ISSN 1440-771X

MONASH University

Australia

Department of Econometrics and Business Statistics

http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/

Forecasting Intraday Time Series with Multiple Seasonal Cycles Using Parsimonious Seasonal Exponential Smoothing

James W. Taylor and Ralph D Snyder

October 2009

Working Paper 9/09

Forecasting Intraday Time Series with Multiple Seasonal Cycles

Using Parsimonious Seasonal Exponential Smoothing

James W. Taylor

Saïd Business School

University of Oxford

and

Ralph D. Snyder

Monash University

2 October 2009

Address for Correspondence:

James W. Taylor
Saïd Business School
University of Oxford
Park End Street
Oxford OX1 1HP, UK

Tel: +44 (0)1865 288927
Fax: +44 (0)1865 288805
Email: james.taylor@sbs.ox.ac.uk

Forecasting Intraday Time Series with Multiple Seasonal Cycles

Using Parsimonious Seasonal Exponential Smoothing

Forecasting Intraday Time Series with Multiple Seasonal Cycles Using Parsimonious Seasonal Exponential Smoothing

Abstract

This paper concerns the forecasting of seasonal intraday time series. An extension of Holt-Winters exponential smoothing has been proposed that smoothes an intraday cycle and an intraweek cycle. A recently proposed exponential smoothing method involves smoothing a different intraday cycle for each distinct type of day of the week. Similar days are allocated identical intraday cycles. A limitation is that the method allows only whole days to be treated as identical. We introduce an exponential smoothing formulation that allows parts of different days of the week to be treated as identical. The result is a method that involves the smoothing and initialisation of fewer terms than the other two exponential smoothing methods. We evaluate forecasting up to a day ahead using two empirical studies. For electricity load data, the new method compares well with a range of alternatives. The second study involves a series of arrivals at a call centre that is open for a shorter duration at the weekends than on weekdays. By contrast with the previously proposed exponential smoothing methods, our new method can model in a straightforward way this situation, where the number of periods on each day of the week is not the same.

Keywords: Exponential smoothing; Intraday data; Electricity load; Call centre arrivals.

Jel Code C22

INTRODUCTION

This paper is concerned with forecasting time series recorded at an intraday frequency. It is common for such data to exhibit seasonal cycles within the week and also within the day. Examples of the type of series of interest are the two samples of half-hourly electricity load presented in Figure 1. Both series show a repeating intraweek cycle of length m_2=336 periods and also, at least for the weekdays, there is a similar intraday pattern of length m_1=48 periods. The forecasting methods that we consider aim to capture this feature of double seasonality. Although electricity load is often modelled in terms of weather variables, the use of univariate methods is common for lead times less than half-a-day, and for longer lead times in locations where weather data is not available (Medeiros and Soares, 2008).Other examples of intraday data exhibiting double seasonality are call centre arrivals (see, for example, Tych et al., 2002), hospital admissions, and internet and transportation traffic counts (see, for example, Lam et al., 2006; Gould et al, 2008). In these applications, explanatory variables are typically not available, and so univariate forecasting methods is required.

A variety of univariate methods have been proposed for forecasting intraday data. These include the exponentially weighted regression of Christiaanse (1971); seasonal ARMA modelling (see, for example, Hagan and Behr, 1987); the forecast combining proposal of Smith (1989); artificial neural networks (see, for example, Darbellay and Slama, 2000); the dynamic harmonic regression of Tych et al. (2002); vector autoregression models, such as that of Cottet and Smith (2003); the random effects models of Weinberg et al. (2007); and Shen and Huang's (2008) approach based on singular value decomposition. In many applications with intraday data, the forecasting method must be implemented within an automated system. A class of univariate methods that have been widely used in automated applications is exponential smoothing (Hyndman et al., 2008). Exponential smoothing methods have performed well in empirical studies with intraday data (see, for example, Taylor et al., 2006). In this paper, we develop the use of exponential smoothing for intraday data.

Two exponential smoothing methods have been presented, in the literature, for intraday data. The Holt-Winter's method has been adapted by Taylor (2003) to give a method that involves the smoothing of an intraday cycle and intraweek cycle. We term this 'HWT exponential smoothing'. Although empirical results for the method have been encouraging (see, for example, Taylor et al., 2006), it has a couple of questionable features. Firstly, the method uses a common intraday cycle for all days of the week. Secondly, the method can be viewed as being of high dimension because the smoothing of an intraday cycle and an intraweek cycle requires the initialisation and updating of ($1+m_1+m_2$) terms. Gould et al. (2008) address these issues by introducing an exponential smoothing method that models the intraweek and intraday cycles through the use of a different intraday cycle for each distinct type of day of the week. For example, if the pattern of demand on Saturdays is different to the rest of the week, it is allocated its own intraday cycle. If all weekdays can be assumed to be the same, a common intraday cycle can be used for these five days. The result is a method with lower dimensionality than the HWT method. Due to its focus on the intraday cycle, the method has been termed 'intraday cycle (IC) exponential smoothing'.

A limiting feature of IC exponential smoothing is that it allows only whole days to be treated as identical. We would argue that it often makes more sense to assume that just parts of days are identical. For example, it could be that demand during daylight hours differs on each day of the week, but that the pattern of demand during night hours can be treated as identical on all days of the week. In this paper, we propose a new exponential smoothing formulation that has the flexibility to allow parts of different days of the week to be treated as identical. The result of this flexibility is a model of lower dimension than the IC method. In view of this, we term the method 'parsimonious seasonal exponential smoothing'. The HWT and IC methods can be viewed as special cases of this new method.

In this paper, we present two empirical studies. The penultimate section of the paper describes the analysis of a half-hourly series of the number of calls arriving at a call centre. The efficient scheduling of call centre staff relies on accurate forecasts for the number of calls (Gans et al., 2003). The call centre is open for a shorter duration at the weekends than on weekdays. This situation, where the number of periods on each day of the week is not the same, cannot be satisfactorily addressed with the HWT and IC methods. By contrast, it can be modelled in a straightforward way using parsimonious seasonal exponential smoothing.

The other empirical study involves two series of electricity load; one for Great Britain and the other for France. Automated electricity demand prediction is required for the control and scheduling of power systems. The liberalisation of electricity markets has meant that load forecasting is also important for market participants to support energy transactions (Bunn, 2000). We use the two load series to illustrate the implementation of methods throughout the earlier sections of the paper, and then, in a later section, we use the series to compare the forecast accuracy of the various methods. The series in Figure 1 are samples of the two electricity load series. The full series spanned the three-year period from Thursday 1 January 2004 to Sunday 31 December 2006, inclusive. We used the first two years of data for estimation of method parameters, and the remaining year for post-sample evaluation. As the estimation sample contains just two years of data, we make no attempt to model annual seasonality. We focussed on lead times from one half-hour-ahead up to one day-ahead.

In the next section, we review previously proposed exponential smoothing methods. We then introduce our new parsimonious seasonal exponential smoothing method. The section that follows uses the load data to compare post-sample forecast accuracy for the exponential smoothing methods, as well as ARMA modelling and an artificial neural network. We then present our empirical study of intraday call centre arrivals. The final section summarises, and provides concluding comments.

