

# Electricity Exchange: Demand Side Unit performance monitoring

Vahid Bokharaie<sup>2</sup>, Paula Carroll<sup>1</sup>, Mel Devine <sup>\*2</sup>, Peter Fennell <sup>\*2</sup>,  
James Gleeson<sup>2</sup>, Kevin Hayes<sup>2</sup>, Geoff Hunter<sup>3</sup>, Jan Idiak<sup>4</sup>, William  
Lee<sup>2</sup>, Joseph Lynch<sup>2</sup>, Joanna Mason<sup>2</sup>, Jakub Nowotarski<sup>4</sup>, Maurice  
O’Connell<sup>2</sup>, David O’Sullivan<sup>2,5</sup>, Jakub Tomczyk<sup>4</sup>, Jonathan  
Ward<sup>6</sup>, Lai To Anson Yan<sup>7</sup>, and Paddy Finn<sup>8</sup>

<sup>1</sup>School of Business, University College Dublin, Ireland

<sup>2</sup>MACSI, Department of Mathematics and Statistics, University of  
Limerick, Ireland

<sup>3</sup>Limerick Institute of Technology, Ireland

<sup>4</sup>Wroclaw University of Technology, Poland

<sup>5</sup>University College Cork, Ireland

<sup>6</sup>School of Mathematics, University of Leeds, UK

<sup>7</sup>University of York, UK

<sup>8</sup>Industry representative, Electricity Exchange

<sup>\*</sup> *Corresponding authors mel.devine@ul.ie, peter.fennell@ul.ie*

August 15, 2013

MACSI, University of Limerick

**Executive summary:** Demand Side Response management encourages electricity demand reduction during peak hours. One avenue for achieving this is through Demand Side Units (DSUs). These are large electricity consumers who can afford to reduce their demand on the electricity grid when required. Issues with DSUs revolve around verification that the correct demand reduction takes place, with limited monitoring capabilities from the electrical grid operator Eir-Grid. This issue is studied here with the current methods thoroughly analysed and new methods proposed. In this report six different forecasting methods are presented, and their accuracy is compared using two different error metrics. Due to inherent stochasticity in demand it is found that there is no one forecasting method which is unequivocally best, but the ‘Keep it simple’ weekly and the temperature dependent models are identified as the most promising models to pursue. Initial investigations suggest that a ‘proxy day’ mechanism may be preferable to the current method of verifying that the correct demand reduction takes place.

*Keywords:* Data Analysis, Prediction Models, Demand Side Response Management

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Demand Side Response Management . . . . .	3
1.2	Current DSU performance monitoring method . . . . .	4
<b>2</b>	<b>Description of the problem</b>	<b>5</b>
<b>3</b>	<b>Data description</b>	<b>6</b>
3.1	Supermarket Data . . . . .	6
3.2	Bank Data . . . . .	6
3.3	Hospital Data . . . . .	7
3.4	Manufacturer Data . . . . .	8
3.5	Data Management Issues . . . . .	12
<b>4</b>	<b>Forecasting Models and Methodologies</b>	<b>13</b>
4.1	Forecasting Models . . . . .	13
4.1.1	Seasonal ARIMA . . . . .	13
4.1.2	KIS models . . . . .	16
4.1.3	Temperature dependent model . . . . .	16
4.1.4	ARIMA II . . . . .	17
4.1.5	Stepwise Linear Regression . . . . .	19
4.1.6	Additional methods and forecasting practices . . . . .	19
4.2	Error Metrics . . . . .	21
<b>5</b>	<b>Results and Analysis</b>	<b>23</b>
5.1	Forecasting Models . . . . .	23
5.2	Improved accuracy for shorter time horizons . . . . .	25
5.3	Difficulties with 5% precision accuracy . . . . .	26
<b>6</b>	<b>Dispatch verification</b>	<b>27</b>
<b>7</b>	<b>Conclusions, Recommendations and Further Work</b>	<b>29</b>
<b>8</b>	<b>Appendix</b>	<b>32</b>

# 1. Introduction

## 1.1 Demand Side Response Management

Demand Side Response can be defined as actions taken by consumers to cut the amount of electricity they use at particular times in response to a prompt from the grid operator. Demand Side Response has existed in Ireland since 2003 in the form of the Winter Peak Demand Reduction Scheme. This scheme was introduced to reduce demand on the electricity grid at peak consumption hours during the winter period. Demand Side Units (DSUs), which are typically industrial customers, received capacity payments to reduce their load between 5.00p.m. and 7.00p.m. In this way the demand on the grid was reduced, cutting the need of the grid operator to call on peak power generators whose limitations include their expense, inflexibility and high level of  $CO_2$  emissions.

The Winter Peak Demand Reduction Scheme ended after the winter of 2012-2013 and is being replaced by the Peak Demand Reduction scheme. Under this new scheme, DSUs can be asked to reduce their load, i.e., to dispatch, at any time. However, dispatch is forecast to occur seldomly and DSUs will be given notification of dispatch at least one day prior. DSUs requested to dispatch will then have to reduce their demand on the grid for the dispatch period which can last from thirty minutes up to two hours (see Figure 1 for an illustrative example of dispatch). Demand on the grid can be reduced by either reducing energy consumption or by using power from back-up generators. DSUs receive financial incentives for being available for dispatch, a key ingredient needed to stimulate participation in the Peak Demand Reduction scheme (SEI, 2008).

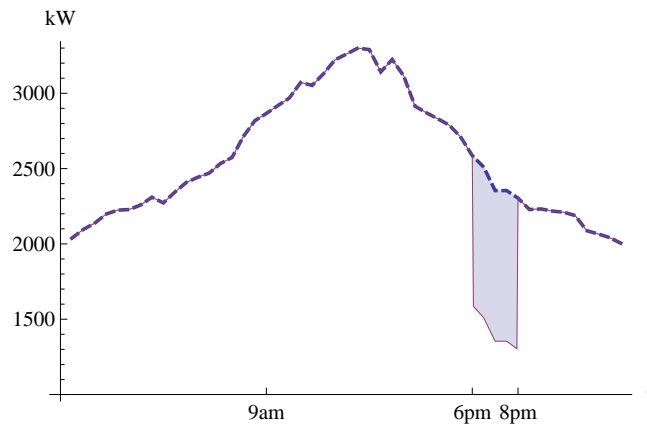


Figure 1: Electricity consumption of a DSU over a 24 hour period. Dashed Blue line is usage with no dispatch. Continuous pink line is usage with 1MW dispatch between 6.00p.m. and 8.00p.m. The shaded region shows the reduction in demand on the grid as a result of the dispatch.

DSUs must comply with the code of national grid operator EirGrid which specifies operational standards and protocols. DSUs are also subject to on-going performance monitoring checks to ensure a range of conditions are adhered to. One of these conditions relates to one day ahead demand forecasts which a DSU must submit on a daily basis. In particular, actual demand must be within  $\pm 5\%$

of forecast values. Poor performance may result in the DSU being declared at a lower availability resulting in a loss of revenue.

The problem brought to MACSI was by a DSU aggregator, Electricity Exchange. They are a third party company that aggregates the demand from a number of DSUs, which include supermarkets, banks and hospitals, and act as a single DSU. Their problem relates to the current DSU performance monitoring regime.

## 1.2 Current DSU performance monitoring method

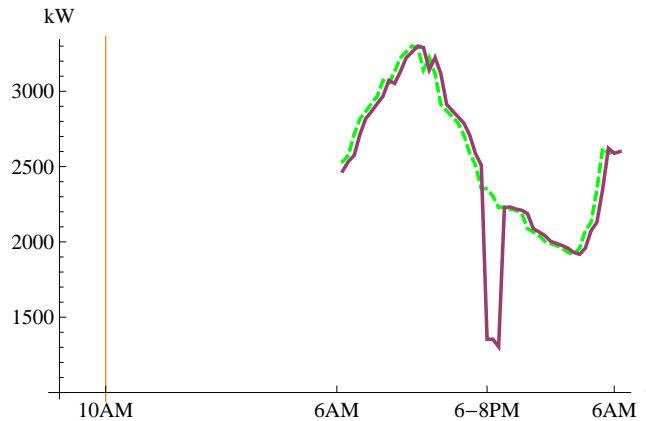


Figure 2: Current performance monitoring method employed by EirGrid. Forecasts (dashed green) are submitted at 10.00a.m. for 24-hour period starting at 6.00a.m. the following day. This is monitored by EirGrid by continuously comparing it to the actual usage (purple). Dispatch is verified by subtracting the actual from forecast usage in the dispatch period.

DSUs submit demand forecasts to EirGrid on a daily basis. Specifically, a prediction is given at 10.00a.m. for the 24-hour period from 6.00a.m. the following morning to 6.00a.m. the morning after that. The accuracy of the prediction is monitored by EirGrid every 30 minutes throughout the 24-hour period. Predicted values that differ from actual values by greater than  $\pm 5\%$  are flagged as poor predictions. As described in Section 1, a high number of poor predictions can result in the DSU being declared at a lower availability with financial implications for the DSU.

The reason for forecast monitoring is to verify dispatch. When a DSU dispatches, its demand on the grid should fall by the amount of electricity it was requested to dispatch. However, EirGrid has no way of knowing what the DSUs demand on the grid would have been had it not been asked to dispatch. Thus calculating how much dispatch actually occurs becomes an issue. EirGrid currently verifies dispatch by measuring the difference between the forecast usage and actual usage in the dispatch period and checking if this difference corresponds to the amount of dispatch requested. See Figure 2 for a graphical illustration of the current DSU performance monitoring method.

## 2. Description of the problem

The current performance monitoring scheme is seen as being quite strict and difficult to adhere to. Early estimates are that DSU forecasts will fall outside of the  $\pm 5\%$  margin a large number of times per day. Electricity Exchange are concerned about the likelihood of poor forecast performance and the effect that will have on its business. With that in mind Electricity Exchange brought the following questions to MACSI:

1. Which prediction algorithms are best suited to the task of predicting customer demand?
2. Is there evidence of an improvement in accuracy if the prediction time horizon is reduced from 48 hours?
3. Can any prediction model be guaranteed to have a forecast error within  $\pm 5\%$  for all customer segments?
4. Are there more optimal ways of verifying dispatch?

The problem is tackled in the following manner. Forecasting models are developed to predict usage for a time horizon of 48 hours ahead. The accuracy of these models are compared for different customer segments, with different models working best for different customers. The gain in accuracy by reducing the forecast time horizon is then examined. Analytical and numerical results are given to highlight the difficulty of keeping the forecasting error within  $\pm 5\%$ . Finally, alternative dispatch verification procedures are examined.

### 3. Data description

Data was supplied by Electricity Exchange on twenty two DSUs (16 supermarkets, 1 bank, 2 hospitals, 1 plastic manufacturer, 1 cement manufacturer, and 1 steel manufacturer). For each data set, the demand at every half hour interval over a period of one year was given. To visually observe trends, in this section some sample time series of demand for the six different customer segments provided are graphed and discussed. All the figures in this section commence on a Monday.

#### 3.1 Supermarket Data

The data provided for twelve of the supermarkets runs for a period of one year from 01/02/2012 to 31/01/2013. The data for the remaining four supermarkets (supermarkets 13 to 16) runs for a period from the 01/08/2011 to the 31/08/2012. In Figure 3 we plot a sample of the weekly demand for one supermarket. It can be seen clearly that there is a prominent daily pattern: electricity usage begins at a low level at midnight, reaches a peak during the middle of the day and falls back to a low level of usage again. This pattern repeats itself each day.

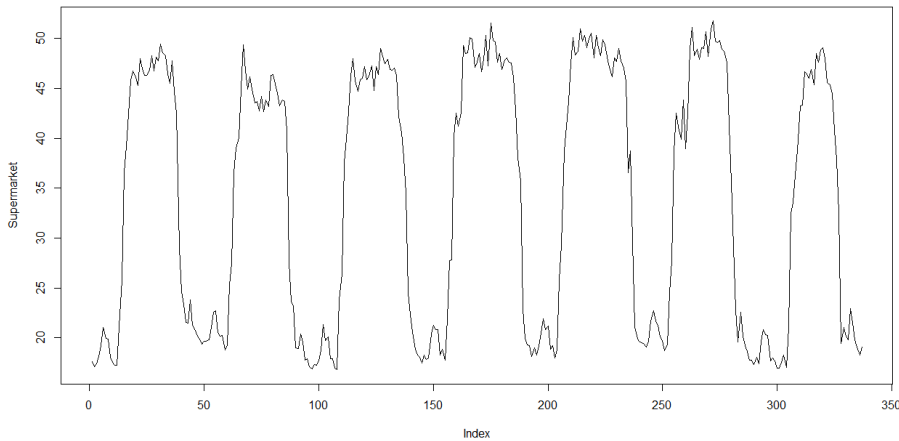


Figure 3: Time series of the energy demand in kWh of one of the supermarkets taken at half-hour intervals over a period of one week (336 half hours).

