2017

Short Term Demand Forecasting for the Integrated Electricity Market

Katie Kavanagh
Dublin Institute of Technology, d15125143@mydit.ie

Follow this and additional works at: http://arrow.dit.ie/sjer

Part of the Electrical and Computer Engineering Commons

Recommended Citation
doi:10.21427/D75G90
Available at: http://arrow.dit.ie/sjer/vol2/iss1/1
Short Term Demand Forecasting for the Integrated Single Electricity Market

Katie Kavanagh
School of Electrical and Electronic Engineering
Dublin Institute of Technology
Dublin, Ireland

Corresponding Author: Katie Kavanagh, D15125143@mydit.ie

ABSTRACT: This paper presents a means for the short term load forecasting (STLF) of electricity. The forthcoming Integrated-Single Electricity Market (I-SEM) diverges from the current market structure (the Single Electricity Market or SEM), with significant impacts on Irish supply companies, creating a need for these companies to be able to accurately forecast their customers’ load in the Day Ahead. Using a Double Seasonal Exponential Smoothing variation of the Holt-Winters method that factors in an error correction, data from the Irish market was trained and used to forecast a supply company’s demand resulting in an average daily MAPE (Mean Absolute Percentage Error) of 2.99% over a period of nearly four weeks. The suite of formulas used employs daily and weekly seasonal components to forecast a full day’s (48 half-hour periods) demand.

KEY WORDS: Holt-Winters, Exponential Smoothing, Electricity Demand Forecasting, Short Term Electricity Forecasting, I-SEM

1 Introduction

In May of 2018, the current electricity market of Ireland (all-island), the Single Electricity Market or the SEM, is set to undergo changes and become the Integrated Single Electricity Market or I-SEM (Gaelectric, 2016). There are many differences in the structure of the SEM and the I-SEM; this paper addresses the change of balance responsibility and what this will mean for Irish supply companies in terms of forecasting their customer loads. A means for accurately forecasting short term electricity demand, suitable to the Irish market, will be presented and critiqued.

In this context, short term load forecasts refer to periods ranging from a one hour lead time to a week ahead. This research considers forecasting loads for half hourly lead times for a day (24hrs) ahead.

1.1 The I-SEM

The new I-SEM market will be based around voluntary Day Ahead and Intraday markets in electrical energy. The I-SEM and the European rules that serve as its foundation call for the creation of a number of new roles and responsibilities as well as changes in the responsibilities of a number of existing license holders in the SEM. These parties include the Market Operator (MO), Transmission System Operators (TSOs) and market participants.

This new market structure was developed as a means for creating a larger, more competitive marketplace and for the purpose of establishing an electricity market that is conducive with the European internal market, the European Target Model (ETM) and the corresponding EU goals. Such goals revolve around the aims of eventually creating a level playing field for the efficient use of cross-border capacity and of harmonised EU electricity markets.

This new market, while fulfilling requirements to be on track toward the ETM, will answer issues that have risen from the existing market, as a result of changes in demand, generation and interconnection.

The SEM Committee has assessed that the I-SEM High Level Design (HLD) will best deliver the benefits of European market integration in terms of:

- security of supply;
- promotion of renewable energy sources;
establishment of a level playing field in which competition can flourish;
maximising the efficient use of interconnectors; and
provision of a sound investment climate that is based upon a stable and predictable regulatory framework.

(SEM Committee, 2014, Decision on High Level Design)

1.2 What this Means for Participants
Under the structure of the I-SEM, there will be multiple markets running across different timeframes to accommodate trading. These markets include a Forwards market, a Day-Ahead market, an Intraday market and a Balancing market and imbalance settlement. As a generator’s or supplier’s position is determined in the single market structure Ireland has now, their disposable position will be determined by their trading activity across the markets of the I-SEM. In the new market, all parties (generators and suppliers) will be balance responsible. This means that the respective parties will be responsible for any discrepancies between the Day-Ahead market and Intra-Day market trades they make, and the actual metered quantities they produce or take. This results in mandatory participation in the Balancing market and any subsequent imbalance settlements to fine-tune committed production and consumption trades. Notably, generators and suppliers will want to manage their bid quantities in the different markets, in order to avoid exposure to the price differences between the Day-Ahead and Imbalance markets. To do so, these companies will have to take an as accurate as possible position in the Day Ahead market (SEM Committee, 2014). This will require a reliable forecasting method for supply companies.

A literature review was carried out to identify a suitable forecasting method for a supply company in the I-SEM. Many methods were identified, but a variation of the Double Seasonal Holt-Winters Exponential Smoothing with Error Correction was deemed likely to be the most suitable and was tested. This is now discussed in detail.

