



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

International Journal of Forecasting 22 (2006) 1–16

*international journal  
of forecasting*

[www.elsevier.com/locate/ijforecast](http://www.elsevier.com/locate/ijforecast)

# A comparison of univariate methods for forecasting electricity demand up to a day ahead

James W. Taylor<sup>a,\*</sup>, Lilian M. de Menezes<sup>b</sup>, Patrick E. McSharry<sup>c</sup>

<sup>a</sup>*Saïd Business School, University of Oxford, Park End Street Oxford OX1 1HP, UK*

<sup>b</sup>*Cass Business School, City University, UK*

<sup>c</sup>*Department of Engineering, University of Oxford, UK*

---

## Abstract

This empirical paper compares the accuracy of six univariate methods for short-term electricity demand forecasting for lead times up to a day ahead. The very short lead times are of particular interest as univariate methods are often replaced by multivariate methods for prediction beyond about six hours ahead. The methods considered include the recently proposed exponential smoothing method for double seasonality and a new method based on principal component analysis (PCA). The methods are compared using a time series of hourly demand for Rio de Janeiro and a series of half-hourly demand for England and Wales. The PCA method performed well, but, overall, the best results were achieved with the exponential smoothing method, leading us to conclude that simpler and more robust methods, which require little domain knowledge, can outperform more complex alternatives.

© 2005 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.

*Keywords:* Electricity demand forecasting; Exponential smoothing; Principal component analysis; ARIMA; Neural networks

---

## 1. Introduction

Electricity-supply planning requires efficient management of existing power systems and optimisation of the decisions concerning additional capacity. Demand prediction is an important aspect in the development of any model for electricity planning.

The form of the demand depends on the type of planning and the accuracy that is required; hence it can be represented as an annual demand (GW), a peak demand (MW), or load duration curves like daily, weekly or annual. Short-term load forecasts are required for the control and scheduling of power systems. The focus varies from minutes to several hours ahead. The predictions are required as inputs to scheduling algorithms for the generation and transmission of electricity. The load forecasts help in determining which devices to operate in a given period, so as to minimise costs and secure

---

\* Corresponding author. Tel.: +44 1865 288927; fax: +44 1865 288805.

E-mail address: [james.taylor@sbs.ox.ac.uk](mailto:james.taylor@sbs.ox.ac.uk) (J.W. Taylor).

demand even when local failures may occur in the system.

In the short run, the load is mainly influenced by meteorological conditions, seasonal effects (daily and weekly cycles, calendar holidays) and special events. Weather related variation is certainly critical in predicting electricity demand for lead times beyond a day-ahead (Chow & Leung, 1996; Taylor & Buizza, 2003). However, when the interest is in shorter lead times, a univariate model will be sufficient. Indeed, univariate models are the norm for lead times up to about six hours ahead, and, due to the lack of readily available weather forecasts, they are sometimes used for longer lead times. This has been the case in Brazil where access to weather variables is often difficult and expensive (Souza & Soares, 2003). In this paper, we compare the accuracy of two simple benchmarks and four more sophisticated univariate methods, including a new regression based method that uses principal component analysis. We evaluate the methods using intra-day data for the state of Rio de Janeiro in Brazil and for England and Wales. We consider lead times up to a day ahead, but given that we consider only univariate methods, prediction for

lead times less than six hours ahead is of particular interest.

Fig. 1 shows hourly demand in Rio for a fortnight in May 1996. A within-day seasonal cycle, of duration 24 periods, is apparent from the similarity of the demand profile from one day to the next, particularly on weekdays. A within-week seasonal cycle, of duration 168 periods, is evident when one compares the demand on the corresponding day of adjacent weeks. Fig. 2 shows half-hourly demand in England and Wales for a fortnight in late March to early April 2000. The plot shows a within-day seasonal cycle of duration 48 periods and a within-week seasonal cycle of duration 336 periods. The short-term forecasts of National Grid Transco, which is the transmission company for England and Wales, are important not only for scheduling load, but also because the one hour ahead forecasts are a key input to the balancing market, which operates on a rolling one hour ahead basis to balance supply and demand after the closure of bi-lateral trading between generators and suppliers.

In the next section, we review the literature on short-term load forecasting. We then describe the two series that we use to compare the different meth-

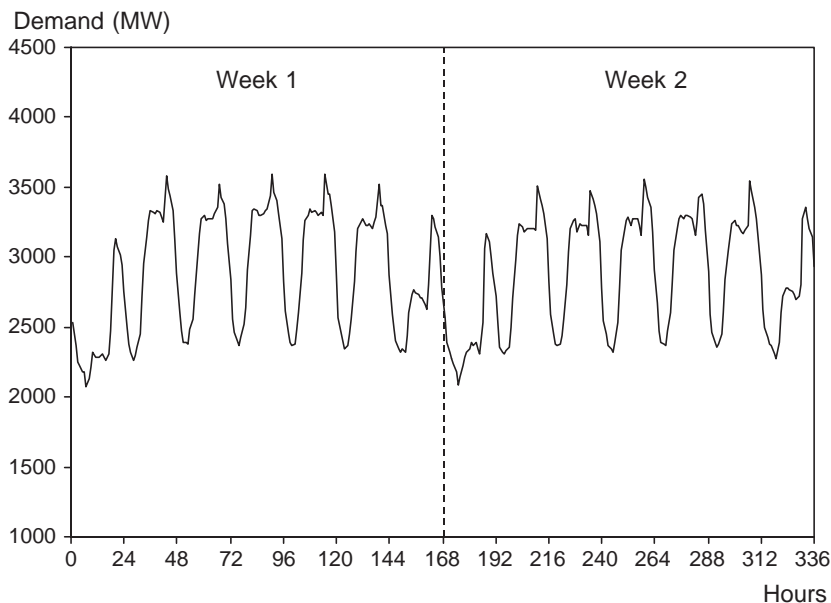


Fig. 1. Hourly electricity demand in Rio de Janeiro from Monday, 5 May 1996 to Sunday, 18 May 1996.

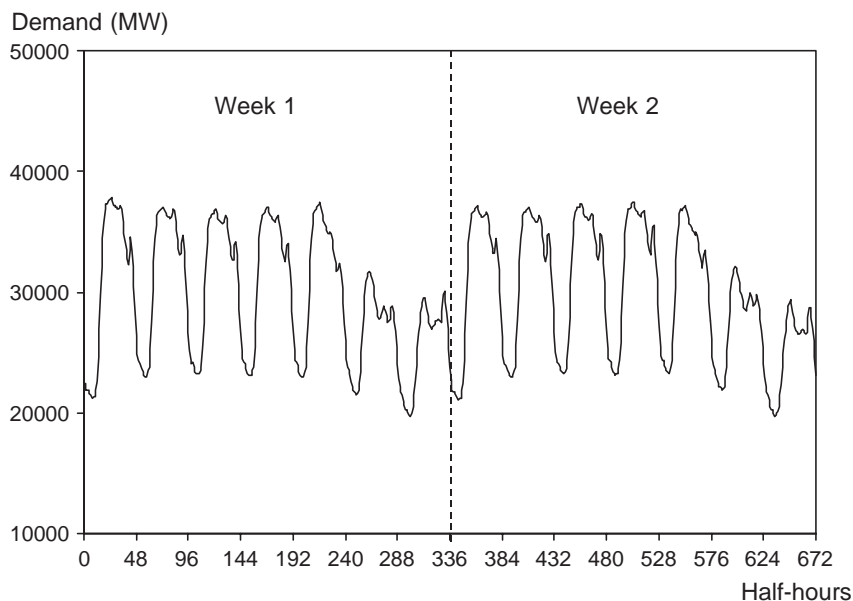


Fig. 2. Half-hourly electricity demand in England and Wales from Monday, 27 March 2000 to Sunday, 9 April 2000.

ods. In Section 4, we present the methods, and in Section 5 we report on their relative performances for the two load time series. In the final section, we provide a summary and conclusion.

