Density Forecasting for the Efficient Balancing of the Generation and Consumption of Electricity

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

International Journal of Forecasting, 2006, Vol. 22, pp. 707-724.

Density Forecasting for the Efficient Balancing of the Generation and Consumption of Electricity

Abstract

The transmitters of electricity in Great Britain are responsible for balancing generation and consumption. Although this can be done in the hour between closure of the market and real-time, off-loading or calling-up electricity at this late stage can be costly. Costs can be substantially reduced if the imbalance can be anticipated ahead of time and balanced by trading on the market. Efficient trading relies on accurate density forecasts for the Net Imbalance Volume, which is defined as the sum of all actions taken to balance the system. Forecasting this density is the focus of this paper. We break down the problem into point and volatility prediction. We evaluate density forecasts in terms of the economic benefit generated from trading advice resulting from the forecasts. Promising results were achieved using a seasonal ARMA model or periodic AR model for point forecasting, with a simplistic approach to volatility forecasting.

Key words: electricity markets; volatility forecasting; density forecasting; seasonality; periodic models; economic value.

1. Introduction

In the wholesale electricity market of Great Britain, National Grid is the company that operates the transmission system, and balances supply and demand. The market involves suppliers undertaking contracts with generators to meet their anticipated requirements. An hour ahead of real time (known as “gate closure”), the market closes and market participants indicate what electricity they intend to produce or consume to meet their contractual obligations. In the hour between gate closure and real-time, National Grid must ensure that the demand is balanced by an equal amount of generation. To do this, they take “bids” and “offers” from generators and suppliers, who indicate the prices at which they would be prepared to change their intended generated output or consumption.

The energy imbalance is known as Net Imbalance Volume (NIV). It is defined for each half-hour as the sum of all actions, in the market and after gate closure, that National Grid undertook to balance the system for that half-hour. The convention is that a negative value for NIV means that contracted generation exceeded consumption, forcing the company to undertake sales of electricity or accept bids from generators to reduce their generation. These actions, by the company, are based on the one hour-ahead prediction of NIV.

Accepting bids or offers, after gate closure, can be very costly for the company. These costs can be substantially reduced if NIV can be anticipated ahead of gate closure and balanced by trading on the market. Indeed, the costs tend to decrease as the lead time increases. Efficient trading relies on accurate forecasts for the probability density function of NIV. This is the focus of this paper. We break down the problem of density forecasting into point and volatility prediction. By making a distributional assumption, the point and volatility forecasts can then be converted into a density forecast for NIV. It is not only National Grid that forecast NIV, but also generators seeking to gain an advantage by being able to anticipate the trades of the company. The problem of forecasting NIV is common to any self-despatching electricity market, of which there are several in Europe and the US.

In Section 2, we describe the NIV time series. Sections 3 and 4 evaluate the accuracy of point and volatility forecasting methods, respectively, using error summary measures. In Section 5, we evaluate the quality of density forecasting methods in terms of economic benefit by deriving the monetary outcome that is generated from trading advice resulting from the density forecasts. The final section provides a summary and concluding comments.

2. Net Imbalance Volume

In this paper, we analyse half-hourly NIV observations for the one-year period from 11 March 2003 to 10 March 2004, inclusive. This series is plotted in Figure 1. We chose not to use data prior to 11 March 2003 because a change was introduced to the electricity trading rules on this day, and the forecasters at National Grid felt it was quite possible that this would have led to a change in the structure of the NIV time series. We considered omitting observations for the days immediately following the potential structural change, but we did not do this because informal inspection of the series around this period did not actually reveal any clear evidence of structural change in the series.

---------- Figure 1 ----------

We used the first nine months of observations (273 days or 13,104 observations) to estimate model parameters, and the remaining three months (93 days or 4,464 observations) for post-sample forecast evaluation. The use of nine months of data for model estimation was recommended by National Grid from previous experience. In our study, we focused on the two forecast origins, and associated lead times, for which the National Grid trading support team are required to supply the company's traders with NIV forecasts. Forecasts from a 3pm origin are required each day for the 48 half-hours from 11:30pm on the next day through to 11pm on the day after. The lead times from this origin are, therefore, from 65 to 112 steps ahead. Forecasts from an 8am origin are required each day for the 48 half-hours from

11:30pm on the same day through to 11pm on the next day, which corresponds to a shorter set of lead times, running from 31 to 78 steps ahead.

The company had considered the use of explanatory variables in the modelling of an earlier period of NIV data. Weather variables were considered but they were found to have no significant impact on NIV. A dummy variable had been used for bank holidays, and daily peak demand forecasts for a range of lead times were also considered. For our more recent data set, we were unable to find evidence of the usefulness of these explanatory variables. Therefore, in this paper, we focus solely on univariate methods.

Figure 2 presents the ACF calculated from the nine-month estimation sample. The confidence interval in the figure is calculated using as standard error the inverse of the square root of the number of observations used to estimate the ACF values. In addition to significant autocorrelation at the early lags, there are strong spikes at the lags that are integer multiples of 48, indicating an intra-day seasonality. Significance at the lags that are integer multiples of 336 would indicate an intra-week seasonality. The very dominant intra-day seasonality makes it difficult to see an intra-week seasonal effect in the ACF. However, in our modelling of the series in Section 3, we show that intra-week seasonality is present in the series.

---------- Figure 2 ----------

Using the nine-month estimation sample, we investigated whether the series was stationary in the mean, based on the following augmented Dickey-Fuller test:

$$\Delta NIV_t = \alpha + \gamma\, NIV_{t-1} + \sum_{i=1}^{336} \beta_i \Delta NIV_{t-i} + u_t$$

where α, γ and β_i are constant parameters and u_t is an error term. We selected 336 lagged difference terms in order to allow for autocorrelation due to both seasonal cycles. We included a deterministic trend term, t, but removed it after finding that it was not significant (p-value = 0.462). The t-statistic for the parameter γ was -4.660, which is comfortably below the 1% critical value of -3.434 leading us to conclude that the series is stationary in the mean.

3. Point Forecasting

In this section, we compare the accuracy of point forecasts from 10 methods, the first three of which are relatively simplistic benchmark approaches.

3.1. Simple Benchmark Methods

Method P1 - Moving average - The forecast for all lead times is the average of all values of NIV, from the nine months immediately prior to and including the forecast origin.

Method P2 - Moving average: time of day - The forecast for each lead time is the average of the values for the same half-hour of the day from the nine months immediately prior to and including the forecast origin.

Method P3 - Moving average: time of week - The forecast for each lead time is the average of the values for the same half-hour of the week from the nine months immediately prior to and including the forecast origin.

3.2. Method P4 - Exponential smoothing

We implemented the double seasonal version of Holt-Winters exponential smoothing, which was developed by Taylor (2003) in order to accommodate the two seasonal cycles in a half-hourly electricity demand series. As the method involves no model specification, it has the appeal of simplicity and robustness. From a theoretical perspective, exponential smoothing methods can be considered to have a sound basis as they have been shown to be equivalent to a class of state space models (see Hyndman *et al*., 2002). Applying Taylor's method to the NIV series, the exponential smoothing formulation is:

$$S_t = \alpha (NIV_t - D_{t-48} - W_{t-336}) + (1-\alpha)(S_{t-1} + T_{t-1})$$
$$T_t = \gamma (S_t - S_{t-1}) + (1-\gamma) T_{t-1}$$
$$D_t = \delta (NIV_t - S_t - W_{t-336}) + (1-\delta) D_{t-48}$$
$$W_t = \omega (NIV_t - S_t - D_{t-48}) + (1-\omega) W_{t-336}$$
$$\hat{NIV}_t(k) = S_t + k T_t + D_{t-48+k} + W_{t-336+k} + \phi^k \left(NIV_t - \left(S_{t-1} + T_{t-1} + D_{t-48} + W_{t-336} \right) \right)$$

where $\hat{NIV}_t(k)$ is the *k* step-ahead forecast made from forecast origin *t*; α, γ, δ and ω are the smoothing parameters; S_t and T_t, are the smoothed level and trend; and D_t and W_t are the seasonal indices for the intra-day and intra-week seasonal patterns. The term involving the parameter ϕ, in the forecast function, is an adjustment for first-order autocorrelation. For simplicity, we have presented the forecast function for $k \leq 48$, but it is straightforward to rewrite the expression for longer lead times.

The initial smoothed values for the level, trend and seasonal components are estimated by averaging the early observations. All the parameters are estimated in a single procedure by minimising the sum of squared 1-step-ahead in-sample errors. We derived the following values: $\alpha = 0.007$, $\gamma = 0.000$, $\delta = 0.203$, $\omega = 0.119$ and $\phi = 0.884$. The low value of α and high value of ϕ reflects the fact that the adjustment for first-order autocorrelation has, to a large degree, made redundant the smoothing equation for the level. The value of zero for γ was accompanied by very small values for the smoothed trend, T_t. This seems reasonable given that in Section 2 we concluded that the series was stationary in the mean.

