Bachelor thesis

Using Exponential Smoothing Methods for Modelling and Forecasting Short-Term Electricity Demand

By

Michelle de Ruiter, 414429

Erasmus School of Economics

Bachelor Econometrics and Operations Research

Bachelor thesis [programme: Business Analytics and Quantitative Marketing]

Project supervisor: Dr. M. Zhelonkin

Second assessor: Dr. M. van de Velden

Rotterdam, 2 July 2017

Abstract

In this paper different exponential smoothing methods are considered for modelling and forecasting short-term electricity demand in England and Wales. The time series contains half-hourly time periods and two seasonalities can be observed – one within each day and one within each week. Both seasonalities are modelled separately with the Holt-Winters methods. The double seasonal Holt-Winters method introduced by Taylor (2003) accommodates both seasonal patterns. An autoregressive model is fitted to the residuals in order to deal with first-order autocorrelation. In addition, a model for multiple seasonal (MS) processes introduced by Hyndman et al. (2008) is used, which divides the seasonal component into several sub-cycles and allows for the seasonal sub-cycles to be combined. This reduces the amount of seed values to be estimated. Moreover, the MS process allows for the seasonal terms to be updated more than once during a seasonal cycle. Although the MS process with three combined sub-cycles reduces the amount of seed values to be estimated and gives accurate forecasts, the double seasonal Holt-Winters method outperforms the other methods in forecasting the short-term electricity demand.

Contents

1 Introduction

Short-term electricity demand forecasting has become increasingly popular over the last decades. Accurate forecasts are very important for energy suppliers in order to operate and schedule such that the costs are minimized.

A data set provided by the National Grid is used that contains twelve weeks of half-hourly demand for electricity in England and Wales. The first eight weeks will be used to estimate the model parameters and the last four weeks are used for the evaluation of the forecasts. The series contains two seasonalities, one within each week and one within each day. These seasonal cycles have to be captured using accurate forecasting methods. The multiplicative seasonal ARIMA model is a widely used benchmark approach for this purpose. Taylor (2003) introduced the double seasonal ARIMA model which is able to capture both seasonalities in the short-term electricity demand data. An alternative and very competitive method for modelling short-term electricity demand are exponential smoothing methods, which have become increasingly popular because of their accuracy. Exponential smoothing methods assign exponentially decreasing weights to the observations, which means that older observations are given relatively less weight than more recent observations. Holt-Winters is a widely used exponential smoothing method for forecasting short-term electricity demand. However, the standard Holt-Winters method can only accommodate one seasonal pattern. In this paper, the double seasonal Holt-Winters method introduced by Taylor (2003) will be used to forecast the time series. Three approaches are given; the standard Holt-Winters for within-day seasonality, the standard Holt-Winters for within-week seasonality, and the double seasonal Holt-Winters method that captures both seasonalities. In addition, these models will be estimated using a simple autoregressive (AR) model that is fitted to the one-step ahead forecast errors. This is suggested by Chatfield (1987) in order to deal with first-order autocorrelated errors.

There are two variations of the Holt-Winters method, the multiplicative and the additive formulation. Taylor (2003) only models the electricity demand with the multiplicative formulation, which is used when the seasonal variation depends on the mean level of the time series. The second formulation is the additive formulation, where the seasonal variation is independent of the current mean level. Because the seasonal variations are roughly constant through the series, the additive formulation of the double seasonal Holt-Winters method will also be used to model and forecast the electricity demand.

The electricity demand data in England and Wales show clearly different patterns for the weekdays and the weekends. When modelling the electricity demand, Taylor (2003) assumes the same intraday cycle for all days of the week. However, the demand for electricity at, for example, Monday through Friday shows clearly different patterns than at Saturday and Sunday. Therefore, the model for multiple seasonal (MS) processes introduced by Hyndman et al. (2008) will be incorporated to model and forecasts the short-term electricity demand. This model can pick up the similarities for time periods between different days, it reduces the number of parameters to be estimated and it allows for more flexibility in updating the seasonal components. Moreover, Hyndman et al. (2008) found that his MS model produced more accurate forecasts than the double seasonal Holt-Winters model when forecasting hourly utility demand.

The main aim of this paper is to use exponential smoothing methods to model the short-term electricity demand in England and Wales and provide forecasts as accurate as possible. Taylor (2003) showed that his double seasonal Holt-Winters method is a very good candidate for forecasting this time series. However, Taylor (2003) uses the same intra-day cycle for all days of the week and therefore, a more efficient approach would be the multiple seasonal processes introduced by Hyndman et al. (2008), which relaxes some of the assumptions of the Holt-Winters methods. Moreover, Hyndman et al. (2008) found that his multiple seasonal process produced more accurate forecasts for his utility load and traffic flows data than the double seasonal Holt-Winters model.

The paper is structured as follows. Section 2 describes the main results of the researches of Taylor (2003) and Hyndman et al. (2008) and some examples of which methods other researches have been used to model complex seasonal patterns. In Section 3 the data of short-term electricity demand is described. The time series is decomposed into the seasonal, trend and remainder component in order to explain the most important features of the data. In Section 4 the different models and techniques are described starting with a basic explanation of exponential smoothing in Section 4.1. In Section 4.2 the Holt-Winters for within-day and within-week seasonality and the double seasonal Holt-Winters method are explained. The additive formulation is also addressed. Section 4.3 explains the multiple seasonal process introduced by Hyndman et al. (2008). The estimation and forecasting procedure of all methods is explained in Section 4.4 and Section 4.5 describes how an AR(1) model is fitted to the residuals. Empirical analysis is done in Section 5, where the models described in Section 4 are incorporated and used to model and forecast the short-term electricity demand. A discussion and conclusion of the research is given in Section 6.

2 Literature review

A lot of research about time series with multiple seasonal patterns has already been done in many published papers, such as Taylor (2003), Hyndman et al. (2008) and Taylor (2010). A widely used benchmark approach for modelling seasonal time series is the ARIMA model of Box, Jenkins, and Reinsel (1993). Taylor (2003) introduced the double seasonal ARIMA model which is able to capture two seasonalities. Furthermore, exponential smoothing methods, which are the central modelling techniques used in this paper, are very popular in order to model time series with multiple seasonalities.

Taylor (2003) modelled the half-hourly electricity demand in England and Wales using four methods, Holt-Winters for within-day seasonality, Holt-Winters for within-week seasonality, double seasonal Holt-Winters and the double seasonal ARIMA model. These first three methods will be explained and incorporated in this paper. Taylor (2003) fitted an autoregressive model to the residuals in order to deal with autocorrelated errors. He concludes that the double seasonal Holt-Winters methods with AR(1) adjustments outperforms the double seasonal ARIMA model and therefore, the double seasonal ARIMA model will not be considered in this paper as exactly the same data is used.

Taylor (2003) uses the multiplicative version of the double seasonal Holt-Winters method, where the seasonality is estimated by smoothing the ratio of the observed value to the product of the local level and the local seasonal index. Also, the level equation makes sure that the series is seasonally adjusted. This multiplicative method is much more common, and is used when the seasonal variations increase with the mean level of the series. However, another formulation involves additive seasonality as explained in Hyndman and Athanasopoulos (2012). In this method, the series is seasonally adjusted by subtracting the seasonality in the level equation and the seasonality is estimated by smoothing the difference of the observed value, the local level and the trend. The additive version is used when the seasonal variations are roughly constant through the series. Because this seems the case for the electricity demand, also the additive version of the double seasonal Holt-Winters method will be investigated in this paper.

Multivariate modelling could be considered as in Taylor and Buizza (2003) which investigate the use of weather ensemble predictions in electricity demand forecasting. However, this paper will only consider univariate modelling approaches. Although weather variables such as temperature, the amount of rainfall and wind speed play a very important role in the demand for electricity according to Taylor and Buizza (2003), Taylor (2003) states that univariate methods seem to be sufficient for modelling such short lead times. This is due to the fact that most of the time weather variables change smoothly over time rather than abrupt.

Other researchers that used exponential smoothing methods to model complex seasonalities are for example De Livera, Hyndman, and Snyder (2010), who introduced a model that is able to handle non-integer periods, high frequency multiple seasonal patterns and dual calender affects. Also Souza, Barros, and de Miranda (2007) use double seasonal exponential smoothing methods that account for holidays and temperature effects. However, for simplicity the data used in this paper contains no irregular days such as holidays.

Also Hyndman et al. (2008) make use of the Holt-Winters and the double seasonal Holt-Winters methods for modelling hourly data for both utility loads and traffic flows. They introduce a new multiple seasonal (MS) process, which allows for each day to have its own hourly pattern by making use of different sub-cycles. An example for the utility loads is that they can generate one common sub-cycle for Monday till Thursday and one common sub-cycle for Friday till Sunday. This will reduce the number of initial values and parameters to be estimated. In contrast to the double seasonal Holt-Winters method introduced by Taylor (2003), this methods allows for the seasonal terms to be updated more than once during a seasonal cycle. Also, there is more flexibility in updating the seasonal components because the model allows for different smoothing parameters for different sub-cycles. Hyndman et al. (2008) concludes that the forecasts with this MS process are more accurate than the double seasonal Holt-Winters method, when applied to the utility load and traffic flows data. Therefore, this paper will also incorporate this MS process to model and forecast the short-term electricity demand.

3 Data

A time series provided by the National Grid is used which contains the demand of short-term electricity in England and Wales from Monday 5 June 2000 to Sunday 27 August 2000, which are exactly 12 weeks. The data set is short-term in the sense that it contains 4032 observations, which means that there are 12 weeks of half-hourly demand for electricity. For simplification, this period is chosen as it contains no special days, such as holidays. The univariate methods discussed in this paper will not be able to give reasonable forecasts for such data, as demand at these special days are very different from "regular" days. The time series is plotted in the top panel of Figure 3.1.

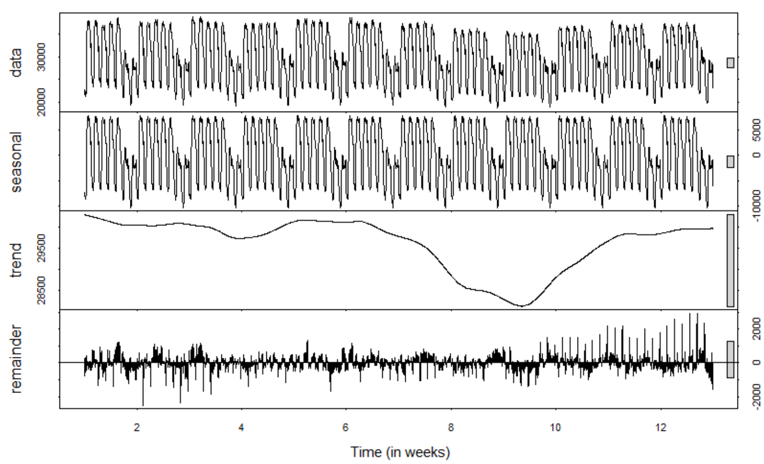

Figure 3.1: STL-decomposition of the time series of short-term electricity demand.

