UNIVERSITY OF LJUBLJANA

FACULTY OF ECONOMICS

MASTER'S THESIS

EMPIRICAL MODELLING OF ELECTRICITY SPOT PRICES AND THEIR VOLATILITIES IN SLOVENIA

Ljubljana, September 2014

SAŠA SAJE WANG

AUTHORSHIP STATEMENT

The undersigned Saša Saje Wang, a student at the University of Ljubljana, Faculty of Economics, (hereafter: FELU), declare that I am the author of the master's thesis entitled EMPIRICAL MODELLING OF ELECTRICITY SPOT PRICES AND THEIR VOLATILITIES IN SLOVENIA, written under supervision of dr. Igor Masten.

In accordance with the Copyright and Related Rights Act (Official Gazette of the Republic of Slovenia, Nr. 21/1995 with changes and amendments) I allow the text of my master's thesis be published on the FELU website.

I further declare

- the text of my master's thesis to be based on the results of my own research;
- the text of my master's thesis to be language-edited and technically in adherence with the FELU's Technical Guidelines for Written Works which means that I
    - cited and / or quoted works and opinions of other authors in my master's thesis in accordance with the FELU's Technical Guidelines for Written Works and
    - obtained (and referred to in my master's thesis) all the necessary permits to use the works of other authors which are entirely (in written or graphical form) used in my text;
- to be aware of the fact that plagiarism (in written or graphical form) is a criminal offence and can be prosecuted in accordance with the Criminal Code (Official Gazette of the Republic of Slovenia, Nr. 55/2008 with changes and amendments);
- to be aware of the consequences a proven plagiarism charge based on the submitted master's thesis could have for my status at the FELU in accordance with the relevant FELU Rules on Bachelor Thesis / Master's Thesis / Doctoral Dissertation.

Ljubljana, September 23rd, 2014 Author's signature:

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

INTRODUCTION

Traditionally, electricity prices were relatively stable with only limited if any volatility in the determination of their prices. Regulated by national governments, electricity prices were largely determined by their cost of production. However, since the 1990s many countries have started to liberalise their energy sectors in order to be more dynamic, where by restructuring and deregulation governments allowed electricity prices to be determined by market forces of supply and demand. With the introduction of wholesale electricity markets, electricity delivery contracts, as well as electricity forwards and other derivatives, are now traded on electricity exchanges or bilaterally over the counter to various market participants as an actively traded commodity. The oldest electricity exchange Noordpool in Scandinavia was established in 1993 and since then many countries have followed with deregulation and reformed their electricity sectors.

According to Deng & Oren (2006, p. 3), the new policy is supported with the argument that a competitive electricity sector can restructure long-term investments in generating capacity more efficiently than under state ownership, where utility companies were allowed to earn a regulated rate of return above their cost of capital. When authorities confirmed the construction of power plants, the costs were passed to consumers for the duration of the investment through regulated electricity prices, independent of the fluctuations in market value of investment over time due to the changes in energy prices, improving technology, and evolving conditions of supply and demand. This meant that the heavy burden of investment risk in generating capacity was allocated to consumers rather than producers. Plants generating electricity had few incentives to optimize excessive costs and were more concerned with improving and maintaining their services, rather than with adjusting to market demands and with adopting and developing to new advancements in generating electricity. Liberalisation of the electricity industry thus transferred most of the investment risk from consumers to producers. This led to increased trading and consequent uncertainty and high volatility in price determination for both the producers and distributers of electricity that trade on electricity exchanges. For such companies that need to be more prudent in the way how they manage their risks toward large movements in electricity prices, new methods for valuation of electricity prices had to be developed.

Demand for such valuations and optionalities arises naturally from the unpredictability of power consumption and from the optionalities inherent in power plants (Burger, Klar, Müller & Schindlmayr, 2004, p. 109). Before deregulation, there rarely was a necessity to precisely forecast and evaluate electricity prices and electricity derivatives because electricity was not a commodity that could easily be traded and because its supply was provided by regulated utility companies. Market participants did not use the flexibility of delivery contracts in a market-oriented way. With liberalisation and the establishment of electricity exchanges, this changed and online trading platforms for electricity contracts have been founded. Market counterparts can now take advantage of the optionality in their contracts by optimizing against market prices and looking for arbitrage opportunities. Therefore, it has become a significant task for participants to develop new pricing models for the contracts they buy and sell and to quantify and manage involved risks (Burger et al., 2004, p. 109).

Along with fixed contracts (forwards, futures) and other derivatives from financial markets that protect market participants against high and low prices, price forecasting became an extremely valuable tool for producers of electricity and companies that trade in energy markets. The pricing of electricity is complicated due to its unique characteristics, as compared to other financial assets. One of the main differences is that the electricity has to be produced in the same quantity as it is consumed in real time. The non-storability of electricity and its unique physical attributes of production and distribution thus make electricity a highly unusual commodity. In other words, the demand for electricity at different times can occasionally cause very large fluctuations in the spot price or distinctive price spikes reflected in the fat tails of the distribution related to instantaneous supply and demand. Given these features, electricity markets are difficult to predict mainly due to their high volatility, significantly affected by supply and demand on the market as well as other external factors, such as weather, seasonality over days, weeks or months, transportation constraints, prices of fossil fuels, etc. In this regard, accurate price forecasts are necessary in order to develop effective bidding strategies to maximize profits.

The purpose of this thesis is to present the dynamics of modern electricity markets. To give a detailed theoretical explanation of the models used for time series modelling in finance, and to empirically examine whether the proposed models can sufficiently explain the conditional variance and future prices of electricity. We propose and examine econometric models based on the autoregressive moving average process (hereafter: ARMA) and stochastic volatility processes (based on general autoregressive conditional heteroscedasticity models, hereafter GARCH) to simulate estimations of future electricity prices and their variances and to further check the goodness-of-fit of forecasts in comparison to their original de-seasonalised prices. The fundamental hypothesis is to verify whether the forecasted ARMA-GARCH models can sufficiently explain price fluctuations in electricity price time series.

The thesis will be made of two parts, theoretical and empirical. For the purpose of the theoretical part, different econometric models from the field of financial time series modelling are presented. This part is mainly descriptive with specifications of the models used in the second part of the thesis. In the second part, the data of electricity price time series from BSP South Pool, a Slovenian electricity exchange platform are collected, while in the following steps the model is identified with the help of econometric software (Eviews, Matlab); In other words, to determine the order of the model required to capture the dynamics of the data, specifically the lags of ARMA model using autocorrelation and partial autocorrelation functions, along with different information criteria functions. The following steps involve the estimation of parameters and model validation or diagnostic checking whether the specified model is adequate (statistical significance and hypotheses tests on the estimated parameters and residuals). If these are not validated, we return to a different model identified in the identification step and estimate and check new model parameters again. When the statistical significance and hypotheses tests on the estimated parameters are validated, forecasting prices can begin. In the final step of forecasting, whether the forecasts are accurate or not is determined using different statistics to measure forecasting errors.

1 ELECTRICITY MARKETS

1.1 General features of electricity markets

There are distinctive differences between electricity and financial markets due to the nature of electricity as a commodity. Financial models for forecasting and pricing derivatives were primarily designed for stocks and bonds and, as such, do not capture the unique physical properties of electricity. Mainly, these are the non-storability (except, to a certain extent, for hydroelectricity and limited storage in generators), the seasonality and price spikes, the difficulties of transportation as well as the necessity for the European Community to define clear rules for cross-border electricity transmission (Geman, 2002 pp. 3-4). Due to these constraints, electricity is treated as a flow commodity, for which absence of arbitrage remains the fundamental principle on which the pricing of electricity and its financial products are based (Kluge, 2006. p. 3). If the relation across time and space provided by arbitrage broke down, spot prices would be expected to be highly dependent on temporal and local supply and demand conditions, as well as affecting the relationship between electricity spot and derivative prices (Naeem, 2010, p. 3).

According to Karaktsani & Bunn (2004, p. 2), induced by physical constraints and perhaps also by generator’s strategic behaviour, irregular volatility occurs frequently during short periods and not only over periods of higher frequency. This can be seen empirically in different markets, where volatility presents significant problems for pricing, but may also suggest profitable strategies to those who understand and are able to efficiently anticipate its complexity. Due to such features and constraints of electricity markets, conventional financial models cannot be applied directly to electricity markets, but adaptations and modifications have to be made (Kluge, 2006, p. 3). Nevertheless, stochastic volatility models are fundamental for trading, production scheduling, derivatives pricing, capacity investment and generation asset valuation (Karaktsani & Bunn, 2004, p. 2).

Kluge (2006, pp. 3-4) describes the main differences and similarities between electricity as a commodity and conventional share markets as:

- Underlying unit: In conventional financial markets, the underlying unit (a stock or a bond) is a share of a certain company, whereas in electricity markets it is a specific unit of energy specified in MWh. Theoretically speaking, if we could omit the storability constraint and electricity could be stored in reservoirs or storages, the units of electricity would be bought like any other commodity and would only involve an electronic financial transaction and an assignment of the purchased energy units into the buyer's portfolio without actual physical delivery (Kluge, 2006, p.3).

- Production and consumption: In share markets, the number of shares remains constant over time (unless the company issues new shares) and gives the owner codetermination rights. However, electricity is produced and consumed in real time, which (even with a hypothetical ability to store) has a profound effect on the price per unit. Based on microeconomic considerations, we can expect the long-term price to revert to the production cost, which is one reason mean reverting models are mostly used in commodity markets (Kluge, 2006, p. 3).

- Inability to store: To understand the behaviour of electricity prices, it must be noted that electricity is practically very difficult to store with current technology. It is virtually impossible to store the amount of electricity a large factory consumes on a single day, let alone the energy of an entire country (Kluge, 2006, p. 4). In most countries, there are only a few reservoir power plants. The non-storability of electricity can thus hardly be ignored. One of the consequences is that the relation between spot prices and futures prices cannot be described by the cost of carry. Eydeland & Geman (1999, p.8) suggest that the non-storability of electricity indicates a breakdown of deriving the basic properties of spot prices from the analysis of forward curves. The most evident consequence of this is enormous price fluctuations and price spikes, which have far-reaching consequences; specifically, electricity, as a pure flow commodity, requires time to transfer a certain amount of energy due to constraints in the transmission grids that can cause congestion.

Therefore, contracts always specify a delivery period. Simultaneously, production and consumption have to constantly be in balance. A small imbalance can be absorbed in voltage changes and, for supply excess, dissipation in the grid and generating plants. Supply dropping below the demand could result in a black out. This real time balance of supply and demand introduces seasonality, related to cyclical fluctuations of demand by the hour in the day, week or month. Additionally, the inelasticity of supply and demand (end users, households and firms receive electricity at a fixed price and do not restrict their needs if electricity prices increase on the exchange) makes electricity highly vulnerable to extreme events such as power outages. In such cases, the maximum supply could drop to levels near the current demand, causing the price to rise considerably. This can be resolved quickly, when the outage is resolved or when spare power station is activated, thus normalizing the situation and bringing the price down to previous levels. Such extreme price occurrences are called spikes. Another aspect of balancing the production and consumption in electricity is that spot prices in wholesale markets can theoretically take any value (even any negative value); according to Branger, Reichmann & Wobben (2010, p. 54), negative electricity prices make technical and economic sense, and can include important incentive signals for load shifting. They allow power plants to pay to consumers when electricity demand is very low, while production is relatively high at the same point in time. Consequently, paying consumers to consume electricity is cheaper than temporarily shutting down production.

1.1.1 Characteristics of electricity markets

Along with abovementioned unique features of electricity, the most important properties of electricity markets observed in all electricity markets are:

- Mean reversion: Toward an equilibrium level in the long run. According to Burger et al. (2004, p. 3), the time for electricity price to revert to its mean level has a magnitude of days or at most weeks and can be explained with the recoveries from power station outages and changes in weather conditions. Electricity prices, like other commodity prices, have a strong tendency to revert quickly to a mean level. Geman (2002, p. 4), reports that Pindyck (1999) analysed a 127-year time series for crude oil and bituminous coal and a 75-year time series for natural gas, and concludes that both indicate mean reversion towards a stochastically fluctuating trend line. The same pattern can be observed in electricity prices for which Eydeland & Geman (1999) analysed several regions of the US. The same is observed also for the Slovenian time series of electricity prices, as can be seen in Figure 3.

- Price spikes: As mentioned earlier, one important aspect of electricity prices is the presence of spikes, specifically of sudden upward jumps followed by a quick reversion to normal levels. This happens as the imbalance of supply and demand is impossible to correct in a very short time span. As such, spikes present difficulties for modelling any underlying unit as they are not consistent with the usual modelling based on normal distribution. However, according to Eydeland and Wolyniec (2003, p. 5), spikes can be useful in two other ways. On the demand side, they can help signal shortages and encourage customers to reduce their usage in times of stress, while on the supply side, they can signal shortages and help bring in more supply.

- Seasonality: Seasonal patterns and periodicities in electricity series are present in every electricity market and can be easily observed in autocorrelation functions of electricity price data.[1] Such patterns occur daily, weekly or monthly and are mostly the consequence of daily business activities, as well as weather conditions. For the purpose of this research, we have taken into account the 24-hour daily and 168-hour weekly periodic seasonality because we are the modelling the hourly data.

- High volatility: Another specific feature of electricity prices is very high volatility uncommon in other financial assets and commodities. It is not uncommon that a standard deviation can be six sigma or higher in a relatively small sample of a few thousand observations. The high volatility is another consequence of non-storability and the limited

[1] Figures 6 and 7 on page 34 show the daily seasonal patterns in autocorrelation and partial autocorrelation functions in the Slovenian hourly electricity time series for the period of the first quarter of 2014.

potential to store electric power efficiently, while setting the equilibrium prices in real time is one of the requirements of the market.

- Regionality or locality of electricity prices: Electricity markets are geographically distinct, with regions through which transmission of electricity is physically impossible or economically inefficient (Geman, 2002, p. 3). There are also significant price differences across different regions, which is another aspect of electricity prices that differs from financial assets. These differences are the result of different generating capacities and differences in transmission grids within regions in a country or internationally as some regions or countries produce surpluses and others shortages of electricity.