3.3. ARMA Modelling

Method P5 - Company's simple autoregression

For an earlier sample of NIV data than ours, the company investigated a variety of forecasting approaches, including nonlinear methods, such as an artificial neural network. The method that produced the most accurate post-sample results was an LS regression of NIV_t on the following two lagged variables: the most recent observation corresponding to the same half-hour period and the corresponding half-hour observation in the previous week. Our implementation of this approach involved estimating three LS regression models. To forecast up to 48 periods ahead, we used NIV_{t-48} and NIV_{t-336} as regressors; for lead times of 49 to 96, we used NIV_{t-96} and NIV_{t-336}; and, for lead times 97 to 144, we used NIV_{t-144} and NIV_{t-336}.

Method P6 - SARMA

Multiplicative seasonal ARMA modelling has often been used for univariate forecasting of intraday load time series (e.g. Laing and Smith, 1987; Darbellay and Slama, 2000; Taylor *et al*., 2006). The multiplicative double seasonal ARMA model (see Box *et al*., 1994, p. 333) applied to the NIV series can be written as

$$\phi_p(L)\,\Phi_{P_1}\left(L^{48}\right)\Omega_{P_2}\left(L^{336}\right)\left(NIV_t - c\right) = \theta_q(L)\,\Theta_{Q_1}\left(L^{48}\right)\Psi_{Q_2}\left(L^{336}\right)\varepsilon_t$$

where c is a constant term; L is the lag operator; ε_t is a white noise error term; and 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. This model can be expressed as $\text{ARMA}(p,q)\times(P_1,Q_1)_{48}\times(P_2,Q_2)_{336}$.

We followed the Box-Jenkins methodology to identify the most suitable SARMA model based on the estimation sample. For the NIV data, differencing was not necessary. We considered lag polynomials up to order three, and based model selection on the Schwartz Bayesian Criterion (SBC), with the requirement that all parameters were significant (at the 5% level). Our final model was the following ARMA(2,2)×(3,3)$_{48}$×(3,3)$_{336}$ model:

$$\begin{aligned}\left(1-\phi_1 L-\phi_2 L^2\right)\left(1-\phi_{96}L^{96}-\phi_{144}L^{144}\right)\left(1-\phi_{336}L^{336}-\phi_{672}L^{672}-\phi_{1008}L^{1008}\right)\left(NIV_t-c\right)\\ =\left(1-\theta_1 L-\theta_2 L^2\right)\left(1-\theta_{48}L^{48}-\theta_{96}L^{96}-\theta_{144}L^{144}\right)\left(1-\theta_{336}L^{336}-\theta_{672}L^{672}-\theta_{1008}L^{1008}\right)\varepsilon_t\end{aligned} \quad (1)$$

The model parameters were estimated using maximum likelihood based on a Gaussian distribution. The resultant parameters and associated standard errors are presented in Table 1. Box *et al*. (chapter 5, 1994) describe how to compute forecasts from such models. Comparison of the residual ACF in Figure 3 with the ACF in Figure 2 (note the change of y-axis scale) for the original NIV series shows that the model successfully accommodated much of the autocorrelation in the NIV series. However, we were unable to eliminate it all, with the Ljung-Box Q-statistic indicating significant autocorrelation (at the 5% level) at lag 22 and higher. A similar problem existed for all other SARMA models that we considered.

---------- Figures 3 and 4, and Table 1 ----------

We also inspected the residuals for heteroskedasticity. Figure 4 presents the ACF of the squared residuals. The squared residuals serve as a proxy for the residual variance because it is unobservable. The figure shows significant and strong autocorrelation, indicating conditional heteroskedasticity. The relatively large spikes occur every 48 periods, indicating the presence of intra-day seasonality in the residual variance. Intra-week seasonality may also be present but it is difficult to see because of the sizeable intra-day effect. The presence of a degree of autocorrelation and clear heteroskedasticity in the residuals from the SARMA model implies that the model's estimation was inefficient and parameter testing was not strictly valid. However, the model seems likely to be an improvement on the company's simple autoregression, Method P5. Therefore, we decided to include it as Method P6 in our analysis. In the next section, we describe Method P7, which involves the modelling of the heteroskedasticity in order to improve the SARMA model.

Method P7 - SARMA-SGARCH

Generalized autoregressive conditional heteroskedasticity (GARCH) models (see Engle, 1982; Bollerslev, 1986) are widely used to forecast volatility in finance. These models express the conditional variance as a linear function of lagged squared error terms and lagged conditional variance terms. A natural seasonal extension of the GARCH model for the intra-day and intra-week seasonality in the residual variance is to include seasonal lag terms. Estimating an ARMA model jointly with an appropriate GARCH model is considered to be more efficient than estimating the ARMA model with an incorrect assumption of homoskedastic errors. In view of this, we included in our comparison of point forecasting methods a SARMA model with seasonal GARCH model for the variance. We considered the same seasonal lags as we had done for the SARMA model. We estimated model parameters using maximum likelihood under the assumption of a Student-*t* distribution, which was proposed by Bollerlev (1987). We based model selection on the SBC, with the requirement

that all parameters were significant (at the 5% level). The formulation of our preferred model is presented in expressions (2) and (3).

$$\begin{aligned}&\left(1-\phi_1 L-\phi_3 L^3\right)\left(1-\phi_{96}L^{96}-\phi_{144}L^{144}\right)\left(1-\phi_{336}L^{336}-\phi_{672}L^{672}-\phi_{1008}L^{1008}\right)\left(NIV_t-c\right)\\&\quad=\left(1-\theta_1 L-\theta_3 L^3\right)\left(1-\theta_{48}L^{48}-\theta_{96}L^{96}-\theta_{144}L^{144}\right)\left(1-\theta_{336}L^{336}-\theta_{672}L^{672}-\theta_{1008}L^{1008}\right)\varepsilon_t\end{aligned} \tag{2}$$

where ε_t/σ_t has a Student-*t* distribution and

$$\begin{aligned}&\left(1-\beta_{48}L^{48}-\beta_{144}L^{144}-\beta_{672}L^{672}\right)\sigma_t^2\\&\quad=\alpha_0+\left(\alpha_1 L+\alpha_2 L^2+\alpha_3 L^3+\alpha_{48}L^{48}+\alpha_{144}L^{144}+\alpha_{336}L^{336}+\alpha_{672}L^{672}\right)\varepsilon_t^2\end{aligned} \tag{3}$$

If the model is able to capture all of the conditional heteroskedasticity in the residuals, there will be no significant autocorrelation evident in the ACF of the squared standardised residuals (residuals divided by the estimated values for σ_t), which is presented in Figure 5. Although significant autocorrelation remains, comparison of the figure with Figure 4 indicates that the SGARCH model has accommodated much of the heteroskedasticity in the residuals. We considered other SARMA-SGARCH models, but we were unable to eliminate all autocorrelation from the squared standardised residuals.

---------- Figure 5 ----------

3.4. Periodic AR

A periodic ARMA model is one in which the parameters change with the seasons (see Franses and Paap, 2004). Empirical evidence for economic data has shown that seasonality is often not satisfactorily modelled by standard time-invariant coefficient models and that periodic models can be more successful (e.g. Osborn et al., 1988; Franses and Romijn, 1993). However, although periodic models can improve explanatory power, the evidence on their usefulness for forecasting is less clear (Ghysels and Osborn, Section 6.6, 2001).

Applications have tended to show that periodic AR models tend to be sufficient, and that MA terms are unnecessary (Franses and Paap, p. 28, 2004). In view of this, we considered only periodic AR models for the NIV series. We found that, in addition to an AR

term of lag 1, the strong seasonality in the data necessitated the inclusion of an AR term of lag 48 and, possibly, also an AR term of lag 336. We considered the possibility of periodicity in the constant term and in all three AR term coefficients.

Figures 6, 7 and 8 show how autocorrelation at lags 1, 48 and 336 varies, dependent upon the half-hour of the day, for the estimation sample. An informal assessment of these figures suggests that periodicity in autocorrelation at lags 48 and 336 is more significant than at lag 1. We also examined how the autocorrelation at these three lags varied across the 336 half-hours of each week. However, with only 39 weeks in our estimation sample, the resulting figures provided inconclusive evidence of such intra-week periodicity.

---------- Figures 6 to 8 ----------

A common approach to modelling the time-varying coefficients in periodic AR models is to specify a separate parameter for each season. This seems reasonable for quarterly or monthly data, but, as the half-hourly NIV data has seasonal cycles consisting of 48 and 336 seasons, this approach implies a highly parameterised model. Faced with a similar problem in their analysis of the volatility in high frequency intra-day financial returns, Andersen and Bollerslev (1997, 1998) and Martens *et al*. (2002) use the flexible Fourier form proposed by Gallant (1981). We used this approach in our periodic AR models for NIV, which we present in the next three sections. The parameters in these models were estimated by OLS regression.