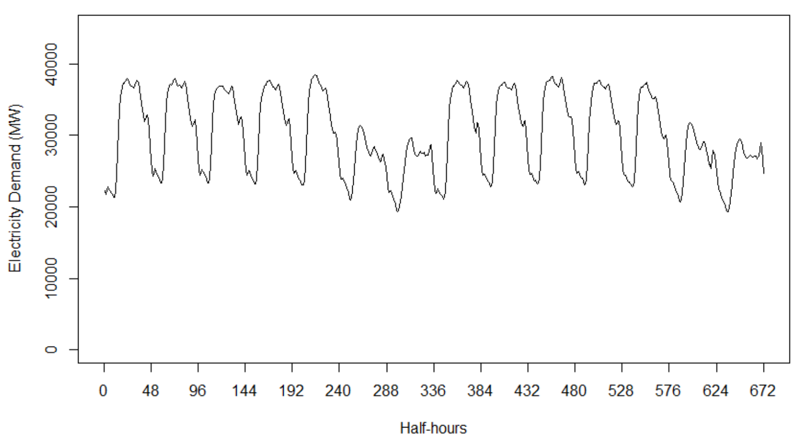

Figure 3.2: Two weeks of half-hourly electricity demand in England and Wales from Monday 5 June 2000 to Sunday 18 June 2000.

Figure 3.2 shows the short-term electricity demand for only a fortnight in June 2000. This figures shows clearly two seasonalities, a within-day seasonality and a within-week seasonality. The within-day seasonal cycle, which has a duration of 48 half-hour periods, can be observed because the demand at a particular time period in one day is similar to the demand at the same time period on another day. Particularly the weekdays show similar daily seasonal patterns. The weekends however show a different pattern than the weekdays. Moreover, every week the same seasonal pattern occurs. This longer weekly seasonal cycle has a duration of 336 half-hour periods and seems to dominate the seasonal variation in the data.

To further explore the time series, a STL-decomposition is performed, which is a method to decompose a time series. The abbreviation STL means "Seasonal and Trend decomposition using Loess" and is developed by Cleveland, Cleveland, McRae, and Terpenning (1990). The decomposition is shown in the last three panels of Figure 3.1. The bars at the right hand side of each graph allow to compare the magnitudes of each component. The bar on the seasonal panel is almost as large as the bar on the data panel, which indicates that the seasonal component dominates the variation in the data. The bar on the trend panel shows that the trend has not much influence on the variation in the data. The trend panel shows a clear downward pattern around weeks eight to ten, which is slightly noticeable in the data panel.

The first eight weeks of the time series will be used for estimation of the model parameters. This amounts to the first 2688 half-hourly time periods as estimation sample. The remaining four weeks will be used for estimation, which amounts to the last 1344 observations of the data set.

4 Methodology

4.1 Exponential smoothing

One can use simple methods to generate forecasts for time series, such as the naïve method, which just takes the last observation of the series as the forecasts, or the average method, which uses the average of all observations as the forecasts. These simple methods thus give equal weights to all of the forecasts. Most of the time, this does not give very reliable predictions. A better approach is usually to assign more weight to more recent observations and less weight to observations further in the past. An example of this kind of forecast is

$$\hat{X}_{T+1|T} = \alpha X_T + \alpha(1-\alpha)X_{T-1} + \alpha(1-\alpha)^2 X_{T-2} + ..., \tag{1}$$

where $0 \leq \alpha \leq 1$ is the smoothing parameter. As can be seen from Equation 1 the weights decrease exponentially and α determines at which rate they decrease. When α is close to 1, relatively more weight is assigned to more recent observations and when α is close to 0, relatively more weight is given to observations further in the past. This is the general idea behind exponential smoothing.

Simple exponential smoothing (SES) can be used for time series that contain no trend nor seasonality. In this paper, the exponential smoothing methods will be formulated in the so-called component form. The component form for SES is

$$\begin{aligned} \text{Forecast equation:} \quad & \hat{X}_{t+1|t} = S_t, \\ \text{Level equation:} \quad & S_t = \alpha X_t + (1-\alpha)S_{t-1}. \end{aligned}$$

This form contains only one component, which is the smoothing equation for the level S_t. This equation can be rewritten in the error correction form, which will make clear what the role of α is. We can rewrite S_t as

$$\begin{aligned} S_t &= \alpha X_t + (1-\alpha)S_{t-1} \\ &= \alpha X_t + S_{t-1} - \alpha S_{t-1} \\ &= S_{t-1} + \alpha(X_t - S_{t-1}) \\ &= S_{t-1} + \alpha e_t, \end{aligned}$$

where $e_t = X_t - S_{t-1}$ is the one-step within-sample forecast error at time t, $t = 1, ..., T$. Now we can see that the level is "smoother" when α comes close to zero, because small weight is given to the error e_t.

This was a short explanation of how the simplest form of exponential smoothing methods works. However, the short-term electricity demand data does contain a trend and seasonality. Holt (1957) and Winters (1960) introduced an exponential smoothing method that accounts for both trend and seasonality, which will be referred to as the standard Holt-Winters method. Taylor (2003) extended this method to the double seasonal Holt-Winters method, which is able to capture both seasonalities. These methods will be explained in Section 4.2.

4.2 Holt-Winters method

4.2.1 Standard Holt-Winters method

In order to deal with the trend and the seasonal components in the time series of short-term electricity demand, the standard Holt-Winters method will be implemented. This method can only accommodate one seasonality. Therefore, the within-day (s=48) and the within-week (s=336) seasonalities will be modelled separately.

The standard Holt-Winters method contains three components, one smoothing equation for the level, one for the trend and one for the seasonality. The component formulation of the multiplicative Holt-Winters method is represented below. The same formulation as in Taylor (2003) is used.

$$\hat{X}_t(h) = (S_t + hT_t)I_{t-s+h}, \tag{2a}$$

$$\text{Level:} \qquad S_t = \alpha\left(\frac{X_t}{I_{t-s}}\right) + (1-\alpha)(S_{t-1} + T_{t-1}), \tag{2b}$$

$$\text{Trend:} \qquad T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1}, \tag{2c}$$

$$\text{Seasonality:} \qquad I_t = \delta\left(\frac{X_t}{S_t}\right) + (1-\delta)I_{t-s}. \tag{2d}$$

The h-step ahead forecasts $X_t(h)$ are calculated using the smoothing equations for the level, the trend and the seasonality. The level in Equation 2b is estimated by smoothing (1) the ratio of the observed value and the seasonal index at exactly one seasonal cycle ago and (2) the sum of the previous level and the previous trend $S_{t-1} + T_{t-1}$, where $0 \leq \alpha \leq 1$ is the smoothing parameter. The trend component in Equation 2c is additive and is estimated by the weighted average of the first differences of the level $S_t - S_{t-1}$ and the previous trend T_{t-1}. The smoothing parameter for the trend is $0 \leq \beta \leq 1$. The seasonal index at time t in Equation 2d is estimated by the weighted average of (1) the ratio of the observed value at time t and the level at time t and (2) the seasonal index exactly one seasonal cycle ago. For example, when the within-day seasonality (s=48) for the electricity demand is modelled, the seasonal index at time t=50 will be estimated using the seasonal index exactly one day ago, at t=2. The smoothing parameter for the seasonality is $0 \leq \delta \leq 1$. The initial values for the level S_0, the trend T_0 and the seasonal index $\mathbf{I}_0 = \begin{bmatrix} I_{-s}, & I_{-s+1}, & \dots, & I_0 \end{bmatrix}^\top$ are chosen the same as in the standard implemented R function HoltWinters.R.

4.2.2 Double seasonal Holt-Winters method

Taylor (2003) proposed an extension of the standard Holt-Winters method in order to deal with two seasonalities. The double seasonal Holt-Winters method contains four components – one for the level, one for the trend and one for each seasonality. The method can be extended to three or more seasonalities. The formulation of the multiplicative model in this paper is same as used by Taylor (2003) and is represented in the following equations:

$$\hat{X}_t(h) = (S_t + hT_t)D_{t-s_1+h}W_{t-s_2+h}, \tag{3a}$$

$$\text{Level:} \qquad S_t = \alpha\left(\frac{X_t}{D_{t-s_1}W_{t-s_2}}\right) + (1-\alpha)(S_{t-1} + T_{t-1}), \tag{3b}$$

$$\text{Trend:} \qquad T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1}, \tag{3c}$$

$$\text{Seasonality 1:} \qquad D_t = \delta\left(\frac{X_t}{S_tW_{t-s_2}}\right) + (1-\delta)D_{t-s_1}, \tag{3d}$$

$$\text{Seasonality 2:} \qquad W_t = \omega\left(\frac{X_t}{S_tD_{t-s_1}}\right) + (1-\omega)W_{t-s_2}. \tag{3e}$$

The h-step-ahead forecasts $\hat{X}_t(h)$ are now calculated using the level component, the trend component and the two seasonal indexes D_t and W_t, where D_t is the smoothing equation for the s_1-period seasonal cycle and W_t is the smoothing equation for the s_2-period seasonal cycle. In this paper, $s_1 = 48$ for the within-day seasonality and $s_2 = 336$ for the within-week seasonality. The level in Equation

3b is estimated using the weighted average of (1) the ratio of the observed value at time t and the multiplication of the two seasonal indices one seasonal cycle ago and (2) the sum of the previous level and the previous trend. The trend component in Equation 3c is estimated in the same way as in the standard Holt-Winters method. The smoothing parameters for the level and the trend are $0 \leq \alpha, \beta \leq 1$ respectively. The within-day seasonal component D_t is estimated using the weighted average of (1) the ratio of the observed value at time t and the level at time t multiplied by the within-week seasonal index exactly one week ago and (2) the within-day seasonal index exactly one day ago. The within-week seasonal component W_t is estimated using the weighted average of (1) the ratio of the observed value at time t and the level at time t multiplied by the within-day seasonal index exactly one day ago and (2) the within-week seasonal index exactly one week ago. The smoothing parameters for the within-day and within-week seasonal components are $0 \leq \delta, \omega \leq 1$ respectively.

The initial values for the trend T_0, the level S_0, the within-day seasonal index

$$\mathbf{D}_0 = \begin{bmatrix} D_{-s_1}, & D_{-s_1+1}, & \ldots, & D_0 \end{bmatrix}^\top$$

and the within-week seasonal index

$$\mathbf{W}_0 = \begin{bmatrix} W_{-s_2}, & W_{-s_2+1}, & \ldots, & W_0 \end{bmatrix}^\top$$

are chosen the same as in Taylor (2003).

4.2.3 The additive formulation

Two variations of this method exist, the additive method and the multiplicative method.

The Holt-Winters methods described above all involved multiplicative seasonality. However, another formulation involves additive seasonal patterns. According to Hyndman and Athanasopoulos (2012) the choice of which method to use depends mostly on the seasonal variations. When these are constant through the series, the additive method should be the best choice. When the seasonal variations change proportional to the level of the time series, the multiplicative method is preferred. The latter is much more used for forecasting seasonal time series, but the series of electricity demand shows roughly constant seasonal variations, also the additive formulation will be investigated for the double seasonal Holt-Winters method. This formulation is represented below.

$$\begin{aligned}
\hat{X}_t(h) &= S_t + hT_t + D_{t-s_1+h} + W_{t-s_2+h}, && \text{(4a)}\\
\text{Level:} \quad S_t &= \alpha(X_t - D_{t-s_1} - W_{t-s_2}) + (1-\alpha)(S_{t-1} + T_{t-1}), && \text{(4b)}\\
\text{Trend:} \quad T_t &= \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1}, && \text{(4c)}\\
\text{Seasonality 1:} \quad D_t &= \delta(X_t - S_t - W_{t-s_2}) + (1-\delta)D_{t-s_1}, && \text{(4d)}\\
\text{Seasonality 2:} \quad W_t &= \omega(X_t - S_t - D_{t-s_1}) + (1-\omega)W_{t-s_2}. && \text{(4e)}
\end{aligned}$$

The level component in Equation 4b is estimated by smoothing the difference of the observed value and the within-day and within-week seasonality one seasonal cycle ago instead of smoothing the ratio of these components. The trend component in Equation 4c is the same as in Equation 3c for the multiplicative formulation. The seasonal indices undergo the same change - instead of the ratio, the difference of the observed value, the level and the other seasonal index is used to estimate the component. Also, the h-step ahead forecasts are calculated using addition instead of multiplication. Again, α, β, δ and ω are the smoothing parameters to be estimated.