- Non-stationarity: Electricity price data exhibits strong non-stationarity, a consequence of seasonal patterns that distort the time series. Burger et al. (2004, p. 4) suggests that, for this reason, conventional stochastic volatility models might be inappropriate and suggest modelling with non-stationary models.

1.2 Spot markets

In this thesis, we focus on day-ahead bidding prices or short-term deliveries of electricity; an overview of spot markets, and their price determination follow. Electricity spot prices are essentially the prices of electricity for short-term physical delivery within wholesale markets, most commonly associated with the day-ahead bidding of market participants. Electricity prices in such markets are determined by the balance of supply and demand and therefore signal when there are problems in the grid if the price increases. This aids in detecting where production is inadequate as consumption exceeds production. As already stated, such factors as weather or generators not producing to their full capacity can also impact how much electricity can be transported through the grid and consequently affect the price. This is known as 'transmission capacity' or, in other words, the volume of electricity that can be transported through the grid. According to Branger et al. (2010, p. 53), the supply and demand curve in these markets is given by the merit order model, in which all available producers offer electricity in an increasing series of their variable costs, so that the market-clearing price corresponds to the variable costs of the marginal power plant in the merit order to satisfy the demand.

In deregulated power markets, electricity exchanges such as Nordpool, EEX, EPEX, BSP Southpool, etc., provide market participants a place for power plants, TSOs (transmission system operating companies), traders, suppliers or retailers of electricity, energy companies and other financial institutions to buy or sell physical electricity. Electricity needs to be transferred to its destination through transmission grids, which are operated in each market by TSOs.[2] According to Kluge (2006, p. 5), this part of the market is monopolistic and tariffs are set by regulators; prices should reflect the maintenance cost and the energy loss, as it is the TSO's liability to buy the electricity lost through the transmission from the spot market. This guarantees that each market participant submits

[2] ELES in Slovenia.

or receives the exact amount of electricity as specified in the spot market contract. Therefore, the total contract cost is composed of the spot price, trading fees and transmission charges.

Spot market contracts are usually specified as day-ahead auctioned hourly bids for the physical delivery of a specified amount of units that each market participant buys or sells.

1.2.1 BSP Southpool spot price determination

Trading on the BSP Southpool electricity exchange can be done for day-ahead auction trading, or continuously during each day (intra-day bidding in real time) by companies that meet the requirements set by BSP Southpool and with access to the transmission grids.

For the day-ahead bidding, market participants can submit anonymous standardized hourly products on the EuroMarket trading platform (electricity prices per MWh are determined using the bidding system), where contracts are made between sellers and buyers for the physical delivery of the electricity the following day after the price is set (BSP website). The day-ahead market is driven by market participant’s forecasts, planning and expectations, meaning that buyers have to estimate how much electricity they require to meet the demands on the following day as well as how much they are willing to pay for it for each hour. The sellers have to estimate how much and at what price will they deliver electricity for each hour of the following day.

According to BSP Southpool, auction trading is divided into the following phases:

- The call phase: call phase runs each day until 9:40 a.m., when BSP Southpool closes the bidding for electricity that will be delivered the next day.

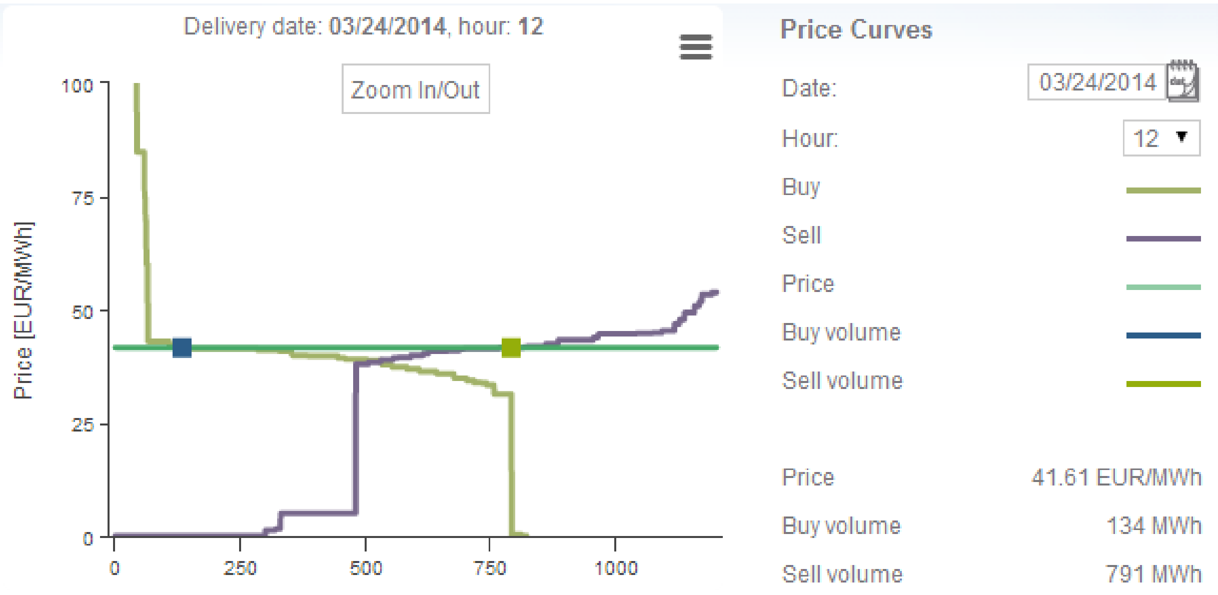
- The freeze phase: runs from 9:40 a.m. to 9:50 a.m., during which the market supervisor can examine the orders and react in case of any irregularities.

- Price determination phase: runs from 9:50 a.m. to 10.30. a.m., when the marginal spot price is calculated for each individual hour by the intersection of the aggregate supply and demand curve (also known as an equilibrium point).

After 9:40 a.m., when the call phase for market participants ends, all bids and offers are aggregated into two curves for each delivery hour, forming an aggregate supply and demand curve. The price for each hour is determined by the intersection of those curves, which represent all bids and offers in the market. Once the marginal market prices are calculated, trades are settled and physically delivered from 00:00 a.m. onward to buyers for each hour of the following day. Based on this price, shown on Figure 1, it is clear how many units all members buy or sell. The average daily price is further calculated as the base load price, which is split into the peak price (denoting the hours of the day with the highest consumption, 08:00–20:00) and the off-peak price (for the remaining hours).

Another type of bidding on the BSP Southpool can be done continuously in real time (one day before the delivery day and up to 60 minutes prior to product expiration on delivery day), where members can adjust to a change of supply and demand during the day. Here, prices are determined on the basis of the price/time priority criterion. Although the majority of trades are made for the day-ahead bids, intra-day bidding is offered to help secure and adjust to a necessary balance between supply and demand on the market (BSP website).

Figure 1. Supply and demand curves for electricity on March 24th from 11 a.m. to 12 a.m. on BSP Southpool day-ahead market.



Source: *Day ahead trading results*, 2014.

2 UNIVARIATE TIME SERIES MODELLING

In this and the next chapter, we present characteristics of various linear and stochastic processes used empirically in Chapter 4, beginning with models for univariate time series modelling and forecasting and continuing with conditional volatility models in Chapter 3.

A time series is a collection of sequential data points, measured at successive points in time, spaced at uniform time intervals. Brooks (2008, p. 206) defined it as ‘univariate time series models are a class of specifications where one attempts to model and to predict financial variables using only information contained in their own past values and possibly current and past values of other (explanatory) variables.’ Macroeconomic models from the 1960s were often composed of simultaneous equations that frequently had poorer forecasting performance than fairly simple univariate models based on only a few parameters and using only past data. Therefore, time series models are used when a more complex structural model is inappropriate. This led George Box and Gwilym

Jenkins to propose autoregressive integrated moving average (ARIMA) models in the 1970s (Greene, 2003, p. 608). These are mathematical models of time series that aim to find the best fit of a forecast to the given past values of a time series. Currently, time series analysis is used in different fields as diverse as economics, statistics, finance, weather forecasting, seismology, pattern recognition, astronomy etc. In the following sections, we specify the notations and present important concepts for modelling ARMA models.

2.1 Important concepts and notations in time series

Time series data, as the name suggests, are data collected over a period of time on one or more variables, where the data are associated with a particular frequency of observation (hourly, daily, weekly, monthly, etc.). In the following sections, we denote y_t as an observation of an asset at time t, in our case the price of electricity at time t. The preceding observation is written as y_{t-1} and succeeding as y_{t+1}.

The following paragraphs are dedicated to several important concepts in time series modelling. The first of these is the notion of whether the series is stationary or not. This is essential to time series analysis, as the stationarity of a series strongly characterizes its behaviour and properties.

2.1.1 White noise process

A white noise process represents a stationary time series or a stationary random process with zero autocorrelation. In an autoregressive model, white noise is the disturbance term or the source of randomness, denoted as ε_t. A process is white noise if it satisfies the following conditions:

$$E(\varepsilon_t) = 0 \tag{1}$$

$$E(\varepsilon_t^2) = \sigma^2 \tag{2}$$

$$Cov(\varepsilon_t, \varepsilon_s) = 0 \text{ for all } s \neq t. \tag{3}$$

Therefore, a white noise process has a constant mean and variance and zero covariance, except at lag zero.

2.1.2 Stationarity

A stationary process is a stochastic process whose joint probability distribution does not change when shifted in time. Therefore, the process has a constant mean, variance and autocorrelation over time for each given lag and does not follow any trends. According to Brooks (2008, pp. 318-320), the stationarity of a series is essential for the following reasons:

- The stationarity of a series can strongly influence its behaviour and properties. For example a ‘shock’ is often used to denote a change in a variable, while for a stationary series, the shocks gradually die away (a shock during time t will have a smaller effect in $t+1$, a still smaller effect in $t+2$, and so on). In contrast, for a non-stationary series, the persistence of a shock will always be infinite (the effect of a shock during time t will not have a smaller effect in time $t+1$, in time $t+2$, etc.).

- The use of non-stationarity data can lead to spurious regressions. If two stationary variables are generated as independent random series, when one of those variables is regressed on the other, the *t*-ratio on the slope coefficient would not be expected to be significantly different from zero, and the value of R^2 [3] would be expected to be very low. This seems obvious for the variables that are not related to one another. However, if two variables are trending over time, a regression of one on the other could have a high R^2 even if the two are totally unrelated. Therefore, if standard regression techniques are applied to non-stationary data, the end result could be a regression that ‘looks’ good under standard measures (significant coefficient estimates and a high R^2), but which is actually valueless.

- If the variables employed in the regression model are not stationary, then it can be proved that the standard assumptions for asymptotic analysis will not be valid. In other words, the usual ‘*t*-ratios’ will not follow a *t*-distribution, and the *F*-statistic will not follow an *F*-distribution, and so on.

There are two forms of stationarity: strict stationarity and its weaker form, both presented below. However, for the purpose of our empirical analysis, we refer to weakly stationarity or second order stationarity.

2.1.2.1 Strictly stationary process

According to Brooks (2008, p. 207), a strictly stationary process is one where, for any $t_1, t_2, ..., t_T \in Z$, any $k \in Z$ and $T = 1, 2, ...$ (T is an arbitrary positive integer and $t_1, t_2, ..., t_T$ is a collection of T positive integers) states that,

$$Fy_{t_1}, y_{t_2}, ..., y_{t_T}(y_1, ..., y_T) = Fy_{t_{1+k}}, y_{t_{2+k}}, ..., y_{t_{T+k}}(y_1, ..., y_T)\,. \tag{4}$$

Where F denotes the joint distribution function of the set of random variables. Thus, a time series is strictly stationary if the joint distribution of ($y_{t_1}, y_{t_2}, ..., y_{t_T}$) is identical to that of ($y_{t_{1+k}}, y_{t_{2+k}}, ..., y_{t_{T+k}}$) for all t. We can say that the probability of y_t is the same as that for y_{t+k}, for all values of k. In other words, a series is strictly stationary if the distribution of its values remains the same as the time progresses, implying that the probability that y falls within a particular interval is the same now as at any time in the

[3] R^2 is the coefficient of determination. It indicates how well the data fit the proposed statistical model.

past or future (Brooks, 2008, p. 208). This is, however, a very strong condition, and a weaker version of stationarity is often assumed.

2.1.2.2 A weak stationary process

The weak stationarity implies that the time plot of data would show that the T values fluctuate with constant variation around a constant level (Tsay, 2008, p. 23). A time series is said to be weakly stationary if for all $t = 1, 2, ..., \infty$, it satisfies the following conditions:

- the series has a constant mean ($E(y_t) = \mu$),

- constant variance ($E(y_t - \mu)(y_t - \mu) = \sigma^2 < \infty$)

- and constant autocovariance structure ($E(y_{t_1} - \mu)(y_{t_2} - \mu) = \gamma_{t_2 - t_1} \forall t_1, t_2$)

where E is the expectation operator. Autocovariance determines how y is linked to its previous values, where for a stationary series they depend only on the difference between t_1 and t_2 , so that the covariance between y_t and y_{t-1} is the same as between y_{t-1} and y_{t-2} , y_{t-2} and y_{t-3} , and so on. If the first condition ($E(y_t) = \mu$) is satisfied then the autocovariance is given by

$$\gamma_s = E(y_t - E(y_t))(y_{t-s} - E(y_{t-s})), s = 0, 1, 2, ... \tag{5}$$

2.1.3 Autocorrelation and partial autocorrelation

In identifying the stochastic process, we are interested in the connection between two random variables of a process at different points in time. One way to measure the linear dynamic of a series and determine the order of p and q is with the use of autocorrelation (acf) and partial autocorrelation function (pacf), as described below. To identify the most appropriate model, we used sample acf and sample pacf to check whether they are similar to theoretical properties of a stationary series. The second approach uses different information criteria and is explained in detail under the identification of ARMA models.