#### 3.2 Bank Data

The data for the bank is provided for a period of thirteen months from 01/01/2012 to 31/01/2013. In Figure 4 a sample of the weekly demand for the bank is shown. During the weekdays a daily pattern is observable: low electricity usage at midnight and that usage peaks during office hours from 7am to 4.30pm. After 4.30pm the electricity usage reduces to a low level, however there is a small peak briefly at around 7.30pm each day. It can be observed that the electricity usage

is much lower during both Saturday and Sunday. From plotting three weeks' demand in Figure 5 it is clear that this pattern continues each week.

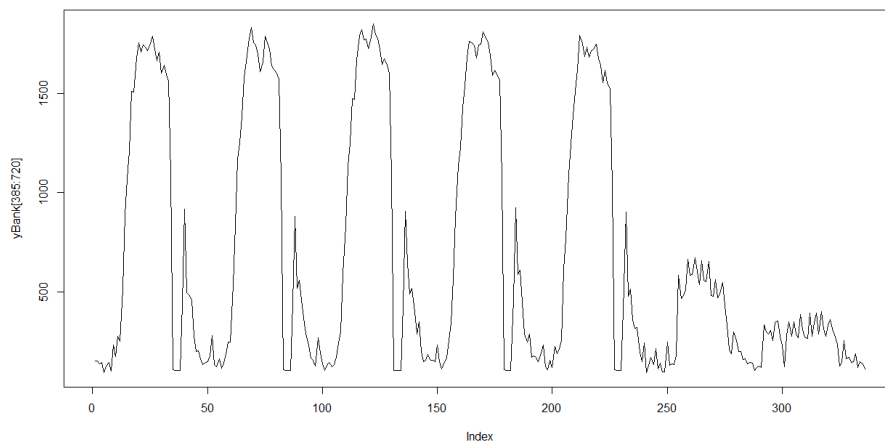


Figure 4: Time series of the energy demand in kWh of the bank taken at half-hour intervals over a period of one week (336 half hours).

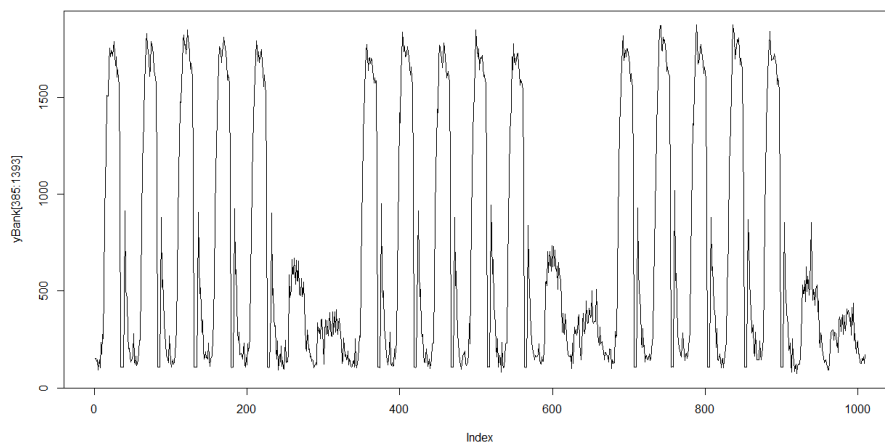


Figure 5: Time series of the energy demand in kWh of the bank taken at half-hour intervals over a period of three weeks (1008 half hours).

### 3.3 Hospital Data

The data for the two hospitals is provided for a period of one year from 01/03/2011 to the 29/02/2012. In Figure 6 we plot a sample of the weekly demand for one of the hospitals. We observe that during the week the electricity usage starts at low usage at midnight, reaches a peak at about 11am each day and declines

to low electricity usage again. There is much lower electricity usage during the weekend. From plotting three weeks' demand in Figure 7 it is clear that this pattern continues each week.

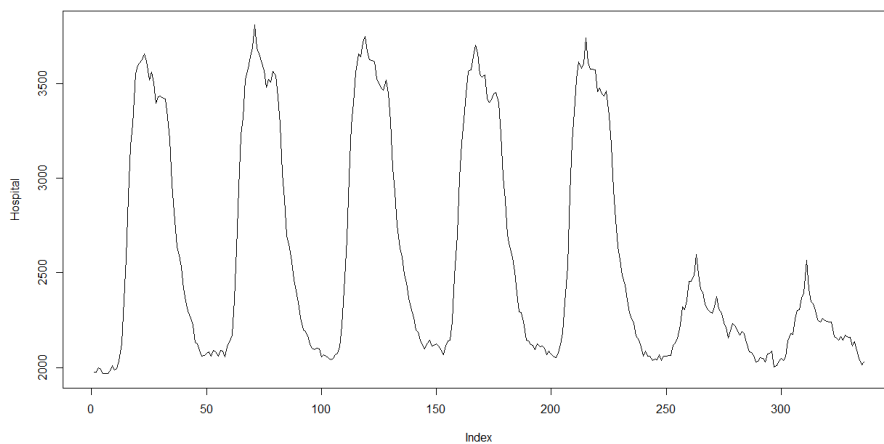


Figure 6: Time series of the energy demand in kWh of one of the hospitals taken at half-hour intervals over a period of one week (336 half hours).

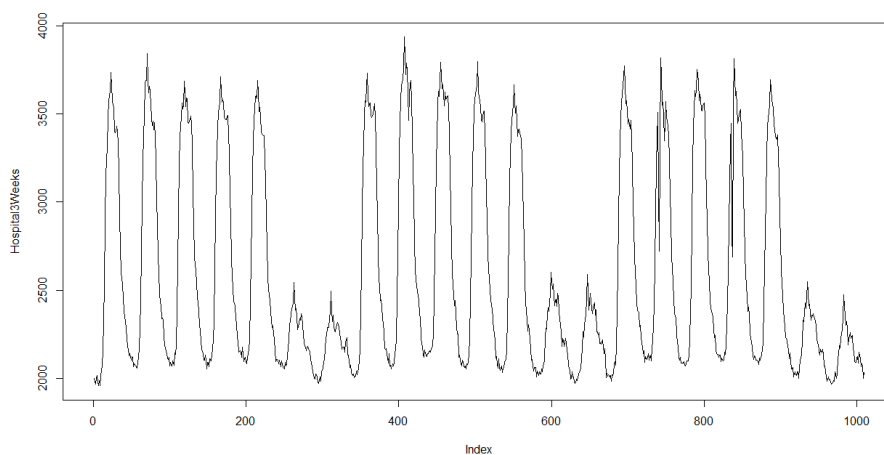


Figure 7: Time series of the energy demand in kWh of one of the hospitals taken at half-hour intervals over a period of three weeks (1008 half hours).

### 3.4 Manufacturer Data

The manufacturer data provided is from three different industrial sectors: plastic, cement, and steel. The data is provided for a period of one year from 01/02/2012 to 31/01/2013.



### Plastic manufacturer data

In Figure 8 a sample of the weekly demand for the plastic manufacturer is shown. It is observable that they do not use any electricity during the weekdays from 8.00am to 9.00pm, instead they use large amounts of electricity overnight from 9.00pm to 8.00am. During the weekend, the plastic manufacturer utilises electricity at all times, but at a lower level. There are no periods of zero electricity usage during the weekend. From plotting three weeks' demand in Figure 9 it is clear that this pattern continues each week.

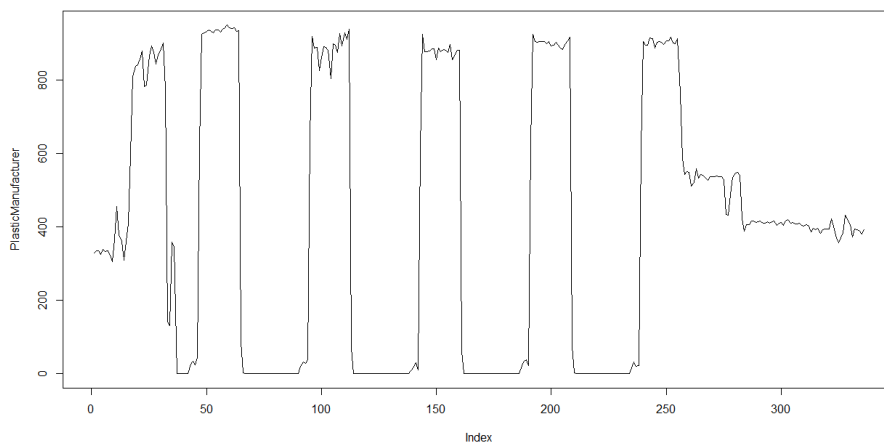


Figure 8: Time series of the energy demand in kWh of the plastic manufacturer taken at half-hour intervals over a period of one week (336 half hours).

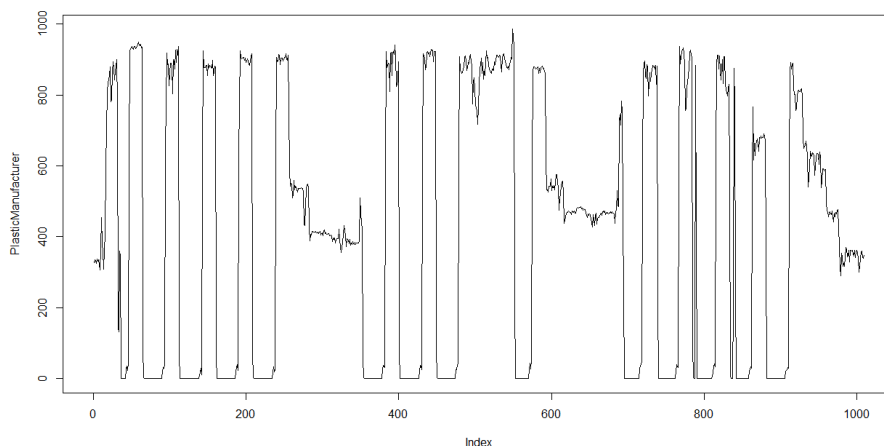


Figure 9: Time series of the energy demand in kWh of the plastic manufacturer taken at half-hour intervals over a period of three weeks (1008 half hours).

### Steel manufacturer data

In Figure 10 a sample of the weekly demand for the steel manufacturer is shown. It is observed that there is a regular pattern of usage from Monday to Friday. On Saturday, the peak electricity usage does not reach the maximum peak reached during the week, and on Sunday the peak electricity usage is lower again. From plotting three weeks' demand in Figure 11 it is clear that this pattern continues each week.

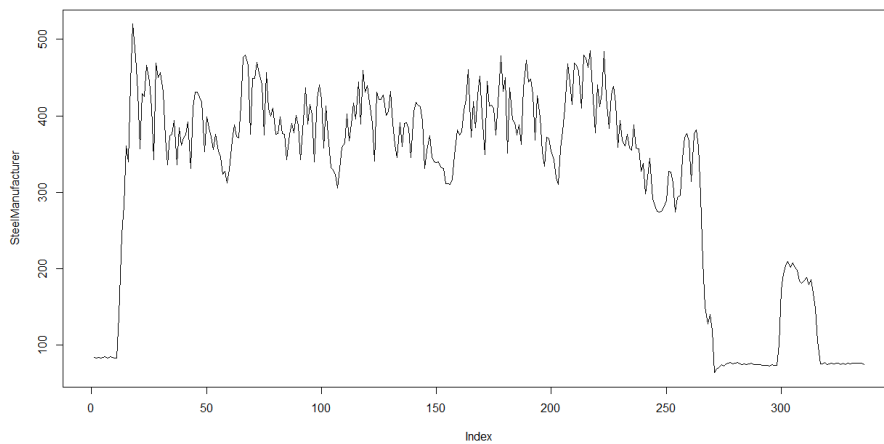


Figure 10: Time series of the energy demand in kWh of the steel manufacturer taken at half-hour intervals over a period of one week (336 half hours).

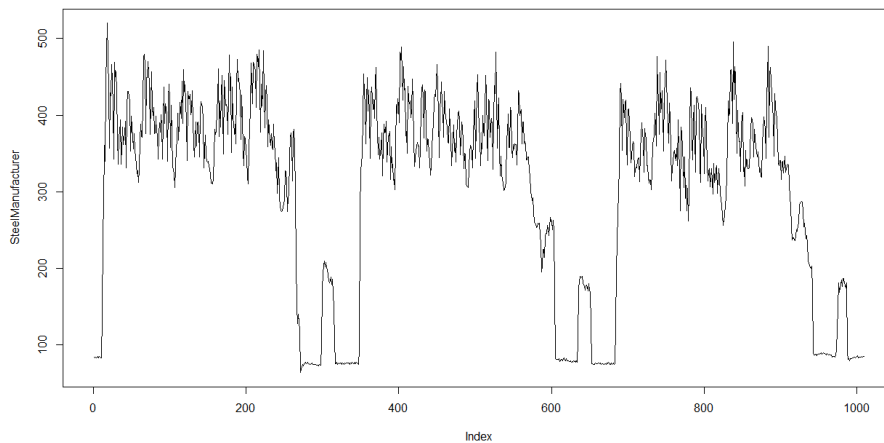


Figure 11: Time series of the energy demand in kWh of the steel manufacturer taken at half-hour intervals over a period of three weeks (336 half hours).