4.3 Multiple seasonal (MS) processes

4.3.1 The seasonal components of the Holt-Winters methods

The Holt-Winters methods explained in Section 4.2 assume the same seasonal intra-day cycle for all days of the week. The standard Holt-Winters method uses the seasonal cycle

$$\mathbf{I}_t = \begin{bmatrix} I_t, & I_{t-1}, & \ldots, & I_{t-s+1} \end{bmatrix}^\top,$$

where s=48 (i.e. within-day seasonality) or s=336 (i.e. within-week seasonality). This model thus needs to estimate s+2 different seed values – one for the level, one for the trend and s for the seasonal component. Every seasonal term is updated once during the seasonal cycle of s time periods. This means that for a daily cycle of length s=48 the seasonal terms are updated once every day and that for a weekly cycle of length s=336 the seasonal terms are updated once every week. Also, the model uses the same smoothing parameter δ for each of the s seasonal terms during a particular cycle.

The double seasonal Holt-Winters method uses two seasonal cycles denoted by

$$\mathbf{D}_t = \begin{bmatrix} D_t, & D_{t-1}, & \ldots, & D_{t-s_1+1} \end{bmatrix}^\top$$

for the within-day seasonality and

$$\mathbf{W}_t = \begin{bmatrix} W_t, & W_{t-1}, & \ldots, & W_{t-s_2+1} \end{bmatrix}^\top$$

for the within-week seasonality. This model thus requires $s_1 + s_2 + 2$ different seed values, where s_1=48 and s_2=336. This means that 336 seasonal terms are updated once every week and an additional 48 seasonal terms are updated once every day. Also in this model, the same smoothing parameter δ for the within-day seasonal terms and the same smoothing parameter ω for the within-week seasonal terms are used.

4.3.2 The seasonal component of the MS process

Hyndman et al. (2008) developed a model for a multiple seasonal process with the aim of updating the seasonal terms more than once during a seasonal cycle. This is done by relaxing some of the assumptions for the seasonal components of the standard and double seasonal Holt-Winters methods described above. The model allows to combine certain daily sub-cycles into one sub-cycle. For example, when we look at the data in the top panel of Figure 3.1, we can distinguish 7 different daily sub-cycles. However, the cycles of the weekdays seem very similar and we could combine these cycles into one sub-cycle. If the two sub-cycles of Saturday and Sunday are also combined, we end up with a seasonal component that is divided into $r = 2$ sub-cycles. This is a large reduction in the amount of seasonal terms to be estimated. How this process exactly works is explained below.

First, the seasonal component with the initial seasonal terms is divided into r sub-cycles as follows

$$\begin{aligned} \mathbf{c}_{i,0} &= (I_{i,-s_1+1}, I_{i,-s_1+2}, \ldots, I_{i,-1}, I_{i,0})^\top \\ &= (I_{-s_1(r-i)-s_1+1}, I_{-s_1(r-i)-s_1+2}, \ldots, I_{-s_1(r-i)-1}, I_{-s_1(r-i)})^\top, \end{aligned} \tag{5}$$

where $i = 1, \ldots, r$ and $r \leq k = \frac{s_2}{s_1}$. If we would choose r equal to k, the seasonal component would be divided into $k = \frac{s_2}{s_1} = \frac{336}{47} = 7$ daily sub-cycles. As we will see later in this paper, under certain restrictions this amounts to the standard Holt-Winters method for within-week seasonality.

The division of the initial seasonal component in Equation 5 can be extended for each time period t as follows:

$$\mathbf{c}_{it} = (I_{i,t}, I_{i,t-1}, \ldots, I_{i,t-s_1+1})^\top, \tag{6}$$

where $i = 1, \ldots, r$. Now, $\mathbf{c}_{it}$ in Equation 6 contains at each time period t the current s_1 seasonal terms for cycle i.

4.3.3 The model for the MS process

The model for multiplicative seasonality

Hyndman et al. (2008) uses the error correction form to represent the model for the multiple seasonal process. This form is explained in Section 4.1. In this paper, all models are represented in the component form. A very similar model with additive trend and multiplicative seasonality that is represented in the component and the error correction form is stated in Hyndman and Athanasopoulos (2012), who provide various tables containing both formulations of various different models. These tables are used to write the error correction formulation of the model of Hyndman et al. (2008) into the component form. The component formulation of the model for the multiple seasonal process with multiplicative seasonality is

$$\hat{X}_t(h) = (S_{t-1} + hT_{t-1})\mathbf{d}_t^\top \mathbf{I}_{t-s_1+h}, \tag{7a}$$

$$\text{Level:} \qquad S_t = \alpha\left(\frac{X_t}{\mathbf{d}_t^\top \mathbf{I}_{t-s_1}}\right) + (1-\alpha)(S_{t-1} - T_{t-1}), \tag{7b}$$

$$\text{Trend:} \qquad T_t = \beta^*(S_t - S_{t-1}) + (1-\beta^*)T_{t-1}, \tag{7c}$$

$$\text{Seasonality:} \qquad \mathbf{I}_t = \mathbf{\Gamma}\mathbf{d}_t\left(\frac{X_t}{S_{t-1} + T_{t-1}}\right) + \mathbf{I}_{t-s_1} - \mathbf{\Gamma}\mathbf{d}_t\mathbf{d}_t^\top \mathbf{I}_{t-s_1}, \tag{7d}$$

where $\mathbf{I}_t = \begin{bmatrix} I_{1,t}, & I_{2,t}, & \ldots, & I_{r,t} \end{bmatrix}^\top$ and $\mathbf{d}_t = \begin{bmatrix} d_{1,t}, & d_{2,t}, & \ldots, & d_{r,t} \end{bmatrix}^\top$, where

$$d_{i,t} = \begin{cases} 1 & \text{if time period } t \text{ occurs within sub-cycle } i; \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

$\mathbf{\Gamma}$ is a matrix that contains for each sub-cycle the smoothing parameters γ_{ij}, where $i, j = 1, \ldots, r$. In this way, the model allows for different smoothing parameters for different sub-cycles. The diagonal elements γ_{ii} are used to update seasonal terms during time periods of the same sub-cycle, whereas the off-diagonal elements γ_{ij} are used to update seasonal terms during time periods in another sub-cycle. We will denote this model by MS(r; s_1, s_2). The h-step ahead forecasts $\hat{X}_t(h)$ are calculated similarly as in Taylor (2003), because this allows for easy comparison of the Holt-Winters methods from Section 4.2 and the MS processes.

The model for additive seasonality

Besides the MS process for multiplicative seasonality, Hyndman et al. (2008) introduced the MS process for additive seasonality. The Holt-Winters methods described above are implemented for both the multiplicative and the additive seasonality, so the MS process will also be adapted for additive seasonal patterns. Instead of the error correction formulation according to Hyndman et al. (2008), the component formulation will be used to represent the model. The derivation from the error correction form used by Hyndman et al. (2008) to the component form can be found in Appendix A. The MS process for additive seasonality is very similar as the model in Equations 7 and can be represented as follows:

$$\hat{X}_t(h) = S_{t-1} + hT_{t-1} + \mathbf{d}_t^\top \mathbf{I}_{t-s_1+h}, \tag{9a}$$

$$\text{Level:} \qquad S_t = \alpha(X_t - \mathbf{d}_t^\top \mathbf{I}_{t-s_1}) + (1-\alpha)(S_{t-1} - T_{t-1}), \tag{9b}$$

$$\text{Trend:} \qquad T_t = \beta^*(S_t - S_{t-1}) + (1-\beta^*)T_{t-1}, \tag{9c}$$

$$\text{Seasonality:} \qquad \mathbf{I}_t = \mathbf{\Gamma}\mathbf{d}_t(X_t - S_{t-1} - T_{t-1}) + \mathbf{I}_{t-s_1} - \mathbf{\Gamma}\mathbf{d}_t\mathbf{d}_t^\top \mathbf{I}_{t-s_1}. \tag{9d}$$

4.3.4 Model restrictions

The matrix of smoothing parameters for the different sub-cycles can become quite large when the amount of different sub-cycles r increases, because $\mathbf{\Gamma}$ is an $r \times r$ matrix and thus contains r^2 elements. Hyndman et al. (2008) imposes three restrictions on $\mathbf{\Gamma}$ to reduce the number of parameters. Some of these restrictions will lead to the standard or double seasonal Holt-Winters method described in Section 4.2.

First of all, Hyndman et al. (2008) divide the smoothing parameters in $\mathbf{\Gamma}$ into diagonal elements and off-diagonal elements as follows:

$$\gamma_{i,j} = \begin{cases} \gamma_1^* & \text{if } i = j; \\ \gamma_2^* & \text{if } i \neq j. \end{cases} \tag{10}$$

The three restrictions of interest according to Hyndman et al. (2008) are

- **Restriction 1:** $\gamma_1^* \neq 0$ and $\gamma_2^* = 0$
- **Restriction 2:** $\gamma_1^* = \gamma_2^*$
- **Restriction 3:** Equivalent to Equation 10

Restriction 1 does not allow the seasonal terms in a particular sub-cycle to be updated during another sub-cycle. Moreover, when $r = k$ this restricted model is the same as the standard Holt-Winters model for within-week seasonality, where $\gamma_1^* = \delta$ in Equation 2d. If the seed values for the different sub-cycles would be equal to each other, the model with restriction 2 would be identical to the standard Holt-Winters for within-day seasonality. However, the seed values of the sub-cycles are rarely identical in modelling this multiple seasonal process. This restricted model will allow the seasonal terms in a sub-cycle to be updated during another sub-cycle, although using the same smoothing parameter. Restriction 3 allows for different smoothing parameters when updating seasonal sub-cycles during the same sub-cycle and other sub-cycles. Moreover, when $r = k$ this restricted model is equivalent to the double seasonal Holt-Winters method, where $\gamma_1^* = \delta + \omega$ and $\gamma_2^* = \delta$ in Equations 3d and 3e. A brief overview of the amount of smoothing parameters and seed values to be estimated for all the models is given in Table 1.

Table 1: Smoothing parameters and seed values required for each model

Model	Smoothing parameters	Seed values
Holt-Winters for within-day seasonality	3	$s_1 + 2$
Holt-Winters for within-week seasonality	3	$s_2 + 2$
Double seasonal Holt-Winters	4	$s_1 + s_2 + 2$
MS(r; s_1, s_2)	$r^2 + 2$	$rs_1 + 2$
MS(r; s_1, s_2) with Restriction 1	3	$rs_1 + 2$
MS(r; s_1, s_2) with Restriction 2	3	$rs_1 + 2$
MS(r; s_1, s_2) with Restriction 3	4	$rs_1 + 2$

4.3.5 Model selection

In order to select the most appropriate model we have to choose the number of different sub-cycles r and the restriction for the smoothing parameters belonging to the seasonal component. The model selection is performed based on the two step selection procedure by Hyndman et al. (2008).

