Using Combined Forecasts with Changing Weights
for Electricity Demand Profiling

James W. Taylor

London Business School

and

Shanti Majithia

National Grid Company

Address for Correspondence:

James W. Taylor
London Business School
Sussex Place
Regent's Park
London   NW1 4SA

Tel: 0171 262 5050 ext 3246
Fax: 0171 724 7875
E-mail: jtaylor@lbs.ac.uk

Using Combined Forecasts with Changing Weights for Electricity Demand Profiling

Abstract

Day-ahead half-hourly demand forecasts are required for scheduling and for calculating the daily electricity pool price. One approach predicts turning points on the demand curve and then produces half-hourly forecasts by a heuristic procedure, called profiling, which is based on a past demand curve. This paper investigates possible profiling improvements. Using a cubic smoothing spline in the heuristic leads to a slight improvement. Often, several past curves could reasonably be used in the profiling method. Consequently, there are often several demand curve forecasts available. Switching and smooth transition forecast combination models are considered. These models enable the combining weights to vary across the 48 half-hours which is appealing as different forecasts may be more suitable for different periods. Several criteria are used to control the changing weights, including weather, and the methodology is extended to the case of more than two forecasts. Empirical analysis gives encouraging results.

Keywords: electricity demand forecasting; combining forecasts

Introduction

The role of electricity demand forecasting has developed with the changing structure of the electricity supply industry in England and Wales. The forecasts of the National Grid Company (NGC), which is responsible for the transmission of electricity, have always been a crucial input to operational planning where the generation output is scheduled to meet customer demand.[1] However, since the re-structuring of the industry in 1990, and the introduction of the daily electricity 'power pool', demand forecasts have also been used to set the price of electricity in the pool. The subsequent economic importance of the day-ahead forecasts has led to interest in NGC's demand predictions, from the industry regulator, generators and suppliers, who now view the predictions as a major influence on the business dynamics of the industry.[2,3] With the anticipated new structure of the industry, accurate short-term demand forecasting will also be required by utilities who will need to predict their customers' demand, and by those wishing to trade electricity on financial markets.

This paper considers the approach currently used by NGC to produce day-ahead forecasts. At 10:30am each day, forecasts must be produced for the total electricity demand for England and Wales at half-hourly intervals throughout the 24 hours of the next day. The approach taken is to first forecast the demand at the curve's 10 or 11 turning points, which are known as *cardinal points*. The half-hourly forecasts are then obtained by a procedure known as *profiling* which involves fitting a curve to the cardinal points. Baker[4] describes how this procedure evolved from the days when computing power limited cost optimisation to particular critical periods of the day, the cardinal points. As computer storage and processing capabilities developed, the approach was augmented with curve fitting to provide load predictions at periods between the cardinal points. The current approach to profiling fits a demand profile from a previous day, which is subjectively selected based on weather and calendar date considerations. There is a tendency for smaller errors to occur at the cardinal points than

between them. Hence, there is strong motivation to try to improve the profiling procedure. This is the aim of this study.

The use of a two-stage procedure which relies heavily on the heuristic approach of profiling may seem unappealing. However, there is no consensus as to the best approach to short-term electricity demand forecasting. Bunn[5] reviews a wide variety of methods and the recent competition organised by the Puget Power Company in Seattle witnessed a range of different approaches including: time-varying splines[6], artificial neural networks[7], multiple regression models[8], judgemental forecasts produced by Puget Power's own personnel, and Box-Jenkins transfer function intervention-noise models. NGC's cardinal point forecasts are produced by separate regression models which are functions of seasonal and weather variables.[4,9] This is similar to the approach of the overall winners of the Seattle competition, Ramanathan *et al.*[8], who produced hourly forecasts by using separate models for each hour of the day.

It has been suggested that the profiling methodology should allow more than one past curve to be used in profiling for the 48 half-hours of the next day. In view of this, there may be value in implementing a switching or combining mechanism within the profiling methodology. Ramanathan *et al.*[8] suggest the use of switching procedures as an extension of their demand forecasting approach. Switching models and the related theme of smooth transition combining models are investigated by Deutsch *et al.*[10] in the context of economic data. This paper adapts and develops the use of these models for load profiling. The next section describes the profiling procedure that is currently used. It is then shown how slight improvements can be achieved through the use of cubic splines. The two sections that follow consider switching and traditional combining models. The next four sections introduce smooth transition combining models which use the shape of the base curve and weather data to set the weights on different forecasts. Improvement in *ex ante* forecast accuracy is evaluated for the new models and, in the final section, a summary and concluding comments are provided.

