

# Forecasting the electricity load from one day to one week ahead for the Spanish system operator

José Ramón Cancelo <sup>a,\*</sup>, Antoni Espasa <sup>b,1</sup>, Rosmarie Grafe <sup>c,2</sup>

<sup>a</sup> *Universidade da Coruña, Dpto. Economía Aplicada II, Facultad Ciencias Económicas, Campus de Elviña, 15008 A Coruña, Spain*

<sup>b</sup> *Universidad Carlos III de Madrid, Statistics Department Universidad Carlos III, Madrid, Spain*

<sup>c</sup> *Red Eléctrica de España, Pº del Conde de los Gaitanes, 177, 28109 Alcobendas (Madrid), Spain*

## Abstract

This paper discusses the building process and models used by *Red Eléctrica de España* (REE), the Spanish system operator, in short-term electricity load forecasting. REE's forecasting system consists of one daily model and 24 hourly models with a common structure. There are two types of forecasts of special interest to REE, several days ahead predictions for daily data, and one day ahead hourly forecasts. Accordingly, the forecast accuracy is assessed in terms of their errors. To do this, we analyse historical, real time forecasting errors for daily and hourly data for the year 2006, and report the forecasting performance by day of the week, time of the year and type of day. Other aspects of the prediction problem, like the influence of the errors in predicting the temperature on forecasting the load several days ahead, or the need for an adequate treatment of special days, are also investigated.

© 2008 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.

**Keywords:** Energy forecasting; Hourly and daily models; Time series; Forecasting practice

## 1. Introduction

Short-term load forecasting has long been a issue of major interest for the electricity industry. Traditionally, hourly forecasts with a lead time of between one hour and seven days are required for the scheduling and control of power systems. From the perspective of the system operators and regulatory agencies, they are a primary input for the safe and reliable operation of the system. For

producers, they are a basic tool for determining the optimal utilization of generators and power stations, as some facilities are more efficient than others. Since the rise of free electricity markets that followed the deregulation of the power industry in the nineties, more agents became interested in load forecasting. Nowadays there are a huge range of market participants, bidding strategies are more and more complex, and a number of financial derivatives have been developed. Commercial success depends on the ability to submit competitive bids, and even marginal improvements in forecasting the load can lead to substantial increases in trading profits. Nevertheless, the fact that a widespread demand for load forecasts exists does not mean that all agents are concerned with the same issues, as each organization has

\* Corresponding author. Tel.: +34 981167000 2512; fax: +34 981167070.

E-mail addresses: [jramon@udc.es](mailto:jramon@udc.es) (J.R. Cancelo),

[espasa@est-econ.uc3m.es](mailto:espasa@est-econ.uc3m.es) (A. Espasa), [rosmarie@ree.es](mailto:rosmarie@ree.es) (R. Grafe).

<sup>1</sup> Tel.: +34 916249803.

<sup>2</sup> Tel.: +34 916502012 2314; fax: +34 916504542.

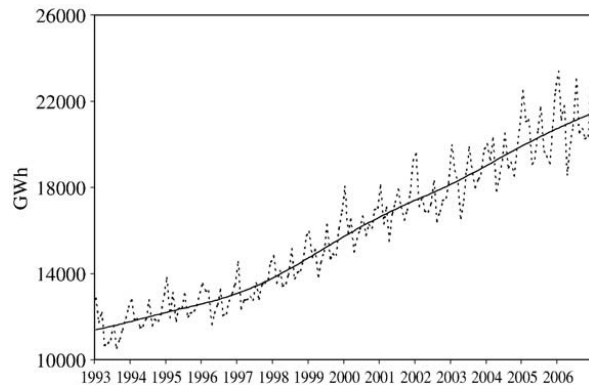


Fig. 1. Monthly electricity consumption in mainland Spain: the original series and the trend computed with a Hodrick-Prescott filter with  $\lambda = 14,400$ .

particular aims, and a prerequisite of setting up fully adequate forecasting devices is explicitly stating the actual information needs of the organization.

This paper describes the forecasting experience at *Red Eléctrica de España* (REE), based on different models developed by the authors since 1988. Load forecasting is a central activity within REE, and the search for a forecasting system able to encompass all short-term load forecasting tasks has long been a priority. Forecasting success is important, and is a necessary condition for validating any model. However, the models are also internal tools that have to adapt to the operational features of REE, and play a leading role in the acquisition, sharing, and use of knowledge within the organization (Nevis, DiBella, & Gould, 1995; Cancelo & Espasa, 2001). As such, they are at the end of a sequence that begins with identifying the needs of REE, the kind of information that forecasts are expected to provide, critical origins and lead times, and ways to implement forecasting tools to promote organizational learning, i.e., the capacity to improve performance based on experience. In this sense, models are not aimed at producing tentative estimates that should be further elaborated by the control operator to derive the final predictions of the load. On the contrary, the forecasting procedures seek to capture the knowledge of the qualified staff of the organization, and hence their input in developing and monitoring the models is a major necessary condition for their success.

The paper is organized as follows. Section 2 summarizes the stylized features of the load data for mainland Spain to give a general view of the forecasting problem. Section 3 reviews the foundations of the mo-

dels that make up REE's forecasting system. After a general introduction, we sketch their main characteristics, report some estimated effects, and compare REE's approach to other procedures that have been suggested in the literature. Section 4 evaluates the forecasting performance and Section 5 concludes.

## 2. Load data in mainland Spain: Stylized facts

The main features of load series have been extensively reported in the literature: trend, superimposed levels of seasonality, short-term dynamics, special days, nonlinear effects of meteorological variables, possible nonlinear time dependence, etc.; see for instance Espasa, Revuelta, and Cancelo (1996) in the general context of daily time series of economic activity. Such features are not constant, as the observed series display a great variety of profiles, according to the relative importance of each underlying component and the way they interact.

Centring on the particular traits of mainland Spain, it is an area of almost 500,000 square kilometres, with more than 42 million people. The average daily electricity consumption in 2005 was 677 GWh. The winter peak was 851 GWh and the summer peak 786 GWh. Figs. 1–5 highlight some of the main characteristics of the load series. Fig. 1 plots the monthly series to obtain a clearer picture of the trend. It shows both the original series and the trend computed with a Hodrick-Prescott filter with  $\lambda = 14,400$ , the default value for monthly data.

Fig. 2 depicts the daily consumption during 2005. Several intra-annual components can be observed: the

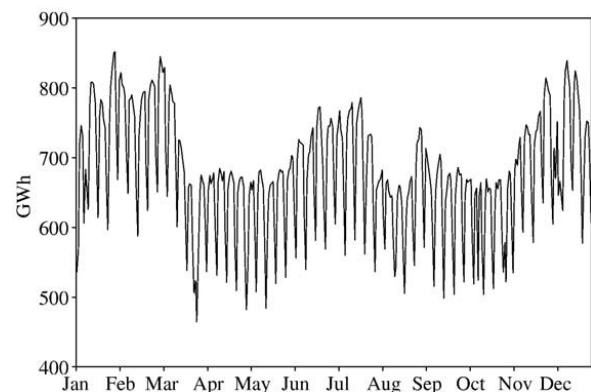


Fig. 2. Daily electricity load in 2005.