The first step involves the decision of which value of r to use. To select possible values of r, the seasonal terms of the last week of the estimation sample, estimated with the double seasonal Holt-Winters model, are plotted. From this plot it will be clear which possible daily seasonal cycles can be combined into one sub-cycle.

From the estimation sample consisting of n=2688 time-periods the last 20% of the time-periods rounded to the nearest multiple of whole weeks are withheld. This amounts to $q = 672$ time-periods or 2 weeks of data for the short-term electricity demand. Next, the parameters of the different MS(r; 48, 336) will be estimated using observations 1 to $n - q$ by minimizing the one-step ahead sum of squared errors. For each model, the MAPE(1) for the one-step ahead forecasts for observations $n - q + 1$ to n is calculated as follows:

$$\text{MAPE}(1) = \frac{100}{q} \sum_{t=n-q+1}^{n} \Big| \frac{X_t - \hat{X}_t}{X_t} \Big|. \tag{11}$$

.

The value of r is picked from the model with the lowest MAPE(1) value.

The second step involves choosing the restrictions on $\mathbf{\Gamma}$. The value of r from the first step is selected and the one-step ahead forecasts for periods $n - q + 1$ till n are made for Restriction 1, 2, 3 and no restriction. The restriction that produces the lowest MAPE(1) is imposed in the MS process.

4.4 Estimation and prediction

The estimation and forecasting procedure is the same for all the methods discussed in Sections 4.2 and 4.3. The first eight weeks of the data are used for the estimation of the model parameters. For the series of half-hourly demand this amounts to the first $n = 2688$ observations. The one-step ahead forecasts $\hat{X_{t+1}}$ are calculated for the estimation sample. These forecasts are used to compute the one-step ahead forecast errors $\hat{e}_{t+1}$ as follows:

$$\hat{e}_{t+1} = X_{t+1} - \hat{X_{t+1}}, \tag{12}$$

where t=0,...,$n - 1$. The smoothing parameters are optimized by minimizing the sum of squared errors (SSE), as in the following equation:

$$\text{SSE} = \sum_{t=1}^{n} \hat{e}_t^2. \tag{13}$$

When the values for the smoothing parameters are optimized, these values can be used to forecast the electricity demand for the remaining four weeks, which is also called the post-sample period. This period consists of observations at $t = n + 1, ..., T$, which makes a total of p=1344 observations. The h-step ahead forecasts are calculated for $h = 1, ..., s_1$, where s_1=48 is the within-day seasonality. These forecasts $\hat{X}_{t+h}$ can be saved in a matrix, where the rows represent h=1,...,48 and the columns represent t=2688,...,4031. This matrix looks as follows:

$$\begin{bmatrix} \hat{X}_{2689} & \hat{X}_{2690} & \hat{X}_{2691} & \dots & \hat{X}_{4031} & \hat{X}_{4032} \\ \hat{X}_{2690} & \hat{X}_{2691} & \hat{X}_{2692} & \dots & \hat{X}_{4032} & - \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \hat{X}_{2736} & \hat{X}_{2737} & \hat{X}_{2738} & \dots & - & - \end{bmatrix}$$

A widely used measure for comparing the post-sample forecast accuracy is the Mean Absolute Percentage Error (MAPE). The matrix of h-step ahead forecasts allows us to calculate the MAPE(h) values for $h = 1, \dots, s_1$ as follows:

$$\text{MAPE(h)} = \frac{100}{p - (h - 1)} \sum_{t=n}^{T-h} \Big| \frac{X_{t+h} - \hat{X}_{t+h}}{X_{t+h}} \Big|. \tag{14}$$

4.5 Residual autocorrelation

According to Gardner (1985) the forecasts from exponential smoothing methods can be improved by introducing an AR(1) model to the one-step ahead forecast errors. Taylor (2003) uses this adjustment in the Holt-Winters methods because he found sizeable first-order autocorrelated errors. The AR(1) model is as follows:

$$e_t = \phi e_{t-1} + \zeta_t, \tag{15}$$

where ϕ is the model parameter to be estimated and ζ_t is white noise. The h-step-ahead forecasts are modified by adding the term $\phi^k e_t$, where ϕ is estimated with the estimation sample. This means that the formulas of the h-step-ahead forecasts for all models change. For example, the h-step ahead forecast for the standard Holt-Winters model in Equation 2a becomes

$$\begin{aligned} \hat{X}_t(k) &= (S_t + kT_t)I_{t-s+k} + \phi^k e_t \\ &= (S_t + kT_t)I_{t-s+k} + \phi^k \{X_t - (S_{t-1} + T_{t-1})I_{t-s}\}. \end{aligned} \tag{16}$$

The modified h-step ahead forecasts for the other models can be found in Appendix B.

5 Results

5.1 Holt-Winters without AR(1) adjustment

The optimal smoothing parameter values are calculated by minimizing the sum of squared one-step ahead forecast errors as explained in Section 4.4. The optimized values for the Holt-Winters method without the AR(1) adjustment are shown in Table 2. The estimated parameter $\hat{\alpha}$ takes values close to 1, which means that all methods give relatively more weight to more recent observations. As explained in Section 3 the variability in the estimation sample is dominated by the seasonality, rather than the trend. This trend constantly takes small values when modelling Holt-Winters for within-week seasonality and double seasonal Holt-Winters. This explains the zero values for $\hat{\beta}$ for these two methods in Table 2. However, in the Holt-Winters for within-day seasonality the optimized parameter $\hat{\beta}$ for the trend takes the larger value of 0.853, which is accompanied by the smoothed trend taking highly varying values. According to Taylor (2003), Holt-Winters for within-day seasonality incorporated this highly varying trend because it is unable to model the within-week seasonality.

Table 2: Holt-Winters smoothing parameters optimized from estimation sample. There is no adjustment for autocorrelation in the residuals of the methods.

	Level α	Trend β	Within-day seasonality δ	Within-week seasonality ω
Holt-Winters for within-day seasonality	0.986	0.853	1.000	-
Holt-Winters for within-week seasonality	0.831	0.000	-	1.000
Double seasonal Holt-Winters	0.881	0.000	0.831	1.000

For each of the three methods the MAPE values for 48 lead times are calculated as explained in Section 4.4. In Figure 5.1 the MAPE results for the different Holt-Winters methods are compared. The MAPE values for Holt-Winters for within-day seasonality increase at such a large rate that only the MAPE values until 2-step ahead forecasts are plotted. This large increase in MAPE values is due to the fact that this method only accommodates the within-day seasonality, while the data in the top panel of Figure 3.1 also shows a within-week seasonal pattern, which dominates the variation in the data. Therefore, Holt-Winters for within-week seasonality shows much better results. The double seasonal Holt-Winters method shows better forecast accuracy for lead times up to 9 and beyond 39, which indicates that taking both seasonalities in account will probably improve the forecast accuracy.

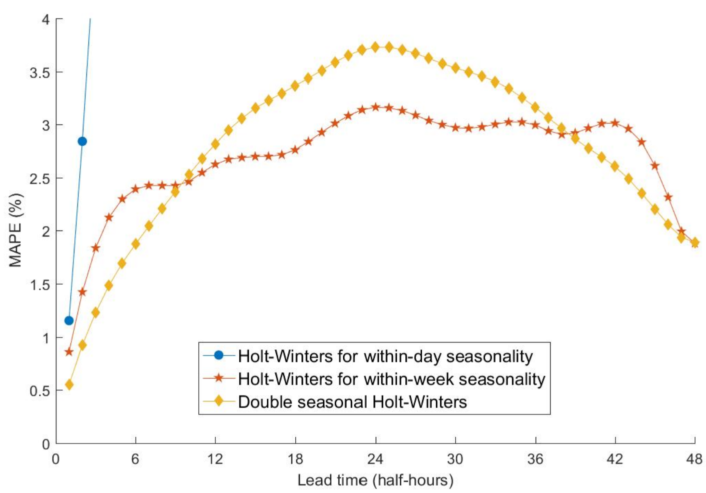

Figure 5.1: Comparison of the out of sample MAPE(h) results for the Holt-Winters methods.

The forecasts can be improved by fitting an AR(1) model to the residuals, which is explained in Section 4.5. The one-step ahead forecast errors for all three Holt-Winters methods show significant autocorrelation, which means that the residuals contain some information that can be exploited to obtain more accurate forecasts. Figure 5.2 shows the sample autocorrelation function (ACF) for the one-step ahead forecast errors in the estimation sample for double seasonal Holt-Winters. Only the double seasonal Holt-Winters method will be considered here because this method is likely to produce the most accurate forecasts of the three Holt-Winters methods. The ACF plots of the Holt-Winters for within-day and within-week seasonality can be found in Appendix C. The ACF plot in Figure 5.2 shows spikes at various lags, which means that the forecasts are not optimal and can be improved.

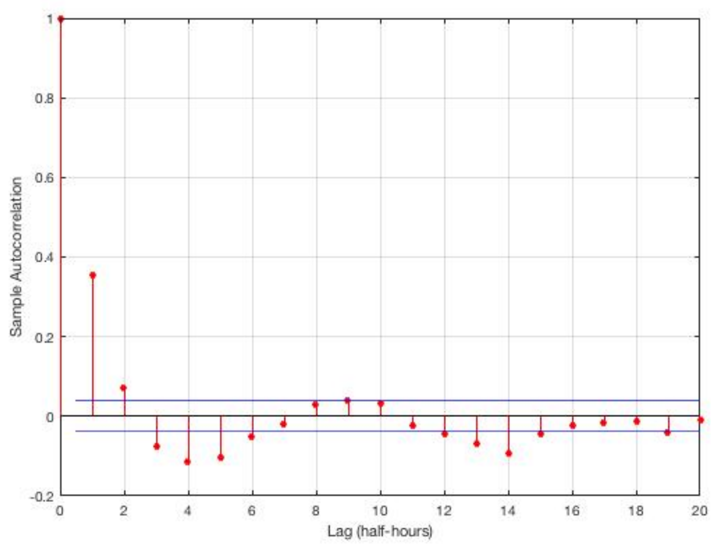

Figure 5.2: The sample autocorrelation plot of the one-step ahead forecast errors obtained from estimating the estimation sample with the double seasonal Holt-Winters method.

5.2 Holt-Winters with AR(1) adjustment

Because of the first-order autocorrelation of the one-step ahead forecast errors, an AR(1) model is fitted to these errors. The estimation of the parameters is done in a single stage, which is more efficient than a two-stage estimation procedure according to Chatfield (1987). The parameters are optimized by minimizing the sum of squared one-step ahead forecast errors from the estimation sample, where an AR(1) model is fitted to the errors. The optimized parameters are shown in Table 3. The values of the estimated level parameter $\hat{\alpha}$ are much smaller for Holt-Winters for within-week seasonality and double seasonal Holt-Winters than the value for $\hat{\alpha}$ in the models without AR(1) adjustment. It seems that the smoothing equation for the level has almost been replaced by the fitted AR(1) model for the residuals. The estimated trend parameter $\hat{\beta}$ now also takes the value zero for Holt-Winters for within-day seasonality. The parameters $\hat{\delta}$ and $\hat{\omega}$ have decreased substantially for all of the three multiplicative seasonal Holt-Winters methods compared to those in Table 2.