The Profiling Heuristic

The current approach used by NGC for profiling proceeds by selecting a past demand curve which is likely to be similar to the demand profile for the next day. The success of the profiling procedure depends heavily on the choice of this *base curve*. Two common choices are the demand curve from the same weekday one year previously and one week previously, provided weather conditions are predicted to be similar. The base curve is fitted by first calculating the ratio of forecast to base curve demand for each cardinal point. These are termed the *scaling ratios*. In addition to the turning points, the first and last half-hours of the day and several strategically positioned points are also chosen as cardinal points. Scaling ratios for the half-hours between two cardinal points are then calculated by linear interpolation of the scaling ratios at the two points. Demand forecasts are then calculated from the product of the base curve's demand and the scaling ratio for each half-hour. The procedure is slightly complicated by the fact that it is common for a forecast curve to result that has cardinal points at different half-hours to the base curve. This would tend to lead to poor forecasts, and so when this occurs, the profiling procedure is repeated a sufficient number of times, with the most recent forecast curve as base curve, until the cardinal points occur at identical periods on the new base and forecast curves. Typically two or three profiling repetitions are required. Locating the cardinal points is thus rather *ad hoc* and is one of the least appealing aspects of the heuristic.

Evaluating half-hourly forecasts for a four week period, 19 May to 15 June 1997, shows slightly greater accuracy at the cardinal points. This supports the view of the company's forecasters, and motivates consideration of improvements to profiling. It is difficult to discern whether the forecast errors at the half-hours between cardinal points are due to profiling error or cardinal point forecast error. In order to study profiling improvements, in the next seven sections cardinal point forecasts are not used. Instead profiling is performed using actual demand at the cardinal points. In the penultimate section of the paper, profiling with cardinal point forecasts is reconsidered in order to evaluate whether the new proposals for

profiling improve *ex ante* forecasting performance. Figure 1 shows the result of profiling for 2 June 1997, using actual demand at the cardinal points, and the same weekday one year previously as base curve. The resultant forecast accuracy is impressive which is perhaps surprising, given the heuristic nature of the profiling procedure.

---------- **FIGURE 1** ----------

Using Cubic Splines Within the Profiling Heuristic

A cubic spline is a continuous piecewise cubic polynomial with continuous first and second derivatives.[11] Splines have been used in a variety of ways for electricity load forecasting.[6,12,13] Using cubic interpolating and smoothing splines to profile between the cardinal points leads to very poor forecasts. However, there is some potential for using cubic splines within the traditional profiling method. Figure 2 shows scaling ratios for 2 June 1997 using the same weekday one year previously as base curve. Fairly constant scaling ratios for all half-hours would indicate that the base curve and forecast curve are of similar shape. The change in the scaling ratio therefore indicates how the similarity changes. Figure 2 shows how a linear interpolation leads to a discontinuous change in the similarity at the cardinal points. This seems intuitively unappealing and so, instead, a cubic smoothing spline could be used. As the location of the cardinal points on the forecast curve often do not coincide with the location of those on the base curve, the spline profiling would be performed a second time using the forecast curve as the new base curve.

---------- **FIGURE 2** ----------

Table 1 compares traditional profiling with the new spline version, using actual demand instead of forecasts at the cardinal points and three different base curves: the same weekday of the previous week, previous year and two years previously. The new version slightly outperforms traditional profiling whilst maintaining its basic simplistic appeal. As in the study of Islam and Meade[14], this paper uses a conventional root mean square error

4

measure in order to be consistent with the objective functions used later for model estimation. The rest of this paper describes how switching and combining methods can be incorporated in the profiling method to achieve improved results. The new spline profiling method is used throughout, although all proposals are applicable to the traditional method.

---------- **TABLE 1** ----------

Switching Models

Rather than choosing one base curve to use in profiling, it often seems that different base curves should be used for different periods of the day. For example, the same weekday one year previously may seem preferable for the morning but the same weekday in the previous week may seem better for the afternoon. This motivates consideration of regime switching models (see Granger and Teräsvirta[15]). A simple switching model for two forecasts, $f_{t,1}$ and $f_{t,2}$, can be expressed as:

$$f_{t,S} = I(t \in I_1) f_{t,1} + (1 - I(t \in I_1)) f_{t,2} \tag{1}$$

where I_1 is the regime. $I(t0I_1) = 1$ if $t0I_1$ and $I(t0I_1) = 0$ if $t \lceil I_1$. The regime is chosen to reflect whether forecast $f_{t,1}$ is likely to be superior to $f_{t,2}$ for each half-hour t. Deutsch et al.[10] investigate these models in the context of economic forecasting using either lagged forecast errors or an economic variable as the regime. The implementation in this study is rather different as forecasts are required for all 48 half-hours of the next day whilst they consider the more standard time series forecasting problem of extrapolation to the next period.

The choice of regime is of crucial importance for the success of the method. The fundamental criterion that is used to judgementally select the base curve for profiling is the similarity of the shape of the base curve to the shape indicated by the cardinal point forecasts. A natural measure of the similarity, between the past base curve used in profiling and the shape mapped out by the cardinal points, is the change in the scaling ratios which result from profiling. Consequently, in this study the switching regime is defined such that the forecast is

5

chosen for a particular half-hour, if that forecast has associated scaling ratios that change less than those associated with the alternative forecast. The change in the scaling ratios at a particular half-hour t can be defined simply as:

$$\Delta\, ratio_t = \tfrac{1}{2}\left|ratio_{t+1} - ratio_t\right| + \tfrac{1}{2}\left|ratio_t - ratio_{t-1}\right| \tag{2}$$

Figure 3 shows the scaling ratios from the new spline profiling method for 2 June 1997 using the same weekday of the previous week and the same weekday one year previously as base curves. The change in the scaling ratios is less for the latter profiling for half-hours 6 to 43, so this forecast would be chosen for this period. In fact, the same weekday of the previous week was a bank holiday which explains why the shape is very inappropriate for profiling for 2 June 1997. Although Figure 3 is useful in illustrating the idea, a bank holiday would not normally be used as a base curve and so the two sets of scaling ratios are unlikely to differ by as much as those in the figure.