2 Literature Review
Many methods of forecasting electricity exist. Due to the nature of electricity and the factors affecting its demand, only short-term load forecasting (STLF) techniques were considered. There is a long history of STLF and a large amount of literature on the subject to accompany it. Many reviews have been conducted considering markets around the world. Methods include soft computing techniques as well as statistical methods; however given the depth of literature in the broader area, this review is limited to statistical methods of STLF.

One such review (Abu-El-Magd and Sinha, 1982) looks at and compares the drawbacks and merits of multiple STLF methods. In the review, a Multiple Regression Approach, Spectral Decomposition, Exponential Smoothing Method, Stochastic Time Series Approach, State Space Approach and Multivariable Load Modelling Approach are examined. Time Series and State Spaces approaches were noted as the most popular at the time, while models such as Multiple Regression and certain Time Series methods were highlighted for requiring long, time-consuming analysis.

More recently, a method review (Alfarez and Nazeeruddi, 2002) looked at techniques under the categories of:

- Multiple Regression;
- Exponential Smoothing;
- Iterative Reweighted Least-Squares Models;
- Adaptive Load Forecasting;
- Stochastic Time Series;
- AMAX based of Genetic Algorithms;
- Fuzzy Logic;
- Neural Networks; and
- Knowledge-Based Expert Systems.

A preference for Fuzzy logic and ANN techniques was concluded, as well as a move towards hybrid approaches. This work also explicated that time series techniques are widely used. The authors also found that exponential smoothing techniques compare well to the more conventional techniques described.

Similarly, further work in the area (Singh, 2012) states that regression is widely used in this context since few parameters are required and because load is easily predicted using previous load data. Multiple regression is stated as the most prevalent of all the traditional techniques since it has the ability to capture a large number of factors affecting load.

In line with this, more current work (Fahad and Arbab, 2014) states that time is the most important factor in short term load forecasting due to its high impact on consumer load; namely that it has “time of day” and “day of week” properties, or seasonalities.
Significant work in the area has been undertaken by Taylor. His work centres mainly on univariate methods of STLF. Using MAPE to measure results, Taylor found that on a national level of demand in England and Wales, a Double Seasonal Exponential Smoothing model gave a MAPE of 0.4 – 1.2%; a Neural Network model resulted in 0.4 – 2.1%; Double Seasonal ARIMA in 0.4 – 1.6%; Regression method in 0.5% - 1.4%; Seasonal Random Walk in 0/4 – 2.25% and Error Modelling Seasonal Random Walk in 0.5 – 1.7%. (Taylor, 2006).

This work was built on to further authenticate the findings by both Taylor and McSharry (Taylor, 2007) and Gould, (Gould, 2007). The new method of exponential smoothing by Gould considers different intraday cycles for different days as a quality of the load, and captures seasonal differences between weekdays and weekends.

Given the positive results of Taylor’s findings outlined above, and the findings of a 2003 paper that are conducive to the data provided (Taylor and Buizza, 2003), that weather data is not required for STLF, the Double Seasonal Holt-Winters Exponential smoothing and the Double Seasonal Holt-Winters Exponential Smoothing with Error Correction methods were selected for this research.

3 Methodology

In order to establish if the Double Seasonal Holt-Winters with Error Correction method is suitable to the Irish market, the method was applied to a supplier’s data.

3.1 Data Description

An Irish supply company’s actual load data for 15 months – from 1st January 2013 to 31st March 2014 – was supplied for this research. This specific supply company deals with 53 commercial and industrial customers and does not work with residential customers.

Nearly four full weeks, 27 days exactly, were used to forecast this method. Five weeks of data were used for parameter estimation and initialisation. The data supplied was issued in kilowatt hours and was reported in half-hourly loads, grouped into blocks of 48 half-hours (comprising a full day) for the 15 months. All data in this paper are in kilowatt hours. The total load for the 15 months was 444,241,529.43 kWh. The total load for the 27 days of forecasts was 27,191,885.67 kWh. The mean half-hourly load for the 27 days was 20981.39 kWh.

While testing this method, it was important that forecasts be made for a period of time free of any bank holidays or special days. For this reason, forecasts were projected from Tuesday, 1st October 2013 to Sunday, 27th October 2013. The 28th was the October Bank Holiday in 2013. Parameter estimation and initialisation took place in the five weeks prior, starting on 26th August 2013.