In Figure 5.3 the MAPE results are shown for the three multiplicative Holt-Winters methods for lead times up to a day-ahead. Again, the MAPE results for the Holt-Winters for within-day seasonality are very poor, which can be explained by the dominance of the within-week seasonality in the variation of the data. However, the results for Holt-Winters for within-week seasonality and the double seasonal Holt-Winters have improved compared to the models without the AR(1) adjustment for the forecast errors. The double seasonal Holt-Winters method outperforms the Holt-Winters for within-day and

Table 3: Holt-Winters smoothing parameters optimized from estimation sample. The methods include an AR(1) model for the residuals.

	Level α	Trend β	Within-day seasonality δ	Within-week seasonality ω	AR ϕ
Holt-Winters for within-day seasonality	0.803	0.000	0.689	-	0.736
Holt-Winters for within-week seasonality	0.013	0.000	-	0.416	0.924
Double seasonal Holt-Winters	0.012	0.004	0.179	0.325	0.935
Additive double seasonal Holt-Winters	0.000	0.000	0.362	0.344	0.986

within-week seasonality for all h-step ahead forecast errors.

The additive formulation of the double seasonal Holt-Winters model gives almost identical performances compared to the multiplicative version of the model. The estimates for the smoothing parameters $\hat{\beta}$ and $\hat{\delta}$ for the additive formulation can be found in Table 3 and are slightly larger than those of the multiplicative version. Also the estimated parameter $\hat{\phi}$ of the AR(1) model is slightly larger. Figure 5.4 shows the MAPE(h) values for both the additive and the multiplicative double seasonal Holt-Winters methods. Most of the time the additive formulation gives more accurate forecasts than the multiplicative formulation, although the improvement is almost nil.

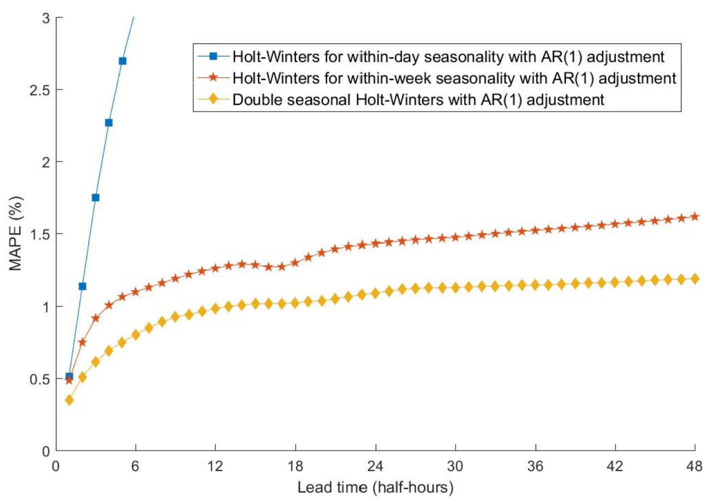

Figure 5.3: Comparison of the out of sample MAPE(h) results for the three Holt-Winters methods including AR(1) model for the residuals.

Figure 5.5 shows the ACF plot of residuals from the multiplicative double seasonal Holt-Winters method with AR(1) adjustment. The plot shows much less spikes at different lags than the ACF plot in Figure 5.2, which indicates that the AR(1) model for the residuals reduced the remaining autocorrelation. However, there still is some autocorrelation in the one-step ahead errors. The ACF plots of the residuals from the Holt-Winters method for within-day and within-week seasonality with AR(1) adjustment can be found in Appendix C.

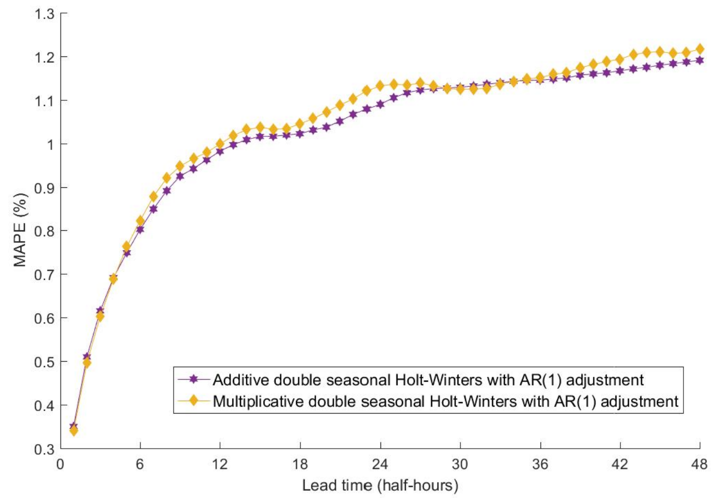

Figure 5.4: Comparison of the MAPE results for the additive and multiplicative double seasonal Holt-Winters methods including AR(1) model for the residuals.

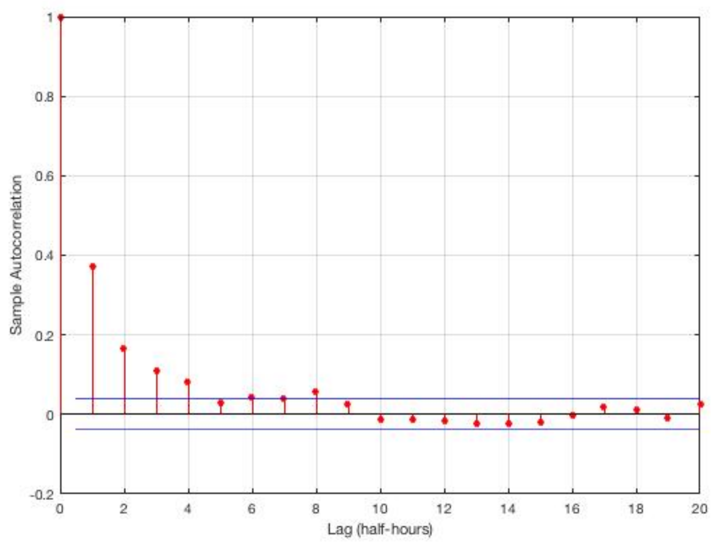

Figure 5.5: The sample autocorrelation plot of the one-step ahead forecast errors obtained from estimating the estimation sample with the double seasonal Holt-Winters method including AR(1) adjustment.

5.3 The MS processes

5.3.1 Model selection

The first step of selecting an appropriate model is to choose the value of r. In order to choose this value, it has to be investigated which seasonal sub-cycles can be combined. To see which seasonal cycles could be combined, the seasonal terms for each day of the last week of the estimation sample are plotted in Figure 5.6. The seasonal terms belong to the double seasonal Holt-Winters method, which is equivalent to the MS(7; 48, 336) model as explained in Section 4.3.4.

From Figure 5.6 it can be seen that the weekdays and the weekends show clearly different seasonal

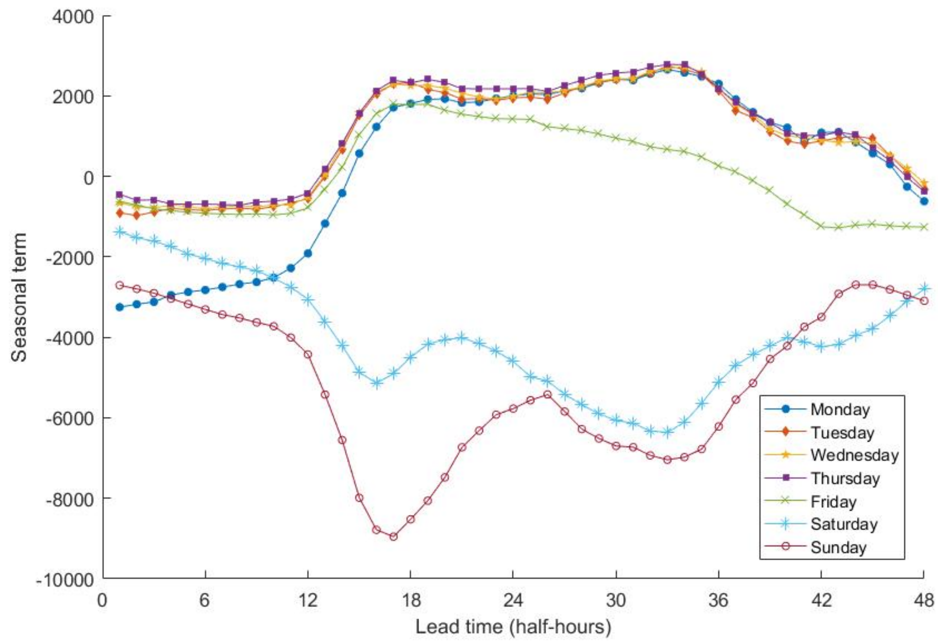

Figure 5.6: The estimated half-hourly seasonal sub-cycles for each day of the last week of the estimation sample. The seasonal term on the y-axis represents the seasonal terms in Equation 6, where $r = 7$.

patterns. Therefore, one possibility for combining the seasonal cycles is one common sub-cycle for Monday through Friday and one common sub-cycle for Saturday and Sunday. However, Saturday and Sunday also show dissimilarities in the patterns. Therefore, a second approach involves one sub-cycle for Monday through Friday and separate sub-cycles for Saturday and Sunday. Another interesting pattern is that of Monday and Friday. When the weekend has ended and Monday starts at 12:00 AM, the seasonal terms take the same values as the weekend until 6:00 AM. After that, the pattern becomes the same as that of the weekdays. Moreover, when the weekend begins at Friday, the seasonal pattern of Friday changes from the same pattern as the weekdays to the seasonal values of the weekend. However, most of the time the patterns for Monday and Friday display similar patterns as Tuesday till Thursday. Therefore, the following MS processes are investigated:

- MS(2; 48, 336): A common sub-cycle for Monday through Friday and a common sub-cycle for Saturday and Sunday.
- MS(3; 48,336): A common sub-cycle for Monday through Friday and two separate sub-cycles for Saturday and Sunday.

The selection procedure for the model with multiplicative seasonality

The smoothing parameters of the multiplicative MS(2; 48, 336) and MS(3; 48, 336) processes are estimated using the first $n - q = 2016$ time periods (i.e. the first 6 weeks). The estimated parameter values can be found in Appendix D.1. The one-step ahead forecasts are produced for observations $n - q + 1 = 2017$ till $n = 2688$. The MAPE(1) values resulting from the one-step ahead forecast errors for both methods are shown in Table 4. The MS(2; 48, 336) process gives a lower MAPE(1) value than the MS(3; 48, 336) process. Moreover, the MS(2; 48, 336) process only needs to estimate 6 parameter values and 98 seed values, whereas the MS(3; 48, 336) process needs 11 parameters estimates and 146 seed values. Therefore, we continue with imposing the restrictions on the MS(2;

48, 336) model. The MAPE(1) values resulting from the one-step ahead forecasts produced by the MS(2; 48, 336) with the various restrictions are shown in Table 4. Imposing any of the three restrictions does not improve the accuracy of the one-step ahead forecasts, because all MAPE(1) values are larger than that of the MS(2; 48, 336) model without any restrictions.