PREVIOUSLY PROPOSED EXPONENTIAL SMOOTHING METHODS FOR INTRADAY DATA

HWT exponential smoothing

Taylor's (2003) HWT exponential smoothing method is a relatively simple extension of Holt-Winters exponential smoothing. Taylor's formulation aims to model both the intraday and intraweek cycles in intraday data. It is presented in expressions (1)-(5):

$$\hat{y}_t(k) = l_t + d_{t-m_1+k_1} + w_{t-m_2+k_2} + \phi^k e_t \quad (1)$$

$$e_t = y_t - \left(l_{t-1} + d_{t-m_1} + w_{t-m_2}\right) \quad (2)$$

$$l_t = \alpha\left(y_t - d_{t-m_1} - w_{t-m_2}\right) + (1-\alpha)l_{t-1} \quad (3)$$

$$d_t = \delta\left(y_t - l_t - w_{t-m_2}\right) + (1-\delta)d_{t-m_1} \quad (4)$$

$$w_t = \omega\left(y_t - l_t - d_{t-m_1}\right) + (1-\omega)w_{t-m_2} \quad (5)$$

where l_t is the smoothed level; d_t is the seasonal index for the intraday seasonal cycle; w_t is the seasonal index for the intraweek seasonal cycle that remains after the intraday cycle is removed; α, δ and ω are smoothing parameters; and $\hat{y}_t(k)$ is the k step-ahead forecast made from forecast origin t;

and k_1=[(k-1) mod m_1]+1 and k_2=[(k-1) mod m_2]+1. The term involving the parameter ϕ, in the forecast function of expression (1), is an adjustment for first-order residual autocorrelation. As shown by Taylor, this adjustment substantially improves the performance of the method, and so should be considered integral to the method. A trend term is not included in any of the exponential smoothing formulations in this paper because, in our empirical work, its inclusion resulted in no change in forecast accuracy. We present all exponential smoothing formulations with additive seasonality. Multiplicative seasonality formulations led to similar results to the analogous additive versions.

Hyndman et al. (2008, Section 2.5) describe how innovations state space models are an attractive class of statistical models for representing exponential smoothing methods. These models contain a single source of error, and Hyndman et al. 2008 (Chapter 13) argue that they have a number of advantages over multiple source of error models. Parameter estimation is simpler for the single source of error models, and, because their updating equations are identical in form to their model equations, interpretation and manipulation is easier. Monte Carlo simulation provides a simple way to generate prediction intervals from innovations state space models. The following innovations state space model captures the essential features of HWT exponential smoothing (see Taylor, 2009):

$$y_t = l_{t-1} + d_{t-m_1} + w_{t-m_2} + \phi e_{t-1} + \varepsilon_t \qquad (6)$$

$$e_t = y_t - \left(l_{t-1} + d_{t-m_1} + w_{t-m_2}\right) \qquad (7)$$

$$l_t = \alpha\left(y_t - d_{t-m_1} - w_{t-m_2}\right) + (1-\alpha)l_{t-1} \qquad (8)$$

$$d_t = \delta\left(y_t - l_t - w_{t-m_2}\right) + (1-\delta)d_{t-m_1} \qquad (9)$$

$$w_t = \omega\left(y_t - l_t - d_{t-m_1}\right) + (1-\omega)w_{t-m_2} \qquad (10)$$

where ε_t ~ NID(0,σ^2). This notation is used to indicate that the ε_t are independent and normally distributed with zero mean and constant variance σ^2. An advantage of the HWT method is its relative simplicity. Indeed, its formulation will seem natural and intuitive to the many people familiar with the standard Holt-Winters method. Turning to the method's disadvantages, Gould et al. (2008) argue that it is unappealing to use the same intraday cycle state variable, d_t, for each of the seven days of the week. Furthermore, the HWT method can be viewed as being of high dimension, as it involves initialising and updating the level, plus the m_1 periods of the intraday cycle, as well as the m_2 periods of the intraweek cycle. For half-hourly load data, this amounts to (1+m_1+m_2)=385 terms.

IC exponential smoothing

The IC exponential smoothing method of Gould et al. (2008) allows the intraday cycle for the different days of the week to be represented by different seasonal components. To obtain a more parsimonious formulation than the HWT method, Gould et al. propose that a common intraday cycle is used for days of the week that exhibit similar patterns of demand. By contrast with the HWT method, there is no representation in the formulation for the intraweek seasonal cycle.

Before presenting the model, let us first consider which days of the week can be considered to have a similar intraday cycle for our electricity load series. To guide us, we follow the approach of Gould et al. by using the plots in Figures 2 and 3, which show, for each day of the week, the average load for each period of the day, calculated using in-sample data. For the British data, Figure 2 shows that Saturday and Sunday should each be treated as having its own distinct cycle. In this figure, the patterns for Monday morning and Friday afternoon suggest that there should be a distinct cycle for each of these days. The patterns on the other three weekdays are very similar, which suggests a common cycle for these days. Similar conclusions can be drawn from Figure 3 for the French series, although the argument for a distinct Friday intraday cycle is less clear. It seems reasonable to implement, for both series, a model with distinct intraday cycles for Friday, Saturday, Sunday and Monday, and a common cycle for the other three days. IC exponential smoothing for the case of five distinct intraday cycles is presented as an innovations state space model in the following expressions:

$$y_t = l_{t-1} + \sum_{i=1}^{5} I_{it}\, d_{i,t-s_1} + \phi e_{t-1} + \varepsilon_t \tag{11}$$

$$e_t = y_t - \left(l_{t-1} + \sum_{i=1}^{5} I_{it}\, d_{i,t-s_1} \right) \tag{12}$$

$$l_t = l_{t-1} + \alpha e_t \tag{13}$$

$$d_{it} = d_{i,t-s_1} + \left(\sum_{j=1}^{5} \gamma_{ij} I_{jt} \right) e_t \qquad (i = 1 \text{ to } 5) \tag{14}$$

$$I_{jt} = \begin{cases} 1 & \text{if time period } t \text{ occurs on a day of type } j \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

where $\varepsilon_t \sim \text{NID}(0,\sigma^2)$; l_t is the smoothed level; d_{it} is the value of the intraday cycle of type i in period t (for i=1 to 5); and α and the γ_{ij} are smoothing parameters. Although not included by Gould et al., the addition of the residual adjustment term involving the parameter ϕ led to noticeable improvement in forecast accuracy. The γ_{ij} can be viewed as a 5×5 matrix of parameters. Through these parameters, the formulation enables the five types of intraday cycle to be updated at different rates. Gould et al. propose several alternative restrictions for the γ_{ij} matrix; one involves estimation with just the basic constraint that the parameters lie between zero and one, and the other involves the additional restrictions of common diagonal elements and common off-diagonal elements. We refer to these two forms of the model as unrestricted and restricted, respectively. As noted by Gould et al., the restricted form is identical to the HWT method, provided seven distinct intraday cycles are used and the seasonal smoothing parameters are written in terms of the HWT parameters as $\gamma_{ii}=\delta+\omega$ and $\gamma_{ij}=\delta$.

In comparison with the HWT method, the IC method has the appeal of lower dimensionality. As noted at the end of the previous section, the HWT method involves the initialisation and updating of $(1+m_1+m_2)=385$ terms. By contrast, the IC method, with five distinct intraday cycles, involves the initialising and updating the level and $5m_1$ seasonal periods, which amounts to $(1+5\times48)=241$ terms.

A NEW PARSIMONIOUS SEASONAL EXPONENTIAL SMOOTHING METHOD

A limiting feature of IC exponential smoothing is that it allows only whole days to be clustered together. We would argue that it often makes more sense to allow parts of days to be clustered together. For example, in the IC method of the previous section, a distinct intraday cycle was used for Fridays. However, for the British data, Figure 2 shows that the average Friday pattern of load prior to about 11am seems to be very similar to that on Tuesday, Wednesday and Thursday, and so it seems inefficient to treat these earlier Friday hours as distinct from the rest of the week.