---------- **FIGURE 3** ----------

Traditional and Switching Combinations

If several forecasts are produced for the same period from different information sources, then a combination of the forecasts may be able to synthesise this information to deliver an improved prediction.[16] This is the main motivation for combining instead of selecting a single preferred prediction. There will be different information in two load curve forecasts produced by profiling which has used different past days as base curve. Combining is therefore a natural alternative to selecting one of the two load curve forecasts. This paper considers three popular traditional combining methods: the optimal method[17], regression[18] and the simple average. The optimal method is equivalent to regression with no constant term and the weights restricted to sum to one. The regression method is particularly useful if the forecasts are biased, although multicollinearity can be a problem. However, bias is unlikely to be an issue for the NGC forecasts as they carefully monitor their day-ahead predictions.

The weights are restricted to be nonnegative in the combining models, as advocated by Gunter[19].

In this paper, the empirical analysis uses a two week period for model estimation (19 May 1997 to 1 June 1997) and a two week period for post-sample evaluation (2 June 1997 to 15 June 1997). With half-hourly data, this implies estimation and evaluation data sets each of length 672 observations. Consider forecasts, $f_{t,1}$, produced by profiling using the same weekday one year previously and forecasts, $f_{t,2}$, produced using the same weekday of the previous week. Table 2 shows the post-sample performance for the various models introduced so far. The "Benchmark" model (Model 1) represents the new spline profiling method using, as base curves, the past demand curves which were actually chosen by forecasters in 1997. The forecasters were not restricted in their choice of base curve which is obviously an advantage over the combination methods. The forecasts, $f_{t,2}$, have a very poor RMSE because for one day in the evaluation period (Monday 2 June) the day used as base curve in profiling was a bank holiday (Monday 26 May). Although, in practice, a bank holiday would not be used as base curve in profiling for a normal weekday, it is useful in illustrating just how poorly profiling can perform given a poor choice of base profile. Table 2 shows that the optimal and regression combining methods (Methods 3 and 4) are preferable to a policy of using only $f_{t,1}$, or only $f_{t,2}$, throughout the fourteen day evaluation period, but that they are poorer than the benchmark forecast (Model 1).

Deutsch *et al*.[10] argue that, since combining is often useful, one could adapt the 'simple' switching model of expression (1) to switch between two forecast combinations.

$$f_{t,S} = I(t \in I_1)(a_1 + a_2 f_{t,1} + a_3 f_{t,2}) + (1 - I(t \in I_1))(a_4 + a_5 f_{t,1} + a_6 f_{t,2})$$

where the a_i are combining parameters. The results in Table 2 suggest that simple switching (Model 5) is no better than the traditional optimal and regression combining methods (Models 3 and 4), but that switching between optimal and regression combinations (Models 6 and 7) leads to improved performance.

7

---------- **TABLE 2** ----------

Smooth Transition Combining Models

Deutsch *et al.*[10] also propose the use of smooth transition combining models. In contrast to switching between two forecasts, these models enable a combination where the weights are changed gradually according to some criterion (see Granger and Teräsvirta[15]). A simple smooth transition model can be defined using a logistic function:

$$f_{t,S} = \frac{1}{(1 + e^{\alpha + \beta V_t})} f_{t,1} + \left(1 - \frac{1}{(1 + e^{\alpha + \beta V_t})}\right) f_{t,2} \tag{3}$$

where α and β are parameters, and V_t is a variable controlling the change in the weights. Consider for illustrative purposes $\beta > 0$. As V_t varies from a large negative value to a large positive value, the weight on $f_{t,1}$ varies from near 1 to near 0, and correspondingly the weight on $f_{t,2}$ varies from near 0 to near 1. The expression in (3) can be rewritten as

$$f_{t,S} = \frac{1}{(1 + e^{\alpha + \beta V_t})} f_{t,1} + \frac{1}{(1 + e^{-\alpha - \beta V_t})} f_{t,2}$$

This expression shows that if the parameters are derived with the two forecasts interchanged, the parameters will simply change their sign, with the result that the respective weights will remain the same. The combining method is therefore symmetrical as between the two forecasts being combined.

