Short-Term Load Forecasting Methods:
An Evaluation Based on European Data

J. W. Taylor and P. E. McSharry, *Senior Member, IEEE*

Abstract-- **This paper uses intraday electricity demand data from 10 European countries as the basis of an empirical comparison of univariate methods for prediction up to a day-ahead. A notable feature of the time series is the presence of both an intraweek and an intraday seasonal cycle. The forecasting methods considered in the study include: ARIMA modeling; periodic AR modeling; an extension for double seasonality of Holt-Winters exponential smoothing; a recently proposed alternative exponential smoothing formulation; and a method based on the principal component analysis (PCA) of the daily demand profiles. Our results show a similar ranking of methods across the 10 load series. The results were disappointing for the new alternative exponential smoothing method and for the periodic AR model. The ARIMA and PCA methods performed well, but the method that consistently performed the best was the double seasonal Holt-Winters exponential smoothing method.**

Index Terms— **Electricity demand forecasting, exponential smoothing, principal component analysis, ARIMA, periodic AR.**

I. INTRODUCTION

ELECTRICITY demand forecasting is of great importance for the management of power systems. Long-term forecasts of the peak electricity demand are needed for capacity planning and maintenance scheduling [1]. Medium-term demand forecasts are required for power system operation and planning [2]. Short-term load forecasts are required for the control and scheduling of power systems. Short-term forecasts are also required by transmission companies when a self-dispatching market is in operation. There are several such markets in Europe and the US. For example, in Great Britain, 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 [3, 4]. More generally, error in predicting electricity load has significant cost implications for companies operating in competitive power markets [5].

It is well recognized that meteorological variables have a very significant influence on electricity demand (see, for example, [6]). However, in online short-term forecasting systems, multivariate modeling is usually considered impractical [7]. In such systems, the lead times considered are less than a day-ahead, and univariate methods can be sufficient because the meteorological variables tend to change in a smooth fashion, which will be captured in the demand series itself. Indeed, univariate models are often used for prediction up to about four to six hours ahead, and, due to the expense or unavailability of weather forecasts, univariate methods are sometimes used for longer lead times.

In a recent study [8], methods for short-term load forecasting are reviewed, and two intraday load time series are used to compare a variety of univariate methods. One of the aims of this paper is to validate the results of that study. It concluded that a double seasonal version of Holt-Winters exponential smoothing was the most accurate method, with a new approach based on principal components analysis (PCA) also performing well. Using time series of intraday electricity demand from 10 European countries, we empirically compare the better methods identified in [8] and also the following two new candidate methods: an intraday cycle exponential smoothing method (see [9]), and a new periodic AR approach, which we believe has not previously been considered for electricity demand forecasting. All of the methods are specifically formulated to deal with the double seasonality that typically arises in demand data. This seasonality involves intraday and intraweek seasonal cycles.

Artificial neural networks (ANNs) have featured prominently in the load forecasting literature (see [10]). Their nonlinear and nonparametric features have been useful for multivariate modeling in terms of weather variables. However, their usefulness for univariate short-term load prediction is less obvious. Indeed, the results in [8] for the ANN were not competitive. Although we accept that a differently specified neural network may be useful for univariate load modeling, in this paper, for simplicity, we do not include an ANN. We would hope that the better performing methods in our study can serve as benchmarks in future studies with ANNs.

The paper is structured as follows: Section II describes the electricity demand data; Section III describes the methods included in the study; Section IV presents the post-sample results of a comparison of the short-term forecasts with lead times up to one day ahead; and Section V summarises the results and concludes the paper.

J. W. Taylor is with the Saïd Business School, University of Oxford, Park End Street, Oxford OX1 1HP, UK (e-mail: james.taylor@sbs.ox.ac.uk).

P. E. McSharry is with the Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, UK (corresponding author; tel: 44-1865-273095; fax: 44-1865-273905; e-mail: patrick@mcsharry.net).

II. THE 10 EUROPEAN ELECTRICITY DEMAND SERIES

Our dataset consisted of intraday electricity demand from 10 European countries for the 30 week period from Sunday 3 April 2005 to Saturday 29 October 2005. We obtained half-hourly data for six of the countries, and hourly data for the remaining four. The first 20 weeks of each series were used to estimate parameters and the remaining 10 weeks to evaluate post-sample accuracy of forecasts up to 24 hours ahead. For the half-hourly series, this implies 6,720 observations for estimation and 3,360 for evaluation. For the hourly series, 3,360 observations were used for estimation and 1,680 for evaluation. We represent the lengths of the intraday and intraweek cycles as s_1 and s_2, respectively. For the half-hourly demand series, $s_1 = 48$ and $s_2 = 336$, and for the hourly series $s_1 = 24$ and $s_2 = 168$.