Table 4: Withheld sample MAPE(1) in MS model selection.

	Model	Restriction	MAPE(1)	Parameters	Seed values
Step 1	MS(2; 48, 336)	none	1.757	6	98
	MS(3; 48, 336)	none	2.308	11	146
Step 2	MS(2; 48, 336)	1	1.937	3	98
	MS(2; 48, 336)	2	2.077	3	98
	MS(2; 48, 336)	3	1.981	4	98

Estimation of the MS(2; 48, 336) process using the estimation sample of the first $n = 2688$ (i.e. the first 8 weeks) gives the following parameter estimates:

$$\hat{\alpha} = 1.000,\ \hat{\beta}^* = 0.000 \text{ and } \hat{\mathbf{\Gamma}} = \begin{bmatrix} 0.497 & 0.000 \\ 0.167 & 0.221 \end{bmatrix}.$$

The estimated parameter for the level takes the maximum value of 1.000 which means that relatively more weight is given to more recent observations. The estimated parameter $\hat{\beta}^*$ for the trend takes the value zero, which can be explained by the seasonality that dominates the variability in the data. The estimated smoothing parameters in $\hat{\mathbf{\Gamma}}$ show that the seasonal sub-cycles are updated during time periods of the same sub-cycle, but also during time periods of the other sub-cycle. This can be explained by writing the multiplicative MS(2; 48, 336) process as follows:

$$\begin{bmatrix} I_{1,t} \\ I_{2,t} \end{bmatrix} = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} d_{1,t} \\ d_{2,t} \end{bmatrix} \left(\frac{X_t}{S_{t-1} + T_{t-1}} \right) + \begin{bmatrix} I_{1,t-s_1} \\ I_{2,t-s_1} \end{bmatrix} - \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} d_{1,t} \\ d_{2,t} \end{bmatrix} \begin{bmatrix} d_{1,t} & d_{2,t} \end{bmatrix} \begin{bmatrix} I_{1,t-s_1} \\ I_{2,t-s_1} \end{bmatrix}.$$

From this formulation it is easier to see that at a time period t the first sub-cycle for the weekdays $I_{1,t}$ can be updated using the sub-cycle for the weekdays $I_{1,t-s_1}$, but also using the sub-cycle for the weekend $I_{2,t-s_1}$. Moreover, the second sub-cycle for the weekend $I_{2,t}$ can be updated using the sub-cycle for the weekend $I_{2,t-s_1}$, but also using the sub-cycle for the weekdays $I_{1,t-s_1}$.

However, the sample autocorrelation function of the residuals from the in-sample one-step ahead forecasts show significant spikes at different lags. The ACF plot is shown in Figure 5.7. This indicates that there is remaining autocorrelation in the one-step ahead forecast errors. Therefore, an AR(1) model is fitted to the one-step ahead errors for the same reasons as with the Holt-Winters models described in Sections 4.5 and 5.1. Again, the model selection procedure as described in Section 5.3.1 is performed for the MS(2; 48, 336) and the MS(3; 48, 336) processes with the implementation of the different restrictions. The results of the model selection procedure for the models with adjustment for autocorrelation are shown in Table 5. The estimated parameter values for each model can be found in Appendix D.2. The MAPE(1) values have reduced substantially for all the considered models from those in Table 4. This means that the forecast accuracy for the withheld sample has improved by adjusting for the autocorrelation in the residuals. The model selection procedure in Table 5 shows that the MS(2; 48, 336) without restrictions produces the lowest MAPE(1) value of 0.528.

Estimation of the MS(2; 48, 336) process without any restrictions and with adjustment for autocorrelated errors is done using the estimation sample of the first $n = 2688$ (i.e. the first 8 weeks) and gives the following parameter estimates:

$$\hat{\alpha} = 0.000,\ \hat{\beta}^* = 0.000,\ \hat{\mathbf{\Gamma}} = \begin{bmatrix} 1.000 & 0.408 \\ 0.917 & 0.505 \end{bmatrix} \text{ and } \hat{\phi} = 0.925.$$

The estimated parameter values are similar to those of the Holt-Winters methods in Table 3. The smoothing equation for the level seems to be replaced by the AR(1) model, because $\hat{\alpha}$ takes the minimum value and $\hat{\phi}$ almost reaches its maximum with a value of 0.925. The estimated parameter for the trend is zero and the seasonal sub-cycles are updated during time periods of the same sub-cycles and during time periods of the other sub-cycle.

The ACF plot of the one-step ahead forecasts errors from the estimation sample is shown in Figure 5.8. The amount of spikes has reduced substantially, which means that the AR(1) model for the residuals reduced the remaining autocorrelation. However, the ACF plot in Figure 5.8 still shows some spikes, especially around lag 14. This indicates that there is still some autocorrelation left in the residuals.

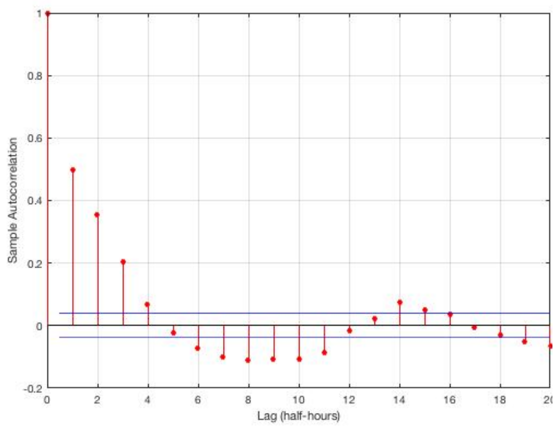

Figure 5.7: The ACF plot of the residuals from the MS(2; 48, 336) process without AR(1) term.

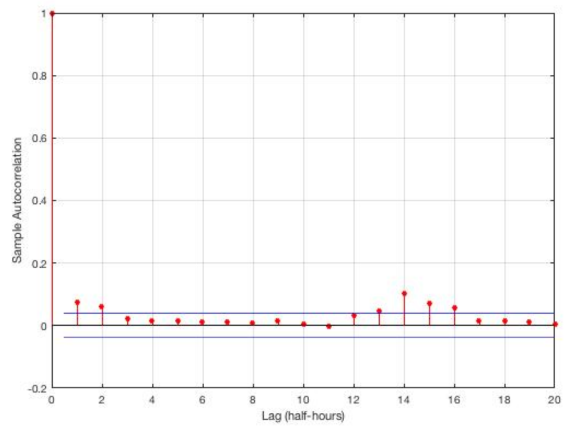

Figure 5.8: The ACF plot of the residuals from the MS(2; 48, 336) process with AR(1) term.

Table 5: Withheld sample MAPE(1) in MS model selection with AR adjustment.

	Model	Restriction	MAPE(1)	Parameters	Seed values
Step 1	MS(2; 48, 336)	none	0.528	7	98
	MS(3; 48, 336)	none	0.614	12	146
Step 2	MS(2; 48, 336)	1	0.551	4	98
	MS(2; 48, 336)	2	0.690	4	98
	MS(2; 48, 336)	3	0.553	5	98

The selection procedure for the model with additive seasonality

So far, only the MS processes that involve multiplicative seasonality have been considered. However, the results in Section 5.2 have shown slight improvements when modelling the electricity demand with the additive formulation of the double seasonal Holt-Winters method rather than the multiplicative version. Therefore, the model selection procedure for the MS process with additive seasonality will also be considered. The models include adjustment for autocorrelated errors by adding an AR(1) term. The additive MS processes without adjustment for autocorrelation are also considered, but they did not provide any improvement compared to the multiplicative version of the models. The models are estimated using the first 6 weeks of observations and the estimated parameter values can be found in Appendix D.3. The MAPE(1) results of the model selection procedure are shown in Table 7.

Table 6: Withheld sample MAPE(1) in MS model selection with AR adjustment.

	Model	Restriction	MAPE(1)	Parameters	Seed values
Step 1	MS(2; 48, 336)	none	0.533	7	98
	MS(3; 48, 336)	none	0.447	12	146
Step 2	MS(3; 48, 336)	1	0.452	4	146
	MS(3; 48, 336)	2	0.795	4	146
	MS(3; 48, 336)	3	0.429	5	146

In the first step of the model selection procedure the MS(3; 48, 336) model produces the lowest MAPE(1) value. However, 12 parameters has to be estimated instead of 7 for the MS(2; 48, 336) model. Also, 146 seed values are needed, which is almost 1.5 times as much as required for the MS(2; 48, 336) model. The second step in the model selection procedure shows that imposing restriction 3 on the MS(3; 48, 336) model produces the most accurate one-step ahead forecasts of all the MS processes considered. This also reduces the amount of parameters substantially. Whereas 12 parameters where needed when estimating the MS(3; 48, 336) model without restrictions, the model with restriction 3 only needs 5 parameters to be estimated. The MS(3; 48, 336) model with restriction 3 is estimated using the estimation sample of 8 weeks. Minimizing the sum of squared one-step ahead errors gives the following parameter estimates:

$$\hat{\alpha} = 0.000,\ \hat{\beta}^{*} = 0.000,\ \hat{\mathbf{\Gamma}} = \begin{bmatrix} 0.512 & 0.297 \\ 0.297 & 0.512 \end{bmatrix} \text{ and } \hat{\phi} = 0.985.$$

5.3.2 Forecasting and comparison

The multiplicative MS(2; 48, 336) process without restriction and the additive MS(3; 48, 336) process with restriction 3 are both used for the out of sample forecasts of the short-term electricity demand data, because these models produced the lowest withheld sample MAPE(1) values when applying the model selection procedure. Only the methods including AR(1) term are used for forecasting out-of-sample, because without AR(1) adjustment the forecast accuracy for longer lead times becomes extremely poor. Moreover, it makes sense to compare the forecast accuracy of all methods discussed in this paper including AR term, because adjusting for autocorrelation have always improved the results.

Figure 5.9 shows the out of sample MAPE(h) values for the multiplicative MS(2; 48, 336) model and the additive MS(3; 48, 336) model with restriction 3. The MS(3; 48, 336) process outperforms the MS(2; 48, 336) model for all lead times. This suggests that the forecast accuracy improves when the sub-cycles for Saturday and Sunday are taken separately instead of taking one common sub-cycle for the weekend.

Table 7: Comparison of the out of sample MAPE(1) values, amount of parameters and seed values for all models with AR adjustment.