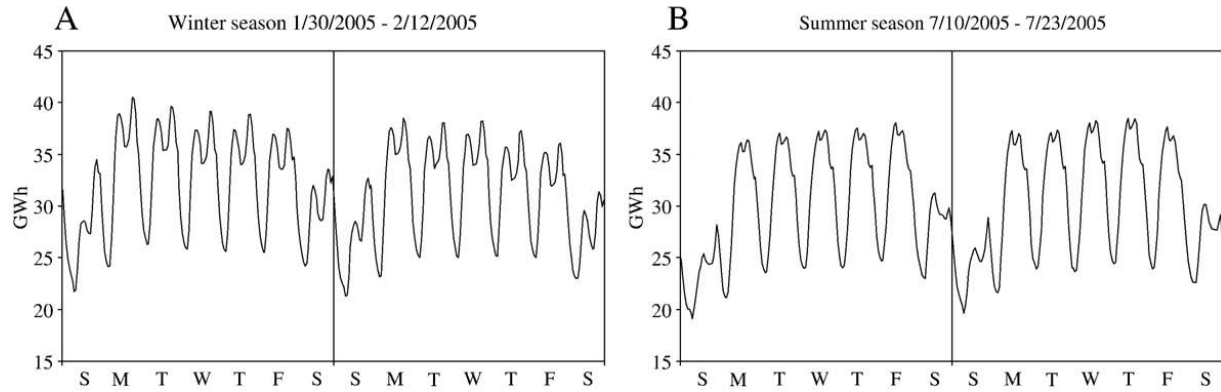


Fig. 3. Intraday electricity load patterns.

weekly seasonal cycle, the weather-sensitive part of the load related to temperature effects in winter and summer, and the influence of the vacation period in August. The annual seasonality is closely related to

meteorological variables and vacation periods. Weekly and annual seasonalities are not constant, as, for instance, the weekly seasonality is less pronounced in August than in any other month.

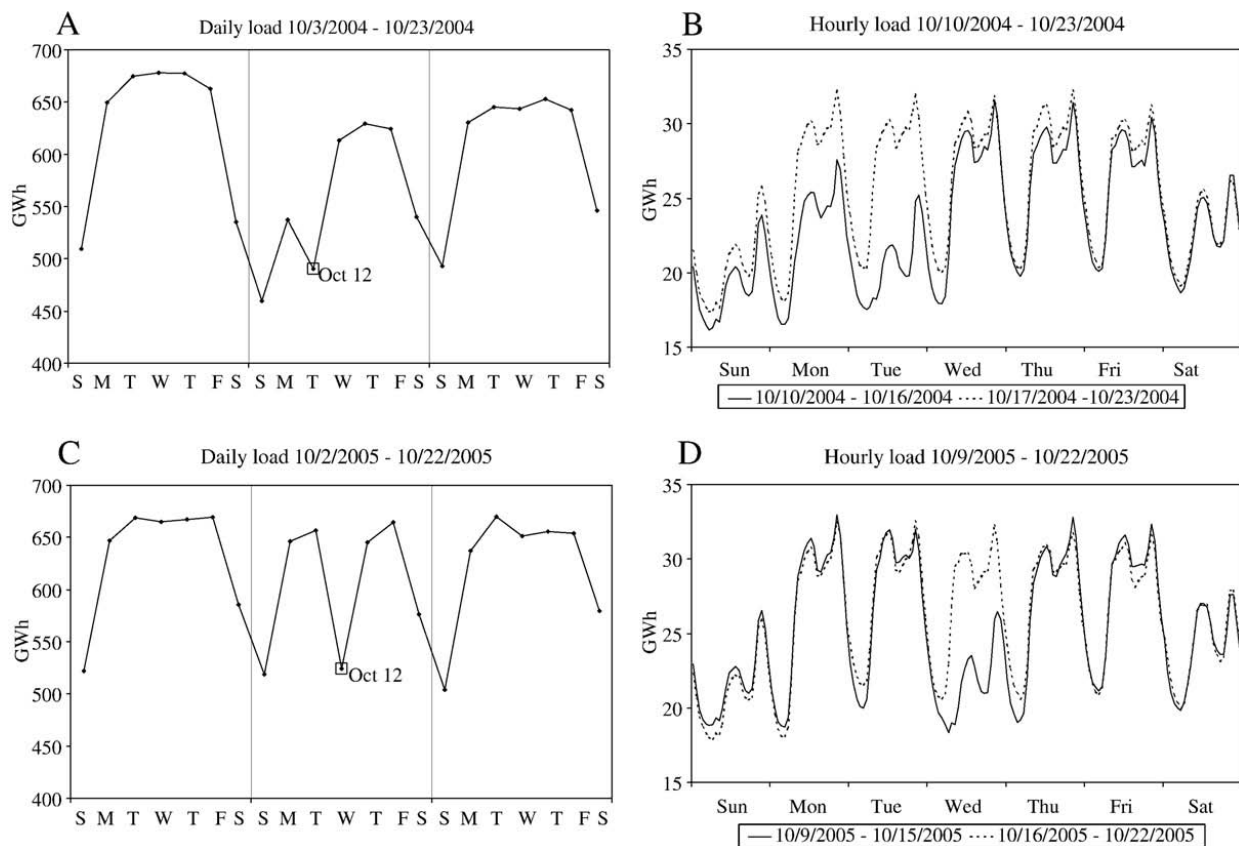


Fig. 4. Effect of a public holiday on the load.

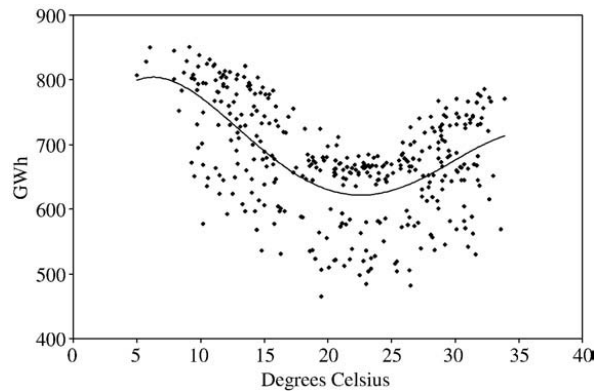


Fig. 5. Daily load and maximum temperature in 2005.

Fig. 3 focuses on intraday patterns. A fortnight of hourly data is displayed in each panel, and the figure shows that the daily periodic behaviour varies with both the day of the week and the time of the year.

Fig. 4 reports the alterations in the load when a special day occurs. The four panels refer to the same public holiday, October 12, on two consecutive years; panels A and B are for the holiday falling on a Tuesday, and panels C and D for the holiday on a Wednesday. It can be seen that the effects of special days are not independent of the weekly seasonal factor, and they can also be shown to be related to the annual seasonality.

Forecasting special days is a key issue for system operators, especially in Spain, due to the idiosyncratic characteristics of the Spanish calendar. In 2006 there were 29 national or regional holidays, each with very different effects on the mainland Spain consumption of electricity. Local holidays in some big cities, notably Madrid (3.1 million people) and Barcelona (1.6 million), must be taken into account too, as they have perceptible effects on the total load. Special days also include long weekends arising from holidays that fall on Tuesday or Thursday/Friday, and three vacation periods (August, Easter and Christmas).

Fig. 5 focuses on the relationship between consumption and maximum temperature for the daily data during 2005; a fourth-order polynomial was fitted to give a tentative indication of its shape. The relationship is U-shaped and (could be) asymmetric, as the left branch is more pronounced than the right wing. There is also some evidence of an exhaustion effect, especially for low temperatures.

### 3. Review of the models

#### 3.1. The basic strategy

The forecasting system currently implemented in REE consists of one daily model for forecasting the daily load up to ten days ahead, and 24 hourly models for computing hourly predictions for horizons up to three days. The daily model is aimed at producing forecasts for network outage planning, while the hourly models are used to derive forecasts for the next-day hourly dispatch.

This combination of 1 + 24 models has proven to be the best approach according to REE's needs. Daily forecasts up to ten days ahead for outage planning could also be obtained from the hourly models, by computing 24 hourly forecasts for each day and adding them up to get the daily total. This procedure, however, produces forecasts with too much uncertainty for lead times up to ten days ahead, while a direct prediction of the daily aggregate provides a more adequate anchor. In a second step, and in order to derive the hourly forecasts that are required to complete the final outage plan, hourly models are used to compute the shape of the load curve for each day, and the daily forecasts are interpolated according to the predicted load curves.