### Cement Manufacturer Data

The cement manufacturer does not have as obvious a pattern as the other industries considered. In Figures 12 and 13 a sample of the weekly demand and a sample of three weeks' demand, respectively, are shown. The complete year's data for the cement manufacturer is shown in Figure 14. Unlike the other DSUs, this DSU has no obvious consumption patterns, and therefore is not incorporated in any of the models developed in Section 4. This DSU is analysed separately in Appendix A Stochastic DSUs.

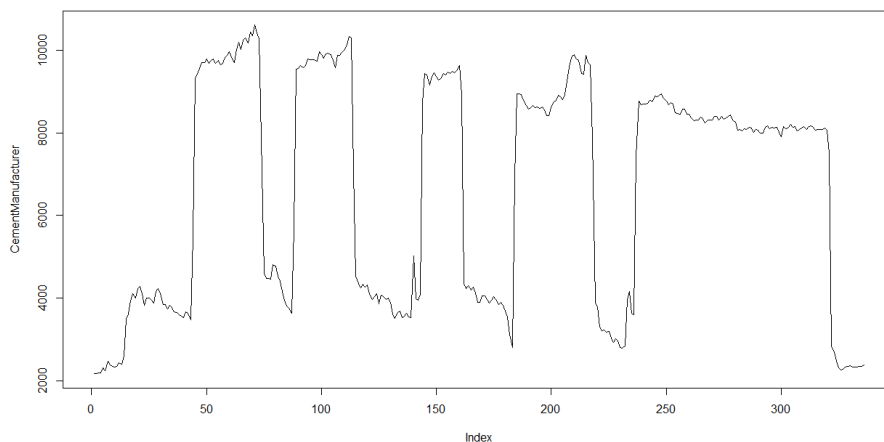


Figure 12: Time series of the energy demand in kWh of the cement manufacturer taken at half-hour intervals over a period of one week (336 half hours).

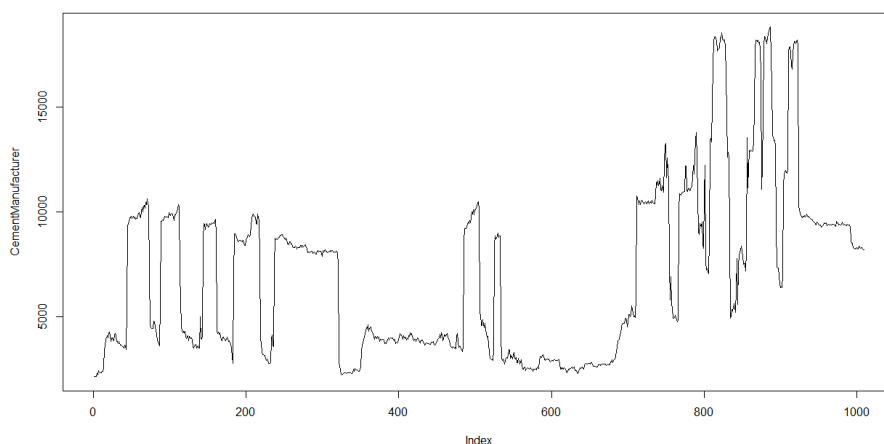


Figure 13: Time series of the energy demand in kWh of the cement manufacturer taken at half-hour intervals over a period of three weeks (1008 half hours).

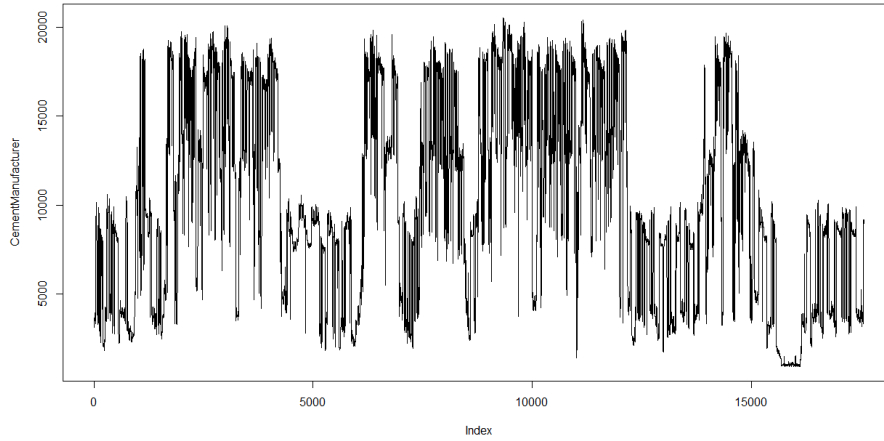


Figure 14: Time series of the energy demand in kWh of the cement manufacturer taken at half-hour intervals over a period of one year (17472 half hours).

### 3.5 Data Management Issues

The data provided was in half-hourly readings of power output in kilowatt hours (kWh). Each data point represents the average power level monitored over the following 30-minute interval, rounded to the nearest kWh.

Some periods of power usage levels were identified as being outliers and not representative of common power consumption patterns. Examples of these periods include the Christmas and New Year period during which all of the DSU's showed uncharacteristically low power levels. Similarly, the two-week period for Builders' holidays (late July to early August) was identified as being non-representative in one DSU. These outlying periods are shown in Figure 15. These identified outliers, along with rare incidences of missing data (possibly due to a power cut), were fully included in the test data sets.

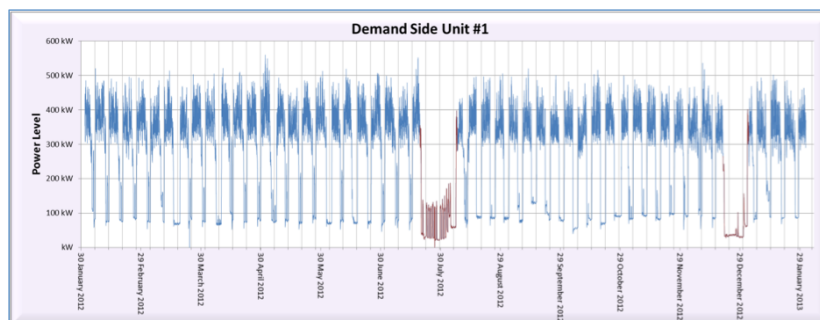


Figure 15: Typical consumption patterns for a steel manufacturer shown in blue. The highlighted brown data shows two outlier periods which correspond to builders' holidays in July and the Christmas period.

## 4. Forecasting Models and Methodologies

The first task was to build forecasting models that could predict demand based on information available at the time of forecast. Data was supplied by Electricity Exchange on each of its twenty two DSUs. For each data set, the demand at every half hour interval over a period of one year was given. Furthermore, weather data was supplied on daily mean temperatures, daily temperature ranges and daily mean rainfall.

The demand data set for each DSU was split into two parts. The first 75%, or 13140 entries, were set aside for the training data set. This was to be used for model construction and fitting parameters. The last 25%, 4380 entries, was the test data set which was to be used to test the quality of the model based on certain metrics.

The data set is a time series of electricity demand, and this along with weather data would be the basis of the models. Each model could take as input the forecasted weather conditions and/or the previous electricity demand and produce a forecast. Different models were examined, and a detailed explanation of each is given here.

### 4.1 Forecasting Models

#### 4.1.1 Seasonal ARIMA

In simple multiplicative models it is assumed that data at time  $t$ ,  $Z_t$ , is described by an underlying trend  $Tr$  which varies because of a seasonal component  $S$  and a residual (growth/decline) component  $R$ :  $Z_t = Tr * S * R$ . The aim is to remove seasonal and residual components so that the underlying trend can be identified and forecast forward. The forecast trend is then re-adjusted to account for seasonal and residual components.

Seasonal ARIMA models Yaffee and McGee (2000) are capable of describing such data. These models consist of seasonal terms, differencing terms, autoregressive terms and moving average terms.

An autoregressive (AR) model of order  $p$ ,  $AR(p)$ , calculates a spot estimate for time point  $t$  as a weighted average of the  $t-1$  to  $t-p$  time series data points. Weights are calculated as regression coefficients using, for example, maximum likelihood estimation.

Consider a time series  $y_1, y_2, \dots, y_n$ . An autoregressive model  $AR(p)$  states that  $y_i$  is the linear function of the previous  $p$  values of the series plus an error term:

$$y_i = \phi_0 + \phi_1 y_{i-1} + \phi_2 y_{i-2} + \dots + \phi_p y_{i-p} + \epsilon_i,$$

where  $\phi_1, \dots, \phi_p$  are weights that have to be determined, and the error  $\epsilon_i$  is normally distributed with zero mean and variance  $\sigma^2$ .

A moving average model of order  $q$ , denoted  $MA(q)$ , calculates a spot estimate for time point  $t$  as a mean term of the series,  $\omega_0$ , and a moving weighted average of current and previous white noise error terms,  $\epsilon_t, \epsilon_{t-1}, \dots, \epsilon_{t-q}$ .

$$y_t = \omega_0 - \omega_1 \epsilon_{t-1} - \omega_2 \epsilon_{t-2} - \dots - \omega_q \epsilon_{t-q} + \epsilon_t,$$

$\omega_0, \omega_1, \dots, \omega_q$  are constant coefficients.

Combining AR and MA models, autoregressive moving average (ARMA( $p,q$ )) models can be defined as:

$$y_t = c + \sum_{i=1}^p \phi_i y_{t-i} - \sum_{j=1}^q \omega_j \epsilon_{t-j},$$

where  $c$  is the combined value of the constants  $\phi_0$  and  $\omega_0$ .

A time series showing a trend  $Tr$  is not stationary in mean, which is a necessary assumption for ARMA models. Differencing is an operation that can be applied to a time series to remove a trend. If after differencing the time series is stationary in mean and variance, then an ARMA( $p,q$ ) model can be used.

Autoregressive integrated moving average (ARIMA) models remove trends in time series by including differencing in the ARMA model. ARIMA( $p,d,q$ ) models have an autoregressive part of order  $p$ , a moving average part of order  $q$  and having applied  $d$  order differencing. The equation for a predicted point  $y_t$  in backshift notation is given by

$$\underbrace{\left(1 - \sum_{i=1}^p \phi_i B^i\right)}_{\text{AR}(p)} \underbrace{(1 - B)^d}_{\text{I}(d)} y_t = c + \underbrace{\left(1 - \sum_{j=1}^q \omega_j B^j\right)}_{\text{MA}(q)} \epsilon_t,$$

where  $B$  is the backshift operator defined as

$$B^k y_t = y_{t-k}.$$

ARIMA models can deal with time series that have trends but need to be extended for time series that display seasonal behaviour. In the case of the latter, a seasonal ARIMA model is required.

ARIMA( $p,d,q$ )( $P,D,Q$ ) $_s$  models incorporate seasonal behaviour into the ARIMA model. They are defined by seven parameters.

$$\begin{aligned} & \underbrace{\left(1 - \sum_{i=1}^p \phi_i B^i\right)}_{\text{AR}(p)} \underbrace{\left(1 - \sum_{k=1}^P \beta_k B^{k*s}\right)}_{\text{ARs}(P)} \underbrace{(1 - B)^d}_{\text{I}(d)} \underbrace{(1 - B^s)^D}_{\text{Is}(D)} y_t \\ & = c + \underbrace{\left(1 - \sum_{j=1}^q \omega_j B^j\right)}_{\text{MA}(q)} \underbrace{\left(1 - \sum_{l=1}^Q \theta_l B^{l*s}\right)}_{\text{MA}s(Q)} \epsilon_t \end{aligned} \quad (1)$$

AR( $p$ ), MA( $q$ ), I( $d$ ) and  $c$  are as previously defined. ARs( $P$ ), MA( $s$ )( $Q$ ) and Is( $D$ ) are seasonal autoregressive, moving average and differencing terms respectively. These take into account the seasonality of the time series which may include daily patterns, weekly patterns and so on.  $s$  is the period of the seasonal pattern appearing. In this case,  $s$  takes the value 48 which corresponds to 48 half-hour time steps or, equivalently, one day. Thus the seasonality incorporated in the model is daily.

The Demand Side Units being considered consisted of sixteen supermarkets, one bank, one steel manufacturing plant, one plastic manufacturing plant and

two hospitals. The data for the DSUs were fitted with a seasonal ARIMA model. The appropriate model was selected by using considerations such as the autocorrelation and partial autocorrelation functions, residual normality tests, data stationarity tests, AIC and prediction errors.