Model	Seasonality	Restriction	MAPE(1)	Parameters	Seed values
HW(48)	Multiplicative	none	0.513	4	50
HW(336)	Multiplicative	none	0.485	4	338
HW(48, 336)	Multiplicative	none	0.350	5	386
HW(48, 336)	Additive	none	0.341	5	386
MS(2; 48, 336)	Multiplicative	none	0.617	7	98
MS(3; 48, 336)	Additive	3	0.396	5	146

Table 7 contains the out of sample MAPE(1) values resulting from forecasting with the Holt-Winters model for within-day seasonality (HW(48)), the Holt-Winters model for within-week seasonality (HW(336)), the double seasonal Holt-Winters models (HW(48, 336)), the MS(2; 48, 336) model and the MS(3; 48, 336) model with restriction 3. The multiplicative MS(2; 48, 336) model produces the

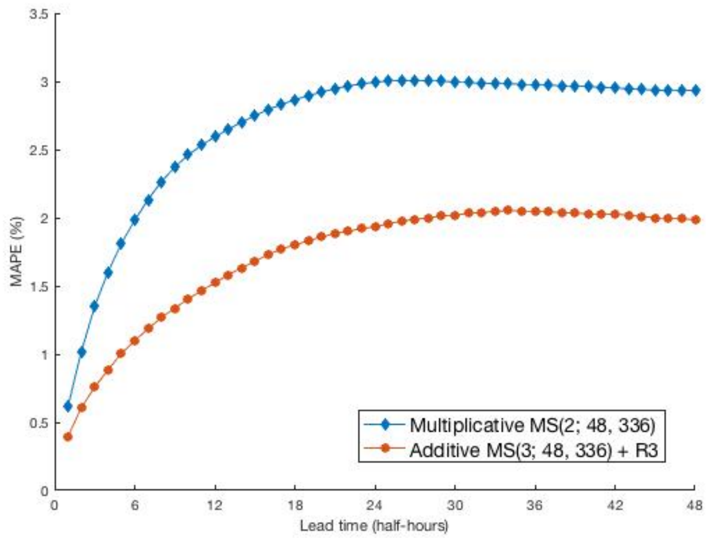

Figure 5.9: Comparison of the out of sample MAPE(h) results for the MS processes including AR(1) model for the residuals.

highest out of sample MAPE(1) value and thus has the poorest one-step ahead forecast accuracy of all models. However, for longer lead times the MAPE(h) results are better than the Holt-Winters model for within-day seasonality. The additive MS(3; 48, 336) model outperforms both the Holt-Winters model for within-day seasonality and the Holt-Winters model for within-week seasonality. The forecast accuracy for lead times up to a half-hour ahead for the MS(3; 48, 336) model is with a value of 0.396 slightly worse than that of the multiplicative and additive double seasonal Holt-Winters models, which produce MAPE(1) values of 0.350 and 0.341 respectively. However, when the MAPE(h) values in Figure 5.9 are compared to those in Figure 5.4, it is clear that the double seasonal Holt-Winters methods outperform the MS(3; 48, 336) process for all lead times. In fact, the additive MS(3; 48, 336) model performs almost twice as bad as the double seasonal Holt-Winters method for lead times beyond 12 hours. Although the performance of the MS(3; 48, 336) process is not as promising as that of the double seasonal Holt-Winters methods, the seed values to be estimated have reduced significantly. Whereas 386 seed values are needed for the double seasonal Holt-Winters methods, the MS(3; 48, 336) process only needs 146 seed values, which is a reduction of more than 60%.

6 Discussion and Conclusion

In this paper a time series of short-term electricity demand in England and Wales is modelled using exponential smoothing methods. The data shows two seasonal patterns, one within each day and one within each week. The Holt-Winters method for within-day seasonality and the Holt-Winters method for within-week seasonality are used to model the data. The Holt-Winters method for within-day seasonality showed very poor forecasting performances. This is due to the fact that this method cannot accommodate the within-week seasonality, which dominates the variation in the data. Therefore, the Holt-Winters method for within-week seasonality showed more accurate forecasts. However, these Holt-Winters methods can only accommodate one seasonal pattern. Taylor (2003) introduced the multiplicative double seasonal Holt-Winters method which is able to capture both seasonalities in the time series. Because the one-step ahead forecast errors showed sizeable first-order autocorrelation, an AR(1) model is fitted to the residuals. When modelling the Holt-Winters methods with AR(1) adjustment, the results of all methods improved and the double seasonal Holt-Winters method outperformed the Holt-Winters method for within-day and within-week seasonality for all lead times. Besides the double seasonal Holt-Winters method with multiplicative seasonality, the model for additive seasonal patterns is also considered and showed more accurate forecasts for almost all lead times. This suggests that the seasonal variation is rather constant and does not depend on the mean level of the time series.

Besides the Holt-Winters methods discussed in Taylor (2003), the MS(r; s_1, s_2) processes introduced by Hyndman et al. (2008) are discussed. This method combines certain seasonal sub-cycles into one sub-cycle, which reduces the amount of seed values that have to be estimated. It also allows for seasonal sub-cycles to be updated during time periods of the same sub-cycle and other sub-cycles with different values for the smoothing parameters. Restrictions are imposed to reduce the number of smoothing parameters and possibly improve the forecast accuracy. Inspecting the seasonal terms for each day of the last week of the estimation sample suggested a MS(2; 48, 336) model with one common sub-cycle for the weekdays and one common sub-cycle for the weekends and a MS(3; 48, 336) model with one common sub-cycle for the weekdays and two separate sub-cycles for Saturday and Sunday. Both the multiplicative and the additive formulations are considered. Because of remaining autocorrelation in the residuals, an AR(1) model is fitted to the one-step ahead errors. This improved the withheld sample MAPE(1) values substantially. The model selection procedure resulted in the MS(2; 48, 336) model without restrictions for the multiplicative formulation and it resulted in the MS(3; 48, 336) model with restriction 3 for the additive formulation. The MS(3; 48, 336) model outperformed the MS(2; 48, 336) model for all lead times. However, the model did not gave more accurate h-step ahead forecasts than the double seasonal Holt-Winters methods. This is probably due to the fact that some information about the seasonal terms is lost when combining all sub-cycles for Monday through Friday in one sub-cycle. However, the amount of seeds to be estimated is reduced substantially. Whereas the double seasonal Holt-Winters model needs to estimate 386 seed values, the MS(3; 48, 336) process needs to estimate 146 seed values and the MS(2; 48, 336) process only needs to estimate 98 seed values. It seems like there is a trade-off between forecast accuracy and the amount of seeds to be estimated. To further investigate this trade-off, more different MS(r; s_1, s_2) processes should be investigated for different values of r and different combinations of common sub-cycles.

Summarizing, the double seasonal Holt-Winters methods including adjustment for autocorrelation produced the most accurate forecasts for the short-term electricity demand data of all models considered in this paper. However, it would be useful to consider other forecasting methods as well. For example, Taylor (2012) uses discount weighted regression, cubic splines and singular value decomposition to model and forecast the British short-term electricity demand. Taylor, de Menezes, and McSharry (2006) use a regression method with principle component analysis to model the electricity demand in England and Wales. Moreover, besides point forecasts one could use prediction intervals for the forecasts as in Hyndman et al. (2008).

Appendices

A From error correction form to component form

The error correction formulation of the additive MS(r; s_1, s_2) model of Hyndman et al. (2008) is as follows:

$$
\begin{aligned}
X_t &= S_{t-1} + T_{t-1} + \mathbf{d}_t^\top \mathbf{I}_{t-s_1} + \epsilon_t, && \text{(17a)}\\
S_t &= S_{t-1} + T_{t-1} + \alpha\epsilon_t, && \text{(17b)}\\
T_t &= T_{t-1} + \beta\epsilon_t, && \text{(17c)}\\
\mathbf{I}_t &= \mathbf{I}_{t-s_1} + \mathbf{\Gamma}\mathbf{d}_t\epsilon_t, && \text{(17d)}\\
\hat{X}_t(1) &= S_{t-1} + T_{t-1} + \mathbf{d}_t^\top \mathbf{I}_{t-s_1}, && \text{(17e)}
\end{aligned}
$$

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

In order to go from error correction representation to component form representation, we first get the expression for ϵ_t and then substitute this in the equations for the level, the trend and the seasonal component. The expression for ϵ_t is

$$
\epsilon_t = X_t - S_{t-1} - T_{t-1} - \mathbf{d}_t^\top \mathbf{I}_{t-s_1}. \tag{18}
$$

Substituting ϵ_t in the level, the trend and the seasonal component equations gives

$$
\begin{aligned}
S_t &= S_{t-1} + T_{t-1} + \alpha(X_t - S_{t-1} - T_{t-1} - \mathbf{d}_t^\top \mathbf{I}_{t-s_1})\\
&= \alpha(X_t - \mathbf{d}_t^\top \mathbf{I}_{t-s_1}) + (1-\alpha)(S_{t-1} + T_{t-1}), && \text{(19a)}\\
T_t &= T_{t-1} + \beta(X_t - S_{t-1} - T_{t-1} - \mathbf{d}_t^\top \mathbf{I}_{t-s_1})\\
&= T_{t-1} + \beta\left(\frac{S_t - S_{t-1} - T_{t-1}}{\alpha}\right)\\
&= T_{t-1} + \beta^*(S_t - S_{t-1} - T_{t-1})\\
&= \beta^*(S_t - S_{t-1}) + (1-\beta^*)T_{t-1}, && \text{(19b)}\\
\mathbf{I}_t &= I_{t-s_1} + \mathbf{\Gamma}\mathbf{d}_t(X_t - S_{t-1} - T_{t-1} - \mathbf{d}_t^\top \mathbf{I}_{t-s_1})\\
&= \mathbf{\Gamma}\mathbf{d}_t(X_t - S_{t-1} - T_{t-1}) + \mathbf{I}_{t-s_1} - \mathbf{\Gamma}\mathbf{d}_t\mathbf{d}_t^\top \mathbf{I}_{t-s_1}, && \text{(19c)}\\
\hat{X}_t(1) &= S_{t-1} + T_{t-1} + \mathbf{d}_t^\top \mathbf{I}_{t-s_1}, && \text{(19d)}
\end{aligned}
$$

where $\beta^* = \frac{\beta}{\alpha}$.

B Formulas for h-step ahead forecasts for models that include AR(1) term

The h-step-ahead forecasts in the multiplicative double seasonal Holt-Winters model in Equation 3a become

$$\begin{aligned}\hat{X}_t(h) &= (S_t + hT_t)D_{t-s_1+h}W_{t-s_2+h} + \phi^h e_t \\ &= (S_t + hT_t)D_{t-s_1+h}W_{t-s_2+h} + \phi^h\{X_t - (S_{t-1} + T_{t-1})D_{t-s_1}W_{t-s_2}\}.\end{aligned} \tag{20}$$

The h-step-ahead forecasts in the additive double seasonal Holt-Winters model in Equation 4a become

$$\begin{aligned}\hat{X}_t(h) &= S_t + hT_t + D_{t-s_1+h} + W_{t-s_2+h} + \phi^h e_t \\ &= S_t + hT_t + D_{t-s_1+h} + W_{t-s_2+h} + \phi^h\{X_t - (S_{t-1} + T_{t-1} + D_{t-s_1} + W_{t-s_2})\}.\end{aligned} \tag{21}$$

The h-step ahead forecasts in the multiplicative MS(r; s_1, s_2) process in Equation 7a become

$$\begin{aligned}\hat{X}_t(h) &= (S_{t-1} + T_{t-1})\mathbf{d}_t^\top \mathbf{I}_{t-s_1+h} + \phi^h e_t \\ &= (S_{t-1} + T_{t-1})\mathbf{d}_t^\top \mathbf{I}_{t-s_1+h} + \phi^h\{X_t - (S_{t-1} + T_{t-1})\mathbf{d}_{t-1}^\top \mathbf{I}_{t-s_1}\}.\end{aligned} \tag{22}$$

The h-step ahead forecasts in additive MS(r; s_1, s_2) process in Equation 9a become

$$\begin{aligned}\hat{X}_t(h) &= S_{t-1} + hT_{t-1} + \mathbf{d}_t^\top \mathbf{I}_{t-s_1+h} + \phi^h e_t \\ &= S_{t-1} + hT_{t-1} + \mathbf{d}_t^\top \mathbf{I}_{t-s_1+h} + \phi^h\{X_t - (S_{t-1} + T_{t-1} + \mathbf{d}_{t-1}^\top \mathbf{I}_{t-s_1})\}.\end{aligned} \tag{23}$$

C Sample autocorrelation plots of the Holt-Winters methods

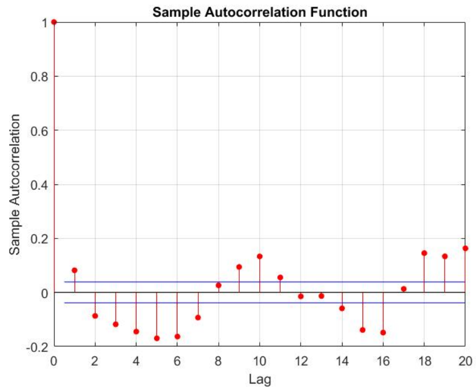

Figure .1: ACF of the residuals from Holt-Winters for within-day seasonality.