Deutsch *et al.*[10] also propose a more general smooth transition model where the transition is between two combined forecasts:

$$f_{t,S} = \frac{1}{(1 + e^{\alpha + \beta V_t})} (a_1 + a_2 f_{t,1} + a_3 f_{t,2}) + \left(1 - \frac{1}{(1 + e^{\alpha + \beta V_t})}\right) (a_4 + a_5 f_{t,1} + a_6 f_{t,2}) \tag{4}$$

Although Deutsch *et al.*[10] do not consider the parameter α, it is included here as it generally improves forecasting performance. The controlling variable V_t can be defined as:

$$V_t = \Delta ratio_{t,1} - \Delta ratio_{t,2} \tag{5}$$

where $\Delta ratio_{t,1}$ and $\Delta ratio_{t,2}$ are the changes in the scaling ratios at period t, for $f_{t,1}$ and $f_{t,2}$ respectively, as defined in expression (2). The intuition for this definition of V_t can be seen by considering again $\beta > 0$ in expression (3). If, at period t, the base curve used to produce $f_{t,2}$ is much more similar to the curve mapped out by the cardinal points than the base curve used to produce $f_{t,1}$, then $\Delta ratio_{t,2}$ will be much less than $\Delta ratio_{t,1}$, and so V_t will be large and positive. A large positive value of V_t implies a low weight on $f_{t,1}$ and a large weight on $f_{t,2}$ which seems sensible since the base curve used for $f_{t,2}$ was superior to that used for $f_{t,1}$.

Table 3 shows the results of fitting the simple transition model given in expression (3) (Model 8) and also combination models as in expression (4). "Transition with optimal combining" (Model 9) refers to models with no constant term and weights restricted to be nonnegative and to sum to 1, and "Transition with regression combining" (Model 10) refers to combinations with a constant term and weights only restricted to be nonnegative. As in the work of Deutsch *et al*. (1994), model parameters are estimated by minimising the sum of squared in–sample forecast errors. The results for the smooth transition combining models are an improvement on those for the switching models and also the traditional fixed weight combinations in Table 2, which suggests that allowing weights to change across the 48 half-hours leads to improved profiling.

---------- **TABLE 3** ----------

Expression (6) shows the simple smooth transition model (Model 8) with the estimated parameters. The general form of the model was presented in expression (3). Importantly, β, the coefficient of V_t, is of the correct sign; the positive value for β implies that a large positive value for V_t will result in a smaller weight on $f_{t,1}$. This is sensible because a large positive value for V_t indicates that the change in the scaling ratios of $f_{t,1}$ is much larger than of $f_{t,2}$. In fact, β was of the correct sign in all of the models considered in our work.

$$\text{Model 8:} \qquad f_{t,S} = \frac{1}{(1+e^{-0.69+53.12V_t})} f_{t,1} + \left(1 - \frac{1}{(1+e^{-0.69+53.12V_t})}\right) f_{t,2} \qquad (6)$$

It is interesting to investigate the combining weights that result from the transition models. Figure 4 shows the cardinal point forecasts for 2 June 1997, as well as the load curve from the same weekday of the previous week (a bank holiday). Figure 5 shows the same cardinal points with the load curve of the same weekday one year previously. These historic load curves can be used to produce profiled forecasts $f_{t,2}$ and $f_{t,1}$, respectively. Figures 4 and 5 show, plotted against the axis on the right, the combining weights produced by the simple smooth transition model in expressions (3) and (6) (Model 8). In Figure 4, for certain periods of the day, such as between half-hours 13 and 20, the shape of the curve mapped out by the cardinal points is radically different to the base profile. By contrast, in Figure 5, between half-hours 13 and 20, the base curve is of the right shape. The combining weights can be seen to vary accordingly.

---------- **FIGURE 4** ----------

---------- **FIGURE 5** ----------

Using Weather Variables in the Transition Models

Forecasters select a base curve for profiling by considering not only the shape of the curve, but also the weather forecasts for the next day. NGC use three weather variables in their various forecasting operations: temperature, illumination and cooling power of the wind.[4,9] Temperature is the most important variable in the cardinal point regression models, but illumination is often the most useful when choosing a base curve for profiling. Of particular concern is nightfall which, during the 4 week period considered in this study, is steadily moving later in the evening. There is thus strong motivation for using weather variables to control the smooth transition combining weights. Ramanathan *et al.*[8] propose the use of switching procedures as an extension to their load forecasting approach, and suggest that different forecasts could be selected depending upon weather considerations.

Deutsch *et al.*[10] use just one variable to control the smooth transition weights. However, it is straightforward to extend the methodology to use more than one variable. In addition to the variable V_t which relates to the shape of the base curves used to produce $f_{t,1}$ and $f_{t,2}$, a variable can be included which relates to the weather conditions on the days from which the base curves were taken. This leads to models of the following form:

$$f_{t,S} = \frac{1}{(1+e^{\alpha+\beta V_t+\gamma W_t})}(a_1+a_2 f_{t,1}+a_3 f_{t,2}) + \left(1 - \frac{1}{(1+e^{\alpha+\beta V_t+\gamma W_t})}\right)(a_4+a_5 f_{t,1}+a_6 f_{t,2})$$

where α, β and γ are parameters and V_t is defined in expression (5). W_t is a variable representing whether the day used as base curve in profiling for $f_{t,1}$ had weather more similar to the forecast for tomorrow than the day used as base curve to produce $f_{t,2}$. V_t and W_t are thus defined in order to influence the weighting of the two forecasts for each half-hour of the day.