In contrast, for horizons up to three days, direct hourly forecasts are more accurate than interpolated daily figures. Although there is some controversy as to whether it is better to treat each hour as a different series (Ramanathan, Engle, Granger, Vahid-Araghi, & Brace, 1997; Cottet & Smith, 2003) or to build a single model to derive the load profile (Smith, 2000; Taylor, de Menezes, & McSharry, 2006), most authors prefer the first approach. REE's experience since 1988 supports that view, as separate models can easily be specified in a very flexible way to capture short-term dynamics and the effects of special days and meteorological variables at different times of day.

All of the models have the same basic structure, related to the traditional decomposition of the observed load into four components: the normal load, the weather sensitive part, special events and a random component. They are given a common type of response function, with parameters that are allowed to vary across series. The full search for a specification was carried out when building the daily model, as the stylized features of the load in the aggregated daily

data are also present in the daily series corresponding to each hour. Let  $C_t$  denote the electricity consumption on day  $t$ , and assume that the model is additive in logarithms:

$$\ln C_t = p_t + s_t + CSD_t + CWEA_t + U_t, \quad (1)$$

where  $p_t + s_t$  is the normal load, with  $p_t$  denoting the trend and  $s_t$  (part of) the seasonality;  $CSD_t$  is the contribution of special days;  $CWEA_t$  is the contribution of meteorological variables; and  $u_t$  is a stationary disturbance that may display some short-term, transitory dynamics. In practice it is not easy to separate annual seasonality from the effect of weather variables, so in the daily model  $s_t$  is basically the weekly seasonality. In fact, most annual seasonality can be explained by weather variables and dummy variables that take vacation periods into account.

In the rest of Section 3.1 we give a rough, general sketch of the way each component is modelled. Next, some alternative approaches are reviewed in Section 3.2, where we also explain why we opted for the specification that follows, and report some estimated effects to give some insight into the dependence of the electricity consumption on special days and temperature.

The basic consumption  $BC_t$  can be defined from Eq. (1) as:

$$\ln BC_t = \ln C_t - CSD_t - CWEA_t = p_t + s_t + u_t. \quad (2)$$

This represents the load that would be observed if there were no further effects from special days or weather variables. Since these additional effects are the major disturbances that alter the observed load,  $BC_t$  is rather smooth, and, at least as an approximation, can be assumed to follow an ARIMA model:

$$\phi(L)\Delta(L)\ln BC_t = \theta(L)a_t, \quad (3)$$

where  $\phi(L)$  and  $\theta(L)$  are polynomials in the lag operator  $L$  with all of their roots outside the unit circle;  $\Delta(L)$  is a polynomial with unit roots; and  $a_t$  is a white noise process. Figs. 1 and 2 suggest that a reasonable specification for  $\Delta(L)$  is  $\Delta\Delta_7 = (1-L)(1-L^7)$ , as the series display a local linear trend and a strong weekly seasonal pattern, which is non-deterministic and is not explained by exogenous variables. If the stationary AR and MA polynomials are specified in a multiplicative form with three factors, a regular factor on the operator  $L$ , a weekly seasonal factor on  $L^7$  and an annual seasonal factor on

$L^{365}$ , the final expression of the ARIMA model for the basic consumption is:

$$\phi(L)\Phi_7(L^7)\Phi_{365}(L^{365})\Delta\Delta_7 \ln BC_t = \theta(L)\Theta_7(L^7)\Theta_{365}(L^{365})a_t. \quad (4)$$

Once the basic load is defined, special days and weather effects are introduced by extending the model to include dummy and weather variables. Their effects on the load are dynamic, and may depend on the day of the week and the time of the year. In the case of dummies for special days, adding leads and lags is a parsimonious way of representing the effect on adjacent days, while the dynamic weather effects reflect the fact that individuals adjust to changing weather conditions with some delay. The nonlinear relationship between consumption and temperature is modelled by computing several degree-days transformations of the observed temperatures, which will be explained in detail in Section 3.2. The model also includes other meteorological variables like cloudiness and the duration of sunlight. Wind speed and relative humidity are also known to affect the load, but are not considered because no reliable forecasts are produced, and therefore they provide no relevant information for load forecasting.

The joint contribution of special days and weather variables can be expressed as:

$$CSD_t + CWEA_t = \sum_{i=1}^m \alpha_i(L)SD_{i,t} + \sum_{j=1}^n \beta_j(L)WEA_{j,t}, \quad (5)$$

where  $SD_{1,t}, SD_{2,t}, \dots, SD_{m,t}$  are  $m$  dummy variables that define the different classes of special days;  $WEA_{1,t}, WEA_{2,t}, \dots, WEA_{n,t}$  represent  $n$  transformations of the observed meteorological variables; and  $\alpha_i(L), \beta_j(L)$ ,  $i=1, \dots, m$ , and  $j=1, \dots, n$ , are lag polynomials.

From Eqs. (4) and (5) it follows that

$$\ln C_t = \sum_{i=1}^m \alpha_i(L)SD_{i,t} + \sum_{j=1}^n \beta_j(L)WEA_{j,t} + \frac{\theta(L)\Theta_7(L^7)\Theta_{365}(L^{365})}{\phi(L)\Phi_7(L^7)\Phi_{365}(L^{365})\Delta\Delta_7} a_t. \quad (6)$$

The final specification of model (6) was derived by starting from an ARIMA model with dummy variables to capture the influence of special days; at this first stage, no effects of weather variables were considered. Next, the model was extended to include a basic response to



temperature, and in subsequent versions this response was allowed to be more and more complex to capture the variety of effects that are reviewed in Section 3.2. Finally, other meteorological variables were tested to assess whether they actually improved the forecasting performance of the model, and the full response to weather conditions was derived. Every effect and variable included in the model was tested to assess its statistical significance, and the usual tests of misspecification were performed to validate the final version.

The search for a specification centred on the daily series; once a final, satisfactory daily model was achieved, its parameters were reestimated for each hourly series to get the final hourly models. As a general rule, it was found that the hourly models are not as complex as the daily model, because many meteorological effects are not significant for specific hours within the day.

### 3.2. Description of the models

#### 3.2.1. Basic load

Several procedures have been proposed to take the trend into account. Some studies get rid of it by taking short samples where the mean level of the load can be assumed to be constant (Smith, 2000). Others consider a variety of deterministic functions of time: linear (Pardo, Meneu, & Valor, 2002), linear and reciprocal (Ramanathan et al., 1997), quadratic (Cottet & Smith, 2003), etc. Deterministic seasonal components include dummy variables (Ramanathan et al., 1997; Soares & Souza, 2006) and time-varying periodic splines (Harvey & Koopman, 1993). Methods for explaining short-term time dependence include linear autoregressions (Pardo et al., 2002; Cottet & Smith, 2003), fractionally integrated processes (Soares & Souza, 2006), artificial neural networks (Lamedica, Prudenzi, Sforza, Caciotta, & Orsolini Cencelli, 1996; Hippert, Bunn, & Souza, 2005; Alves da Silva, Ferreira, & Velasquez, 2008-this issue), double seasonal Holt-Winters' exponential smoothing adjusted for error correlation (Taylor, 2003), etc.

Our proposal for the basic load is simple and yet highly effective. The two real, positive unit roots embedded in the  $\Delta\Delta_7$  operator generate a local linear trend that is flexible enough to approximate a wide variety of situations, and a large sample can be used for estimating the model without a need to worry about the shape of the trend. Furthermore, it does not impose any constraint on

the forecasts, so any changes in the underlying level are automatically accommodated in a natural way.

Weekly unit roots allow for the type of evolving weekly seasonality that is displayed in Figs. 2 and 3. These figures show that the weekly seasonality depends on the time of the year, and the seasonal unit roots within the  $\Delta_7$  operator act like an adaptive mechanism that adjusts this week's pattern to accommodate the yearly variation that is already incorporated in the previous week's observations. In addition, this approach does not impose the restriction that the intraannual evolution of the weekly seasonality is the same year after year, as would be the case if such interactions between seasonalities were modelled by parametric functions of time.