## 2. Review of methods for short-term univariate prediction

The stochastic nature of demand as a function of time has frequently been modelled with seasonal ARIMA and state space models. The latter became very attractive in the 1980s because of the computational efficiency of the Kalman filter (e.g., Campo & Ruiz, 1987) and good performances are still reported in more recent work (Infield & Hill, 1998). ARIMA modelling is used by many, within a univariate framework, as a sophisticated benchmark for evaluating alternative proposals (see, for example, Abraham & Nath, 2001; Darbellay & Slama, 2000; Laing & Smith, 1987; Taylor, 2003). Simpler methods, such as nonparametric regression (Charytoniuk, Chen, & Van Olinda, 1998) and general exponential smoothing (Christiaanse, 1971), are always attractive due to the small number of parameters involved, which make them easy to implement.

The 1990s brought computer power into load forecasting and with it the opportunity to explore the complexity in the load data. Rule-based and fuzzy logic expert systems have been used to model the complexity in the data using domain knowledge. Although these methods are promising, they rely on rules that are extracted from experts' and operators' experience, which are subject to inconsistencies and are thus unreliable. Artificial neural networks have made it possible to experiment with theoretically poor, but data rich, models that can identify the complex non-linear relationships in the data and infer future behaviour. The basic idea is that the networks learn through examples, which consist of the input signals and desired output. The result is that neural networks have the potential to model a far greater range of relationships than models that have a pre-specified form like ordinary least squares linear regressions.

In forecasting electricity demand with neural networks, most work has used multi-layer perceptron networks. Early studies concentrated on forecasting the daily peak, valley or total load, or on forecasting the whole daily profile at once. Many have involved non-linear adaptations of the Box–Jenkins methodology (Hippert, Pedreira, & Souza, 2001). Asar and McDonald (1994) present a univariate network for

predicting peak loads. Darbellay and Slama (2000) present a neural network, built only from historical hourly load, which matched an ARIMA model for prediction from an hour ahead to 36 hours ahead. Several studies focused on short-term load forecasting and took one of two approaches. In the first, a neural network is constructed for each type of day or subset of hours (e.g., 1:00–9:00, 10:00–19:00, 20:00–24:00). The network is then fed with the relevant input variables for the corresponding type (Lee, 1992). In the second approach, a single neural network is developed, but within the input variables information concerning the specific type of day or hour is required (e.g., Daneshdoot, Lotfalian, Bumroonggit, & Ngoy, 1998).

Among those who have focused on univariate models, Liu et al. (1996) address very short-term load forecasting. They describe three approaches: a fuzzy logic-based method, an artificial neural network and an AR model-based approach to predict from one minute to 30 minutes ahead based on the previous 30 minutes of metered data. They conclude that both the fuzzy logic and neural network systems outperformed the AR model. The AR model, however, did not require pre-training and was fairly accurate up to five minutes ahead.

Although some success has been reported with the use of applied neural networks for univariate short term load forecasting (e.g., Hippert et al., 2001), there remains little systematic evidence in favour of a particular architecture, nor on the amount of training data that should be used (Rui & El-Keib, 2004). Many

recent papers in applying a neural network on its own or combined with fuzzy logic present relatively small MAPEs for real and/or simulated data, but these are often not compared to the performance of standard approaches on the same data. Consequently, it is not clear whether neural networks and more complex non-linear models outperform simpler and more standard forecasting procedures, such as ARIMA modelling.

In spite of the numerous methods that have been reported, we have to agree with Piras and Buchenel (1999) that no method has been found to be clearly better than others. In this paper, we compare the forecast accuracy of six univariate methods. The four main methods that we consider are: (1) multiplicative seasonal ARIMA; (2) an exponential smoothing alternative because of its robustness which makes it attractive for online demand forecasting; (3) an artificial neural network implementation that showed good performance for a similar type of data; and (4) a principal component analysis approach, which draws inspiration from decomposition and regression alternatives. In addition, we include a seasonal version of the standard naïve random walk benchmark model and we develop a benchmark that is better suited to the nature of load data.

### 3. The two load time series

To evaluate forecasting performance, we used 30 weeks of hourly observations for electricity demand in Rio and 30 weeks of half-hourly observations for

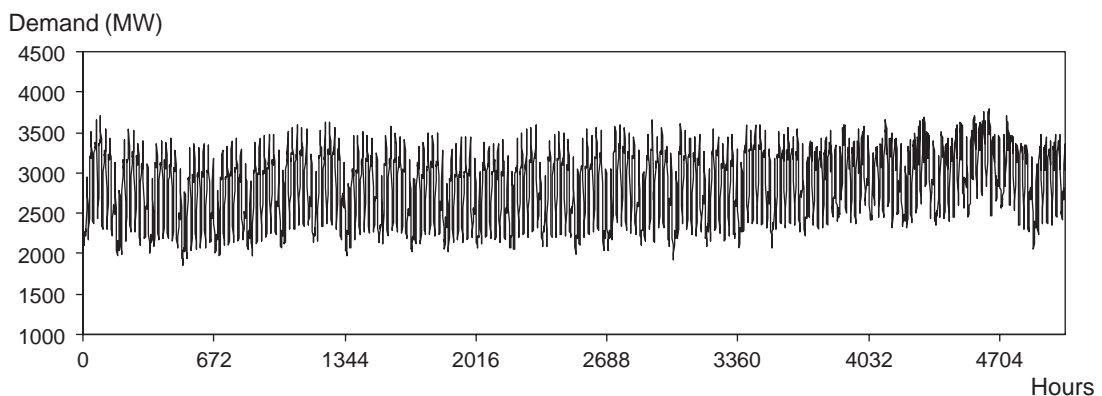


Fig. 3. Hourly electricity demand in Rio de Janeiro for Sunday, 5 May 1996 to Saturday, 30 November 1996.

demand in England and Wales. Hourly data has traditionally been used for short-term forecasting in Brazil, and half-hourly in England and Wales. The Rio data covers the period from Sunday 5 May 1996 to Saturday 30 November 1996. The England and Wales data was from the period Monday 27 March 2000 to Sunday 22 October 2000. The series are shown in Figs. 3 and 4.

As in many other studies of electricity demand, prior to fitting and evaluating methods, we elected to smooth out the ‘special days’, such as bank holidays, as their inclusion is likely to be unhelpful in our comparison of methods. Demand on these days is so very unlike the rest of the year that online univariate methods are generally unable to produce reasonable forecasts. In practice, interactive facilities tend to be used for special days, which allow operator experience to supplement or override the system offline. If a forecasting method is unable to tolerate gaps in the historical series, the special days can be smoothed over, leaving the natural periodicities of the data intact (Laing & Smith, 1987). An alternative to this would be to treat the special days as missing observations.

For each series, we used the first 20 weeks of data to estimate method parameters and the remaining 10 weeks to evaluate post-sample accuracy of forecasts up to 24 hours ahead. For the Rio series, this amounted to 3360 hourly observations for estimation and 1680 for evaluation. For the England and Wales series, this amounted to 6720 half-hourly observations for estimation and 3360 for evaluation.

