Stockholms universitet

Forecast evaluation of dynamic regression and sarima models applied to Electricity Spot Prices - Time Series Analysis

Christopher Hellesnes

Kandidatuppsats i matematisk statistik
Bachelor Thesis in Mathematical Statistics

Kandidatuppsats 2017:10
Matematisk statistik
Juni 2017

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Mathematical Statistics
Stockholm University
Bachelor Thesis **2017:10**
http://www.math.su.se

Forecast evaluation of dynamic regression and sarima models applied to Electricity Spot Prices - Time Series Analysis

Christopher Hellesnes*

June 2017

Abstract

Deregulation changed the dynamics of the Swedish power market. The deregulated market opened up for the power exchange, Nord Pool. In this thesis we will examine if a dynamic regression model including daily temperature data provides a more powerful model in terms of forecasting than a univariate time series model solely based on the electricity spot price. In this thesis a SARIMA model is deemed suitable due to the weekly seasonality found in the electricity spot price data. The analysis indicates that a dynamic regression model with the inclusion of an exogenous variable can provide a more powerful model, in terms of forecasting.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: chris.hellesnes@gmail.com. Supervisor: Filip Lindskog and Mathias Lindholm.

Acknowledgments

I would like to thank my pet hamster.

Contents

1 Introduction

In today's western society we are all dependent on electricity, without a properly working power market parts our infrastructure would collapse. The power markets impact on our infrastructure is believed to be one of the reasons why it was kept a government owned monopoly for so long. However a new political era with new ideas arose during the 90's and the Swedish power market was deregulated. A deregulated market with different market participants with different motives creates a dynamic market where estimates about future prices becomes increasingly important. This has lead to a upswing in electricity price forecasting, resulting in different modeling techniques with varying degree of complexity. Many models are stochastic time series models in the core with additionally modifications such as jump-diffusion [10] or regime switching and stochastic volatility models [11]. In this thesis we will limit ourselves to the use of more traditional univariate time series models and then allow for the inclusion of exogenous variables. The purpose of this thesis is to answer the question "Does a model with a exogenous variable predict the future spot price better than a model without?", even though many potential suitable exogenous variables exists only historical temperature data will be used. Previous work suggests that temperature has a significant impact on the electricity price and SMHI provides reliable temperature data from 1756 [9]. In warm areas the use of electricity should reach it's peak during summer time when extensive use of air condition is utilized, in the Nordics the opposite relation exist due to extensive use of heating during winter time. As this relationship is assumed to exist in our data a model with temperature data included should have stronger forecasting abilities.

2 Background

2.1 Deregulation of the Swedish Power Market

During the 1990's a wave of deregulation swept through the Nordic countries, and the power markets was not unaffected. The electricity market was deregulated in 1991 in Norway, 1995 in Finland, and the Swedish market was deregulated in 1996 [3]. Deregulation changed the dynamics of the Swedish power market as production and sale became an open, competitive market. However, transmission and distribution continued to be a monopoly market. As a result of the deregulation, companies that conducted business in both areas were divided into separate entities. The deregulation of the power market was seen as a natural step after the successful deregulation and privatization of telecom, postal, domestic aviation, taxi and railway markets. These political acts were carried out in the belief that it would lead to greater social benefit, an idea known already from the 18th century when Adam Smith coined the phrase "the invisible hand" in his famous book "The Wealth of Nations" from 1776. The deregulations later lead to the establishment of Nord Pool, the Nordic power exchange [3].

2.2 Nord Pool

The power exchange was established in 1992 after the deregulation of the power markets in Norway and in 1996 the exchange became a mutual market place for Norway and Sweden under the name Nord Pool. Today Nord Pool is a multinational platform where countries such as Norway, Sweden, Denmark, Finland, Latvia, Estonia, and Lithuania participate. At Nord Pool there's a market for physical delivery (Elspot and Elbas) and financial derivatives, e.g. futures and options with cash settlement (Eltermin). The marketplace consists of approximately 360 participants, such as producers, brokers, large consumers, and TSOs (Transmission System Operators). During 2016 a total amount of 505 TWh were traded on Nord Pool [3].

2.3 Elspot

Elspot is Nord Pool's instrument that enables market participants to buy physical electricity with Day Ahead delivery. Every day until 12 o'clock hourly bid and offer prices are sent in by the participants interested in buying or selling electricity at a certain price the next day. The hourly bid and offer prices are aggregated into supply and demand curves, the intersection of the curves becomes the equilibrium price for each respectively hour, and this is shown graphically in Figure 1. Since this method is applied, it's theoretically possible that the supply and demand curves do not have an intersection which would lead to an absence of price, this has so far never occurred for this market.

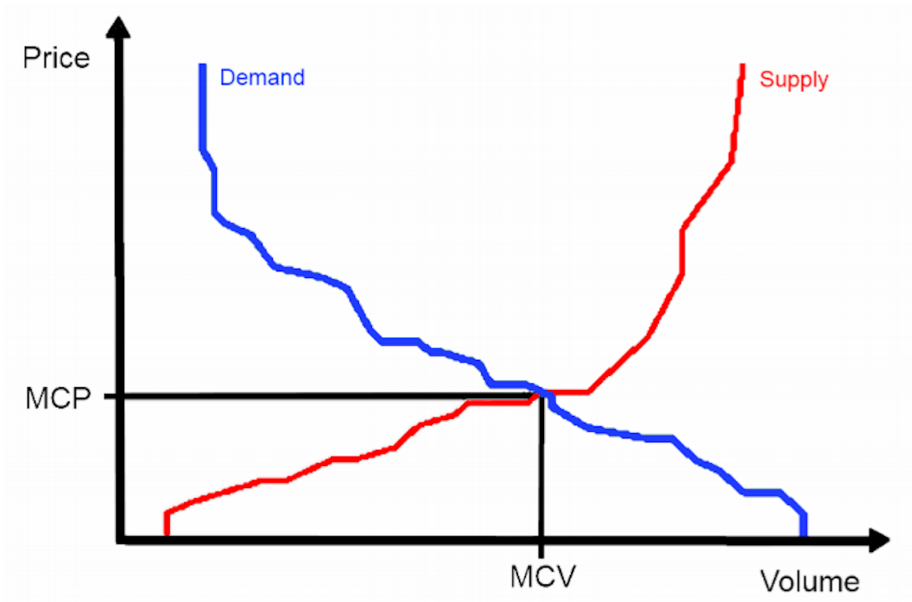

Figure 1: Aggregated supply and demand curves. Figure taken from Løkken Walter [4].

There are three different order types available for the participants active on Elspot. The most common order type is hourly orders where the participant sends in bids/offers and desired volume for each hour of the day. Another type of order is Block orders which means that the participant sends an order for at least 1 MW during minimum three consecutive hours. Block orders are of the type fill-or-kill, which means that the order only is executed if the whole specified volume can be filled. The third order type is called flexible sell orders, this order type uses a complex algorithm in order to achieve an optimal solution for the participant where price has the highest priority and time is second.

2.4 Local Pricing

The notion that different countries have different prices on electricity is easy to accept, demand and supply drives the prices and these might differ throughout. The idea of different prices within a country sounds more distant, however this is sometimes the case. This market dynamic can be explained with Sweden as a base case. Sweden's power grid is divided into four main areas; SE1 - Luleå, SE2 - Sundsvall, SE3 - Stockholm and SE4 - Malmö. It is not always in the interest of the power generators to supply electricity. For example a company that generates electricity via hydropower might be better off by refraining to generate power and instead let the water reservoir levels rise, which in fact is one of the few ways to "store" electricity in an effective manner. As hydropower is an important part of the power ecosystem in Sweden, this might lead to different pricing. Another reason for local pricing are bottle necks in the power grid, for example when cheaper electricity from one area might be blocked due to finite capacity in the grid transmission which leads to different pricing. This inherent feature of the electricity price makes it an interesting commodity to model, this thesis will however only use data collected from SE3 - Stockholm, making the prices comparable.

3 Theory

3.1 Time Series

This section is written in order to make this thesis more accessible and understandable for readers, the following theory is gathered from [1] unless stated otherwise. Time series are defined as observations of a variable, often at equally spaced time points, examples of common time series are the hourly air temperature, the daily closing price of a stock, or the monthly revenue of a grocery store. In this thesis, we will focus on time dependent times series, i.e. observations at time t are correlated with the previous observations. The end goal of time series analysis is often to be able to forecast the future in a sufficient manner. Companies could for example use reliable estimates about the future in order to optimize their revenues and earnings.

3.2 Stationarity

Stationarity is an important concept in time series analysis. A time series y_t is said to be strictly stationary if $(y_{t_1},...,y_{t_n})$ has the same joint distribution as $(y_{t_1+s},...,y_{t_n+s})$, $\forall s$ where $n > 0$. That is $(y_{t_1},...,y_{t_n})$ is required to be time-invariant. This requirement is however hard to prove in practice and a weaker form of stationary is defined. A time series y_t is said to be weakly stationary if E$[y_t]$=μ and Cov(y_t,y_{t+s}) does not depend on t. From the definitions we see that if y_t is strictly stationary with finite first- and second moment y_t is also weakly stationary, however the converse is not true in general [2].