2.1.3.1 Autocorrelation

When normalising autocovariance by diving it by variance σ^2, and assuming the above conditions, we obtain a correlation coefficient between times t and s , commonly denoted as ρ_s which can be expressed as

$$\rho_s = \frac{E(y_t - E(y_t))(y_{t-s} - E(y_{t-s}))}{\sqrt{E(y_t - \mu)^2 E(y_{t-s} - \mu)^2}} = \frac{Cov(y_t, y_{t-s})}{Var(y_t)} = \frac{\gamma_s}{\gamma_o} \tag{6}$$

Dividing autocovariance with variance puts the range of autocorrelation within the range of ± 1. By the definition, the autocorrelation at lag 0 is always $\rho_0 = 0$. If ρ_s is plotted against $s = 0, 1, 2, ...$ a graph known as autocorrelation function (acf) is obtained (used to capture the linear dynamics of the data). Autocorrelation function is essentially a set of correlation coefficients between the series and lags of itself over time. In the above equation (6), if ρ_s decreases to zero quickly the series is said to be stationary. If the series has a trend, ρ_s will decline toward zero slowly. If the series has a seasonal pattern, like in the hourly electricity data, ρ_s will be significantly different from zero at lag 24 (daily seasonal pattern), lag 168 (weekly seasonal pattern), etc. In the case of daily seasonality for electricity prices, this can be observed in Figure 6.

2.1.3.2 Partial autocorelation

The partial autocorrelation function (pacf) is the partial correlation of coefficients between different series and lags of itself over time. In other words, pacf measures the correlation between an observation k periods ago and the current observation, after controlling for observations at intermediate lags (all lags $< k$), i.e. the correlation between y_t and y_{t-k}, after removing the effects of $y_{t-k+1}, y_{t-k+2}, ..., y_{t-1}$. For example, the pacf for lag 3 would measure the correlation between y_t and y_{t-3}, after controlling for the effects of y_{t-1} and y_{t-2} (Brooks, 2008, p. 222). The coefficients of acf and pacf at lag 1 are always equal as there is no intermediate lag effect to eliminate; pacf is denoted as ρ_{ss}. For example, the pacf at lag 2 can be expressed as:

$$\rho_{22} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

where ρ_1 is the autocorrelation at lag 1 and ρ_2 is the autocorrelation at lag 2. A more general form of pacf can be written as:

$$\rho_{ss} = \frac{\rho_s - \rho_{s-1}^2}{1 - \rho_{s-1}^2} \tag{7}$$

2.2 Autoregressive model

A model that depends on only previous values of itself is called an autoregressive model. In the time series literature, an autoregressive model of order 1 denoted as AR(1) is a model in which the dependent variable y_t (in our case the price of electricity) depends only upon its previous value y_{t-1} (explanatory variables) and an error term. The term 'autoregressive' comes from the fact that y_t is regressed on its lagged values. The basic first order autoregressive model, denoted as AR(1) can be written as:

$$y_t = \mu + \phi y_{t-1} + \varepsilon_t \tag{8}$$

The variable y_t is said to be autoregressive (or self regressive) because under certain assumptions it holds that:

$$E[y_t \mid y_{t-1}] = \mu + \phi y_{t-1}.$$

If y_t has a statistically significant lag 1 autocorrelation coefficient, it indicates that the lagged observation y_{t-1} might be useful in forecasting y_t. A more general pth-order autoregression or autoregressive AR(p) process can be expressed as:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \varepsilon_t \tag{9}$$

Where y is the series and p is the order of the process, representing the number of parameters that need to be estimated. μ is a constant, while $\phi_1 ... \phi_p$ are the parameters of the model that describe the effects of a unit change between consecutive observations (y_{t-1} on y_t, y_{t-2} on y_{t-1} and so on). ε_t is a white noise disturbance term, assumed to be normally and identically distributed with no significant autocorrelation, a zero mean and a variance constant over time $(\varepsilon_t \sim N(0, \sigma^2))$. For the purposes of this thesis, the above model (6) can be written more compactly using sigma notation as

$$y_t = \mu + \sum_{i=1}^{p} \phi_i y_{t-1} + \varepsilon_t \,. \tag{10}$$

2.3 Moving average model

Another class of time series models used in finance for predicting future values is known as the moving-average process. Similarly to the AR model, the effect is to represent y_t as a function of its own past values. Like in the AR model, let ε_t be a white noise process with $(\varepsilon_t \sim N(0, \sigma^2))$. The series, denoted as MA(q), can be expressed as

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + ... + \theta_q \varepsilon_{t-q} \tag{11}$$

Where μ is a constant, $\varepsilon_t ... \varepsilon_{t-q}$ are white noise disturbance terms sometimes also labelled as innovations in the model. q represents the order of the moving-average mode, while $\theta_1 ... \theta_q$ are moving-average parameters that need to be estimated in order to describe the effects of past errors. Using sigma notation, it can be expressed as

$$y_t = \mu + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \,. \tag{12}$$

Put simply, MA is a linear combination of white noise processes, where y_t depends on the current and previous values of a white noise disturbance terms or, in other words, that y_t is a weighted average of the current and past shocks $\varepsilon_t ... \varepsilon_{t-i}$.

2.4 Transforming non-stationary data

As stationarity is a desired property of an estimated AR and MA models, the time series often requires us to transform a non-stationary series into a stationary one. This can be achieved using a logarithmic transformation, square root transformation or by de-trending by differencing a series into a stationary one. Time series with seasonal components like the electricity price data can also be de-seasonalised using additive decomposition to remove the seasonal components and thus to smooth a series into a stationary one. With regards to the electricity time series, y_t is a sum of three independent components, namely: the seasonal S_t, the trend T_t, and the stochastic irregular I_t, where each of three components have the same units as the original series:

$$y_t = S_t + T_t + I_t. \tag{13}$$

A seasonally adjusted series can be obtained by estimating and removing the seasonal effect from the original time series. By denoting the estimated seasonal component as $\hat{S}_t$, the seasonally adjusted series can then be expressed as:

$$SA_t = y_t - \hat{S}_t = T_t + I_t. \tag{14}$$

The difference between the original and the deseasonalised series can be observed by comparing Figures 6 (original series) and 9 (daily and weekly deseasonalised series).

2.5 Testing the stationarity of a series

Because of the reasons for stationarity mentioned earlier, we have to perform a series of statistical tests to examine whether a unit root[4] is present in the autoregressive model. The pioneering work on testing for stationarity was developed by David Dickey and Wayne Fuller in 1979. Their test checks the null hypothesis that $\phi = 1$(the process contains a unit root) against the one-sided alternative that ϕ<1 (the process is stationary) in the (AR) model such as:

$$y_t = \phi y_{t-1} + \varepsilon_t. \tag{15}$$

Therefore the test hypotheses are:

H_0: series contains a unit root

H_1: series is stationary.

According to Brooks (2008, p. 327), for the ease of computation and interpretation, the following regression is often employed instead of (15):

$$\Delta y_t = \psi y_{t-1} + \varepsilon_t \quad . \tag{16}$$

So that a test of $\phi = 1$is equivalent to a test of $\psi = 0$(since $\phi - 1 = \psi$). The test statistic for the original DF test is then defined as:

[4] A process has a unit root if 1 is a root of the process's characteristic equation.

$$DF = \frac{\hat{\psi}}{\hat{SE}(\hat{\psi})} \quad (17)$$

Where $\hat{SE}$ denotes the estimate of the standard error. The Dickey-Fuller test statistic does not follow the usual t-distribution under the null hypotheses but rather follows a non-normal distribution, where critical values were derived from simulations in various experiments. When the test statistic is more negative than the critical value, the null hypotheses is rejected in favour of the alternative hypotheses that the series is stationary. The test is however valid only if ε_t is white noise. According to Brooks (2008, p. 328), 'ε_t is assumed not to be autocorrelated, but would be so if there was autocorrelation in the dependent variable of the regression Δy_t which has not been modelled. If this is the case, the test would be "oversized", meaning that the true size of the test would be higher than the nominal size used. The solution is to "augment" the test using p lags of the dependent variable.' The augmented Dickey-Fuller (ADF) test can now be expressed as:

$$\Delta y_t = \psi y_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta y_{t-i} + \varepsilon_t \quad (18)$$

ADF is still conducted on ψ, and relies on the same critical values from the DF tables. Another type of unit root test was developed by Peter Phillips and Pierre Perron, known as the Phillips-Perron test. Like the ADF test, it incorporates an automatic correction to the DF test to allow for autocorrelated residuals. Both the ADF test and PP test, however, suffer for the same limitation, i.e. their power is low if the process is stationary, but with a root close to non-stationarity boundary. In other words, the process is rejected as non-stationary when the coefficient under hypotheses ϕ is close to 1, especially when the sample size is small. To overcome this problem of ADF and PP tests for which the data appear as stationary by default if there is little information in the sample, there is another unit root test developed by Kwiatkowski et al., known as the KPSS test. This test is not discussed here, as our sample size is big enough; however, we employ it in the empirical part to compare its results with the ADF and PP procedures, to see if the same conclusions on stationarity is obtained. The test statistics employ an opposite null hypotheses to the above tests, with the null hypotheses as the stationary process. Specifically:

H_0: series is stationary

H_1: series contains a unit root.

2.6 Autoregressive moving average model

The application of above AR(p) and MA(q) models can sometimes become quite complicated because a high order model with many parameters may be needed to adequately describe the dynamic structure of the data. To alleviate this difficulty, George E.P. Box and Gwilyn Jenkins suggested combining the AR(p) and MA(q) models in order to construct an autoregressive moving-average (ARMA(p,q)) model. Such modifications with a relatively low values of p and q have proven quite successful in forecasting models (Greene, 2003, p. 610). According to Tsay (2002, p. 48), an ARMA model combines the concepts of AR and MA models into a compact form, so that the number of parameters used is kept small. The joint model can be written as

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + ... + \theta_q \varepsilon_{t-q} + \varepsilon_t \,, \tag{19}$$

or more compactly as

$$y_t = \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t \,. \tag{20}$$

According to Greene (2003, p. 619), ‘there is no underlying economic theory that sates why a compact ARMA(p,q) representation should adequately describe the movement of a given economic time series. Nonetheless, as a methodology for building forecasting models, this set of tools and its empirical counterpart have proved as good as and even superior to much more elaborate specifications.’ The first to approach to estimate and forecast such ARMA models in a systematic manner were Box and Jenkins (1976). Their approach to modelling a stochastic process is considered to be practical and pragmatic, and consists of the following steps:

1. Identification

2. Estimation

3. Model Checking

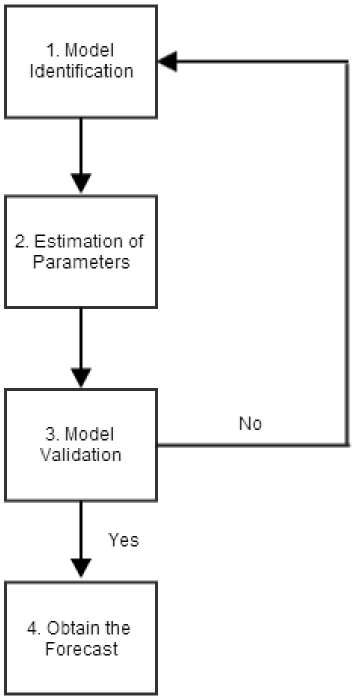
4. Forecasting

Figure 2 shows the flowchart of how we obtain our ARMA model. In the first step, we identify a subset of models based on the observed data. In the second step, after the model is identified, we estimate its parameters, and in the third step validate the model using various diagnostic checks by applying statistical hypothesis testing. If the chosen model is validated, we continue with forecasting; otherwise, we return to Step 1 and refine or identify a new model.

2.6.1 Identification of an ARMA model

The first step involves determining the order of the ARMA(p,q) model to acquire the dynamic features of the series. In application, the order of p and q of an ARMA model is unknown and must be specified empirically. This is referred to as ‘the order determination of ARMA models’. Two general methods exist for determining the orders of p and q. The first method uses graph plots of acf and pacf, while the second approach uses different information criteria functions.

Figure 2. Flowchart of the ARMA methodology



Source: R. C. Garcia, J. Contreras, M. Akeren & J. B. C. Garcia, *A GARCH Forecasting model to Predict Day-Ahead Electricity Prices,* 2005, Figure 1.

2.6.1.1 Using acf and pacf to identify an ARMA model

By plotting the sample acf and pacf, we use the graphical representations to determine the most appropriate specification. However, to identify the appropriate model, the data first have to be stationary. Therefore, we have to first transform the data to obtain stationarity by taking logs, first differences or both, as well as identify the trend and seasonality and deseasonalise the entire series accordingly. Once the stationarity, trend and seasonality have been addressed, we need to identify the order of ARMA lags by using plots of acf and pacf and comparing them to their theoretical characteristics of a stationary series. In the table below, we list the defining characteristics of AR, MA and ARMA process:

Table 1. Defining characteristics of AR, MA and ARMA process

Model	ACF	PACF
AR(p)	geometrically decaying acf	number of non-zero points of pacf = AR order (p)
MA(q)	Number of non-zero points of acf = MA order (q)	geometrically decaying pacf
ARMA(p,q)	geometrically decaying acf	geometrically decaying pacf

Source: C. Brooks, *Introductory econometrics for finance,* 2008, p.225.

2.6.1.2 Information criteria for ARMA model identification

Modelling time series with real data often delivers 'messy' acf and pacf plots, and thus consequently complicates its identification. One way to overcome this subjectivity in the interpretation of acf and pacf is to use information criteria functions. According to Brooks (2008, p. 232), information criteria embody two factors: residual sum of squares (RSS)[5], and some penalty for the loss of degrees of freedom from adding extra parameters. Therefore, adding a new variable or an additional lag to a model will have two competing effects on the information criteria: the RSS will fall but value of the penalty term will increase. The object is to choose the number of parameters that minimises the value of the information criteria. Thus, adding an extra lag will reduce the value of the criteria only if the fall in RSS is sufficient to more than outweigh the increased value of the penalty term (Brooks, 2008, p. 232). Many different information criteria are used to determine the order of the model required, which vary according to how strict the penalty term is. For the purpose of this thesis, we employ Akaike's information criterion (AIC) and Schwarz's Bayesian information criterion. According to Brooks (2008, p. 233), these are defined algebraically as:

$$AIC = \ln(\sigma^2) + \frac{2k}{T} \tag{21}$$

$$SBIC = \ln(\sigma^2) + \frac{k}{T}\ln T \tag{22}$$

Where σ^2 is the residual variance (equivalent to RSS divided by the number of observations T), $k = p + q + 1$ is the total number of parameters estimated and T is the sample size. When using the criteria based on the estimated parameters, the criterion with the lowest parameters should be used. In this thesis, we compute the AIC and SBIC using Eviews software, which derives the test statistic from the log-likelihood function based on maximum likelihood estimation.[6] The corresponding Eviews formulae are:

$$AIC_l = \frac{-2l}{T} + \frac{2k}{T} \tag{23}$$

$$SBIC_l = \frac{-2l}{T} + \frac{k}{T}(\ln T) \tag{24}$$

where

$$l = \frac{-T}{2}(1 + \ln(2\pi) + \ln(\frac{\hat{u}'\hat{u}}{T}))$$

It is difficult to determine which of the above statistics delivers better results, as SBIC is strongly consistent but inefficient as it embodies a much stricter penalty term than AIC,

[5] RSS is the measure of discrepancy between the data and the estimated model.
[6] see estimation of GARCH models

while AIC is not consistent, but is generally more efficient. In other words, SBIC asymptotically delivers the correct model order, while AIC generally delivers too large models. To correctly determine the correct models, we further check the acf and pacf of the selected models by 'trial and error' of different models identified by AIC and SBIC, using the Ljung Box statistics described in the next section.