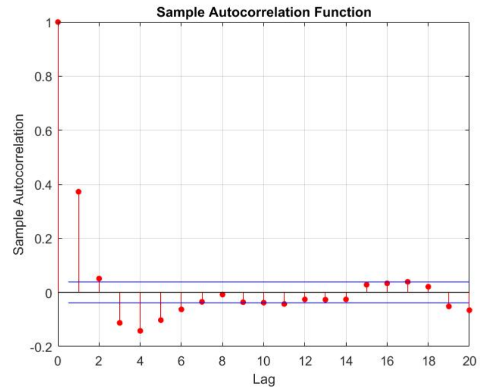

Figure .2: ACF of the residuals from Holt-Winters for within-week seasonality.

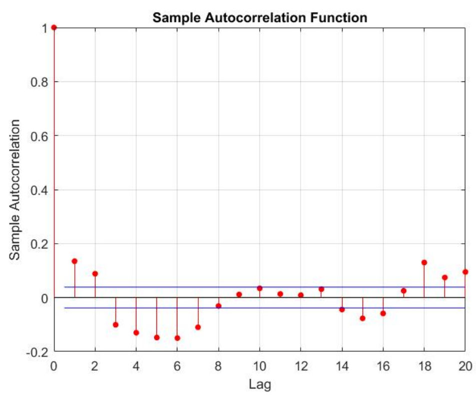

Figure .3: ACF of the residuals from Holt-Winters for within-day seasonality with AR(1) adjustment.

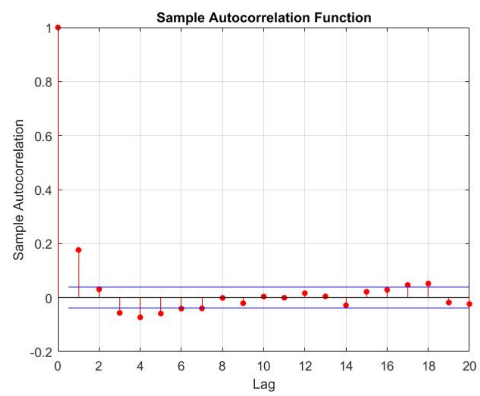

Figure .4: ACF of the residuals from Holt-Winters for within-week seasonality with AR(1) adjustment.

D Parameter estimates of model selection procedures

D.1 The MS process with multiplicative seasonality and no adjustment for autocorrelation.

The MS(2; 48, 336) process without restrictions:

$$\alpha = 1.000,\ \hat{\beta} = 0.000 \text{ and } \hat{\mathbf{\Gamma}} = \begin{bmatrix} 0.489 & 0.000 \\ 0.167 & 0.212 \end{bmatrix}.$$

The MS(3; 48, 336) process without restrictions:

$$\alpha = 1.000,\ \hat{\beta} = 0.000 \text{ and } \hat{\mathbf{\Gamma}} = \begin{bmatrix} 0.414 & 0.504 & 0.243 \\ 0.372 & 0.436 & 0.305 \\ 0.500 & 0.500 & 0.361 \end{bmatrix}.$$

The MS(2; 48, 336) process with restriction 1:

$$\alpha = 1.000,\ \hat{\beta} = 0.000 \text{ and } \hat{\mathbf{\Gamma}} = \begin{bmatrix} 0.336 & 0.000 \\ 0.000 & 0.336 \end{bmatrix}.$$

The MS(2; 48, 336) process with restriction 2:

$$\alpha = 1.000,\ \hat{\beta} = 0.000 \text{ and } \hat{\mathbf{\Gamma}} = \begin{bmatrix} 0.362 & 0.362 \\ 0.362 & 0.362 \end{bmatrix}.$$

The MS(2; 48, 336) process with restriction 3:

$$\alpha = 1.000,\ \hat{\beta} = 0.000 \text{ and } \hat{\mathbf{\Gamma}} = \begin{bmatrix} 0.421 & 0.258 \\ 0.258 & 0.421 \end{bmatrix}.$$

D.2 The MS process with multiplicative seasonality and adjustment for autocorrelation.

The MS(2; 48, 336) process without restrictions:

$$\hat{\alpha} = 0.000,\ \hat{\beta} = 0.000,\ \hat{\mathbf{\Gamma}} = \begin{bmatrix} 0.492 & 0.371 \\ 0.305 & 0.383 \end{bmatrix} \text{ and } \hat{\phi} = 0.970.$$

The MS(3; 48, 336) process without restrictions:

$$\hat{\alpha} = 0.000,\ \hat{\beta} = 0.001,\ \hat{\mathbf{\Gamma}} = \begin{bmatrix} 0.511 & 0.432 & 0.316 \\ 0.383 & 0.515 & 0.428 \\ 0.500 & 0.500 & 0.428 \end{bmatrix} \text{ and } \hat{\phi} = 0.967.$$

The MS(2; 48, 336) process with restriction 1:

$$\hat{\alpha} = 0.000,\ \hat{\beta} = 0.000,\ \hat{\mathbf{\Gamma}} = \begin{bmatrix} 0.430 & 0.000 \\ 0.000 & 0.430 \end{bmatrix} \text{ and } \hat{\phi} = 0.973.$$

The MS(2; 48, 336) process with restriction 2:

$$\hat{\alpha} = 0.000,\ \hat{\beta} = 0.000,\ \hat{\mathbf{\Gamma}} = \begin{bmatrix} 0.350 & 0.350 \\ 0.350 & 0.350 \end{bmatrix} \text{ and } \hat{\phi} = 0.963.$$

The MS(2; 48, 336) process with restriction 3:

$$\hat{\alpha} = 0.000,\ \hat{\beta} = 0.000,\ \hat{\mathbf{\Gamma}} = \begin{bmatrix} 0.417 & 0.265 \\ 0.265 & 0.417 \end{bmatrix} \text{ and } \hat{\phi} = 0.970.$$

D.3 The MS process with additive seasonality and adjustment for autocorrelation.

The MS(2; 48, 336) process without restrictions:

$$\hat{\alpha} = 0.014,\ \hat{\beta} = 0.000,\ \hat{\mathbf{\Gamma}} = \begin{bmatrix} 0.522 & 0.055 \\ 0.369 & 0.290 \end{bmatrix} \text{ and } \hat{\phi} = 0.980.$$

The MS(3; 48, 336) process without restrictions:

$$\hat{\alpha} = 0.016,\ \hat{\beta} = 0.001,\ \hat{\mathbf{\Gamma}} = \begin{bmatrix} 0.521 & 0.149 & 0.132 \\ 0.403 & 0.780 & 0.016 \\ 0.361 & 0.538 & 0.523 \end{bmatrix} \text{ and } \hat{\phi} = 0.978.$$

The MS(3; 48, 336) process with restriction 1:

$$\hat{\alpha} = 0.001,\ \hat{\beta} = 0.000,\ \hat{\mathbf{\Gamma}} = \begin{bmatrix} 0.614 & 0.000 & 0.000 \\ 0.000 & 0.614 & 0.000 \\ 0.000 & 0.000 & 0.614 \end{bmatrix} \text{ and } \hat{\phi} = 0.979.$$

The MS(3; 48, 336) process with restriction 2:

$$\hat{\alpha} = 0.000,\ \hat{\beta} = 0.000,\ \hat{\mathbf{\Gamma}} = \begin{bmatrix} 0.387 & 0.387 & 0.387 \\ 0.387 & 0.387 & 0.387 \\ 0.387 & 0.387 & 0.387 \end{bmatrix} \text{ and } \hat{\phi} = 0.992.$$

The MS(3; 48, 336) process with restriction 3:

$$\hat{\alpha} = 0.000,\ \hat{\beta} = 0.000,\ \hat{\mathbf{\Gamma}} = \begin{bmatrix} 0.554 & 0.315 & 0.315 \\ 0.315 & 0.554 & 0.315 \\ 0.315 & 0.315 & 0.554 \end{bmatrix} \text{ and } \hat{\phi} = 0.985.$$

References

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1993). *Time Series Analysis: Forecasting and Control.* Prentice Hall, New Jersey.

Chatfield, C. (1987). The Holt-Winters Forecasting Procedure. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *27*, 264–279. doi: 10.2307/2347162

Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A Seasonal-Trend Decomposition Procedure Based on Loess. *Journal of Official Statistics*, *6*, 3–73.

De Livera, A. M., Hyndman, R. J., & Snyder, R. D. (2010). Forecasting Time Series with Complex Seasonal Patterns using Exponential Smoothing. *Department of Econometrics and Business Statistics Working Paper Series 15/09*.

Gardner, E. S. (1985). Exponential Smoothing: The State of the Art. *International Journal of Forecasting*, *4*, 1–28. doi: 10.1002/for.3980040103

Hyndman, R. J., & Athanasopoulos, G. (2012). *Forecasting: Principles and Practice.* OTexts.

Hyndman, R. J., Gould, P. G., Koehler, A. B., Ord, J. K., Snyder, R. D., & Vahid-Araghi, F. (2008). Forecasting Time Series with Multiple Seasonal Patterns. *European Journal of Operational Research*, *191*, 207–220. doi: 10.1016/j.ejor.2007.08.024

Souza, R. C., Barros, M., & de Miranda, C. V. C. (2007). *Short Term Load Forecasting Using Double Seasonal Exponential Smoothing and Interventions to Account for Holidays and Temperature Effects* (Tech. Rep.). Pontifícia Universidade Católica do Rio de Janeiro.

Taylor, J. W. (2003). Short-Term Electricity Demand Forecasting Using Double Seasonal Exponential Smoothing. *The Journal of the Operational Research Society*, *54*, 799–805. doi: 10.1057/palgrave.jors.2601589

Taylor, J. W. (2010). Exponentially Weighted Methods for Forecasting Intra-day Time Series with Multiple Seasonal Cycles . *International Journal of Forecasting*, *26*, 627–646.

Taylor, J. W. (2012). Short-Term Load Forecasting with Exponentially Weighted Methods. *IEEE Transactions of Power Systems*, *27*, 458–464.

Taylor, J. W., & Buizza, R. (2003). Using Weather Ensemble Predictions in Electricity Demand Forecasting . *International Journal of Forecasting*, *19*, 57–70.

Taylor, J. W., de Menezes, L. M., & McSharry, P. E. (2006). A Comparison of Univariate Methods for Forecasting Electricity Demand up to a Day Ahead. *International Journal of Forecasting*, *22*, 1–16.