Short-term time dependence is assumed to be generated by an ARMA process, because it is more general than the residual autocorrelation adjustments used by Ramanathan et al. (1997), Pardo et al. (2002) and Taylor (2003), and can handle a wider variety of dynamic effects with fewer parameters. It is also preferred to artificial neural networks because there is no sound empirical evidence that the load data display the type of nonlinear behaviour driven by ANNs (Darbellay & Slama, 2000; Hippert et al., 2005; Taylor et al., 2006).

#### 3.2.2. Special days

Most applied studies separate normal and special days and build different models for each category (Ramanathan et al., 1997; Darbellay & Slama, 2000). This strategy is easy to implement for typical working days and weekends, but it is not so straightforward for public holidays, long weekends, vacation periods, etc. Public holidays may be smoothed (Taylor et al., 2006), treated as Sundays (Smith, 2000), or replaced by the load observed on a similar day in the week before (Hippert et al., 2005). The number and types of special days are usually specified on a priori grounds, although some authors have used pattern recognition analysis to classify day types empirically (Lamedica et al., 1996).

Within the single model approach, common practice is to use dummy variables to estimate the changes in the load due to special days. The number of total parameters varies across studies: one in Soares and Souza (2006), three in Pardo et al. (2002), and six in Cottet and Smith (2003). REE's experience indicates that the number required is much higher, at least for mainland Spain. Although a detailed presentation of REE's procedure is

well beyond the scope of this paper, we briefly review the case of public holidays to sketch the basic ideas. There are two classes of holidays, common holidays and special holidays. Common holidays include some national holidays and all regional and local holidays. Besides these, there are three special holidays, all of them nationwide: January 6, May 1 and August 15. It was found that each special holiday has a specific and separate influence on the load, which is statistically different from the influence of any other special day.

To highlight the importance of allowing for different effects for each class of public holiday, Fig. 6 displays the estimated decrease of the load on the day of the holiday. Although some particular coefficients could be restricted to take the same value, as a whole the response varies in magnitude and shape. The estimated standard deviation for each estimator is about 0.5%, and broadly speaking all of the differences between estimated responses above 1% are statistically significant.

### 3.2.3. Temperature effects

Meteorological variables play a leading role in the model, especially temperature. In this brief review we focus on the response to the daily maximum temperature, and assume for simplicity that the other weather variables have no additional effects on the load. The daily maximum temperature for mainland Spain is computed as a weighted average of the maximum temperatures recorded between 6 A.M. and 6 P.M. at ten observatories that represent different climatic regions;

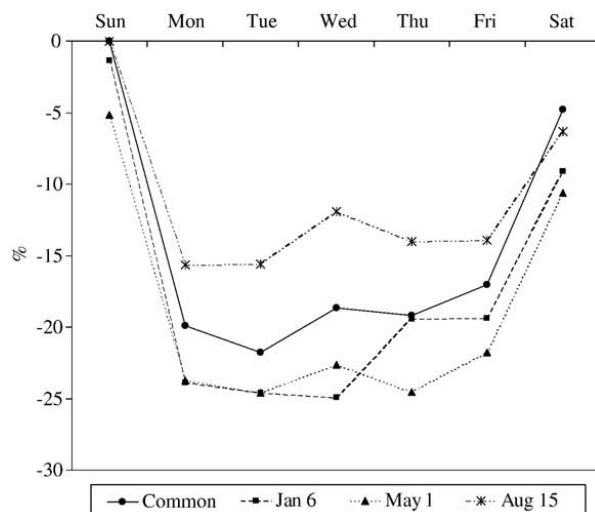


Fig. 6. Estimated decrease of the daily load on a public holiday.

within a typical year it varies from a low of 8 °C to a high of 34 °C. Territorial weights are derived from the statistics of the Electric Power Industry (source Ministry of Industry, Tourism and Trade), that report annual data for the Spanish provinces.

The estimated relationship between load and temperature is highly complex for a number of reasons. On a priori grounds, the literature suggests that it is non-linear and U-shaped as long as there is enough variation of temperatures in the sample (Engle, Granger, Rice, & Weiss, 1986). The response is asymmetric, in the sense that a one-degree increase when the temperature is high and a one-degree decrease when the temperature is low need not have the same impact on the load (Valor, Meneu, & Caselles, 2001). The relationship is different for working and nonworking days, as heating and cooling systems equipment and operation are very different in workplaces from private residences (Smith, 2000); it also changes with the time of the year (Hyde & Hodnett, 1997). There is a dynamic effect, as the load at  $t$  does not depend solely on the temperature at  $t$ , but also on the temperatures on the previous days (Le Comte & Warren, 1981). Exhaustion effects may appear because of the limited capacity of existing electricity appliances, so there is no further increase in electricity consumption when the temperature exceeds some saturation level (Henley & Peirson, 1997).

In regard to the actual need of considering such a variety of temperature effects, it is true that some of them are seldom necessary, as temperature usually evolves smoothly and the load adapts to its changes. They become very important, however, when temperature changes abruptly. In this situation, a model with a complex response is better prepared to predict the load, and hence to reduce the risk of very marginal scheduling that increases the likelihood of failures or equipment damage.

The particular shape of the response is specific to each set of data, and issues such as the degree of non-linearity, the order of the dynamics, and the presence and location of the saturation levels, for instance, must be empirically determined. In what follows, we focus on the main characteristics of the estimated relationship as it is currently used to forecast the load. Further details on the modelling strategy can be found in Cancelo and Espasa (1996).

The response to the maximum temperature has two branches, one for temperatures below 20 °C and one for temperatures higher than 24 °C. The interval from

20 °C to 24 °C is a comfort zone of moderate temperatures with no effect on electricity consumption. The general expression of the response to low temperatures is given by:

$$\alpha_0 HDDS9_t + \beta_0 HDDS11_t + \gamma_0 HDDS14_t + \sum_{i=0}^5 \pi_i HDDS20_{t-i}, \quad (7)$$

with

$$HDDSx_t = \begin{cases} 0 & \text{if } T_{\max_t} > x \\ x - T_{\max_t} & \text{if } 9 < T_{\max_t} < x, \\ x - 9 & \text{if } T_{\max_t} < 9 \end{cases} \quad (8)$$

where  $T_{\max_t}$  denotes the observed maximum temperature and  $HDDSx_t$  is a heating degree-days variable with reference temperature  $x$  and a saturation level of 9 °C.

Expression (7) captures the basic shape of the relationship for temperatures below 20 °C, but it was found that its magnitude depends on the time of the year, and, in some seasons, on the day of the week. Two seasons are distinguished, which will be referred to as summer and non-summer, or rest of the year. In the non-summer season, the response is different for normal weekdays and for weekends/special weekdays. In the summer season the impact does not depend on the day of the week and the response is much simpler: temperatures below 16 °C are seldom observed and the dynamics go from  $t$  to  $t-2$ , so the only non-zero coefficients in (7) are  $\pi_0$ ,  $\pi_1$  and  $\pi_2$ .

At the other branch of the load/temperature response curve, the response to high temperatures is given by

$$\sum_{i=0}^2 \omega_i CDDS24_{t-i}, \quad (9)$$

where  $CDDS24_t$  is a cooling degree-days variable, with threshold 24 °C and saturation level 33 °C:

$$CDDS24_t = \begin{cases} 0 & \text{if } T_{\max_t} < 24 \\ T_{\max_t} - 24 & \text{if } 24 < T_{\max_t} < 33. \\ 9 & \text{if } T_{\max_t} > 33 \end{cases} \quad (10)$$

Temperatures above 24 °C only affect the load in the summer season, and there are no differences between normal weekdays and weekends/special days. The response increases linearly from 24 °C up to 33 °C, and

remains constant above that level. The load at  $t$  depends on the observed temperatures at  $t$ ,  $t-1$  and  $t-2$ .