2.6.2 Estimation and diagnostic checking of ARMA models

The second step in building an ARMA model involves the estimation of parameters of the chosen model, generally by non-linear least squares or by maximum likelihood. In the third step, whether the model specified and estimated is adequate is determined. Specifically, the fitted model must be examined by applying various diagnostics checks to determine whether or not the model applicably represents the data. If the model is adequate, residuals should behave as white noise; if non-stationarity or any other inadequacies are found, the whole process of identification, estimation and diagnostic checking must be repeated with a different model. To test the residual diagnostics, a Ljung Box test with a joint hypothesis that all m lags of autocorrelations ρ_s are all simultaneously equal to zero is employed. In other words, that the data are white noise and thus independently distributed is verified. The Ljung Box test statistic is defined as:

$$Q_{LB} = T(T+2)\sum_{s=1}^{m}\frac{\hat{\rho}_s}{T-s} \tag{25}$$

Where T is the sample size, $\hat{\rho}_s$ is the sample autocorrelation at lag s, and m is the number of lags being tested. The test follows a χ^2 distribution with m degrees of freedom, most commonly at a 5% or 1% significance level. Therefore, the test statistic that joint autocorrelations are zero is rejected at the significance value α if the calculated test statistic Q_{LB} exceeds its critical value $\chi^2_{1-\alpha,m}$.

2.6.3 Forecasting with ARMA models

Forecasts for an ARMA model are generated by calculating the conditional expectations for separate AR(p) and MA(q) processes and then combining those results into a general ARMA(p,q) forecast. The general formula for an ARMA(p,q) forecast at time t for s steps ahead, denoted as $f_{t,s}$ is defined as

$$f_{t,s} = \sum_{i=1}^{p}\phi_i f_{t,s-1} + \sum_{j=1}^{q}\theta_j \varepsilon_{t+s-j} \tag{26}$$

Where $f_{t,s} = y_{t+s}, s \leq 0$ for AR(p) process and $f_{t,s} = \varepsilon_{t+s}, s \leq 0$ for MA(q) process, while $\varepsilon_{t+s} = 0$ when $s > 0$. The forecast produced by an ARMA(p,q) process depends on past and current values of the response as well as the past and current values of the residuals.

In the next section, we briefly discuss the model representations for an AR(p) and MA(q) processes.

2.6.3.1 Forecasting representation of the autoregressive process

To present an AR forecast, suppose that we estimate the following AR(2) model:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t \tag{27}$$

According to Brooks (2008, pp. 249-251), parameter constancy over time is assumed. If this relationship holds for the series y at time t, it is also assumed to hold for y at time $t+1$, $t+2$,…, so that 1 can be added to each of the time subscripts in (27) for one period ahead, 2 for two periods ahead and so on, to arrive at the following:

$$y_{t+1} = \mu + \phi_1 y_t + \phi_2 y_{t-1} + \varepsilon_{t+1} \tag{28}$$

$$y_{t+2} = \mu + \phi_1 y_{t+1} + \phi_2 y_t + \varepsilon_{t+2} \tag{29}$$

Assuming all the information up to and including that at time t is known, we can produce forecasts for y at time $t+1$, $t+2$,…, $t+s$. As past observations up to time t are known, we can produce a forecast by applying the conditional expectations operator to equation (28). We also set the expected residual term to zero ($E(\varepsilon_{t+1}) = 0$), as it is unknown at time t. In the case of the one-step-ahead forecast, this can be expressed as:

$$f_{t,1} = E(y_{t+1|t}) = E(\mu + \phi_1 y_t + \phi_2 y_{t-1} + \varepsilon_{t+1} \mid \Omega_t) \tag{30}$$

Where $E(y_{t+1|t})$ is a short-hand notation for $E(y_{t+1|t} \mid \Omega_t)$. Deriving further, we obtain the one-step-ahead forecast as expressed in (32):

$$f_{t,1} = E(y_{t+1|t}) = \mu + \phi_1 E(y_t \mid t) + \phi_2 E(y_{t-1} \mid t) \tag{31}$$

$$f_{t,1} = E(y_{t+1|t}) = \mu + \phi_1 y_t + \phi_2 y_{t-1} \tag{32}$$

Applying the same process by using conditional expectations for (29) and using the expected forecast for y_{t+1} we can produce a two-step-ahead forecast:

$$f_{t,2} = E(y_{t+2|t}) = E(\mu + \phi_1 y_{t+1} + \phi_2 y_t + \varepsilon_{t+2} \mid \Omega_t) \tag{33}$$

$$f_{t,2} = E(y_{t+2|t}) = \mu + \phi_1 E(y_{t+1} \mid t) + \phi_2 E(y_t \mid t) \tag{34}$$

$$f_{t,2} = E(y_{t+2|t}) = \mu + \phi_1 f_{t,1} + \phi_2 y_t$$

(35) Continuing with the same procedure, we can generate forecasts for 3, 4,…, s-steps ahead:

$$f_{t,3} = E(y_{t+3|t}) = E(\mu + \phi_1 y_{t+2} + \phi_2 y_{t+1} + \varepsilon_{t+3} \mid \Omega_t) \tag{36}$$

$$f_{t,3} = E(y_{t+3|t}) = \mu + \phi_1 E(y_{t+2} \mid t) + \phi_2 E(y_{t+1} \mid t) \tag{37}$$

$$f_{t,3} = E(y_{t+3|t}) = \mu + \phi_1 f_{t,2} + \phi_2 f_{t,1} \tag{38}$$

The general formula for an AR(2) process forecast with s-steps ahead can thus be expressed as:

$$f_{t,s} = \mu + \phi_1 f_{t,s-1} + \phi_2 f_{t,s-2}$$

(40)

2.6.3.2 Forecasting representation of the moving average process

Unlike an autoregressive process, a moving average has a memory only of length q, and this limits its forecast horizon (Brooks, 2008, p. 249). To illustrate this, suppose that an MA(2) model has been estimated:

$$y_t = \mu + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t \tag{41}$$

Again, assuming constancy over time, equation (41) will hold for times $t+1$, $t+2$, and so on.

$$y_{t+1} = \mu + \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-1} + \varepsilon_{t+1} \tag{42}$$

$$y_{t+2} = \mu + \theta_1 \varepsilon_{t+1} + \theta_2 \varepsilon_t + \varepsilon_{t+2} \tag{43}$$

Using conditional expectations for equation (42) and setting $E(\varepsilon_{t+1}) = 0$, we can generate the forecast for one step ahead:

$$f_{t,1} = E(y_{t+1|t}) = E(\mu + \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-1} + \varepsilon_{t+1} \mid \Omega_t) \tag{44}$$

$$f_{t,1} = E(y_{t+1|t}) = \mu + \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-1} \tag{45}$$

Similarly, applying the same rules we can produce the forecast for two steps ahead:

$$f_{t,2} = E(y_{t+2|t}) = E(\mu + \theta_1 \varepsilon_{t+1} + \theta_2 \varepsilon_t + \varepsilon_{t+2} \mid \Omega_t) \tag{46}$$

$$f_{t,2} = E(y_{t+2|t}) = \mu + \theta_2 \varepsilon_t$$

(47)

Like ε_{t+1}, ε_{t+2} is also unknown at time t, thus like for ε_{t+1}, the conditional expectation for ε_{t+2} is also set to zero ($E(\varepsilon_{t+2}) = 0$). Similarly, for 3, 4,…, s-steps ahead, the forecasts are given by:

$$f_{t,3} = E(y_{t+3|t}) = E(\mu + \theta_1 \varepsilon_{t+2} + \theta_2 \varepsilon_{t+1} + \varepsilon_{t+3} \mid \Omega_t) \tag{48}$$

$$f_{t,3} = E(y_{t+3|t}) = \mu \tag{49}$$

$$f_{t,s} = \mu \;\forall\; s \geq 3 \tag{50}$$

As the MA(2) model has a memory of only two periods, all forecasts three or more steps ahead collapse to the intercept. ARMA(p,q) forecasts can now easily be obtained in the same way by applying the rules for AR(p) and MA(q) parts, and using the general formula (26).

2.6.3.3 Checking whether the forecast is accurate

Determining the accuracy of a forecast is important because it clarifies the difference between the actual future value and its forecast. In practice, forecasts are mostly generated for the entire forecasted period; therefore, these values are compared to their actual values, while the difference between them is aggregated. To test the adequacy of forecasts, we use different statistical tests, which measure the forecast error, i.e. a difference between the actual value and the predicted forecast value. Brooks (2008, p. 251) argues that the forecast error defined in this way will be positive if the forecast was too low and vice versa. Therefore, we cannot simply sum the forecast errors, because the positive and negative errors will cancel each one out. For this, we square them or take their absolute values. For the purpose of the thesis, we use the following measures of the forecast error, mean square error (MSE), defined as:

$$MSE = \frac{1}{T-(T_1-1)}\sum_{t=T_1}^{T}(y_{t+s}-f_{t,s})^2 \tag{51}$$

Where T_1 represents the first forecasted observation and all other variables are defined as before. MSE provides a quadratic loss function making it more useful when large forecast errors are disproportionately more serious than smaller forecast errors (Brooks, 2008, p. 252). The next statistic, the mean absolute error (MAE) measures the mean absolute forecast error.

$$MAE = \frac{1}{T-(T_1-1)}\sum_{t=T_1}^{T}\left|y_{t+s}-f_{t,s}\right| \tag{52}$$

The mean absolute percentage error (MAPE) corrects for the problem of asymmetry between the actual and forecast values as well as interpreting the forecast error in percentages, which bounds it from below with 0.

$$MAPE = \frac{100}{T-(T_1-1)}\sum_{t=T_1}^{T}\left|\frac{y_{t+s}-f_{t,s}}{y_{t+s}}\right| \tag{53}$$

3 CONDITIONALLY HETEROSCEDASTIC MODELS

3.1 Characteristics of volatility

Thus far, all the models discussed have been linear in their parameters. Linear models such as ARMA(p,q) have many qualities, namely that their properties are very well understood, that models that appear non-linear, can be transformed linearly, etc. However, according to Brooks (2008, p. 380), it is unlikely that many relationships in finance are intrinsically non-linear. Brooks (2008, p. 380) lists the limitations of the linear time series models such as ARMA(p,q), which are unable to explain the following significant characteristics common to economic and financial data, including:

- Leptokurtosis – predisposition for financial data to exhibit distributions with fat tails and high peaks at the mean.

- Volatility clustering – volatility may be high for certain periods and low for other periods, meaning volatility appears in bursts, so that large returns are expected to follow large returns, and small returns are expected to follow small returns. A likely interpretation for this phenomenon is that the arrival of information regarding price changes appears in bunches rather than being uniformly allocated over time.

- Leverage affects – the tendency for volatility to rise more following a large price fall than following a price rise of the same magnitude.

Numerous kinds of non-linear models exist, though only a few are applicable for modelling financial data. For the purpose of our analysis, we model processes based on the ARCH model and its modifications, such as GARCH, GJR-GARCH and EGARCH, which suggest a form of heteroscedasticity in which the variance of the disturbance depends on the size of the preceding disturbance as well as allow the behaviour of a series to follow different processes at different points in time.

3.2 ARCH model

The first model that offers a systematic context for modelling volatility is the ARCH model proposed by Engle in 1982. ARCH stands for autoregressive conditional heteroscedasticity. Linear ARMA models, as expressed in (20), rely on the basic assumption that the errors are homoscedastic, namely that the errors are normally distributed with zero mean and constant variance. However, this assumption does not hold for many of the financial time series data, where the variance of the errors is heteroscedastic, i.e. the errors are not constant. This is reflected in the flaws of the standard error estimates, when data is modelled as homoscedastic, but the errors are heteroscedastic. To overcome this, it is reasonable to propose a process that does not assume a constant variance and describes how the variance of the errors evolves. To understand how the ARCH model works, we have to first define the conditional variance of the error term. The conditional variance of ε_t, denoted as σ^2 is expressed as:

$$\sigma_t^2 = \operatorname{var}(\varepsilon_t \mid \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) = E\left[(\varepsilon_t - E(\varepsilon_t))^2 \mid \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots\right] \qquad (54)$$

Where it is assumed that $E(\varepsilon_t) = 0$, so that

$$\sigma_t^2 = \operatorname{var}(\varepsilon_t \mid \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) = E\left[\varepsilon_t^2 \mid \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots\right] \qquad (55)$$

The above equation states that the conditional variance of a zero mean normally distributed variable ε_t is equal to the conditional expected value of the square of ε_t.