A significant difficulty is that NGC only receive day-ahead weather forecasts for five different times in the following day: 9am, 12pm, 3pm, 5pm and 9pm. Given the limited weather forecast information, it seems wise to keep the definition of W_t simple. For example, one variable that gives reasonable results is defined as a dummy variable taking a value of 1 for half-hours 1 to 21, if and only if the day used as base curve in profiling for forecast $f_{t,1}$ had illumination at 9am more similar to tomorrow's forecast for 9am, than the day used as base curve to produce forecast $f_{t,2}$. In a similar way, the variable takes a value of 1 for half-hours 22 to 27, depending on the illumination information for 12pm, and a value of 1 for half-hours 28 to 32, depending on the illumination information at 3pm, etc. Table 4 reports the performance of the simple smooth transition model (Model 11) and two combining versions (Models 12 and 13) using W_t as just defined, and V_t. The results are an improvement on those for the transition models without weather given in Table 3.

---------- **TABLE 4** ----------

Expression (7) shows the simple smooth transition model (Model 11) with the estimated parameters. Importantly, the coefficients of V_t and W_t are of the correct sign. In the

previous section, we discussed how the parameter β should take a positive value. Reassuringly, this is the case in expression (7). The definition of the dummy variable W_t implies that if it takes a value of 1, the weight on $f_{t,1}$ should be greater than if W_t takes a value of 0. The coefficient, γ, should thus be negative. This is indeed the case, as can be seen from expression (7). In fact, γ was of the correct sign in all of the models considered in our work.

$$\text{Model 11:} \qquad f_{t,S} = \frac{1}{(1 + e^{0.05 + 82.33V_t - 0.96W_t})} f_{t,1} + \left(1 - \frac{1}{(1 + e^{0.05 + 82.33V_t - 0.96W_t})}\right) f_{t,2} \qquad (7)$$

The parameters in all of the smooth transition models should be re-estimated periodically, as the optimal values are likely to vary throughout the year. For example, when using the smooth transition model with regression combining (Models 13) for a four week Autumn period, the best results were achieved using two weather variables, temperature and illumination. Apart from this, the relative performance of the various models was broadly similar for this Autumn period to the results obtained for the Summer period presented in this paper.

A Flexible Combining Model

The combining models described so far combine forecasts $f_{t,1}$ and $f_{t,2}$ which are produced using the same weekday one year and one week previously. However, in practice, the choice of base curve varies greatly from day to day. The benchmark forecast (Model 1), included in Table 2, was constructed using the base curves that were actually judgementally selected by the company in 1997. For most of the evaluation period, the forecasters chose the same weekday of the previous year, however, for several days they selected the same weekday 51 weeks previously, and also a different weekday of the previous week. It seems likely that the combining models described so far in this paper would benefit by the use of subjective judgement in the choice of base curves used to produce the forecasts. In fact, if the parameter α is omitted from the simple smooth transition model in expression (3), then this

model can be used to combine forecasts produced using any base curve, regardless of the forecasts used to estimate the model parameters. This is because the weights will adjust to reflect the relative suitability of the base curves used. The simple smooth transition model with shape and weather variables can also be used, provided α is omitted.

As an illustration, consider the situation where parameters β and γ are estimated using the forecasts $f_{t,1}$ and $f_{t,2}$, defined earlier in the paper. For the evaluation period, $f_{t,1}$ can be replaced by forecasts constructed using the base curves that were judgementally selected in 1997. The results in Table 5 for these 'flexible' models (Models 14 and 15) are the best so far. The weather variable used in Model 15 is the same as that used in the models of Table 4.

---------- **TABLE 5** ----------

Using More than Two Forecasts

Deutsch *et al.*[10] investigate the combination of only two forecasts. Since forecasters often consider more than two past curves when profiling, it seems sensible to adapt the switching and combining models presented so far to the case of three forecasts. The switching models can be extended to three forecasts by using the same regime that was used for the two forecast case. The smooth transition models can be extended by using the multinomial logit function (see Greene[20] and Kamstra and Kennedy[21]). The model is then written as:

$$f_{t,S} = \left(\frac{1}{1+e^{\alpha_2+\beta_2 V_{t,2}} + e^{\alpha_3+\beta_3 V_{t,3}}}\right)f_{t,1} + \left(\frac{e^{\alpha_2+\beta_2 V_{t,2}}}{1+e^{\alpha_2+\beta_2 V_{t,2}} + e^{\alpha_3+\beta_3 V_{t,3}}}\right)f_{t,2} + \left(\frac{e^{\alpha_3+\beta_3 V_{t,3}}}{1+e^{\alpha_2+\beta_2 V_{t,2}} + e^{\alpha_3+\beta_3 V_{t,3}}}\right)f_{t,3}$$

where α_2, α_3, β_2 and β_3 are parameters, and $V_{t,2}$ and $V_{t,3}$ are variables controlling the change in the weights. $V_{t,2}$ must convey the quality of $f_{t,2}$ relative to $f_{t,1}$, and $V_{t,3}$ must convey the quality of $f_{t,3}$ relative to $f_{t,1}$. $V_{t,2}$ and $V_{t,3}$ can be conveniently defined as:

$$V_{t,2} = \Delta ratio_{t,1} - \Delta ratio_{t,2} \qquad\qquad V_{t,3} = \Delta ratio_{t,1} - \Delta ratio_{t,3}$$

where $\Delta ratio_{t,1}$, $\Delta ratio_{t,2}$ and $\Delta ratio_{t,3}$ are the changes in the scaling ratio at period t, for forecasts $f_{t,1}$, $f_{t,2}$ and $f_{t,3}$ respectively, as defined in expression (2).