Electricity demand on "special days", such as bank holidays, is very different to normal days and can give rise to problems with online forecasting systems. In practice, interactive facilities tend to be used for special days, which allow operator experience to supplement or override the standard forecasting system. In our study, we smoothed out these unusual observations by taking averages of observations from the corresponding period in the two adjacent weeks.

TABLE I
MEAN LOAD AND POPULATION FOR THE 10 COUNTRIES.

	Mean load (GW)	Population (million)	Mean load (W) per capita
Half-Hourly			
Belgium	9.4	10.4	906
Finland	8.3	5.2	1,588
France	47.8	60.9	786
Great Britain	36.0	58.9	611
Ireland	2.9	4.1	711
Portugal	5.2	10.6	492
Hourly			
Italy	32.7	58.1	563
Norway	12.0	4.6	2,599
Spain	25.5	40.4	630
Sweden	14.5	9.0	1,608

Table I lists the 10 countries along with the mean electricity demand for the 30 week period and, in order to give a feel for the size of the countries, we also list their respective populations. For illustrative purposes, in Fig. 1, we plot the time series for France, Finland and Ireland. The decrease in French demand during August is due to the summer vacation. The Irish series maintains a relatively constant level throughout the 30 week period. By contrast, the Finnish data shows a temporary level shift in the first half of the series. This was due to inactivity in the paper industry, which was caused by a large conflict in contract negotiations between the workers and employers.
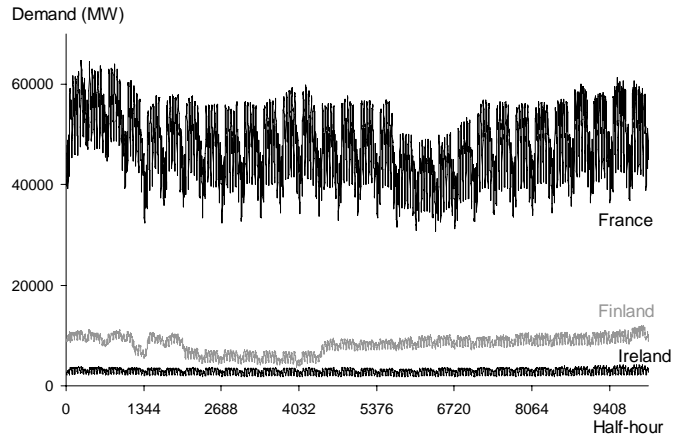

Fig. 1. Half-hourly electricity demand in France, Finland and Ireland from Sunday, 3 April 2005 to Saturday, 29 October 2005.

Fig. 2 shows the French series for the fortnight in the middle of the 30 week period. This graph shows a within-day seasonal cycle of duration $s_1 = 48$ periods and a within-week seasonal cycle of duration $s_2 = 336$ periods. The weekdays show similar patterns of demand, whereas Saturday and Sunday have different levels and profiles. The intraweek and intraday features in Fig. 2 are typical of those in all 10 series.
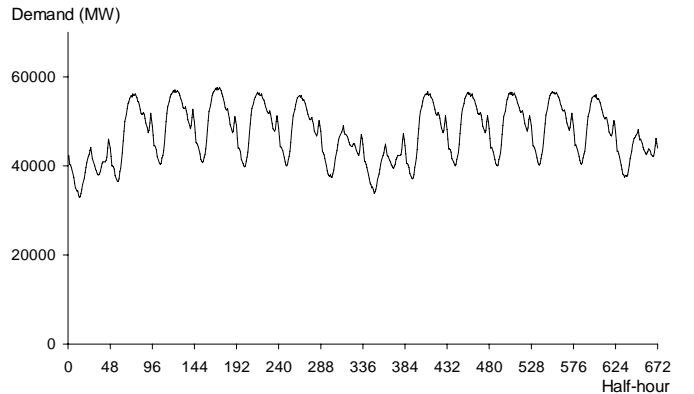

Fig. 2. Half-hourly electricity demand in France from Sunday, 10 June 2005 to Saturday, 23 June 2005.

III. FORECASTING METHODS

A. Simplistic Benchmark Methods

We implemented two naïve benchmark methods. The first was a 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, 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. The forecast function is written as $\hat{y}_t(k) = y_{t+k-s_2}$, where y_t is the demand in period t, and k is the forecast lead time ($k \le s_2$). The second simplistic benchmark that we used was the simple average of the corresponding observations in each of the previous four weeks. For this method, the forecast function is

$$\hat{y}_t(k) = \left(y_{t+k-s_2} + y_{t+k-2s_2} + y_{t+k-3s_2} + y_{t+k-4s_2} \right)/4.$$

