Erasmus University Rotterdam

Thesis BSc2 Econometrics/Economics

# Frequency domain models versus Holt-Winters, short-term electricity demand forecasting in England and France

*Author*

Jeroen de Groot (466 544)

*Supervisor*

Mikhail Zhelonkin

*Second Assessor*

Karel de Wit

July 12, 2021

**Abstract**

This paper evaluates the predictive accuracy of different Holt-Winters methods for a twelve week time series of electricity demand in England and Wales. Double seasonal Holt-Winters outperforms Holt-Winters and a well defined multiplicative seasonal ARIMA model. Moreover, this paper evaluates the performance of the best Holt-Winters method, double seasonal Holt-Winters, with models utilising the frequency domain. Frequency domain models are not able to outperform double seasonal Holt-Winters for electricity demand in England and Wales. However, they are able to outperform double seasonal Holt-Winters for a more complex, five year, electricity demand time series in France. Short-term electricity demand forecasting utilising the frequency domain is promising for longer time series with more seasonal variance.

The views stated in this thesis are those of the author and not necessarily those of the supervisor, second assessor, Erasmus School of Economics or Erasmus University Rotterdam.

# Contents

# 1 Introduction

Accurate predictions of demand are in general vital across industries. Accurate predictions are beneficial since it reduces the number of times that inventory problems occur for companies, empty taxis that congest the city for example, Yao et al. (2018). Demand forecasting is especially of importance for goods where storage is not possible, electricity for example, where forecast errors have significant impact on profits, Fan and Chen (2006), and where an increase of 1% in the forecast error could lead to the loss of millions of dollars, Al-Musaylh et al. (2018). Accurate demand predictions for electricity can greatly reduce financial but also environmental costs. Ghalehkhondabi et al. (2017) show that much research regarding electricity demand forecasting has been done and different methods have been proposed to accurately predict the demand for electricity. They further show that research is still able to continuously improve existing electricity demand methods and successfully introduce new methods. This is also the goal of this paper.

A key characteristic of many demand functions is seasonality. This is the case for electricity demand as well. Electricity demand shows three seasonal cycles, a within-day, within-week and within-year seasonal cycle. Taylor (2003) shows that double seasonal exponential smoothing is a viable method to accurately predict electricity demand and that double seasonal exponential smoothing outperforms a multiplicative seasonal autoregressive integrated moving average (ARIMA) for predicting electricity demand for a time series with a within-day and within-week seasonal pattern. Electricity demand with a within-year seasonal cycle is analysed in Taylor (2010) and it is concluded that a triple seasonal specification of an autoregressive moving average (ARMA) and Holt-Winters outperform their double seasonal specification for this time series. These models are of the class time domain models.

Time domain models display changes in a signal, electricity demand in this case, over time. Changes in a signal can however also be displayed in the frequency domain, where how much of a signal is apparent in a frequency band is shown. The two time series discussed in this paper are well suited for analysis in the frequency domain since they show strong seasonal patterns. Forecasting utilising the frequency domain means that the time series needs to be expressed as a Fourier series. Bloomfield (2004) explains that Fourier series that represent time series data can be expressed by a family of trigonometric functions. This transformed data can be analysed in the frequency domain since the family of trigonometric functions map the variation of the time series on a frequency band. A benefit of this approach is that trigonometric functions are able to model complex seasonalities that are not able to be modelled with exponential smoothing methods, De Livera et al. (2011). For analysis in the frequency domain many parameters have to be estimated. Therefore, the least absolute shrinkage and selection operator (LASSO) from Tibshirani (1996) is useful. This criterion shrinks parameter estimates, reduces the parameter set and therefore improves forecast accuracy and ease of parameter interpretation.

To my knowledge there is no existing research that compares the performance of time domain models with frequency domain models in forecasting short-term electricity demand for a time series with three strong seasonal patterns. The aim of this paper is to contribute to existing literature by answering the following research question: *Can models based on Fourier series that incorporate seasonal patterns outperform a double seasonal Holt-Winters method in forecasting short-term electricity demand?* In this paper I find that a model based on Fourier series that incorporates three seasonal patterns, which I refer to as a Triple Fourier model, is able to outperform double seasonal Holt-Winters for a five year time series of electricity demand in France, if error terms are correctly modelled. A Double Fourier model does not outperform a double seasonal Holt-Winters method.

This paper is split in six sections. Section 1 contains the introduction. In Section 2 I discuss the current standing in the literature for short-term electricity demand forecasting with time- and frequency domain models. In Section 3 I give a concise overview of key characteristics of the data that I use in this research. Then, in Section 4 I introduce the models used in Taylor (2003) that I use for analysis of electricity demand in England and Wales first. In the second part of Section 4 I introduce a modification of these models for analysis of electricity demand in France and I conclude Section 4 by introducing the Fourier models, LASSO estimation, and the metrics I use for evaluation. I discuss my results in Section 5. The last section, Section 6, consists of the discussion and conclusion. Here, I conclude my research, evaluate my methods and suggest recommended future research.

## 2 Literature Review

This paper discusses univariate forecasting of short-term electricity demand in the time domain and in the frequency domain. Short-term electricity forecasting in the time domain is discussed in many papers and hence literature regarding forecasting in the time domain is discussed first.

This paper mainly builds upon the paper Taylor (2003), therefore that paper will be the starting point for the literature review in the time domain. Taylor (2003) forecasts half-hourly short-term electricity demand with the standard and an adapted Holt-Winters exponential smoothing formulation and a multiplicative seasonal ARIMA model. Holt-Winters exponential smoothing is introduced in Winters (1960). The method is introduced to be able to forecast sales that show seasonality. It is the first method that could accurately do this without the need for extensive storage space at the time. Holt-Winters is still a popular forecasting procedure for time-series that show seasonality. The Holt-Winters exponential smoothing procedure introduced in Winters (1960) can only incorporate one seasonal pattern. Many time series show more seasonal patterns however, one example being electricity demand. Taylor (2003) therefore introduces double seasonal Holt-Winters to incorporate a second seasonal pattern. But, electricity demand shows a third seasonal pattern as well, one on yearly basis. To incorporate that seasonal pattern Taylor (2010) introduces triple seasonal Holt-Winters. Holt-Winters methods are recursive and initial values for the parameters are important for the forecast accuracy. Trull et al. (2020) provide a framework for initialization methods and which method to use for the different parameters in Holt-Winters with multiple seasonalities. Another recent paper by Jiang et al. (2020) shows how Holt-Winters is enhanced for situations with insufficient training data and that Holt-Winters performs well in such situations. Thus, the Holt-Winters exponential smoothing formulation is flexible, can forecast time series with multiple seasonal patterns and is particularly useful for forecasting sales and short-term (electricity) demand, Chatfield and Yar (1988).

The Holt-Winters method belongs to the class of exponential smoothing methods. A key characteristic of all methods that fall under the umbrella of exponential smoothing is that the forecasts are based on a weighted average of past observations. The smoothing methodology is called 'exponential' smoothing as the weights decrease exponentially over time, Hyndman et al. (2008). The exponential smoothing method was first introduced by Brown and Holt in the early 50s. Exponential smoothing became popular since the method gave good results in comparison with more complicated forecasting methods such as Box-Jenkins, Gardner Jr (1985). Smooth transition methods are used for forecasting financial time series as well, one example being a recent paper by Liu et al. (2020). They provide further empirical evidence, on already existing literature, that exponential smoothing methods are among the most robust and accurate models for forecasting daily volatility. Lidiema (2017), De Oliveira and Oliveira (2018), Smyl (2020) and Yang et al. (2018) show other areas in which (Holt-Winters) exponential smoothing is applied. Thus, exponential smoothing methods are widely applicable for forecasting different time series, of which Holt-Winters for electricity demand is one of the many possible applications of exponential smoothing.

Beside Holt-Winters there are other approaches for short-term electricity demand forecasting. Taylor (2003) mentions that ARIMA models are widely used benchmark models for short-term electricity demand forecasting. Therefore, he implements an ARIMA model for forecasting short-term electricity demand as well. ARIMA models were introduced in the 1970s by George Box and Gwilym Jenkins and they are popular for their statistical properties and well explained methodology, according to Box et al. (2015). In addition, the models are flexible in that they can represent multiple types of time series, purely autoregressive (AR), purely moving average (MA) but also combined (ARIMA). A major limitation is that the model assumes as linear correlation structure for the time series observations, Zhang (2003).

Finally, it is also possible to forecast time series data using the frequency domain instead of the time domain. To achieve this, data from the time domain is transformed to the frequency domain through trigonometric functions. The expression of data in the frequency domain via trigonometric functions is called Fourier analysis. The reason trigonometric functions are mostly used for analysing time series data is due to their simple behavior when the scale of time changes, Bloomfield (2004). The foundations for Fourier analysis were laid in Fourier (1822). Fourier methods were however only widely applicable after Cooley and Tukey (1965) introduced an algorithm that significantly reduced the computational time. Two well cited books discuss the applications of Fourier series for time series, Fuller (2009) and Bloomfield (2004). Both books explain well how periodic time series can be analysed in the frequency domain by extracting cycles from the data via trigonometric functions. Trigonometric functions are also crucial in the paper by De Livera et al. (2011). They introduce a new state space modeling framework utilising Fourier series for forecasting time series with complex seasonalities. De Livera et al. (2011) show that the trigonometric formulation, based on Fourier series, can decompose the seasonal time series in non-integer cycles, i.e. 365.25. Existing univariate models such as Holt-Winters can only incorporate integer cycles, i.e. 365. Together with Box-Cox transformations and ARMA error correction this modelling framework incorporating trigonometric functions leads to better out-of-sample performance than traditional seasonal exponential smoothing methods. Karabiber and Xydis (2019) implement trigonometric functions for time series analysis as well. They forecast the Danish electricity price. Fourier analysis of time series is also present in other academic fields. Bush et al. (2017) analyse the effect of climate change on phenological activity in ecosystems for example. Thus, Fourier analysis is applicable for time series in different academic fields. In this paper I focus on electricity demand time series. De Livera et al. (2011) consider electricity demand in Turkey where forecasting based on trigonometric functions proves to provide accurate forecasts. In this paper I analyse two different electricity demand time series.