3.3 Backshift Operator

The Backshift Operator is a mathematical operator often used to transform non-stationary series. The process of transforming time series using the backshift operator is referred to as differencing. We define the first order differencing as

$$\Delta y_t = y_t - y_{t-1} = (1 - B)y_t \ .$$

Where B is the backshift operator. Differencing can be applied multiple times to obtain a stationary time series, but sometimes differencing of higher order, i.e. with lag > 1, is needed to handle seasonality in the data. Seasonal differencing is defined as

$$\Delta_s y_t = y_t - y_{t-s} = (1 - B^s)y_t \ .$$

3.4 Autocorrelation Function

The Autocorrelation Function (ACF) is defined as the linear dependency between y_t and its past value y_{t-s}

$$\rho_s = \frac{Cov(y_t, y_{t-s})}{\sqrt{var(y_t)var(y_{t-s})}} \ .$$

Under the assumption of weak stationarity this becomes

$$\rho_s = \frac{Cov(y_t, y_{t-s})}{var(y_t)} \ .$$

With $-1 \leq \rho_s \leq 1$. For a given sample series, the Sample Autocorrelation Function for y_t with lag=s is defined as

$$\hat{\rho}_s = \frac{\sum_{t=s+1}^{T}(y_t - \bar{y})(y_{t-s} - \bar{y})}{\sum_{t=1}^{T}(y_t - \bar{y})^2} \ .$$

3.5 Partial Autocorrelation Function

The Partial Autocorrelation Function (PACF) gives the correlation of different lags for the times series conditioned on the shorter lags. We will allow ourselves to introduce this concept through a AR(p) model, analogously with the description in [1]. We have the following set of AR(p) models, with increasing quantity of parameters.

$$y_t = \phi_{0,1} + a_t + \phi_{1,1}y_{t-1}$$

$$y_t = \phi_{0,2} + a_{2t} + \phi_{1,2}y_{t-1} + \phi_{2,2}y_{t-2}$$

$$y_t = \phi_{0,3} + a_{3t} + \phi_{1,3}y_{t-1} + \phi_{2,3}y_{t-2} + \phi_{3,3}y_{t-3}$$

$$y_t = \phi_{0,4} + a_{4t} + \phi_{1,4}y_{t-1} + \phi_{2,4}y_{t-2} + \phi_{3,4}y_{t-3} + \phi_{4,4}y_{t-4}$$

$$\vdots$$

The equations above are in the form of linear regression and can be estimated through the least square method. Furthermore the estimate of $\hat{\phi_{i,i}}$ is called the lag-i PACF sample of y_t. This can be interpreted as that the lag-i PACF is the added contribution of y_{t-i} to y_t in an AR(i-1) model. Therefore $\hat{\phi_{i,i}}$ should be close to zero for all $i > p$ for an AR(p) process. This can be used to determine the order of an AR model.

3.6 AR

A simple model that functions as a building block for more complex models is the Autoregressive (AR) model. The AR(1) model can be written as

$$y_t = \phi_0 + \phi_1 y_{t-1} + a_t \ ,$$

where a_t is a white noise series. Generalized, the AR model of order p can be written as

$$y_t = \phi_0 + \sum_{i=1}^{p} \phi_i y_{t-i} + a_t \ .$$

3.7 MA

Another building block for more complex models is the moving average model. In its simplest form, the MA(1) model can be written as

$$y_t = \phi_0 + a_t - \theta_1 a_{t-1} \ ,$$

where a_t is a white noise series. Generalized, the MA model of order q can be written as

$$y_t = \phi_0 + a_t - \sum_{i=1}^{q} \theta_i a_{t-i} \ .$$

3.8 ARMA

The ARMA model combines the Autoregressive and Moving Average models into one model, one benefit of this idea is that the number of parameters in the model can be kept small. A time series y_t is said to be a ARMA(1,1) model if it satisfies

$$y_t - \phi_1 y_{t-1} = \phi_0 + a_t - \theta_1 a_{t-1} \ ,$$

where a_t is a white noise series. The left-hand side of the equation is the AR component of the model and the right-hand side is the MA component. The model can be re-written as

$$y_t = \phi_1 y_{t-1} + \phi_0 + a_t - \theta_1 a_{t-1} \ .$$

The general ARMA(p,q) can be written as

$$y_t = \phi_0 + \sum_{i=1}^{p} \phi_i y_{t-i} + a_t - \sum_{i=1}^{q} \theta_i a_{t-i} \ ,$$

where a_t is a white noise series and p and q are non-negative integers.

3.9 ARIMA

ARIMA model is an extension of the ARMA model where I stands for integrated. The model takes care of time series that needs to be differenced before any model can be applied. By utilizing the backshift operator the general ARIMA(p,d,q) can be written as

$$(1 - \sum_{i=1}^{p} \phi_i B^i)(1-B)^d y_t = (1 - \sum_{i=1}^{q} \theta_i B^i) a_t \ ,$$

where a_t is a white noise series.

3.10 SARIMA

SARIMA stands for Seasonal Autoregressive Integrated Moving Average. The model is useful when modeling data that exhibits seasonal effects, i.e. when the data is differenced with a lag greater than one. The SARIMA model therefore has a set of parameters for the seasonal and the non-seasonal components. SARIMA(p,d,q)(P,D,Q)$_s$ can be written as

$$(1-\sum_{i=1}^{P}\phi_i B^{si})(1-B^s)^D(1-\sum_{i=1}^{p}\phi_i B^i)(1-B)^d y_t = (1-\sum_{i=1}^{q}\theta_i B^i)(1-\sum_{i=1}^{Q}\theta_i B^{si})a_t \ ,$$

where a_t is a white noise series and p,q,P and Q are non-negative integers. Since the model handles seasonal effect it can be useful in modeling the electricity price. As spot prices in the Nordic are said to exhibit seasonal effects on a daily, weekly and yearly basis, where the weekly effect is the most pronounced [4] [7].

3.11 Dynamic Regression

The univarite time series above play an important role in time series analysis, however they do not allow for the inclusion of exogenous variables that could influence the spot price such as consumption of electricity, temperature, water levels in reservoirs for hydropower generators etc. In order to include these variables that might influence the price we define a family of models called dynamic regression models, in this thesis we will limit ourselves to only discuss regression models with time series errors, a generic model can be written as

$$y_t = \beta x_{t-1} + e_t \ ,$$

where e_t follows a time series model, such as ARIMA and SARIMA defined above.

3.12 AIC

There are several ways to determine the suitability of an estimated model, one of which is the Akaike Information Criterion (AIC). A good model should have as low AIC-value as possible, even though aspects such as over fitting always should be kept in mind when choosing a model. AIC is defined as

$$AIC = 2k - 2ln(\hat{L}) \ .$$

Where k is the number of free parameters and $ln(\hat{L})$ is the maximised value of the log-likelihood function.

3.13 MAPE

Since our end goal is to answer the question "does a model with a exogenous variable predict the future spot price better than a model without?" we need one, or more, metrics to determine forecasting capabilities. The first metric that we will introduce is mean absolute percentage error (MAPE). MAPE, which is usually expressed in terms of percentage, can be defined in the following way

$$MAPE = \frac{100}{n} \sum_{t=1}^{n} \mid \frac{y_t - f_t}{y_t} \mid ,$$

where n is the number of point forecasts, y_t is the actual value and f_t is the forecasted value. One limitation of MAPE is that division by zero can occur if $y_t = 0$, however we will not run into this problem since all our observation is non-zero.

3.14 MPE

Another useful forecasting metric is mean percentage error, which is defined as

$$MPE = \frac{100}{n} \sum_{t=1}^{n} \frac{y_t - f_t}{y_t} .$$

Since MPE does not use absolute values positive and negative forecasting errors will cancel each other out, making MPE a complementary metric to MAPE and a good way to evaluate bias in the forecast.

3.15 ADF

In order to verify the existence of a unit root in an AR(p) process, one may perform a test known as Augmented Dickey Fuller (ADF). This is done by test of $\mathrm{H}_0 : \beta = 1$ against $\mathrm{H}_1 : \beta < 1$ using the regression

$$X_t = c_t + \beta X_{t-1} + \sum_{i=1}^{p-1} \Delta X_{t-i} + e_t$$

where c_t is a deterministic function of the time index t and $\Delta X_j = X_j - X_{j-1}$ is the differenced series of X_t. Therefor the t-ratio of $\hat{\beta} - 1$ is known as the ADF-test. Mathematically this is expressed as

$$\text{ADF-test} = \frac{\hat{\beta} - 1}{std(\hat{\beta})}$$

where $\hat{\beta}$ is the least-squares estimate of β. The results of a ADF-test is easily interpreted. If $\mathrm{H}_0 : \beta = 1$ is rejected, the times series is considered stationary.

4 Data

This thesis aims to answer questions regarding if time series models through inclusion of exogenous variables can improve the prediction of the electricity spot price. The main data series considered is therefore the historical electricity spot prices from 01/01/1996 to 31/12/2016 on a daily resolution, gathered from Nord Pool [3]. More specifically it is the historical SE3 spot price, i.e. the price from the Stockholm region in Sweden. The data provided by Nord Pool can be seen in Figure 2. At first glance the spot price seems to endure high volatility and volatility clustering. In matter of fact electricity is indeed highly volatile compared to other commodities such as gold or oil, this makes the power market dynamics complex and a challenge in terms of modeling procedures. From Figure 2 we also draw the conclusion that the spot price is far from stationary, which is an assumption for the models described in the theory section. We make a remark and conclude that we need to achieve stationarity before any modeling work can begin. Regarding seasonality no strong assessments can be made ocular and thus statistical methods will have to be applied later to discover for which time frame seasonality can be found, if any.

