include the main characteristics of the price data. For example, to include multiple seasonality, factors of the form $(1-\phi_{24}B^{24})$, $(1-\phi_{168}B^{168})$, and/or $(1-\theta_{24}B^{24})$, $(1-\theta_{168}B^{168})$, and perhaps $(1-B^{24})$, $(1-B^{168})$ can be added to the model.

It can be observed that the model in (4) relates actual prices to past prices through function $\phi(B)$, and actual errors to past errors through function $\theta(B)$.

Finally, certain hypotheses on the three models must be assumed. These hypotheses are imposed on the error term, ε_t. In Step 0, this term is assumed to be a randomly drawn series from a normal distribution with zero mean and constant variance σ^2. In Step 3, a diagnostic checking is used to validate these model assumptions, as explained in subsection 2.4.

2.2 Step 1

Depending on the selected approach, a trial model must be identified for the price data. First, in order to make the underlying process stationary (more homogeneous mean and variance), a transformation of the original price data may be necessary. In this step, if a logarithmic transformation is applied to the price data, a more stable variance is attained for the three models.

For the dynamic regression and the transfer function approaches, all the initial parameters are set to zero. Particularly, in the ARIMA formulation, the inclusion of factors of the form $(1-B^s)$ may be necessary to make the process more stationary. And, to attain a more stable mean, factors of the form $(1-B)$, $(1-B^{24})$, $(1-B^{168})$ may be necessary, depending on the particular type of electricity market, as explained at the end of this section. The initial selected parameters for the ARIMA formulation are based on the observation of the autocorrelation and partial autocorrelation plots. In successive trials, the same observation of the residuals obtained in Step 3 (observed values minus predicted values) can refine the structure of the functions in the model.

2.3 Step 2

After the functions of the models have been specified, the parameters of these functions must be estimated. Good estimators of the parameters can be found by assuming that the data are observations of a stationary time series (Step 1), and by maximizing the likelihood with respect to the parameters [10].

The SCA System [12] is used to estimate the parameters of the corresponding model in the previous step. The parameter estimation is based on maximizing a likelihood function for the available data. A conditional likelihood function is selected in order to get a good starting point to obtain an exact likelihood function, as described in [10].

2.4 Step 3

In this step, a diagnosis check is used to validate the model assumptions of Step 0. This diagnosis checks if the hypotheses made on the residuals (actual prices minus predicted prices by estimated model in Step 1) are true. Residuals must satisfy the requirements of a white noise process: zero mean, constant variance, uncorrelated process and normal distribution. These requirements can be checked by taking tests for randomness, such as the one based on the Ljung-Box statistic, and observing plots, such the autocorrelation and partial autocorrelations plots.

If the hypotheses on the residuals are validated by tests and plots, then, the corresponding model can be used to forecast prices. Otherwise, the residuals contain a certain structure that should be studied to refine the model in Step 1. This study is based on a careful inspection of the autocorrelation and partial autocorrelation plots of the residuals.

2.5 Step 4

In Step 4, the corresponding model from Step 2 can be used to predict future values of prices (typically 24 hours ahead). Due to this requirement, difficulties may arise because predictions can be less certain as the forecast lead time becomes larger.

The SCA System is again used to compute the 24-hour forecast. The exact likelihood function is selected to obtain a very accurate prediction.

As a result of these five steps, the final models for the Spanish and Californian electricity markets are the following:

2.5.1 Dynamic regression

Final selected parameters ω_l^p of function $\omega^p(B)$ in (1) that are different from zero are those corresponding to indices l = 1, 2, 3, 24, 25, 48, 49, 72, 73, 96, 97, 120, 121, 144, 145, 168, 169, 192, 193.