Under the ARCH model, the autocorrelation in volatility is processed by allowing the conditional variance of the error term σ_t^2, to be time dependent on the previous value of the squared error (Brooks, 2008, pp. 387-388). The ARCH(q) model thus takes the following form:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 \tag{56}$$

where α_0 is the long-term average value and α_i the parameter that measures the information about volatility during the previous periods and needs to be estimated using the maximum likelihood method. For the purpose of our analysis, we model the conditionally heteroscedastic models under the conditional mean equation in the form of an ARMA(p,q) model. Therefore, by jointly taking the conditional mean model of an ARMA(p,q) process to estimate the conditional mean and ARCH(q) equation to estimate the conditional variance, we model the conditional heteroscedasticity by augmenting a dynamic equation to a time series model to govern the time evolution of the conditional variance of the shock (Tsay, 2003, p. 81). An example of such mixed general ARMA(p,q)-ARCH(q) can be expressed as:

$$y_t = \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t \qquad \varepsilon_t \sim N(0, \sigma_t^2) \tag{57}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 \tag{58}$$

3.3 GARCH model

ARCH models have seldom been used in practice, due to their limitations. Various difficulties include ambiguity in deciding the value of q, the number of lags of squared residuals in the processes, as well as the fact that it often requires a large value of q, resulting in a large model that is not parsimonious. Another problem is that non-negativity constraints might be violated, meaning that by expanding the model to include more parameters, the more likely it will be that some of them will have negative estimated values. To overcome these problems, Bollerslev (1986) proposed an extended ARCH(q) model, known as GARCH or the generalized ARCH model, which enables the conditional variance to be dependent on its own previous lags. The GARCH (1,1) is written as:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \tag{59}$$

where σ_t^2 is the conditional variance, as it is a one-period-ahead estimate for the variance obtained on any of the relevant past information, meaning that the current fitted variance is a weighted function of long-term average (dependent on α_0), information about volatility during the previous period ($\alpha_1\varepsilon_{t-1}^2$) and the fitted variance from the model during the previous period ($\beta\sigma_{t-1}^2$). The conditional variance in (55) is changing, although the unconditional variance of ε_t is constant and given by:

$$\text{var}(\varepsilon_t) = \frac{\alpha_0}{1-(\alpha_1+\beta)} \quad (60)$$

To obtain a stationary model, the condition $\alpha_1 + \beta < 1$ must hold. If $\alpha_1 + \beta \geq 1$, it would mean that the unconditional variance of ε_t is not defined, and would be known as non-stationary in variance, with some highly undesirable features, especially in forecasting variance with such a model. An extended and more general form of GARCH(p,q), in which the current conditional variance depends on q lags of the squared error and p lags of the conditional variance is given by:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \quad (61)$$

However, in most cases a GARCH(1,1) model is sufficient to capture the volatility clustering in the data. According to Brooks (2008, p. 394), rarely is any higher order GARCH model estimated or even entertained in the academic finance literature. Thus, for modelling electricity spot prices, we process a mixed ARMA(p,q)-GARCH(1,1) model.

3.4 GJR-GARCH model

The GJR-GARCH model is an extension to the GARCH model proposed by Glosten et al. (1993) that allows for corrections in the symmetric response of volatility to positive and negative shocks in a GARCH model. This occurs as the conditional variance in (61) is a function of the magnitude of the lagged residuals and not their signs, a consequence of squaring the lagged residual in (61). According to Brooks (2008, p. 404), it has been argued that a negative shock to a financial time series is likely to cause volatility to rise by more than a positive shock of the same magnitude; this is known as the leverage effect. Conditional variance in a GJR model is given by:

$$\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma\varepsilon_{t-1}^2 L_{t-1}$$

(62)

Where $L_{t-1} = 1$ if $\varepsilon_{t-1} < 0$,

$L_{t-1} = 0$ otherwise.

Everything else equal to the GARCH model, the additional term $\gamma\varepsilon_{t-1}^2 L_{t-1}$ measures asymmetric responses, where for a leverage effect we would see $\gamma > 0$.

3.5 EGARCH model

Similar to the GJR-GARCH model, the exponential GARCH (EGARCH) model proposed by Nelson (1991) also allows for possible asymmetric effects between price increases and falls. According to Brooks (2008, p. 406), there are different way to express the conditional variance of the EGARCH equation. One of possible specification of EGARCH(1,1) is given by:

$$\ln \sigma_t^2 = \omega + \beta \ln(\sigma_{t-1}^2) + \gamma \frac{\varepsilon_{t-1}}{\sqrt{\sigma_{t-1}^2}} + \alpha \left[\frac{|\varepsilon_{t-1}|}{\sqrt{\sigma_{t-1}^2}} - \sqrt{\frac{2}{\pi}} \right] \tag{63}$$

The model is similar to the GARCH model, although it is distinct from it in several ways. First, it uses logged conditional variance, which relaxes the positiveness constraint of the model coefficients, hence there is no need to artificially impose non negativity constraints on the model coefficients. Second, the γ coefficient enables to measure for asymmetric responses, where if the relationship between the volatility and movements is negative, γ will also be negative.

3.6 Estimation of conditional heteroscedastic models

Unlike in the conditional mean models, GARCH models are not in linear form, thus the common OLS estimation cannot be applied for the estimation of GARCH parameters. The main reason for this is that OLS minimises the residual sum of squares, which depend only on the parameters in the conditional mean equation, and not the conditional variance; therefore, RSS is no longer appropriate (Brooks, 2008, p. 395). The most widely used method for GARCH models is the maximum likelihood approach. In general, when we estimate the parameters using maximum likelihood, we form a likelihood function that is essentially a joint probability density function for a given time series; however, instead of thinking of it as a function of the data given the set of parameters, $f(y_1, y_2, ...y_n | \theta)$, we think of the likelihood function as a function of the parameters given the data, $LF(\theta | y_1, y_2, ...y_n)$ (Reider, 2009, p. 7). The method works by finding the values of the parameters by maximizing the probability of obtaining the observed data from the actual data or, in other words, by finding the most likely values of the parameters given the actual data. The purpose is to form a log likelihood function and to estimate the parameters by first differentiating the LLF with respect to the parameters and further equating the partial derivatives to zero. An example of LLF for the basic ARMA(1,1)-GARCH(1,1) is given as:

$$y_t = \mu + \phi_1 y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \sigma_t^2) \tag{64}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \tag{65}$$

$$LLF = -\frac{T}{2}\log(2\pi) - \frac{-1}{2}\sum_{t=1}^{T}\log(\sigma_t^2) - \frac{-1}{2}\sum_{t=1}^{T}\frac{(y_t - \mu - \phi_1 y_{t-1} - \theta_1\varepsilon_{t-1})^2}{\sigma_t^2}$$

(66)

In the above model, we can derive the maximum likelihood estimators, by first differentiating the LLF with respect to the following parameters μ, ϕ_1, θ_1 and the initial volatility σ_t^2. After the first derivatives are obtained, we can further derive the maximum likelihood estimators, generally denoted by placing hats above the parameters ($\hat{\mu}$).

3.7 Volatility forecasting with conditional heteroscedastic models

GARCH type models are basically models for forecasting volatility where the variance rate follows a mean reverting process and the variance rate estimated at the end of the observation $t-1$ is processed for the succeeding observation t. In order to estimate the forecasts of electricity spot prices, we consider a hybrid model composed of an ARMA and GARCH equations to account for a number of different characteristics of the electricity time series at the same time. According to Brooks (2008, p. 411), a GARCH model is essentially used to describe movements in the conditional variance of the error term ε_t that may not seem very useful; however, it is possible to show that

$$\mathrm{var}(y_t \mid y_{t-1}, y_{t-2}, ...) = \mathrm{var}(\varepsilon_t \mid \varepsilon_{t-1}, \varepsilon_{t-2}, ...)$$

(67)

meaning that the conditional variance of y_t, given its previous values, is the same as the conditional variance of ε_t, given its previous values. Thus modelling σ_t^2, gives forecasts for the variance of y_t as well. Forecasts of a GARCH model can be obtained using methods similar to those of an ARMA model. Consider the following example by Brooks (2008, pp. 412-414) for an ARMA(1,1)-GARCH(1,1) model:

$$y_t = \mu + \phi_1 y_{t-1} + \theta_1\varepsilon_{t-1} + \varepsilon_t$$

(68)

$$\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

(69)

A forecast can be produced by generating forecasts of $\sigma_{T+1}^2 \mid \Omega_T$, $\sigma_{T+2}^2 \mid \Omega_T$, ..., $\sigma_{T+s}^2 \mid \Omega_T$, where Ω_T denotes all information available up to and including that of observation T, for which the conditional variance is given in equation (69). Adding one, two and three to each of the time subscripts of the original equation, we can write the one-, two- and three-steps ahead equations as:

$$\sigma_{T+1}^2 = \alpha_0 + \alpha_1 \varepsilon_T^2 + \beta \sigma_T^2$$

(70)

$$\sigma_{T+2}^2 = \alpha_0 + \alpha_1 \varepsilon_{T+1}^2 + \beta \sigma_{T+1}^2$$

(71)

$$\sigma_{T+3}^2 = \alpha_0 + \alpha_1 \varepsilon_{T+2}^2 + \beta \sigma_{T+2}^2$$

(72)

Note that σ_{T+1}^2, σ_{T+2}^2, σ_{T+3}^2 are not the one-, two-, three-step ahead-forecasts, as these values are not known. To produce forecasts, we have to take their conditional expectations. In the case of the one step-ahead forecast, we take the conditional expectation of (70) and denote the forecast as $\sigma_{1,T}^{f^2}$. The one-step-ahead forecast $\sigma_{1,T}^{f^2}$ is made at time T, where all the values on the right-hand side of (70) are known.

$$\sigma_{1,T}^{f^2} = \alpha_0 + \alpha_1 \varepsilon_T^2 + \beta \sigma_T^2$$

(73)

The two-steps-ahead forecast is different, as we do not know the actual observation of $T+1$; however, we can model it by using the one-step forecast to construct it. This can be done by first rewriting (71) as

$$\sigma_{2,T}^{f^2} = \alpha_0 + \alpha_1 E(\varepsilon_{T+1}^2 \mid \Omega_T) + \beta \sigma_{T+1}^2$$

(74)

Where $E(\varepsilon_{T+1}^2 \mid \Omega_T)$ is the expectation of ε_{T+1}^2, made at time T. We can find it by using the expression for the variance of the residual ε_t. The model assumes that the series ε_t has a zero mean, so that the variance can be written as:

$$\text{var}(\varepsilon_t) = E\left[(\varepsilon_t - E(\varepsilon_t))^2\right] = E(\varepsilon^2) \quad .$$

(75)

Thus, the conditional variance of ε_t is σ_t^2, so that

$$\sigma_t^2 \mid \Omega_T = E(\varepsilon_t)^2$$

(76)

Turning equation (75) around and sending it one period forward to apply it to the problem, we obtain:

$$E(\varepsilon_{T+1} \mid \Omega_T)^2 = \sigma_{T+1}^2$$

(77)

However, σ^2_{T+1} is not known at time T, so that it is replaced by the forecast ($\sigma^{f^2}_{1,T}$) for it, so that (74) becomes:

$$\sigma^{f^2}_{2,T} = \alpha_0 + (\alpha_1 + \beta)\sigma^{f^2}_{1,T}$$

(78)

Similarly, we can produce the three-steps-ahead forecast:

$$\sigma^{f^2}_{3,T} = E_T(\alpha_0 + \alpha_1\varepsilon^2_{T+2} + \beta\sigma^2_{T+2})$$

(79)

$$\sigma^{f^2}_{3,T} = \alpha_0 + (\alpha_1 + \beta)\sigma^{f^2}_{2,T} \tag{80}$$

$$\sigma^{f^2}_{3,T} = \alpha_0 + (\alpha_1 + \beta)\left[\alpha_0 + (\alpha_1 + \beta)\sigma^{f^2}_{1,T}\right]$$

(81)

$$\sigma^{f^2}_{3,T} = \alpha_0 + \alpha_0(\alpha_1 + \beta) + (\alpha_1 + \beta)^2\sigma^{f^2}_{1,T} \tag{82}$$

Further any s-step-ahead forecast would be produced using the following equation

$$\sigma^{f^2}_{s,T} = \alpha_0\sum_{i=1}^{s-1}(\alpha_1 + \beta)^{i-1} + (\alpha_1 + \beta)^{s-1}\sigma^{f^2}_{1,T} \tag{83}$$

For any value of $s \geq 2$. When the s-step-ahead forecast $\sigma^{f^2}_{s,T}$ goes towards the unconditional variance as s goes to infinity, we can see that the variance forecast approaches the unconditional variance of ε_t. From the s-step-ahead forecast, we see that in the second term of the equation (83), $(\alpha_1 + \beta)$ determines how quickly the forecast converges to the unconditional variance. In other words, when $\alpha_1 + \beta < 1$, the final term in (83) becomes progressively smaller as s increases. In the GARCH model, the variance rate exhibits mean reversion with a reversion level of the unconditional variance of ε_t and a reversion rate of $1-(\alpha_1 + \beta)$. Therefore, the forecast of the future variance rate converges towards the unconditional variance of ε_t as we look further ahead. In the opposite case, when $\alpha_1 + \beta > 1$, the weight given to the long-term average variance is negative and the process would be termed 'mean fleeing' instead of 'mean reverting' (Hull, 2012, pp. 509-510).

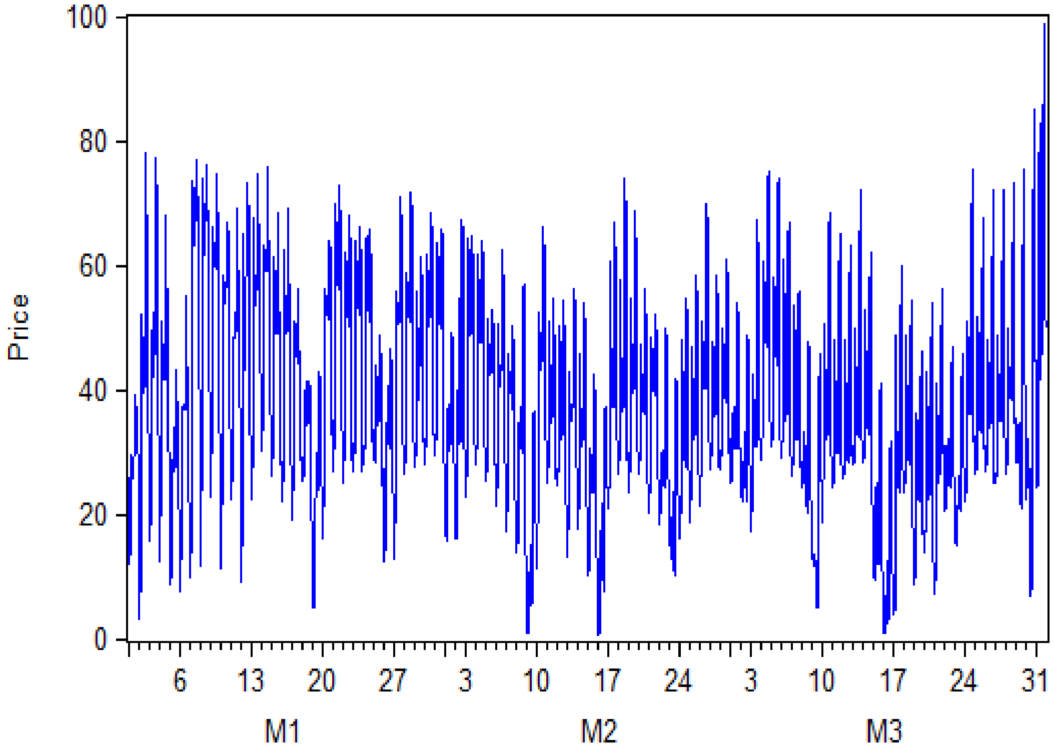
4 MODELLING

In this chapter, we apply the proposed models to the Slovenian BSP Southpool hourly spot prices using different models, including a pure ARMA(p,q) model, ARMA(p,q)-GARCH(p,q), ARMA-EGARCH(p,q) and ARMA(p,q)-GJR(p,q).