Method P8 - Periodic AR: time of day and week

Despite there being no clear evidence of intra-week periodicity, we felt it would be interesting to consider periodicity with period lengths of both 48 and 336 half-hours. We estimated the following periodic model:

$$NIV_t = \phi_0(t) + \phi_1(t)\, NIV_{t-1} + \phi_{48}(t)\, NIV_{t-48} + \phi_{336}(t)\, NIV_{t-336} + \varepsilon_t$$

where

$$\phi_p(t) = \omega_p + \sum_{i=1}^{4} \left(\lambda_{pi} \sin\left(2i\pi \tfrac{d(t)}{48}\right) + \nu_{pi} \cos\left(2i\pi \tfrac{d(t)}{48}\right) + \kappa_{pi} \sin\left(2i\pi \tfrac{w(t)}{336}\right) + \upsilon_{pi} \cos\left(2i\pi \tfrac{w(t)}{336}\right)\right)$$

$d(t)$ and $w(t)$ are repeating step functions that number the half-hours from 1 to 48 within each day, and from 1 to 336 within each week, respectively. The choice of summing from i=1 to 4 was made arbitrarily.

Method P9 - Periodic AR: time of day

Given the lack of evidence for intra-week periodicity, we implemented the following model that aims to capture only the intra-day periodicity:

$$NIV_t = \phi_0(t) + \phi_1(t)\, NIV_{t-1} + \phi_{48}(t)\, NIV_{t-48} + \phi_{336}(t)\, NIV_{t-336} + \varepsilon_t$$

where $$\phi_p(t) = \omega_p + \sum_{i=1}^{4}\left(\lambda_{pi} \sin\left(2i\pi \tfrac{d(t)}{48}\right) + \nu_{pi} \cos\left(2i\pi \tfrac{d(t)}{48}\right)\right)$$

Method P10 - Periodic AR: time of day with no AR(1)

In Section 2, we described how the forecast lead times of interest in this paper are from 30 to 112 periods ahead. This raises the importance of the lag 48 and lag 336 AR terms, and rather questions the benefit of the lag 1 AR term. Indeed, the company's simple autoregression approach, Method P5, makes no attempt to model short-lag autocorrelation. In terms of in-sample fit and quality of residuals, the exclusion of the lag 1 term clearly led to a poorer model. Nevertheless, we elected to implement the periodic model in expression (4), which has no lag 1 AR term. The parameters for this model are presented in Table 2.

$$NIV_t = \phi_0(t) + \phi_{48}(t)\, NIV_{t-48} + \phi_{336}(t)\, NIV_{t-336} + \varepsilon_t \qquad (4)$$

where $$\phi_p(t) = \omega_p + \sum_{i=1}^{4}\left(\lambda_{pi} \sin\left(2i\pi \tfrac{d(t)}{48}\right) + \nu_{pi} \cos\left(2i\pi \tfrac{d(t)}{48}\right)\right)$$

---------- Table 2 ----------

3.5. Point Forecasting Results

We calculated the mean absolute error (MAE) and root mean squared error (RMSE) for the post-sample forecast errors from the 10 methods for each of the 48 forecast horizons from

each of the two forecast origins. In Table 3, we present the MAE results for the 3pm forecast origin. For conciseness, rather than present the MAE value for each lead time, we show the MAE averaged across groups of eight lead times. Such averages can be dominated by those lead times for which prediction is particularly difficult. In view of this, as a summary of the relative performances of the methods across all lead times, in the final column of Table 3, we show the mean value of a Theil-U measure calculated for each lead time as the ratio of the MAE for that method to the MAE for the SARMA method, Method P6. In Table 4, we provide the MAE results for the 8am forecast origin. We do not present the RMSE results, as the relative performances of the methods for this measure were very similar to that for the MAE.

---------- Tables 3 and 4 , and Figures 9 and 10 ----------

Reassuringly, the results show a tendency for the more sophisticated methods, Methods P4 to P10, to outperform the simple benchmark methods, Methods P1 to P3. A clear exception to this is the poor performance of Method P8, which is the periodic AR model with intra-day and intra-week periodicity in the coefficients of autoregressive terms with lags 1, 48 and 336. Perhaps not surprisingly, given the lack of clear evidence of intra-week periodicity in the autocorrelations, the results for Method P9 show that the accuracy of the periodic AR approach improves when just the intra-day periodicity is accommodated. More surprising, to us, was the noticeable improvement in the approach when the AR term of lag 1 was removed from the model to produce Method P10. In terms of the mean Theil measure, this method and the SARMA model, Method P6, were the best performing of all the 10 methods. The accuracy of the SARMA model was not improved when the model was estimated with a GARCH model for the conditional variance, Method P7. This finding and the relative performance of the three periodic AR models leads us to conclude that, for the lead times considered in our study, the quality of in-sample diagnostics did not relate well to post-sample forecast accuracy.

To provide a little more insight into the variability within the forecast errors, in Figures 9 and 10, we present box-plots, corresponding to the two different forecast origins, for the AE results from the SARMA model of Method P6. The two plots are very similar, with the values being generally slightly larger for the earlier forecast origin.

4. Volatility Forecasting

In this section, we evaluate methods for forecasting the volatility in the NIV series. More specifically, we compare the accuracy of methods for forecasting, for the various lead times, the variance of the forecast error from the SARMA model of Method P6 in Section 3. We focus on a single method from Section 3 in order to provide a fair and simple comparison of the variance forecasting methods. Estimating the variance is non-trivial because it is not constant and it contains seasonality. We can deduce this from our assessment, in Section 3.3, of the ACF of the squared residuals from the SARMA model. We consider six methods; the first three of which are simple benchmark approaches.

4.1. Simple Benchmark Methods Applied to Historical Forecast Errors

Method V1 - Historical variance - This approach assumes the variance is constant across half-hours of the week. For each lead time, *k*, we calculated the variance of all *k* step-ahead in-sample forecast errors made for all half-hours in the nine month period immediately prior to the forecast origin being considered. The resulting variances were used as variance forecasts for both the 3pm and 8am forecast origins.

Method V2 - Historical variance: time of day - This approach assumes that the variance is constant for any given half-hour of the day, but that it can be different for different half-hours of the day. For the 3pm forecast origin, the *k* step-ahead variance forecast was calculated as the variance of the errors resulting from the *k* step-ahead forecasts, made from the 3pm origin, on all

days in the nine month period immediately prior to the forecast origin being considered. In an analogous way, variance forecasts were constructed for the 8am forecast origin.

Method V3 - Historical variance: time of week - This approach assumes that the variance is constant for any given half-hour of the week, but that it can be different for different half-hours of the week. For the 3pm forecast origin on a particular day of the week, the *k* step-ahead variance forecast was calculated as the variance of the errors resulting from the *k* step-ahead forecast, made from the 3pm origin, on all similar days of the week in the nine month period immediately prior to the forecast origin being considered. Variance forecasts were constructed similarly for the 8am forecast origin.

4.2. Methods Based on the Theoretical SARMA Error Variance Formula

The three methods in this section are based on the theoretical formulae for the standard error of the SARMA point forecast at the various lead times (see chapter 5, Box *et al*., 1994). These formulae are functions of the variance of the SARMA model error term, ε_t, at all lead times up to, and including, the lead time of interest. The three methods in this section present alternative approaches to forecasting the variance of ε_t in these periods.

Method V4 - SARMA constant error variance

This approach assumes the variance is constant. We estimated the variance of ε_t in all future periods as the variance of the SARMA estimation sample residuals.

Method V5 - SGARCH

We fitted an SGARCH model to the SARMA model residuals. We considered the same seasonal lags as we had done for the SGARCH model of Method P7 in Section 3.3, and. estimated parameters using maximum likelihood based on the Student-*t* distribution. Indeed, all GARCH models in this paper were estimated in this way. We based model selection on

the SBC, with the requirement that all parameters were significant (at the 5% level). Expression (5) presents our selected model, and Table 5 provides the estimated parameters.

$$\left(1-\beta_{48}L^{48}-\beta_{144}L^{144}-\beta_{672}L^{672}\right)\sigma_t^2 = \alpha_0+\left(\alpha_1 L+\alpha_2 L^2+\alpha_{48}L^{48}+\alpha_{144}L^{144}+\alpha_{336}L^{336}+\alpha_{672}L^{672}\right)\varepsilon_t^2 \quad (5)$$

---------- Tables 5 and 6 ----------

Method V6 - Periodic GARCH

In the GARCH literature, the tendency has been to accommodate seasonality through the use of periodic GARCH terms, which were first considered by Bollerslev and Ghysels (1996). Franses and Paap (2000) employ a periodic GARCH model to capture the day-of-the-week effect in daily stock index return volatility. Martens *et al*. (2002) use a periodic GARCH model to capture the intra-day and intra-week seasonality in half-hourly exchange rate returns. They used a flexible Fourier form for the periodically time-varying parameters.