## 3 Data

Two different electricity demand time series are considered in this paper. I use half-hourly electricity demand from England and Wales and half- and bi-hourly electricity demand from France. Data for England and Wales is retrieved from the `forecast` package in R written by Hyndman and Khandakar (2008). The data from England and Wales, Figure 1, is for a period of twelve weeks and includes 4032 observations. Within this period two seasonal patterns are apparent, a within-day seasonal cycle of 48 half-hour periods and a within-week seasonal cycle of 336 half-hour periods. These cycles are visible in the seasonal decomposed plot in Figure 10 in Appendix A, where the third box on the left shows the within-day seasonal cycle and the first box on the right shows the within-week seasonal cycle. The sample is split into an estimation and evaluation sample. The estimation samples contains the first eight weeks of observations, namely 2688 observations. The evaluation sample contains the remaining four weeks of in total 1344 observations.

The data of electricity demand in France, Figure 2, is for a five year period and includes 87648 observations considering half-hourly electricity demand. The estimation sample includes the first four years and the evaluation sample the last year of observations. Bi-hourly demand is constructed by taking the average of four consecutive observations. The half- and bi-hourly data show a third seasonality, namely a within year seasonal cycle of 336 * 52 half-hour periods. Data for France is retrieved from the website of *Réseau de Transport d'Électricité*, RTE (2020). The `mstl()` function from the `forecast` package from R does not detect any seasonalities in the time series with electricity demand in France. Therefore, the time series is more complex than the electricity demand time series for England and Wales, the output of the `mstl()` function is presented in Appendix A in Figure 11. Summary statistics of both time series are shown in Table 1. From the summary statistics and time series plots it is clear that electricity demand is quite constant during late spring, summer and early autumn, except for August. Electricity demand in August is consistently lower in France since most people go on vacation during that month. Electricity demand is less consistent during the winter months since demand is higher and therefore general changes in electricity demand, such as the Christmas holidays, have a larger influence on overall demand.

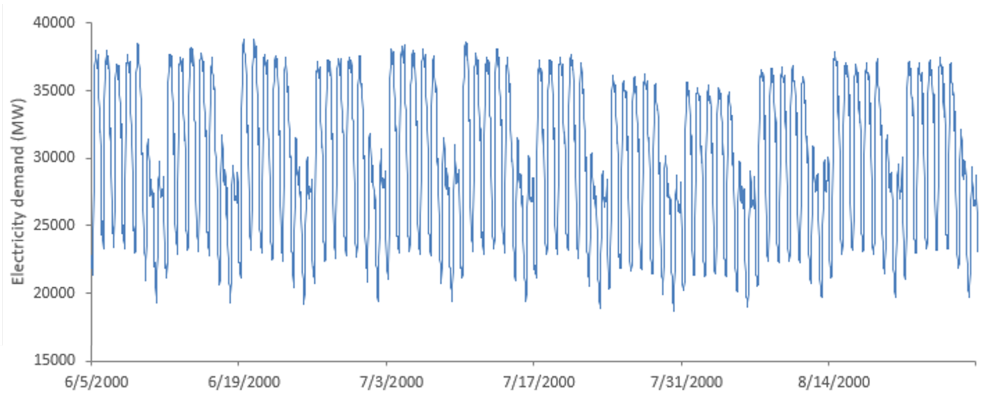


Figure 1: Half-hourly electricity demand in England and Wales from Monday 5th of June 2000 till Sunday 27th of August 2000.

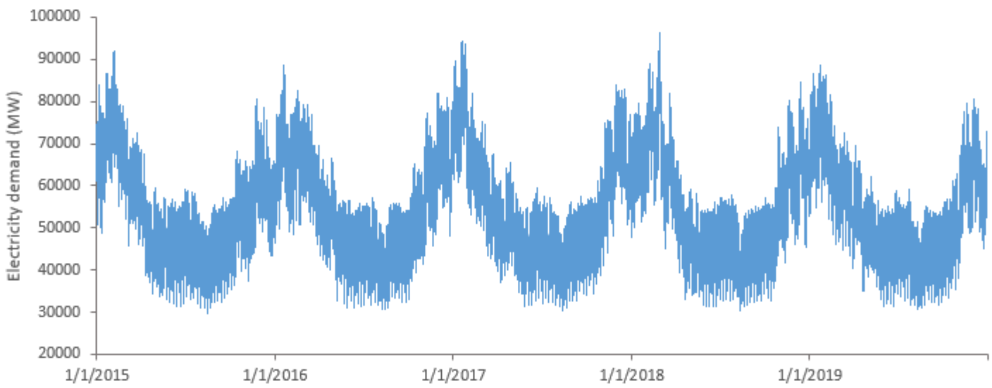


Figure 2: Half-hourly electricity demand in France from 1st of January 2015 till 31st of December 2019.

Table 1: Summary statistics for the whole sample of England and Wales and France.

| | Electricity Demand in England and Wales (MW) | Electricity Demand in France (MW) |
|---|---|---|
| Mean | 29617 | 54274 |
| Maximum | 38777 | 96272 |
| Minimum | 18640 | 29590 |
| Std. Dev. | 5567 | 11911 |
| Skewness | -0.077 | 0.470 |
| Kurtosis | 1.659 | 2.611 |
| Jarque-Bera | 306.285 | 3774.500 |

# 4 Methodology

The methodology section is split in three parts. First, the replication part where all models used in Taylor (2003) are reviewed. Second, the extension part where the adaption of the Holt-Winters method for yearly data and the approach towards modelling in the frequency domain is explained. Finally, the metrics that I use for evaluation are clarified.

## 4.1 Methods for replication

In this section I review the methods used in Taylor (2003). First, I introduce the multiplicative seasonal ARIMA model and then I review multiple Holt-Winters methods.

### 4.1.1 Multiplicative seasonal ARIMA

Within this section the electricity demand at time $t$ is referred to as $y_t$. A multiplicative seasonal ARIMA model with just one seasonal pattern is written as:

$$\phi_p(L)\Phi_P(L^s)\nabla^d\nabla_s^D y_t = \theta_q(L)\Theta_Q(L^s)\varepsilon_t, \tag{1}$$

where $L$ is the lag operator, $\nabla$ is the difference operator, $(1-L)$, $s$ is the number of periods within a seasonal cycle, thus $\nabla_s$ is the seasonal difference operator, $d$ and $D$ are the orders of differencing, $\varepsilon_t$ is a white noise error term, and $\phi_p$, $\Phi_p$, $\theta_q$ and $\Theta_Q$ are polynomials of orders $p$, $P$, $q$ and $Q$ respectively, Taylor (2003). The model is multiplicative in the sense that the lag operators $L$ and $L^s$ are multiplied on each side of the equation resulting in a vast collection of lag operators. Within the literature this model is often referred to as an ARIMA$(p,d,q)$ x $(P,D,Q)_s$ model.

Following Box et al. (2015), the multiplicative seasonal ARIMA model can be extended to incorporate multiple seasonalities. In the case of two seasonalities the model is called multiplicative double seasonal ARIMA and is written as:

$$\phi_p(L)\Phi_{P_1}(L^{s_1})\Omega_{P_2}(L^{s_2})\nabla^d\nabla_{s_1}^{D_1}\nabla_{s_2}^{D_2} y_t = \theta_q(L)\Theta_{Q_1}(L^{s_1})\Psi_{Q_2}(L^{s_2})\varepsilon_t, \tag{2}$$

where the new symbols $s_1$ and $s_2$ refer to the number of periods within the two seasonal cycles, and $\Omega_{P_2}$ and $\Psi_{Q_2}$ are polynomial functions of orders $P_2$ and $Q_2$, respectively. This model is referred to as ARIMA$(p,d,q)$ x $(P_1,D_1,Q_1)_{s_1}$ x $(P_2,D_2,Q_2)_{s_2}$ in Taylor (2003). Just like the extension for two seasonalities the model can easily be adjusted to include a third seasonality.

In Taylor (2003) the Bayesian information criterion (BIC) from Schwarz et al. (1978) is compared for an extensive range of ARIMA models and the model with the lowest BIC and satisfactory residuals is an ARIMA$(2,0,0)$ x $(2,0,2)_{48}$ x $(2,0,2)_{336}$ model. Therefore, I use this model in the empirical analysis. This model is referred to as the double seasonal ARIMA model in Taylor (2003), in this paper I use the same notation. I estimate the parameters of the double seasonal ARIMA for electricity demand in England and Wales and I implement the double seasonal ARIMA model in the R programming language. Parameter estimates are obtained with maximum likelihood estimation via the `msarima()` function from the `smooth` package, written by Svetunkov (2021).

### 4.1.2 Holt-Winters

Besides ARIMA, I use Holt-Winters methods as well. In this section electricity demand at time $t$ is referred to as $y_t$. The standard Holt-Winters method is used for series with one seasonal pattern and its multiplicative version is written as:

$$\begin{aligned} \text{Level } S_t &= \lambda(y_t/I_{t-s}) + (1-\lambda)(S_{t-1}+T_{t-1}), \\ \text{Trend } T_t &= \gamma(S_t - S_{t-1}) + (1-\gamma)T_{t-1}, \\ \text{Seasonality } I_t &= \delta(y_t/S_t) + (1-\delta)I_{t-s}, \\ \text{Forecast } \hat{y}_t(h) &= (S_t + hT_t)I_{t-s+h}, \end{aligned} \tag{3}$$

where $\lambda$, $\gamma$ and $\delta$ are smoothing parameters, and $\hat{y}_t(h)$ is the $h$-step ahead forecast. Initial values for the standard Holt-Winters method are estimated according to methodology of Wheelwright et al. (1998):

$$\begin{aligned} \text{Level } S_0 &= \frac{(y_1 + y_2 + ... + y_s)}{s}, \\ \text{Trend } T_0 &= \frac{(y_{s+1} + y_{s+2} + ... + y_{s+s}) - (y_1 + y_2 + ... + y_s)}{s^2}, \\ \text{Seasonality } I_0 = (I_1, I_2, ..., I_s) &= \left(\frac{y_1}{S_0}, \frac{y_2}{S_0}, ..., \frac{y_s}{S_0}\right), \end{aligned} \tag{4}$$

where $s = 48$ for within-day seasonal Holt-Winters and $s = 336$ for within-week seasonal Holt-Winters. The daily cycle is of length 48 and the weekly cycle of length 336 since I consider half-hourly demand. This results in 48 observations per day and 336 per week. After initialisation parameter estimates can be obtained via equation (3). The Holt-Winters method is multiplicative in the sense that the underlying level of the series is multiplied with the seasonal index. This is appropriate if the seasonal variation depends on the level of the series, Taylor (2003). Following Taylor (2003) I implement the multiplicative version of the model.