2.5.2 Transfer function

Final selected parameters different from zero for function $\omega^d(B)$ in (2) are those corresponding to indices l = 0, 1, 2, 3, 24, 25, 48, 49, 72, 73, 96, 97, 120, 121, 144, 145, 168, 169, 192, 193. Selected parameters ϕ_l of function $\phi(B)$ in (2) that are different from zero are those corresponding to indices l = 1, 2, 3, 24, 25, 48, 49, 72, 73, 96, 97, 120, 121, 144, 145, 168, 169, 192, 193. Due to seasonality, function $\theta(B)$ has been divided into two functions $\theta(B) = \theta_1(B)\,\theta_2(B)$ where $\theta_1(B) = 1 - \sum_{l=1}^{\Theta} \theta_{1l} B^l$ and $\theta_2(B) = 1 - \sum_{l=1}^{\Theta} \theta_{2l} B^l$. Selected parameters θ_{1l} for the first factor are different from zero for indices l = 1, 2, 3, 24. The only parameter different from zero for the second factor is at index l = 168.

2.5.3 ARIMA

Final models for the Spanish and Californian electricity markets are, respectively:

$$
\begin{aligned}
&(1-\phi_1 B^1 - \phi_2 B^2 - \phi_3 B^3 - \phi_4 B^4 - \phi_5 B^5) \\
&(1-\phi_{23} B^{23} - \phi_{24} B^{24} - \phi_{47} B^{47} - \phi_{48} B^{48} - \\
&-\phi_{72} B^{72} - \phi_{96} B^{96} - \phi_{120} B^{120} - \phi_{144} B^{144}) \\
&(1-\phi_{168} B^{168} - \phi_{336} B^{336} - \phi_{504} B^{504}) \log p_t = \\
&c + (1-\theta_1 B^1 - \theta_2 B^2)(1-\theta_{24} B^{24}) \\
&(1-\theta_{168} B^{168} - \theta_{336} B^{336} - \theta_{504} B^{504})\varepsilon_t
\end{aligned} \tag{6}
$$

$$
\begin{aligned}
&(1-\phi_1 B^1 - \phi_2 B^2)(1-\phi_{23} B^{23} - \phi_{24} B^{24} - \phi_{47} B^{47} - \\
&-\phi_{48} B^{48} - \phi_{72} B^{72} - \phi_{96} B^{96} - \phi_{120} B^{120} - \phi_{144} B^{144}) \\
&(1-\phi_{167} B^{167} - \phi_{168} B^{168} - \phi_{169} B^{169} - \phi_{192} B^{192}) \\
&(1-B)(1-B^{24})(1-B^{168}) \log p_t = c + (1-\theta_1 B^1 - \theta_2 B^2) \\
&(1-\theta_{24} B^{24} - \theta_{48} B^{48} - \theta_{72} B^{72} - \theta_{96} B^{96})(1-\theta_{144} B^{144}) \\
&(1-\theta_{168} B^{168} - \theta_{336} B^{336} - \theta_{504} B^{504})\varepsilon_t
\end{aligned} \tag{7}
$$

Note that, as mentioned in Step 0, the proposed formulation extends the standard ARIMA model by including more than two factors in (6) and (7), and a special polynomial structure of the overall function.

It should be noted that model (6) needs the previous 5 hours to predict the next hour, whereas (7) just needs the previous two hours. Also, the model in (6) does not use differentiation, and the one in (7) uses hourly, daily and weekly differentiations: $(1-B)(1-B^{24})(1-B^{168})$. This is related to the stationarity property of the series, and it can be traced by inspecting the autocorrelation and partial autocorrelation plots.

3 NUMERICAL RESULTS

3.1 Case Studies

The three models described in the previous section have been applied to predict the electricity prices of mainland Spain and Californian markets.

For the Spanish electricity market two weeks have been selected to forecast and validate the performance for each of the three models. The first week corresponds to the second week of May 2001 (from days May 11th to 17th), which is typically a high demand week. The second one corresponds to the fourth week of August 2001 (from 25th to 31st) which is a typically a low demand week. The hourly data used to forecast the first week are from January 1st to May 10th, 2001; and the hourly data used to forecast the second week are from June 1st to August 24th, 2001.

For the Californian electricity market, the week of April 3rd to 9th 2000 has been chosen. This week is prior in time to the beginning of the dramatic price volatility period. The hourly data used to forecast this week are from January 1st to April 2nd, 2000.

The ARIMA models (6)-(7), corresponding to the Spanish and Californian markets, are used in the three proposed study cases. The dynamic regression and transfer function models are unique for both markets and all the case studies.