The model can also include weather variables $W_{t,2}$ and $W_{t,3}$ defined in a similar way to the weather variable for the two forecast case in the previous section. $W_{t,2}$ conveys the quality of $f_{t,2}$ relative to $f_{t,1}$, and $W_{t,3}$ conveys the quality of $f_{t,3}$ relative to $f_{t,1}$. Better results are achieved by imposing the restrictions $\alpha_2 = \alpha_3$, $\beta_2 = \beta_3$ and $\gamma_2 = \gamma_3$. Using base curve shape and weather as criteria, the simple smooth transition model is then written as:

$$f_{t,S} = \left(\frac{1}{1 + e^{\alpha + \beta V_{t,2} + \gamma W_{t,2}} + e^{\alpha + \beta V_{t,3} + \gamma W_{t,3}}} \right) f_{t,1} + \left(\frac{e^{\alpha + \beta V_{t,2} + \gamma W_{t,2}}}{1 + e^{\alpha + \beta V_{t,2} + \gamma W_{t,2}} + e^{\alpha + \beta V_{t,3} + \gamma W_{t,3}}} \right) f_{t,2}$$
$$+ \left(\frac{e^{\alpha + \beta V_{t,3} + \gamma W_{t,3}}}{1 + e^{\alpha + \beta V_{t,2} + \gamma W_{t,2}} + e^{\alpha + \beta V_{t,3} + \gamma W_{t,3}}} \right) f_{t,3}$$

Table 6 reports post-sample results for combining and switching models involving three forecasts. $f_{t,1}$ and $f_{t,2}$ are the forecasts defined and used earlier, and $f_{t,3}$ is produced using the same weekday two years previously as base curve in profiling. The weather variable used in several of the models is the same as that used in the models of Table 4. The flexible smooth transition models (Models 14 and 15) use forecasts $f_{t,1}$, $f_{t,2}$ and $f_{t,3}$ to estimate model parameters; and then, as in the previous section, for the evaluation period, $f_{t,1}$ is replaced by forecasts constructed using the base curves that were judgementally selected in 1997. For almost all the models, the results are better than those achieved with two forecasts. As for the two forecast case, the transition models (Models 8 to 15) outperform the switching models (Models 5 to 7), and there appears to be some potential for the use of weather variables. All the smooth transition models comfortably outperform the benchmark forecasts (Model 1) which represents the performance achieved by the company using the traditional approach.

---------- **TABLE 6** ----------

Using Demand Forecasts at the Cardinal Points

So far in this paper, actual demand has been used at the cardinal points in order to focus solely on profiling accuracy. However, to establish the robustness of the results, it is important to repeat the analysis with demand forecasts at the cardinal points. With forecasts at the cardinal points, almost identical forecasting performance results from the traditional profiling approach and from the new spline version which applies a cubic smoothing spline to the scaling ratios. It seems that the improvement in profiling that was seen when profiling with actual demand at the cardinal points has been clouded by the error in the cardinal point forecasts. The results for switching and combination models involving two forecasts are given in Table 7. The general pattern of the results is similar to the case where actual demand was used at the cardinals. The smooth transition regression model with the shape and weather variables (Model 13) performs the best.

---------- **TABLE 7** ----------

---------- **TABLE 8** ----------

With actual demand used at the cardinal points, the models involving three forecasts were more accurate than those using two. However, as Table 8 testifies, when using forecasts at the cardinal points, the results for the three forecast case are poorer. It is interesting to see that the superiority of the smooth transition models (Models 8 to 15) over the switching models (Models 5 to 7) is repeated, and that, as with two forecasts, the smooth transition regression combining model with shape and weather variables (Model 13) performs the best.

We performed nonparametric sign tests on the difference between the magnitude of the NGC benchmark forecast errors and those from Models 13, 14 and 15. Under the null hypothesis that there is no difference, 50% of the signs would be positive and 50% negative. We found that for the models involving two forecasts, the magnitude of the forecast errors from Models 13 and 14 (in Table 7) were significantly lower than those of the NGC benchmark forecasts (at the 5% significance level). The magnitude of the errors from Model

15 were not significantly lower. For the models involving three forecasts, the magnitude of the errors of Model 13 (in Table 8) were significantly lower, whilst those of models 14 and 15 were not. We have reservations regarding the application of this test, as it assumes that each set of forecast errors has no autocorrelation. This seems unlikely for this day-ahead profiling application, since an overestimation at half-hour t will often be followed by an overestimation at half-hour $t+1$.