Since the standard Holt-Winters method is only able to accommodate one seasonal pattern I extend the method to incorporate the second seasonal pattern that is apparent from the electricity demand time series of England and Wales, Figure 1. This seasonal pattern is included in the double seasonal Holt-Winters method and this method is written as:

$$\begin{aligned} \text{Level } S_t &= \lambda\{y_t/(D_{t-s_1}W_{t-s_2})\} + (1-\lambda)(S_{t-1}+T_{t-1}), \\ \text{Trend } T_t &= \gamma(S_t - S_{t-1}) + (1-\gamma)T_{t-1}, \\ \text{Seasonality 1 } D_t &= \delta\{y_t/(S_t W_{t-s_2})\} + (1-\delta)D_{t-s_1}, \\ \text{Seasonality 2 } W_t &= \omega\{y_t/(S_t D_{t-s_1})\} + (1-\omega)W_{t-s_2}, \\ \text{Forecast } \hat{y}_t(h) &= (S_t + hT_t)D_{t-s_1+h}W_{t-s_2+h}, \end{aligned} \tag{5}$$

where $\lambda$, $\gamma$, $\delta$ and $\omega$ are the smoothing parameters, and $\hat{y}_t(h)$ is the h-step ahead forecast. When this method is applied to the data of England and Wales, $s_1 = 48$ and $s_2 = 336$. In that case $D_t$ would represent the within-day and $W_t$ the within-week seasonality. Initial values for Trend and Level are constructed as:

$$\begin{aligned} \text{Trend } T_0 &= \frac{1}{2}\left\{\frac{1}{N}\left(\frac{1}{N}\sum_{t=1}^{N} y_t - \frac{1}{N}\sum_{t=N+1}^{2N} y_t\right) + \frac{1}{N}\sum_{t=1}^{N}\Delta y_t\right\}, \\ \text{Level } S_0 &= \left(\frac{1}{2N}\sum_{t=1}^{2N} y_t\right) - (N+0.5)T_0, \end{aligned} \tag{6}$$

where $N = 336$. Initial values for the within-day seasonal index, $D_t$, are set as the average of the ratios of the actual observation at time t relative to their 48-point centred moving average, taken from the same half-hour for each day of the first week. For example, the first of 48 initial values for the within-day seasonal index is:

$$D_{0,1} = \frac{1}{7} \sum_{k=0}^{6} \frac{y_{1+48k}}{M^{48}_{1+48k}}, \tag{7}$$

where $M^{48}$ is the 48-point centred moving average at time $t = 1 + 48k$, and $k$ runs from 0 to 6 such that the average of the same half-hour for each day is computed. Initial values for the within-week seasonal index, $W_t$, are set as the average of the ratios of the actual observation at time t relative to their 336-point centred moving average, taken from the same half-hour period on the same day of the first two weeks, divided by the initial within-day seasonal index at time t. For example, the first of 336 initial values for the within-week seasonal index is:

$$W_{0,1} = \frac{1}{2}\left(\frac{y_1}{M^{336}_1} + \frac{y_{337}}{M^{336}_{337}}\right)\frac{1}{D_{0,1}}, \tag{8}$$

where $M^{336}$ is the 336-point centred moving average at time $t = 1$ and $t = 337$. Taylor (2003) sets the initial values accordingly.

The Holt-Winters methods are adjusted for autocorrelation by incorporating an AR(1) model, $\varepsilon_t = \eta\varepsilon_{t-1} + \zeta_t$ to the one-step-ahead forecast errors, $\varepsilon_t$. The h-step ahead forecasts from the origin $\tau$ is also to be adjusted by adding a term, namely: $\eta^h\varepsilon_\tau$. This adjustment is originally proposed by Reid (1975) and Gilchrist (1976). The forecast for Holt-Winters and the double seasonal Holt-Winters with an AR(1) model for the one-step-ahead forecast errors is respectively written as:

$$\begin{aligned} \hat{y}_t(h) &= (S_t + hT_t)I_{t-s+h} + \eta^h\varepsilon_t, \\ \hat{y}_t(h) &= (S_t + hT_t)D_{t-s_1+h}W_{t-s_2+h} + \eta^h\varepsilon_t. \end{aligned} \tag{9}$$

Parameters of the Holt-Winters exponential smoothing method are estimated by minimizing the sum of squared one-step-ahead forecast errors, which is to minimize the following:

$$\hat{\lambda} = \arg\min_t \sum_t (y_t - \hat{y}_t)^2. \tag{10}$$

For double seasonal Holt-Winters without AR(1) adjustment and within-week Holt-Winters without AR(1) adjustment the parameter estimates are obtained from the R programming language with functions from the `forecast` package, Hyndman and Khandakar (2008). These functions follow the same estimation procedure as I have described Section 4.1.2. Parameter values of the other Holt-Winters methods are estimated with the `fmincon()` function of Matlab according to Section 4.1.2 as well.

## 4.2 Methods for extension

I this section I discuss the methods I use for the extension. I first discuss the adaption of double seasonal Holt-Winters for electricity demand in France, which is an annual time series with an additional seasonal cycle. Then, I introduce the Fourier models that I use for analysis in the frequency domain. I introduce Fourier models for electricity demand in England and Wales and how to extend those Fourier models for electricity demand in France.

### 4.2.1 Double seasonal Holt-Winters with rolling window

I adjust the double seasonal Holt-Winters method from Section 4.1.2 for analysis of electricity demand in France by implementing a rolling window. Taylor (2010) considers two annual electricity demand time series with similar seasonal cycles as I observe for electricity demand in France. In Taylor (2010) additive and multiplicative formulations of double seasonal Holt-Winters led to similar results. Therefore, I again use the multiplicative seasonality formulation. The rolling window includes observations for two months and moves one week at a time. Initial values for the first window are estimated according to Section 4.1.2 equations (6) till (8). However, from the second window on-wards the initial values for Trend, Level and the seasonal indices are estimated differently. The initial values for Trend and Level are the estimated Trend and Level values at observation 336 from the previous window. The initial values for the within-day seasonal index are the estimated values from the within-day seasonal index from the previous window at observations 289 till 336. Similarly, the initial values for the within-week seasonal index are the estimated values from the within-week seasonal index from the previous window at observations 1 till 336. This procedure is motivated by observing that the window moves with 336 half-hour periods at a time and therefore window $k$ essentially starts at observation 337 from window $k-1$.

Forecasts are constructed similarly to double seasonal Holt-Winters in Section 4.1.2. However, since the window moves one week at a time forecasts are essentially constructed on weekly basis since new forecasts are made for week two till eight of the forecast period of the previous window. Only for the last window the forecast period is different. The last forecast period consists of one week plus one additional day since the year consists of 365 days which are 52 weeks plus one day.

### 4.2.2 Fourier series models

The essence of Fourier series is to express a time series as a function of weighted trigonometric functions, Bloomfield (2004). Fourier proved that a time series can be completely represented by weighted average of trigonometric functions. A model that can completely represent the time series is a Full Fourier series model, based on trigonometric functions from Bloomfield (2004):

$$y_t = c + \sum_{k=1}^{K} \gamma_k^{(1)} \sin\left(\frac{2\pi kt}{N}\right) + \gamma_k^{(2)} \cos\left(\frac{2\pi kt}{N}\right), \tag{11}$$

where $y_t$ is electricity demand at time $t$, $c$ denotes a constant, $\gamma_k^{(1)}$ and $\gamma_k^{(2)}$ denote the coefficients that weight the different harmonics, $N$ is the sample size and by setting $K = N/2$, the model perfectly fits the data. Since there is no stochastic dependence all h-step ahead forecasts will be the same.

I estimate $c$, $\gamma_k^{(1)}$ and $\gamma_k^{(2)}$ in equation (11) with the least absolute shrinkage and selection operator (LASSO) from Tibshirani (1996):

$$\sum_{i=1}^{n}\Big(y_i - \sum_j x_{ij}\beta_j\Big)^2 + \lambda \sum |\beta_j|, \tag{12}$$

where $y_i$ is the response variable for predictors $x_{ij}$ and $\lambda$ is the tuning parameter for the second term, which is the penalty term for the number of included variables. Hence, when $\lambda$ is small, LASSO is essentially equal to OLS. For the Fourier models this would entail that $y_i$ is electricity demand and $x_{ij}$ are the different cycles with their corresponding complexities with estimated coefficients $\gamma_k^{(1)}$ and $\gamma_k^{(2)}$, which are presented in equation (12) by $\beta_j$. I employ the LASSO instead of ridge regression and OLS since the Fourier models deal with a large amount of parameters that have to be estimated. OLS estimates tend to have low bias but high variance in this case. Ridge regression is able to reduce the variance of OLS estimates but does not shrink parameter estimates to zero. Because LASSO does shrink parameter estimates to zero, the parameter set is reduced and interpretation of the parameter estimates is easier. By using LASSO bias will be higher but variance will be substantially lower which results in more accurate forecasts.

One difficulty with LASSO in this setting is that the choice of the tuning parameter, $\lambda$, is often based on cross-validation. Cross-validation for time-series tends to prove difficult since time series characteristics are often lost. Therefore, I do selection of the tuning parameter, $\lambda$, with LASSO based on the BIC. This criterion is tailored towards forecasting and is well suited for time series analysis.

I estimate the parameters with the `HDeconometrics` package, written by Vasconcelos (2021), which is able to do selection of the tuning parameter, $\lambda$, of LASSO estimation via the `glmnet` package, written by Simon et al. (2011), based on the BIC instead of cross-validation. Since LASSO selects the most important parameters by setting others equal to zero the model from equation (11) does not perfectly fit the data and residuals are left, the new model is written as:

$$y_t = c + \sum_{k=1}^{K} \gamma_k^{(1)} \sin\left(\frac{2\pi kt}{N}\right) + \gamma_k^{(2)} \cos\left(\frac{2\pi kt}{N}\right) + \varepsilon_t. \tag{13}$$

The residuals of the model in equation (13) can be utilised to improve forecast accuracy of electricity demand by fitting an AR($p$) model to the residuals. The Full Fourier model is fitted with an AR(1) and AR(48) model to the error terms: $\varepsilon_t = \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-48} + u_t$.