We propose a new exponential smoothing formulation that allows parts of different days of the week to be treated as identical, with the remaining parts of these days treated as distinct. In comparison with the IC method, the new formulation offers greater flexibility and parsimony. This prompts us to term the method ‘parsimonious seasonal exponential smoothing’. In this method, we use the term ‘season’ to refer to a set of periods in the intraweek cycle for which demand is assumed to be identical. The method proceeds by clustering the m_2 periods of the week so that each belongs to one of M distinct seasons, where $M \leq m_2$. (We discuss the clustering for the load data in the next section.) A seasonal state s_{it} is defined for each season i, and exponential smoothing is used to update the M seasonal states. The simplest version (Version 1) of parsimonious seasonal exponential smoothing is presented in the form of an innovations state space model in expressions (16)-(19).

Parsimonious seasonal exponential smoothing – Version 1

$$y_t = \sum_{i=1}^{M} I_{it} s_{i,t-1} + \phi e_{t-1} + \varepsilon_t \tag{16}$$

$$e_t = y_t - \sum_{i=1}^{M} I_{it} s_{i,t-1} \tag{17}$$

$$s_{it} = s_{i,t-1} + (\alpha + \omega I_{it}) e_t \qquad (i = 1, 2, \ldots, M) \tag{18}$$

$$I_{it} = \begin{cases} 1 & \text{if period } t \text{ occurs in season } i \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

where $\varepsilon_t \sim \mathrm{NID}(0,\sigma^2)$; α and ω are smoothing parameters; and ϕ is the parameter of a residual autoregressive term. Each seasonal state, s_{it}, is updated using expression (18), with the degree of updating depending on whether or not the current period t is in season i. There is no separate smoothing of the level, as in the HWT and IC methods. Instead, the level is captured within the seasonal states, and updated each period through the parameter α. This gives a more concise formulation, and potentially simplifies initialisation of the states. The smoothing of a trend can be incorporated by the addition of a constant term in expression (18). However, this term was not beneficial for the series considered in this paper.

Clustering periods of the week into distinct seasons

Drawing on the approach used to specify IC exponential smoothing, we now use Figures 2 and 3 to dictate the clustering of the periods of the week into seasons for parsimonious seasonal exponential smoothing. For the British plot in Figure 2, we make the following observations:

(i) As the average demand on each weekend period is so very different to the other periods of the week, we define 96 distinct seasons for the weekend periods.

(ii) The average patterns of demand on Tuesdays, Wednesdays and Thursdays are so similar that they should be treated as having a common intraday cycle. For this, we define 48 distinct seasons.

(iii) After about 8.30am, the average pattern of demand on Mondays is extremely similar to that on Tuesdays, Wednesdays and Thursdays. This implies that we should specify these 31 Monday periods to be identical to the corresponding periods on Tuesdays, Wednesdays and Thursdays. However, for each of the first 17 periods on Mondays, we specify distinct seasons, as they would seem to be noticeably different from the other periods of the week.

(iv) Up to about 11am, the average pattern of demand on Fridays is very similar to that on Tuesdays, Wednesdays and Thursdays, and so we specify these 22 Friday morning periods to be identical to the corresponding periods on Tuesdays, Wednesdays and Thursdays. However, we specify distinct seasons for the last 26 periods on Fridays.

This assessment of Figure 2 leads us to define (96+48+17+26)=187 distinct seasons for the British demand data. This can be compared with the number of terms that require initialisation and updating in the HWT and IC methods. As we discussed earlier, the HWT method involves $(1+m_1+m_2)$=385 terms, and our implementation of the IC method has (1+5×48)=241 terms.

Let us now consider the French data. A similar set of observations to those made above regarding Figure 2 can be made for the French average intraday cycles plotted in Figure 3:

(i) 96 distinct seasons are defined for the weekend periods.

(ii) 48 distinct seasons are used to represent the common intraday cycle occurring on Tuesdays, Wednesdays and Thursdays.

(iii) The average pattern of demand up to midday on Mondays is rather different to the other days, and so we define distinct seasons for the first 24 Monday periods. The other 24 periods are defined to be identical to the corresponding periods on Tuesdays, Wednesdays and Thursdays.

(iv) For Fridays, the average demand is somewhat different to the other days beyond about 1pm. Therefore, we specify 22 distinct seasons for these Friday periods. The other 26 periods are defined to be identical to the corresponding periods on Tuesdays, Wednesdays and Thursdays.

The result of this is (96+48+24+22)=190 distinct seasons for the parsimonious seasonal exponential smoothing model applied to the French data. We acknowledge that, based on Figures 2 and 3, other clusters of periods into seasons could be considered, but we have opted here for simplicity. As we have presented it, the selection of clusters relies on the judgement of the user. We also considered statistical approaches to clustering that require only minimal judgemental input. We considered hierarchical and k-mean cluster analysis based on the sum of the Euclidean distances between observations for different periods of the same historical week in the estimation sample. We implemented an unconstrained clustering that allowed any periods of the week to be clustered in the same season, and hence treated as identical. We also experimented with a constrained version that only permitted clusters to be formed from periods of different days that corresponded to the same half-hourly period of the day. However, none of the statistical clustering approaches led to improved accuracy over the judgemental selections of clusters. In view of this, we rely on a purely judgemental specification, which is the approach taken by Gould et al. for IC exponential smoothing.

Extending the parsimonious seasonal exponential smoothing model

The HWT and IC exponential smoothing methods are special cases of extended versions of parsimonious seasonal exponential smoothing. To obtain the HWT method, we first set M=m_2, so that each period of the week is treated as a distinct season. We then extend the parsimonious seasonal exponential smoothing model to the formulation of expressions (20)-(24), which we term 'Version 2' of the model. In comparison with Version 1, Version 2 includes an additional term in the updating equations of expression (22), and a corresponding additional smoothing parameter δ. The updating equations cause the seasonal states to be updated to differing degrees depending on whether or not the current period t falls in season i or in the same period of the day as season i.

Parsimonious seasonal exponential smoothing – Version 2

$$y_t = \sum_{i=1}^{M} I_{it} s_{i,t-1} + \phi e_{t-1} + \varepsilon_t \tag{20}$$

$$e_t = y_t - \sum_{i=1}^{M} I_{it} s_{i,t-1} \tag{21}$$

$$s_{it} = s_{i,t-1} + \left(\alpha + \delta I_{it}^{period\,of\,day} + \omega I_{it}\right) e_t \qquad \left(i = 1, 2, \ldots, M\right) \tag{22}$$

$$I_{it} = \begin{cases} 1 & \text{if period } t \text{ occurs in season } i \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

$$I_{it}^{period\,of\,day} = \begin{cases} 1 & \text{if period } t \text{ occurs at same period of day as season } i \\ 0 & \text{otherwise} \end{cases} \tag{24}$$

where $\varepsilon_t \sim \text{NID}(0,\sigma^2)$.

The restricted version of the IC method is also a special case of Version 2 of the parsimonious seasonal exponential smoothing model. Let us consider the restricted method with the same five intraday cycles specified in an earlier section for this method. We obtain this model if, in Version 2 of the parsimonious seasonal exponential smoothing model, we define M=(5×m_1) distinct seasons as being the periods on Friday, Saturday, Sunday, Monday and Tuesday, and we specify that Wednesday and Thursday are identical to Tuesday. With the same M=(5×m_1) distinct seasons, a further extension of parsimonious seasonal exponential smoothing delivers the unrestricted form of the IC method. This extension is given in expressions (25)-(30). In addition to requiring that each period of the week is assigned to one of the M seasons, this version of the model also requires that each period of the week is classified as belonging to one of five day types. As in the unrestricted IC method, this parsimonious seasonal exponential smoothing formulation contains a smoothing parameter α for the level and a matrix of parameters γ_{jk} to dictate the updating of the seasonality.