B. Seasonal ARMA Modeling

Double seasonal ARMA models are often used as benchmarks in load forecasting studies (e.g. [1], [8], [11], [12]). For each of the 10 load series, we followed the Box-Jenkins methodology to identify the most suitable model based on the estimation sample of 20 weeks. We considered differencing, but the resultant models had weaker diagnostics than models fitted with no differencing. The multiplicative double seasonal ARMA model (see [13], p. 333) can be written as

$$\phi_p(L)\,\Phi_{P_1}\!\left(L^{s_1}\right)\Omega_{P_2}\!\left(L^{s_2}\right)\!\left(y_t - c\right)=\theta_q(L)\,\Theta_{Q_1}\!\left(L^{s_1}\right)\Psi_{Q_2}\!\left(L^{s_2}\right)\varepsilon_t$$

where c is a constant term; L is the lag operator; ε_t is a white noise error term; ϕ_p, Φ_{P_1}, Ω_{P_2}, θ_q, Θ_{Q_1} and Ψ_{Q_2} are polynomial functions of orders p, P_1, P_2, q, Q_1 and Q_2, respectively. The model can be expressed as $\mathrm{ARMA}\,(p,q)\times(P_1,Q_1)_{s_1}\times(P_2,Q_2)_{s_2}$.

We estimated the models using maximum likelihood with the likelihood function based on the standard Gaussian assumption. We considered lag polynomials up to order three. This choice was made arbitrarily, but it is consistent with other load forecasting studies, and it was supported by experimentation with several of the series. We based model selection on the Schwarz Bayesian Criterion, with the requirement that all parameters were significant (at the 5% level).

C. Periodic AR Models

In intraday electricity demand time series, the intraday seasonal cycle is usually reasonably similar for the five weekdays, but quite different for the weekends. This implies that the autocorrelation at a lag of one day is time-varying across the days of the week. Such time-variation cannot be captured in the seasonal ARMA model described in the previous section. A class of models that can capture this feature is periodic ARMA models. In these models, the parameters are allowed to change with the seasons [14]. Such models have been shown to be useful for modelling economic data (e.g. [15]). In the electricity context, periodic models have been considered with some success in studies investigating methods for forecasting intraday net imbalance volume [4] and daily electricity spot prices [16].

To assess the potential for periodic ARMA models, we examined whether the autocorrelation at a specified lag exhibited variation across the periods of the day or the week. For example, for the half-hourly French series, Fig. 3 shows how the autocorrelation at lag $s_1=48$ varies across the $s_2=336$ half-hours of the week. The first period of the x-axis corresponds to the first period on a Sunday. The autocorrelation values were calculated from just the 20-week in-sample period. Although the sample size is not sufficiently large to conclude with confidence, the variation in the autocorrelation in this plot, and in similar graphs for the other series, suggested to us that there was some appeal in estimating periodic ARMA models for our data.

Studies have shown that periodic MA terms are unnecessary (see [14], p. 28), and so for simplicity, in our work, we considered only periodic AR models. More specifically, we estimated models with periodicity in the coefficient of AR terms of lag s_1. The formulation for this method is presented in the following expressions:

$$\left(1-\phi_1 L\right)\!\left(1-\phi_{s_1}(t)L^{s_1}\right)\!\left(1-\phi_{s_2}L^{s_2}\right)\!\left(y_t - c\right)=\varepsilon_t$$

where

$$\phi_{s_1}(t)=\omega+\sum_{i=1}^{4}\left(\begin{array}{c}\lambda_i \sin\!\left(2i\pi \frac{d(t)}{s_1}\right)+v_i \cos\!\left(2i\pi \frac{d(t)}{s_1}\right)\\ +\kappa_i \sin\!\left(2i\pi \frac{w(t)}{s_2}\right)+\upsilon_i \cos\!\left(2i\pi \frac{w(t)}{s_2}\right)\end{array}\right)$$

$d(t)$ and $w(t)$ are repeating step functions that number the periods within each day and week, respectively. For example, for the half-hourly series, $d(t)$ counts from 1 to 48 within each day, and $w(t)$ counts from 1 to 336 within each week. ω, λ_i, v_i, κ_i, υ_i, ϕ_1 and ϕ_{s_2} are constant parameters. The periodic parameter, $\phi_{s_1}(t)$, uses a similar flexible fast Fourier form to that employed in an analysis of the volatility in intraday financial returns [17]. For simplicity, we arbitrarily chose to sum from $i=1$ to 4 for all 10 series. The parameters were estimated using maximum likelihood.
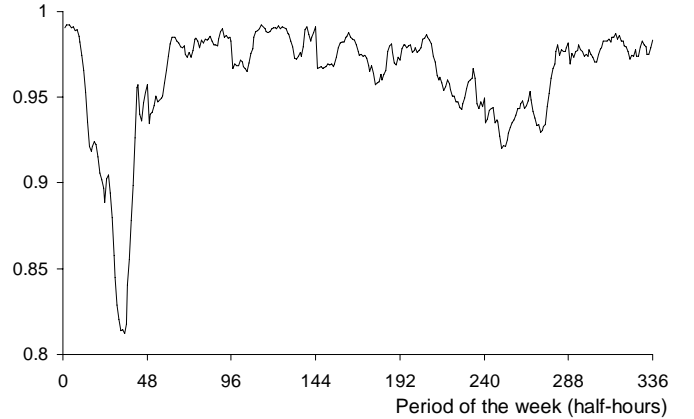

Fig. 3. For the in-sample period of the France series, lag 48 autocorrelation estimated separately for each period of the week.