Fourier models can also be estimated by manually constructing cycles and their corresponding complexities, I call such a Fourier model a Partial Fourier model and this model is written as:

$$y_t = c + \sum_{i=1}^{P} \sum_{k=1}^{K_i} \gamma_{i,k}^{(1)} \sin\left(\frac{2\pi kt}{p_i}\right) + \gamma_{i,k}^{(2)} \cos\left(\frac{2\pi kt}{p_i}\right) + \varepsilon_t, \tag{14}$$

where $c$ denotes a constant, $P$ the number of cycles that are manually constructed, $K_i$ the complexity for cycle $p_i$, $\gamma_{i,k}^{(1)}$ and $\gamma_{i,k}^{(2)}$ denote the coefficients that weigh harmonic $i$ with complexity $k$ and $\varepsilon_t$ is the error term at time $t$. I model a Partial Fourier model with $p_1 = 48$ and $p_2 = 336$ corresponding to a within-day cycle and within-week cycle for electricity demand in England and Wales. I call this model a Double Fourier model. I estimate two different Double Fourier models regarding their complexity $K_i$. First, I estimate a Double Fourier model with $K_1 = 5$ and $K_2 = 10$. I choose $K_1 = 5$ because from the periodogram in Figure 12 in Appendix B I observe that the daily cycle has multiples up to five. Motivation for $K_2 = 10$ is similar since from Figure 13 in Appendix B I observe that the weekly cycle has multiples up to ten. Secondly, I estimate a Double Fourier model with $K_1 = 100$ and $K_2 = 100$, motivation for this is that the daily and weekly cycle have higher multiples but those are not visible from the periodogram. This means they explain a small portion of the variance. Including many of them may however in sum explain a significant portion of the variance. I fit the Double Fourier model with two AR models for the error terms, an AR(1), AR(48) model: $\varepsilon_t = \phi_1\varepsilon_{t-1} + \phi_2\varepsilon_{t-48} + u_t$, and an AR(1), AR(48), AR(49), AR(50) model: $\varepsilon_t = \phi_1\varepsilon_{t-1} + \phi_2\varepsilon_{t-48} + \phi_3\varepsilon_{t-49} + \phi_4\varepsilon_{t-50} + u_t$

For electricity demand in France I model a Triple Fourier model where $p_1 = 48$, $p_2 = 336$ and $p_3 = 17520$ corresponding to a within-day, within-week and within-year seasonal cycle where all complexities $K_i$ are equal to 100. Lastly, since the data of electricity demand in France includes a leap year, I model a fourth cycle, $p_4 = 17568$, to take the extra day into account. I model the Triple Fourier model with two different AR models. One fits an AR(1) and AR(48) model to the error terms and the second fits an AR(1), AR(2), AR(48), AR(49) and AR(50) model to the error terms.

I estimate the coefficients of all Partial Fourier models with LASSO from Tibshirani (1996). LASSO has a penalty term for the number of parameters and will therefore set coefficients equal to zero for complexities of cycles that do not contribute towards forecasting accuracy. Consequently, I expect the parameter estimate for $\gamma_{2,7}^{(1)}$ and $\gamma_{2,7}^{(2)}$ for the double Fourier model with $K_1 = 5$ and $K_2 = 10$ to be zero since this is the same cycle as the daily cycle with $k_1 = 1$.

### 4.3 Metrics for evaluation

All models are evaluated on the mean absolute percentage error (MAPE). Following Taylor (2003) this is the most widely used evaluation criteria in electricity demand forecasting, and is written as:

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right|. \tag{15}$$

Motivated by the recommendation from Hippert et al. (2001) three other evaluation criteria are used in the Taylor (2003) namely, mean absolute error (MAE), root-mean-square error (RMSE), and root-mean-square percentage error (RMSPE):

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t|, \tag{16}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}, \tag{17}$$

$$\text{RMSPE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(\frac{\hat{y}_t}{y_t} - 1\right)^2}. \tag{18}$$

Since Taylor (2003) and Taylor (2010) mention that performance for the Holt-Winters methods is very similar for all four of the criteria the MAE, RMSE and RMSPE are only mentioned when performance of the Fourier Models is discussed.

## 5 Results

The results are split in two sections. In the first section the results for the methods discussed in Taylor (2003), the replication part, are shown and in the second section the results for double seasonal Holt-Winters with rolling window and the Fourier models, the extension part, are shown.

## 5.1 Replication

In this section I first discuss the parameter estimates and MAPE performance of the models without AR adjustment and afterwards I discuss the performance of the methods with AR adjustment to the error terms.

For Double Seasonal ARIMA I retrieved the following parameters:

$$
\begin{aligned}
&(1-0.57L-0.38L^2)(1+0.28L^{48}+0.32L^{96})(1-0.52L^{336}-0.47L^{672})y_t \\
&\qquad = (1+0.19L^{48}+0.33L^{96})(1-0.01L^{336}+0.24L^{672})\varepsilon_t + 0.56.
\end{aligned}
$$

The derived parameter values for Holt-Winters are presented in Table 2. From Table 2 it is observed that Holt-Winters with within-week seasonality and double seasonal Holt-Winters have high values for $\lambda$ accompanied with zero or very close to zero values for the trend, $\gamma$. This seems logical as the variation in the data is dominated by seasonality, Figure 1. This is also reflected by the parameter estimates of the seasonal indices. For Holt-Winters with within-day seasonality the value for the trend is quite high as it is unable to pick up the within-week seasonality and includes this seasonal variance in the trend.

Table 2: Holt-Winters parameters calculated from the 8-week estimation sample of England and Wales

| | Level $\lambda$ | Trend $\gamma$ | Within-Day seasonality $\delta$ | Within-Week seasonality $\omega$ |
|---|---|---|---|---|
| Holt-Winters for Within-Day Seasonality | 0.98 | 0.80 | 1.00 | - |
| Holt-Winters for Within-Week Seasonality | 0.82 | 0.00 | - | 1.00 |
| Double Seasonal Holt-Winters | 0.77 | 0.01 | 0.99 | 0.99 |

No adjustment for autocorrelation in the residuals

The fact that the time series is dominated by within-week seasonality is clearly visible from the performance of the methods in Figure 3. Performance of Holt-Winters with within-day seasonality is poor, hence I decided to leave more than three lead times ahead forecasts out of the plot. Holt-Winters with within-week seasonality performs quite well. Clearly, within-week seasonality dominates the time series and is essential in models to be able to accurately forecast short-term electricity demand. Double seasonal Holt-Winters outperforms within-week Holt-Winters for 18 of the 48 lead times, including the first twelve. This indicates that for the short-run there is benefit in using a method that is able to pick up two seasonalities. Performance of both Holt-Winters for within-week seasonality and double seasonal Holt-Winters tends to improve after 12 hours. Consequently, a forecast for 12 hours ahead is more accurate when it is made from an origin 12 hours previous to the current period. The last method, double seasonal ARIMA, outperforms all Holt-Winters methods for 39 of the 48 lead times. The (double) seasonal Holt-Winters methods are not able to outperform the double seasonal ARIMA model.

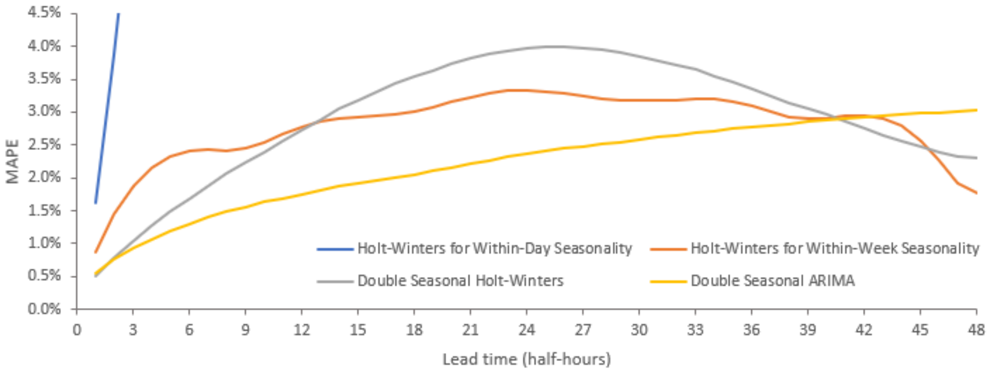


Figure 3: Comparison of MAPE for 4 week post sample period. No adjustment for autocorrelation in the residuals of Holt-Winters.

Taylor (2003) observes that the 1-step-ahead forecast errors in the estimation sample show sizeable first-order autocorrelation for all three Holt-Winters methods. Hence, I include an AR(1) model fitted to the 1-step-ahead forecast errors. Results for the three Holt-Winters methods with adjustment for autocorrelation are shown in Table 3. Comparing parameter estimates of Table 3 with parameter estimates of Table 2 shows that the trend estimate for Holt-Winters for within-day seasonality is now zero and that the model tries to capture the within-week seasonality with the AR-term, $\eta$. For Holt-Winters for within-week seasonality and double seasonal Holt-Winters the parameter estimate of the level, $\lambda$, and the seasonal indices, $\delta$ and $\gamma$, drop substantially and the AR-term shows a very high parameter estimate for both models, indicating that the modelled first-order autocorrelation is able to explain a high portion of the variance in the time series.

Table 3: Holt-Winters parameters calculated from the 8-week estimation sample of England and Wales

| | Level $\lambda$ | Trend $\gamma$ | Within-Day seasonality $\delta$ | Within-Week seasonality $\omega$ | AR-term $\eta$ |
|---|---|---|---|---|---|
| Holt-Winters for Within-Day Seasonality | 0.80 | 0.00 | 1.00 | - | 0.62 |
| Holt-Winters for Within-Week Seasonality | 0.01 | 0.00 | - | 0.46 | 0.92 |
| Double Seasonal Holt-Winters | 0.03 | 0.00 | 0.18 | 0.27 | 0.93 |

Adjustment for autocorrelation with an AR(1) model for the residuals.

The importance of modelling the first-order autocorrelation in the Holt-Winters methods is reflected in the performance of the models in Figure 4, where all Holt-Winters methods have improved forecast accuracy for all lead times compared with Figure 3. Relative performance of the methods is still intact except that double seasonal ARIMA is now outperformed by double seasonal Holt-Winters and Holt-Winters for within-week seasonality for all 48 lead times. Double seasonal Holt-Winters is the best performing model for short-term electricity demand forecasting in England and Wales for every single lead time.