The load on bank holidays is irregular as demand is generally less. This holds true especially for commercial and industrial consumers whose businesses would typically be closed on a bank holiday. For example, this supply company had a total demand of 650,644.43 kWh on the October bank holiday Monday in 2013 (the 27th). The previous three Mondays (the 21st, 14th and 7th) had total demands of 1,010,526.16, 972,528.66 and 1,049,213.24 kWh, respectively. Figure 1 below shows how the October bank holiday’s load compared to the previous Mondays’.

In Figure 1, one can see that the demand on a bank holiday is flatter than a typical day, showing less inclination towards peaks and troughs in the demand. The demand is fairly consistent throughout the day. It is also evident that the overall load for the day was less than a typical Monday of that time of the year.

3.2 Data Analysis

The data provided shows the seasonalities described in the Literature Review. There are clear patterns on a weekly cycle and on a daily cycle. Throughout the
course of a week, each day’s load is clearly outlined by starting in a trough at 00:00, reaching a peak during the middle of the day, and descending back into a trough with the deepest point around 23:30. This is illustrated below in Figure 2.

Figure 2: Historical Weekly Customer Load, Saturday to Friday

Figure 2 shows that weekly, the load moves in peaks and troughs that hold a fairly smooth and consistent shape from Monday to Friday. This pattern is representative of an average week for this load portfolio.

Looking more closely, it is evident in Figure 3 that on a weekday (Monday – Friday) basis, the load of the days goes through its own gentle trough and peaks.

The Sunday night trough that Monday morning picks up on is much lower than any of the other weekdays. Similarly, the peaks of each day build up between Monday and Wednesday, and then recede slightly from Wednesday to Friday. The daily maximum load increased from 27,140.68 kWh on Monday, to 28,883.73 on Tuesday, 29,529.35 on Wednesday and then decreased to 29,343.13 on Thursday, down to a weekday peak load low of 28,924.55 on Friday. Individual daily load profiles are shown in Figure 4.

Figure 3: Historical Week 14/10/13 - 18/10/13 Daily Loads

The data shows daily peak load hours of 08:00 – 17:00, with the peak half hour load typically occurring at 12:00. The data demonstrates slight dips during the peak hours at 10:00 and again around 13:00, before finally making the decent for the evening at 15:30.

Saturday and Sunday also start in a trough, increases to a peak and descends back into a trough, but typically does this in a far less consistent and pronounced manner. As illustrated in Figure 5, sometimes outliers occur in the data. For instance the sharp peak on Sunday the 27th of October 2013 from 01:00 – 02:00 is uncharacteristic of the data. The load jumps from 13,086.72 kWh at 00:30 to 19,304.42 at 01:00, climbing to 25,879.91 at 01:30 before declining to 19,372.02 at 02:00 and finally falling back to a more typical load of 12,865.92 at 02:00. It is unknown what caused this sharp spike, but with a closer look at the company’s customer portfolio and weather data from the day an answer could be produced.

Figure 4: Historical October Weekend Loads
Most often, as demonstrated by the first three weekends in the forecast period, the data presents a higher peak on Sunday than on Saturday. The same holds true for the troughs, with a higher load Saturday night/Sunday morning than on Sunday night/Monday morning. The maximum load on a Saturday for the four weeks was 19,581.04 kWh, while the maximum load for a Sunday in the same period – excluding the sharp spike – was 18,894.39. The minimum load for a Saturday was 12,969.98 at 23:00 and 11,632.70 for a Sunday during the same period, at 23:30.

The peak hours during the weekend occur from 09:00 to 15:00 on a Saturday and from 11:00 to 06:00 on a Sunday. The lowest points in the troughs are typically between 23:00 and 06:00.

The load profiles show clear time factors in the form of daily and weekly patterns. Similarly, it demonstrates that these cycles are mostly predictable with the exception of bank holidays. Forecasting for special days or bank holidays is outside the scope of this research. Given the weekly and daily patterns, a forecasting method taking the cyclical patterns into consideration is thought best.

### 3.3 Collating Data

For this particular research, the start of the week has been set on Tuesday to accommodate the initialisation and estimation periods. There are 48 half-hour periods per a day, and 336 half-hour periods per week.

### 3.4 Model Selection

The data provided for this research is historical data. Given this, a time series method reliant on historical data to estimate the model parameters is appropriate.