4.1 Descriptive statistics and description of data

The dataset used to analyse the hypotheses consists of 2160 hourly observations from January 1st to March 31st 2014, collected from the BSP Southpool exchange. The electricity spot price hourly time series for the first quarter of 2014 can be seen in Figure 3.

Figure 3. 2014 Q1 BSP Southpool hourly dataset



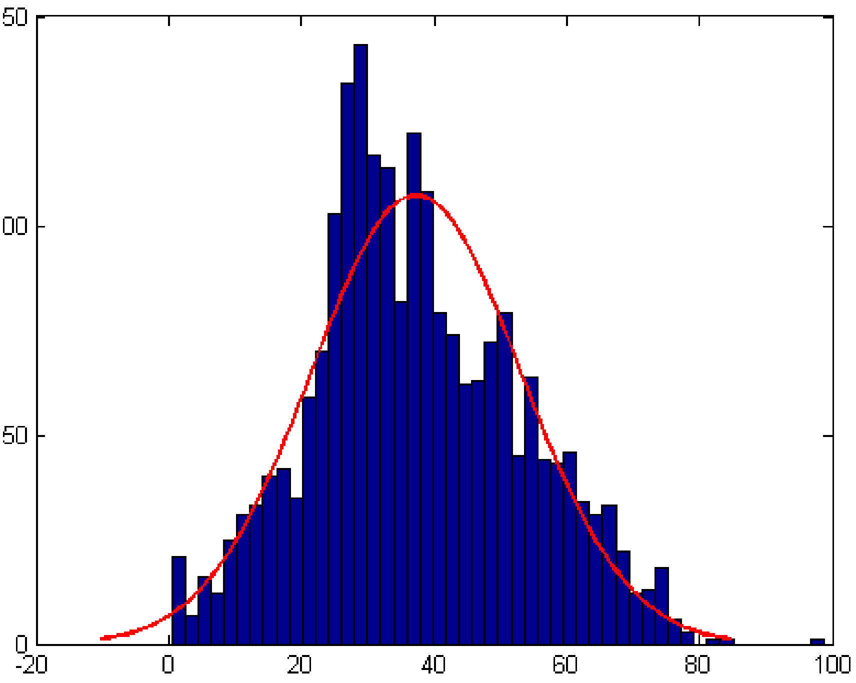
From the above plot, it is clearly evident that the series exhibits a strong presence of daily and weekly seasonality. Volatility clustering is also present, though it appears at first sight that the volatility occurring in bursts is also correlated with seasonality. We further check the descriptive statistics of the series presented in Table 2.

Table 2. Descriptive statistics of 2014 Q1 BSP Southpool hourly dataset

Statistic	Value
Observations	2160
Mean	37.30
Median	35.69
Maximum	98.85
Minimum	0.50
Standard Deviation	15.81
Skewness	0.2555
Kurtosis	2.7424

The above table shows high volatility in the series with a mean price 35.695 and a maximum and minimum of 98.5 and 0.5 respectively. Expecting that the data is not normally distributed, we nevertheless have to verify how far the distribution of the dataset from the normal distribution and stationarity is. Therefore, we plot the normalized histogram of the dataset in Figure 4.

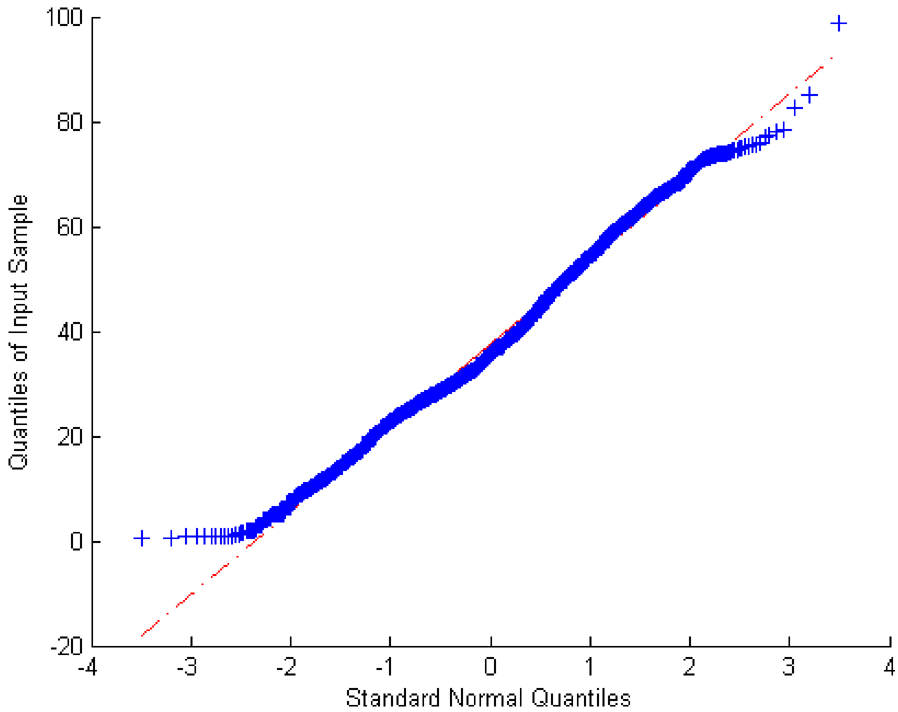
Figure 4. Normalized histogram of 2014 Q1 BSP Southpool hourly dataset



From the normalized histogram, it seems that the distribution of the dataset is close to normal, with some spikes above the normal distribution for prices just below the median price and an outlier at the right tail of the distribution. Almost normal distribution of the dataset is confirmed by the two parameters for comparison of a certain probability distribution with the normal distribution, i.e. skewness and kurtosis. Skewness measures the asymmetry of the distribution of a certain variable around its mean, while kurtosis measures its peakness. In order for the probability distribution to be normally distributed, the values of skewness and kurtosis should be 0 and 3, respectively. Table 2 shows that skewness and kurtosis of our dataset are 0.255 and 2.7424, which is very close to their theoretical values. Following the normalized histogram, we also plot the QQ-plot[7] in Figure 5.

Figure 5. QQ-plot of 2014 Q1 BSP Southpool hourly dataset

[7] The QQ plot is a plot of quantiles of two distributions, in our case of the dataset and the normal distribution, where the pattern of points compare the distributions; the distributions are similar when its points lie approximately on the line $y = x$ of the dataset against the normal distribution.



The QQ-plot of the dataset also confirms that the data is close to normal distribution, with some diverging points from normality at the tails.

4.2. Stationarity and unit root tests of the dataset

Following the reasons for the stationarity of a series and problems when using non-stationary data as explained in the previous chapter, we have to test the series for stationarity. We test the data with an ADF test, PP test and KPSS test. Recall that the hypotheses for the ADF test and PP test are:

H_0: series contains a unit root

H_1: series is stationary.

While KPSS test employs an opposite null hypotheses;

H_0: series is stationary

H_1: series contains a unit root.

The results of the test with test statistics are given in Table 3. Note that the critical value at 1% significance is -3.43 for ADF test and PP test, and 0.739 for KPSS test. Moreover, in the case of the ADF test and PP test, the test is performed with up to 14 lags of the dependent variable in the regression equation with a constant but no trend in the test equation.

Table 3. Unit root tests for 2014 Q1 BSP Southpool hourly dataset

Data	ADF	t-statistic	PP	t-statistic	KPSS	t-statistic
2014 Q1 Hourly dataset	Reject (at 1%)	-10.84874	Reject (at 1%)	-9.345637	Not reject (at 1%)	1.461686

The ADF test and PP test reject the hypothesis that the series contains a unit root, while the KPSS test does not reject its null hypotheses that the series is stationary. Therefore, all the tests confirm that the series is not non-stationary.

4.3 Identification and seasonal adjustment of the data

Apart from observing the time plot of the series, seasonal patterns and the time-varying nature of volatility can also be spotted with the plots of the acf and pacf. Using graphical representations of the acf and pacf, we analyse the most appropriate specification and deseasonalise the series according to the seasonal patterns in the data. In the first step, we plot the original price series over time,and plot the acf and pacf in Figures 6 and 7.

From the plots, it is evident that both series are not dying away quickly, which is a clear sign of non-stationarity in the series. Also highly significant are the correlations after every 24th lag, a consequence of the daily cyclical effect. Therefore, to smooth the series in order to obtain a stationary series, we further deseasonalise the series for the 24- and 168-hour periodic seasonal effects; 168-hour correlations can be seen in Figure 3, where we observe a cyclical pattern after every 7th day or every 168th observation. By applying a stable seasonal filter, we now deseasonalise for the 24-hour daily and 168-hour weekly periodicity using additive decomposition. A deseasonalised series for the first quarter of 2014 is plotted in Figure 8, while the acf and pacf of the deseasonalised series are shown in Figure 9 and Figure 10.