We fitted a periodic GARCH model to the residuals of the SARMA model from Method P6. We used a time-invariant coefficient for lag variance terms, which has been the approach used in all the periodic GARCH models that we have seen. Given the results of our comparison of point forecasting periodic AR models in Section 3, we implemented a model that aimed to capture only the intra-day periodicity in the coefficients and we did not include a lag 1 term. The resultant periodic GARCH model is presented in expression (6) with parameters and standard errors given in Table 6.

$$\sigma_t^2 = \alpha_0(t)+\alpha_{48}(t)\,\varepsilon_{t-48}^2+\alpha_{336}(t)\,\varepsilon_{t-336}^2+\beta_{48}\,\sigma_{t-48}^2+\beta_{336}\,\sigma_{t-336}^2 \quad (6)$$

where $$\alpha_p(t)=\omega_p+\sum_{i=1}^{4}\left(\lambda_{pi}\sin\left(2i\pi\tfrac{d(t)}{48}\right)+\nu_{pi}\cos\left(2i\pi\tfrac{d(t)}{48}\right)\right)$$

4.3. Volatility Forecasting Results

We calculated the MAE and RMSE for the post-sample variance estimation error from the six methods for each of the 48 forecast horizons from each of the two forecast origins. We used the square of the post-sample SARMA model forecast error as a proxy for variance, so that the MAE was calculated, for each forecast horizon, *k*, as follows:

$$\text{MAE} = \frac{1}{m} \sum_{i} \left| e_{i,k}^2 - \hat{\sigma}_{i,k}^2 \right|$$

where $e_{i,k}^2$ is the square of the *k* step-ahead SARMA model forecast error made from day *i*; $\hat{\sigma}_{i,k}^2$ is the corresponding variance forecast; and *m* is the number of days in the post-sample period. In Tables 7 and 8, we present the MAE results for the 3pm and 8am forecast origins, respectively. The mean Theil-U measure in the final column was calculated for each method as the mean of the ratios of the MAE for that method to the MAE for Method V6, periodic GARCH. Tables 7 and 8 show that, overall, the best performing method was periodic GARCH. Its superiority over seasonal GARCH is consistent with its predominant use in the financial volatility literature. By contrast with the MAE results, the RMSE results in Tables 9 and 10 indicate that the simpler methods are preferable. The impressive performance of the simpler methods may be due to the unstable nature of the volatility in NIV. Simpler methods have the appeal of robustness, which is particularly important for unstable series.

---------- Tables 7 to 10 ----------

5. Evaluating Economic Value of NIV Density Forecasts

At National Grid, a NIV density forecast is required in order to provide trading advice. In Section 5.2, we describe nine density forecasting methods that are based on the point and volatility forecasting methods presented in Sections 3 and 4, respectively. The resulting density forecasts could be evaluated using statistical procedures (see, for example, Diebold *et al*., 1998, and Christoffersen, 1998). However, this would not address whether the

difference in accuracy between forecasting approaches had any economic significance. For the majority of forecasting applications, evaluating the economic benefit of an improvement in forecast accuracy is not straightforward. However, in the case of NIV, if we make some reasonable simplifying assumptions, we can derive the economic benefit by evaluating the monetary outcome that results from trading advice provided by the various density forecasting methods. We discuss this in more detail in the next section.

5.1. Trading Advice and the Resulting Cost Benefit

The net balancing cost that results from trading an amount of electricity in period *t* in order to balance the impact of NIV in period *t+k* is presented in expression (7).

$$Cost(t,k,Trade) = Trade \times MktPrice_t(k) + \left(NIV_{t+k} - Trade\right) \times \begin{cases} BidPrice_{t+k} & if \quad NIV_{t+k} < Trade \\ OfferPrice_{t+k} & if \quad NIV_{t+k} > Trade \end{cases} \qquad (7)$$

where *Trade* is the amount of electricity traded; $MktPrice_t(k)$, is the price at which the market is trading electricity in period *t* for exchange in period *t+k* (i.e. it is the forward price); and NIV_{t+k} is the actual value of NIV in period *t+k*; $BidPrice_{t+k}$ is the price at which National Grid can balance an excess of generation, for period *t+k*, after gate closure; and $OfferPrice_{t+k}$ is the price at which National Grid can balance an excess of demand, for period *t+k*, after gate closure. The cost benefit due to the trade is given by expression (8):

$$Benefit(t,k,Trade) = Cost(t,k,0) - Cost(t,k,Trade) \qquad (8)$$

As the industry regulator does not permit National Grid to trade speculatively, the optimal trade is the amount that minimises the expectation in period *t* of the net balancing cost in period *t+k*. Differentiating the expectation of expression (7) with respect to *Trade* delivers the rule that the optimal trade is the value that satisfies the following expression:

$$MktPrice_t(k) = \hat{F}_{t,t+k}(Trade) \times BidPrice_{t+k} + \left(1 - \hat{F}_{t,t+k}(Trade)\right) \times OfferPrice_{t+k} \qquad (9)$$

where $\hat{F}_{t,t+k}(Trade)$ is the forecast at period *t* of the cdf for NIV_{t+k}, evaluated at the traded quantity, *Trade*. (In other words, $\hat{F}_{t,t+k}(Trade)$ is the estimate in period *t* of the probability of NIV_{t+k} being less than the traded quantity.) Rearranging expression (9), it then follows that the optimal trade is given by:

$$\hat{F}^{-1}_{t,t+k}\left(\frac{OfferPrice_{t+k} - MktPrice_t(k)}{OfferPrice_{t+k} - BidPrice_{t+k}}\right) \tag{10}$$

We set realistic caps on the amount that is traded, so that, if the market price is greater than the offer price, or the market price is lower than the bid price, these caps are enforced. The trading advice and cost benefit formulae rely on the following simplifying assumptions: the trades do not affect market price; liquidity is never a problem; and we can forecast the bid and offer prices perfectly. Despite these simplifications, we feel the essence of the use of the NIV density forecast is captured in the trading advice of expression (10).

The quality of the trading advice in expression (10) relies on the quality of the NIV density forecasts. In Section 5.3, we show the results of substituting into the cost benefit formulae, in expression (8), the trading advice produced by nine different density forecasting approaches, which we now present in Section 5.2.

5.2. Density Forecasting Methods

Methods D1 to D6 - Methods based on SARMA point forecasts

We generated density forecasts by using the SARMA point forecasts with the variance forecasts from the six methods considered in Section 4. In order to be consistent with the distribution type used in their estimation, for the two GARCH approaches, we used Student-*t* distributions with degrees of freedom set to the values optimised in the maximum likelihood estimation. For the other four methods, we used a Gaussian distribution.

Method D7 - Simplistic benchmark

This approach was based on a Gaussian distribution, centred at the point forecasts from Method P2, with variance estimated using Method V2. Methods P2 and V2 are simple benchmark approaches that performed reasonably in Sections 3 and 4, respectively.

Method D8 - Company benchmark

We constructed a density forecast based on the company's simple autoregression point forecasting method, Method P5. A Gaussian distribution was used with variance forecasts created using the simplest method from Section 4, Method V1. We felt that this was a reasonable representation of the company's past approach to density estimation.

Method D9 - SARMA-SGARCH

We produced density forecasts using the SARMA-SGARCH model of expressions (2) and (3), in Section 3.3, for which the SARMA and SGARCH components had been estimated in a single stage.

5.3. Cost Benefit Results

Tables 11 and 12 show the cost benefit resulting from trading based on the nine density forecasting methods. The cost benefit is calculated using expression (8) with optimal trade calculated from expression (10). The final columns in the tables present the sum of the cost benefits for each method. This value represents the total benefit resulting from National Grid trading once for each half-hour period of the three-month evaluation period.

The tables show that Methods D1 to D6, which are based on the SARMA model point forecasts from Method P6, generally led to greater cost benefit than the other three methods. The best performing method in both tables is Method D2, which uses the SARMA point forecasts with forecast error variance estimated as the variance of in-sample errors with the same lead time and from the same time of day. Interestingly, the results are not competitive for Method D6, the method based on periodic GARCH variance forecasts, which according to the MAE measure was the most accurate variance forecasting approach in Section 4. In that section, RMSE highlighted the simpler variance forecasting methods as performing the best. This suggests that, for this application, RMSE is a better indicator of which variance forecasting method will deliver density forecasts with the greatest economic benefit.