D. Double Seasonal Holt-Winters Exponential Smoothing

Exponential smoothing has found widespread use in automated applications, such as inventory control. The seasonal Holt-Winters method has been adapted in order to accommodate the two seasonal cycles in electricity demand series [18]. This involves the introduction of an additional seasonal index and an extra smoothing equation for the new seasonal index. The multiplicative formulation for the double seasonal Holt-Winters method is given in the following expressions:

$$l_t = \alpha\,(y_t/(d_{t-s_1}w_{t-s_2})) + (1-\alpha)\,l_{t-1} \tag{1}$$

$$d_t = \delta\,(y_t/(l_t w_{t-s_2})) + (1-\delta)\,d_{t-s_1} \tag{2}$$

$$w_t = \omega\,(y_t/(l_t d_{t-s_1})) + (1-\omega)\,w_{t-s_2} \tag{3}$$

$$\hat{y}_t(k) = l_t d_{t-s_1+k} w_{t-s_2+k} + \phi^k\!\left(y_t - \left(l_{t-1}\,d_{t-s_1}\,w_{t-s_2}\right)\right) \tag{4}$$

where l_t is the smoothed level; d_t and w_t are the seasonal indices for the intraday and intraweek seasonal cycles, respectively; α, δ and ω are the smoothing parameters; and

$\hat{y}_t(k)$ is the k step-ahead forecast made from forecast origin t (where $k \leq s_1$). The term involving the parameter ϕ, in the forecast function (4), is a simple adjustment for first-order autocorrelation. A trend term was included in the original formulation, but we found it not to be of use for our 10 series.

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

The initial smoothed values for the level and seasonal components are estimated by averaging the early observations. The parameters are estimated in a single procedure by minimizing the sum of squared one step-ahead in-sample errors. We constrained the parameters to lie between zero and one. The resultant parameters for the 10 load series are presented in Table II. For many of the series, the value of ϕ is very high and the value of α is very low indicating that the adjustment for first-order autocorrelation has, to a large degree, made redundant the smoothing equation for the level. It is also interesting to note that, for a given series, the values are similar for the two smoothing parameters, δ and ω, for the seasonal indices. We also implemented a version of the method with the optimized values of δ and ω constrained to be identical, and with $\alpha{=}0$ so that the level was set as a constant value equal to the mean of the in-sample observations. This formulation delivered predictions only marginally poorer than the full method given in expressions (1) to (4). This is somewhat surprising, given that this reformulation of the method involves just two parameters.

TABLE II
FOR EACH OF THE 10 LOAD SERIES, PARAMETERS OF THE
HOLT-WINTERS METHOD FOR DOUBLE SEASONALITY.

	α	δ	ω	ϕ
Half-Hourly				
Belgium	0.043	0.146	0.175	0.820
Finland	0.000	0.083	0.153	0.996
France	0.004	0.249	0.231	0.987
Great Britain	0.002	0.316	0.168	0.970
Ireland	0.009	0.227	0.153	0.910
Portugal	0.094	0.201	0.210	0.771
Hourly				
Italy	0.039	0.271	0.281	0.944
Norway	0.039	0.126	0.151	0.863
Spain	0.036	0.193	0.217	0.871
Sweden	0.022	0.223	0.134	0.928

From a theoretical perspective, exponential smoothing methods can be considered to have a sound basis as they have been shown to be equivalent to a class of state space models [19]. The double seasonal Holt-Winters formulation of expressions (1) to (4) can be expressed as a single source of

error state-space model. This model can be used as the basis for producing prediction intervals. The motivation that led us to consider periodic AR models prompted us to also consider periodicity in the parameters of the double seasonal Holt-Winters method. Disappointingly, this did not lead to improved accuracy, and so for simplicity we do not report these further results in Section IV. The issue of periodicity is addressed in the next section in an alternative exponential smoothing formulation that has recently been proposed.