Parsimonious seasonal exponential smoothing – Version 3

$$y_t = \sum_{i=1}^{M} I_{it} s_{i,t-1} + \phi e_{t-1} + \varepsilon_t \tag{25}$$

$$e_t = y_t - \sum_{i=1}^{M} I_{it} s_{i,t-1} \tag{26}$$

$$s_{i,t} = s_{i,t-1} + \left(\alpha + I_{it}^{period\,of\ day} \sum_{j=1}^{5} \sum_{k=1}^{5} \gamma_{jk} I_{jt}^{period\,day\,type} I_{ki}^{season\,day\,type} \right) e_t \qquad (i = 1, 2, \ldots, M) \tag{27}$$

$$I_{it}^{period\,of\ day} = \begin{cases} 1 & \text{if period } t \text{ occurs at same period of day as season } i \\ 0 & \text{otherwise} \end{cases} \tag{28}$$

$$I_{jt}^{period\,day\ type} = \begin{cases} 1 & \text{if period } t \text{ occurs on a day of type } j \\ 0 & \text{otherwise} \end{cases} \tag{29}$$

$$I_{ki}^{season\,day\ type} = \begin{cases} 1 & \text{if season } i \text{ occurs on a day of type } k \\ 0 & \text{otherwise} \end{cases} \tag{30}$$

where ε_t ~ NID(0,σ^2).

We have introduced Versions 2 and 3 of the parsimonious seasonal exponential smoothing model as being equivalent, with appropriate choices of the M distinct seasons, to the HWT and IC methods. However, our proposal is that these additional versions of the new model could be used with other choices for the M distinct seasons. In the next section, we consider the three versions of the method with the judgementally selected clusters that earlier led to M=187 for the British series, and

M=190 for the French series. To implement Version 3 of the model for the British series, we defined the following five different ‘day types’: Saturday is ‘day type 1’; Sunday is ‘day type 2’; Tuesday, Wednesday, Thursday, Monday periods after 8.30am, and Friday periods before 11am are all ‘day type 3’; Monday morning periods up to 8.30am are ‘day type 4’; and Friday periods from 11am onwards are ‘day type 5’. For the French series, we applied the same ‘day type’ definitions, except that 8.30am was replaced by midday, and 11am was replaced by 1pm.

EVALUATION OF FORECASTS FOR THE LOAD TIME SERIES

In this section, we use the two load series to compare the post-sample forecast accuracy of the exponential smoothing methods and several benchmark methods. Estimation was performed once using the first two years of the data, and the forecast origin was then rolled forward through the one-year post-sample evaluation period to produce a collection of forecasts for lead times up to a day ahead. Before evaluating forecast accuracy, we first describe the estimation of the exponential smoothing methods and introduce the benchmark methods.

Exponential smoothing methods

We implemented the HWT method; the restricted and unrestricted forms of the IC method; and the three versions of parsimonious seasonal exponential smoothing. Initial smoothed values for the level and seasonal components were estimated by averaging the observations from the first three weeks of data. We constrained the parameters to lie between zero and one, and estimated them by minimizing the sum of squared in-sample forecast errors (SSE). For each method, the optimisation proceeded by first generating 10^5 vectors of parameters from a uniform random number generator between 0 and 1. For each of the vectors, we then evaluated the SSE. The 10 vectors that produced the lowest SSE values were used, in turn, as the initial vector in a quasi-Newton algorithm. Of the 10 resulting vectors, the one producing the lowest SSE value was chosen as the final parameter vector.

For the British series, the optimized parameters for the HWT method were α=0.003, δ=0.295, ω=0.397 and ϕ=0.968, and for the restricted form of the IC method, we derived α=0.013, γ_{ij}=0.281, γ_{ii}=0.640 and ϕ=0.953. The low values of α and high values of ϕ for the HWT method are consistent with previous studies with intraday load data (see, for example, Taylor et al., 2006). Earlier, we noted that the restricted form of the IC method is identical to the HWT method, provided seven distinct intraday cycles are used and the parameters for the two methods obey $\gamma_{ii}=\delta+\omega$ and $\gamma_{ij}=\delta$. Our derived parameters, for both the British and French data, approximately satisfy these two equations. Even though our IC method implementation involves five distinct intraday cycles, it seems likely that the HWT and restricted IC methods will deliver similar results. For Version 1 of the parsimonious seasonal exponential smoothing method, the optimised parameter values for the British data were α=0.035, ω=0.739 and ϕ=0.936, and for Version 2, we obtained α=0.018, δ=0.272, ω=0.346 and ϕ=0.947. The values for Version 2 are similar to those derived for the HWT method, which is perhaps a little surprising because the two methods have such different dimensionality.

Simplistic benchmark methods

We implemented two simplistic benchmark methods. The first takes as a forecast the most recently observed value for the same half-hour of the week as the period to be predicted. The second produces forecasts as the average of the observations occurring at the same half-hour of the week in the most recent four weeks.

Double seasonal ARMA

Double seasonal ARMA models are often used as sophisticated benchmarks in load forecasting studies (e.g. Darbellay and Slama, 2000; Pedregal and Young, 2008). The multiplicative double seasonal ARMA model 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 ε_t ~ NID(0,σ^2); 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 lack of clear evidence from unit root tests initially led us not to apply differencing. However, for similar data to ours, some researchers have applied the differencing operator $(1-L)(1-L^{s_1})(1-L^{s_2})$ (see, for example, Darbellay and Slama, 2000; Pedregal and Young, 2008). In view of this, we implemented two versions of double seasonal ARMA modelling; one with this differencing operator and one with no differencing. We estimated model parameters using maximum likelihood. We considered lag polynomials up to order three, and selected AR and MA terms using the Schwarz Bayesian Criterion, with the requirement that all parameters were significant (at the 5% level).

Artificial neural networks

Much research has focused on the use of ANNs to model electricity load in terms of weather variables (see Hippert et al., 2001). In that context, the nonlinear and nonparametric features of ANNs have strong appeal. Their usefulness for univariate modelling is less clear, and relies on the nonlinear structures in the evolution of the load series. Nevertheless, we felt it would be interesting to include an ANN as a second sophisticated benchmark method in our study. Darbellay and Slama (2000) and Taylor et al. (2006) build univariate ANNs to model intraday load data with load at various lags as inputs. They both use a single model, and generate multi-step-ahead predictions by using forecasts for earlier lead times. Our approach differed from this by employing a separate ANN for each lead time. Using a hold-out sample of six months of the estimation sample, we found that forecast accuracy was improved by applying the differencing operator $(1-L^{s_1})(1-L^{s_2})$ prior to ANN modelling.

We used a single hidden layer feedforward network with k inputs, x_{it}, connected to each of m units in a single hidden layer, which, in turn, were connected to a single output, y_t. The model is written

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

where the functions $g_1(\cdot)$ and $g_2(\cdot)$ are chosen as sigmoidal and linear respectively, and w_{ji} and v_j are the weights. The output was specified as the load variable, differenced as described above, and normalised by subtracting the mean and dividing by the standard deviation. In the model for h step-ahead prediction, as inputs, we used this variable at the forecast origin and at the following lags: 1, 2, s_1-h, $2s_1$-h, $3s_1$-h, s_2-h, $2s_2$-h, $3s_2$-h, s_3-h, $2s_3$-h and $3s_3$-h. The network weights were estimated using least squares. To penalise complexity and avoid overfitting, we added to the objective function the sum of the weights squared multiplied by a regularisation parameter λ (see Bishop, 1997, §9.2). The minimisation was performed using backpropagation with learning rate η and momentum parameter μ (see Bishop, 1997, §7.5).