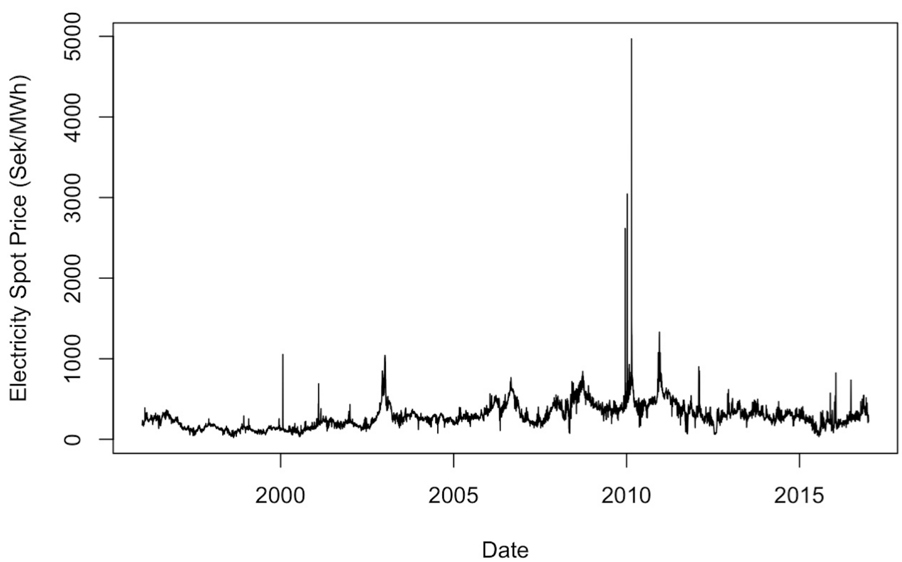

Figure 2: Daily electricity spot closing prices.

The exogenous data series to be included in a dynamic regression model is temperature data from the Stockholm region in Sweden. The data was gathered from the website of SMHI [9]. The data contains historical daily mean temperatures at the Stockholm Old Astronomical Observatory between 1756-2015. Data from 2015 and later are obtained from SMHI "Öppna Data" [9]. The data was adjusted in order to stretch from 01/01/1996 to 31/12/2016; the data is plotted in Figure 3 below.

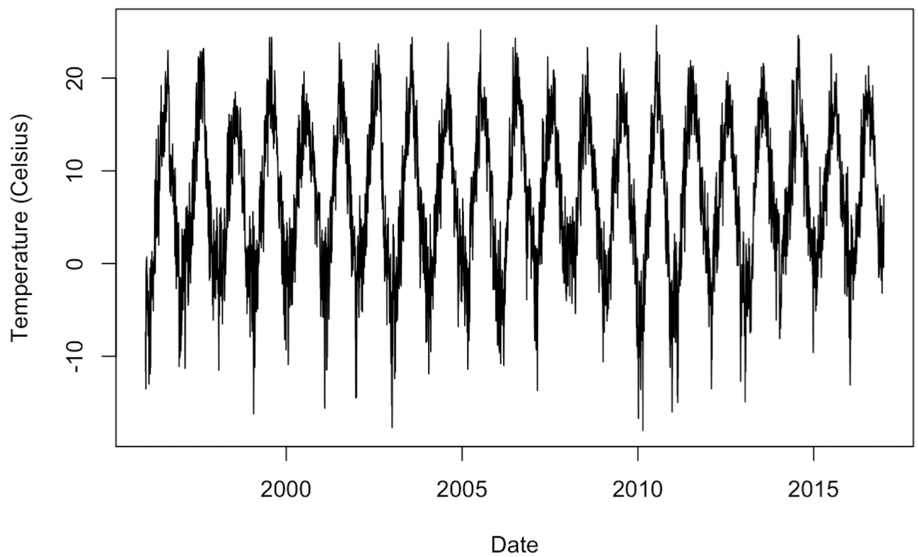

Figure 3: Daily mean temperatures at the Stockholm Old Astronomical Observatory.

The temperature data oscillates in a tight range between -10 and 20 Celsius. Furthermore, one can see clear seasonal effects, mainly on a yearly basis. But our knowledge about temperature data tells us that seasonality should be found on an hourly an monthly basis as well.

4.1 Data Preprocessing

Before any models can be fitted to the underlying data, the basic assumptions of normality and weakly stationarity has to be assessed. The spikes in electricity spot price data (see Figure 2) could be viewed as outliers because price spikes could be a result of non-recurring events such as transmissions congestions in the power supply, however removing data point is never desirable due to loss of useful information. Instead a logarithmic transformation is applied to the data in order to control the spikes. The impact of the transformation is examined by plotting the sample quantiles of the data against the theoretical normal quantiles for both the spot price and the transformed spot price in Figure 4. As expected the transformation handles the heavy tails created by the price spikes in such a manner that the transformed data can be viewed as normally distributed, even though some heavy tails still persist.

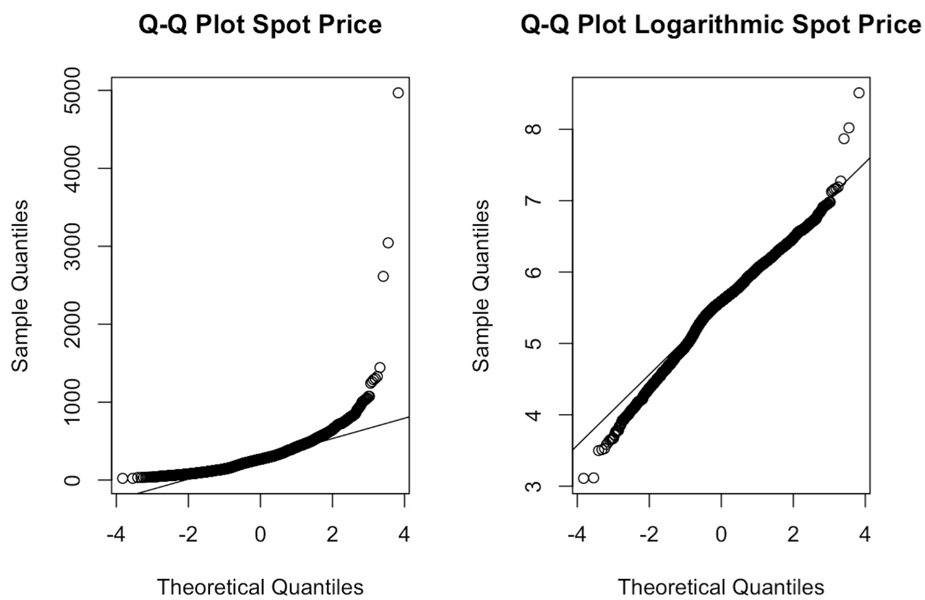

Figure 4: Q-Q plot for Spot Price and Logarithmic Spot Price.

After normality can be assumed, tests are performed to evaluate if the data can be considered stationary. Via an ACF plot and augmented Dickey–Fuller test it can be concluded that the time series is not stationary [Figure 19 in Appendix]. Previous work demonstrates that the spot price data usually show clear signs of weekly seasonality. The weekly effect could be derived from the fact that large electricity consumers, e.g. industry, use significant less electricity during the weekend and since electricity cannot be stored in a scalable way, prices should drop. The data is therefore differenced with a lag of seven. Electricity spot price often also shows signs of daily and yearly effects, however clear signs of this could not be found and therefore only weekly seasonality was taken into consideration. The transformed data are again evaluated with respect to stationarity. The null hypothesis is H0: = non-stationary against H1: = stationary, our significance level is set to 0.05, i.e. the p-value has to be lower than 0.05 for the data to be considered stationary.

```
##
##  Augmented Dickey-Fuller Test
##
## data:  diff(log(elpris), lag = 7)
## Dickey-Fuller = -21.324, Lag order = 19, p-value = 0.01
## alternative hypothesis: stationary
```

Figure 5: Augmented Dickey-Fuller Test for differenced logarithmic Spot Price.

With a test value of -21.324 and a p-value of 0.01, the differenced data passes the augmented Dickey–Fuller test and graphically it also seems stationary. Since the data through logarithmic transformation and seasonal differencing now fulfills the assumptions, a model can be fitted to the data.

5 Analysis

5.1 Model

The logarithmic and seasonally differenced data passes the model assumptions and suitable models that explain the dependencies in the data can be fitted. In Figure 6 the ACF of the processed data is plotted, the pattern in the ACF exhibits exponentially decreasing lags, a feature that is typical for AR models. Since the ACF exhibits these features it can be concludes that the transformed data is not white noise. It therefore contains stochastic dependencies that need to be modeled.