To summarize the estimated effects, Fig. 7 displays the long-run gain for each temperature value. The long-run gain for  $T_{\max} = T_{\max}^*$  is defined as the percent increase of the daily load on day  $t$ , assuming that the temperature stays constant at  $T_{\max}^*$  for all of the relevant lags in Eqs. (7) and (9), with respect to the consumption that would have been registered if the temperature had been within the interval 20 °C to 24 °C. As an example, the gain for  $T_{\max} = 26$  °C is equal to 2.48%, which means that if the observed maximum temperature on days  $t-2$ ,  $t-1$  and  $t$  is 26 °C, the daily consumption at  $t$  will be 2.48% higher than if the temperature had been within the interval 20 °C to 24 °C during that period.

#### 4. Forecast performance during 2006

In Section 3 it was stated that there are two types of forecasts of special interest to REE, several days ahead predictions from daily data, and one day ahead hourly forecasts. Accordingly, the forecast accuracy should be assessed in terms of their errors, so in this section we discuss the performance of the models with regard to the two activities during 2006.

##### 4.1. Prediction up to ten days ahead using the daily model

Daily forecasts are used to set up the weekly network outage plan. They are computed by the middle of the week, usually on Wednesday morning, with information up to Tuesday, for the seven-day period beginning the following Saturday. The relevant lead times go from 4 to 10 days ahead, although there are some minor modifications; for example, the origin of the forecast changes when a public holiday falls on a Wednesday (four times in 2006). We ignore this for the sake of simplicity, and act as if there were a one-to-one relationship between the day of the week and the lead time. The forecasts are based on model parameters that were estimated using daily data from January 1, 1993 to March 31, 2005 (4473 observations). No re-estimation was performed during 2006, as it was found that the estimates remain rather stable for moderate increases in the sample size (less than two or three years).



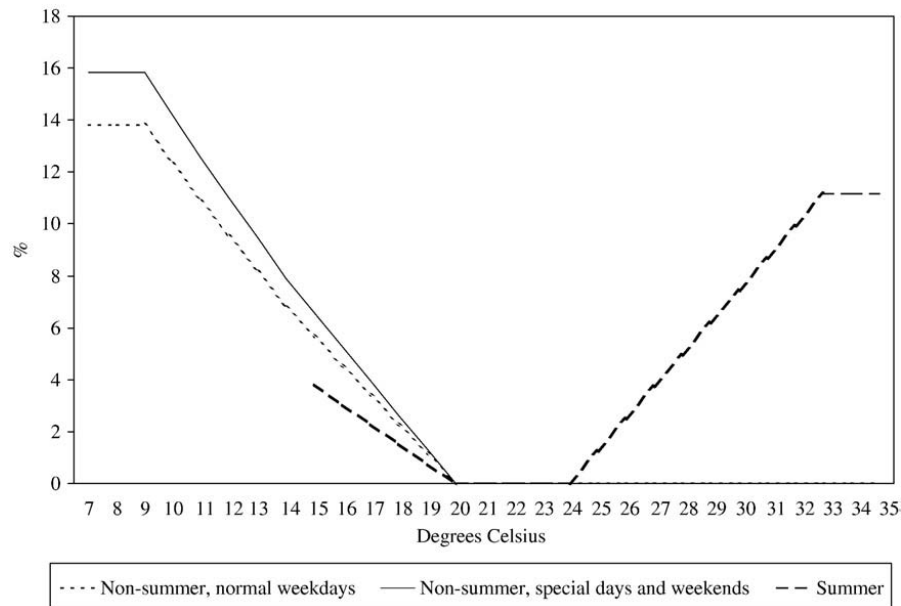


Fig. 7. Long-run gain in maximum temperature: percent increase of the daily load with respect to the load for temperatures within the interval 20 °C to 24 °C.

Two types of errors are computed, historical errors and model errors. By historical errors we mean true post-sample forecast errors, i.e., the errors that were actually observed. These arise either because of model misspecification or as a consequence of the errors made in forecasting meteorological variables; temperature forecasts are highly accurate for horizons up to three or four days, but they deteriorate as the lead time goes further. Hence, and in order to assess the influence of meteorological predictions, the load forecasts were recalculated by inserting the true values of the meteorological variables in the model, and the resulting errors are referred to as model errors.

Table 1 reports mean absolute percentage errors (MAPE) for historical errors and model errors. It is well known that each set of data deserves particular consideration, and no general indications of how effective a forecasting procedure is can be given unless an explicit comparative evaluation is carried out on the same time series. In spite of this, the real time performance of the model looks to be satisfactory, in the sense that the errors are within the bounds that guarantee the electricity supply security and quality. In the full year, the MAPEs of the historical errors are always below 2.9%, even for lead times of ten days ahead. They increase with the lead time, as expected, although

Table 1  
Mean absolute percentage errors (MAPE) in daily forecasting

Day	Lead time	Historical errors				Model errors			
		All	Non-summer	Summer	Difference	All	Non-summer	Summer	Difference
Sat	4	1.34	1.33	1.35	0.02	1.35	1.32	1.40	0.08
Sun	5	1.97	2.03	1.86	−0.17	1.74	1.80	1.62	−0.18
Mon	6	2.13	2.16	2.07	−0.09	1.53	1.53	1.52	−0.01
Tue	7	2.06	1.79	2.63	0.84	1.53	1.18	2.26	1.08
Wed	8	2.36	1.95	3.19	1.24	1.72	1.32	2.54	1.22
Thu	9	2.59	2.08	3.61	1.53	1.81	1.45	2.58	1.13
Fri	10	2.82	2.23	3.95	1.72	2.01	1.74	2.53	0.79

there is some evidence that the forecasts are slightly worse on Sundays.

Using temperature predictions instead of true values has no perceptible effect on Saturday's forecasts. From Sunday on, part of the actual error in forecasting the daily load is due to unsatisfactory temperature predictions, and that part increases with the lead time. The MAPEs of the model errors increase with the lead time too, but do not exceed 2% in any case. Model errors confirm that the forecasting performance is less satisfactory on Sunday than on weekdays.

To assess the performance by time of the year, the database of errors was split into two parts, summer and rest-of-the-year. Table 1 indicates that the forecasts deteriorate during the summer months, especially for weekdays. Errors in forecasting temperature are the cause of this deterioration for the longest horizons, but it basically reflects a weakness of the model rather than of any of its inputs. The most important errors are in August, a major period of vacation in mainland Spain: according to official estimates, the industrial production index fell by more than 30% in August 2005 due to seasonal effects. The model includes some dummy variables to capture the effects of the vacation on the load, but at this stage the results have not been as satisfactory as for the rest of the year.

As for the type of day, it is not straightforward to compare the forecasting performance for weekdays that are normal days, weekdays that are special days and weekends, because of the changing lead time. In order to give some indication of this issue, we concentrated on weekdays, and separated normal days (214 observations in the forecasting period) from special days (46 observations). For historical errors the MAPE is 2.34% for normal weekdays and 2.63% for special weekdays, while for model errors it is equal to 1.65% and 2.06%, respectively.

Decision makers are especially concerned about large errors, as they involve major rises in costs due to either unnecessary start-ups and energy purchases, or very marginal scheduling that increases the likelihood of failures or equipment damage. There is not a general agreement as to what a large error is, as each organization determines its own acceptable limit, given its activities and aims. From the point of view of a utility, for instance, that limit depends heavily on its characteristics and operational features, but it also varies accor-

ding to the day of the week and hour of the day, the type of load (hourly, daily or peak), etc.