To derive values for λ, η, μ and m, and the choice of lags to include as inputs, we evaluated one step-ahead prediction accuracy using a hold-out sample of six months of the estimation sample. We then used the same values and inputs for the models for all other lead times. For the British series, this led to the inclusion of all inputs except lag 2, and λ=0.0001, η=0.1, μ=0.9 and m=4. For the French data, we found that all inputs should be included, and we obtained the same parameter values, with the exception that m=8.

Post-sample load forecasting results

In Figure 4, for each method, we present the MAPE results averaged over the two series. Presenting just the average seems reasonable because, for most of the methods considered, their relative performances were similar for the two load series. The two ARMA methods were an exception to this, with differencing being preferable for the British series and no differencing being better for the French data. We also evaluated the root mean squared percentage error (RMSPE), mean absolute error (MAE) and root mean squared error (RMSE), but we do not report these results here because the rankings of the methods for these measures were similar to those for the MAPE.

The results for the restricted IC method were similar to those of the HWT method and Version 2 of the parsimonious seasonal exponential smoothing method, and the unrestricted IC method results were similar to those of Version 3 of the parsimonious seasonal exponential smoothing method. In view of this, for simplicity, we have omitted the IC method results from Figure 4. The results for both of the simplistic benchmark methods were so relatively poor that, with our choice of scaling of the axes, they are not shown in Figure 4. With regard to the sophisticated benchmark methods, Figure 4 shows the ARMA approach with differencing slightly outperforming the ARMA method with no differencing. The ANN is relatively uncompetitive for the early lead times, but outperforms both ARMA methods beyond about 15 hours ahead. The HWT method matches or beats the ARMA and ANN methods at all lead times. Turning to parsimonious seasonal exponential smoothing, we see that Version 1 performed relatively poorly, Version 2 produced similar results to the HWT method, and Version 3 produced the best results across all lead times.

EMPIRICAL STUDY OF CALL CENTRE TIME SERIES

In this section, we consider the Israeli bank data introduced and analysed by Brown et al. (2005), and also used by Taylor (2008). This data is taken from a small call centre that is open from 7am to midnight on weekdays (Sunday to Thursday in Israel), from 7am to 2pm on Fridays, and from 8pm to midnight on Saturdays. Our analysis focuses on the single time series of total arrivals minus internet assistance calls for the period 1 August to 25 December 1999, inclusive, with the first 14 weeks used for estimation, and the remaining seven weeks used for post-sample evaluation. We smoothed out the 'special days' in the series prior to method estimation. To be consistent with our analysis of the load data, we analyzed half-hourly arrivals and we chose to focus on lead times up to a day ahead. Forecasts for such short lead times are used for dynamic intraday updating of the deployment of call centre agents (Hur et al., 2004).

In the call centre literature, several authors have applied a variance stabilizing transformation prior to model fitting. Brown et al. (2005) present one such transformation, which draws on the count data nature of call centre arrivals. They explain that if y is a Poisson random variable, then the transformed variable $\sqrt{y+\frac{1}{4}}$ is asymptotically Gaussian as $\lambda\rightarrow\infty$. We applied the transformation prior to the use of each forecasting method.

For most of the methods considered in this paper, implementation is not straightforward for series in which the number of periods is not the same on each day of the week, which is the case for the Israeli bank series. An exception to this is the parsimonious seasonal exponential smoothing method. Before discussing this further, we describe how, in order to be able to apply the other methods, we converted the Israeli bank series into a new series with the same number of periods in each day. We achieved this by combining the two weekend days into one single day consisting of the same number of periods as the weekdays. The result is a series with five weekdays plus a sixth day consisting of Friday arrivals from 7am to 2pm, zero values from 2pm to 8pm, and Saturday arrivals from 8pm to midnight. We refer to this as the 'manipulated' arrivals series. Figure 5 presents the final four weeks of the new manipulated series. Each day of this series consists of m_1=34 half-hourly periods, and each week contains m_2=6×34=204 half-hourly periods. In our description of the forecasting methods below, we present separately the methods applied to the manipulated series, and those applied to the original series.

Methods applied to the manipulated arrivals series

For the manipulated arrivals series, we implemented all of the methods used with the load data, except the parsimonious seasonal exponential smoothing method. To specify the IC method, we plotted the average intraday patterns for each day of the week, as shown in Figure 6. As the estimation sample is considerably smaller for the call centre series than for the load series, the average intraday patterns in Figure 6 are subject to more sampling error than those for the load data in Figures 2 and 3. In view of this, although there are differences in the average intraday patterns for the Israeli weekdays, Sunday to Thursday, it seems sensible to assume the weekdays possess a common intraday cycle in the IC method. The average intraday patterns on Fridays and Saturdays appear to be lower than those on the other days, and so, for the IC method, we specified a distinct intraday cycle for the day that is made up of the opening hours on Fridays and Saturdays.

Methods applied to the original arrivals series

The original arrivals series consisted of five days with 34 half-hourly periods, one day with 14 periods, and one day with just eight periods. This implies an intraweek cycle of length m_2=192 periods. Due to the unequal number of periods on the days of the week, we were unable to apply to the original series, the HWT method with its double seasonality features. However, by setting d_t=0, for all t, in the HWT method formulation, we have a simpler method that is able to model just the intraweek cycle. We applied this 'single seasonal HWT' method to the original arrivals series. (Note that this method is not the same as standard Holt-Winters exponential smoothing because of the presence of the residual autocorrelation term in the single seasonal HWT formulation.)

The weakness of single seasonal HWT is that it does not capture the commonality in the intraday cycles for the five weekdays. A method that can model this common feature, as well as the intraweek cycle, is the parsimonious seasonal exponential smoothing method. For this method, we specified 34 seasons to represent the intraday cycle that is common to the five weekdays, 14 seasons for the Friday periods, and a further eight seasons for the Saturday periods. This implies a total of M=56 seasons. We found that Version 2 of the method offered no improvement over Version 1. This suggested to us that Version 3 would not be useful. Furthermore, we felt the simplicity of the choice of seasons and the relatively short length of the series did not warrant use of Version 3 of the method.

Post-sample call centre forecasting results

We evaluated the various methods using MAE and RMSE. Percentage measures could not be used due to the existence of periods where the series was close or equal to zero. It is important to note that, regardless of whether a method was applied to the original or manipulated arrivals series, post-sample forecast accuracy was recorded for the same set of periods, consisting of just the opening hours of the call centre. For simplicity, we present and discuss only the MAE results, as the relative rankings of the methods were similar when the RMSE was used. These results are presented in Figure 7. We do not show the results for the ANN method, as they were not competitive. This is presumably at least in part due to the relatively short length of the time series. We also do not present the results for either the restricted or unrestricted IC methods because the results for these methods were very similar to those of the HWT method.

In Figure 7, the simplistic benchmark methods are outperformed by all the sophisticated methods, except the ARMA method that involved no differencing. Using differencing can be seen to lead to improvement for the ARMA method at all lead times. However, the figure shows that this ARMA method is, in turn, beaten at all lead times by the HWT method. Let us now turn to the two methods that were applied to the original arrivals series. The figure shows that single seasonal HWT offers a small improvement on the best of the methods applied to the manipulated arrivals series. The other method that was applied to the original series was the parsimonious seasonal exponential smoothing method. For simplicity, we show just the results for Version 1 of this method because Version 2 offered no improvement. The figure shows that the best results at all lead times were achieved by this parsimonious seasonal exponential smoothing method. It is impressive that the method beats not only those methods applied to manipulated arrivals series, but also the single seasonal HWT method applied to the original data.