Following these consideration, it was found that the order of the optimal model was  $ARIMA(3,0,1)(0,1,1)_{48}$ . Inserting these parameter choices into Equation 1 and expanding out the backshift operator gives the following linear equation for the predicted demand  $y_t$ :

$$y_t = \phi_1(y_{t-1} - y_{t-49}) + \phi_2(y_{t-2} - y_{t-50}) + \phi_3(y_{t-3} - y_{t-51}) \\ + \epsilon_t - \omega_1\epsilon_{t-1} - \theta_1(\epsilon_{t-48} - \omega_1\epsilon_{t-49}) + c$$

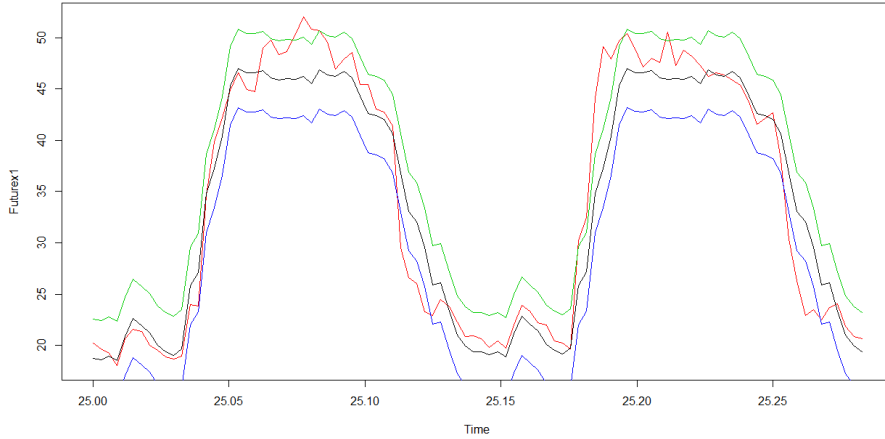


Figure 16: Seasonal ARIMA predictions for one supermarket over a 48-hour time period. The red line is actual electricity usage. The black line is predicted electricity usage. The green and blue lines are the upper and lower 95% confidence intervals.

As an example of the seasonal ARIMA model, details are here given of the model applied to one of the supermarkets. The model with parameters fitted was as follows:

$$y_t = 0.9265(y_{t-1} - y_{t-49}) - 0.0778(y_{t-2} - y_{t-50}) - 0.0106(y_{t-3} - y_{t-51}) \\ + \epsilon_t - 0.0979\epsilon_{t-1} + 0.9663(\epsilon_{t-48} - 0.0979\epsilon_{t-49})$$

The coefficients for the other fifteen supermarkets are given in Table 1.

Figure 16 shows the seasonal ARIMA predictions for the supermarket. Included in the plot is the actual electricity usage, the predicted electricity usage under the seasonal  $ARIMA(3,0,1)(0,1,1)_{48}$  model, and the 95% percent confidence interval of the estimation. This confidence interval is based on the assumption that the residuals are normally distributed, and is given by

$$y_t \pm 1.96\sigma,$$

DSU	ARIMA coefficients				
	$\phi_1$	$\phi_2$	$\phi_3$	$\omega_1$	$\theta_1$
Supermarket 1	0.9265	0.0778	0.0106	0.0979	0.9663
Supermarket 2	0.1021	0.8223	0.2533	0.9784	1
Supermarket 3	1.4511	0.4138	0.0446	0.9044	0.9579
Supermarket 4	1.2182	0.5234	0.2019	0.1190	0.9547
Supermarket 5	0.539	0.1374	0.1479	0.4397	0.9593
Supermarket 6	1.182	0.4576	0.1362	0.1176	0.8879
Supermarket 7	0.9795	0.1031	0.0296	0.0226	0.9603
Supermarket 8	0.1595	0.5758	0.0964	0.6602	0.9821
Supermarket 9	0.3729	0.3630	0.0052	0.6850	0.9586
Supermarket 10	1.0979	0.2601	0.0065	0.1152	0.9176
Supermarket 11	1.8739	1.0304	0.1533	0.9504	0.9233
Supermarket 12	1.0867	0.0319	0.1945	0.1520	0.9570
Supermarket 13	0.5156	0.2113	0.1694	0.3799	0.9889
Supermarket 14	1.6079	0.6420	0.0250	0.6609	0.7753
Supermarket 15	1.1346	0.4216	0.1831	0.0328	0.9658
Supermarket 16	0.5208	0.3400	0.0431	0.6860	0.9891

Table 1: ARIMA coefficients (see Equation 4.1.1) for the sixteen supermarkets. Note that  $c = 0$  for all models.

where  $\hat{y}_t$  is the predicted value using the seasonal ARIMA model, 1.96 is the normal distribution point for a 95% confidence interval and  $\sigma$  is the maximum likelihood estimate of the innovations standard deviation based on the fitted seasonal ARIMA model.

#### 4.1.2 KIS models

For the purpose of benchmarking predictions from more sophisticated prediction models, such as ARIMA, a series of simple forecasts dubbed the ‘Keep It Simple’ (KIS) models are used. These forecasts rely on the highly seasonal variation of the data, particularly over daily and weekly timescales. In the first case, *KIS median*, the weekday half-hour median of the training data is used as the forecast for the corresponding weekday half-hour in the test data. Note that the weekly forecast is the same for all weeks in the future. In the second case, *KIS weekly*, last week’s energy use is used as a forecast for this week, i.e. at half-hour  $t$  the forecast is

$$y_t = y_{t-336}. \quad (2)$$

The use of more involved prediction methods can only be justified if they can outperform simple predictors such as the KIS models. An example of the output from the KIS weekly method is shown in Figure 17.

#### 4.1.3 Temperature dependent model

The results of the KIS models were promising, so avenues for a possible extension to these models were sought. The best possible extension was found to be that of including weather data.

A temperature-demand dependence can be quite evident, as illustrated in Figure 18. This relationship is strongest during periods of high demand. This



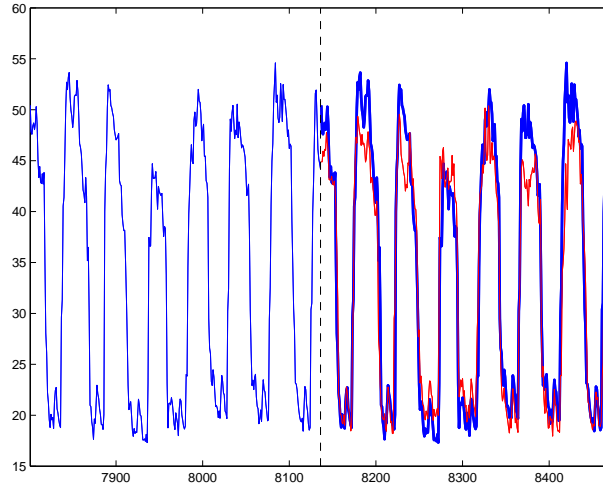


Figure 17: Actual demand for week 1 (blue), predicted demand using the KIS weekly model for week 2 (blue bold), and actual demand for week 2 (red) for one of the supermarkets.

has a simple explanation, that DSUs require more electricity when the temperature is higher. This requirement could be in the form of heat related consumption, such as refrigeration, heating etc.

The KIS weekly model was extended to include temperature dependence. An extra term was included in Equation 2 which takes into account the difference in temperature between the time point to be predicted and one week previous. The formula for the extended model is

$$y_t = y_{t-336} + \Delta T \cdot \alpha,$$

where  $\Delta T$  is difference between the temperature at time  $t$  (in this case it is forecast) and time  $t - 336$ , and  $\alpha$  is the slope of the linear function.  $\alpha$  was fitted using ordinary least squares on all available historical temperature data.

#### 4.1.4 ARIMA II

The ARIMA model (as described in Section 4.1.1) uses values from the previous day to predict today's electricity demand. In contrast, the KIS models use values from the previous week. In order to study this contrast in more detail, a correlation analysis was performed to see which days are more correlated with their previous days and which days are more correlated with the same day from the previous week.

Table 2 shows the percentage of time that each day was more correlated with the same day from the previous week versus their previous day. For example, it was found that Mondays were more correlated with the previous Monday 84% of the time while Mondays were more correlated with their previous day (Sunday) only 16% of the time. Table 2 shows that Mondays, Thursdays, Saturdays and Sundays are more correlated with the same day from the previous week while Tuesdays and Fridays are more correlated with their previous days (Mondays

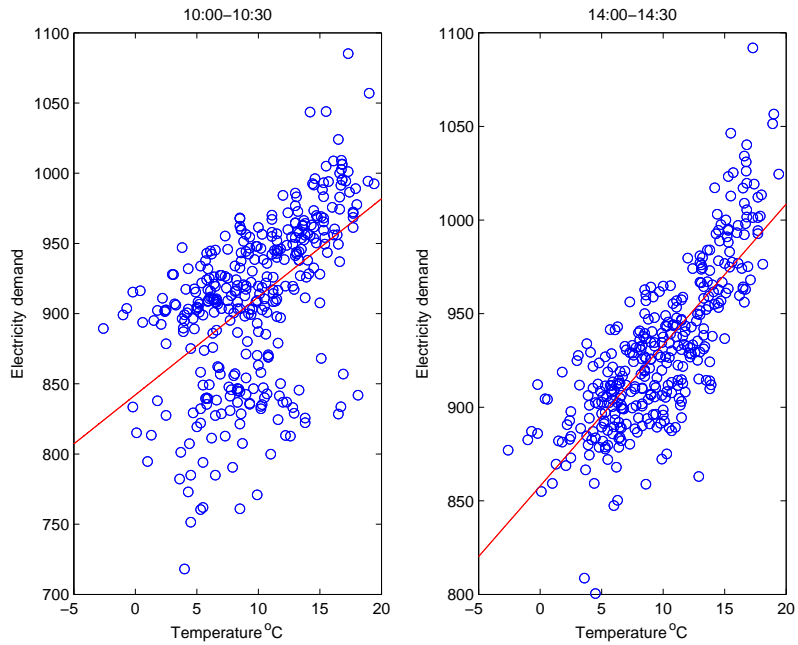


Figure 18: Demand at 10.00a.m. to 10.30a.m. (left) and 2.00p.m. to 2.30p.m. (right) versus temperature for one of the supermarkets. A significant linear relationship exists.

Day	With same day previous week	With previous day
Monday	84%	16%
Tuesday	45%	55%
Wednesday	51%	49%
Thursday	78%	22%
Friday	37%	63%
Saturday	82%	18%
Sunday	86%	14%

Table 2: The percentage of time that each day is more correlated with the same day from the previous week versus its previous day.

and Thursdays respectively). It is not clear which day Wednesdays are more correlated to.

As a result of this analysis, the ARIMA model (described in Section 4.1.1) was used with manipulated data. Instead of using the conventional daily pattern of Monday, Tuesday, Wednesday etc., if one is predicting electricity demand for a Monday, then data for Mondays only is fed into the model. If one is predicting for a Tuesday then data for Tuesdays only is fed into the model and so on. Because the ARIMA model looks at previous day's values, when, for example, data for Sundays only is being used, the previous day is in fact the previous Sunday. This data manipulation incorporates seasonality into the ARIMA model. When this is done in this report, the model is known as ARIMA II.

#### 4.1.5 Stepwise Linear Regression

Stepwise linear regression (SLR) is a procedure that uses a systematic approach to building a model with a large number of predictor variables, McClave et al. (1997). In the first step of the procedure a linear regression model is fitted using the best predictor only. The regression with the best predictor is the one with the smallest error defined by

$$\sum_t (\hat{y}_t - y_t)^2, \quad (3)$$

where  $\hat{y}_t$  represents the value predicted by the regression and  $y_t$  the actual value. At each of the next steps, another linear regression model is fitted, each time with the next best predictor being added to the regression of the previous step. This procedure continues as long as the error in equation (3) decreases at each step. Once this error stops decreasing the procedure stops and no more predictors are added. In this analysis 1000 predictor variables are used to forecast  $y_t$ . These are the demands  $y_{t-1}, y_{t-2}, \dots, y_{t-1000}$ , i.e., the last 23 days of data. Therefore, this method incorporates both daily and weekly patterns. Each linear regression is fitted using the first 5000 data entries. In order to prevent over fitting to this data, the linear regressions found at each step of the procedure are also cross-validated with the next 5000 data entries. The linear regression with the smallest error (according to equation (3)) on the cross-validation dataset is chosen as the model.

For each DSU a model consisting of  $N$  predictors is developed taking the form

$$\hat{y}_t = \sum_{i=1}^N \beta_i y_{t-t_i}, \quad (4)$$

where  $\beta_i$  and  $t_i$  represent the coefficient and lag associated with predictor  $i$ , respectively. For supermarket 1 a best fit linear regression model was found with 12 predictors. The lags and coefficients associated with these predictors are given in Table 3. Lags and coefficients associated with all other DSUs are given in Appendix C.