For simplicity, in this paper we assume a constant limit equal to 5%, so any error above that level in absolute value is termed a large error. This is a rough estimate that is often used at REE, and has also been suggested as an adequate benchmark in the literature; see for instance [Ranaweera, Karady, & Farmer \(1997\)](#). Table 2 shows that in 2006 there were 32 large historical errors (out of 365; that is, 8.8%). The number of large errors increases with the lead time, as expected. Interestingly, most of them are caused by errors in predicting the weather variables: when model errors are considered there are only 7 large errors, less than 2% of the total.

#### 4.2. Forecasting the hourly load for the next-day dispatch schedule

The hourly forecasts that determine the final dispatch schedule for the next day constitute the second set of critical forecasts for REE. Such forecasts are computed every day, and the lead time is always one day. Unlike the previous exercise, the lead time is constant and the origin changes along the week. The models are reestimated daily, so the forecast for day  $t$  is based on parameters that were estimated using data from January 1, 2001 to the last available observation.

Table 3 reports mean absolute percentage errors of the historical hourly absolute percentage errors of the forecasts for the next day that were made public on REE's web page during 2006. They refer to 365 consecutive days, with no exceptions, no corrections for public holidays, vacation periods, cold snaps, unex-

Table 2  
Absolute percentage errors higher than 5% in daily forecasting

Day	Lead time	Historical errors		Model errors	
		Number	% Total	Number	% Total
Sat	4	1	1.9	0	0.0
Sun	5	3	5.7	1	1.9
Mon	6	3	5.8	1	1.9
Tue	7	4	7.7	0	0.0
Wed	8	6	11.5	1	1.9
Thu	9	5	9.6	2	3.8
Fri	10	10	19.2	2	3.8
Total	–	32	8.8	7	1.9

Table 3

Mean absolute percentage errors for one day ahead hourly forecasting: all days and details by day of the week

Hours	All days	Day of the week						
		Sun	Mon	Tue	Wed	Thu	Fri	Sat
1 A.M.	1.34	1.55	1.28	1.47	1.36	1.47	1.04	1.23
2 A.M.	1.37	1.70	1.05	1.52	1.33	1.45	1.17	1.37
3 A.M.	1.35	1.78	1.08	1.66	1.33	1.33	1.06	1.23
4 A.M.	1.35	1.80	1.19	1.61	1.23	1.28	1.02	1.32
5 A.M.	1.39	1.80	1.22	1.69	1.41	1.26	0.99	1.34
6 A.M.	1.31	1.61	1.17	1.51	1.34	1.30	0.99	1.25
7 A.M.	1.33	1.47	1.24	1.53	1.33	1.22	1.18	1.33
8 A.M.	1.45	1.64	1.56	1.52	1.39	1.30	1.28	1.46
9 A.M.	1.66	2.03	1.87	1.78	1.66	1.42	1.31	1.52
10 A.M.	1.63	2.08	2.14	1.57	1.60	1.43	1.32	1.26
11 A.M.	1.60	2.06	2.01	1.46	1.55	1.45	1.40	1.24
12 P.M.	1.60	1.89	2.02	1.52	1.53	1.45	1.52	1.26
1 P.M.	1.61	1.79	1.86	1.52	1.59	1.49	1.58	1.41
2 P.M.	1.62	1.77	1.78	1.57	1.58	1.52	1.60	1.50
3 P.M.	1.63	1.79	1.75	1.54	1.57	1.65	1.55	1.53
4 P.M.	1.77	1.91	1.95	1.60	1.63	1.76	1.74	1.77
5 P.M.	1.91	1.97	1.97	1.68	1.88	1.72	2.05	2.09
6 P.M.	1.90	1.82	1.81	1.71	2.03	1.65	1.95	2.32
7 P.M.	1.85	1.91	1.66	1.66	2.08	1.56	1.69	2.39
8 P.M.	1.74	1.84	1.67	1.68	1.89	1.55	1.53	2.03
9 P.M.	1.63	1.93	1.70	1.49	1.81	1.36	1.39	1.72
10 P.M.	1.40	1.66	1.57	1.22	1.47	1.09	1.16	1.60
11 P.M.	1.39	1.40	1.59	1.21	1.62	1.21	1.15	1.57
12 A.M.	1.55	1.80	1.61	1.44	1.67	1.24	1.35	1.73
All hours	1.56	1.79	1.62	1.55	1.58	1.42	1.38	1.56

pected events, etc., other than those implemented in the models, and no supplementary subjective adjustments by the control operator or any other expert. Summarising all days of the week together, the MAPE remains below 2% for all 24 hours. It has the lowest values at late night hours and early in the morning, up to 8 A.M.; within this period it lies between 1.3% and 1.5%. It then increases to about 1.6% for the first half of the business hours, from 10 A.M. to 3 P.M., and rises further to more than 1.7% from 4 P.M. to 8 P.M. A closer inspection of Table 3 reveals that this pattern of intraday evolution is not stable, as it changes according to the day of the week. In the beginning of the week, from Sunday to Tuesday, the maximum MAPE is attained around 9 A.M. to 10 A.M., while on Saturdays it is not observed until 7 P.M.

When the database of errors is split by day of the week, it is found that the most difficult period to predict is Saturday afternoon (from 5 P.M. to 8 P.M.), followed by Sunday morning (from 9 A.M. to 11 A.M.) and Monday morning (from 10 A.M. to 12 P.M.). Saturday

afternoon behaves like a Sunday in some aspects, and like a weekday in regard to commercial activities, especially retail trading, and that duality increases the forecast uncertainty to its maximum. In any case there are only 13 values of the MAPE higher than 2% (out of 168, or 7.7%), and none above 2.5%.

Table 4 reports MAPEs computed separately by time of the year and by type of day. Contrary to what was found in daily forecasting, the hourly models perform better in the summer season for almost all hours, although the difference is greater early in the morning. With regard to the types of day, there are differences in the magnitude of the errors among normal weekdays, weekends and special weekdays. For normal weekdays MAPEs vary from 1.14% at 7 A.M. to 1.64% at 6 P.M. The intraday evolution is rather similar on weekends, but now the MAPEs range from a low of 1.40% at 7 A.M. to a high of 2.15% at 7 P.M. Finally, for weekdays that are special days the MAPES go from 1.39% at 4 A.M. to 2.90% at 5 P.M.

The number of errors above 5% in absolute value is 218 (out of 8160, or 2.5%). Large errors are more likely to occur between 5 P.M. and 8 P.M. and on weekends: almost one third are within that four-hour interval, and 44% are on either Sunday or Saturday. It should be noted, however, that the differences among the MAPEs reported in Table 3 are not due solely to these large observations, as the pattern of unpredictability remains roughly the same when the means are recalculated without considering the errors above 5%.

To conclude this review of the forecasting performance of the hourly models, we assess their relative performance with respect to two benchmarks. Some years ago REE contracted two external organizations to develop alternative forecasting procedures; both of them are proprietary and we do not have permission to describe their details in this paper. The first alternative (benchmark 1) combines seasonal decomposition, exponential smoothing and neural networks; it includes temperature as an explanatory variable, but does not make any allowance for special days. The second procedure (benchmark 2) combines seasonal ARIMA models and artificial neural networks, includes temperature variables, and models special days explicitly.