3.2 Numerical Results

Numerical results for the three proposed models are presented. Fig. 1 to 9 show the forecasted prices for each of the three models and for each of the three weeks studied, two for the Spanish electricity market and one for the Californian market, together with the actual prices.

Figure 1 corresponds to the selected week of May using a dynamic regression model for the Spanish market.

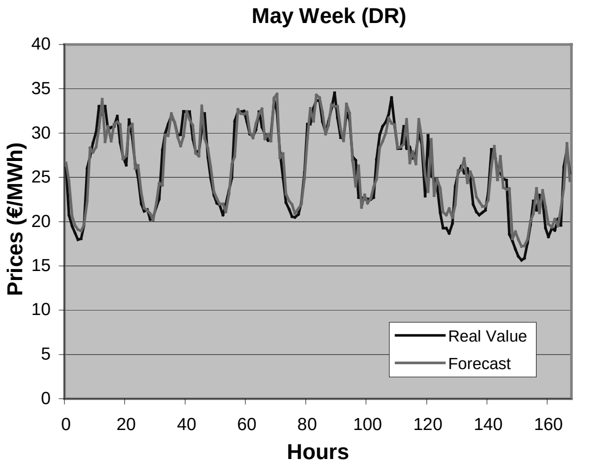

Figure 1: Forecast of May week using a dynamic regression model in the Spanish market. Prices in €/MWh.

The seven daily mean errors for this week appear in Table 1. The daily mean errors are around 5%. It can be observed that the hours with higher prediction errors are those corresponding to weekend days (Saturday and Sunday).

Days	1	2	3	4	5	6	7
Mean(%)	5.24	3.94	3.06	3.34	5.19	7.23	7.55

Table 1: Daily mean errors of May week using a dynamic regression model in the Spanish market.

Figure 2 corresponds to the selected week of May using a transfer function model for the Spanish market.

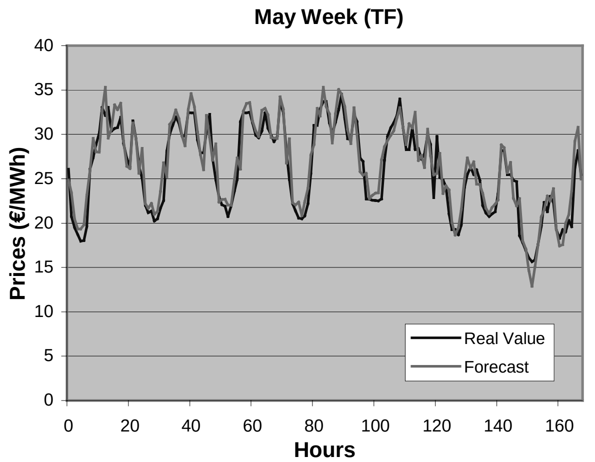

Figure 2: Forecast of May week using a transfer function model in the Spanish market. Prices in €/MWh.

The seven daily mean errors for this week are shown in Table 2. The daily mean errors are around 5.2%. On Sunday the error is greater.

Days	1	2	3	4	5	6	7
Mean(%)	5.50	4.82	4.07	4.85	5.09	4.81	7.25

Table 2: Daily mean errors of May week using a transfer function model in the Spanish market.

Figure 3 corresponds to the same week in May using the ARIMA model (6) for the Spanish market.

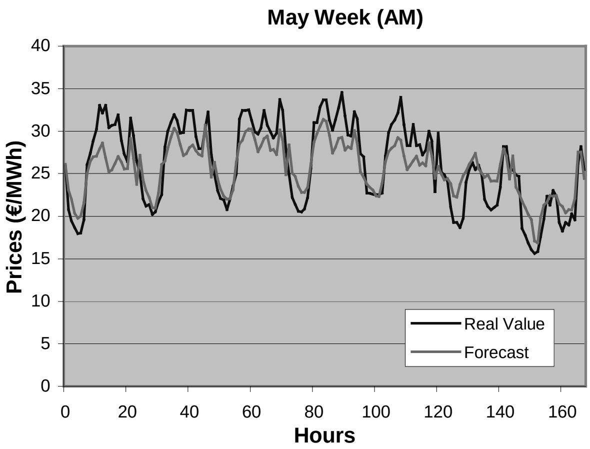

Figure 3: Forecast of May week using an ARIMA model in the Spanish market. Prices in €/MWh.