Summary and Conclusion

This paper has addressed the problem of profiling, which is essentially a heuristic used to fit a past demand curve to day-ahead forecasts for the 10 or 11 turning points of the next day's demand curve, in order to produce half-hourly forecasts. The heuristic was adapted to include a cubic smoothing spline. Although the resulting method is more intuitively appealing, *ex ante* forecasting performance remained the same. The study investigated the use of the switching and smooth transition combining models proposed by Deutsch *et al*.[10] These models were applied to the profiling heuristic, using the shape of the past curve and weather as criteria in selecting between two or three sets of half-hourly day-ahead forecasts. The models gave improved post-sample forecasting performance. The forecasts actually used in practice had a RMSE of 417 MW; the best of the new models had a RMSE of 384 MW. The relatively moderate size of the improvement indicates the power of the traditional human heuristic in choosing past demand curves in the profiling procedure. Nevertheless, any small improvement in day-ahead electricity demand forecasts has a significant impact. Electricity cannot be stored and so even small reductions in forecast error will lead to savings in scheduling and despatch commitment of generating plant.

Although, in this study, switching and smooth transition combining models have been used within the profiling heuristic, there is strong potential for their use in the more standard context where forecasts are produced using different methodologies. For example, they could

be used, as an alternative to the framework of Smith[22], to combine very short-term online electricity demand forecasts.

References

1   Majithia S, Kiernan L and Hannan J (1997). Intelligent systems for demand forecasting. In: Warwick K (ed). *Artificial Intelligence Techniques in Power Systems 1997*. Institute of Electrical Engineers: London, pp 259-279.

2   Trade and Industry Committee (1995). Aspects of the Electricity Supply Industry, House of Commons, Session 1994-95, Eleventh Report.

3   Littlechild S (1996) Operational Research and Regulation: Theory and Practice. *Journal of the Operational Research Society* **47**, 601-611.

4   Baker AB (1985). Load forecasting for scheduling generation on a large interconnected system. In: Bunn DW and Farmer ED (eds). *Comparative Models for Electrical Load Forecasting 1985*. Wiley: New York, pp 57-67.

5   Bunn DW (1982). Short-term forecasting: a review of procedures in the electricity supply industry. *Journal of the Operational Research Society* **33**: 533-545.

6   Harvey AC and Koopman SJ (1993). Forecasting Hourly Electricity Demand Using Time-Varying Splines. *Journal of the American Statistical Association* **88**: 1228-1236.

7   Connor JT (1996). A robust neural network filter for electricity demand prediction. *Journal of Forecasting* **15**: 437-458.

8   Ramanathan R, Engle R, Granger CWJ, Vahid-Araghi F and Brace C (1997). Short-run Forecasts of Electricity Loads and Peaks. *International Journal of Forecasting* **13**: 161-174.

9   Majithia S (1996). Electricity Demand Forecasting in the Operational Timescale. Paper presented at the 16th International Symposium on Forecasting in Istanbul, Turkey, June 23-26, 1996.

10   Deutsch M, Granger CWJ and Teräsvirta T (1994). The combination of forecasts using changing weights. *International Journal of Forecasting* **10**: 47-57.

11   Silverman BW (1985). Some Aspects of the Spline Smoothing Approach to Non-parametric Regression Curve Fitting. *Journal of the Royal Statistical Society*, *Series B* **47**: 1-52.

12   Hendricks W, Koenker R and Poirier DJ (1979). Residential Demand for Electricity: An Econometric Approach. *Journal of Econometrics* **9**: 33-57.

13   Pelikán E, Eben K and Petrák L (1996). One-day Prediction of Electricity Load Reflecting Future RCS Schedule. *Journal of Forecasting* **15**: 427-435.

14   Islam T and Meade N (1996). Forecasting the development of the market for business telephones in the UK. *Journal of the Operational Research Society* **47**: 906-918.

15   Granger CWJ and Teräsvirta T (1993). *Modelling Nonlinear Economic Relationships, Advanced Texts in Econometrics*. Oxford University Press: New York.

16   Bunn DW (1989). Forecasting with more than one model. *Journal of Forecasting* **8**: 161-166.

17   Bates JM and Granger CWJ (1969). The combination of forecasts. *Operations Research Quarterly* **20**: 451-468.

18   Granger CWJ and Ramanathan R (1984). Improved methods of combining forecasts. *Journal of Forecasting* **3**: 197-204.

19   Gunter SI (1992). Nonnegativity restricted least squares combinations. *International Journal of Forecasting* **8**: 45-59.

20   Greene WH (1997). *Econometric Analysis*, 3rd edn. Prentice-Hall: New Jersey.

21   Kamstra M and Kennedy P (1998). Combining Qualitative Forecasts Using Logit. *International Journal of Forecasting* **14**: 83-93.

22   Smith DGC (1989). Combination of Forecasts in Electricity Demand Prediction. *Journal of Forecasting* **8**: 349-356.

		Last year	Last week	Two years ago	Overall
Traditional profiling	RMSE	233	421	206	302
New spline profiling RMSE		223	407	205	293

Table 1: Comparison of forecast errors (measured in MW) for traditional and the new spline profiling. Actual demand used at the cardinal points.