E. Intraday Cycle Exponential Smoothing Model for Double Seasonality

A feature of the double seasonal Holt-Winters method is that it assumes the same intraday cycle for all days of the week, and that updates to the smoothed intraday cycle are made at the same rate for each day of the week. An alternative form of exponential smoothing for double seasonality is presented by Gould et al. in [9]. It allows the intraday cycle for the different days to be represented by different seasonal components. In addition, it allows the different seasonal components to be updated at different rates by using different smoothing parameters.

Our implementation of the method of Gould et al. involves the seasonality being viewed as consisting of the same intraday cycle for the five weekdays and a distinct intraday cycle for Saturday and another for Sunday. The days of the week are thus divided into three types: weekdays, Saturdays and Sundays. By contrast with double seasonal Holt-Winters, there is no representation in the formulation for the intraweek seasonal cycle. Due to its focus on intraday cycles, we term this method 'intraday cycle exponential smoothing'. For any period t, the latest estimated values of the three distinct intraday cycles are given as c_{1t}, c_{2t} and c_{3t}, respectively. The formulation requires three corresponding dummy variables, x_{1t}, x_{2t} and x_{3t}, defined as follows:

$$x_{jt} = \begin{cases} 1 & \text{if time period } t \text{ occurs in a day of type } j \\ 0 & \text{otherwise} \end{cases}$$

Gould et al. present their approach in the form of a state space model, and we follow this convention in our presentation of the model in expressions (5) to (7):

$$y_t = l_{t-1} + \sum_{i=1}^{3} x_{it} c_{i,t-s_1} + \varepsilon_t \tag{5}$$

$$l_t = l_{t-1} + \alpha \varepsilon_t$$
(6)

$$c_{it} = c_{i,t-s_1} + \left(\sum_{j=1}^{3} \gamma_{ij} x_{jt} \right) \varepsilon_t \qquad (i = 1, 2, 3) \tag{7}$$

where l_t is the smoothed level; ε_t is an error term; and α and the γ_{ij} are the smoothing parameters. (Expressions (6) and (7) can easily be rewritten as recursive expressions, which is the more widely used form for exponential smoothing methods.) As with the double seasonal Holt-Winters method, we estimated the initial smoothed values for the level and seasonal components by averaging the early observations. The parameters were estimated in a single procedure by minimizing the sum of squared one step-ahead in-sample

errors. All parameters were constrained to lie between zero and one.

The γ_{ij} can be viewed as a 3×3 matrix of parameters that enables the three types of intraday cycle to be updated at different rates. It also enables intraday cycle of type i to be updated even when the current period is not in a day of type i.

Several restrictions have been proposed for the matrix of γ_{ij} parameters (see [9]). We included in our empirical study two forms of the method; one involved estimation of the matrix of γ_{ij} parameters with the only restriction being that the parameters lie between zero and one, and the other involved the additional restrictions of common diagonal elements and common off-diagonal elements. Gould et al. note that these additional restrictions lead to the method being identical to the double seasonal Holt-Winters method of expressions (1) to (4), provided seven distinct intraday cycles are used, instead of three, as in our study. In our discussion of the post-sample forecasting results in Section IV, we refer to this second form of the model as the restricted form.

We found that the results were substantially improved with the inclusion of the adjustment for first-order autocorrelation that was used in expression (4) of the double seasonal Holt-Winters method. In Section IV, we report only the results for this improved form of the intraday cycle method.

F. A PCA-Based Method

PCA provides a means of reducing the dimension of a multivariate data set to a smaller set of orthogonal variables. These new variables are linear combinations of the original variables. They are uncorrelated and explain most of the variation in the data, and, for this reason, they are commonly refereed to as principal components. A method based on PCA has recently been proposed for short-term load forecasting [8]. The method aims to capture the intraday variation in electricity demand, and it can be viewed as a development of the approach where a separate model is built for each of the s_1 periods of the day (see, for example, [20]). The method exploits the similarity between intraday observations in order to reduce the number of models to be considered. Note that this approach could easily be extended to the multivariate case, if weather related variables were available.

In this section, we present only an overview of the method, as details are provided in [8]. The method proceeds by arranging the observations as an ($n_d \times s_1$) matrix, Y, where n_d is the number of days in the estimation sample. Each column contains observations for a particular intraday period. PCA is applied to the columns of Y to deliver components that are columns of a new ($n_d \times s_1$) matrix. For each component, a regression model is built using day of the week dummies and quadratic trend terms. The models are then used to deliver a day-ahead forecast for each component. Load predictions are created by projecting forecasts of the components back onto the Y space. The method is refined, and speeded up, by focusing attention on just the *principal* components. Cross-validation is used to optimise two parameters: the number of

principal components and the length of the training period used in the PCA. In our study, the cross-validation employed the first half of the 20 week in-sample period for estimation and the second half for evaluation. We set, as optimal parameters, those delivering the minimum sum of squared one step-ahead forecast errors for the training data.