Figure 6. ACF of the original 2014 Q1 BSP Southpool hourly dataset

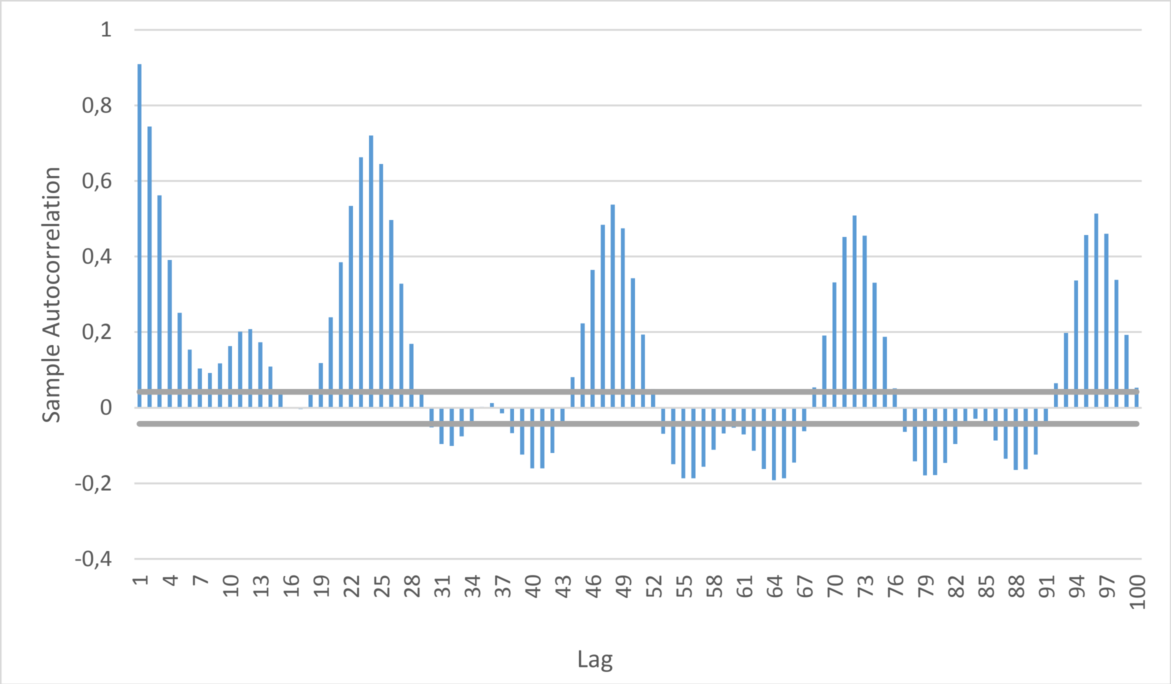


Figure 7. PACF of the original 2014 Q1 BSP Southpool hourly dataset

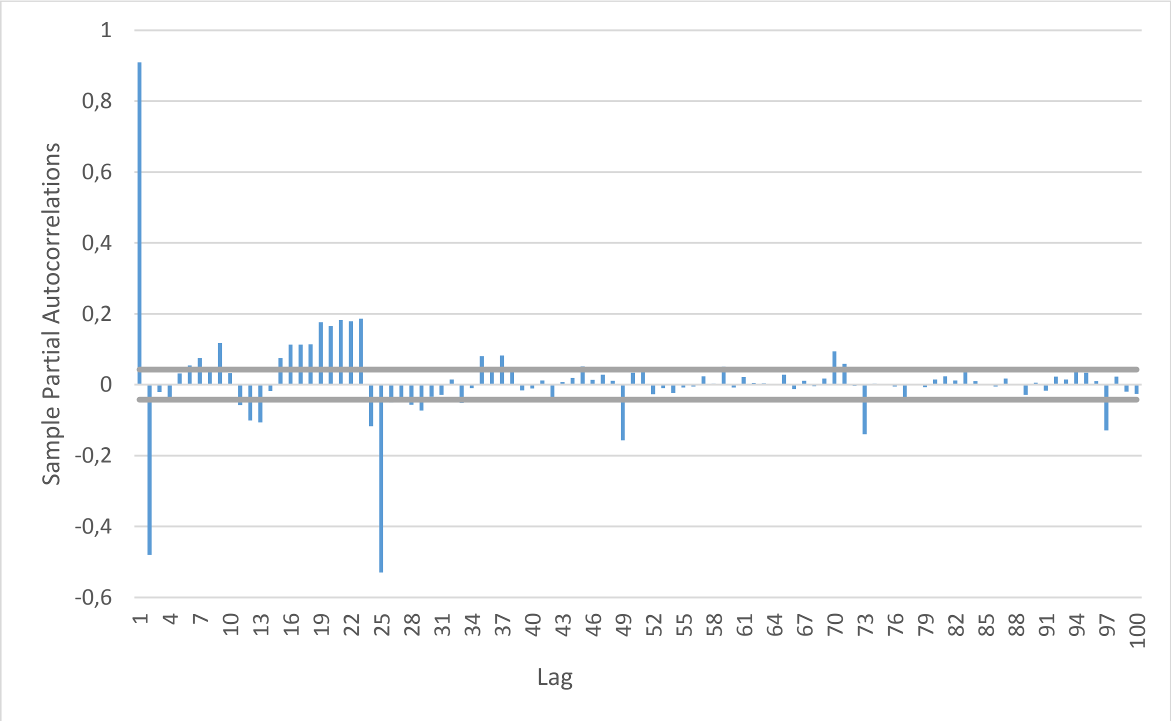
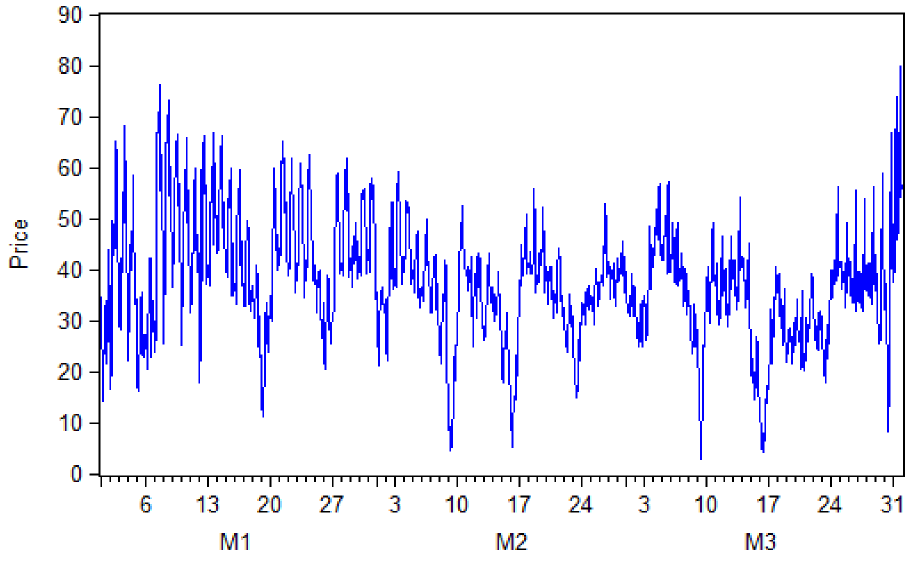


Figure 8. 2014 Q1 Deseasonalised BSP Southpool hourly dataset



Observing the deseasonalised plot, we can observe that a significant portion of intraday volatility has been removed, though the cyclical daily and weekly pattern remain clearly visible.

Figure 9. ACF of hourly 2014 Q1 series after seasonal adjustment for daily and weekly periodicity.

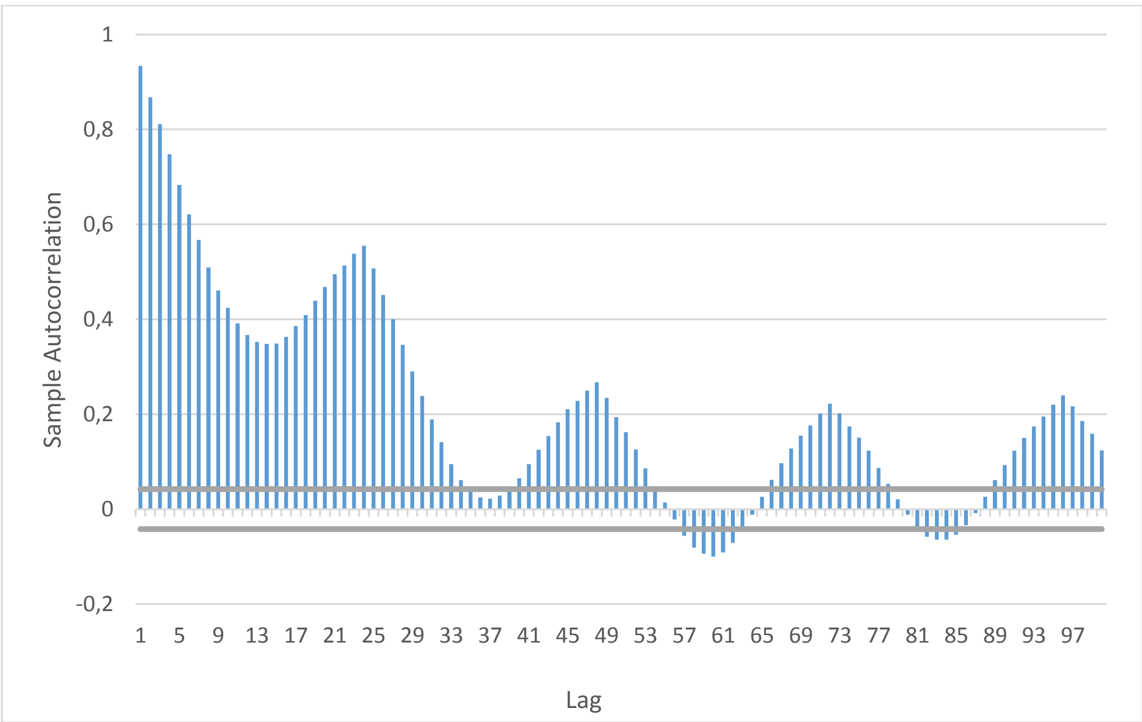
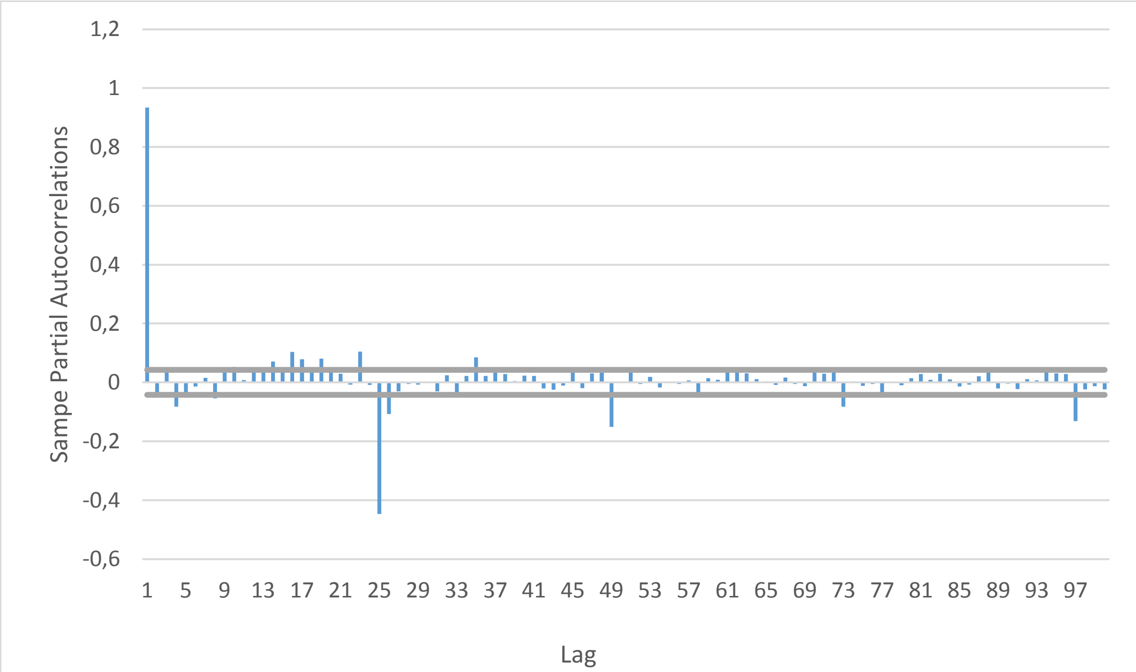


Figure 10. pacf of hourly 2014 Q1 series after seasonal adjustment for daily and weekly periodicity.



Comparing the original and the deseasonalised acf, we can see that the latter becomes much smoother after accounting for the seasonal components, and we can now describe the series as weakly or shortly stationary. Since short stationarity has been obtained, we can proceed further with the next stage and identify the order of the ARMA process.

4.4 ARMA MODEL

4.4.1 IDENTIFICATION OF THE ORDER OF ARMA MODEL

The patterns exhibited in the graphical plots of the deseasonalised series are difficult to interpret in terms of the identification of the order of the model. To overcome this, we use AIC and SBIC information criteria to decide which of the models is most appropriate. Tables 4 and 5 show all values of the AIC and SBIC information criteria, calculated for lags until p=10 and q=10.

Table 4. AIC values for 2014 Q1 deseasonalised series

p/q	0	1	2	3	4	5	6	7	8	9	10
0	7.745746	6.787062	6.348398	6.121744	5.974413	5.865674	5.830440	5.777634	5.738522	5.728824	5.705874
1	5.675287	5.674213	5.672703	5.668230	5.667633	5.667264	5.668122	5.665164	5.666023	5.663960	5.664776
2	5.674343	5.670569	5.671087	5.665971	5.666892	5.667597	5.668486	5.666180	5.659706	5.665530	5.638587
3	5.673482	5.671305	5.672201	5.667299	5.668023	5.648197	5.647695	5.618199	5.648582	5.653909	5.660582
4	5.668611	5.667125	5.667975	5.594047	5.665013	5.647099	5.656629	5.656653	5.648728	5.645299	5.628273
5	5.668172	5.668331	5.669075	5.572252	5.623629	5.623674	5.552644	5.553544	5.522620	5.656302	5.656751
6	5.668712	5.668361	5.647405	5.644519	5.623516	5.541378	5.525299	5.632296	5.632171	5.641282	5.546572
7	5.669777	5.669239	5.674032	5.645229	5.649163	5.557373	5.518977	5.519412	5.604326	5.548807	5.636628
8	5.666200	5.666873	5.618987	5.619753	5.611901	5.608168	5.588489	5.589645	5.572991	5.529207	5.563538
9	5.665957	5.648479	5.620183	5.610490	5.609266	5.610177	5.589177	5.589339	5.487263	5.579350	5.472083
10	5.663260	5.647888	5.642490	5.562756	5.561541	5.521632	5.508551	5.502867	5.579352	5.528957	5.522209

Table 5. SBIC values for 2014 Q1 deseasonalised series

p/q	0	1	2	3	4	5	6	7	8	9	10
0	7.748375	6.792319	6.356284	6.132258	5.987556	5.881446	5.848841	5.798663	5.762180	5.755111	5.734789
1	5.680547	5.682102	5.683222	5.681378	5.683411	5.685671	5.689159	5.688830	5.692319	5.692886	5.696331
2	5.682235	5.681091	5.684240	5.681755	5.685306	5.688642	5.692162	5.692486	5.688643	5.697098	5.672785
3	5.684009	5.684463	5.687991	5.685720	5.689076	5.671882	5.674011	5.647147	5.680162	5.688121	5.697425
4	5.681775	5.682921	5.686403	5.615108	5.688707	5.673426	5.685588	5.688244	5.682952	5.682156	5.667763
5	5.683974	5.686767	5.690144	5.595955	5.649966	5.652645	5.584248	5.587782	5.559492	5.695807	5.698890
6	5.687155	5.689439	5.671117	5.670866	5.652497	5.572994	5.559550	5.669181	5.671691	5.683437	5.591362
7	5.690862	5.692961	5.657317	5.674222	5.680791	5.591637	5.555876	5.558947	5.646497	5.593613	5.684070
8	5.689931	5.693240	5.647991	5.651393	5.646178	5.645082	5.628040	5.631832	5.617815	5.576667	5.613635
9	5.692334	5.677494	5.651835	5.644780	5.646194	5.649742	5.631380	5.634180	5.534742	5.629467	5.524837
10	5.692285	5.679552	5.676793	5.599698	5.601122	5.563851	5.553409	5.550364	5.629487	5.581731	5.577621

The lowest value for both information criteria is present in the form of the ARMA(9,10) model; however, the model contains a number of estimated parameters that are not significant, specifically in their t-statistics and p-values. Following this, we estimate 10 models with the lowest value of AIC and SBIC information criteria and test the significance of their parameters as well as perform the Ljung Box test for autocorrelation in the residuals in order to see which of them fits best. Table 6 lists 11 ARMA models chosen by AIC and SBIC method.

Table 6. ARMA models with the lowest values of AIC and SBIC information criteria.

(p,q)	AIC	(p,q)	SBIC
(9,10)	5.472083	(9,10)	5.524837
(9,8)	5.487263	(9,8)	5.534742
(10,7)	5.502867	(10,7)	5.550364
(10,6)	5.508551	(10,6)	5.553409
(7,6)	5.518977	(7,6)	5.555876
(7,7)	5.519412	(7,7)	5.558947
(10,5)	5.521632	(5,8)	5.559492
(10,10)	5.522209	(6,6)	5.559550
(5,8)	5.522620	(10,5)	5.563851
(6,6)	5.525299	(6,5)	5.572994

We can see that the AIC and SBIC choose similar models, with only one different model among the 10 models with the lowest information criteria value. We further applied diagnostic checks and significance tests to the estimated parameters of all of the above models, and determined that the models that fits best is ARMA(10,7), which we now examine in more detail.

4.4.2 Estimation of parameters, diagnostic checking and interpretation for ARMA model

Table 7 shows the values of estimated coefficients for the ARMA(10,7) process. We can see that all coefficients for both AR and MA processes are statistically significant, except for ϕ_{10}, whose t-Statistic is lower than 2.