It was found that the existing models were more accurate than both benchmarks, so the latter were discarded for on-line operation. However, it was also decided that the two benchmarks would be used off-

Table 4

Mean absolute percentage errors for one day ahead hourly forecasting: details by time of the year and types of day

Hours	Time of the year		Types of day			
	Non-summer	Summer	Weekdays, normal	Weekends + Weekdays, special	Weekends	Weekdays, special
1 A.M.	1.39	1.25	1.28	1.44	1.40	1.54
2 A.M.	1.44	1.23	1.25	1.54	1.54	1.55
3 A.M.	1.48	1.10	1.23	1.53	1.51	1.56
4 A.M.	1.47	1.11	1.24	1.51	1.56	1.39
5 A.M.	1.55	1.07	1.28	1.54	1.57	1.49
6 A.M.	1.47	0.99	1.17	1.52	1.43	1.70
7 A.M.	1.47	1.05	1.14	1.59	1.40	2.03
8 A.M.	1.56	1.24	1.20	1.80	1.55	2.39
9 A.M.	1.81	1.35	1.38	2.06	1.78	2.71
10 A.M.	1.76	1.37	1.40	1.95	1.67	2.59
11 A.M.	1.67	1.46	1.40	1.88	1.65	2.39
12 P.M.	1.66	1.47	1.41	1.86	1.58	2.51
1 P.M.	1.66	1.51	1.42	1.87	1.61	2.48
2 P.M.	1.70	1.46	1.42	1.90	1.64	2.49
3 P.M.	1.70	1.48	1.43	1.90	1.66	2.44
4 P.M.	1.88	1.55	1.49	2.16	1.84	2.88
5 P.M.	2.02	1.67	1.64	2.29	2.03	2.90
6 P.M.	1.98	1.73	1.64	2.26	2.07	2.71
7 P.M.	1.94	1.67	1.53	2.31	2.15	2.67
8 P.M.	1.82	1.58	1.49	2.10	1.94	2.47
9 P.M.	1.69	1.50	1.37	1.99	1.82	2.37
10 P.M.	1.46	1.27	1.17	1.72	1.63	1.94
11 P.M.	1.38	1.41	1.26	1.59	1.48	1.82
12 A.M.	1.55	1.55	1.41	1.74	1.76	1.69
All hours	1.65	1.38	1.36	1.84	1.68	2.20

line to help in monitoring the performance of the preferred models. For doing so, hourly forecasts are computed from the three methods at the same time and based on the same information, so that they are fully comparable. They are not comparable, however, to the true historical errors reported in previous tables for a variety of reasons: they are usually computed on normal weekdays at 10 A.M. for the entire next day, so in practice they are closer to a two-day ahead forecast; on Friday the full weekend and Monday are forecasted; when there is a public holiday, the forecasts are computed at the previous day up to the next working day; on vacation periods there are some missing data (39 full days in 2006); etc.

The results are summarized in Table 5 and Fig. 8. Table 5 sketches the main features by focusing on daily figures. The forecasting performance of each method is measured by the daily MAPE and the number of daily errors above 5%. The daily MAPE is the average of the 24 hourly MAPEs; the number of large errors is the number of days with a daily absolute percentage error

(APE) above 5%, daily APE being computed as the average of the 24 hourly APEs. The results are reported both for the whole set of data and separately by type of day. In addition, Fig. 8 compares the 24 hourly MAPEs for each type of day, and gives a more detailed picture of the relative performance of the methods.

Table 5 and Fig. 8 show that the current models outperform the benchmarks in all possible comparisons. They confirm that modelling special days is a major need in forecasting the load, by comparing the performance of benchmark 1 with that of the two other procedures. With regard to the comparison between the current models and benchmark 2, it can be seen from Table 5 that the daily MAPEs are between 0.6% and 1% lower in the current models. In addition, in Fig. 8 the four curves of the current models are always below the related curves of the second benchmark. Finally, it was noted in Table 4 that the forecasting performance of the current models is not as satisfactory for special days as it is for normal weekdays. The differences, however, are smaller for the current



Table 5

Forecasting comparison among the current hourly models and two third-party benchmarks: daily mean absolute percentage errors and number of errors above 5% by type of day

Type of day	Number of observations	Mean Absolute Percentage Errors (MAPE)			Number of errors above 5%		
		Current models	Benchmark 1	Benchmark 2	Current models	Benchmark 1	Benchmark 2
All days	326	1.66	3.21	2.32	2	39	26
Weekdays, normal	196	1.50	2.15	2.09	0	8	10
Weekends + weekdays, special	130	1.91	4.81	2.68	2	31	16
Weekends	92	1.86	2.91	2.52	2	10	9
Weekdays, special	38	2.04	9.42	3.06	0	21	7

Note: The daily MAPE is computed as the average of the 24 hourly MAPEs; the number of errors above 5% is based on the daily average percentage error, which is computed as the average of the 24 hourly APEs for that day. The actual lead time depends on a number of factors, but is always the same for the three competing procedures.

models than for benchmark 2, so it seems that the former are more successful in detecting the pattern of systematic variation of special days, and in using it to improve the forecasts.

## 5. Summary and conclusions

In this paper we have presented the experience of *Red Eléctrica de España (REE)*, the Spanish

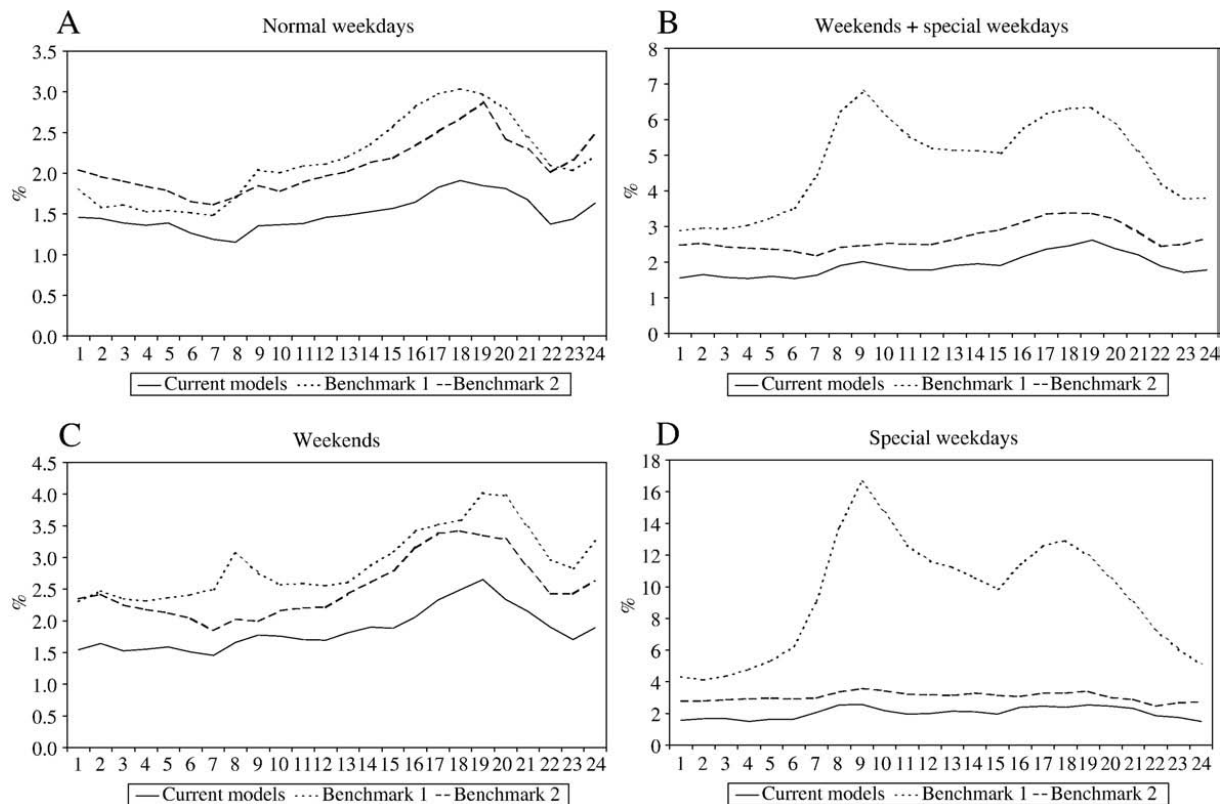


Fig. 8. Forecasting comparison among the current hourly models and two third-party benchmarks: hourly mean absolute percentage errors by type of day.

transmission system operator, in forecasting the electricity load. Adopting the point of view of a specific agent within the power system is one of the distinguishing features of the paper. It leads to an explicit statement of the actual needs of the organization and the kind of information forecasts are asked to provide, and both of these factors determine the way the forecasting activity should be approached and how it should be evaluated.