The errors resulting from this method exhibit serial correlation. As for the Holt-Winters exponential smoothing method, the method benefits by the addition of an AR model of the error process. It is worth noting that in [20] an error model was also employed in order to correct for serial correlation resulting from the use of separate models for each hour of the day. Let $E_t(k)$ be the prediction error associated with a k step-ahead forecast made from origin t. The error-correction model is of the following form:

$$E_t(k) = \alpha_0(k) + \alpha_1(k)E_{t-s_1}(k) + \alpha_2(k)E_{t-1}(1)$$

where the $\alpha_l(k)$ are parameters estimated separately for each lead time, k, using LS regression applied to the estimation sample. Finally, with this model specification that now includes the error correction term above, cross-validation is used to optimise the number of principal components and the length of the training period used in the PCA. For each of our 10 load series, the optimal values are presented in Table III.

TABLE III
FOR THE PCA METHOD, THE OPTIMAL NUMBER OF TRAINING WEEKS AND NUMBER OF PRINCIPAL COMPONENTS FOR THE 10 LOAD SERIES.

	Number of training weeks	Number of principal components
Half-Hourly		
Belgium	7	12
Finland	9	9
France	4	9
Great Britain	8	12
Ireland	8	10
Portugal	8	9
Hourly		
Italy	3	5
Norway	8	6
Spain	10	7
Sweden	8	6

IV. POST-SAMPLE FORECASTING RESULTS

We evaluated post-sample forecasting performance from the various methods using the mean absolute percentage error (MAPE) and the mean absolute error (MAE). Having calculated the MAPE for each method at each forecast horizon, we then summarized each method's performance by averaging the MAPE across the 10 load series. For the half-hourly series, MAPE values were available for 48 half-hour lead times, while, for the hourly data, forecasts were obviously only available for 24 hourly lead times. In order to average across all the series, we focused only on the 24 hourly

lead times. The resulting Mean MAPE values are presented in Fig. 4. Averaging MAE values across the 10 series did not seem sensible because the values tended to be substantially higher for the series corresponding to higher levels of electricity demand. In view of this, for each method, we summarized the MAE performance across the 10 series by averaging, for each lead time, the ratio of the method's MAE to the MAE of the seasonal random walk benchmark method. The relative performances of the methods according to this measure were very similar to those for the Mean MAPE. In view of this, we do not plot this additional measure here. We also evaluated the methods using the root mean squared percentage error and root mean squared error, but we also do not report these results because the relative performances of the methods for these measures were very similar to those for the MAPE.
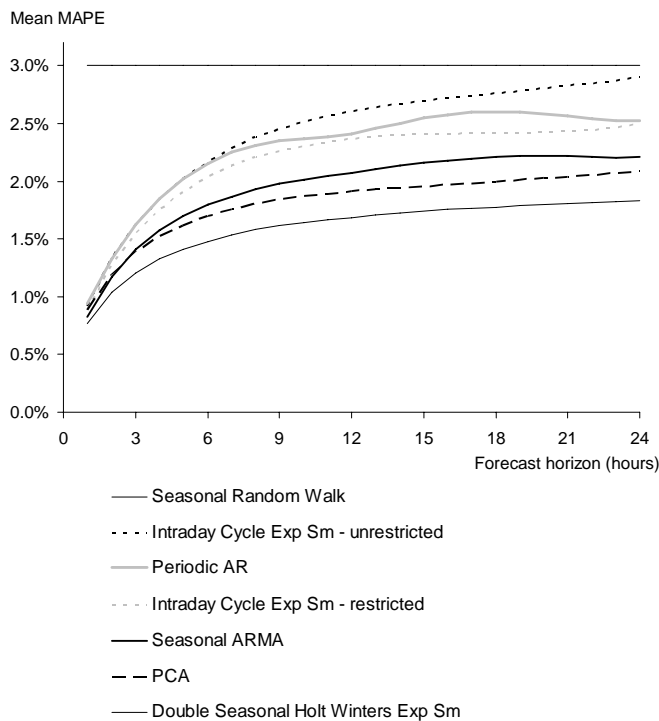

Fig. 4. For the 10 load series, mean MAPE plotted against lead time.

The first point to note is that Fig. 4 does not show the results for the second simplistic benchmark method that involved the simple average of the corresponding observations in each of the previous four weeks. The results for this method were poorer than those for the seasonal random walk, and so for simplicity we opted to omit the results from the figure. Turning to the more sophisticated methods, the figure shows the double seasonal Holt-Winters method performing the best, followed by the PCA method and then seasonal ARMA. Of the two versions of the intraday cycle exponential smoothing method, the restricted form appears to be better, which is consistent with the results in [9]. However, the results for both forms of this method are disappointing. This is also the case for the periodic AR method. Our view is that there is strong potential for the use of some form of period model, but that a longer time series may be needed to estimate the periodicity in

the parameters. With regard to Fig. 3, the use of only 20 weeks of data implies that only 19 observations were available to estimate the intraweek cyclical pattern in the autocorrelation for a given lag. In a similar way, 20 weeks of data is perhaps too little to provide adequate estimates of periodic model parameters.