The seven daily mean errors for this week appear in Table 3. The daily mean errors are around 8%.

Days	1	2	3	4	5	6	7
Mean(%)	10.18	7.35	6.47	8.96	6.78	9.53	8.94

Table 3: Daily mean errors of May week using an ARIMA model in the Spanish market.

Figure 4 corresponds to the selected week of August using a dynamic regression model for the Spanish market.

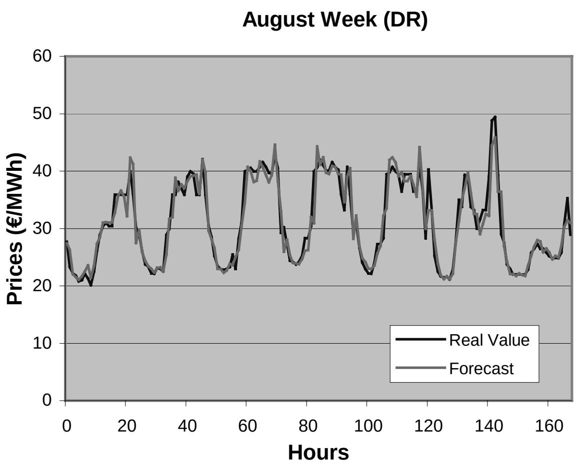

Figure 4: Forecast of August week using a dynamic regression model in the Spanish market. Prices in €/MWh.

The seven daily mean errors for this week appear in Table 4. The daily mean errors are around 4.5%.

Days	1	2	3	4	5	6	7
Mean(%)	4.58	3.61	4.18	5.14	4.62	5.84	3.62

Table 4: Daily mean errors of August week using a dynamic regression model in the Spanish market.

Figure 5 corresponds to the same week of August using a transfer function model for the Spanish market.

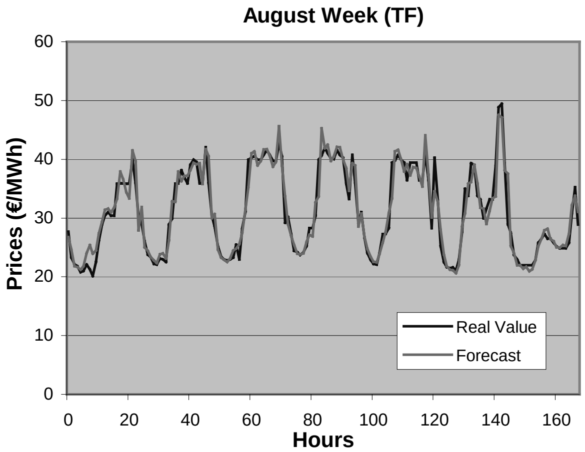

Figure 5: Forecast of August week using a transfer function model in the Spanish market. Prices in €/MWh.

The seven daily mean errors for this week appear in Table 5. The daily mean errors are around 4.4%. A good performance of the prediction method can be observed.

Days	1	2	3	4	5	6	7
Mean(%)	5.81	4.32	3.60	3.97	3.78	5.31	3.96

Table 5: Daily mean errors of August week using a transfer function model in the Spanish market.

Figure 6 corresponds to the same week of August using the ARIMA model (6) for the Spanish market.

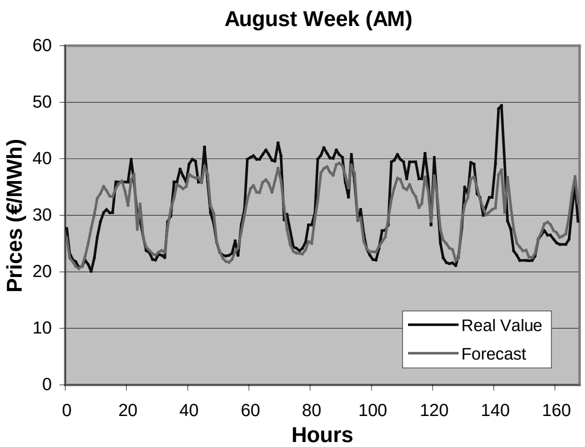

Figure 6: Forecast of August week using an ARIMA model in the Spanish market. Prices in €/MWh.