#### 4. Forecasting methods

##### 4.1. Double seasonal ARMA modelling

A short-term forecasting method that has remained popular over the years, and often appears as a benchmark approach, is multiplicative seasonal ARIMA modelling. The multiplicative double seasonal ARIMA model (see Box, Jenkins, & Reinsel, 1994, p. 333) can be written as

$$\begin{aligned} \phi_p(L)\Phi_{P_1}(L^{s_1})\Omega_{P_2}(L^{s_2})\nabla^d\nabla_{s_1}^{D_1}\nabla_{s_2}^{D_2}(y_t - c) \\ = \theta_q(L)\Theta_{Q_1}(L^{s_1})\Psi_{Q_2}(L^{s_2})\varepsilon_t \end{aligned} \quad (1)$$

where  $y_t$  is demand in period  $t$ ;  $c$  is a constant term;  $s_1$  and  $s_2$  are the number of periods in the different seasonal cycles;  $L$  is the lag operator;  $\nabla$  is the difference operator;  $\nabla_{s_2}$  and  $\nabla_{s_1}$  are seasonal difference operators (e.g.,  $\nabla_{s_1} = (1 - L^{s_1})$ );  $d$ ,  $D_1$  and  $D_2$  are the orders of differencing;  $\varepsilon_t$  is a white noise error term; and  $\phi_p$ ,  $\Phi_{P_1}$ ,  $\Omega_{P_2}$ ,  $\theta_q$ ,  $\Theta_{Q_1}$ , and  $\Psi_{Q_2}$  are polynomial functions of orders  $p$ ,  $P_1$ ,  $P_2$ ,  $q$ ,  $Q_1$ , and  $Q_2$ , respectively. This model can be expressed as ARIMA  $(p,d,q) \times (P_1,D_1,Q_1)_{s_1} \times (P_2,D_2,Q_2)_{s_2}$ . Applying the model to the hourly Rio demand series, we set  $s_1=24$  to model the within-day seasonal cycle of 24 hours, and  $s_2=168$  to model the within-week cycle of 168 hours. For the England and Wales half-hourly series, we set  $s_1=48$  and  $s_2=336$ .

The process of model identification is impractical in an online demand forecasting system, and so the model is chosen offline. We followed the Box–Jenkins methodology to identify the most suitable ARIMA

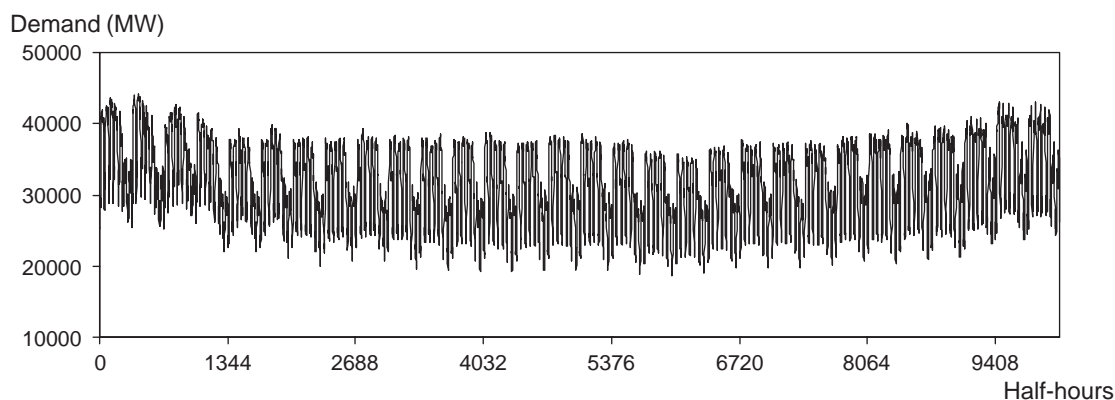


Fig. 4. Half-hourly electricity demand in England and Wales for Monday, 27 March 2000 to Sunday, 22 October 2000.

model based on the 20-week estimation sample for each series. The autocorrelation function and partial autocorrelation function were used to select the order of the model, which was then estimated by maximum likelihood. The residuals were inspected for any remaining autocorrelation. Laing and Smith (1987) explain that, in the multiplicative double seasonal ARIMA formulation in expression (1), polynomials of order greater than two are rarely necessary when fitting a model to half-hourly data for England and Wales. We also found for our England and Wales data that polynomials of order two were sufficient. However, this was not sufficient for the Rio data because significant autocorrelation remained in the residuals. By using polynomials up to order three, we eliminated the residual autocorrelation. We compared the Schwartz Bayesian information Criterion (SBC) for an extensive range of different ARIMA models. For the Rio series, the methodology resulted in the following ARIMA(3,0,3) × (3,0,3)<sub>24</sub> × (3,0,3)<sub>168</sub> model:

$$\begin{aligned} & (1 - 2.16L + 1.39L^2 - 0.23L^3) \\ & \times (1 - 0.22L^{24} - 0.12L^{48} - 0.36L^{72}) \\ & \times (1 - 0.20L^{168} - 0.45L^{336} - 0.34L^{504}) \\ & \times (y_t - 2,753) = (1 - 1.34L + 0.25L^2 + 0.09L^3) \\ & \times (1 - 0.36L^{72})(1 - 0.27L^{336} - 0.32L^{504})\varepsilon_t. \end{aligned}$$

Using the same methodology for the England and Wales series, we produced the following ARIMA(2,0,1) × (2,0,1)<sub>48</sub> × (1,0,2)<sub>336</sub> model:

$$\begin{aligned} & (1 - 1.40L + 0.42L^2)(1 - 1.04L^{48} + 0.09L^{96}) \\ & \times (1 - 0.97L^{336})(y_t - 26,441) \\ & = (1 - 0.20L)(1 - 0.68L^{48}) \\ & \times (1 - 0.61L^{336} - 0.09L^{672})\varepsilon_t. \end{aligned}$$

#### 4.2. Exponential smoothing for double seasonality

The widespread use of exponential smoothing in automated applications, such as inventory control, led Taylor (2003) to consider the use of the approach for online electricity demand forecasting. The application requires an extension of the standard Holt–Winters exponential smoothing formulation to accommodate the two seasonal cycles in the electricity demand series. This involves the introduction of an additional

seasonal index and an extra smoothing equation for the new seasonal index. The formulation for double multiplicative seasonality is given in the following expressions:

$$S_t = \alpha(y_t / (D_{t-s_1} W_{t-s_2})) + (1 - \alpha)(S_{t-1} + T_{t-1}) \quad (2)$$

$$T_t = \gamma(S_t - S_{t-1}) + (1 - \gamma)T_{t-1} \quad (3)$$

$$D_t = \delta(y_t / (S_t W_{t-s_2})) + (1 - \delta)D_{t-s_1} \quad (4)$$

$$W_t = \omega(y_t / (S_t D_{t-s_1})) + (1 - \omega)W_{t-s_2} \quad (5)$$

$$\begin{aligned} \hat{y}_t(k) = & (S_t + kT_t)D_{t-s_1+k}W_{t-s_2+k} \\ & + \phi^k(y_t - ((S_{t-1} + T_{t-1})D_{t-s_1}W_{t-s_2})). \quad (6) \end{aligned}$$

$S_t$  and  $T_t$ , are the smoothed level and trend;  $D_t$  and  $W_t$  are the seasonal indices for the intraday and intraweek seasonal cycles, respectively;  $\alpha$ ,  $\gamma$ ,  $\delta$  and  $\omega$  are the smoothing parameters; and  $\hat{y}_t(k)$  is the  $k$  step-ahead forecast made from forecast origin  $t$ . The term involving the parameter  $\phi$ , in the forecast function expression Eq. (6), is a simple adjustment for first-order autocorrelation. All the parameters in the method,  $\alpha$ ,  $\gamma$ ,  $\delta$ ,  $\omega$  and  $\phi$ , are estimated in a single procedure by minimizing the sum of squared one step-ahead in-sample errors. The initial smoothed values for the level, trend and seasonal components are estimated by averaging the early observations. As with the double-seasonal ARIMA model, applying the method to the Rio hourly data, we set  $s_1=24$  and  $s_2=168$ , while for the half-hourly England and Wales series,  $s_1=48$  and  $s_2=336$ .