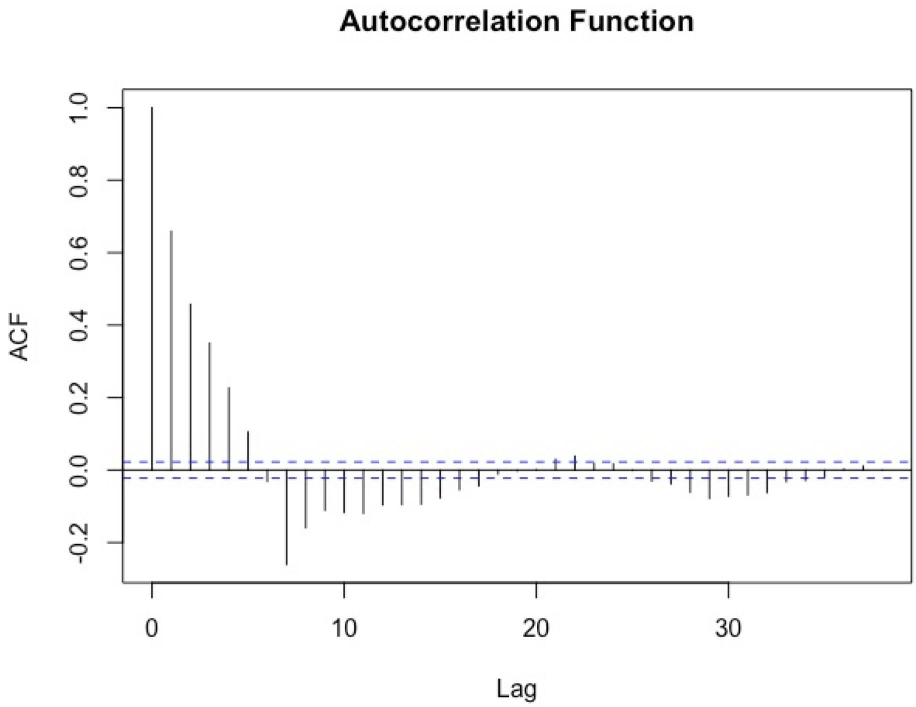

Figure 6: Autocorrelation Function for transformed data.

Due to the seasonal differencing a SARIMA(p,0,q)(P,1,Q)$_7$ model is initially chosen, where adequate non-negative integers for p,q,P, and Q needs to be found. A simple SARIMA(1,0,0)(0,1,0)$_7$ model without any seasonal terms is fitted based on the AR behavior found in Figure 6 and in order to examine if the seasonal effect can be detected in the data. Since the model doesn't include any seasonal terms we expect a spike at every 7th lag in the ACF and PACF plots. The residuals of SARIMA(1,0,0)(0,1,0)$_7$ model display spikes at every 7th lag, which confirms the weekly seasonality in the data [Figure 20 in Appendix]. Due to the AR-like behavior in Figure 6 and the detected weekly seasonality a SARIMA(1,0,0)(0,1,1)$_7$ model is fitted. The model is chosen due to its simplicity, and we want to examine how much of the dependencies in the data that can be captured with it.

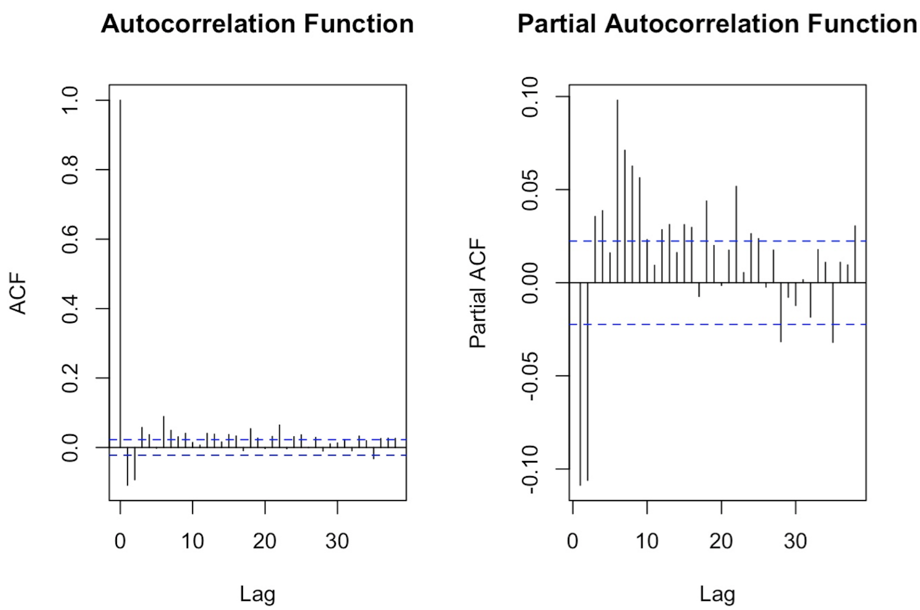

Figure 7: Autocorrelation Function and Partial Autocorrelation Function for model 1.

The residuals of the model are plotted in an ACF and PACF plot in Figure 7. The plots show that model captures most of the dependencies, especially the seasonality is taken care of through the seasonal MA term in the model. The residuals are accepted as white noise even though some significant lags do exist. Therefore we test other models to find a model that explains dependencies in an even more satisfactory manner. All the contemplated models are summed up in Table 1 where they are sorted based on their AIC values.

SARIMA(p,d,q)(P,D,Q)$_s$	AIC
$(2,0,2)(1,1,2)_7$	-10448.48
$(2,0,2)(0,1,2)_7$	-10370.79
$(1,0,1)(1,1,1)_7$	-10123.71
$(1,0,1)(0,1,1)_7$	-10076.86
$(2,0,0)(0,1,2)_7$	-10028.79
$(2,0,0)(0,1,1)_7$	-9996.81
$(1,0,0)(0,1,1)_7$	-9871.88

Table 1: AIC value for the contemplated models.

The model with the lowest AIC value is SARIMA$(2,0,2)(1,1,2)_7$. Since the AIC is designed to penalize models with extra parameters we draw the conclusion that the model is indeed interesting for further analysis, we keep both

SARIMA(1,0,0)(0,1,1)$_7$ and SARIMA(2,0,2)(1,1,2)$_7$ as possible options for the final model. The ACF and PACF for the residuals of SARIMA(2,0,2)(1,1,2)$_7$ are plotted below, the model seems to explain all the dependencies in the data and the remaining residuals are considered to be white noise.

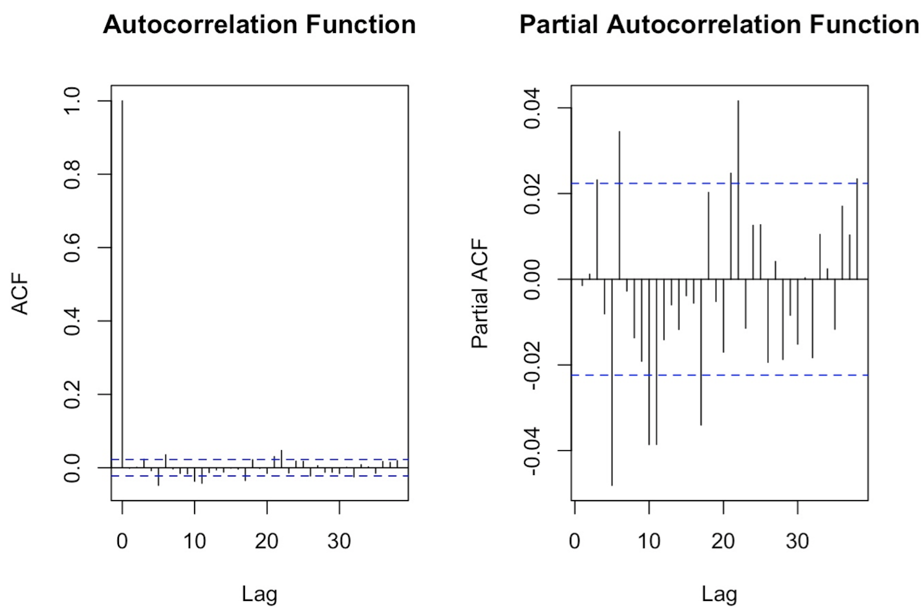

Figure 8: Autocorrelation Function and Partial Autocorrelation Function for model 2.

When examining the residual plots [Figure 21 in Appendix] of the models above tendencies of spikes and volatility clustering appear. This is clear signs of GARCH effects, however this is determined to be outside of the scope of this thesis and we move on to model validation.