The final column in Table 11 shows that, for the predictions made from the 3pm forecast origin, Method D2 led to a benefit of £104,000 more than achieved by the company benchmark approach, Method D8. As this is for just a three-month period, it implies an improvement in the benefit of £416,000 over a year. For the predictions made from the 8am forecast origin, Table 12 indicates a difference in the benefit of £205,000, between these two methods, for the 3-month period, which translates into an improvement in trading benefit of £820,000 over a year.

---------- Tables 11 and 12 ----------

6. Summary and Concluding Comments

This paper has investigated methods for NIV density forecasting. We decomposed the problem down into point forecasting and volatility forecasting. In our evaluation of point forecasting methods, the two best performing methods were a multiplicative seasonal ARMA model and a periodic AR model. Comparing volatility forecasting methods, we found that MAE and RMSE measures delivered a contradictory ranking of methods. We evaluated density forecasting in terms of economic benefit by deriving the monetary outcome that is

generated from trading advice resulting from a variety of density forecasting methods. The best results were achieved by a method that constructed density forecasts from seasonal ARMA point forecasts with forecast error variance estimated using a simplistic approach that accounted for the intra-day seasonality in the variance. The economic benefit in using this approach, rather than one based on the company's point forecasting method, was calculated as being of the order of £416,000 for the earlier forecast origin and £820,000 for the later origin.

There are a number of potential areas for further study. Combining the forecasts produced by different methods has been found to be useful in a wide variety of applications, so it would be interesting to consider combinations of point, variance or, perhaps, density forecasts. Another possibility is to combine the trading advice resulting from different density forecasts. A different form of synthesis is prompted by the result that the simpler volatility forecasting methods were more successful than the GARCH models. There may be scope for filtering off the seasonal components of the heteroskedasticity, using a relatively simple approach, and then applying a statistical model to capture the localised structure in the data. If several years of data were available, the modelling of an annual seasonal cycle could be considered. In this paper, we have used a Gaussian or Student-*t* distribution to construct density forecasts from point and volatility predictions. It may be beneficial to try to use, instead, some form of empirical distribution of historical values. It would also be interesting to consider the more direct modelling of the inverse of the cdf, given in expression (10), because this is ultimately what is needed in order to compute the trading advice. As the inverse of the cdf is the quantile function, this implies modelling the quantiles of NIV.

Acknowledgements

We are very grateful to Matthew Roberts, Andrew Ryan and Shanti Majithia of National Grid for data and information regarding NIV and the models previously considered at the company. We are also grateful for the helpful comments of three anonymous referees.

References

Andersen, T.G. & Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets, *Journal of Empirical Finance*, 4, 115-158.

Andersen, T.G. & Bollerslev, T. (1998). DM-dollar volatility: Intraday activity patterns, macroeconomic announcements and longer run dependencies, *Journal of Finance*, 53, 219-265.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31, 307-327.

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *Review of Economics and Statistics*, 69, 542-547.

Bollerslev T. & Ghysels, E. (1996). Periodic autoregressive conditional heteroskedasticity, *Journal of Business and Economic Statistics*, 14, 139-151.

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

Christoffersen, P.F. (1998). Evaluating interval forecasts, *International Economic Review*, 39, 841-862.

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

Diebold, F.X., Gunther, T.A., & Tay, A.S. (1998). Evaluating density forecasts with applications to financial risk management, *International Economic Review*, 39, 863-883.

Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50, 987-1008.

Franses, P.H. & Paap, R. (2000). Modelling day-of-the-week seasonality in the S&P 500 index. *Applied Financial Economics*, 10, 483-488.

Franses, P.H. & Paap, R. (2004). *Periodic Time Series Models*, Oxford: Oxford University Press.

Franses, P.H., & Romijn, G. (1993). Periodic integration in quarterly UK macroeconomic variables, *International Journal of Forecasting*, 9, 467-476.

Gallant, A.R. (1981). On the bias in flexible functional forms and an essentially unbiased form: The Fourier flexible form, *Journal of Econometrics*, 15, 211-245.

Ghysels, E. & Osborn, D.R. (2001). *The Econometric Analysis of Seasonal Time Series*, Cambridge: Cambridge University Press.

Hyndman, R.J., Koehler, A.B., Snyder, R.D. & Grose, S. (2002). A state space framework for automatic forecasting using exponential smoothing methods, *International Journal of Forecasting*, 18, 439-454.

Laing, W.D. & Smith D.G.C. (1987). A comparison of time series forecasting methods for predicting the CEGB demand. Proceedings of the Ninth Power Systems Computation Conference.

Martens, M., Chang, Y.-C. & Taylor, S.J. (2002). A comparison of seasonal adjustment methods when forecasting intraday volatility, *The Journal of Financial Research*, 25, 283-299.

Osborn, D.R., Heravi, S., & Birchenhall, C.R. (1988). Seasonality and the order of integration for consumption, *Oxford Bulletin of Economics and Statistics*, 50, 361-377.

Taylor, J.W. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of Operational Research Society*, 54, 799-805.

Taylor, J.W., M. de Menezes, L., & McSharry, P.E. (2006). A Comparison of Univariate Methods for Forecasting Electricity Demand Up to a Day Ahead, *International Journal of Forecasting*, forthcoming.

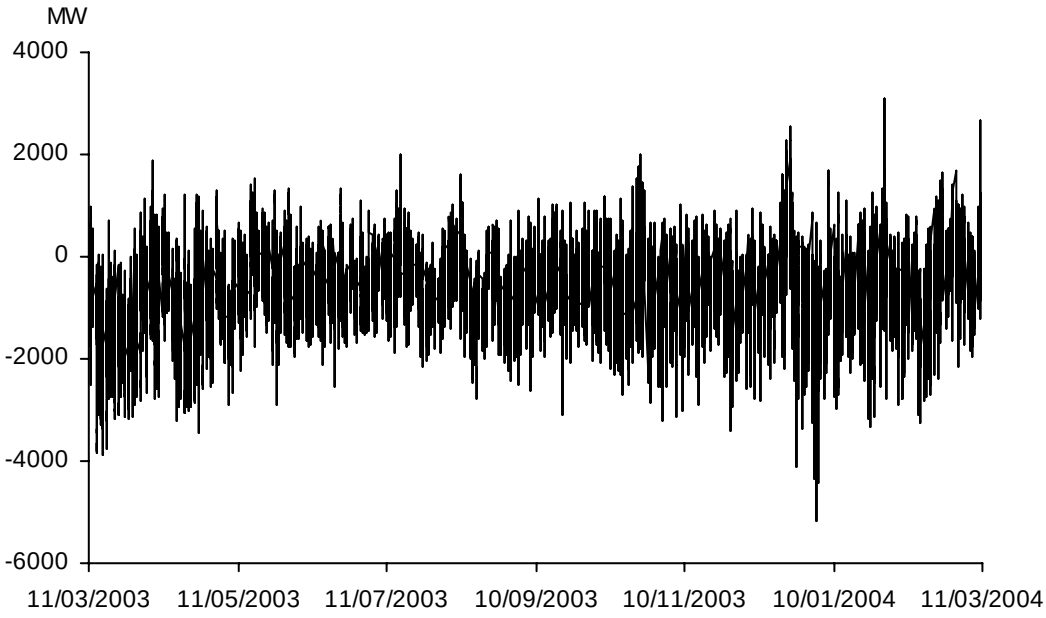

Figure 1. Half-hourly Net Imbalance Volume
for the period 11 March 2003 to 10 March 2004.

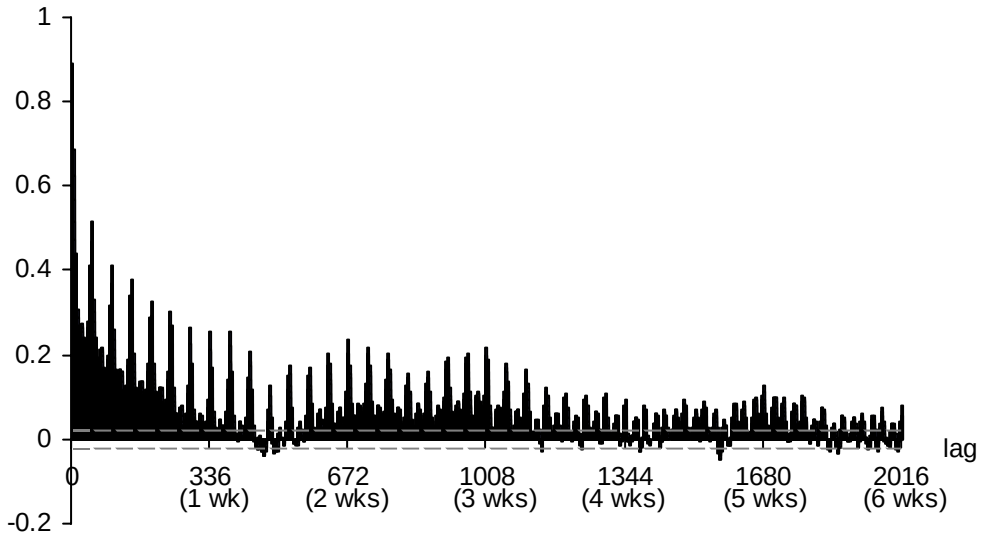

Figure 2. Autocorrelation function for the nine-month estimation sample
(13,104 observations). Dotted lines indicate 95% confidence interval.