In our implementation of the method, we calculated initial smoothed values for the level, trend and seasonal components using the approach described by Taylor (2003), which involves averaging the early observations. Table 1 shows the parameters that we

Table 1

Parameters estimated from the 20-week estimation samples for double seasonal exponential smoothing with an AR(1) model for the residuals

	Level $\alpha$	Trend $\gamma$	Within-day seasonality $\delta$	Within-week seasonality $\omega$	AR $\lambda$
Rio	0.01	0.00	0.09	0.15	0.88
England and Wales	0.02	0.04	0.32	0.15	0.98

estimated. The parameters for the England and Wales series are reasonably similar to those reported by Taylor for a shorter series of England and Wales demand. The high value of  $\phi$  and low value of  $\alpha$ , for both series, reflects the fact that the adjustment for first-order autocorrelation has, to a large degree, made redundant the smoothing equation for the level. The low values for the trend smoothing parameter,  $\gamma$ , were accompanied by very small values for the smoothed trend,  $T_t$ . This seems reasonable since variation in the 20-week estimation periods are dominated by seasonality.

An important point to note regarding the double seasonal exponential smoothing approach is that, by contrast with ARIMA modelling and the majority of other approaches to short-term demand forecasting, there is no model specification involved. This gives the method strong appeal in terms of simplicity and robustness.

#### 4.3. Artificial neural network

We felt that it was important to include a neural network in our study because this approach has received a great deal of attention in the load forecasting literature in recent years. However, a basic problem that we faced was which form of neural network to use. The neural network load forecasting literature contains many different approaches and designs, but there is no consensus as to the preferred form. We opted to implement a similar neural network modelling approach to the one described relatively recently in this journal by Darbellay and Slama (2000) for their forecasting of hourly demand in the Czech Republic. We chose this approach because it was used for a univariate application involving similar data to ours. We are, therefore, providing more evidence in an area where replication with similar data is rare.

We used a single hidden layer feedforward network, which is the most widely-used neural network for forecasting (Zhang, Patuwo, & Hu, 1998). It consists of a set of  $k$  inputs, which are connected to each of  $m$  units in a single hidden layer, which, in turn, are connected to an output. In regression terminology, the inputs are explanatory variables,  $x_{it}$ , and the output is the dependent variable,  $y_t$ . In this study, the output is demand and the inputs

are lag demand. The resultant model can be written as

$$f(\mathbf{x}_t, \mathbf{v}, \mathbf{w}) = g_2 \left( \sum_{j=0}^m v_j g_1 \left( \sum_{i=0}^k w_{ji} x_{it} \right) \right)$$

where  $g_1(\cdot)$  and  $g_2(\cdot)$  are activation functions, which we chose to be sigmoidal and linear, respectively, and  $w_{ji}$  and  $v_j$  are the weights (parameters). We estimated the weights using the following minimisation

$$\min_{\mathbf{v}, \mathbf{w}} \left( \frac{1}{n} \sum_{t=1}^n (y_t - f(\mathbf{x}_t, \mathbf{v}, \mathbf{w}))^2 + \lambda_1 \sum_{j=0}^m \sum_{i=0}^k w_{ji}^2 + \lambda_2 \sum_{j=0}^m v_j^2 \right)$$

where  $n$  is the number of in-sample observations, and  $\lambda_1$  and  $\lambda_2$  are regularisation parameters which penalise the complexity of the network and thus avoid overfitting (see Bishop, 1997, §9.2). We established suitable values for  $\lambda_1$  and  $\lambda_2$  and for the number,  $m$ , of units in the hidden layer using a hold out method with a third of the estimation sample used for testing (see Bishop, 1997, §9.8).

Neural networks will perform poorly if the time series is nonstationary (Cottrell, Girard, Girard, Mangeas, & Muller, 1995). In our ARIMA modelling of the estimation samples of data for both the Rio and England and Wales data, we found no justification for differencing. Indeed, we found that the best fit, assessed in terms of SBC, resulted when no differencing was used. However, Fig. 4 shows that the within-year cycle in the England and Wales data appears in our 30-week sample as a trend through the data. The plot in Fig. 3 for the Rio data also shows a slight rise in the level of the data, which is due to demand rising in response to the need for air conditioning in the warmer months. In their ARMA analysis, apparent nonstationarities led Darbellay and Slama (2000) to apply to their hourly demand data,  $y_t$ , the following differencing operator:  $(1-L)(1-L^{24})(1-L^{168})y_t$ . Prior to fitting the neural network, we applied the same operator to our hourly Rio data, and for the England and Wales half-hourly data, we used  $(1-L)(1-L^{48})(1-L^{336})y_t$ .

We drew inspiration from the study of Darbellay and Slama (2000) for our choice of lagged demand to

use as input variables. For the hourly Rio data, we considered the following lags: 1 to 6, 23, 24, 25, 48, 72, 96, 120, 144, 168, 192, 216, 240, 264, 288, 312 and 336. The hold out method led us to use  $\lambda_1 = \lambda_2 = 0.005$  and  $m = 12$ , and to use the following lags in our model: 1, 2, 24, 25, 48, 72, 96, 120, 144, 168, 192, 216, 240, 264, 312 and 336. For the half-hourly England and Wales data, we considered the following lags: 1 to 12, 47, 48, 49, 96, 144, 192, 240, 288, 336, 384, 432, 480, 528, 576, 624 and 672. The hold out method led us to use  $\lambda_1 = \lambda_2 = 0.005$  and  $m = 14$ , and the following lags: 1, 48, 49, 96, 144, 192, 240, 288, 336, 384, 432, 480, 624 and 672.

Note that a single model was estimated for each of the two demand series. As with the other models fitted in this paper, parameters were estimated using the estimation sample of data for each series, and the parameters were not updated after this. One step-ahead forecasts were iteratively used as inputs in order to generate multi-step predictions.

#### 4.4. A regression method with principal component analysis

Principal component analysis (PCA) is a standard statistical method that is used for reducing the dimension of multivariate data sets, where variables are highly correlated, to a smaller set of variables. These are linear combinations of the original variables; they are uncorrelated and explain most of the variation in the data and are thus called principal components. In a regression context, one can therefore focus on this smaller number of independent variables, rather than dealing with the large number of original variables with complex interrelationships (Chatterjee, Hadi, & Price, 1999).

Here, we employ PCA to capture the intra-day variation in electricity demand. The method can be viewed as a development of the approach of Ramathan, Engle, Granger, Vahid-Araghi, and Brace (1997) and Carnero, Koopman, and Ooms (2003) and others who model each hour (or half-hour) separately. The essence of our PCA method is that it exploits the similarity between certain hours (or half-hours) in order to reduce the number of models to be considered. We note that the approach could easily be extended to the multivariate context, if weather related variables were available.

For each day  $i$  that is considered, let the  $Y_{ij}$  denote the electricity demand on day  $i$  at intra-day period  $j$ , where  $i = 1, \dots, N$  and  $j = 1, \dots, M$ . Note that  $M = 24$  for the hourly Brazil data and  $M = 48$  for the half-hourly England and Wales data. The matrix  $\mathbf{Y}$  is then used to construct a matrix  $\mathbf{C}$  by summing over the time coordinate:  $\mathbf{C} = (N - 1)^{-1/2} \mathbf{Y}^T \mathbf{Y}$ . The standard eigenvalue decomposition yields  $\mathbf{C} = \mathbf{V} \mathbf{\Sigma}^2 \mathbf{V}^T$ , where  $\mathbf{V}$  is an orthonormal matrix satisfying  $\mathbf{V}^T \mathbf{V} = \mathbf{I} = \mathbf{V} \mathbf{V}^T$ , and  $\mathbf{\Sigma}^2$  is a positive definite diagonal matrix; the positive square roots  $\sigma_j$  of its diagonal are known as the singular values of  $\mathbf{C}$  (Strang, 1993). These singular values are typically arranged in decreasing order so that  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_M$ . The columns of  $\mathbf{V}$ ,  $\mathbf{v}_j$ , are the orthogonal basis functions, each of which contributes  $\sigma_j^2$  to the total variance. The principal components are obtained by projecting the daily electricity demand profiles onto these basis functions or reduced dimension space. That is:  $\mathbf{P} = \mathbf{Y} \mathbf{V}$ , where the  $i$ th column of  $\mathbf{P}$  is the component corresponding to the  $i$ th basis function. These elements in the column describe patterns that are associated with the different days of the week. For example, the lowest loads should be associated with Sunday and the highest with midweek.