SUMMARY AND CONCLUDING COMMENTS

In contrast to the HWT exponential smoothing method, the IC exponential smoothing method avoids the use of a common intraday cycle for all days of the week, and instead specifies a distinct intraday cycle for days that cannot be assumed to be the same. Dimensionality of the method is reduced and efficiency increased by allowing the same intraday cycle to be used for days that can be assumed to be identical. However, an unappealing feature of the IC method is that it allows only whole days to be treated as identical. We feel that it often makes more sense to assume that just parts of days are identical. In this paper, we have introduced a new parsimonious seasonal exponential smoothing model for intraday data that has the flexibility to allow parts of different days of the week to be treated as identical. Post-sample forecasting results for two electricity load series showed that the method competes well with the HWT and IC methods.

Our second empirical study involved a series of call volumes from a call centre that is open for a shorter duration at the weekends than on weekdays. The HWT and IC methods cannot satisfactorily address the situation where the number of periods on each day of the week is not the same. However, it can be modelled in a straightforward way using parsimonious seasonal exponential smoothing. Post-sample forecasting results supported the use of this new method.

In terms of future work regarding the parsimonious seasonal exponential smoothing method, it may be useful to have some form of statistical measures or procedures to support the clustering of the periods of the week into distinct seasons. In this paper, we used the average historical intraday patterns to assist judgmental clustering. Although our experimentation with statistical cluster analysis did not produce promising results, a fresh look at this with alternative data may be more fruitful. It would also be interesting to consider the use of information criteria to select among different specifications of the method. Finally, although the method has been used in this paper for intraday data that exhibits repeating intraday and intraweek seasonal cycles, the method may also be useful for data that possesses just a single repeating seasonal cycle.

Acknowledgements

We would like to thank Matthew Roberts of National Grid for supplying the British load data, and RTE for making the French load data available on their website. We acknowledge Avi Mandelbaum for making available the Israeli bank data.

References

Bishop CM. 1997. *Neural Networks for Pattern Recognition*. Oxford University Press: Oxford, UK.

Brown LD, Gans N, Mandelbaum A, Sakov A, Shen H, Zeltyn S, Zhao L. 2005. Statistical analysis of a telephone call center: A queuing science perspective. *Journal of the American Statistical Association* **100**: 36-50.

Bunn DW. 2000. Forecasting loads and prices in competitive power markets. *Proceedings of the IEEE* **88**: 163-169.

Christiaanse WR. 1971. Short-term load forecasting using general exponential smoothing. *IEEE Transactions on Power Apparatus and Systems* **PAS-90**: 900-910.

Cottet R, Smith M. 2003. Bayesian modeling and forecasting of intraday electricity load. *Journal of the American Statistical Association* **98**: 839–849.

Darbellay GA, Slama M. 2000. Forecasting the short-term demand for electricity – Do neural networks stand a better chance? *International Journal of Forecasting* **16**: 71-83.

Gans N, Koole G, Mandelbaum A. 2003. Telephone call centers: Tutorial, review, and research prospects. *Manufacturing and Service Operations Management* **5**: 79-141.

Gould PG, Koehler AB, Ord JK, Snyder RD, Hyndman RJ, Vahid-Araghi F. 2008. Forecasting time-series with multiple seasonal patterns. *European Journal of Operational Research* **191**: 207-222.

Hagan MT, Behr SM. 1987. The time series approach to short-term load forecasting. *IEEE Transactions on Power Systems* **PWRS-2**: 785-791.

Hippert HS, Pedreira CE, Souza RC. 2001. Neural networks for short-term load forecasting: a review and evaluation. *IEEE Transactions on Power Systems* **16**: 44-55.

Hur D, Mabert VA, Bretthauer KM. 2004. Real-time work schedule adjustment decisions: An investigation and evaluation. *Productions and Operations Management* **13**: 322-339.

Hyndman RJ, Koehler AB, Ord JK, Snyder RD. 2008. *Forecasting with Exponential Smoothing: The State Space Approach*. Springer-Verlag: Berlin Heidelberg, Germany.

Lam WHK, Tang YF, Chan KS, Tam M-L. 2006. Short-term hourly traffic Forecasts using Hong Kong Annual Traffic Census. *Transportation* **33**: 291-310.

Medeiros MC, Soares LJ. 2008. Modeling and forecasting short-term electricity load: A comparison of methods with an application to Brazilian data. *International Journal of Forecasting* **24**: 630-644.

Pedregal DJ, Young PC. 2008. Development of improved adaptive approaches to electricity demand forecasting. *Journal of the Operational Research Society* **59**: 1066-1076.

Shen H, Huang JZ. 2008. Interday forecasting and intraday updating of call center arrivals. *Manufacturing and Services Operations Management*, **10**: 391-410.

Smith DGC. 1989. Combination of forecasts in electricity demand prediction. *Journal of Forecasting* **8**: 349-356.

Taylor JW. 2003. Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of the Operational Research Society* **54**: 799-805.

Taylor JW. 2008. A comparison of time series methods for forecasting intraday arrivals at a call center. *Management Science* **54**: 253-265.

Taylor JW. 2009. Density forecasting of intraday call center arrivals using models based on exponential smoothing. Working paper, University of Oxford.

Taylor JW, de Menezes LM, McSharry PE. 2006. A comparison of univariate methods for forecasting electricity demand up to a day ahead. *International Journal of Forecasting* **22**: 1-16.

Tych W, Pedregal DJ, Young PC, Davies J. 2002. An unobserved component model for multi-rate forecasting of telephone call demand: The design of a forecasting support system. *International Journal of Forecasting* **18**: 673-695.

Weinberg J, Brown LD, Stroud JR. 2007. Bayesian forecasting of an inhomogeneous Poisson process with applications to call center data. *Journal of the American Statistical Association* **102**: 1185-1198.

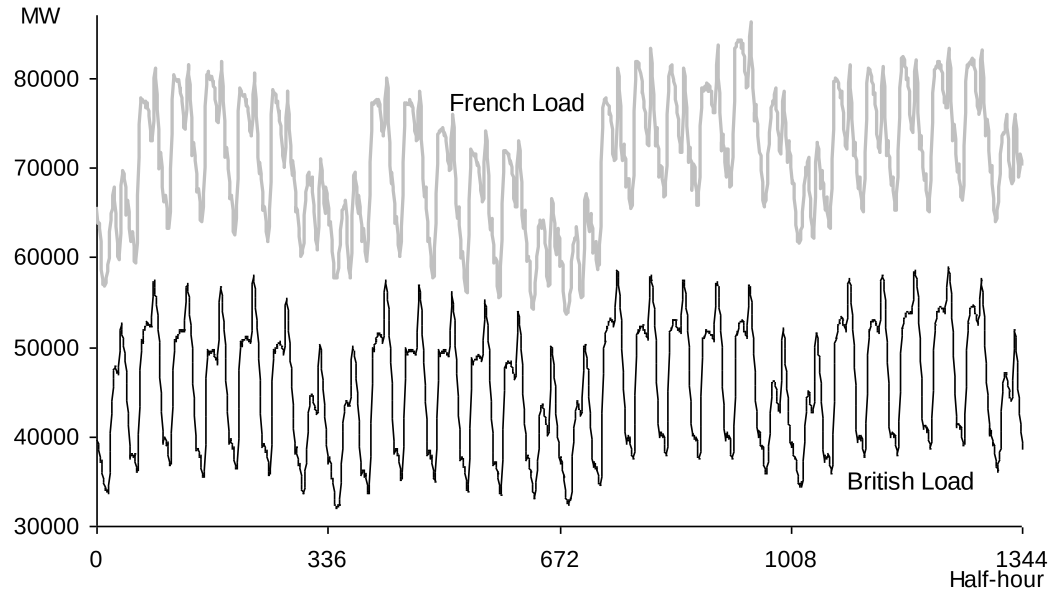

Figure 1. Half-hourly load in Great Britain and France for the four-week period from Sunday 8 January to Saturday 4 February 2006. Each week contains 7×48=336 periods.