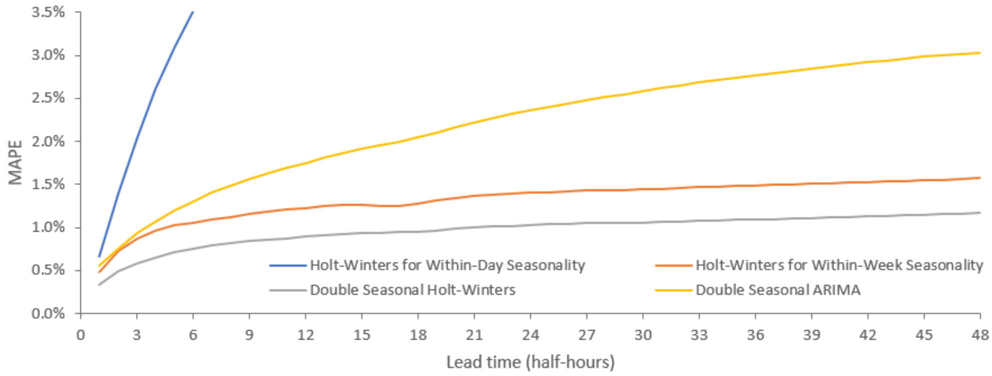


Figure 4: Comparison of MAPE for 4 week post sample period. Holt-Winters is adjusted for autocorrelation with an AR(1) model for the residuals.

## 5.2 Extension

In this section I first present results for the Fourier models for electricity demand in England and Wales and afterwards I discuss results for electricity demand in France. In this section all the Fourier models are compared with standard double seasonal Holt-Winters for England and Wales and with double seasonal Holt-Winters with a rolling window for France. These models are chosen as benchmark since double seasonal Holt-Winters is the best performing model of Section 5.1.

### 5.2.1 England and Wales

The benefit of incorporating LASSO instead of OLS is clear from the parameter estimates of the Double Fourier model. The parameter estimate for $\gamma_{2,7}^{(1)}$ and $\gamma_{2,7}^{(2)}$ are equal to zero, indicating that indeed the weekly cycle with complexity $k_2 = 7$ is equivalent to the daily cycle with $k_1 = 1$ and therefore does not contribute towards forecasting accuracy. OLS would not have been able to accurately detect this and shrink one coefficient much more than the other. From Figure 5 it is clear that a Fourier model with manual modelling of the cycles is more accurate than a Full Fourier model. Furthermore, fitting an AR(1) and AR(48) model to the error term greatly improves forecasting accuracy.

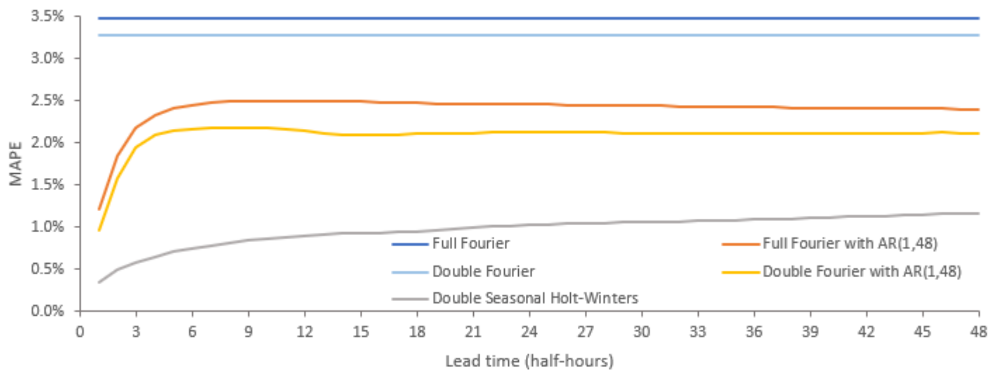


Figure 5: Comparison of MAPE for 4 week post sample period. Holt-Winters is adjusted with an AR(1) model for the error terms and the Fourier Models are adjusted with an AR(1) and an AR(48) term for the error terms. For the Double Fourier model $K_1 = 5$ and $K_2 = 10$.

When I increase the complexity of both seasonal cycles to $K_1 = 100$ and $K_2 = 100$ for the Double Fourier model the accuracy improves as is clear from Figure 6. Analysis of the (partial) autocorrelation plots, Figure 14 and Figure 15 in Appendix C, shows that, additional to the AR(1) and AR(48) model for the error terms, fitting an AR(49) and AR(50) model to the error terms as well may improve accuracy. Clearly, implementing an AR(49) and AR(50) model for the error terms improves the accuracy of the Double Fourier model further, Figure 6. The Double Fourier model is however still not able to outperform the double seasonal Holt-Winters method for any lead time.

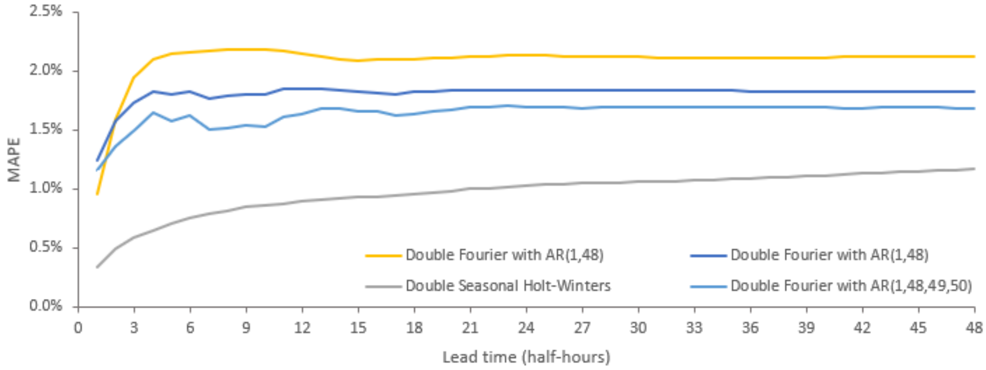


Figure 6: Comparison of MAPE for 4 week post sample period. Holt-Winters is adjusted with an AR(1) model for the error terms and the Fourier Models are adjusted with an AR model for the error terms. Complexity of the yellow Double Fourier model is 5 for the daily cycle and 10 for the weekly cycle. The blue Double Fourier have complexity 100 for both cycles previous to lasso estimation.

### 5.2.2 France

Estimation of a Full Fourier model for half-hourly electricity demand in France requires the initialization of 87648 parameters. This results in a matrix of 87648 by 87648 for the LASSO estimation. This makes the computation in the programming language R infeasible since the memory size of the aforementioned matrix is too large. Therefore, the Full Fourier model is estimated on a bi-hourly data set. From Figure 7 it is clear that the Triple Fourier model greatly outperforms the Full Fourier model for every single lead time. Just as with electricity demand in England and Wales the Fourier model with manually constructed cycles performs best and performance is greatly improved by fitting an AR(1) and AR(12) model to the error terms.

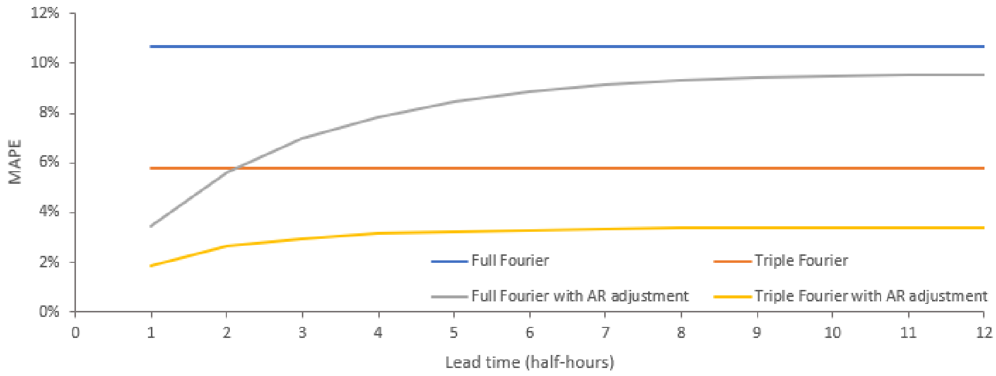


Figure 7: Comparison of MAPE for 1 year post sample period. The error terms of the Fourier Models are adjusted with an AR(1) and an AR(12) model since the data set is bi-hourly.

I estimate the Triple Fourier model from Figure 7 for the half-hourly data as well and results are visible in Figure 8. From Figure 8 I conclude that adding a leap year to the Fourier model specification does not improve forecasting accuracy since I am unable to distinguish the difference between the MAPE values for the Triple and Quadruple Fourier model. Also, applying a moving window of five years to the error terms to allow for different parameter estimates through time for the AR model does not increase forecasting accuracy. However, decreasing the window to two months shows an increase in forecasting accuracy for lead times further than 39 half-hours ahead and a slight decrease for lead times less than 39 half-hours ahead. These Fourier models are not able to outperform the double seasonal Holt-Winters method.

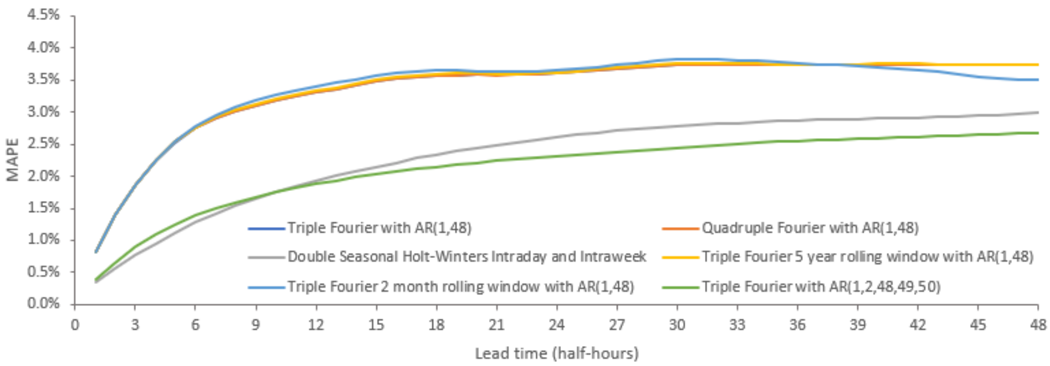


Figure 8: Comparison of MAPE for 1 year post sample period. The error terms of the Fourier Models are adjusted with an AR($p$) model and the error terms of the double seasonal Holt-Winters method are adjusted with an AR(1) model.