The seven daily mean errors for this week appear in Table 6. The daily mean errors are around 7.8%.

Days	1	2	3	4	5	6	7
Mean(%)	10.03	4.34	8.67	6.15	7.91	9.31	8.16

Table 6: Daily mean errors of August week using an ARIMA model in the Spanish market.

Figure 7 corresponds to the selected week in April using a dynamic regression model for the Californian market.

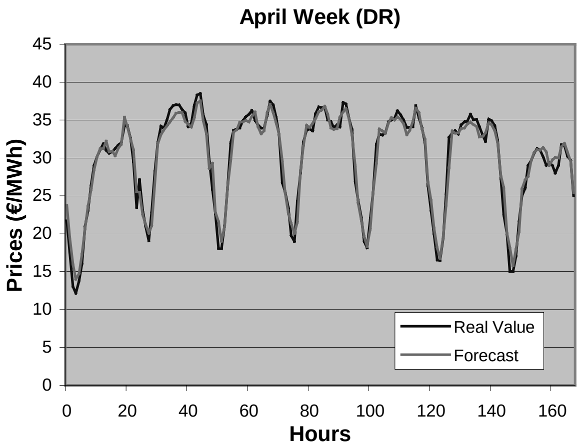

Figure 7: Forecast of April week using a dynamic regression model in the Californian market. Prices in \$/MWh.

The seven daily mean errors for this week appear in Table 7. The daily mean errors are around 3.3%.

Days	1	2	3	4	5	6	7
Mean(%)	4.68	3.12	3.51	2.8	1.96	2.89	4.14

Table 7: Daily mean errors of April week using a dynamic regression model in the Californian market.

Figure 8 corresponds to the same week in April using a transfer function model for the Californian market.

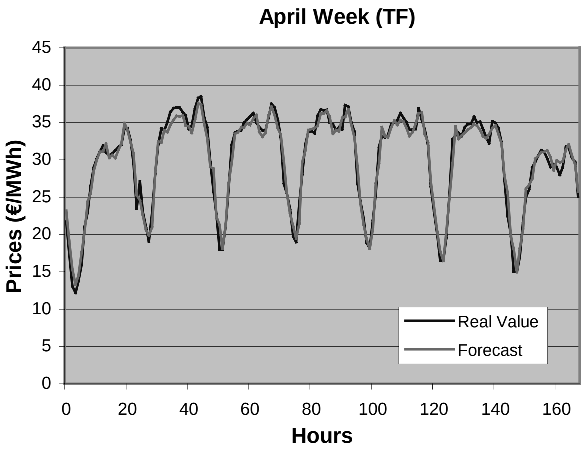

Figure 8: Forecast of April week using a transfer function model in the Californian market. Prices in \$/MWh.

The seven daily mean errors for this week appear in Table 8. The daily mean errors are around 3.25%.

Days	1	2	3	4	5	6	7
Mean(%)	4.27	3.32	3.39	2.97	2.36	2.61	3.87

Table 8: Daily mean errors of April week using a transfer function model in the Californian market.

Figure 9 corresponds to the same week in April using the ARIMA model (7) for the Californian market.

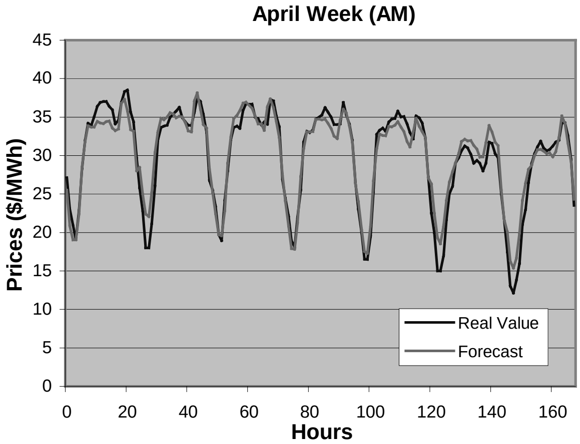

Figure 9: Forecast of April week using an ARIMA model in the Californian market. Prices in \$/MWh.