While PCA provides a technique for modelling the intra-day pattern, the weekly seasonality has yet to be incorporated. This is achieved by regressing each principal component on six dummy variables that indicate the day of the week (Monday to Saturday with Sunday arbitrarily chosen as the baseline) plus a linear and a quadratic term. The latter two terms capture the amount of linear and quadratic growth (or decay) of the electricity demand over time and are justified within a short range by scatterplots of the data. If we had several years of data for estimating this regression, the quadratic term could easily be replaced by a harmonic function that could be identified from the time series pattern of each component. The regression model for the  $q$ th principal component is, therefore, of the following form:

$$p_q(i) = \alpha_{q_0} + \alpha_{q_1} d_1 + \dots + \alpha_{q_6} d_6 + \beta_q i + \lambda_q i^2 + error(i)$$

where  $i$  is a label for the  $i$ th day in the dataset;  $d_l$  (for  $l = 0, \dots, 6$ ) are dummy variables indicating the day of the week; and  $\alpha_{q_l}$ ,  $\beta_q$  and  $\gamma_q$  are constant parameters. This model can then be used to forecast



each of the components,  $p_q(i+1)$ , for up to one day ahead.

There are two parameters which are required for specifying the model: (i) the number of weeks,  $W$ , used to compute the principal components, and (ii) the number of principal components,  $Q$ , that are kept ( $Q \leq M$ ). The second half of the 20 week estimation sample of data for both the England and Wales and Rio data sets was used to estimate the model parameters  $W$  and  $Q$  and the coefficients in the error model. The aim was to find the model with the best mean absolute percentage error (MAPE), averaged over all the intra-day samples, for the training data. Both  $W$  and  $Q$  were varied until a minimum value of MAPE was found. The results given in Section 5 were obtained using values  $W=10$  and  $Q=24$  for the England and Wales data and  $W=10$  and  $Q=8$  for the Rio data.

Having selected  $Q$ , this model approximates the matrix of demand profiles and the forecasts for intra-day period  $j$  of day  $i+1$  are given by

$$\sum_{q=1}^Q p_q(i+1)V_{jq},$$

where the first  $Q$  columns of the matrix  $\mathbf{V}$  contain the most relevant empirical orthogonal functions.

After applying this model, we found that the errors were serially correlated. We therefore followed a similar approach to that of Taylor (2003) for his exponential smoothing method, and adjusted the model by adding an AR model of the error process. As our PCA approach is an extension of the approach of Ramanathan et al. (1997), it is interesting to note that they also incorporate an error model in order to correct for serial correlation resulting from their use of separate models for each hour of the day. In the particular case of the hourly Rio data, we define  $E_t(k)$  as the prediction error associated with a  $k$  step-ahead forecast made from origin  $t$ , which is observed to be highly correlated with both the corresponding prediction error that occurred 24 hours before,  $E_{t-24}(k)$ , as well as the last known prediction error,  $E_{t-1}(1)$ . Consequently, each future prediction error,  $E_t(k)$ , is estimated by an expression of the form:

$$E_t(k) = \alpha_0(k) + \alpha_1(k)E_{t-24}(k) + \alpha_2(k)E_{t-1}(1)$$

where the  $\alpha_i(k)$  are parameters estimated separately for each lead time,  $k$ , using LS regression applied to

the estimation sample. For the England and Wales data, a similar pattern was observed and thus the lag 24 term in the model is replaced by a lag 48 term.

#### 4.5. Two simplistic benchmark methods

The random walk is the most widely used and simplest naïve benchmark method used in forecasting studies. However, for multi-step-ahead forecasting of seasonal data, the method is likely to perform poorly. A more natural benchmark forecast is provided by the seasonal version of the random walk, which takes as a forecast the observed value for the corresponding period in the most recent occurrence of the seasonal cycle. With two seasonal cycles within the data, it seems sensible to focus on the longer cycle, so that the prediction is constructed simply as the observed value for the corresponding period in the previous week. For the hourly Rio data, the forecast function is  $\hat{y}_t(k) = y_{t+k-168}$ , where  $y_t$  is the demand in period  $t$ , and  $k$  is the forecast lead time ( $k \leq 168$ ). For the half-hourly England and Wales data,  $\hat{y}_t(k) = y_{t+k-336}$  (where  $k \leq 336$ ). In our discussion of the results in Section 5, we refer to this forecasting approach as the ‘naïve benchmark’.

We used this forecasting approach as the basis of a second benchmark method. We refined the forecasts from the method by using the error modelling approach that was described in Section 4.4 for modelling the errors from the PCA method. In our discussion of the results in Section 5, we refer to this forecasting approach as the ‘naïve benchmark with error model’.

## 5. Results

Figs. 5 and 6 show the post-sample forecasting accuracy of the six methods for lead times up to a day ahead for each of the two series. The error summary measure reported in the figures is the mean absolute percentage error (MAPE), which has been traditionally used to measure accuracy in load forecasting. Because the MAPE captures the proportionality between the forecast error and the actual load, it is preferred and easily interpreted by those in the industry. Hippert et al. (2001) write that, although the MAPE has become an industry standard, it is

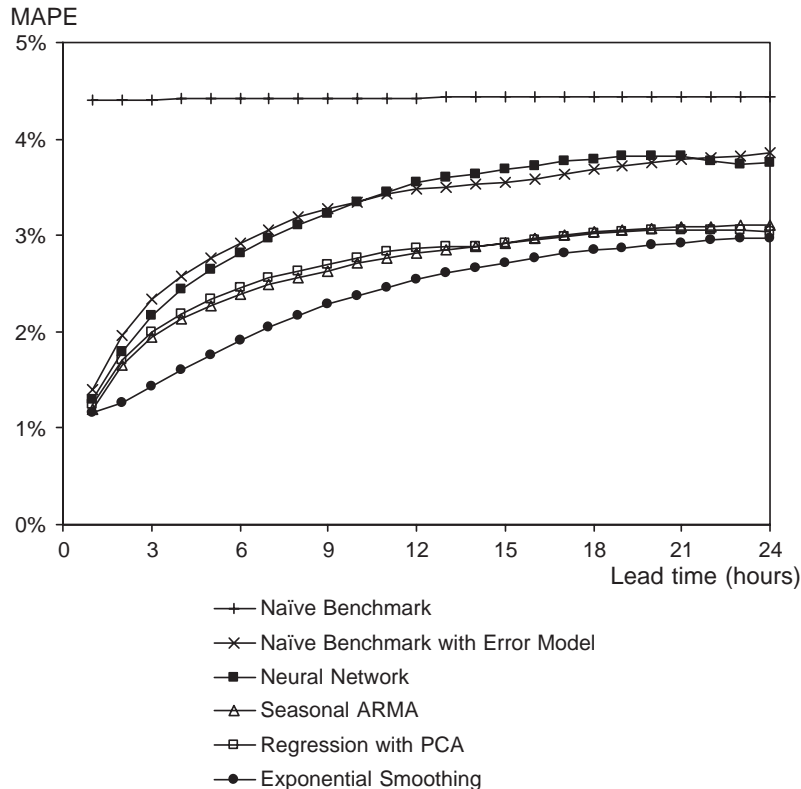


Fig. 5. MAPE results plotted against lead time for the Rio 10-week post-sample period.

advisable to also consider other summary measures. We also calculated the mean absolute error, median absolute percentage error, median absolute error, root mean squared percentage error and root mean squared error, but we do not report these results here because the relative performances of the methods for these measures were very similar to those for the MAPE.