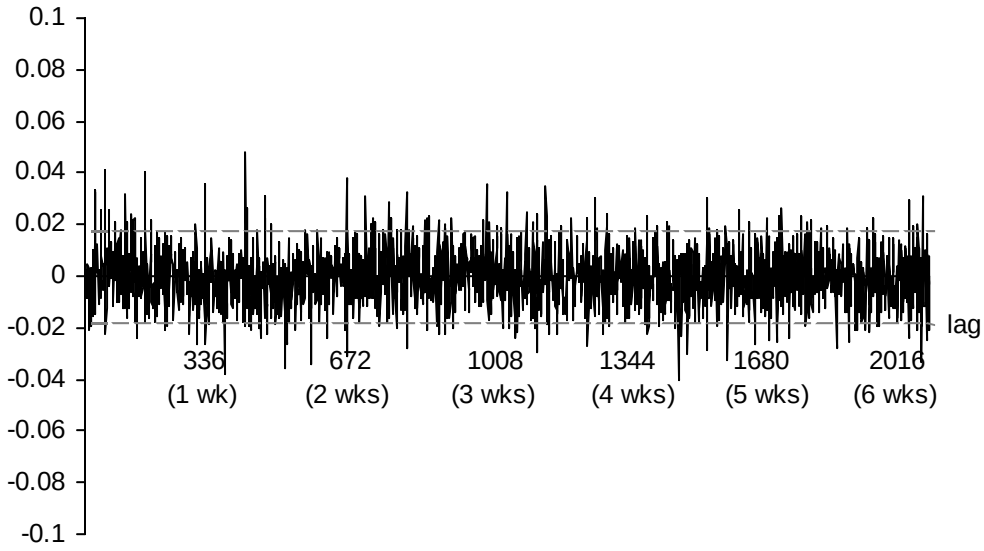

Figure 3. Autocorrelation function for the residuals from the SARMA model, Method P6. Dotted lines indicate 95% confidence interval.

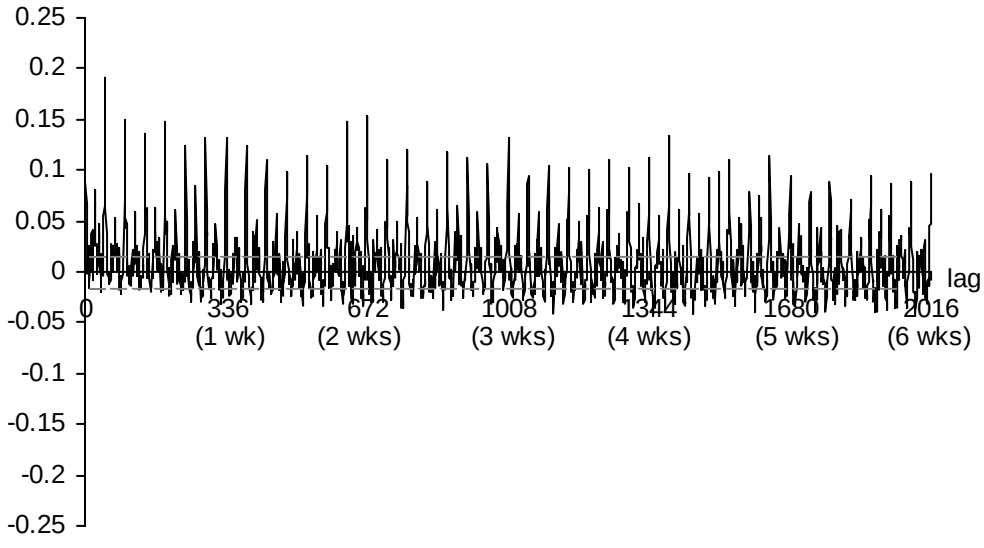

Figure 4. Autocorrelation function for the squared residuals from the SARMA model, Method P6. Dotted lines indicate 95% confidence interval.

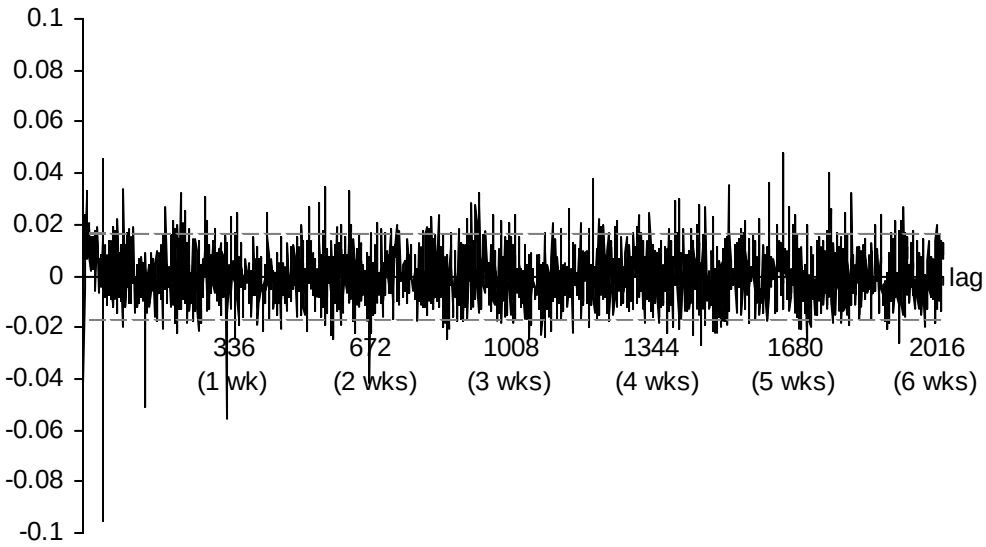

Figure 5. Autocorrelation function for the squared standardised residuals from the SARMA-SGARCH model, Method P7. Dotted lines indicate 95% confidence interval.

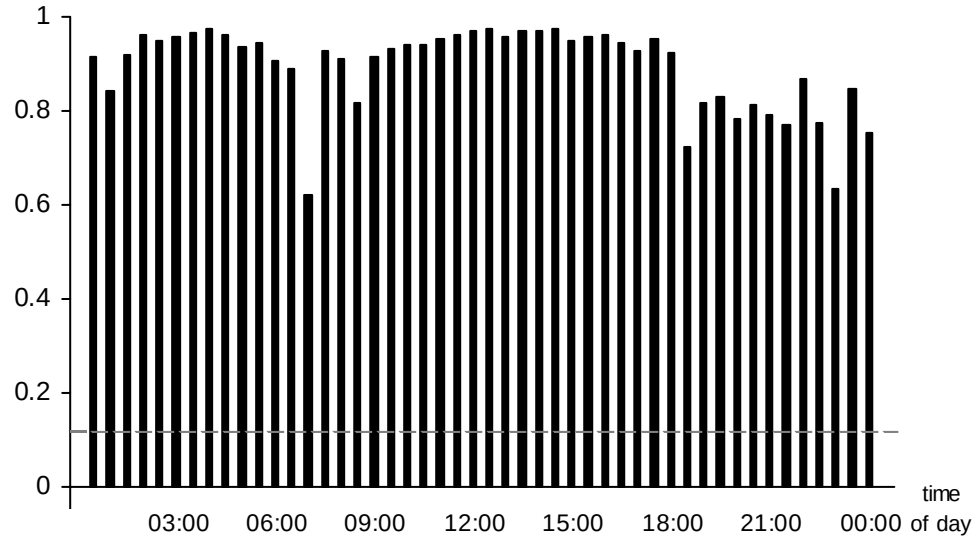

Figure 6. Lag 1 autocorrelation at the 48 half-hours of the day, calculated using observations from only the nine-month estimation sample. Dotted line indicates upper bound of 95% confidence interval.