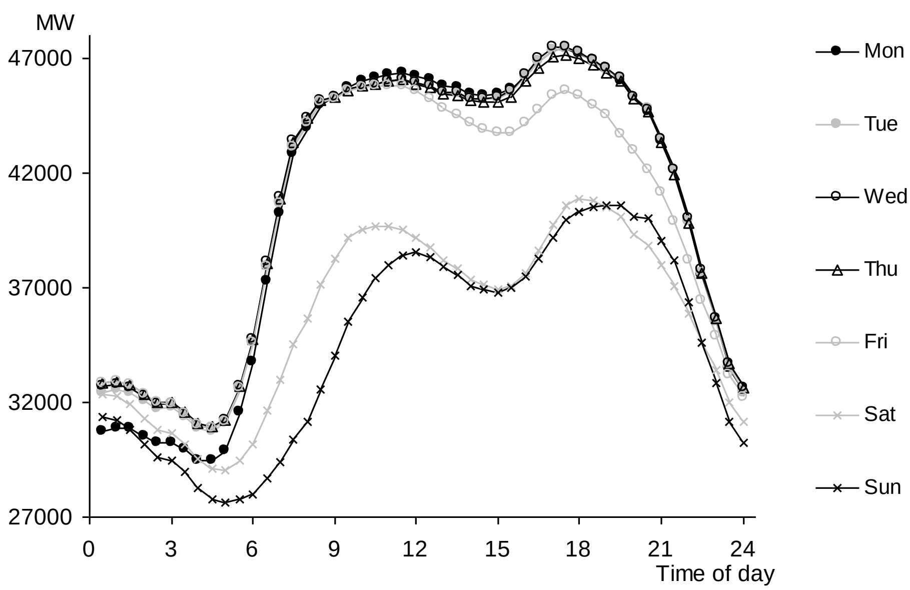

Figure 2. Average intraday cycle for each day of the week for the British load data. Calculated using the two-year estimation sample.

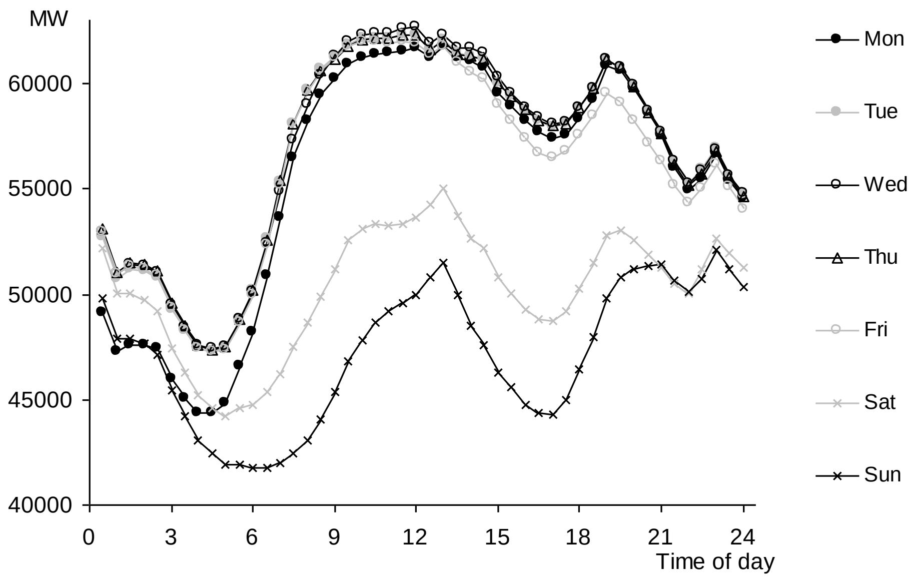

Figure 3. Average intraday cycle for each day of the week for the French load data. Calculated using the two-year estimation sample.

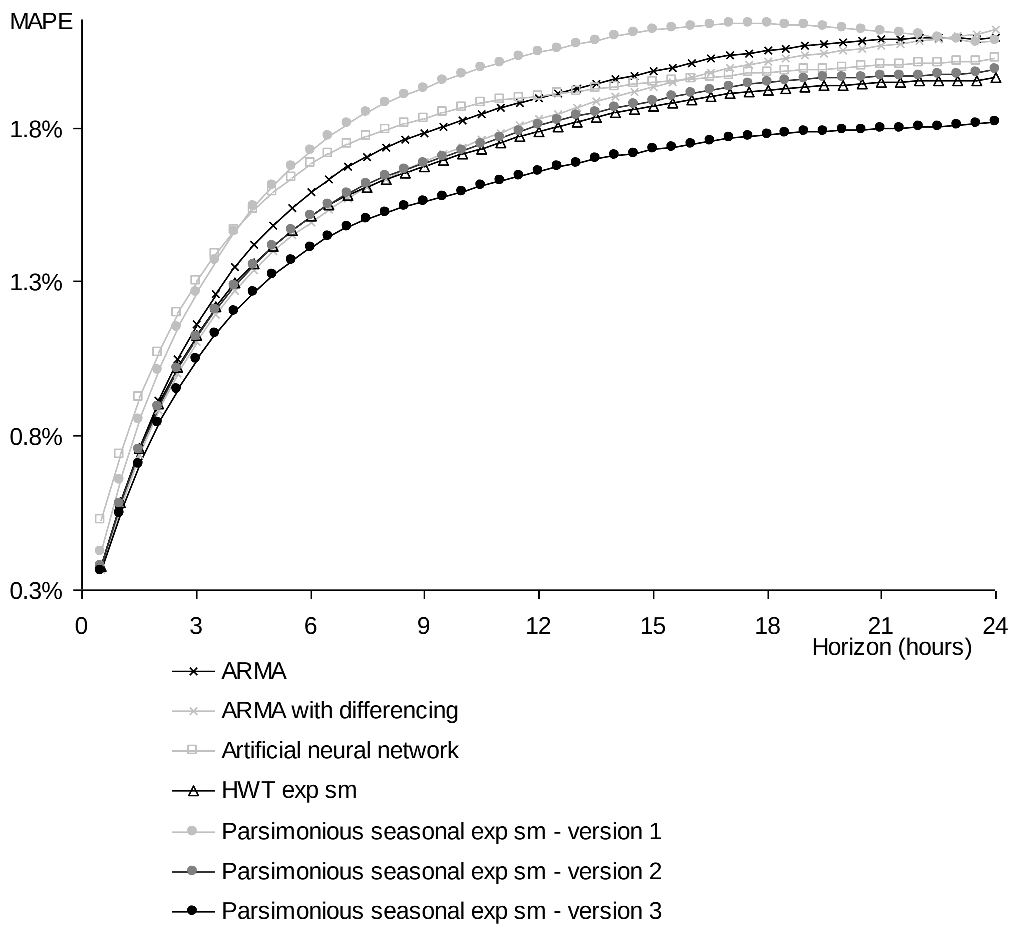

Figure 4. Average post-sample MAPE for the British and French load series.