5.2 Model Validation

From the modeling procedure two models that explain the electricity spot price in a sufficient way is developed. In order to continue this process in a prudent manner the models has to be validated. The models should have variables with significant coefficients, and the residuals should follow a symmetric distribution such as the normal or t-distribution. Model 1 is SARIMA(1,0,0)(0,1,1)$_7$, i.e. a model with one AR term and a seasonal MA term. In order to test the significance of the coefficients a z test is performed. The z test evaluates if respective coefficient is non-zero, thus the null hypothesis is $H_0 := Coefficient_i = 0$ against the alternative hypothesis $H_1 := Coefficient_i \neq 0$. As displayed in Figure 9 both estimated coefficients for model 1 has a p-value < 0.01 and the null hypothesis is rejected.

```
##
## z test of coefficients:
##
##         Estimate Std. Error z value  Pr(>|z|)
## ar1   0.8584365  0.0074700 114.918 < 2.2e-16 ***
## sma1 -0.8676863  0.0096006 -90.378 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 9: Z test for model 1.

We therefore move on and examine the residuals. Firstly a Q-Q plot with the residuals quantiles versus the theoretical normal quantiles is plotted. The plot show signs of heavy, yet symmetrical, tails and we conclude that the residuals are not normally distributed. As the tails are to be considered symmetrical a t-distribution is considered instead. Empiricism concludes that a t-distribution with three degrees of freedom is the most suitable candidate and the Q-Q plot is deemed satisfactory even though some deviant and heavy tails persists.

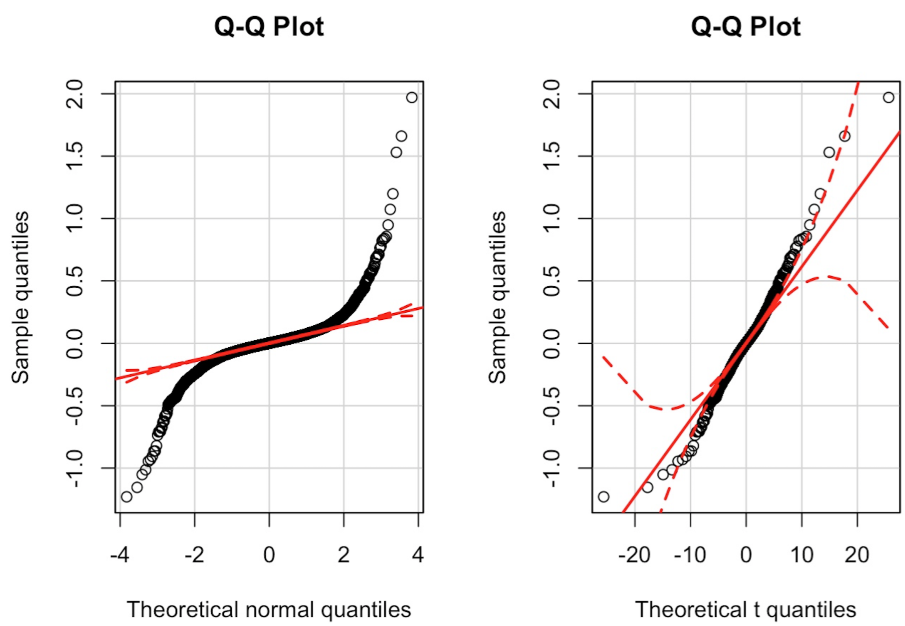

Figure 10: Q-Q plots for the residuals of model 1.

The second model, model 2, is SARIMA(2,0,2)(1,1,2)$_7$ which was selected because it yielded the lowest AIC value of all the contemplated models. The model consists of two AR terms, two MA terms, one seasonal AR terms, and two seasonal MA terms. We start by conducting a z test to test the null hypothesis is $H_0 := Coefficient_i = 0$. All coefficients are significant on a 0.05 significance level, however the p-value of the second AR terms is 0.042 which means its significance would be rejected on a chosen significance level of < 0.4. Since we try to avoid over fitting and unnecessary complexity by inclusion of extra parameters, the second AR term is excluded from the model. Model 2 is then SARIMA(1,0,2)(1,1,2)$_7$ where all parameters are significant on a significance level of 0.01. In accordance with the procedure for model 1, the sample quantiles for the residuals of model 2 are plotted against the theoretical normal quantiles. As with model 1 the residuals have heavy tails and we draw the conclusion that the residuals are not normally distributed. The theoretical quantiles of the t-distribution with three degrees of freedom leaves us with a more satisfactory result as shown below in Figure 11, we accept that the residuals can be considered to be t-distributed and the models are deemed good enough to continue the modeling procedure.

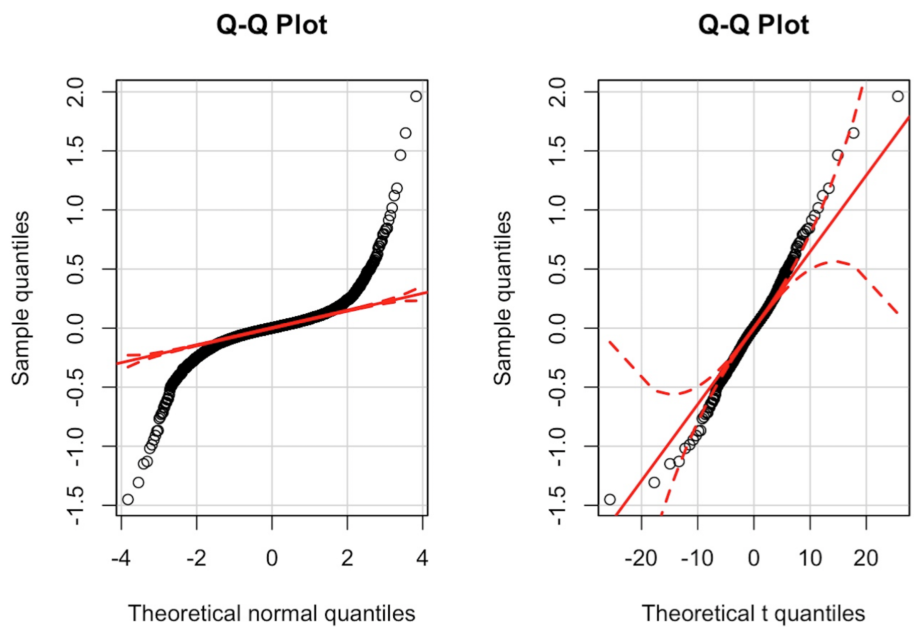

Figure 11: Q-Q plots for the residuals of model 2.

5.3 Exogenous Variable

The main question we want answer with this thesis is if the inclusion of exogenous variables, here through a regression model with time series errors (dynamic regression), creates a model with stronger predicting power. In practice, we will keep our two models and extend them by adding the exogenous variable as an explanatory variable. Possible exogenous variables that could increase the prediction power of the models include, but are not limited to, electricity consumption, electricity production, water reservoir levels, and temperature. In this thesis only temperature data will be considered in the question if exogenous variable can improve the model in terms of forecast ability. In order to extend the model, the lag number and the number of terms to include should be determined. For the sake of simplicity and to avoid the common pitfall of over fitting, which would lead to a model with low predicting power, the number of additional terms is simply set equal to one. In order to determine the lag number to be used for the exogenous term we study the cross-correlation function between the electricity spot price and the temperature data. Electricity spot price is set as x_t and the temperature as y_t. The cross-correlation function shows the dependencies between x and y_t for different time lags t. Below, in Figure 12, the cross-correlation function between the endogenous and exogenous variable is plotted for different lags, bear in mind that each lag represent a time jump of one day. The correlation function reaches its peak at lag equal to zero, i.e. the spot price correlates best with the temperature the same day. However, in a forecasting procedure inclusion of an exogenous variable with lag zero is not possible since that value itself is unknown. A possible remedy for this is to forecast the exogenous variable or to set the lag equal to one. In this thesis, the second option is chosen and the lag is set equal to one.

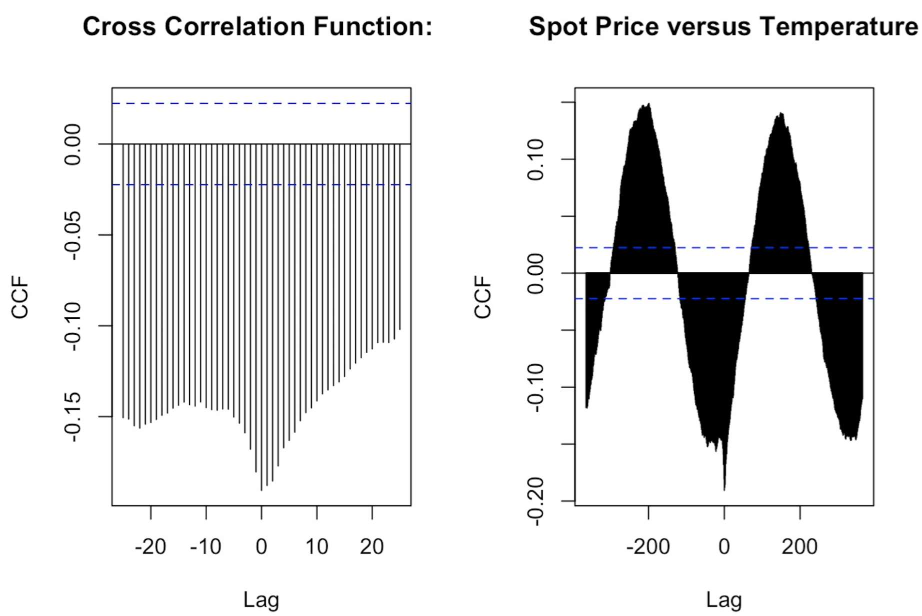

Figure 12: Cross Correlation Function between electricity Spot Price and mean Temperature.

We therefore choose to include one exogenous variable to model 1 and model

2 and to set the lag equal to one, this gives us two new models in the form of

$$y_t = \beta x_{t-1} + e_t \ ,$$

where e_t is the time series error. When e_t is modeled the order of model 1 and 2 is used to make the models comparable, the coefficients are however re-estimated.

The two new models, from here on referred to as model 3 and model 4, also have to be validated before we can move on to forecasting. We will start with model 3, that is the SARIMA(1,0,0)(0,1,1)$_7$ model with the inclusion of an explanatory variable for temperature. Model 3 ACF and PACF show similar behavior as for model 1, we therefore conclude that model 3 according to the ACF and PACF explains the dependencies in the data in a sufficient way [Figure 22 in Appendix]. The Q-Q plots also follow the same pattern as for model 1, a t-distribution with three degrees therefore is fitted and seems to describe the residuals in a sufficient manner. Furthermore a z test of the coefficients for model 3 are performed in order to evaluate the significance of the coefficients, the null hypothesis is $H_0 := Coefficient_i = 0$. The result of the z test is summarized in Figure 13, all coefficients are significant on a significance level of 0.01 and the model explain the variations in the spot price well.