#### 4.1.6 Additional methods and forecasting practices

Further techniques such as Ensemble Methods and Support Vector Machines are described in the literature but were not investigated by the study group due

Predictor $i$	$\beta_i$	$t_i$
1	0.32	336
2	0.23	672
3	0.22	1008
4	0.16	97
5	0.06	288
6	-0.10	433
7	0.10	717
8	-0.06	380
9	0.05	240
10	-0.07	865
11	0.06	193
12	0.04	626

Table 3: Coefficients and lags associated with the Stepwise Linear Regression (SLR) model for supermarket 1.

to time constraints. Neural Networks were studied, but the results were too inconsistent to be considered for what needs to be a reliable model.

Javed et al. (2012) discuss approaches to forecasting domestic load and note that weather in this sector has the strongest influence on demand. The use of Short Term Load forecasting models to predict individual household demand is investigated. Results are summarised using Multiple Linear Regression and Neural Networks. Data includes time series usage, anthropological and structural information. The issues of accuracy when forecasting at the individual house level compared to forecasting at the aggregate level is discussed, with aggregate level forecasting preferred.

AEIC (2009) describe various methodologies used in the dispatch verification process of demand response. Methods are proposed to estimate the difference between what the customer actually used and what that customer would have used if the dispatch call had not been issued, or the baseline. It is noted that new methods of establishing baselines and measuring their accuracies are constantly evolving as Smart Metering technologies evolve. Techniques used to calculate baselines are day matching, regression analysis and a proxy day approach. Day matching attempts to select a baseline day that most accurately matches the dispatch event day. Regression analysis involves using statistical regression methods to create a model that best matches. A proxy day is one that has the same characteristics as a dispatch event day; this is here examined in Section 6. A comparison of the baseline techniques is shown in Figure 19.

Measurement and verification of demand reduction schemes are also discussed in Goldberg and Agnew (2013). A number of recommendations are made about the characteristics that affect measurement and verification accuracy. Baseline adjustment methodologies are suggested. For example, in certain sectors weather has a strong influence on load so day-of-event adjustments should be made to remove this bias.

Furthermore, a number of recommendations are put forward to limit gaming opportunities of participants in demand reduction schemes. It is noted that only consumption can be metered directly, not reduction in consumption. However, forms of load simulation can be used to assess how well a particular baseline

Baseline Methodology	Pro	Con
Previous Day	<ul style="list-style-type: none"> <li>• Most likely the same usage pattern as the event day</li> <li>• Easy method for customer to understand</li> </ul>	<ul style="list-style-type: none"> <li>• Does not take into account the effects of weather on load</li> <li>• The need for a baseline adjustment</li> </ul>
Average Daily Usage	<ul style="list-style-type: none"> <li>• Easy method for customer to understand</li> <li>• Averaging takes out the variability in load for the days used to create the average day</li> </ul>	<ul style="list-style-type: none"> <li>• An average load shape created from multiple day load shapes will not totally capture the usage pattern for an event day</li> <li>• The need for a baseline adjustment</li> </ul>
Proxy Day	<ul style="list-style-type: none"> <li>• Matches a day based on defined variables uniform with event day</li> </ul>	<ul style="list-style-type: none"> <li>• Finding a day based on the defined variables</li> <li>• The need for a baseline adjustment</li> <li>• There might not be a day to use as the proxy day</li> </ul>
Regression Model	<ul style="list-style-type: none"> <li>• Concept of variable relationship is easy to understand</li> </ul>	<ul style="list-style-type: none"> <li>• Customer understanding of the process used</li> <li>• Selecting the correct variables to use in the model</li> </ul>

Figure 19: Baseline techniques, AEIC (2009)

method represents what would have happened in the absence of a demand reduction event.

## 4.2 Error Metrics

Error metrics are an important part of the analysis in both judging the strength of each model and providing evidence that forecasts become less accurate as the time horizon of a forecast increases. Three metrics are used here, with each having having the absolute percentage error at their core. For a prediction  $p_t$  of the demand at time point  $t$  in a forecast, the *absolute prediction error* is defined as:

$$\varepsilon_t = \left| \frac{p_t - a_t}{a_t} \right|,$$

where  $a_t$  is the actual demand at time point  $t$ .

Two metrics are used to compare the forecasting models. The first,  $E_{5\%}$ , is based on EirGrid's performance monitoring mechanism, which flags predictions if the absolute percentage error is greater than 5%. A dummy variable  $F_t$  is introduced to quantify this:

$$F_t = \begin{cases} 0 & \text{if } \varepsilon_t < 0.05 \\ 1 & \text{if } \varepsilon_t \geq 0.05 \end{cases}$$

The forecasting models predict demand every half hour for a period of 24-hours. The fraction of these half hour predictions that lie outside  $\pm 5\%$  for a forecast

on a given day  $d$  is

$$\frac{1}{48} \sum_{t=1}^{48} F_t^d$$

$E_{5\%}$ , the average of this fraction over the whole test period of  $N$  days, is the first metric used:

$$E_{5\%} = \frac{1}{N} \sum_{d=1}^N \left( \frac{1}{48} \sum_{t=1}^{48} F_t^d \right)$$

The second metric  $E_M$  is based on the *mean absolute percentage error (MAPE)*. On a given day  $d$ , the MAPE is defined as the average absolute percentage error over the prediction time period:

$$M^d = \frac{1}{48} \sum_{t=1}^{48} \varepsilon_t$$

$E_M$  is then the average of this quantity over the whole test data set.

$$E_M = \frac{1}{N} \sum_{d=1}^N M^d$$

Note that the two metrics are inherently different and should not be confused with each other.  $E_{5\%}$  gives the fraction of times that predictions in a forecast do not lie within 5%. On the other hand,  $E_M$  gives the average percentage inaccuracy of a forecast over the whole length of the forecast. It is a standard measure of accuracy for time series forecasting models.

Finally, a third measure  $E_{5\%}^T$  is introduced which is a modification of  $E_{5\%}$ . This metric is used to show that models become more inaccurate as they forecast further ahead.

At every half-hour  $h$  in the data set, of which there are  $N_h$ , a  $T$ -ahead prediction is made.  $E_{5\%}^T$  is the fraction of times the final point  $T$  in the forecast lies outside the  $\pm 5\%$  margin:

$$E_{5\%}^T = \frac{1}{N_h} \sum_{h=1}^{N_h} F_T^h \quad (5)$$

This metric is used in the analysis is to show how the accuracy of forecasts can be improved by reducing the time horizon  $T$  from 48 hours.

## 5. Results and Analysis

### 5.1 Forecasting Models

Each of the forecasting models outlined in Section 4 were compared by their ability to forecast under the current Eirgrid forecasting scheme. The models were tested on every one of the DSUs using the metrics  $E_{5\%}$  and  $E_M$ . The results are shown in Table 4.<sup>1</sup>

The results differ depending on the DSUs sector. The supermarkets are quite predictable, with the  $E_M$  values for the best models always falling between 3% and 10%. For the first set of supermarkets, the temperature dependent model outperforms the other models in the majority of cases. Interestingly, the only model that is better than it for some cases is the KIS weekly model. The fact that these two models outperform the ARIMA model with daily seasonality and the KIS median model indicates that weekly seasonality, along with temperature, is the important factor.

The models for the hospitals and the steel manufacturer also perform quite well. In the case of the former, the weekly models perform best, while the KIS median provides a good fit for the steel.

On the other hand, the demand of the bank and plastic manufacturer are more difficult to predict. The errors associated with these are quite large, with no one model providing good predictions across all the different customer segments.

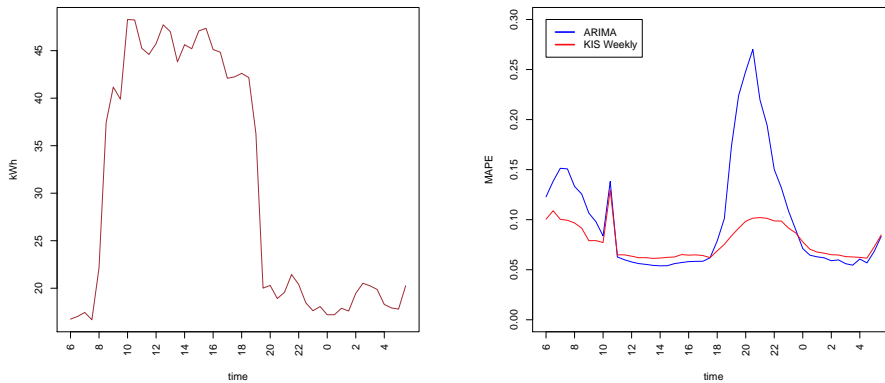


Figure 20: Left: Demand profile of supermarket 1 for a single day. Demand rapidly increases from 7.00a.m. to 10.00a.m. and rapidly decreases from 6.00p.m. to 9.00p.m. Right: The MAPE at every half hour from 6.00a.m. to 6.00a.m. averaged over all supermarkets throughout the test period for the ARIMA model (blue) and the KIS weekly model (red)

By analysing the error, it can be seen that there are periods throughout the day that are more difficult to predict. Figure 20 shows the average error for each half hour time point from 6.00a.m. to the following 6.00a.m. for both the

<sup>1</sup>Values listed for the ARIMA II model are averaged over the seven days of the week, daily accuracy under both error metrics is given in Tables 5 and 8.

DSU	$E_{5\%}$						$E_M$					
	ARIMA	KIS <sub>M</sub>	KIS <sub>W</sub>	T	ARIMA II	SLR	ARIMA	KIS <sub>M</sub>	KIS <sub>W</sub>	T	ARIMA II	SLR
Supermarket 1	0.60	0.61	0.56	0.50	<b>0.38</b>	0.52	0.13	0.10	0.12	<b>0.08</b>	0.07	0.09
Supermarket 2	0.55	0.39	0.45	<b>0.37</b>	0.45	0.41	0.14	0.09	0.12	<b>0.06</b>	0.06	0.08
Supermarket 3	0.52	0.59	0.51	0.46	0.84	<b>0.45</b>	0.12	0.13	0.13	<b>0.07</b>	0.08	0.08
Supermarket 4	0.63	0.78	0.52	<b>0.50</b>	0.70	0.55	0.12	0.15	0.12	<b>0.10</b>	0.10	0.10
Supermarket 5	0.47	0.43	<b>0.40</b>	0.42	0.48	0.42	0.10	0.08	0.09	0.07	<b>0.06</b>	0.07
Supermarket 6	0.56	0.86	0.46	<b>0.40</b>	0.61	0.48	0.14	0.18	0.14	<b>0.08</b>	0.09	0.09
Supermarket 7	0.67	0.73	0.53	<b>0.41</b>	0.54	0.52	0.16	0.14	0.15	0.09	<b>0.08</b>	0.10
Supermarket 8	0.60	0.68	0.51	<b>0.49</b>	0.69	0.51	0.17	0.16	0.17	<b>0.09</b>	0.09	0.10
Supermarket 9	0.68	0.90	<b>0.47</b>	0.51	0.59	0.52	0.16	0.25	0.14	0.10	0.11	<b>0.09</b>
Supermarket 10	0.53	0.83	<b>0.41</b>	0.43	0.51	0.51	0.15	0.17	0.15	<b>0.08</b>	0.08	0.10
Supermarket 11	0.57	0.63	0.49	0.47	<b>0.44</b>	0.49	0.13	0.12	0.12	0.08	<b>0.07</b>	0.08
Supermarket 12	0.52	0.47	0.43	<b>0.40</b>	0.48	0.45	0.14	0.12	0.12	<b>0.06</b>	0.08	0.08
Supermarket 13	0.53	0.69	0.53	<b>0.47</b>	*	0.48	0.08	0.09	0.08	0.08	*	<b>0.07</b>
Supermarket 14	<b>0.15</b>	0.50	0.26	0.47	*	0.27	<b>0.03</b>	0.05	0.04	0.08	*	0.05
Supermarket 15	0.44	<b>0.31</b>	0.32	<b>0.31</b>	*	0.34	0.08	<b>0.04</b>	0.05	0.06	*	0.06
Supermarket 16	0.56	0.41	<b>0.40</b>	0.43	*	0.46	0.12	<b>0.06</b>	<b>0.06</b>	0.09	*	0.09
Bank	*	0.80	<b>0.68</b>	0.76	*	0.83	*	0.72	<b>0.34</b>	0.42	*	0.39
Steel	*	0.40	0.53	<b>0.33</b>	*	0.61	*	<b>0.07</b>	0.12	0.14	*	0.10
Plastic	*	<b>0.76</b>	0.78	<b>0.76</b>	*	0.82	*	0.90	0.69	<b>0.47</b>	*	0.52
Hospital 1	*	0.95	0.39	<b>0.29</b>	*	0.48	*	0.38	0.11	<b>0.10</b>	*	0.10
Hospital 2	*	0.72	0.16	0.23	*	<b>0.14</b>	*	0.07	<b>0.04</b>	0.05	*	0.06

Table 4: The accuracy of each model for each DSU under both the  $E_{5\%}$  and  $E_M$  metrics. Each number represents the fraction of times the model is outside the  $\pm 5\%$  bounds. The models are the ARIMA model, KIS median (KIS<sub>M</sub>), KIS weekly (KIS<sub>W</sub>), temperature dependent (T), ARIMA II and stepwise linear regression (SLR) models. The most accurate model for each DSU for both metrics is highlighted in bold.