A closer look at the (partial) autocorrelation plots reveals that forecasting accuracy may be improved with additional terms in the AR model. Motivation for adding an AR(2), AR(49) and AR(50) model to the error terms is clear from Figure 16 and Figure 17 in Appendix C where I observe that the autocorrelations are significant and show a slowly declining but seasonal pattern and the partial autocorrelations are clearly significant for the first, second and 48th till 50th lag. By including these terms in the AR model for the error terms the performance of the Triple Fourier model improves up to the point that it is more accurate than the double seasonal Holt-Winters method for predicting electricity demand for lead times greater than 10 half-hours ahead. For lead times below 10 half-hours ahead the double seasonal Holt-Winters method outperforms the Triple Fourier model.

When considering the MAE, RMSE and RMSPE the Triple Fourier model outperforms the double seasonal Holt-Winters method for similar horizons. From Figure 9 it is clear that the Triple Fourier model outperforms the double seasonal Holt-Winters method for 40 of the 48 lead times considering MAE and for 38 of the 48 lead times considering the RMSE. Finally, for the RMSPE, Figure 18 in Appendix D, the Triple Fourier is more accurate for 36 of the 48 lead times.

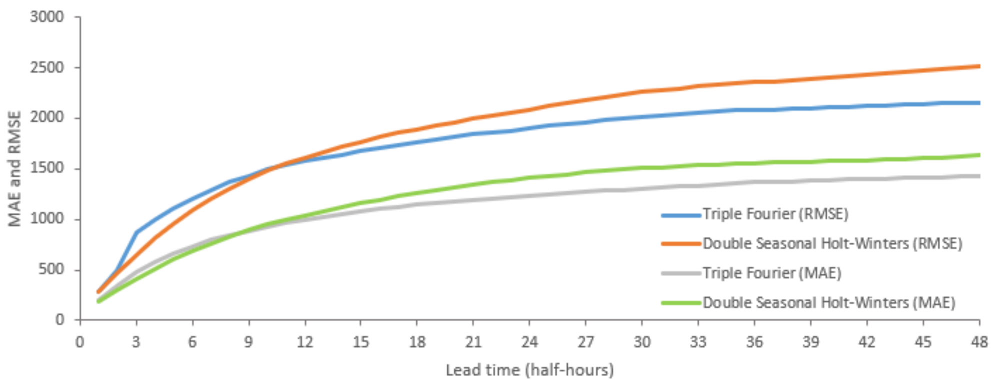


Figure 9: Comparison of MAE and RMSE for 1 year post sample period. The Triple Fourier model is adjusted with an AR(1), AR(2), AR(48), AR(49) and AR(50) model for the error terms and the double seasonal Holt-Winters method is adjusted with an AR(1) model for the error terms.

# 6 Discussion and Conclusion

Models based on Fourier series that incorporate seasonal patterns can outperform a double seasonal Holt-Winters method depending on the number of seasonal patterns included in the Fourier model and the number of strong seasonal patterns in the time series.

I introduce three types of Fourier models, the Double Fourier model with two seasonal cycles, the Triple Fourier model with three seasonal cycles and the Quadruple Fourier model with four seasonal cycles. For electricity demand in England and Wales, where I consider a time series of three months of half-hourly electricity demand during summer, the Double Fourier model is unable to outperform the double seasonal Holt-Winters method. The second time series that I consider regards electricity demand in France. This time series contains half-hourly electricity demand observations for 5 years. For this time series the Triple Fourier model with five AR terms fitted to the error terms is able to outperform the double seasonal Holt-Winters method. The time series in France includes national public holidays, substantially more observations and more seasonal variation since electricity demand changes substantially throughout the year. Hence, based on this research Fourier models seem to perform relatively better for long, complex time series than for short, simple time series. I advise to use double seasonal Holt-Winters for short time series with two seasonal patterns and to incorporate a Triple Fourier model for time series that include an annual seasonality that the double seasonal Holt-Winters method is unable to incorporate.

Furthermore, I find that including a leap year in the Fourier model, i.e. the Quadruple Fourier model, for electricity demand in France did not improve the accuracy of the predictions. Next, I find that a two month rolling window improved forecasting accuracy of the Fourier model with three AR terms for lead times greater than 39 half-hours ahead. Hence, forecasting accuracy of the Triple Fourier model that currently outperforms double seasonal Holt-Winters may be further improved by incorporating a rolling window for the AR model that is fitted to the error terms.

Forecasting accuracy of the double seasonal Holt-Winters method may be further improved as well since the double seasonal Holt-Winters method with rolling window shows unstable parameter estimates. These are shown in Figure 19 in Appendix D. The parameter estimates are quite volatile over time which may indicate that finding the optimum of the objective function is difficult using my programming setup. I would expect the parameters to change over time but in Figure 19 they sometimes change with over 100% on weekly basis which seems too much. One consequence of the unstable parameter estimates could be that the forecasting accuracy of the double seasonal Holt-Winters method for electricity demand can be improved by improving the consistency of the parameter estimates.

Another venue to improve forecasting performance of the double seasonal Holt-Winters method is to include a third seasonality for electricity demand in France. Taylor (2010) introduces a triple seasonal Holt-Winters method which outperforms the double seasonal Holt-Winters method.

In this research I consider electricity demand in France. France has public holidays based on only the Gregorian calendar. Countries such as Turkey have public holidays for both the Gregorian and Hijri calendar and I therefore expect the Triple Fourier model to show relatively stronger performance than the Holt-Winters methods for such a time series. Empirical research regarding this observation was however outside the scope of this paper but in my opinion is interesting for future research.

Beside holidays based on multiple calendars, moving holidays can results in complex seasonal patterns as well. In my methodology I have not implemented a procedure for indicating moving holidays such as Easter. McElroy et al. (2018) discuss the effects and implementation of moving holidays. Implementation may further improve forecast accuracy.

Finally, for the shortest lead times the double seasonal Holt-Winters method performs best whilst for longer lead times the Triple Fourier model performs best. Smith (1989) discusses the implementation of combining forecasts. Taylor (2010) shows a simple average of two univariate procedures already improves MAPE values, Taylor (2008) shows combining a weather-based Holt-Winters with double seasonal Holt-Winters improves accuracy and Yang et al. (2016) show that combining methods that deal with linear and non-linear data is superior to their individual performance and that the combined method has extensive applicability. Hence, much research indicates that combining forecasts improves accuracy and therefore this may very well be the case with combining Fourier and Holt-Winters as well.

## References


Al-Musaylh, M. S., Deo, R. C., Adamowski, J. F., and Li, Y. (2018). Short-term electricity demand forecasting with MARS, SVR and ARIMA models using aggregated demand data in Queensland, Australia. *Advanced Engineering Informatics*, 35:1–16.

Bloomfield, P. (2004). *Fourier analysis of time series: An introduction.* John Wiley & Sons, New York.

Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time series analysis: Forecasting and control.* John Wiley & Sons, San Francisco.

Bush, E. R., Abernethy, K. A., Jeffery, K., Tutin, C., White, L., Dimoto, E., Dikangadissi, J.-T., Jump, A. S., and Bunnefeld, N. (2017). Fourier analysis to detect phenological cycles using long-term tropical field data and simulations. *Methods in Ecology and Evolution*, 8(5):530–540.

Chatfield, C. and Yar, M. (1988). Holt-Winters forecasting: Some practical issues. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 37(2):129–140.

Cooley, J. W. and Tukey, J. W. (1965). An algorithm for the machine calculation of complex fourier series. *Mathematics of Computation*, 19(90):297–301.

De Livera, A. M., Hyndman, R. J., and Snyder, R. D. (2011). Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association*, 106(496):1513–1527.

De Oliveira, E. M. and Oliveira, F. L. C. (2018). Forecasting mid-long term electric energy consumption through bagging ARIMA and exponential smoothing methods. *Energy*, 144:776–788.

Fan, S. and Chen, L. (2006). Short-term load forecasting based on an adaptive hybrid method. *IEEE Transactions on Power Systems*, 21(1):392–401.

Fourier, J. B. J. (1822). *Théorie analytique de la chaleur.* Firmin Didot.

Fuller, W. A. (2009). *Introduction to statistical time series*, volume 428. John Wiley & Sons, New York.

Gardner Jr, E. S. (1985). Exponential smoothing: The state of the art. *Journal of Forecasting*, 4(1):1–28.

Ghalehkhondabi, I., Ardjmand, E., Weckman, G. R., and Young, W. A. (2017). An overview of energy demand forecasting methods published in 2005–2015. *Energy Systems*, 8(2):411–447.

Gilchrist, W. (1976). Statistical forecasting. *Biometrical Journal*, 20(3).

Hippert, H. S., Pedreira, C. E., and Souza, R. C. (2001). Neural networks for short-term load forecasting: A review and evaluation. *IEEE Transactions on power systems*, 16(1):44–55.

Hyndman, R., Koehler, A. B., Ord, J. K., and Snyder, R. D. (2008). *Forecasting with exponential smoothing: The state space approach*. Springer, Berlin.

Hyndman, R. J. and Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software*, 26(3):1–22.

Jiang, W., Wu, X., Gong, Y., Yu, W., and Zhong, X. (2020). Holt-Winters smoothing enhanced by fruit fly optimization algorithm to forecast monthly electricity consumption. *Energy*, 193:116779.

Karabiber, O. A. and Xydis, G. (2019). Electricity price forecasting in the Danish day-ahead market using the TBATS, ANN and ARIMA methods. *Energies*, 12(5):928.

Lidiema, C. (2017). Modelling and forecasting inflation rate in Kenya using SARIMA and Holt-Winters triple exponential smoothing. *American Journal of Theoretical and Applied Statistics*, 6(3):161–169.

Liu, M., Taylor, J. W., and Choo, W.-C. (2020). Further empirical evidence on the forecasting of volatility with smooth transition exponential smoothing. *Economic Modelling*, 93:651–659.

McElroy, T. S., Monsell, B. C., and Hutchinson, R. J. (2018). Modeling of holiday effects and seasonality in daily time series. *Statistics*, 1.

Reid, D. J. (1975). A review of short-term projection techniques. *Practical aspects of forecasting*, 8:25.

RTE (2020). Electricity consumption in France, retrieved from https://www.rte-france.com/en/eco2mix/electricity-consumption-france.

Schwarz, G. et al. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6(2):461–464.