Even though there is a wide, systematic production of forecasts at specific moments of the day or week at REE, there are two types of forecasts that are of particular interest: the daily forecasts that are calculated on Wednesday morning for the seven-day period beginning on the following Saturday, which are used for outage planning; and every day's hourly forecasts, which are used to set up the final hourly dispatch schedule for the next day. Both categories are critical inputs for REE's activities, to the extent that forecasting success should be mainly assessed in terms of their errors. A point that has been stressed throughout the paper is that in applied forecasting we need to identify the needs of the particular agent, set up adequate devices specifically designed to meet such needs, and monitor performance by focusing on the errors that have significant economic impacts according to the activities of the organization.

The forecasting system currently implemented at REE consists of one daily model for forecasting the daily load up to ten days ahead, and 24 hourly models for computing hourly predictions for horizons up to three days. All of the models have the same basic structure, although their parameters are allowed to adapt to the particular features of each time series.

The forecasting performance was evaluated by analysing different types of errors for the year 2006. The forecasting performance on special weekdays is worse than on normal weekdays, although the difference between the MAPEs is small, and indicates that the model has been quite successful in explaining most of the distortions induced by the occurrence of a special day. In regard to the hourly forecasts for the next-day dispatch schedule, it was found that the most difficult period to predict is Saturday afternoon, as it behaves like a Sunday in some aspects and like a weekday in others.

The performance of the hourly models was compared to two benchmarks. A distinguishing feature of this paper is that the benchmarks are competing methods developed by external organizations to challenge the

existing procedure, and even to become the main forecasting devices. Nevertheless, the current models outperformed the benchmarks in all comparisons. In particular, the exercise showed that the major improvement in error reduction comes from understanding how the load reacts to special days, and integrating their influence, together with the effects of weather variables, in a comprehensive model that captures the main determinants of the electricity load.

### Acknowledgments

We gratefully acknowledge the invaluable help and support we have always received from REE in the development of this project, with special mentions for Agustín Maure, Technical Secretary; Ventura Rodríguez, head of the Department of Statistics and Information; and Juan Moreno, Operation Director. We are also grateful to an anonymous referee and the guest editor James Taylor for their very useful comments. The second author gratefully acknowledges financial support from the Spanish Ministry of Education with the SEJ2006-03919/ECON grant.

### References

- Alves da Silva, A. P., Ferreira, V. H., & Velasquez, R. (2008). Input space to neural network based load forecasting. *International Journal of Forecasting*, 24, 616–629 (this issue), doi:10.1016/j.ijforecast.2008.07.006.
- Cancelo, J. R., & Espasa, A. (1996). Modelling and forecasting daily series of electricity demand. *Investigaciones Económicas*, 20, 359–376.
- Cancelo, J. R., & Espasa, A. (2001). Using high-frequency data and time series models to improve yield management. *International Journal of Services Technology and Management*, 2, 59–70.
- Cottet, R., & Smith, M. (2003). Bayesian modeling and forecasting of intraday electricity load. *Journal of the American Statistical Association*, 98, 839–849.
- Darbellay, G. A., & Slama, M. (2000). Forecasting the short-term demand for electricity. Do neural networks stand a better chance? *International Journal of Forecasting*, 16, 71–83.
- Engle, R. F., Granger, C. W. J., Rice, J., & Weiss, A. (1986). Semiparametric estimates of the relation between weather and electricity sales. *Journal of the American Statistical Association*, 81, 310–320.
- Espasa, A., Revuelta, J. M., & Cancelo, J. R. (1996). Automatic modelling of daily series of economic activity. In A. Prat (Ed.), *Proceedings in computational statistics COMPSTAT 1996* (pp. 51–63). Heidelberg: Physica-Verlag.

- Harvey, A., & Koopman, S. J. (1993). Forecasting hourly electricity demand using time-varying splines. *Journal of the American Statistical Association*, 88, 1228–1236.
- Henley, A., & Peirson, J. (1997). Non-linearities in electricity demand and temperature: parametric versus non-parametric methods. *Oxford Bulletin of Economics and Statistics*, 59, 149–162.
- Hippert, H. S., Bunn, D. W., & Souza, R. C. (2005). Large neural networks for electricity load forecasting: are they overfitted? *International Journal of Forecasting*, 21, 425–434.
- Hyde, O., & Hodnett, P. F. (1997). An adaptable automated procedure for short-term electricity load forecasting. *IEEE Transactions on Power Systems*, 12, 84–94.
- Lamedica, R., Prudenzi, A., Sforna, M., Caciotta, M., & Orsolini Cencelli, V. (1996). A neural network based technique for short-term forecasting of anomalous load periods. *IEEE Transactions on Power Systems*, 11, 1749–1756.
- Le Comte, D. M., & Warren, H. E. (1981). Modeling the impact of summer temperatures on national electricity consumption. *Journal of Applied Meteorology*, 20, 1415–1419.
- Nevis, E. C., DiBella, A. J., & Gould, J. M. (1995). Understanding organizations as learning systems. *Sloan Management Review*, 36, 73–85.
- Pardo, A., Meneu, V., & Valor, E. (2002). Temperature and seasonality influences on Spanish electricity load. *Energy Economics*, 24, 55–70.
- Ramanathan, R., Engle, R., Granger, C. W. J., Vahid-Araghi, F., & Brace, C. (1997). Short-run forecasts of electricity loads and peaks. *International Journal of Forecasting*, 13, 161–174.
- Ranaweera, D. K., Karady, G. G., & Farmer, R. G. (1997). Economic impact analysis of load forecasting. *IEEE Transactions on Power Systems*, 12, 1388–1392.
- Smith, M. (2000). Modeling and short-term forecasting of New South Wales electricity system load. *Journal of Business and Economic Statistics*, 18, 465–478.
- Soares, L. J., & Souza, L. R. (2006). Forecasting electricity demand using generalized long memory. *International Journal of Forecasting*, 22, 17–28.
- Taylor, J. W. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of the Operational Research Society*, 54, 799–805.
- Taylor, J. W., de Menezes, L. M., & McSharpy, P. E. (2006). A comparison of univariate methods for forecasting electricity demand up to a day ahead. *International Journal of Forecasting*, 22, 1–16.
- Valor, E., Meneu, V., & Caselles, V. (2001). Daily air temperature and electricity load in Spain. *Journal of Applied Meteorology*, 40, 1413–1421.

**José Ramón Cancelo** holds an MSc. in Economics from the Universidad de Santiago, and a Ph.D. in Economics from the Universidad Autónoma de Madrid. Currently he is full professor of Business and Economic Statistics at the Universidade da Coruña. His research interests include time series analysis, dynamic models, forecasting and signal extraction.

**Antoni Espasa** is Professor of Econometrics at the Universidad Carlos III de Madrid and director of the research institute Flores de Lemus at the same university. He obtained his Ph.D. from the LSE in 1975. He has been chief economist at the Research Unit of the Bank of Spain. In 1991 he was awarded the national prize in economics Rey Jaime I. He is a fellow of the European Economic Association, and is director of the publication *Bulletin of EU and US Inflation and Macroeconomic Analysis*. His published research includes studies in spectral econometrics, dynamic econometric models, forecasting and modelling daily and hourly series.

**Rosmarie Grafe** has an MSc. in Economics from the Universidad del País Vasco at Bilbao, Spain. She has spent most of her career with Red Eléctrica de España (REE), working at the Operational Direction in the national electricity demand area. She is responsible for the forecasting models at REE. Her interest in forecasting topics started with a master on “Time Series Analysis and Dynamic Econometric Models” at the Centro de Estudios del Banco de España, Madrid.