Turning first to the Rio results in Fig. 5, we can see that the naïve benchmark was substantially outperformed by all the other methods at all lead times. Interestingly, however, Fig. 5 shows that a substantial improvement resulted from the application of the heuristic ‘error model’ to correct for autocorrelation in the forecast errors from this method. Perhaps the most surprising aspect of the results was the relatively poor performance of the neural network. It was unable to beat the naïve benchmark with error model, and it was comfortably beaten by the ARMA, exponential smoothing and PCA methods. Of these three methods, the exponential smoothing approach performed particularly well. The plot of the full sample of observations

in Fig. 3 suggests a partial explanation for this. Between periods 3360 and 4032, there seems to be a slight change in the seasonal pattern. This four-week period is the first four weeks of the 10-week post-sample evaluation period, and it is shown in Fig. 7. This figure seems to indicate a slight change in the seasonal pattern occurring a fortnight into the evaluation period. The change occurred around 6 October 1996, which was the start of summer time in Rio, and on this day the clocks were moved forward one hour. (In order to maintain 24 periods in this day, the two demand values at 00:00 were averaged.) Therefore, it would seem that all the methods have been somewhat challenged by being evaluated on data with slightly different characteristics to the data in the estimation sample. In this context, the robustness of the methods becomes important. The popularity of exponential smoothing in other applications, such as sales forecasting and financial volatility forecasting, is at least partly due to its robustness. We, therefore, conclude that this is a reason for the success of

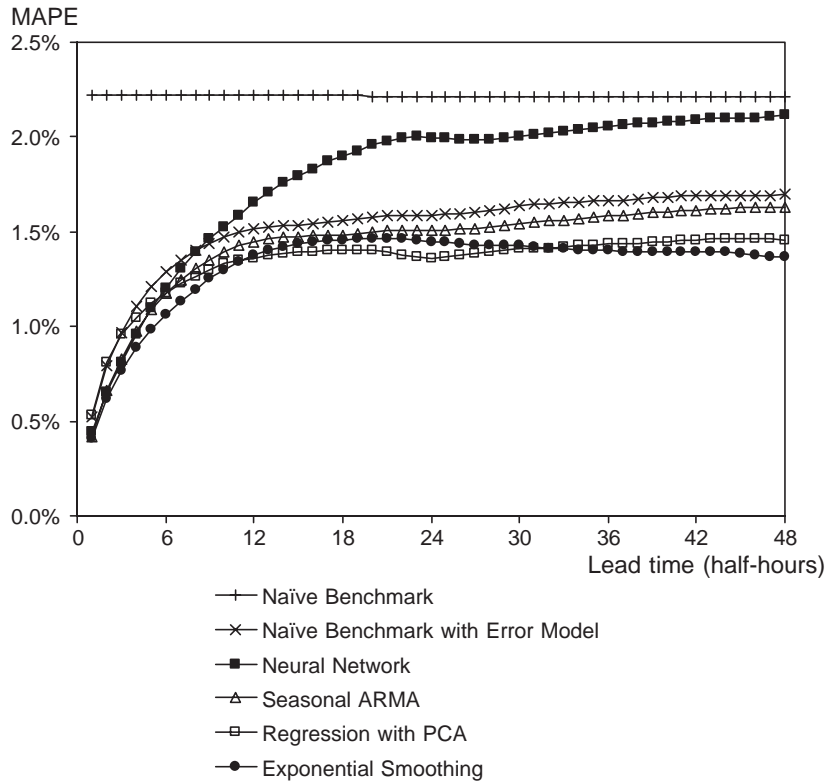


Fig. 6. MAPE results plotted against lead time for the England and Wales 10-week post-sample period.

this approach over the ARMA and PCA methods. For both series, we did see a slight deterioration in the performance of each method as the method was passed through the 10-week evaluation period, which motivates an updating of the methods. However, in our study, we opted not to re-specify or re-estimate methods because we felt that in practice it would be impractical to update, on a frequent basis, the ARIMA and neural network models.

By contrast with the Rio data, the England and Wales series, shown in Fig. 4, does not seem to change its characteristics in the post-sample evaluation period. Our results for this series, therefore, provide insight into whether the exponential smoothing method is able to offer anything more than just simplicity and robustness. The results in Fig. 6 show that the exponential smoothing method did in fact perform well for the England and Wales series. It outperformed all the other methods at all lead times, except the PCA method, which has lower MAPE for lead times 12 to 32. As with the Rio data, the naïve benchmark was

outperformed by all the other methods, but was vastly improved by the inclusion of the heuristic error model. For the England and Wales data, the neural network was competitive for the early lead times, but beyond about 10 steps ahead (5 hours ahead), it performed poorly.

We were surprised that the performance of the neural network relative to the ARMA model was so much worse than that reported by Darbellay and Slama (2000) for their very similar neural network. Our explanation for this is that their data possessed different characteristics to ours. Firstly, an obvious point is that they used data from a different country and this data may have had stronger non-linearities than our series, which would have helped the neural network. The data in the Czech Republic study certainly looks more ‘noisy’ than the two series that we have analysed; some of this apparent ‘noise’ may be non-linearity. Secondly, they only report results for models fitted and evaluated on weekday observations. The omission of the weekend data leaves a much simpler modelling

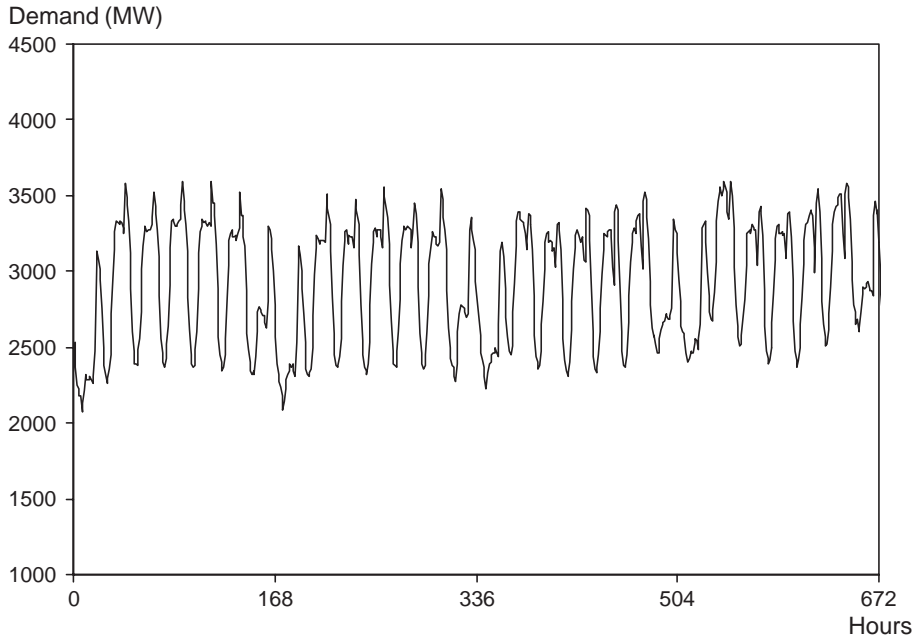


Fig. 7. First four weeks of the 10-week post-sample period of hourly electricity demand in Rio de Janeiro. This period is Monday, 22 September 1996 to Sunday, 19 October 1996.