Simon, N., Friedman, J., Hastie, T., and Tibshirani, R. (2011). Regularization paths for Cox's proportional hazards model via coordinate descent. *Journal of Statistical Software*, 39(5):1–13.

Smith, D. G. (1989). Combination of forecasts in electricity demand prediction. *Journal of Forecasting*, 8(3):349–356.

Smyl, S. (2020). A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting. *International Journal of Forecasting*, 36(1):75–85.

Svetunkov, I. (2021). *Smooth: Forecasting Using State Space Models*. R package version 3.1.1.

Taylor, J. W. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of the Operational Research Society*, 54(8):799–805.

Taylor, J. W. (2008). An evaluation of methods for very short-term load forecasting using minute-by-minute british data. *International Journal of Forecasting*, 24(4):645–658.

Taylor, J. W. (2010). Triple seasonal methods for short-term electricity demand forecasting. *European Journal of Operational Research*, 204(1):139–152.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288.

Trull, O., García-Díaz, J. C., and Troncoso, A. (2020). Initialization methods for multiple seasonal Holt-Winters forecasting models. *Mathematics*, 8(2):268.

Vasconcelos, G. F. R. (2021). *HDeconometrics: Implementation of several econometric models in high-dimension.* R package version 0.1.0.

Wheelwright, S., Makridakis, S., and Hyndman, R. J. (1998). *Forecasting: Methods and applications.* John Wiley & Sons, New York.

Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science*, 6(3):324–342.

Yang, H.-F., Dillon, T. S., Chang, E., and Chen, Y.-P. P. (2018). Optimized configuration of exponential smoothing and extreme learning machine for traffic flow forecasting. *IEEE Transactions on Industrial informatics*, 15(1):23–34.

Yang, Y., Chen, Y., Wang, Y., Li, C., and Li, L. (2016). Modelling a combined method based on ANFIS and neural network improved by DE algorithm: A case study for short-term electricity demand forecasting. *Applied Soft Computing*, 49:663–675.

Yao, H., Wu, F., Ke, J., Tang, X., Jia, Y., Lu, S., Gong, P., Ye, J., and Li, Z. (2018). Deep multi-view spatial-temporal network for taxi demand prediction. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1).

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50:159–175.

# A Time series data

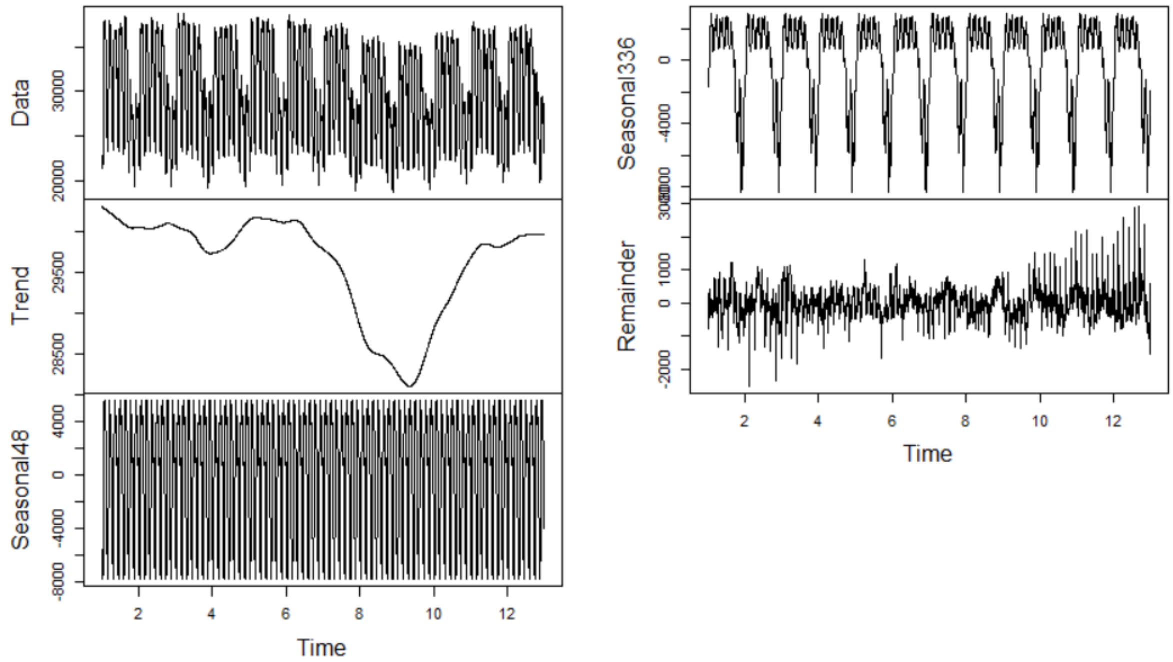


Figure 10: Seasonal decomposition of electricity demand in England and Wales based on Loess.

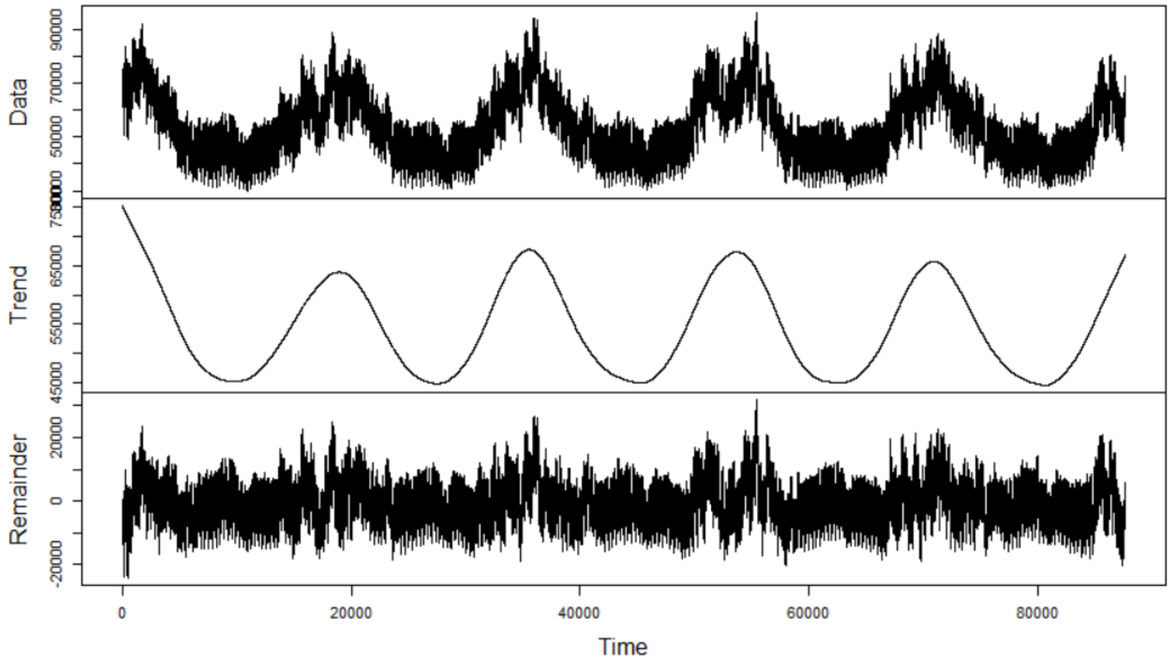


Figure 11: Seasonal decomposition of electricity demand in France based on Loess.

# B Periodograms

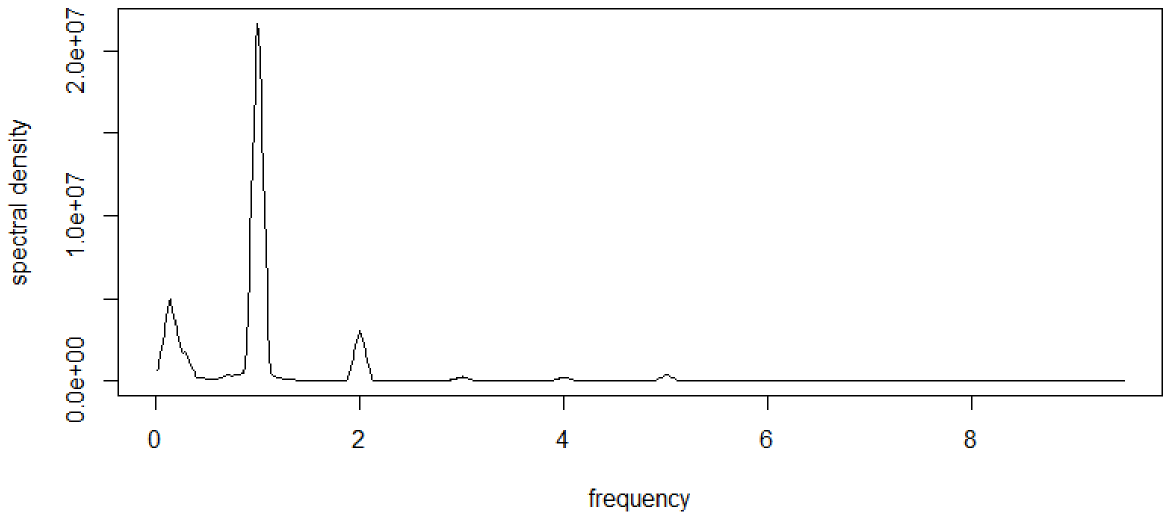


Figure 12: Spectral density for electricity demand of England and Wales where frequency of 1 is equal to the daily cycle of length 48.