The seven daily mean errors for this week appear in Table 9. The daily mean errors are around 5%. During the weekend the errors are greater.

Days	1	2	3	4	5	6	7
Mean(%)	4.35	6.17	2.6	2.53	3.57	8.46	7.44

Table 9: Daily mean errors of April week using an ARIMA model in the Californian market.

Several statistical measures have been used to verify the prediction ability of the proposed models.

For all the study cases, the average prediction error of the 24 hours has been computed for each day. Then, the average of the daily mean errors has been calculated and called Mean Week Error (MWE). Finally, the Forecasted Mean Square Error (FMSE) for the 168 hours of each week has been calculated as:

$$FMSE = \sum_{i=1}^{168} (p_t - \hat{p}_t)^2 \tag{8}$$

where p_t and $\hat{p}_t$ are the actual and forecasted prices, respectively.

An index of uncertainty in the models is the variability of what is still unexplained after fitting the models. That can be measured through the variance of the error term (σ^2). The smaller σ^2 the more precise the prediction of prices. Normally, the value of σ is not known, thus an estimate is used instead. The standard deviation of the error terms, $\hat{s}_R$, can be used as such an estimate. This estimate is useful when the true values of the series are unknown.

These measures can be observed in Table 10. First column indicates the month, and in parenthesis the model used: DR for Dynamic Regression, TF for Transfer Function and AM for ARIMA model. The second column shows the percentage Mean Week Error (MWE), the third one presents the standard deviation of the error terms ($\hat{s}_R$), and the fourth column shows the square root of the Forecasted Mean Square Error (FMSE).

Note that prices, $\hat{s}_R$, and $\sqrt{FMSE}$ are measured in €/MWh and \$/MWh in the Spanish and Californian markets, respectively.

	MWE (%)	$\hat{s}_R$	$\sqrt{FMSE}$
Spain May 2001 (DR)	5.07	0.223	21.48
Spain May 2001(TF)	5.19	0.210	22.15
Spain May 2001 (AM)	8.31	0.215	32.52
Spain August 2001 (DR)	4.51	0.090	28.05
Spain August 2001(TF)	4.39	0.081	25.61
Spain August 2001(AM)	7.79	0.081	42.99
California April 2000 (DR)	3.30	0.056	14.94
California April 2000 (TF)	3.25	0.055	14.51
California April 2000 (AM)	5.01	0.060	21.19

Table 10: Statistical Measures.

Finally, Tables 11 and 12 show the MWE for the last week of the first ten months of the year 2000 in Spain, and the same week for the months of August and November in California. After April 2000 the Californian market entered a highly unstable period that later provoked the collapse at the end of the year. Explanatory variables such as: demand, water storage and the available production of hydro units, are considered in the ARIMA case.

MWE (%)	DR	TF	ARIMA	ARIMA-EX[1]
January 2000	6.93	6.85	12.06	9.97
February 2000	5.54	4.73	8.05	8.13
March 2000	7	7.11	11.28	10.5
April 2000	6.08	5.38	19.37	14.68
May 2000	3.81	2.76	4.99	7.75
June 2000	6.87	6.69	9.97	10.8
July 2000	5.49	5.7	9.39	8.83
August 2000	5.1	5.17	8.17	9.39
September 2000	6.81	7.05	12.01	10.72
October 2000	8.06	7.08	13.63	13.69

Table 11: Mean Week Error for the last week of the first ten months of 2000 in the Spanish Market.

MWE (%)	DR	TF	ARIMA	ARIMA-EX[2]
August 2000	7.08	7.07	15.65	21.03
November 2000	4.63	4.68	13.6	13.68

Table 12: Mean Week Error for the last week of August and November, 2000 in the Californian Market.

[1] With explanatory variables: demand, water storage and available hydro production.

[2] With explanatory variable: demand.