We should also comment that the ranking of the methods was really quite stable across the 10 series, and that the double seasonal Holt-Winters method was consistently the best regardless of the error measure used for evaluation. For the Finland load series, which, as shown in Fig. 1, contains level shifts in the estimation period, the post-sample MAPE values were relatively high for all methods. However, it is interesting to note that, for this series, the ranking of the methods was similar to that shown in Fig. 4.

As we explained in Section I, univariate methods tend to only be used for predicting load up to lead times of about four to six hours. In Fig. 5, we focus more closely on the post-sample three hour-ahead results for the three methods that performed the best in Fig. 4. Fig. 5 shows the three-hour ahead Mean MAPE results plotted against time of day. The largest Mean MAPE values occur for all three methods around 8am, and this is because, around this time of day, demand tends to be changing more rapidly than at other periods of the day. The plot shows the double seasonal Holt-Winters method dominating at almost all periods of the day. The results for the other two methods are much closer, with the seasonal ARMA method matching the PCA method except for the periods around 8am.
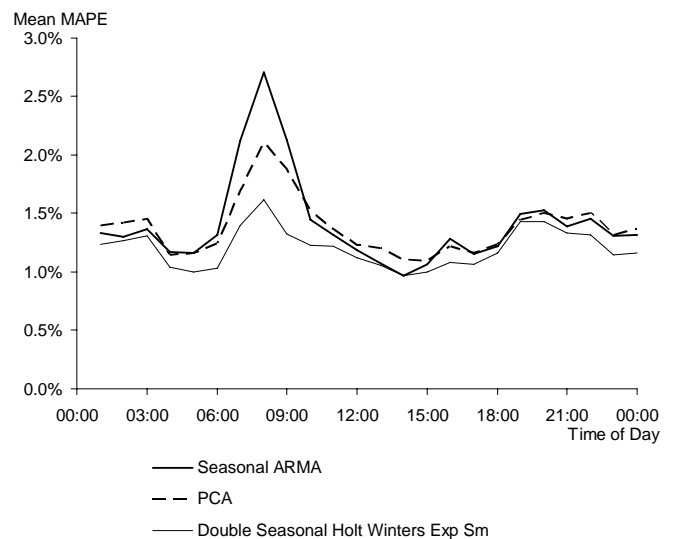

Fig. 5. For three hour-ahead prediction, Mean MAPE for the 10 load series plotted against time of day.

V. SUMMARY AND CONCLUDING COMMENTS

In this paper, we have used 10 time series of intraday electricity demand to compare empirically a number of univariate short-term forecasting methods. One of the aims of the paper has been to validate the findings in [8] using a substantially larger dataset. In addition to the methods that performed well in that study, we have also considered the intraday cycle exponential smoothing method (see [9]) and a

new periodic AR approach. Our results confirm the findings in [8]. All the sophisticated methods outperformed the two naïve benchmark methods, and the best performing method was double seasonal Holt-Winters exponential smoothing, followed by the method based on PCA, and then seasonal ARMA. The results for the new intraday exponential smoothing method and for the periodic AR model were a little disappointing. We suspect that the performance of the periodic AR model may improve with use of a longer time series. Our reasoning is that just 20 weeks of data may well be insufficient to capture the intraweek periodicity in a parameter. The same comment can also be applied to the intraday cycle exponential smoothing method because the approach involves the estimation of a relatively complex parameterization, which enables a form of periodic smoothing parameter.

The success of the double seasonal Holt-Winters exponential smoothing method is impressive, particularly in view of the method's simplicity. Ongoing work is aiming to gain insight into the method, with particular focus on the implication of including the autoregressive error correction term within the formulation. In terms of advising practitioners, the double seasonal Holt-Winters method would seem to be very attractive, as it is simple to understand and implement, and it has been shown to be accurate for short-term load prediction. Furthermore, the method is also appealing because of the existence of an underlying statistical model, which enables the calculation of prediction intervals. Finally, we should acknowledge that, if weather predictions are available, weather-based load forecasting methods may well be more accurate beyond about four to six hours ahead. However, for shorter lead times, the better of the univariate methods considered in this paper should be competitive. In addition, the univariate methods have strong appeal, in terms robustness, for online load prediction.

VI. Acknowledgements

VII. References

[1] P. E. McSharry, S. Bouwman, and G. Bloemhof, "Probabilistic forecasts of the magnitude and timing of peak electricity demand," *IEEE Transactions Power Systems*, vol. 20, pp. 1166-1172, 2005.