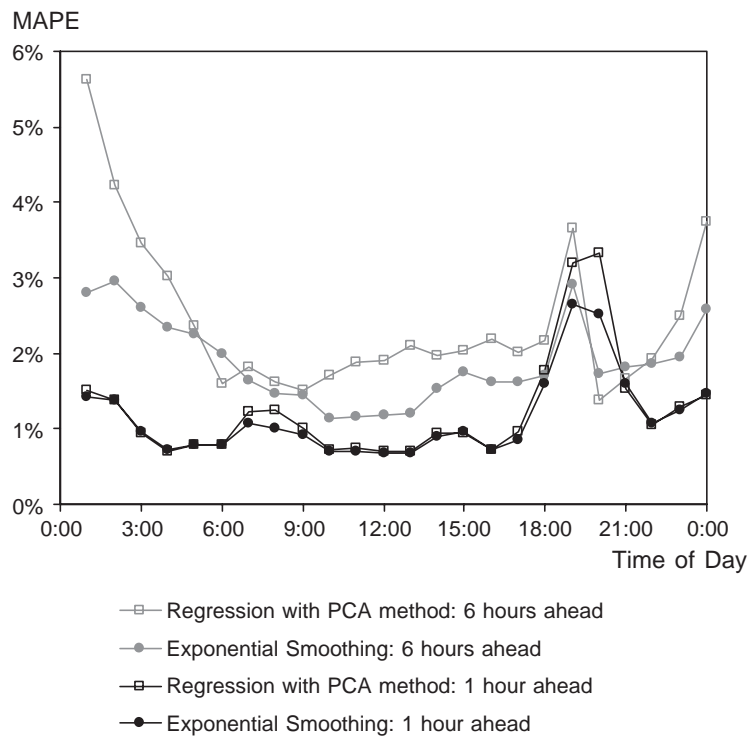


Fig. 8. MAPE results for lead times of one and six hours, plotted against time of day, for the Rio 10-week post-sample period.

task. Thirdly, in contrast to our use of 30 weeks of data, Darbellay and Slama used one full year of data to estimate their models and a second complete year for evaluation. The additional observations and the exposure to the full within-year seasonal cycle may have greatly benefited the Darbellay and Slama network. We should also comment that if a particular neural network model does not perform well, as in our paper, one cannot be certain whether neural networks are, in general, poor predictors because a different specification of neural network might have performed better. Another reason for the poor performance of the neural network could be that we used forecasted values as inputs for the multi-step predictions. Non-linear models can spread errors very dramatically. A neural network model developed directly for given horizons (e.g., 24 neural network models or a model with 24 outputs) could perhaps produce better results.

In the introduction, we explained that univariate methods are mainly used for predicting load up to lead times of about six hours. In Figs. 8 and 9, we focus more closely on the MAPE results for lead

times of one and six hours for the two best performing methods in Figs. 5 and 6. Fig. 8 shows the MAPE results for these two lead times plotted against time of day for the Rio 10-week post-sample period, and Fig. 9 presents the results for the England and Wales series. For six hour-ahead prediction for the Rio series, Fig. 8 shows the exponential smoothing method outperforming the PCA method for most hours of the day, and particularly so for the early hours of the morning. The results are much closer for the one hour-ahead prediction. For the England and Wales series, Fig. 9 shows that the performance of the two methods is similar for one hour-ahead prediction, except for those periods when the errors are relatively large. For these periods, the PCA method seems to perform noticeably worse. The figure shows that, for the six hour lead time, the performance of the methods is very similar for approximately half of the periods, but for the other half, exponential smoothing is better.

To provide a little more insight into the performance of the PCA and exponential smoothing meth-

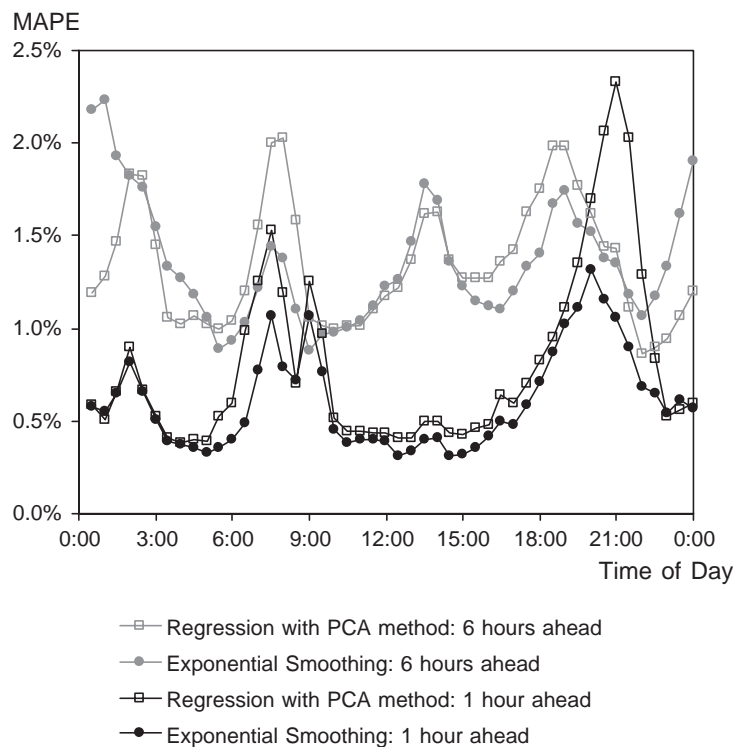


Fig. 9. MAPE results for lead times of one and six hours, plotted against time of day, for the England and Wales 10-week post-sample period.

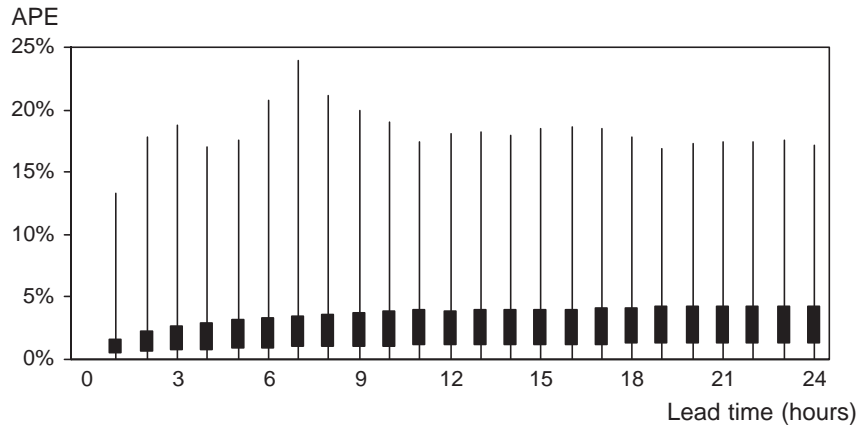


Fig. 10. Box-plots for APE results from the regression with PCA method, plotted against lead time, for the Rio 10-week post-sample period.

ods for the Rio data, in Figs. 10 and 11, we present box-plots for the APE results from the two methods. The inter-quartile ranges for the two methods do not appear to differ greatly. The notable difference between the two figures is in the size of the maximum APE at each lead time. The maximum values for the exponential smoothing method are smaller and more smoothly evolving across the lead times than those for the PCA method.

## 6. Summary and conclusions

In this paper we compared the performance of six methods for forecasting electricity demand. Of the four sophisticated methods, the seasonal ARIMA and

neural network approaches have been widely applied. By contrast, the double seasonal exponential smoothing and PCA methods are newer formulations, which have been specifically designed for high frequency load series. We used series from two different sources, one hourly and the other half-hourly. The two series both show the weekly and daily seasonal patterns that are common to load data, but the series exhibit different degrees of stability in the seasonality and thus present somewhat different forecasting problems.