As discussed in the Literature Review, recent studies in the area demonstrate that the Holt-Winters Exponential Smoothing method of forecasting tends to have positive results (less than 3% MAPE) for data demonstrating trend and seasonal components. This method produces a smoothed time series by using weighted averages of the past data and assigning decreasing importance to the observations as they get older. As this data demonstrates trend and seasonality, it was decided to apply the Double Seasonal Holt-Winters Exponential Smoothing method, and the Double Seasonal Holt-Winters Exponential Smoothing method with Error Correction, which builds upon the latter method. A Naïve Benchmark was also included for comparison.

#### 3.4.1 Double Seasonal Holt-Winters Exponential Smoothing

While a standard Holt-Winters method can be used to forecast seasonal time series, it only accounts for one season. Taylor built on this work with the Double Seasonal Holt-Winters Exponential Smoothing method.

As demonstrated by Taylor (Taylor, 2003), the Holt-Winters method for double multiplicative seasonality is given by equation (1) – (5):

- **Level**
  \[ S_t = \alpha \left( \frac{X_t}{D_{t-s1}W_{t-s2}} \right) + (1-\alpha) (S_{t-1} + T_{t-1}) \] (1)
- **Trend**
  \[ T_t = \gamma (S_t - S_{t-1}) + (1-\gamma) T_{t-1} \] (2)
- **Season 1**
  \[ D_t = \delta \left( \frac{X_t}{S_t W_{t-s2}} \right) + (1-\delta) D_{t-s1} \] (3)
- **Season 2**
  \[ W_t = \omega \left( \frac{X_t}{S_t D_{t-s1}} \right) + (1-\omega) W_{t-s2} \] (4)

- **Equation for Forecasting**
  \[ X_t (k) = (S_t + kT_t) D_{t-s1+k} W_{t-s2+k} \] (5)

Where \( \alpha, \gamma, \delta \) and \( \omega \) are smoothing parameters, and \( S_1 = 48 \) and \( S_2 = 336 \).

The two seasonal indices are \( D_s \) daily, and \( W_w \) weekly.

The first \( s_1 \)-period seasonal index of the data set, \( D_s \), is estimated by smoothing the ratio of observed value, \( X_t \), to the product of the local level, \( S_t \), and local \( s_2 \)-period seasonal index, \( W_{s_2} \). The first \( s_2 \)-period seasonal index, \( W_s \), is estimated by smoothing the ratio of observed value, \( X_t \), to the product of the local level, \( S_t \), and local \( s_1 \)-period seasonal index, \( D_{s_1} \). These are first calculated and then the formulas are applied to the data to prepare them for forecasting.

The initial Level (\( S_0 \)) and initial Trend (\( T_0 \)) were calculated using the first two weeks of the estimation period for a total of 672 (half-hour) observations. The average of the first 336 (week one) observations were calculated. Then the differences (Demand_{s_1} – Demand_{s_2}) between the first 336 observations were calculated. Following this, the differences between week two observations were found. Finally, the average of the week one differences and the average of week two differences were computed (separately for each week). Initial trend \( T_0 \) was equated to (1) \( \frac{1}{2} \) of the average of week one differences plus the average of week two differences and (2) the average of the first
differences for the first 336 observations. The initial level, $S_0$, was selected as the mean of the first 672 (week one & week two) observations minus 336.5 times the initial trend.

The first 336 observations were used to find the initial within-day seasonal indices ($D_t$), while the first 672 observations were used to estimate the initial within-week seasonal indices ($W_t$). The initial values for the $D_t$ were calculated as the ratio of the actual observation to the 48-point centred moving average. The initial value for $W_t$ was calculated as a ratio of the actual observation to the 336-point centred moving average.

Once the data was initialised, a parameter estimation period was run to further improve the accuracy of the forecasts. The parameter estimation period was run for four weeks. During this time, $k$ was set equal to 1. For the actual forecast, $k$ was set equal to the corresponding half-hour of forecast.

The parameters were estimated by using 00.30 as an initial value for the smoothing parameters. These were then estimated by calculating the sum of squared errors (SSE) for the four weeks of parameter estimation forecasts. Then, the Solver add-on of Excel was used to minimise (GNG Nonlinear method) the SSE through the smoothing parameters, subject to the smoothing parameters being greater than 0 but no more than 1. With these resulting parameters, the Day-Ahead forecast was made.

3.4.2 Double Seasonal Holt-Winters Exponential Smoothing with Error Correction

The Double Seasonal Holt-Winters Exponential Smoothing method was further built on by adding an Error Correction.

This method can only be carried out once forecasts have been determined in the parameter estimation phase using the Double Seasonal Holt-Winters Exponential Smoothing method (detailed above) and once the errors have been calculated in this phase. The error term $\lambda e_t$ (see Equation 6) is then applied to the same estimation period and the model parameters recalculated again using the Solver GNG Nonlinear method. This new formula, taking previous errors into consideration, is used to make the actual forecasts.