```
##
## z test of coefficients:
##
##         Estimate  Std. Error z value  Pr(>|z|)
## ar1    0.84509610  0.00764333 110.567 < 2.2e-16 ***
## sma1 -0.86406151  0.00928887 -93.021 < 2.2e-16 ***
## temp -0.00902867  0.00064554 -13.986 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 13: Z test for model 3.

The fourth model, model 4, is SARIMA(1,0,2)(1,1,2)$_7$ with the inclusion of an explanatory variable for temperature. We begin the model validation by studying the ACF and PACF of the residuals. The inclusion of the exogenous variable doesn't change pattern of the ACF and PACF notably with respect to the ACF and PACF of model 2. Therefore it's concluded that the model is satisfactory given our data. Examining the residuals gives the same results, the Q-Q plot with residuals quantiles versus theoretical quantiles is far from satisfactory and instead a t-distribution with three degrees therefore is fitted and seems to describe the residuals in a sufficient manner. Additionally in accordance with previous models a z test is performed to test the significance of the coefficients

```
##
## z test of coefficients:
##
##          Estimate   Std. Error z value  Pr(>|z|)
## ar1   0.97441547  0.00393122 247.866 < 2.2e-16 ***
## ma1  -0.27944582  0.01194816 -23.388 < 2.2e-16 ***
## ma2  -0.19520109  0.01216930 -16.041 < 2.2e-16 ***
## sar1  0.81564348  0.02592043  31.467 < 2.2e-16 ***
## sma1 -1.71579170  0.02903834 -59.087 < 2.2e-16 ***
## sma2  0.71857850  0.02878161  24.967 < 2.2e-16 ***
## temp -0.01051417  0.00063574 -16.538 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 14: Z test for model 4.

Figure 14 shows that all coefficients, including the exogenous coefficient, are significant on a significance level of 0.01 and the model seems to explain the variations in the spot price well. All four models have residuals that can be considered white noise and significant coefficients, it is therefore time to put the models to test through forecasting.

5.4 Simulation

A complementary model validation method to the ones used above is simulation. Via simulation from a specified model it can be determined if a model is suitable given the underlying data and if the coefficient estimate seems reasonable. To demonstrate this procedure we will simulate from model 2, i.e. SARIMA(1,0,2)(1,1,2)$_7$. When the coefficients are estimated with the logarithmic spot price as underlying data the following estimates are obtained: $AR_1 = 0.97$, $MA_1 = -0.246$, $MA_2 = -0.188$, Seasonal $AR_1 = 0.82$, Seasonal $MA_1 = -1.718$, and Seasonal $MA_2 = 0.72$. 1000 time series are then simulated, each series containing 10000 data points. The specified model, i.e. SARIMA(1,0,2)(1,1,2)$_7$, is fitted for each simulation. The coefficient estimates of this procedure is summarized below in Figure 15.

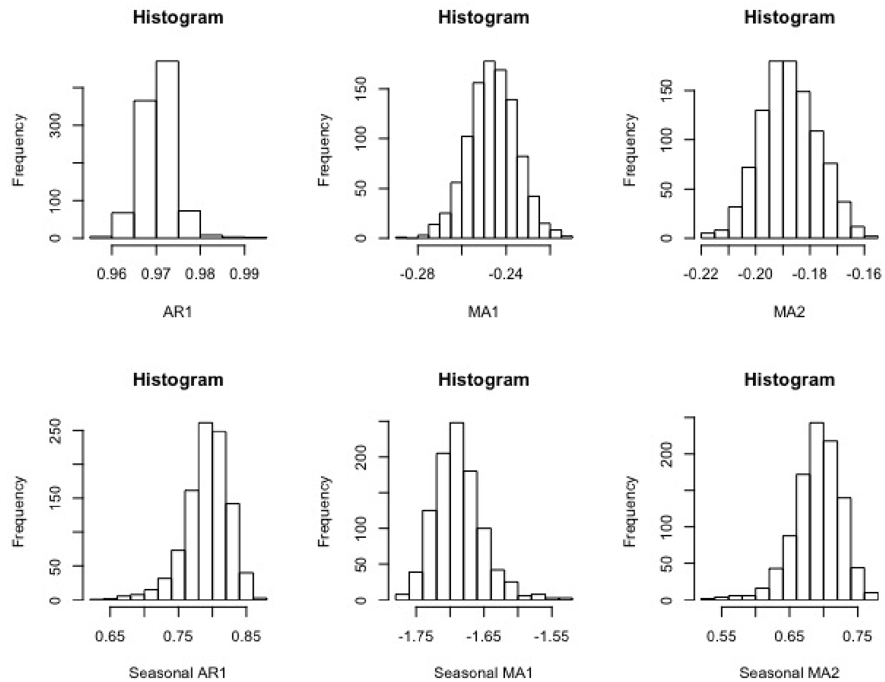

Figure 15: Simulated coefficients from model 2.

As this process is computer intensive only 1000 time series are simulated. At 1000 estimated coefficients per histogram the plots appear to be normally distributed and centered around the estimates obtained with the logarithmic spot price as underlying data. However it is noted that the seasonal histograms do endure a mild skewness, this could be a result of the estimation process used which is to use conditional-sum-of-squares to find starting values, and then maximum likelihood. We conclude that our estimated coefficients are indeed reasonable. In order to examine the trajectories of the simulated time series, 10 series are plotted below beside the trajectory of the observed spot price.

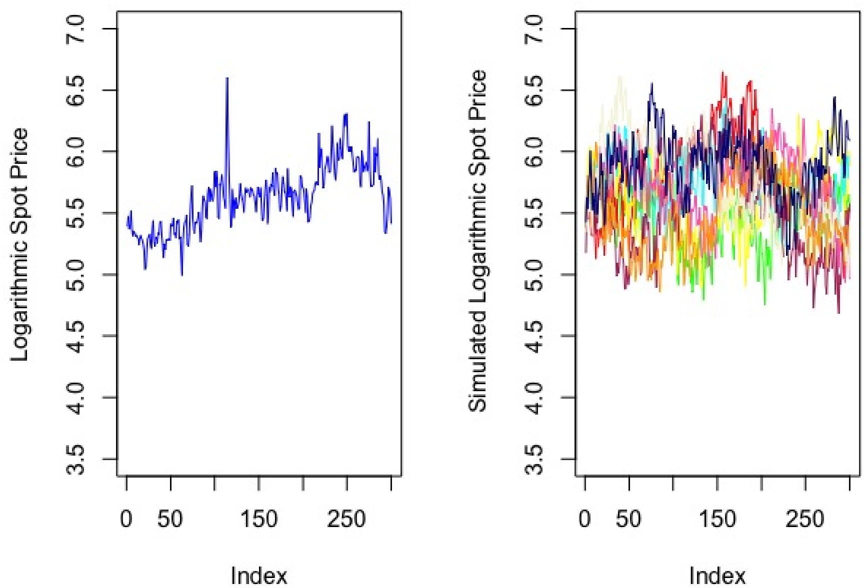

Figure 16: Observed trajectory (left) and Simulated trajectories from model 2 (right).

The simulated time series moves in a range in-line with the observed values, furthermore the trajectories exhibit features that reminds of the ones found in the underlying data. Our conclusion from the simulations is that the fitted model describes the dynamics of the data in a adequate manner.

6 Forecast

In order to answer the question "does a model with an exogenous variable predict the future spot price better than a model without?" we have to forecast electricity spot prices for each model respectively and compare them with adequate metrics. In the modeling procedure, we divided our data in to two parts. One part for training and estimating the parameters and one part for testing the models predicted values against observed values. The total dataset consist of 7671 data points, we will somewhat arbitrary use 7000 observations for training and 671 for forecasting. The forecasting period is therefore from 2nd of March 2015 to and including 31st of December 2016. The length of the testing period is from our point of view long enough to test the models under different market conditions, seasons, temperatures etc. The models have coefficients estimated based on the logarithmic spot price and the forecasted values therefore has to be transformed back by taking the exponential function of each