ARIMA and KIS Weekly models. This error is averaged over supermarkets 1 - 12.

The plot shows that certain periods throughout the day have higher forecasting errors. These periods are from 6.00a.m. to 10.00a.m. and from 6.00p.m. to 11.00p.m. Interestingly, these periods do not correspond to periods of high usage. Rather, they are start-up and shut-down periods when the supermarket opens and closes. Thus it is not the high usage period during the opening hours that is difficult to predict, but the start-up and shut-down periods.

The plots also indicate that while the models have very similar accuracy for most of the 24-hour period, the KIS Weekly model is superior at predicting the more volatile period. The spike in the ARIMA model line around 8.00p.m. shows that it is not effective at predicting the shut down period.

## 5.2 Improved accuracy for shorter time horizons

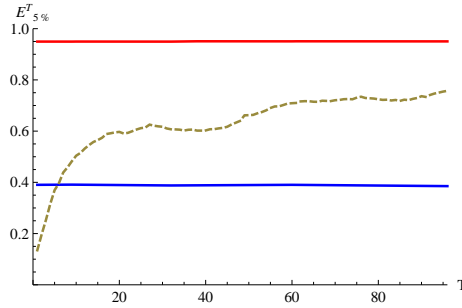


Figure 21:  $E_{5\%}^T$  as a function of the forecast time horizon  $T$  for one of the hospitals.  $T$  increases from 30 minutes ahead to 2 days ahead in 30 minute increments. The KIS weekly (blue) and KIS median (red) models accord no notable decrease in accuracy as the time horizon  $T$  increases. The ARIMA model (dashed mustard) does suffer significant loss of accuracy, especially during the initial increase in  $T$ .

The effect of the length of the forecasting time horizon  $T$  on the accuracy of the predictions was examined. The results for one of the hospitals is illustrated in Figure 21. The metric to assess the accuracy was the  $E_{5\%}^T$  metric as given in Equation 5 in Section 4.2.

It was found that  $T$  does not have a significant effect on the KIS models. This result is easily explained. The weekly KIS models are a function of the time point from one week ago, and so the prediction for a given point will be the same as long as the value of  $T$  is less than a week. For the median KIS model, the prediction for a time point is the median of that time point over the whole year, and so will be the same no matter how long ago  $T$  it is calculated.

On the other hand, the ARIMA model suffers a significant loss of accuracy as  $T$  increases. The initial decline of accuracy is quite sharp, yet after this it continues to decline but at a slower rate.

Similarly, the temperature dependent model will become more inaccurate as the time horizon increases. This is because the forecast temperature for the next half hour will be significantly more accurate than the forecast temperature

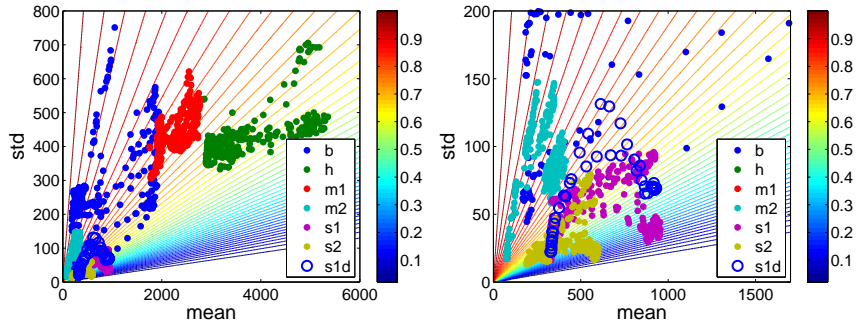


Figure 22: Contours of constant likelihood of exceeding the 5% precision threshold for a forecast based on a random sample from a Gaussian distribution with mean  $\mu$  and standard deviation  $\sigma$ . Markers indicated  $(\mu, \sigma)$  pairs calculated from weekday half-hourly data for each of the datasets. Right panel is an enlarged region of the left. The s1d data corresponds to *daily* half-hour averages and standard deviations of the s1 (supermarket) data.

in 2 days time. The model is fit to exact actual temperatures and so the error in the model will increase as the error in the forecast temperature increases.

### 5.3 Difficulties with 5% precision accuracy

Forecasting within the 5% accuracy required by Eirgrid is clearly difficult at the level of individual DSUs. From the data, it can be seen that the DSUs' energy usage is quite volatile and so when attempting to forecast, it may be just noise that is reproduced. The aim of the analysis was to quantify how often one might expect to exceed the 5% accuracy threshold if each weekday half-hour forecast was a random sample from a Gaussian distribution with the same mean  $\mu$  and variance  $\sigma$  as the corresponding data<sup>2</sup>. This provides another benchmarking test for our models and allows us to identify *a priori* particular datasets that would be difficult to forecast. The random forecast will exceed the 5% threshold less than  $n$  times per 48 forecasts (i.e. every 24hrs) if the inequality

$$2 \int_{1.05\mu}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx < \frac{n}{48}$$

is satisfied. Since the integral on the left-hand side is a function only of the parameter combination  $\mu/\sigma$ , contours of fixed  $n$  correspond to straight lines in  $(\mu, \sigma)$  space.

In Figure 22 contours of fixed  $n$  along with the various  $(\mu, \sigma)$  pairs for each weekday half-hour of each dataset are shown. From this figure, it can be seen that for the hospital (h) and steel (m1) data sets one would expect to exceed the 5% threshold more than 50% of the time. Even in the supermarket datasets, which are much easier to forecast, one would still expect to exceed the 5% threshold at least once a day.

<sup>2</sup>This assumes that the data is well approximated by a Gaussian distribution, which isn't necessarily the case, but the approximation is useful to provide insight into the source of the prediction error

## 6. Dispatch verification

The fundamental goal of EirGrid is to verify that dispatch occurred and that the correct amount of energy was dispatched. The current way of doing this is based on one day ahead forecasts as described in Section 1.2. DSUs submit a forecast, and if that forecast adequately remains within the  $\pm 5\%$  error bounds then the prediction is considered accurate. When a dispatch occurs, then verification that the amount dispatched is correct is easily calculated by subtracting the actual usage from the forecast usage during the dispatch period.

There are pros and cons of the current dispatch verification method. The main benefit to EirGrid is that the method inhibits potential gaming by DSUs. This gaming would be in the form of DSUs increasing their forecast during the dispatch period, making it appear that they dispatched more than they actually did.

As it stands, DSUs are required to submit their forecasts at 10a.m. on the previous day. DSUs are only made aware that they may be required for dispatch *after* the forecast has been submitted, in which case it is too late to alter their forecast.

The downside of the dispatch verification method is the difficulty for DSUs to give an accurate 48-hour forecast which remains within the  $\pm 5\%$  error bounds all the time. As explained in Section 5.3, most realizable forecasting methods will fall outside of the error bounds at least once during the forecast period, causing a DSUs to be flagged. This defeats the purpose of the current system to flag only DSUs with poor forecasts.

Alternative methods of verifying dispatch are known; for a review see AEIC (2009). Like the approach of EirGrid, they involve estimating a baseline which is what the DSU would have used had they not been called for dispatch. The dispatch is then verified by subtracting the actual use from the baseline.

One of these prediction methods briefly analysed here is the method of proxy days. Here, the half-hourly loads of a DSU on the dispatch day prior to dispatch are compared to loads on the same half-hours from previous days. These loads can be compared using various metrics, but in this case the correlation coefficient is used. A number of days with the highest correlation coefficients are chosen, and the average hourly consumption of these days is used to predict what would have happened during the dispatch period on the dispatch day. Note that a single proxy day can be chosen by picking only the day with the highest correlation coefficient.

Results of the Proxy Day formalism indicate that it could be a successful avenue to pursue. Figure 23 shows the actual usage of one DSU on 01/02/2012 and the proxy day and average daily mean prediction. To calculate this proxy day estimate, the 10 days whose usage correlates best with the usage on 01/02/2012 before 3.00p.m. are chosen. The average demand of the 10 chosen days *after* 3.00p.m. adjusted so that the average demand at 3.00p.m. exactly matches the actual demand at that time, is then used as the proxy day forecast for the rest of the day.

It can be seen that the proxy day forecast generated in this manner lies within  $\pm 5\%$  for 4 hours after the forecast. Recalling that dispatch can last for a maximum of two hours, this gives the required level of accuracy. Furthermore, this prediction can be made even more accurate by choosing just the single proxy day with the highest correlation coefficient (Figure 23). The prediction in this

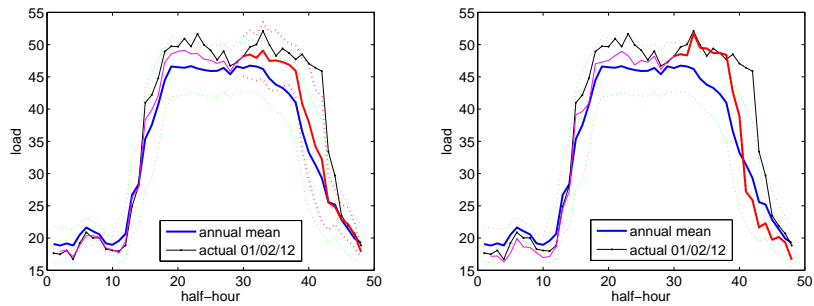


Figure 23: Proxy Day forecast (solid red line) made 3.00p.m. using average of top ten highest correlated proxy days (left) and highest correlated proxy day (right). The black line is the actual demand. The blue line is the demand for a half hour period averaged over all the year. The pink line is the average demand of the proxy day(s) prior to the forecast.

case is remarkably close to the actual usage for four hours after the forecast is made.

Note however that more detailed analysis is required. Further examination of the proxy day mechanism must be carried out, including the optimal number of proxy days to take and the number of hours prior to dispatch that should be taken to be correlated with other days.

## 7. Conclusions, Recommendations and Further Work

Three of the goals of the proposed project have been achieved. Firstly, models were built to forecast the usage of DSUs. Secondly, these models were used to show that the forecasting accuracy decreases as the time horizon of the forecast increases. Thirdly, it was shown analytically that the natural variability of usages leads to errors in forecast models that frequently fall outside the  $\pm 5\%$  margin, thus raising doubts of the current prediction flagging system.

A start was also made on alternate methods of dispatch verification. Provisional results on the proxy day method look promising, but further in-depth examination of this procedure will have to be carried out.

Suggested further work should focus on finding the optimal dispatch verification scheme. The current one day ahead forecast system of EirGrid should be compared to the proxy day and 6-hour rolling forecast methods. The latter is a method put forward by EirGrid and Electricity Exchange where forecasts are submitted every half hour on a rolling basis and the forecast time horizon is six hours. While this method opens the possibility for gaming by DSUs, it is hoped the short time horizon will give the opportunity for accurate forecast that are constantly inside the  $\pm 5\%$  margin, thus renewing the effectiveness of a flagging type system.

### Recommendations

In this report six different forecasting methods are presented, and their accuracy is compared using two different error metrics. Due to inherent stochasticity in demand there is no one forecasting method which is unequivocally best, but the ‘Keep it simple’ weekly and the temperature dependent models are identified as the most promising models to pursue.

Initial investigations suggest that a ‘proxy day’ mechanism may be preferable to the current method of verifying that the correct demand reduction takes place.

### Ideas for further work

Parameters for the ARIMA and SLR models were obtained using data from 2012. As time goes on this data will become outdated, and hence an area of further work is to update and refit these models using the most recent data. The statistical package R, contains a function called `auto.arima` which fits a Seasonal ARIMA model automatically. This could provide a fast automatic approach to fitting a forecast model for new data or new customers. In general, this approach will not produce as good a fit as a statistician who manually fits a forecast model. Currently, approximately 1 year’s worth of data is available for each DSU. When more data becomes available, yearly patterns, e.g., Christmas could be incorporated.

The correlation analysis in Section 4.1.4 (see further Table 2) showed that some days are more correlated to the same day the previous week versus the previous day. An area of future work could be the development of separate models for different days of the week.

Initial findings indicate that aggregating the 12 supermarkets may give a more aggregate forecast. However, a larger number of customers is required for

this analysis to be conclusive.

## **Acknowledgments**

The authors acknowledge support of the Mathematics Applications Consortium for Science and Industry ([www.macsi.ul.ie](http://www.macsi.ul.ie)) funded by the Science Foundation Ireland (SFI) Investigator Award 12/IA/1683, the SFI conference and workshop grant 13/CW/12578, and the Enterprise Ireland Innovation Voucher scheme.