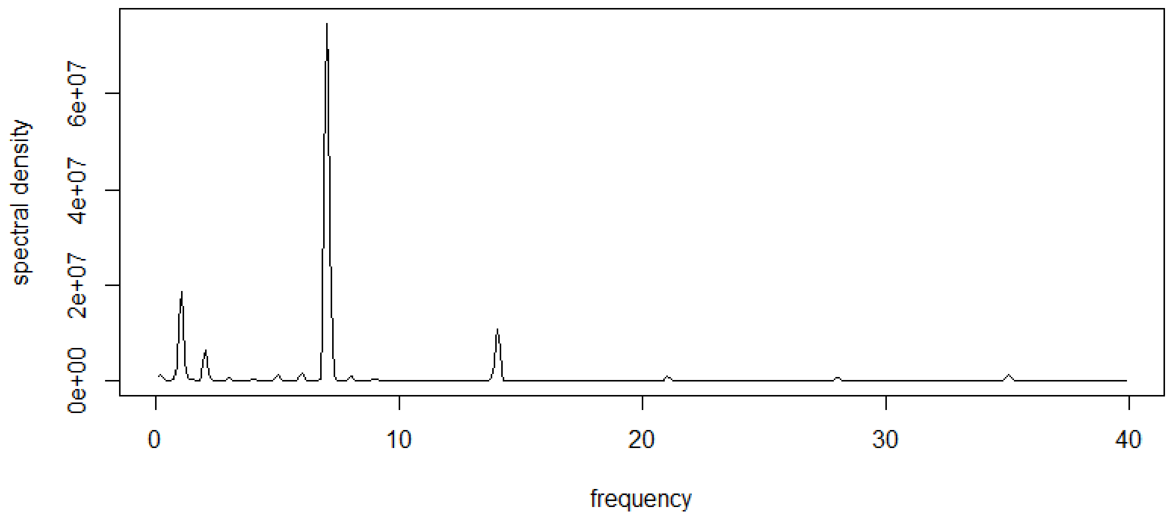


Figure 13: Spectral density for electricity demand of England and Wales where frequency of 1 is equal to the weekly cycle of length 336.

# C (Partial) autocorrelation plots

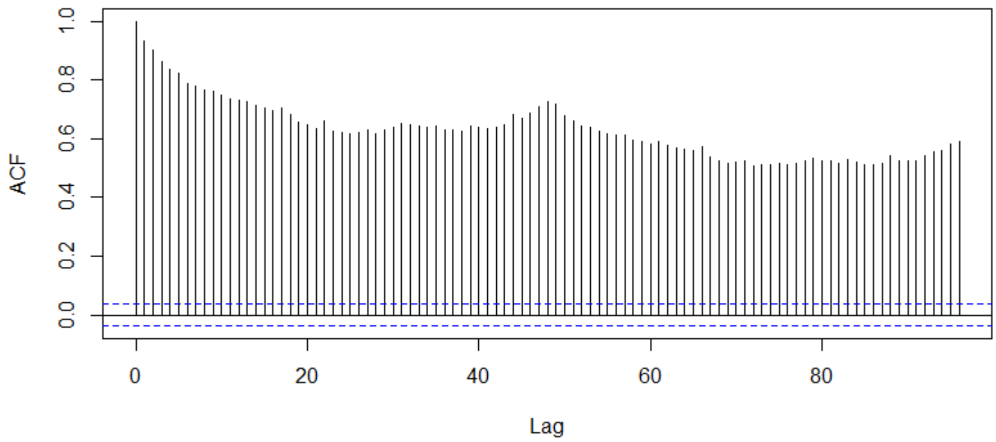


Figure 14: Autocorrelation plot for the 2 month estimation period of the Double Fourier model for electricity demand in England and Wales.

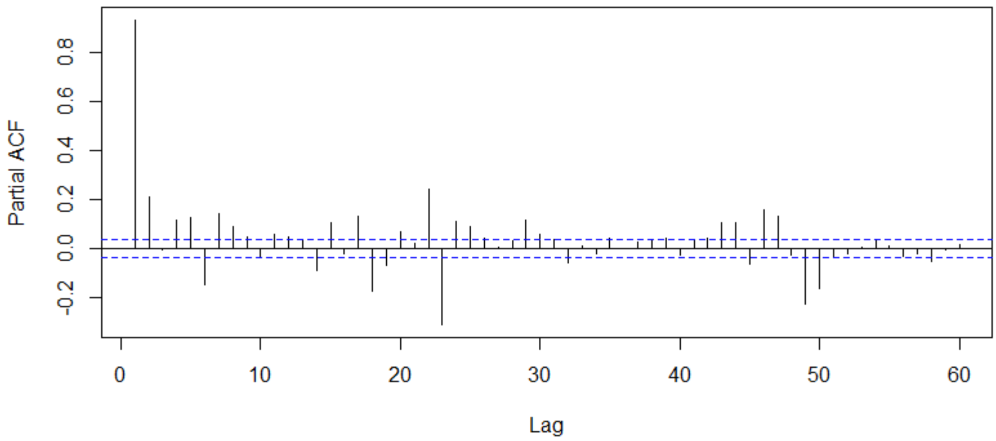


Figure 15: Partial autocorrelation plot for the 2 month estimation period of the Double Fourier model for electricity demand in England and Wales.

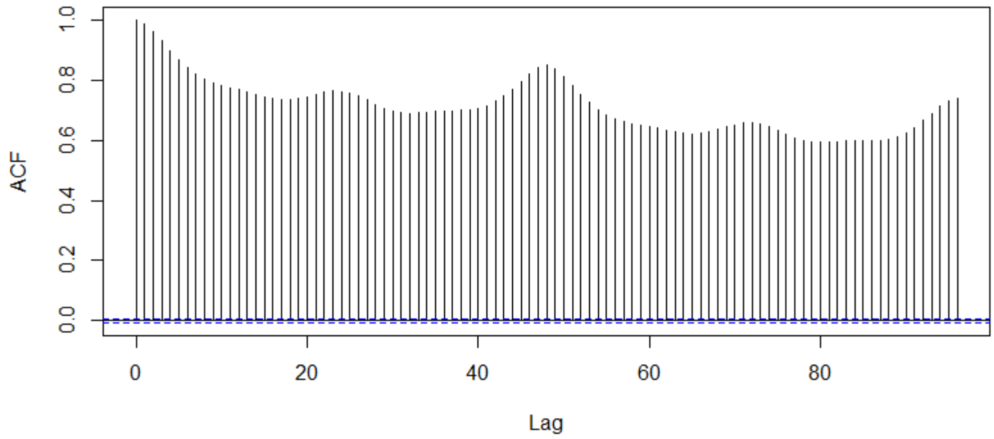


Figure 16: Autocorrelation plot for the 4 year estimation period of the Double Fourier model for electricity demand in France.

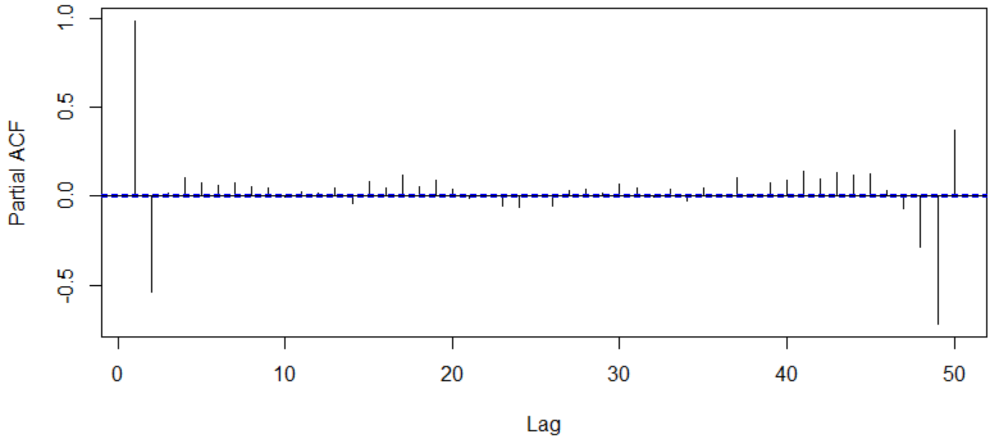


Figure 17: Partial autocorrelation plot for the 4 year estimation period of the Double Fourier model for electricity demand in France.

# D Additional evaluation criteria and parameter estimates

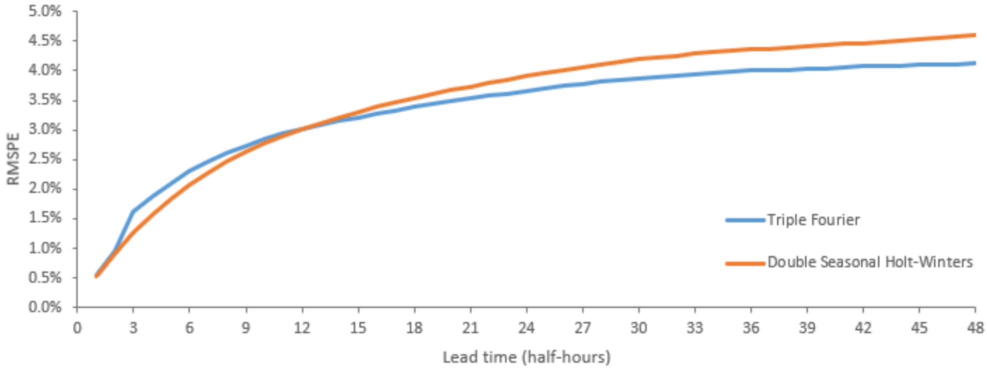


Figure 18: Comparison of RMSPE for 1 year post sample period. The Triple Fourier model is adjusted with an AR(1), AR(2), AR(48), AR(49) and AR(50) model for the error terms and the Double Seasonal Holt-Winters method is adjusted with an AR(1) model for the error terms.

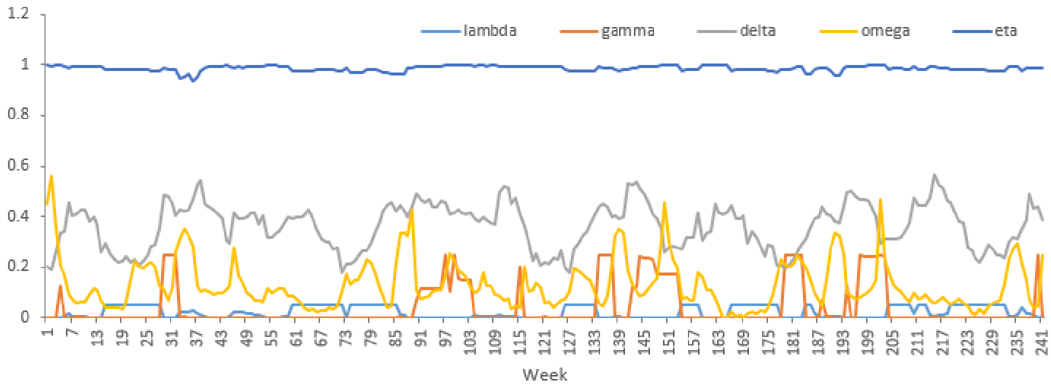


Figure 19: Parameter estimates for double seasonal Holt-Winters with a moving window of two months taking 1 week steps at a time for electricity demand in France.