value. The forecasting technique used will be limited to day ahead point forecast, i.e. the models will be feed with the last days true closing price and temperature in the case of model 3 and model 4. To compare and draw conclusion from the forecasting procedure we will mainly use MAPE and MPE as metrics for evaluation. We believe that MAPE and MPE are good complements where MAPE will be used to compare forecasting precision and MPE will be utilized in order to spot eventual bias, i.e. if the models systematically are over or under-predicting relative to the true value. To create a valid reference point the models will be compared to the Naïve approach. In Naïve forecasting the true value from time t will be used as a forecasted value for time $t+1$, even though this approach may seem simplistic it is quite powerful, however for the models to be useful in practice they should be able to forecast the spot price in a more sophisticated manner [8]. The result of the forecasts is summarized in Table 2 below.

Model	MAPE	MPE
Naïve model	14.06963	-2.31758
SARIMA$(1,0,0)(0,1,1)_7$	12.65822	-1.48914
SARIMA$(1,0,0)(0,1,1)_7$+Temp	12.62637	-1.44148
SARIMA$(1,0,2)(1,1,2)_7$	11.99884	-2.04787
SARIMA$(1,0,2)(1,1,2)_7$+Temp	11.92454	-2.01045

Table 2: Result of forecasting in terms of MAPE and MPE.

All four developed models perform better than the Naïve model with respect to MAPE, this tells us that the models do capture dependencies in the data in a sufficient manner. Also by including temperature as an exogenous the MAPE value was lowered both between model 1 and model 3, and between model 2 and model 4. Our conclusion from this is that the inclusion of exogenous variables does generate a model with higher forecasting abilities. Comparing SARIMA$(1,0,0)(0,1,1)_7$ to SARIMA$(1,0,2)(1,1,2)_7$ we observe a difference of 0.56 in MPE, and the same behavior is observed when adding a exogenous variable to respectively model. As MPE can be seen as a metric of bias SARIMA$(1,0,2)(1,1,2)_7$ is to be viewed as a more conservative model in the sense that it seems to underestimate the price systematically even though the model is more precise in absolute terms given our test data. To give the reader a more non-mathematical, and therefore more intuitive to some, sense of the models predicting power predicted values are plotted against observed values.

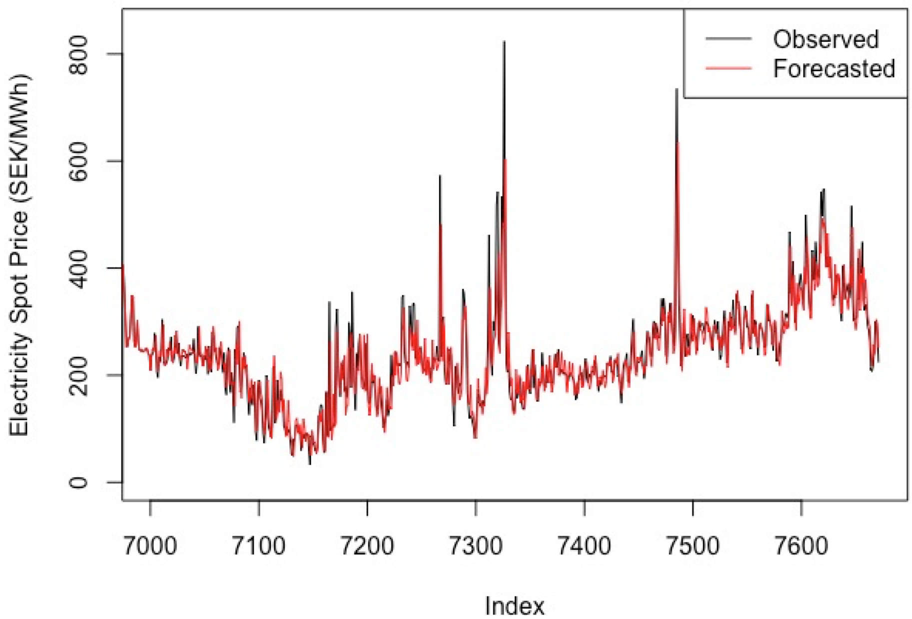

Figure 17: Forecasted Spot Price versus observed Spot Price.

Figure 17 show predicted values versus observed values for SARIMA(1,0,2)$(1,1,2)_7$+$Temp$, the model is able to predict the true values well except when it comes to sudden price spikes. Furthermore, we are interested to view the difference between observed and predicted values on a more granular level, Figure 18 displays the two lines for the last 70 data points of the time series.

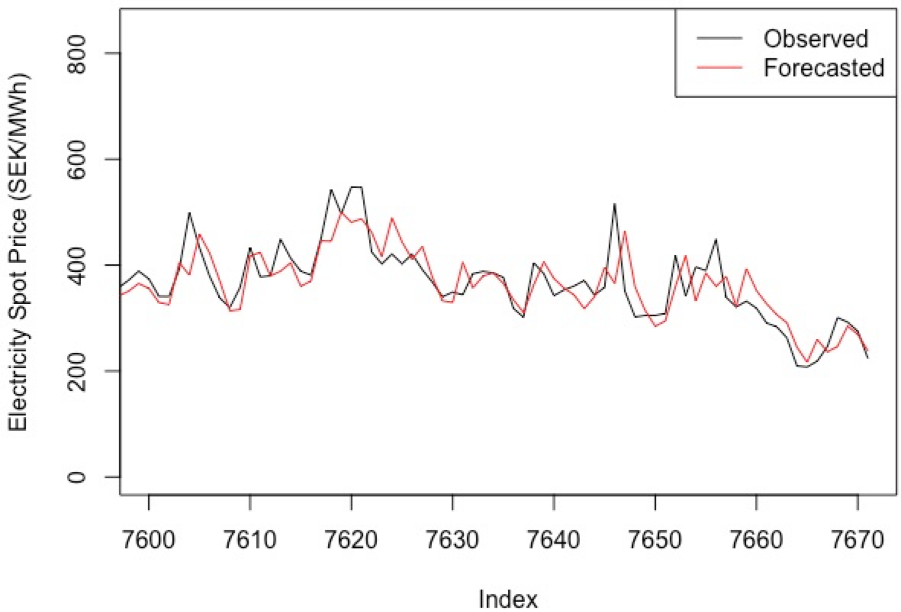

Figure 18: Forecasted Spot Price versus observed Spot Price.

7 Conclusion

The purpose of this thesis was two conduct a thorough analysis of the dynamics of the Swedish power market and to investigate if the inclusion of an exogenous variable could improve the predicting power of our models. Via the data management and modeling procedure four adequate models where found, validated and subsequently applied trough forecasting. The vision throughout the modeling process was to keep the model simple, and minimize the number of parameters to avoid over fitting and potential autocorrelation. Table 2 showed that the inclusion of temperature data indeed lowered the MAPE values and we therefore consider this to be a more powerful model from a forecasting point of view. Furthermore, we conclude that AIC value was a good indication of the model even though one parameter was dismissed due to insignificance. The top performing model in the forecasting procedure is considered to be SARIMA$(1,0,2)(1,1,2)_7$ + *exogenous variable* with the estimated coefficients AR=0.9743, $\text{MA}_1=-0.2790$, $\text{MA}_2=-0.1820$, Seasonal AR=0.7982, Seasonal $\text{MA}_1=-1.6981$, Seasonal $\text{MA}_2=0.7014$ and Temperature=-0.0095, where the coefficients are estimated based on the logarithmic price. We also concluded during the model validation that none of the models residuals were to be considered normally distributed and a t-distribution with three degrees of freedom was used as a substitute and accepted as satisfactory although the fit was not perfect and some deviant tails persisted for all model residuals. When plotting the residuals of the final model [Figure 24 in Appendix] the residuals exhibit spiky behavior, which indicates that some effect is not taken into account during the modeling work. This type of spikes and volatility clustering often is a sign of GARCH effects, however this was considered outside the scope of this thesis and is left as a possible option for further analysis.

8 Discussion