A highlight of this study was the success of the exponential smoothing method. In addition to its forecasting performance, it is important to note that, of the four sophisticated methods that we considered, this method is comfortably the simplest and quickest to

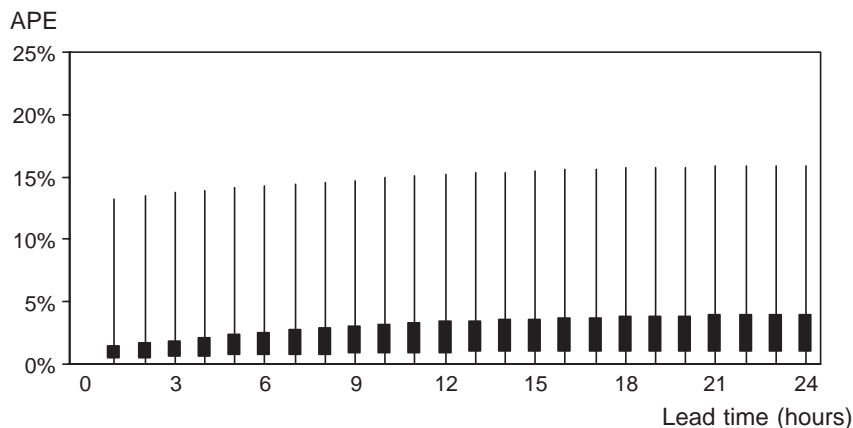


Fig. 11. Box-plots for APE results from the exponential smoothing method, plotted against lead time, for the Rio 10-week post-sample period.

implement. On the disappointing side, the neural network did not perform well. We do not believe that this was due to overfitting or an overly complex architecture, which are two common pitfalls in neural network modelling. Instead, the poor performance may have resulted because we did not separate the data into weekday and weekend observations and then build separate models for each dataset. Other possibilities are that we used less data than previous authors or, quite simply, the non-linearity in our data may be negligible.

The PCA method was certainly competitive for the England and Wales data, which prompted us to consider a combination. For this series, a slight improvement was achieved by combining the exponential smoothing method and the PCA method using a simple average. This motivates further consideration of combining (see Smith, 1989; Taylor & Majithia, 2000), and of the development of a method that is able to synthesise the useful features of the exponential smoothing and PCA methods.

## Acknowledgements

Patrick E. McSharry acknowledges the support of a Research Fellowship from the Royal Academy of Engineering and Physical Sciences Council (EPSRC). We are grateful to three anonymous referees for their very useful comments. We are grateful to three anonymous referees for their very useful comments. We are also grateful to National Grid Transco and ONS-Light for supplying the demand data.

## References

- Abraham, A., & Nath, B. (2001). A neuro-fuzzy approach for modelling electricity demand in Victoria. *Applied Soft Computing Journal*, 1, 127–138.
- Asar, A., & McDonald, R. (1994). A specification of neural network applications in the load forecasting problem. *IEEE Transactions on Control Systems Technology*, 2, 135–141.
- Bishop, C. M. (1997). *Neural networks for pattern recognition*. Oxford: Oxford University Press.
- Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). *Time series analysis: Forecasting and control* (3rd ed.). Englewood Cliffs, New Jersey: Prentice Hall.
- Campo, R., & Ruiz, P. (1987). Adaptive weather-sensitive short-term load forecasting. *IEEE Transactions on Power Systems*, 3, 592–600.
- Carnero, A., Koopman, S. J., & Ooms, M. (2003). *Periodic heteroskedastic RegARFIMA models for daily electricity spot prices*. Discussion paper 03-071/4. Amsterdam: Tinbergen Institute.
- Charytoniuk, W., Chen, M. S., & Van Olinda, P. (1998). Nonparametric regression based short-term load forecasting. *IEEE Transactions on Power Systems*, 13, 725–730.
- Chatterjee, S., Hadi, A. S., & Price, B. (1999). *Regression analysis by example* (3rd ed.). New York: Wiley.
- Chow, T. W. S., & Leung, C. T. (1996). Neural network based short-term load forecasting system using weather compensation. *IEEE Transactions on Power Systems*, 11, 1736–1742.
- Christiaanse, W. R. (1971). Short-term load forecasting using general exponential smoothing. *IEEE Transactions on Power Apparatus and Systems, PAS-90*, 900–902.
- Cottrell, M., Girard, B., Girard, Y., Mangeas, M., & Muller, C. (1995). Neural modelling for time series: A statistical stepwise method for weight elimination. *IEEE Transactions on Neural Networks*, 6, 1355–1364.
- Daneshdoot, M., Lotfalian, M., Bumroonggit, G., & Ngoy, J. P. (1998). Neural network with fuzzy set-based classification for short term load forecasting. *IEEE Transactions on Power Systems*, 13, 1386–1391.
- 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.
- Hippert, H. S., Pedreira, C. E., & Souza, R. C. (2001). Neural networks for short term load forecasting: A review and evaluation. *IEEE Transactions on Power Systems*, 16, 44–55.
- Infield, D. G., & Hill, D. C. (1998). Optimal smoothing for trend removal in short-term electricity demand forecasting. *IEEE Transactions on Power Systems*, 13, 1115–1120.
- Laing, W. D., & Smith, D. G. C. (1987). A comparison of time series forecasting methods for predicting the CEGB demand. *Proceedings of the ninth power systems computation conference*.
- Lee, K. Y. (1992). Short-term load forecasting using an artificial neural network. *IEEE Transactions on Power Systems*, 7, 124–131.
- Liu, K., Subbarayan, S., Shoults, R. R., Manry, M. T., Kwan, C., Lewis, F. L., et al. (1996). Comparison of very short term load forecasting techniques. *IEEE Transactions on Power Systems*, 11, 877–882.
- Piras, A., & Buchenel, B. (1999). A tutorial on short term load forecasting. *Engineering Intelligent Systems*, 1, 41–47.
- 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.
- Rui, Y., & El-Keib, A. A. (2004). *A review of ANN-based short-term load forecasting models, mimeo*, Department of Electrical Engineering, University of Alabama.
- Smith, D. G. C. (1989). Combination of forecasts in electricity demand prediction. *Journal of Forecasting*, 8, 349–356.
- Souza, L. R., & Soares, L. J. (2003). Forecasting electricity load demand: Analysis of the 2001 rationing period in Brazil. *Ensaio Economicos da EPGE, vol. 491*. (ISSN: 0104-8910).
- Strang, G. (1993). *Introduction to linear algebra* (2nd ed.). Wellesley College.

- 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., & Buizza, R. (2003). Using weather ensemble predictions in electricity demand forecasting. *International Journal of Forecasting*, 19, 57–70.
- Taylor, J. W., & Majithia, S. (2000). Using combined forecasts with changing weights for electricity demand profiling. *Journal of the Operational Research Society*, 51, 72–82.
- Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, 14, 35–62.

**James W. Taylor** is a Reader in Decision Science at the Saïd Business School. His research interests include exponential smoothing, prediction intervals, quantile regression, combining forecasts, volatility forecasting, electricity demand forecasting and weather ensemble predictions.

**Lilian M. de Menezes** is a Senior Lecturer in Quantitative Methods at Cass Business School, City University, London. She completed her Ph.D. at London Business School (1993) and has held posts at the London School of Economics and Goldsmiths College, University of London. Her research interests include time series forecasting, nonlinear models and load forecasting. She has published in a number of journals including the *European Journal of Operational Research*, *International Journal of Forecasting* and the *Journal of the Royal Statistical Society*.

**Patrick E. McSharry** is a Royal Academy of Engineering/EPSRC Research Fellow at the Department of Engineering Science and the Oxford Centre for Industrial and Applied Mathematics in the University of Oxford. His research interests include time series analysis, quantitative data analysis, mathematical modelling, forecasting and biomedical signal processing.