		RMSE
$f_{t,1}$		208
$f_{t,2}$		421
Model 1:	Benchmark	164
Model 2:	Simple average	240
Model 3:	Fixed weight optimal combination	192
Model 4:	Fixed weight regression combination	192
Model 5:	Simple switching - shape	194
Model 6:	Switching with optimal combining – shape	184
Model 7:	Switching with regression combining - shape	182

Table 2: Post-sample evaluation of combining and switching models for two forecasts. Base curve shape used as switching criterion. Actual demand used at the cardinal points.

		RMSE
Model 8:	Simple smooth transition - shape	169
Model 9:	Transition with optimal combining - shape	169
Model 10:	Transition with regression combining - shape	176

Table 3: Post-sample evaluation for smooth transition combining models involving two forecasts using base curve shape as a criterion.

		RMSE
Model 11:	Simple smooth transition - shape & weather	162
Model 12:	Transition with optimal combining - shape & weather	165
Model 13:	Transition with regression combining - shape & weather	172

Table 4: Post-sample evaluation for smooth transition models involving two forecasts, using base curve shape & weather as criteria.

		RMSE
Model 14:	Flexible simple smooth transition - shape	152
Model 15:	Flexible simple smooth transition - shape & weather	152

Table 5: Post-sample evaluation for flexible smooth transition models. Actual demand used at the cardinal points.

	RMSE
$f_{t,1}$	208
$f_{t,2}$	421
$f_{t,3}$	197
Model 1: Benchmark	164
Model 2: Simple average	191
Model 3: Fixed weight optimal combination	166
Model 4: Fixed weight regression combination	166
Model 5: Simple switching - shape	196
Model 6: Switching with optimal combining - shape	161
Model 7: Switching with regression combining - shape	165
Model 8: Simple smooth transition - shape	152
Model 9: Transition with optimal combining - shape	157
Model 10: Transition with regression combining - shape	161
Model 11: Simple smooth transition - shape & weather	148
Model 12: Transition with optimal combining - shape & weather	151
Model 13: Transition with regression combining - shape & weather	150
Model 14: Flexible simple smooth transition - shape	146
Model 15: Flexible simple smooth transition - shape & weather	145

Table 6: Post-sample evaluation for smooth transition models involving three forecasts, using base curve shape alone or shape & weather as criterion.

	RMSE
$f_{t,1}$	413
$f_{t,2}$	596
Model 1: Benchmark	417
Model 2: Simple average	456
Model 3: Fixed weight optimal combination	412
Model 4: Fixed weight regression combination	426
Model 5: Simple switching - shape	415
Model 6: Switching with optimal combining - shape	419
Model 7: Switching with regression combining - shape	406
Model 8: Simple smooth transition - shape	393
Model 9: Transition with optimal combining - shape	393
Model 10: Transition with regression combining - shape	388
Model 11: Simple smooth transition - shape & weather	398
Model 12: Transition with optimal combining - shape & weather	398
Model 13: Transition with regression combining - shape & weather	384
Model 14: Flexible simple smooth transition - shape	396
Model 15: Flexible simple smooth transition - shape & weather	397

Table 7: Post-sample evaluation for models involving two forecasts. Demand forecasts used at the cardinal points.

	RMSE
$f_{t,1}$	413
$f_{t,2}$	596
$f_{t,3}$	434
Model 1: Benchmark	417
Model 2: Simple average	431
Model 3: Fixed weight optimal combination	406
Model 4: Fixed weight regression combination	406
Model 5: Simple switching - shape	415
Model 6: Switching with optimal combining - shape	419
Model 7: Switching with regression combining - shape	449
Model 8: Simple smooth transition - shape	405
Model 9: Transition with optimal combining - shape	425
Model 10: Transition with regression combining - shape	415
Model 11: Simple smooth transition - shape & weather	404
Model 12: Transition with optimal combining - shape & weather	424
Model 13: Transition with regression combining - shape & weather	397
Model 14: Flexible simple smooth transition - shape	401
Model 15: Flexible simple smooth transition - shape & weather	407

Table 8: Post-sample evaluation for models involving three
forecasts. Demand forecasts used at the cardinal points.

Figure 1: Actual demand and predicted demand for 2 June 1997, using
the same weekday one year previously as base curve in profiling. Actual
demand used as forecasts at the cardinal points.

Figure 2: Linear interpolation of scaling ratios from the traditional NGC profiling method for 2 June 1997 using 3 June 1996 as base curve. Actual demand used as forecasts at the cardinal points.

Figure 3: Scaling ratios for 2 June 1997 resulting from two different base curves in the new spline profiling procedure. Actual demand used at the cardinal points.

Figure 4: Smooth transition combining weights for forecasts produced
for 2 June 1997 using 26 May 1997 (a bank holiday) as base curve.
Actual demand used at the cardinal points.

Figure 5: Smooth transition combining weights for forecasts
produced for 2 June 1997 using 3 June 1996 as base curve.
Actual demand used at the cardinal points.