While we arrived at a well performing final model, with a MAPE of 11.92 percent it does have its flaws. In the conclusion section we argued that during the model validation the residuals were accepted as being t-distributed even if further investigation might be in place. The residuals also exhibit GARCH type of effects that were not included in the modeling procedure, instead this was deemed to be outside the scope of the thesis, we do however believe that the inclusion of this type of model might lead to a more robust and precise model in terms of forecasting. With that being said price spikes in the electricity spot price is often an effect of non-recurring events such as grid congestion, which is hard to anticipate and therefore model. During the data management, the data was transformed through applying the logarithm of base 10 to the data, this procedure was applied as an alternative to outlier management techniques discussed in other papers [5] [6]. The benefit of this alternative is that none of the data points were altered or removed and we make use of all the information in the original dataset. Seasonality is often a hot topic in papers regarding electricity price modeling, in the Nordics the spot price is said to exhibit seasonality on a daily, weekly, and yearly basis [7]. In this thesis where our data solely was on a daily resolution only a weekly effect were considered. We did however end up with a model that performed well in forecasting and seemed to explain the dependencies in the data with regards to the ACF and PACF and therefore it was concluded that seasonality on a yearly basis could be ignored. The main question of this thesis was to examine if our potential models would benefit in terms of predicting power by the inclusion of an exogenous variable. In the "Conclusion" section we argued that by including a variable for the local temperature in the Stockholm region a higher predicting power was achieved. This result could be extended by inclusion of other variables that might have an impact on the electricity price, one popular exogenous variable used in other papers is the electricity consumption in the relevant area. Further studies could therefore be to include more exogenous variables that correlate with the electricity spot price and to take volatility clustering into account in the modeling process.

9 References

[1] Tsay, Ruey S. 2010. *Analysis of Financial Time Series 3rd ed*, John Wiley Sons, Inc.

[2] Sheldon M.Ross. 2014. *Intoduction to Probability Models 11th ed*, Academic Press.

[3] Nord Pool, About Us, viewed 16 April 2017, http://www.nordpoolspot.com/.

[4] Erik Løkken Walter. 2011. Time Series Analysis of Electricity Prices: *A comparative study of power markets.* https://tinyurl.com/k4mkfzc. Viewed 10 April 2017.

[5] Johan Lindberg. 2012. A Time Series Forecast of the Electrical Spot Price: *Time series analysis applied to the Nordic power market.* http://www.diva-portal.org/smash/get/diva2:408045/FULLTEXT01. Viewed 15 April 2017.

[6] Trueck S, Weron R and Wolff R. 2007. *Outlier Treatment and Robust Approaches for Modeling Electricity Spot Prices*, Paper presented at the 56th Session of the International Statistical Institute.

[7] Weron R. 2004. Electricity price forecasting: *A review of the state-of-the-art with a look into the future*, International Journal of Forecasting Volume 30, Issue 4, http://www.sciencedirect.com/science/article/pii/S0169207014001083. Viewed 20 April 2017.

[8] Misiorek, A. 2005. *Forecasting spot electricity prices with time series models.* The European Electricity Market EEM-05, International conference.

[9] SMHI, Data, viewed 25 April 2017, http://www.smhi.se/klimatdata/meteorologi/temperatur.

[10] Kaminski, V . 1997. *The challenge of pricing and risk managing electricity derivatives, the US power market*, Barber, P. (ed.), Risk Books.

[11] Bunn, D., Karakatsani, N. 2003. *Forecasting Electricity Prices.* London Business School Working Paper.

10 Appendix

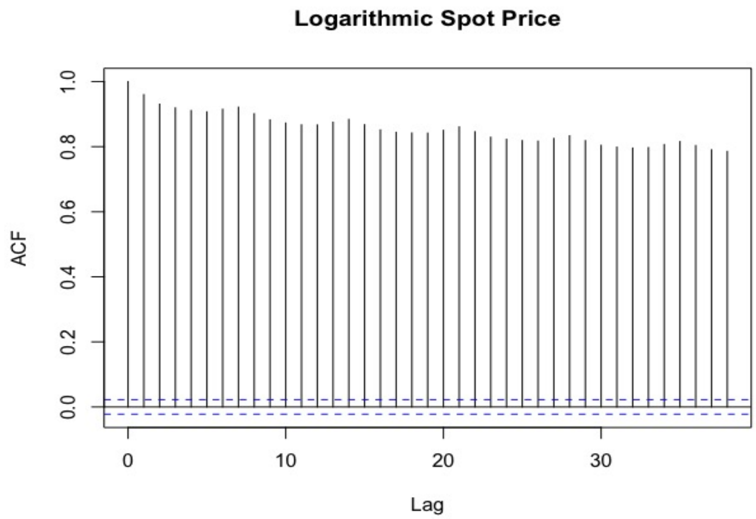

Figure 19: ACF plot for the logarithmic electricity spot price. The plot displays non-stationary data.

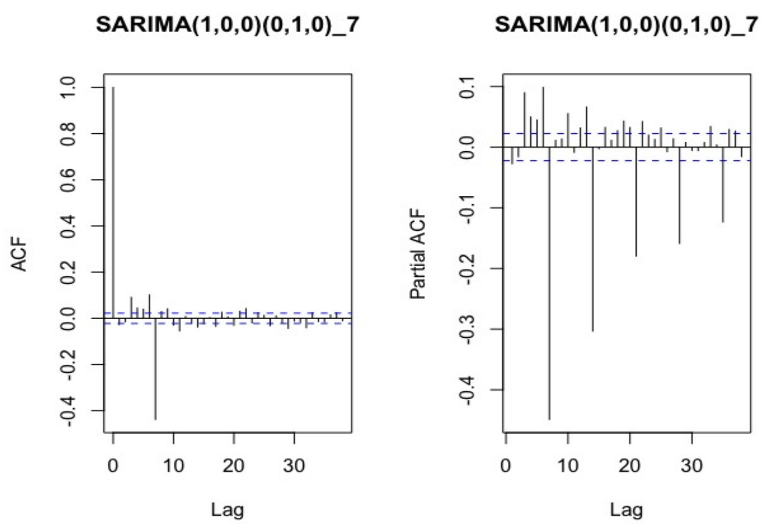

Figure 20: ACF and PACF plot for the model residuals. The plot displays seasonality on a weekly basis.

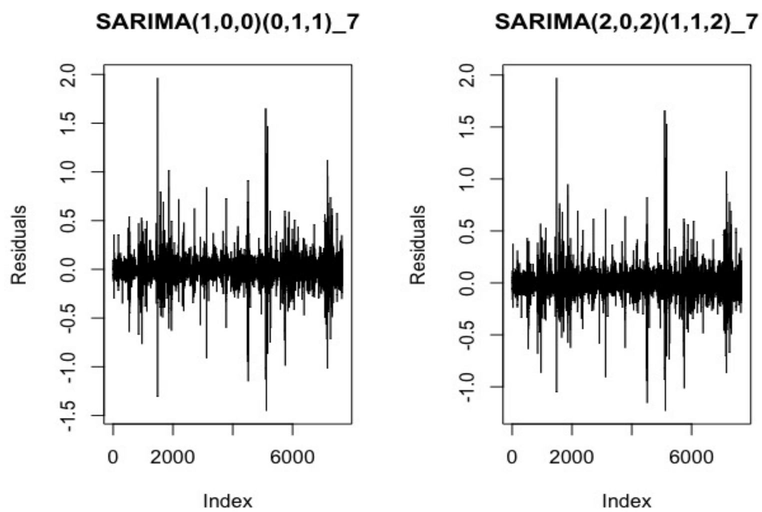

Figure 21: Residual plots for the contemplated models. Both models endure spikes and volatility clustering which typically is signs of GARCH effects.

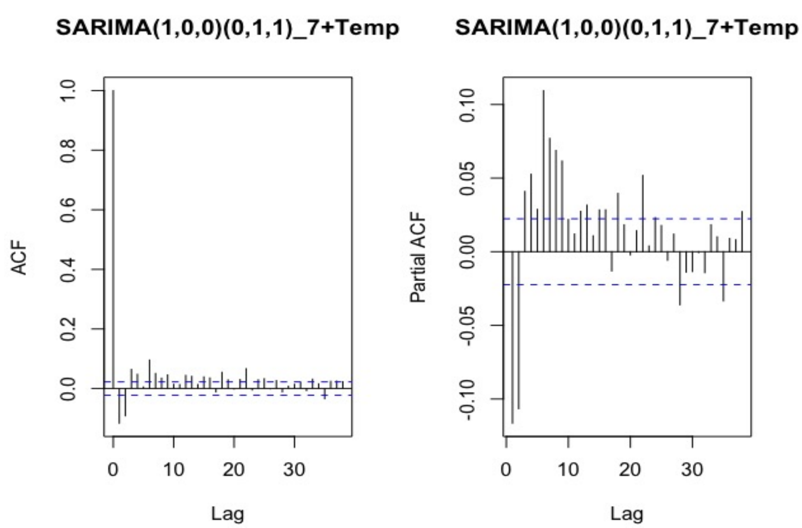

Figure 22: ACF and PACF plot for model 3. The model explains the dependencies in the data in a sufficient way, even though some significant lags persists.The residuals are considered white noise.

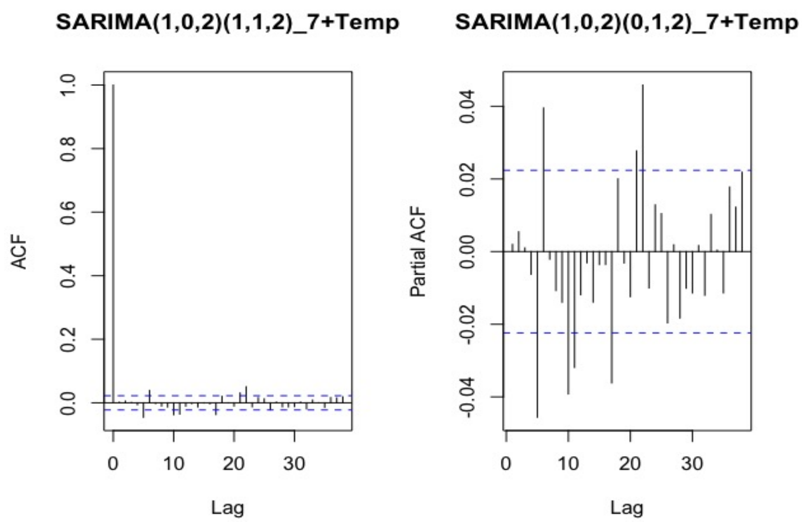

Figure 23: ACF and PACF plot for model 4. The model explains the dependencies in the data in a sufficient way. The residuals are considered white noise.

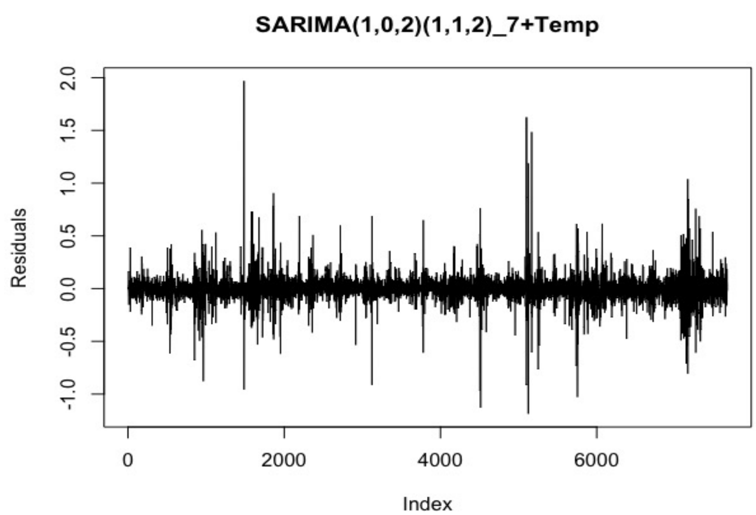

Figure 24: Residual plot for model 4. The plot exhibit spiky behavior, which indicates that some effect is not taken into account during the modeling work. This type of spikes and volatility clustering often is a sign of GARCH effects.