## References

- AEIC. 2009. Demand response measurement and verification. White paper, AEIC. URL [http://www.naesb.org/pdf4/dsmee\\_group2\\_040909w5.pdf](http://www.naesb.org/pdf4/dsmee_group2_040909w5.pdf).
- Goldberg, Miriam L., G. Kennedy Agnew. 2013. Measurement and verification for demand response. Tech. rep., National Forum of the National Action Plan on Demand Response. URL <http://emp.lbl.gov/sites/all/files/napdr-measurement-and-verification.pdf>.
- Javed, Fahad, Naveed Arshad, Fredrik Wallin, Iana Vassileva, Erik Dahlquist. 2012. Forecasting for demand response in smart grids: An analysis on use of anthropologic and structural data and short term multiple loads forecasting. *Applied Energy* **96** 150 – 160. doi:<http://dx.doi.org/10.1016/j.apenergy.2012.02.027>. URL <http://www.sciencedirect.com/science/article/pii/S0306261912001213>.
- McClave, JT, FH Dietrich, T Sincich. 1997. Statistics, 7th edn .
- SEI. 2008. Demand side management in Ireland - evaluating the energy efficiency opportunities. Tech. rep., SEI.
- Yaffee, Robert A, Monnie McGee. 2000. *An Introduction to Time Series Analysis and Forecasting: With Applications of SAS® and SPSS®*. Access Online via Elsevier.

## 8. Appendix

### Appendix A: Stochastic DSUs

One of the DSUs in Electricity Exchange’s portfolio is a cement manufacturer. This DSU has energy consumption patterns that were neither correlated with time or with weather. This DSU is process driven, where energy demand can change in jumps by large amounts due to the turning on or off of a process. The large scale of the demand, between five and twenty MWh every half hour, also meant that failure to accurately predict this DSU meant that the aggregate prediction of all DSUs would be affected.

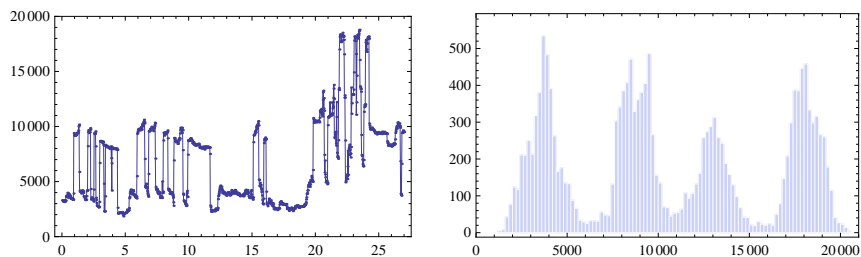


Figure 24: Left: Time series of the energy demand in kWh of the cement manufacturer taken at half-hour intervals over a period of 27 days. Right: Demand histogram for the full period of one year.

Figure 24 show the energy demand profile of the cement manufacturer. The stochasticity of the demand is clear. Large jumps in usage occur which are not correlated with time. Furthermore, the size of these jumps can be as much as 15 MWh. The usage histogram for the whole year helps to give an insight into this jumping process. Each data point falls into one of four clusters, which are hereby referred to as states. These states correspond to low usage (2-6MWh), medium usage (6-10MWh), high usage (10-15MWh) and very high usage (15-20MWh).

In forecasting ahead, the most important consideration is the state that will precede the current state, and when this switch will occur. Fluctuations within each state are of the order 100 kWh, which are insignificant in comparison to the large jumps between states.

Tables are constructed quantifying the probability of going to a state from the current state i.e.

$$p(X_j|X_i) \quad (6)$$

These give definite behaviour for the low usage and very high usage state. Medium usage will precede low usage with certainty, while high usage will precede very high usage with probability 0.96. However, the states following the medium and high usage states are inconclusive. This problem remains unresolved when the previous two or more states are considered, i.e.

$$p(X_j|X_i, X_k) \quad (7)$$

While the states following the high or medium states become more predictable, there is still a significant probability that they will not always go to these states, thereby removing the possibility of a deterministic model.



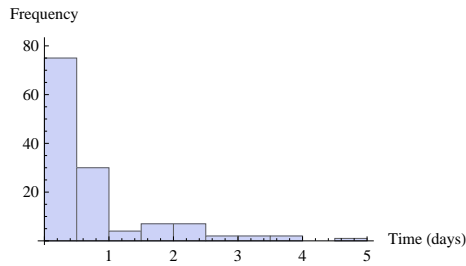


Figure 25: Histogram of the amount of time spent in the low usage state before switching state.

Further to the issues regarding predictions of preceding states, the time at which the process switches states is highly variable. The process can remain in a state for anything between 30-minutes to the order of days. Figure 25 shows the waiting times in the low state before switching. These values range from 30-minutes to 5 days, and follow an exponential distribution.

The conclusion of the analysis is that the cement manufacturer's demand profile cannot be accurately predicted with a deterministic model. Liaison is required between the manufacturer and Electricity Exchange to provide information about exact state switching times. These correspond to the turning on or off of a process. The inter state behaviour was not examined here, but it is hypothesized that ARIMA model could be suitable to model this behaviour.

## Appendix B: ARIMA II model

DSU	ARIMA II $E_{5\%}$						
	Mon	Tues	Wed	Thurs	Fri	Sat	Sun
Super 1	0.15	0.25	0.35	0.35	0.60	0.50	0.45
Super 2	0.40	0.30	0.20	0.55	0.70	0.35	0.65
Super 3	0.90	0.90	0.90	0.85	0.65	0.90	0.75
Super 4	0.90	0.80	0.75	0.70	0.70	0.60	0.45
Super 5	0.40	0.35	0.50	0.65	0.75	0.40	0.30
Super 6	0.75	0.45	0.45	0.75	0.80	0.70	0.40
Super 7	0.45	0.60	0.25	0.90	0.85	0.40	0.35
Super 8	0.80	0.65	0.75	0.70	0.70	0.80	0.45
Super 9	0.65	0.55	0.60	0.65	0.55	0.45	0.65
Super 10	0.45	0.50	0.25	0.85	0.50	0.55	0.50
Super 11	0.45	0.55	0.45	0.60	0.65	0.15	0.25
Super 12	0.55	0.45	0.40	0.50	0.55	0.50	0.45

Table 5: The accuracy of the same day previous week model for each DSU under the  $E_{5\%}$  metric.

DSU	ARIMA II $E_M$						
	Mon	Tues	Wed	Thurs	Fri	Sat	Sun
Super 1	0.07	0.07	0.07	0.07	0.07	0.07	0.07
Super 2	0.07	0.06	0.05	0.07	0.07	0.06	0.07
Super 3	0.09	0.09	0.08	0.08	0.09	0.07	0.07
Super 4	0.13	0.12	0.11	0.11	0.10	0.09	0.10
Super 5	0.06	0.06	0.06	0.06	0.06	0.05	0.05
Super 6	0.10	0.09	0.08	0.08	0.09	0.10	0.07
Super 7	0.09	0.07	0.07	0.07	0.08	0.08	0.08
Super 8	0.12	0.10	0.09	0.09	0.09	0.08	0.07
Super 9	0.10	0.09	0.24	0.09	0.09	0.08	0.08
Super 10	0.09	0.08	0.07	0.08	0.09	0.08	0.08
Super 11	0.07	0.07	0.07	0.06	0.07	0.06	0.08
Super 12	0.07	0.06	0.19	0.06	0.06	0.06	0.06

Table 6: The accuracy of the same day previous week model and the previous day model for each DSU under the  $E_M$  metric.

## Appendix C: Stepwise Linear Regression Model

Predictor $i$	$\beta_i$	$t_i$
1	0.42	336
2	0.23	1008
3	0.19	97
4	0.22	672
5	-0.13	433
6	0.08	239
7	-0.07	912
8	0.05	192

Table 7: Coefficients and lags associated with the Stepwise Linear Regression (SLR) model for supermarket 2.

Predictor $i$	$\beta_i$	$t_i$
1	0.25	336
2	0.30	97
3	0.10	1007
4	0.10	240
5	0.11	672
6	-0.04	436
7	0.06	192
8	0.08	288
9	0.08	1010
10	0.06	144
11	-0.04	915
12	-0.05	433

Table 8: Coefficients and lags associated with the SLR model for supermarket 3.

Predictor $i$	$\beta_i$	$t_i$
1	0.41	336
2	0.26	97
3	0.18	1008
4	0.06	189
5	-0.15	433
6	0.07	674
7	0.10	144
8	0.07	671

Table 9: Coefficients and lags associated with the SLR model for supermarket 4.

Predictor $i$	$\beta_i$	$t_i$
1	0.39	336
2	0.20	1008
3	0.19	97
4	0.19	672
5	0.09	190
6	-0.12	433
7	0.05	242

Table 10: Coefficients and lags associated with the SLR model for supermarket 5.

Predictor $i$	$\beta_i$	$t_i$
1	0.31	336
2	0.28	97
3	0.05	1007
4	0.17	672
5	0.10	240
6	-0.11	433
7	0.07	288
8	0.14	1008
9	0.05	668
10	-0.05	864

Table 11: Coefficients and lags associated with the SLR model for supermarket 6.

Predictor $i$	$\beta_i$	$t_i$
1	0.35	336
2	0.20	1008
3	0.27	97
4	0.05	670
5	0.18	240
6	-0.15	432
7	0.16	672
8	-0.05	577
9	0.04	524
10	-0.05	911

Table 12: Coefficients and lags associated with the SLR model for supermarket 7.

Predictor $i$	$\beta_i$	$t_i$
1	0.31	336
2	0.24	1008
3	0.25	97
4	0.07	239
5	0.11	671
6	-0.09	433
7	0.08	288

Table 13: Coefficients and lags associated with the SLR model for supermarket 8.

Predictor $i$	$\beta_i$	$t_i$
1	0.47	336
2	0.20	1008
3	0.25	97
4	-0.20	433
5	0.09	287
6	0.15	672
7	0.14	148
8	-0.11	820

Table 14: Coefficients and lags associated with the SLR model for supermarket 9.

Predictor $i$	$\beta_i$	$t_i$
1	0.32	336
2	0.26	672
3	0.24	97
4	0.07	1007
5	0.06	191
6	-0.09	768
7	0.09	287
8	-0.08	433
9	0.05	146
10	-0.07	481
11	0.10	1009
12	0.05	383
13	-0.08	337
14	0.05	144

Table 15: Coefficients and lags associated with the SLR model for supermarket 10.

Predictor $i$	$\beta_i$	$t_i$
1	0.36	336
2	0.15	1008
3	0.21	97
4	0.01	671
5	0.09	240
6	-0.16	433
7	0.12	1010
8	0.06	192
9	-0.06	912
10	0.15	672
11	0.04	812
12	0.06	288
13	-0.04	340

Table 16: Coefficients and lags associated with the SLR model for supermarket 11.

Predictor $i$	$\beta_i$	$t_i$
1	0.48	336
2	0.15	1008
3	0.22	97
4	-0.17	433
5	0.17	671
6	0.11	240
7	-0.07	575
8	0.10	1010
9	0.02	189
10	-0.06	863
11	0.11	192
12	-0.06	529
13	-0.04	341
14	0.06	775
15	-0.03	104

Table 17: Coefficients and lags associated with the SLR model for supermarket 12.

Predictor $i$	$\beta_i$	$t_i$
1	0.31	336
2	0.22	1008
3	0.24	97
4	-0.02	671
5	0.05	239
6	0.20	672
7	-0.07	769
8	0.03	862
9	-0.09	433
10	0.05	288
11	0.04	480
12	0.04	384

Table 18: Coefficients and lags associated with the SLR model for supermarket 13.

Predictor $i$	$\beta_i$	$t_i$
1	0.31	144
2	0.14	336
3	0.13	1008
4	0.38	97
5	-0.30	145
6	0.13	288
7	0.12	672
8	0.16	240
9	0.06	915
10	-0.09	242
11	-0.05	385

Table 19: Coefficients and lags associated with the SLR model for supermarket 14.

Predictor $i$	$\beta_i$	$t_i$
1	0.45	336
2	0.22	1008
3	0.05	144
4	0.17	672
5	0.14	97
6	-0.09	433
7	0.12	286
8	-0.08	622
9	0.10	192
10	-0.07	385
11	-0.07	190
12	0.02	909
13	0.02	723

Table 20: Coefficients and lags associated with the SLR model for supermarket 15.

Predictor $i$	$\beta_i$	$t_i$
1	0.38	336
2	0.24	1008
3	0.21	97
4	0.25	672
5	-0.06	434
6	0.12	287
7	-0.09	768
8	0.10	1056
9	-0.11	385
10	-0.05	623
11	0.03	145
12	-0.04	959
13	0.02	862

Table 21: Coefficients and lags associated with the SLR model for supermarket 16.