The Error Term is applied to the forecast as (6):

$$\hat{X}_t (k) = (S_t + k_1) D_{cal+k} W_{cal+k} + \lambda e_t$$

where $s_2$ is equal to 336 (the number of half-hours in the weekly cycle) and $e$ is the error term.

3.5 How Results Were Measured:

For all three methods (including the Naïve Benchmark) results were measured by calculating the MAPE (Mean Absolute Percentage Error) for each of the 27 days in the test period.

In this method, the SSE of the actual Demand minus the Corrected Error Forecasts was minimised and an additional smoothing parameter $\lambda$ was incorporated.

3.4.3 Naïve Benchmark

A Naïve Benchmark was selected for this research to see how the forecasting methods compare. Based off of prior work in the area, a practical and easy to apply method was used. The Naïve Benchmark used (Equation 7) accounts for the seasonality by using averages. Although two seasonalities have been identified in the data set, this method only considers one seasonality. The within week seasonality was considered since it covers a longer period of time. The forecast is calculated by averaging the data for the corresponding half hours in the four previous weeks, and then adding in the error of the previous half hour.

$$\hat{X}_t (k) = (y_{t+k-2s} + y_{t+k-2s} + y_{t+k-3s} + y_{t+k-4s}/4) + e_{t+k-1}$$

Where $s_2$ is equal to 336 (the number of half-hours in the weekly cycle) and $e$ is the error term.

The Naïve Benchmark equates to the sum of the Load at t-336, t-672, t-1008 and t-1344, divided by four, plus the Error of the Corrected Forecast at t-1.

From literature, it was established that the accuracy of a next day forecast should be within one and three percent to constitute a good result (Feinberg and Genethliou, 2005).
4 Results and Discussion

Testing the three methods (including the Naïve Benchmark) on the data resulted in the best MAPE for the Double Seasonal Holt-Winters Exponential Smoothing with Error Correction method. The methods and resulting MAPEs are listed below in Table 1.

<table>
<thead>
<tr>
<th>Method</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS HWES</td>
<td>8.20%</td>
</tr>
<tr>
<td>DS HWES EC</td>
<td>2.99%</td>
</tr>
<tr>
<td>Naïve Benchmark</td>
<td>3.56%</td>
</tr>
</tbody>
</table>

Table 1: Method Results by MAPE

As expected, the Naïve Benchmark yielded a satisfactory result, while the Double Seasonal Holt-Winters with Exponential Smoothing gave an acceptable result, as defined in the Methodology, i.e. between 1 and 3%. The Double Seasonal Holt-Winters method resulted in an undesirable forecast with a comparatively high MAPE.

Figure 7 shows the forecast methods in comparison to the actual load for the month. Figure 8 shows a closer examination of the results for the first week. The graphs show that certain times are less accurate for particular methods, such as the Double Seasonal Holt-Winters with exponential Smoothing method during peak loads. However, overall it is apparent the Double Seasonal Holt-Winters Exponential Smoothing with Error Correction method is the superior method.

4.1 Further Research

Although acceptable results were produced using the Double Seasonal Holt-Winters Exponential Smoothing method with Error Correction, further work could be undertaken. Additional research will be implemented to try and improve forecast accuracy in the context of the Irish market. 2.99% is at the upper limit of an acceptable MAPE for the accuracy of a next day forecast, as outlined in the Methodology, so adjustments will be made to see can this MAPE be improved upon. The dates for initialising the data and estimating the parameters should be considered in the same period as the forecast period, since data from summer months was used to forecast load for an autumn month. This resulted in a MAPE that is just below the acceptable threshold. Data for a similar season should be used to initialise and forecast and see whether this improves the MAPE.

5 Conclusions

In conclusion, the Double Seasonal Holt-Winters Exponential Smoothing method with Error Correction is demonstrated to be an acceptable method of short term load forecasting for a supply company operating in the Irish electricity market. The results for this method were positive in comparison to a simpler Double Seasonal Holt-Winters Exponential Smoothing method and to the Naïve Benchmark. While the MAPE (for the 27 days inclusive) of the DS HWES EC method is 2.99%, the corresponding MAPE for the DS HWES is 8.20% and 3.56% for the Naïve Benchmark.
6 References


7 Appendix
Method Results

Figure 7: Forecasted Load - Full Forecast Period Results by Method

Figure 8: Forecasted Load - Method Results Week 1