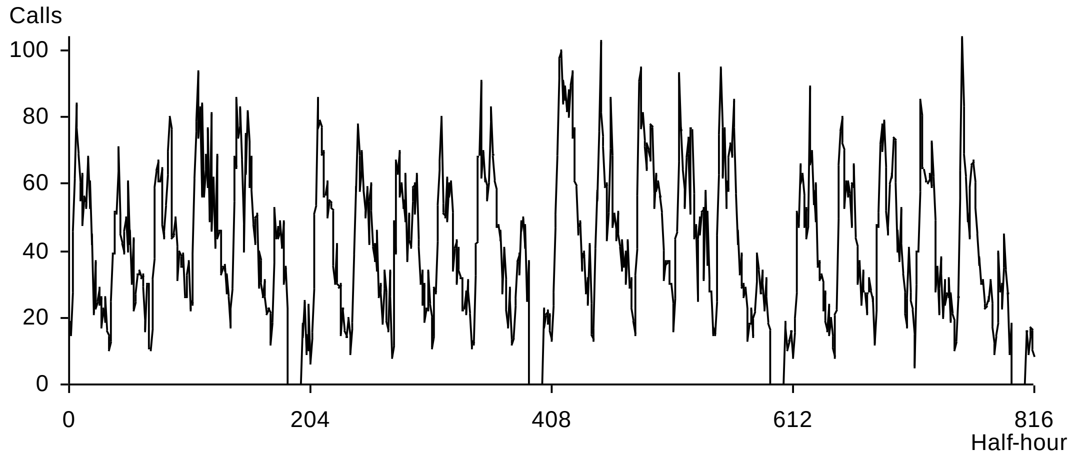

Figure 5. Half-hourly arrivals at the Israeli bank call centre for the four-week period from Sunday 28 November to Saturday 25 December 1999. 'Manipulated' arrivals series is plotted. In this series, each weekend is presented as one day, and each week consists of 6×34=204 half-hourly periods.

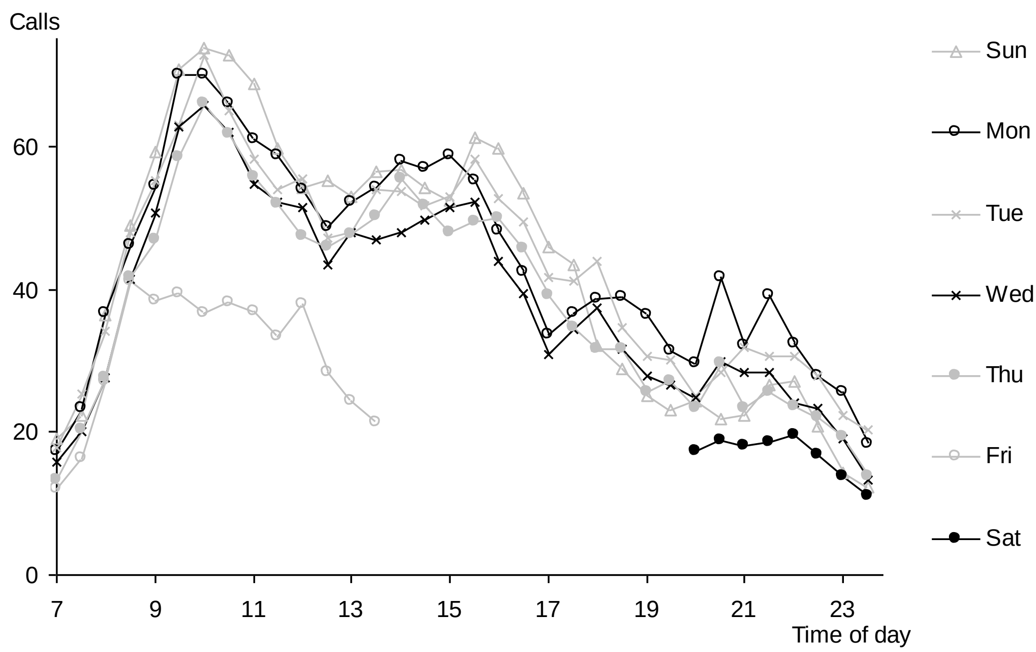

Figure 6. Average intraday cycle for each day of the week for the Israeli bank call centre data. Calculated using the 14-week estimation sample.

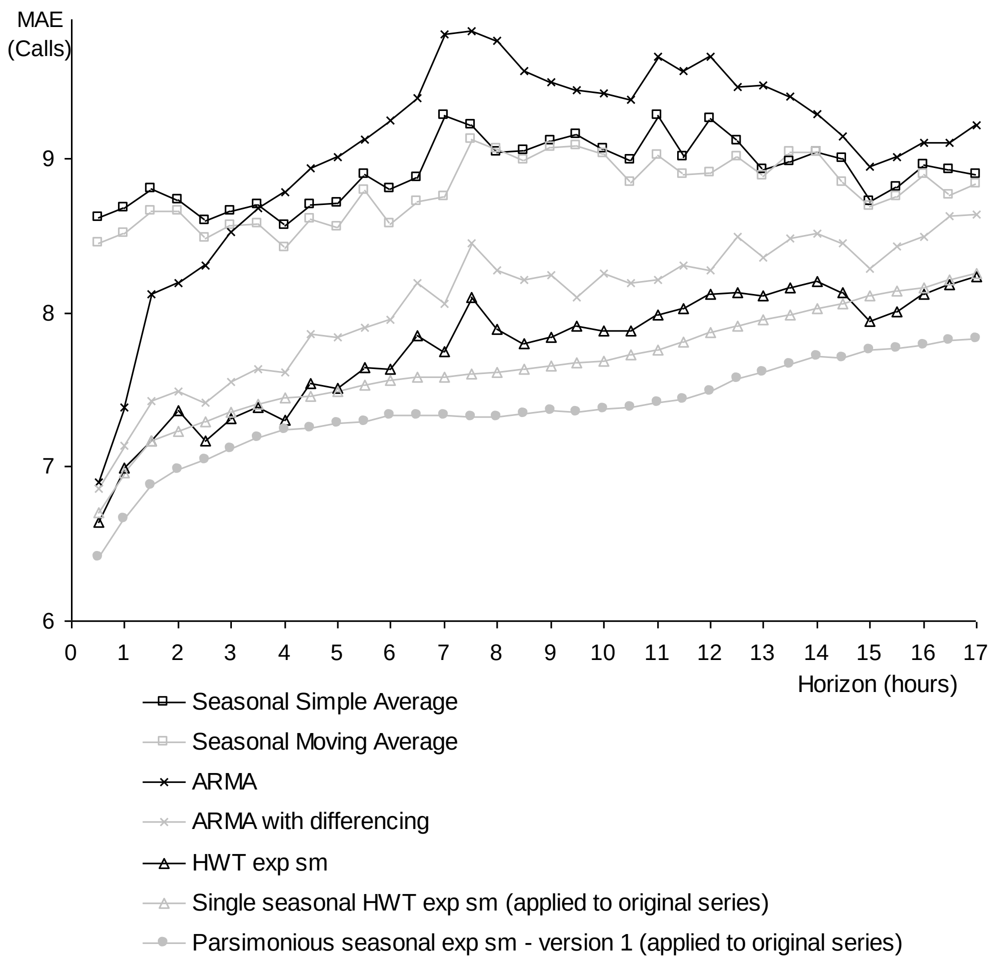

Figure 7. Post-sample MAE for the Israeli bank call centre series. Two methods were applied to the original arrivals series, and the rest to the 'manipulated' arrivals series.