All the study cases have been run on a DELL Precision 620 Workstation with two processors Pentium III, 1 Gb of RAM, and 800 MHz. Running time, including estimation and forecasting, has been under three minutes in all cases.

4 CONCLUSIONS

This paper has proposed three forecasting models: dynamic regression, transfer function and ARIMA, to predict hourly electricity prices in the Spanish and Californian day-ahead markets. These models are based on time series analysis. For all markets and case studies, the dynamic regression and transfer function models have performed better than the ARIMA model, though the three techniques provide reasonable predictions. The difference between the ARIMA model and the other two may be due to the lack of flexibility of the ARIMA formulation when including multiple seasonality terms.

The effect of choosing a low versus a high demand week has been negligible to the forecasts. However, the effect of the explanatory variables has improved the ARIMA predictions, albeit not always.

The forecasted prices have shown better behavior in the Californian market before the crash that took place in the summer of 2000. This could be due to the fact that that market showed less volatility and a lower proportion of outliers before the turmoil.

The Spanish model has needed 5 hours to predict future prices, instead of the 2 hours needed in the Californian one. These differences may reflect different bidding behaviors and that will be subject of future research.

ACKNOWLEDGMENT

This work was partially supported by the Ministry of Science and Technology (Spain) and the European Union through grant FEDER-CICYT 1FD97-1598. We are also grateful to Mr. Bill Lattyak, from SCA, for his help and advice.

REFERENCES

[1] J. M. Arroyo and A. J. Conejo, "Optimal Response of a Thermal Unit to an Electricity Spot Market," IEEE Transactions on Power Systems, vol. 15, no. 3, pp. 1098-1104, August 2000.

[2] A. J. Conejo, J. Contreras, J. M. Arroyo and S. de la Torre, " Optimal Response of an Oligopolistic Generating Company to a Competitive Pool-Based Electric Power Market," to appear in IEEE Transactions on Power Systems, paper 2000TR225.

[3] G. Gross and F. D. Galiana, "Short-Term Load Forecasting," Proc. of the IEEE, vol. 75, no. 12, pp. 1558-1573, December 1987.

[4] O. B. Fosso, A. Gjelsvik, A. Haugstad, M. Birger and I. Wangensteen, "Generation Scheduling in a Deregulated System," IEEE Transactions on Power Systems, vol. 14, no. 1, pp. 75-81, February 1999.

[5] B. Ramsay and A. J. Wang, "A Neural Network Based Estimator for Electricity Spot-Pricing with Particular Reference to Weekend and Public Holidays," Neurocomputing, vol. 23, pp. 47-57, 1998.

[6] B. R. Szkuta, L. A. Sanabria and T. S. Dillon, "Electricity Price Short-Term Forecasting using Artificial Neural Networks," IEEE Transactions on Power Systems, vol. 14, no. 3, pp. 851-857, August 1999.

[7] J. D. Nicolaisen, C. W. Richter, Jr. and G. B. Sheblé, "Signal Analysis for Competitive Electric Generation Companies," Proc. of the Conference on Electric Utility Deregulation and Restructuring, London, United Kingdom, 4-7 April, 2000.

[8] J. J. González and P. Basagoiti, "Spanish Power Exchange and Information System Design Concepts, and Operating Experience," Proc. of the 21st Power Industry Computer Applications International Conference, PICA'99, Santa Clara, USA, May 1999.

[9] F. Albuyeh and Z. Alaywan, "Implementation of the California Independent System Operator," Proc. of the 21st Power Industry Computer Applications International Conference, PICA'99, Santa Clara, USA, May 1999.

[10] G. E. P. Box and G. M. Jenkins, *Time Series Analysis Forecasting and Control*, 2nd ed., Holden-Day, San Francisco, USA, 1976.

[11] F. J. Nogales, J. Contreras, A. J. Conejo and R. Espínola, "Forecasting Next-Day Electricity Prices by Time Series Models," to appear in IEEE Transactions on Power Systems, paper 2001TR212.

[12] L. Liu and G. P. Hudak, Forecasting and Time Series Analysis using the SCA Statistical System, Scientific Computing Associated, 1994.