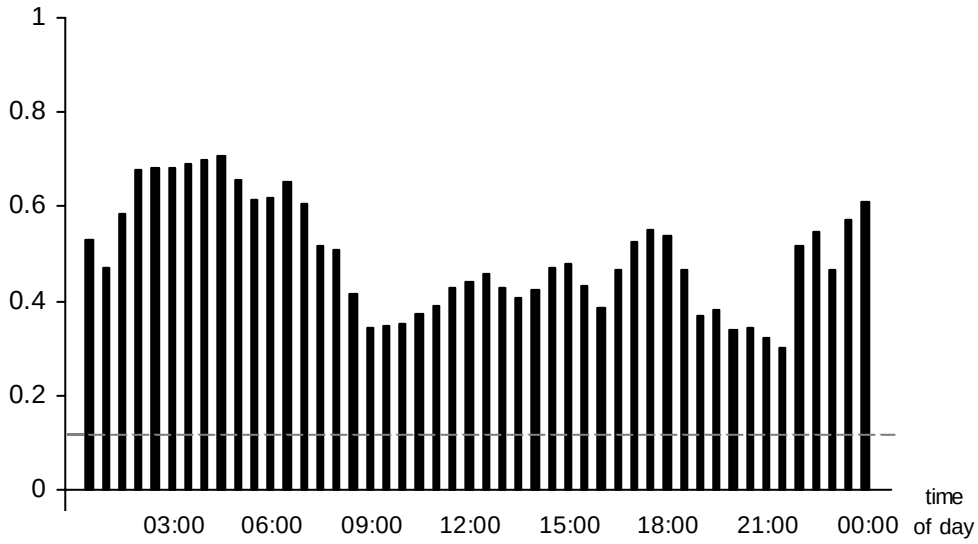

Figure 7. Lag 48 autocorrelation at the 48 half-hours of the day, calculated using observations from only the nine-month estimation sample. Dotted line indicates upper bound of 95% confidence interval.

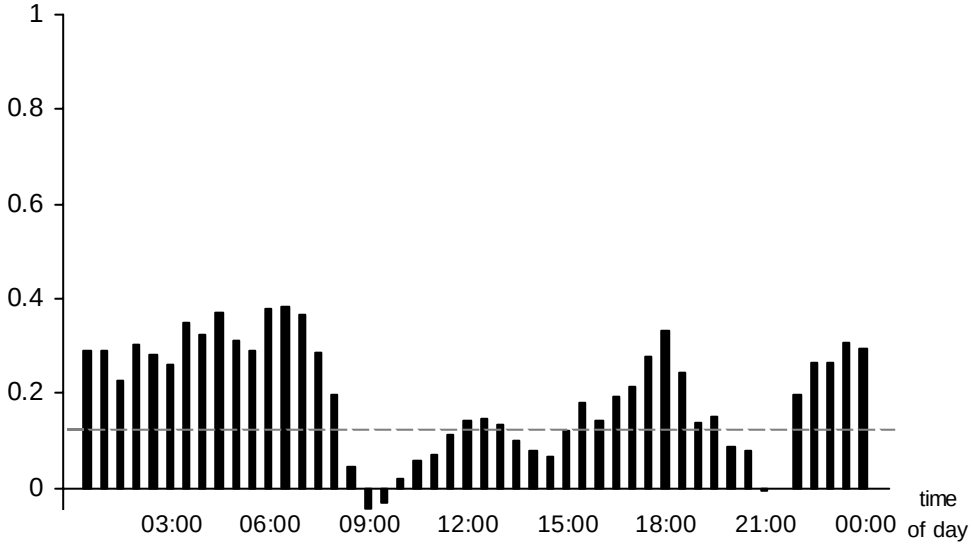

Figure 8. Lag 336 autocorrelation at the 48 half-hours of the day, calculated using observations from only the nine-month estimation sample. Dotted line indicates upper bound of 95% confidence interval.

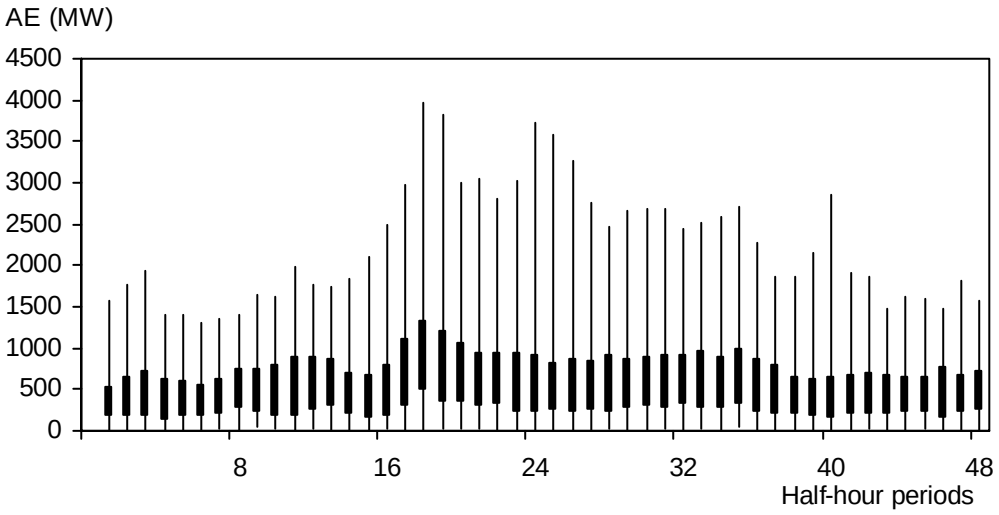

Figure 9. Box-plots for AE results from the SARMA model, Method P6, for point forecasting from forecast origin 3pm for the 48 half-hour periods commencing from 11:30pm on the next day.

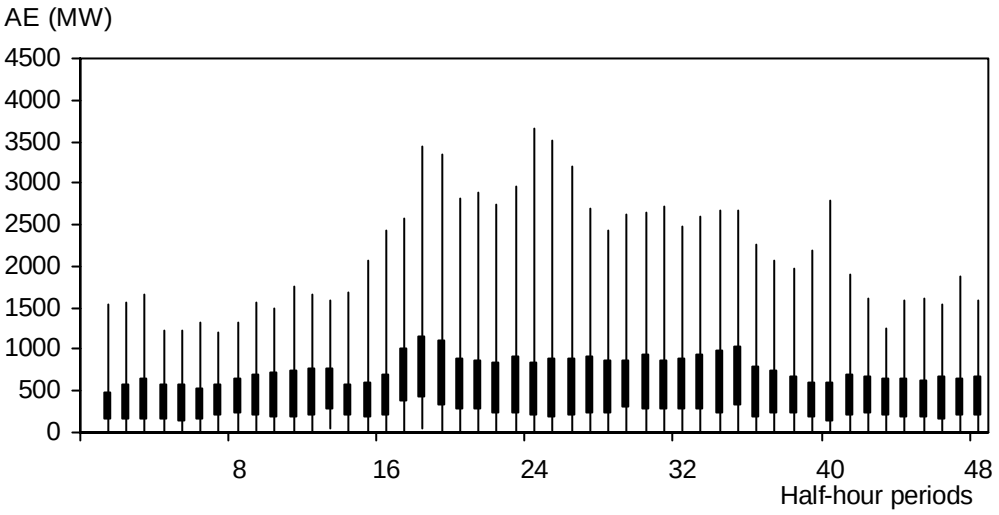

Figure 10. Box-plots for AE results from the SARMA model, Method P6, for point forecasting from forecast origin 8am for the 48 half-hour periods commencing from 11:30pm on the same day.

lag		1	2	3	48	96	144	336	672	1008
ϕ_i		1.397 (0.096)	-0.438 (0.086)			0.475 (0.059)	0.290 (0.051)	0.312 (0.037)	0.165 (0.035)	0.269 (0.027)
θ_i		0.437 (0.096)	0.115 (0.011)		-0.223 (0.009)	0.302 (0.059)	0.225 (0.040)	0.197 (0.038)	0.103 (0.034)	0.209 (0.026)
c	-589.5 (205.3)									

Table 1. Parameters for the SARMA model in expression (1), Method P6, with standard errors in parentheses.

	ω	λ_1	λ_2	λ_3	λ_4	ν_1	ν_2	ν_3	ν_4
$\phi_0(t)$	-279.0 (8.51)	58.7 (12.31)	101.5 (11.80)	5.7 (12.06)	9.4 (12.02)	23.3 (12.07)	-13.5 (11.89)	-4.6 (11.93)	39.0 (11.94)
$\phi_{48}(t)$	0.470 (0.008)	0.0933 (0.0109)	0.0703 (0.0123)	0.0628 (0.0115)	-0.0259 (0.0118)	-0.0054 (0.0118)	-0.0043 (0.0112)	-0.0068 (0.0112)	0.0188 (0.0114)
$\phi_{336}(t)$	0.254 (0.008)	0.0002 (0.0109)	-0.0586 (0.0122)	0.0846 (0.0114)	-0.1426 (0.0118)	-0.0550 (0.0117)	0.0673 (0.0112)	0.0027 (0.0111)	0.0437 (0.0113)

Table 2. Parameters for the periodic AR model in expression (4), Method P10, with standard errors in parentheses.