Predictor $i$	$\beta_i$	$t_i$
1	0.33	336
2	0.25	1007
3	0.24	672
4	0.30	97
5	-0.16	433
6	-0.11	769
7	0.13	144
8	-0.06	480
9	-0.08	814
10	0.03	290
11	0.07	722
12	-0.07	1057
13	0.05	135
14	-0.12	1096
15	0.06	425
16	0.05	713
17	0.10	1010
18	0.05	959
19	-0.05	628

Table 22: Coefficients and lags associated with the SLR model for the bank.

Predictor $i$	$\beta_i$	$t_i$
1	0.13	337
2	0.39	97
3	0.07	1005
4	-0.05	433
5	0.02	670
6	-0.07	766
7	0.11	289
8	-0.03	576
9	-0.11	1096
10	0.05	856
11	0.05	1050
12	0.08	130
13	-0.09	558
14	0.09	675
15	0.05	328
16	-0.08	156
17	0.08	1011
18	0.03	214
19	-0.03	422
20	0.08	344
21	-0.07	779
22	0.05	105
23	0.05	504
24	-0.06	446
25	0.10	669
26	0.05	256
27	-0.06	593
28	0.05	836
29	-0.04	799
30	0.05	660
31	0.08	333
32	0.07	114
33	-0.06	162
34	-0.05	406
35	0.03	733

Table 23: Coefficients and lags associated with the SLR model for the steel company.

Predictor $i$	$\beta_i$	$t_i$
1	0.30	336
2	0.25	1008
3	0.26	97
4	-0.16	433
5	0.09	293
6	-0.06	821
7	0.09	149
8	0.19	672
9	-0.04	629
10	0.11	165
11	-0.07	837
12	-0.05	486
13	0.07	332
14	0.07	1024
15	-0.07	449
16	0.10	308
17	-0.07	877
18	0.06	208
19	-0.07	644

Table 24: Coefficients and lags associated with the SLR model for the plastic company.

Predictor $i$	$\beta_i$	$t_i$
1	0.32	336
2	0.27	672
3	0.29	97
4	-0.13	433
5	0.20	1008
6	-0.14	769
7	0.12	191
8	-0.04	863
9	0.12	287
10	-0.11	1096
11	0.09	760
12	-0.07	912
13	0.07	385

Table 25: Coefficients and lags associated with the SLR model for hospital 2.

Predictor $i$	$\beta_i$	$t_i$
1	0.19	336
2	0.26	1008
3	0.20	672
4	0.05	97
5	-0.08	433
6	0.10	288
7	-0.01	960
8	0.05	335
9	0.05	1057
10	0.15	144
11	-0.12	816
12	-0.05	911
13	0.09	192
14	-0.08	529
15	0.10	337
16	0.05	720
17	-0.04	430
18	0.06	334
19	-0.07	436
20	0.09	99

Table 26: Coefficients and lags associated with the SLR model for hospital 2.

## Appendix D: Pseudocode

# Let  $y_t$  represent actual demand at time  $t$ .  
# Let  $\hat{y}_t$  represent predictions at time  $t$ .

### ARIMA Model

# Load data for DSU of interest

$$\hat{y}_t = \phi_1(y_{t-1} - y_{t-49}) + \phi_2(y_{t-2} - y_{t-50}) + \phi_3(y_{t-3} - y_{t-51}) \\ + \epsilon_t - \omega_1\epsilon_{t-1} - \theta_1(\epsilon_{t-48} - \omega_1\epsilon_{t-49}) + c$$

# To determine the coefficients  $\phi_1 - \phi_3, \theta_1, \omega_1$  and  $c$  for this type of model  
# use Table 1, or else use the  $R$  function  
# "model=arima(data,order=c(3,0,1),seasonal=list(order=c(0,1,1),period=48))".  
# Each  $\epsilon_t$  is determined from the standard normal distribution  
# If you wish to develop a new ARIMA model, i.e., one with different  
# orders to above, use the  $R$  function "auto.arima".  
# It may be appropriate to do this for new customers.

### ARIMA Weekly Model

# The same as above but if, for example, you're predicting a Saturday,  
# then load data for Saturdays only

### KIS Weekly Model

# Load data for DSU of interest

$$\hat{y}_t = y_{t-336}$$

# For each time point use the previous week's value

### KIS Median Model

```
# Load data for DSU of interest
for  $h = 1 \rightarrow 48$  do
    Arrange the values  $y_{h+48}, y_{h+96}, \dots, y_{h+48N}$  in ascending order.
    if  $N$  is odd then
        |  $\hat{y}_h = ((n+1)/2)^{th}$  value in this ordered set
    end
    else
        |  $a = ((n+1)/2)^{th}$  value in this ordered set
        |  $b = (n/2)^{th}$  value in this ordered set
        |  $\hat{y}_h = \frac{a+b}{2}$ 
    end
end
# This model uses the median demand, over a number of previous days  $N$ ,
# to predict each half-hour time point.
```

### Temperature dependent model

```
# Load data for DSU of interest.
# Let  $T_t$  represent temperature at time  $t$ .
# Load temperature data.
 $\hat{y}_t = y_{t-336} + \alpha(T_t - T_{t-336})$ 
# where  $\alpha$  is determined using ordinary least squares on all available
# historical temperature data, using the 'polyfit' or 'regress' function in
# MATLAB.
```

### Aggregation

```
# To aggregate forecasts, repeat the above for any of the different formula
# and sum up predictions.
```

### Error metrics

```
# For more details on error metrics, see Section 4.2.
```

$$\epsilon_t = \frac{|y_t - \hat{y}_t|}{y_t}$$

$$F_t = \begin{cases} 0 & \text{if } \epsilon_t < 0.05 \\ 1 & \text{if } \epsilon_t \geq 0.05 \end{cases}$$

```
# Let  $N$  represent number of days
```

```
# Eirgrid metric
```

$$E_{5\%} = \frac{1}{48 \times N} \sum_{t=1}^{48 \times N} F_t$$

```

# MAPE metric
 $E_M = \frac{1}{48 \times N} \sum_{t=1}^{48 \times N} \epsilon_t$ 
# Time point average The index  $h$  represents half hour period
 $M^h = \frac{1}{N} \sum_{i=h}^N \epsilon_{i+48}$ 

# Rolling metric Let  $T$  represent the length of forecast period
 $E_{5\%}^T = \frac{1}{T} \sum_{t=1}^T F_t$ 

```

**Proxy day code**

```

# let  $y_t^d$  be the demand on day  $d$  at time  $t$  where  $y_1^d$  is demand at 12:00 AM
# let  $t = \hat{h}$  be the beginning of the dispatch for day  $D$ 
if dispatch occurs then
    # find correlation ( $r_{D,d}$ ) between actual demand on day  $D$  and all days
    # from  $d = 1$  to  $d = D - 1$  using time periods from  $y_1$  to  $y_{\hat{h}-1}$  only.
    for  $d = 1 \rightarrow D-1$  do
        |
        |
        | 
$$r_{D,d} = \frac{\sum_{t=1}^{\hat{h}-1} (y_t^D - \overline{y^D})(y_t^d - \overline{y^d})}{\sqrt{\sum_{t=1}^{\hat{h}-1} (y_t^D - \overline{y^D})^2 \sum_{t=1}^{\hat{h}-1} (y_t^d - \overline{y^d})^2}} \quad (8)$$

        |
        |
    end
    #  $\overline{y^d}$  represents the average demand on day  $d$  over the time period
    # from  $t = 1$  to  $t = \hat{h} - 1$ .
    # The day with the largest correlation ( $r_{D,d}$ ) is the proxy day.
    # There are in built functions in both MATLAB and R to
    # calculate Equation 8.
end

```

**Algorithm 1:** Pseudocode for algorithms

```

analyse_all.m
#####
## Data about data
#####
## Names of data files
names={"supermarket1.txt";
      "supermarket2.txt";
      "bank.txt";
      "manufacturing1.txt";
      "manufacturing2.txt";
      "hospital.txt"};
## Numbers of columns in each file
sizes=[12; 4; 1;1; 1; 2];

```

```

## Output files
tablefile=fopen("fitting.dat","w");
modelfile=fopen("models.dat","w")

## Some constants
day=48; week=7*48; ## Number of half hours in a day and week
N=5000; M=1000; ## Size of datasets to fit
##N=10; M=5;
#####
## Main Program Loops
#####
## Loop over datafiles
for ii=1:length(names)
    ## load data
    filename=names(ii){1};
    aaa=load(filename);
    ## Reverse entries
    aaa=aaa(end:-1:1,:);
    for jj=1:sizes(ii)
        ## extract column of data
        y=aaa(:,jj);
        ## Split into parts
        ## Create matrices to hold data
        y1=ones(N,1); xxx1=ones(N,M); ## Model is fitted with this data
        y2=ones(N,1); xxx2=ones(N,M); ## Cross validation: goodness of fit
        ## is calculated with this dataset
        y3=ones(N,1); xxx3=ones(N,M); ## Spare.
        ## Fill matrices
        for i=1:N
            y1(i,1)=y(3*i-2);
            y2(i,1)=y(3*i-1);
            y3(i,1)=y(3*i-0);
            for j=1:M
                xxx1(i,j)=y(3*i-2+2*day+j); ## xxx's contain data delayed by 2days
                xxx2(i,j)=y(3*i-1+2*day+j); ## 2days + 1/2 hour, 2days+2*1/2 hour,
                xxx3(i,j)=y(3*i-0+2*day+j); ## ..., 2days+M*half hour. (M
                ## M is defined above.
            endfor
        endfor
#####
## Fit model to data by calling code in stepwise_regression.m
#####
    stepwise_regression
#####
## Calculate goodness of fit measures
#####
    ## Fraction of fitted results INSIDE +/-5% envelope
    ## Note it is fraction of results OUTSIDE envelope written to file.
    five_percent=sum((y2fit>0.95*y2).*(y2fit<1.05*y2))/length(y2);
    ## Mean absolute fractional deviation

```



```

    mean_abs=mean(abs(y2fit-y2)./y2);
#####
## Write results to file
#####
    fprintf(tablefile,"%20s %3d %f %f\n", filename,jj,1-five_percent,mean_abs)
    fprintf(modelfile,"=====\n")
    fprintf(modelfile,"%s\n",filename)
    fprintf(modelfile,"%20s %20s %20s\n", "Coefficient",
    "Delay (1/2 hour)", "Delay (day)")
    for k=1:length(b)
        fprintf(modelfile, "%20.10e %20d %20.6f\n",
        b(k),2*day+x_index(k),2+x_index(k)/day)
    endfor

    endfor
endfor

## Close output files
fclose(tablefile);
fclose(modelfile);

```

```

stepwise_regression.m

xx=xxx1; ## Holds columns not used in the model, initially full

xx_index=1:size(xx)(2); ## Holds indices of columns not used in model

disp("Stepwise linear regression with cross validation")

x=[]; x_index=[]; ## Columns and indices of columns used in model
                    ## initially empty
chi_sq=[];

## Keep fitting as long as fitting extra columns of data
## to dataset 1 yeilds a model which reduces chisq
## for dataset 2.
fitting=true;
while(fitting);
    chi_sq_test=[];
    for i=1:size(xx)(2)
        ## Generate a trial model by adding each unused column to the data
        ## for each trial model calculate fitting parameters and chisq
        ## using dataset 1
        x_trial=[x,xx(:,i)];
        a=x_trial\y1;
        y1fit=x_trial*a;
        this_chi_sq=sum( (y1-y1fit).^2 );
        chi_sq_test=[chi_sq_test; this_chi_sq];
        ##disp([100,i,this_chi_sq])
    endfor
    ## Find the trial model with the lowest chisq
    [b,k]=min(chi_sq_test);
    ## Add the new column of data and its index to the model
    x=[x,xx(:,k)]; x_index=[x_index,xx_index(k)];
    ## Remove the column of data and its index from the unused data arrays
    xx(:,k)=[]; xx_index(k)=[];
    ## Calculate fitting parameters for the new model
    b=xxx1(:,x_index)\y1;
    ## Calculate fitted values anc chisq values for dataset 2
    y2fit=xxx2(:,x_index)*b;
    this_chi_sq=sum( (y2-y2fit).^2 );
    chi_sq=[chi_sq;this_chi_sq];
    bounded=sum((y2fit>0.95*y2).*(y2fit<1.05*y2))/length(y2);
    disp([chi_sq(end),bounded])
    ## If the new model has reduced chisq for dataset 2
    ## carry on fitting
    ## otherwise set fitting to false to break ot of fitting loop.
    if (length(chi_sq)>1)
        fitting=(chi_sq(end)<chi_sq(end-1));
    endif
endwhile

```

```
endwhile
## Remove last index which increased chisq
x_index(end)=[];

## Recalculate model (without discarded last index
b=xxx1(:,x_index)\y1;
y2fit=xxx2(:,x_index)*b;
## Plot results.
plot(y2,y2fit,'r+',y2,0.95*y2,'b+',y2,1.05*y2,'b+')
sum((y2fit>0.95*y2).*(y2fit<1.05*y2))/length(y2)
```