[2] E. Gonzalez-Romera, M.A. Jaramillo-Moran, D. Carmona-Fernandez, "Monthly Electric Energy Demand Forecasting Based on Trend Extraction," *IEEE Transactions on Power Systems*, vol. 21, pp. 1946-1953, 2006.

[3] M.P. Garcia, and D.S. Kirschen, "Forecasting system imbalance volumes in competitive electricity markets," *IEEE Transactions on Power Systems*, vol. 21, pp. 240- 248, 2006.

[4] J.W. Taylor, "Density forecasting for the efficient balancing of the generation and consumption of electricity," *International Journal of Forecasting*, vol. 22, pp. 707-724, 2006.

[5] D.W. Bunn, "Forecasting loads and prices in competitive power markets," *Proceedings of the IEEE*, vol. 88, pp. 163-169, 2000.

[6] J. W. Taylor, and R. Buizza, "Neural network load forecasting with weather ensemble predictions," *IEEE Transactions on Power Systems*, vol. 17, pp. 626-632, 2002.

[7] D.W. Bunn, "Short-term forecasting: A review of procedures in the electricity supply industry," *Journal of the Operational Research Society*, vol. 33, pp. 533-545, 1982.

[8] J.W. Taylor, L.M. de Menezes, and P.E. McSharry, "A comparison of univariate methods for forecasting electricity demand up to a day ahead," *International Journal of Forecasting*, vol. 22, pp. 1-16, 2006.

[9] P.G. Gould, A.B. Koehler, F. Vahid-Araghi, R.D. Snyder, J.K. Ord, and R.J. Hyndman, "Forecasting Time-Series with Multiple Seasonal Patterns", Working Paper, Monash University, Melbourne, Australia, 2007.

[10] H.S. Hippert, C.E. Pedreira, R.C. Souza, "Neural Networks for Short-Term Load Forecasting: A Review and Evaluation," *IEEE Trans. on Power Systems*, vol. 16, pp. 44-55, 2001.

[11] W.D. Laing, and D.G.C. Smith, "A comparison of time series forecasting methods for predicting the CEGB demand". Proceedings of the Ninth Power Systems Computation Conference, 1987.

[12] G.A. Darbellay, and M. Slama, "Forecasting the short-term demand for electricity – Do neural networks stand a better chance?", *International Journal of Forecasting*, vol. 16, pp. 71-83, 2000.

[13] G.E.P. Box, G.M. Jenkins, and G.C. Reinsel, *Time Series Analysis: Forecasting and Control*, third edition, New Jersey: Englewod Cliffs, Prentice Hall, 1994.

[14] P.H. Franses, and R. Paap, *Periodic Time Series Models*, Oxford University Press, Oxford, UK, 2004.

[15] D.R. Osborn, S. Heravi, and C.R. Birchenhall, "Seasonality and the order of integration for consumption," Oxford Bulletin of Economics and Statistics, vol. 50, pp. 361-377, 1988.

[16] S.J. Koopman, M. Ooms, and M.A. Carnero, "Periodic seasonal Reg-ARFIMA-GARCH models for daily electricity spot prices", *Journal of the American Statistical Association*, vol. 102, pp. 16-27, 2007.

[17] T.G. Andersen, and T. Bollerslev, "DM-dollar volatility: Intraday activity patterns, macroeconomic announcements and longer run dependencies," *Journal of Finance*, vol. 53, pp. 219-265, 1998.

[18] J.W. Taylor, "Short-term electricity demand forecasting using double seasonal exponential smoothing," *Journal of Operational Research Society*, vol. 54, pp. 799-805, 2003.

[19] Hyndman, R.J., Koehler, A.B., Snyder, R.D. and Grose, S, "A state space framework for automatic forecasting using exponential smoothing methods", *International Journal of Forecasting*, vol. 18, pp. 439-454, 2002.

[20] R. Ramanathan,, R. Engle, C.W.J. Granger, F. Vahid-Araghi, and C. Brace, "Short-run forecasts of electricity loads and peaks," *International Journal of Forecasting*, vol. 13, pp. 161-174, 1997.

VIII. Biographies

James W. Taylor is a Reader in Decision Science at the Saïd Business School, University of Oxford. His research interests include exponential smoothing, prediction intervals, quantile regression, volatility forecasting, energy forecasting and weather ensemble predictions.

Patrick E. McSharry is a Royal Academy of Engineering/EPSRC Research Fellow at the Department of Engineering Science, University of Oxford, and a Senior Member of IEEE. He is currently supported by a Marie Curie Research Fellowship, funded by the European Union's Sixth Framework. His research interests include signal processing, data analysis, complex dynamical systems, mathematical modelling and forecasting.