Half-hour periods	1-8	9-16	17-24	25-32	33-40	41-48	Mean Theil
P1 - Moving average	543	682	871	656	658	537	1.12
P2 - Moving average: time of day	505	636	859	654	628	515	1.08
P3 - Moving average: time of week	505	633	797	655	629	514	1.07
P4 - Exponential smoothing	467	575	778	712	613	506	1.04
P5 - Company's simple autoregression	453	551	818	674	625	517	1.03
P6 - SARMA	454	558	**775**	661	**598**	**485**	**1.00**
P7 - SARMA-SGARCH	467	564	805	660	600	499	1.02
P8 - Periodic AR: time of day and week	484	627	809	708	809	523	1.12
P9 - Periodic AR: time of day	485	609	836	675	703	512	1.08
P10 - Periodic AR: time of day, no AR(1)	**443**	**539**	814	**648**	612	497	**1.00**

Table 3. MAE (MW) for point forecasting from forecast origin 3pm for the 48 half-hour periods commencing from 11:30pm on the next day.

Half-hour periods	1-8	9-16	17-24	25-32	33-40	41-48	Mean Theil
P1 - Moving average	541	681	864	658	655	534	1.20
P2 - Moving average: time of day	505	635	852	656	623	512	1.15
P3 - Moving average: time of week	505	632	791	656	625	511	1.14
P4 - Exponential smoothing	421	504	740	699	605	499	1.05
P5 - Company's simple autoregression	411	473	773	679	609	480	1.03
P6 - SARMA	414	495	**703**	641	590	**470**	**1.00**
P7 - SARMA-SGARCH	430	497	725	**635**	**588**	482	1.01
P8 - Periodic AR: time of day and week	439	548	713	768	859	510	1.15
P9 - Periodic AR: time of day	435	515	723	687	723	495	1.08
P10 - Periodic AR: time of day, no AR(1)	**395**	**449**	768	650	598	482	**1.00**

Table 4. MAE (MW) for point forecasting from forecast origin 8am for the 48 half-hour periods commencing from 11:30pm on the same day.

lag		1	2	3	48	96	144	336	672	1008
α_i	1071.7 (514.5)	0.050 (0.010)	0.030 (0.007)		0.107 (0.014)		0.031 (0.013)	0.043 (0.014)	0.024 (0.010)	
β_i					0.296 (0.074)		0.239 (0.070)		0.170 (0.037)	
degrees of freedom	11.21 (1.13)									

Table 5. Parameters for the seasonal GARCH model in expression (5), with standard errors in parentheses.

		ω	λ_1	λ_2	λ_3	λ_4	ν_1	ν_2	ν_3	ν_4
$\alpha_0(t)$		19664.2 (1495.8)	-2750.4 (526.7)	-153.1 (596.3)	-3228.0 (569.2)	1076.9 (541.5)	1054.5 (549.1)	2829.1 (512.8)	1505.5 (508.9)	704.7 (506.4))
$\alpha_{48}(t)$		0.0381 (0.0055)	0.0154 (0.0076)	0.0040 (0.0081)	0.0139 (0.0069)	-0.0045 (0.0061)	-0.0052 (0.0050)	-0.0148 (0.0065)	-0.0046 (0.0055)	-0.0042 (0.0054)
$\alpha_{336}(t)$		0.0456 (0.0064)	0.0021 (0.0082)	0.0035 (0.0094)	0.0081 (0.0085)	-0.0027 (0.0078)	-0.0006 (0.0078)	-0.0025 (0.0073)	0.0011 (0.0073)	-0.0025 (0.0068)
β_{48}	0.1492 (0.0424)									
β_{336}	0.2786 (0.0395)									
degrees of freedom	239.1 (76.3)									

Table 6. Parameters for the periodic GARCH model in expression (6), with standard errors in parentheses.

Half-hour periods	1-8	9-16	17-24	25-32	33-40	41-48	Mean Theil
V1 - Historical variance	358	468	**903**	662	596	385	1.07
V2 - Historical variance: time of day	**294**	446	927	694	622	360	1.04
V3 - Historical variance: time of week	295	447	932	695	636	374	1.05
V4 - SARMA constant error variance	362	469	**903**	662	596	385	1.08
V5 - SGARCH	425	451	940	668	676	460	1.18
V6 - Periodic GARCH	309	**439**	906	**652**	**570**	**340**	**1.00**

Table 7. MAE (MW^2) for forecasting variance of SARMA point forecast errors corresponding to forecasts made from origin 3pm for the 48 half-hour periods commencing from 11:30pm on the next day. MAE values have been divided by 10^3.

Half-hour periods	1-8	9-16	17-24	25-32	33-40	41-48	Mean Theil
V1 - Historical variance	311	386	758	630	585	360	1.09
V2 - Historical variance: time of day	**246**	357	779	664	612	334	1.04
V3 - Historical variance: time of week	256	373	812	697	635	355	1.09
V4 - SARMA constant error variance	322	393	759	632	587	363	1.10
V5 - SGARCH	391	370	806	641	671	439	1.22
V6 - Periodic GARCH	267	**349**	**751**	**616**	**561**	**316**	**1.00**

Table 8. MAE (MW^2) for forecasting variance of SARMA point forecast errors corresponding to forecasts made from origin 8am for the 48 half-hour periods commencing from 11:30pm on the same day. MAE values have been divided by 10^3.

Half-hour periods	1-8	9-16	17-24	25-32	33-40	41-48	Mean Theil
V1 - Historical variance	475	721	1852	1285	1062	**533**	**0.97**
V2 - Historical variance: time of day	482	745	1806	1257	**1040**	535	**0.97**
V3 - Historical variance: time of week	483	747	1820	**1252**	1053	547	0.98
V4 - SARMA constant error variance	476	**720**	1851	1285	1062	**533**	**0.97**
V5 - SGARCH	509	735	**1792**	1283	1063	569	1.00
V6 - Periodic GARCH	**473**	772	1897	1334	1091	539	1.00

Table 9. RMSE (MW^2) for forecasting variance of SARMA point forecast errors corresponding to forecasts made from origin 3pm for the 48 half-hour periods commencing from 11:30pm on the next day. RMSE values have been divided by 10^3.

Half-hour periods	1-8	9-16	17-24	25-32	33-40	41-48	Mean Theil
V1 - Historical variance	399	591	1552	1230	1060	**495**	**0.98**
V2 - Historical variance: time of day	398	609	1519	**1202**	**1036**	496	**0.98**
V3 - Historical variance: time of week	398	610	1521	1230	1056	521	0.99
V4 - SARMA constant error variance	404	**590**	1548	1227	1058	496	**0.98**
V5 - SGARCH	449	598	**1496**	1225	1060	539	1.02
V6 - Periodic GARCH	**390**	622	1585	1273	1086	498	1.00

Table 10. RMSE (MW^2) for forecasting variance of SARMA point forecast errors corresponding to forecasts made from origin 8am for the 48 half-hour periods commencing from 11:30pm on the same day. RMSE values have been divided by 10^3.

Half-hour periods	1-8	9-16	17-24	25-32	33-40	41-48	Average all 48	Sum all 48
Methods based on SARMA point forecasts								
D1 - Historical variance	26	46	31	62	392	49	101	4857
D2 - Historical variance: time of day	**27**	45	35	64	392	49	**102**	**4904**
D3 - Historical variance: time of week	**27**	45	35	64	393	49	**102**	4901
D4 - SARMA constant error variance	26	46	31	62	392	49	101	4857
D5 - SGARCH	**27**	46	**36**	62	388	47	101	4841
D6 - Periodic GARCH	**27**	44	29	57	392	50	100	4785
Methods not based on SARMA point forecasts								
D7 - Simple benchmark	15	34	10	**71**	**428**	46	101	4832
D8 - Company benchmark	26	**48**	17	59	410	40	100	4800
D9 - SARMA-SGARCH	24	40	12	53	410	46	97	4674

Table 11. Economic benefit (£000s) of trading based on density forecasts from forecast origin 3pm for the 48 half-hour periods commencing from 11:30pm on the next day.

Half-hour periods	1-8	9-16	17-24	25-32	33-40	41-48	Average all 48	Sum all 48
Methods based on SARMA point forecasts								
D1 - Historical variance	33	52	51	81	411	55	**114**	5456
D2 - Historical variance: time of day	33	51	54	**83**	411	54	**114**	**5494**
D3 - Historical variance: time of week	33	51	**55**	82	410	53	**114**	5467
D4 - SARMA constant error variance	33	52	51	81	411	55	**114**	5458
D5 - SGARCH	33	51	55	81	407	52	113	5431
D6 - Periodic GARCH	33	50	50	76	411	56	113	5419
Methods not based on SARMA point forecasts								
D7 - Simple benchmark	16	36	15	80	**460**	50	109	5249
D8 - Company benchmark	**35**	**54**	28	66	421	**57**	110	5289
D9 - SARMA-SGARCH	30	47	37	76	434	53	113	5419

Table 12. Economic benefit (£000s) of trading based on density forecasts from forecast origin 8am for the 48 half-hour periods commencing from 11:30pm on the same day.