Table 7. Estimated parameters for ARMA(10,7) process

Parameter	Estimated value	Standard error	t-Statistic	p-values
μ	37.78478	1.866763	20.24081	0.0000
ϕ_1	1.327056	0.098965	13.40938	0.0000
ϕ_2	0.213967	0.073947	2.893510	0.0038
ϕ_3	0.195112	0.063667	3.064549	0.0022
ϕ_4	-0.936265	0.080320	-11.65669	0.0000
ϕ_5	-0.301336	0.048197	-6.252170	0.0000
ϕ_6	-0.221218	0.066335	-3.334884	0.0009

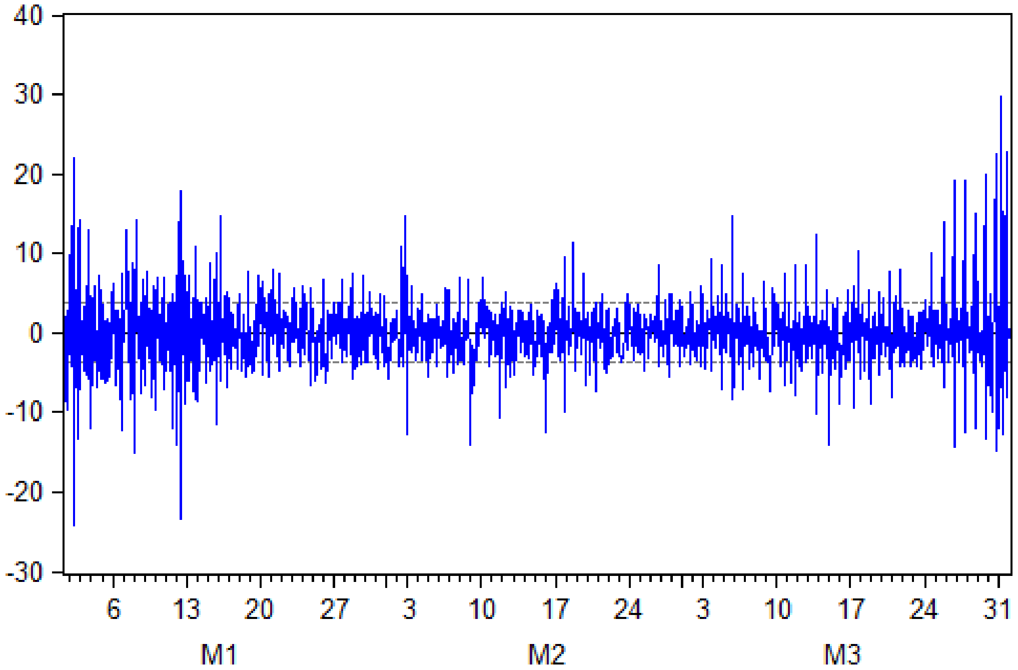
(table continues)

(continues)

ϕ_7	1.322129	0.081605	16.20149	0.0000
ϕ_8	-0.702259	0.079892	-8.790122	0.0000
ϕ_9	0.092411	0.040207	2.298375	0.0216
ϕ_{10}	0.002755	0.030568	0.090116	0.9282
θ_1	-0.419419	0.096667	-4.338808	0.0000
θ_2	-0.679852	0.032005	-21.24218	0.0000
θ_3	-0.717403	0.044178	-16.23881	0.0000
θ_4	0.325928	0.102121	3.191593	0.0014
θ_5	0.648687	0.043518	14.90620	0.0000
θ_6	0.710277	0.034432	20.62855	0.0000
θ_7	-0.693227	0.091602	-7.567841	0.0000

Figure 11 shows the graphical representation of residuals in ARMA(10,7) process, which seem to be stationary and volatility clustering.

Figure 11. Residuals of ARMA(10,7) process



We also plot the acf and pacf of ARMA(10,7) process in Figures 12 and 13.

Figure 12. acf of residuals for ARMA(10,7) process

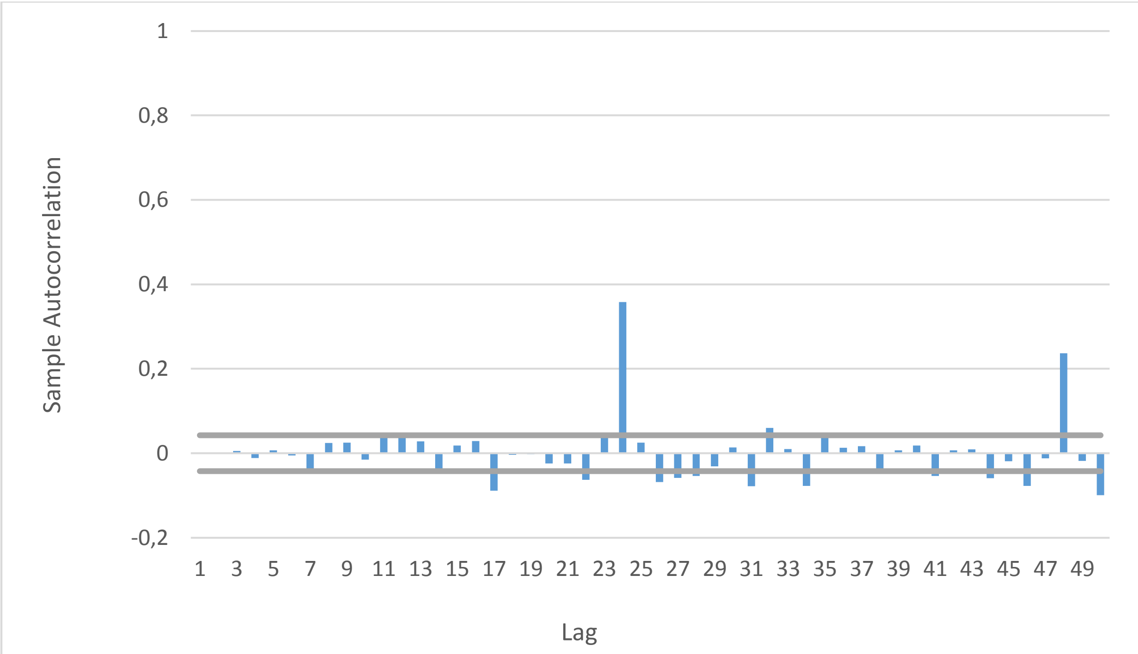
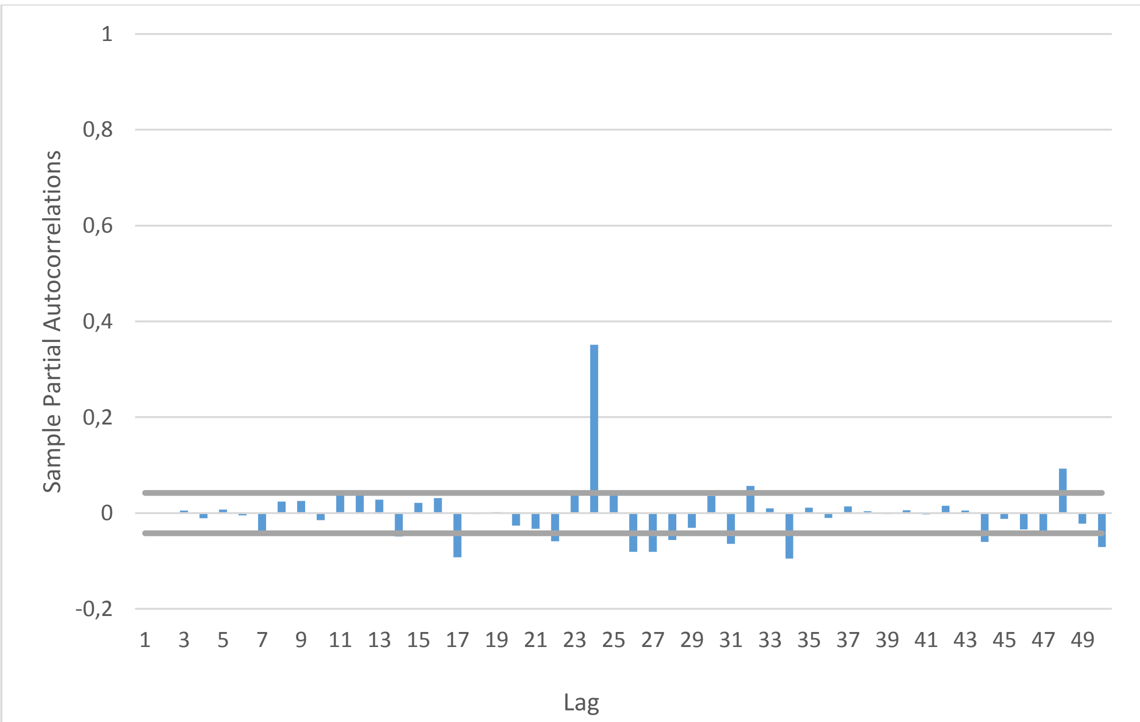


Figure 13. pacf of residuals for ARMA(10,7) process



In both acf and pacf, we can observe a significant correlation at every 24th lag, confirming that the model still exhibits daily seasonality and some intra-day correlation in the residuals. This suggests that variance is close to non-stationary and indicates that a further GARCH process may fit better. To verify for the autocorrelations in the residuals of the ARMA(10,7) process, we perform a Ljung Box test for residuals in a given series, testing the acf at 5% significance. Tables 8 and 9 show the values of the Ljung Box test for residuals and squared residuals of ARMA(10,7) and the relevant critical values from a χ^2 distribution for different degrees of freedom (lags of acf) at the 5% level.

Table 8. Ljung Box test for residuals of ARMA(10,7) model

Lag	AC coefficient	Q-statistics	Critical value (at 5%)
5	0.007	0.3964	11.070
10	-0.015	7.2906	18.307
15	0.021	21.497	24.996
16	0.029	23.336	26.296
17	-0.089	40.403	27.587

Table 9. Ljung Box test for squared residuals of ARMA(10,7) model

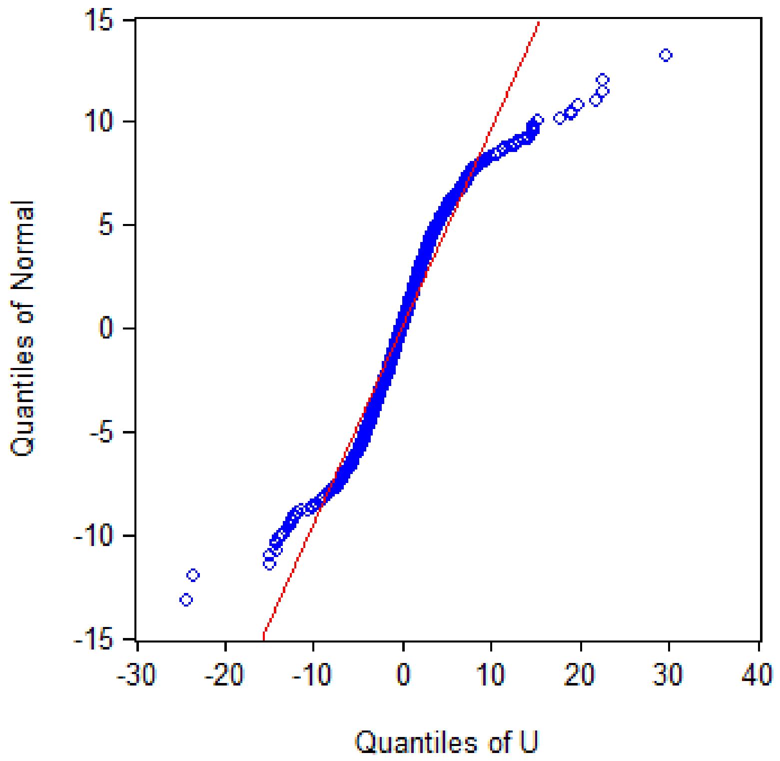
Lag	AC coefficient	Q-statistics	Critical value (at 5%)
1	0.261	147.12	3.841
5	0.032	214.63	11.070

As can be observed from Figure 12, the Ljung Box test confirms that the series is not autocorrelated till the 16th lag. After the 17th lag, the Ljung Box test statistic exceeds its critical value. We can also construct a 95% confidence interval for the acf to verify this by:

$$\pm 1.96 \times \frac{1}{\sqrt{T}} = 0.04217$$

where T =2160 in our case. In the acf and pacf plots, confidence intervals are represented by the grey lines. In Figure 12, the autocorrelation coefficient at the 17th lag clearly exceeds this line as do many subsequent lags; therefore, we can conclude that the series is not autocorrelated until the 16th lag. The squared residuals, however, do not exhibit this, as the Ljung Box test for squared residuals indicates that its test statistic exceeds the critical value for the autocorrelation of residuals at the very first lag, suggesting that modelling variance with a GARCH model might yield better results. In the final step of examining ARMA(10,7) we plot the QQ-plot for residuals to check whether they follow normal distribution.

Figure 14. QQ-plot for residuals of ARMA(10,7)



From Figure 14, we can observe that the residuals of ARMA(10,7) are a bit far from being normally distributed at the tails of the distribution. Therefore, we can conclude that univariate time series model of ARMA(10,7) is shortly stationary in its levels and non-stationary in variance, while the distribution of its residuals is not normally distributed, with the tails of its distribution diverging from normality. To correct for this, we now propose different mixed models, composed of the conditional mean equation based on ARMA specifications and conditional variance equation based on GARCH(p,q) models and its extensions (EGARCH, GJR-GARCH) to account for volatility.

4.5 Mixed ARMA-GARCH model

Using ARMA(10,7) for the conditional mean equation, we propose a dynamic ARMA-GARCH model with one lag for the squared residual and one lag of the conditional variance in the conditional variance equation. Therefore, by modelling a mixed

ARMA(10,7)-GARCH(1,1) process, we estimate the following parameters, as presented in Table 10.

Table 10. Estimated parameters for ARMA(10,7)-GARCH(1,1) process

Parameter	Estimated value	Standard	t-Statistic	p-values
μ	40.59962	1.499416	27.07696	0.0000
ϕ_1	1.096624	0.021539	50.91433	0.0000
ϕ_2	0.625435	0.033521	18.65789	0.0000
ϕ_3	-0.448887	0.012505	-35.8975	0.0000
ϕ_4	0.003725	0.027525	0.135339	0.8923
ϕ_5	-0.886535	0.001898	-466.999	0.0000
ϕ_6	0.358661	0.026963	13.30184	0.0000
ϕ_7	0.326159	0.020724	15.73789	0.0000
ϕ_8	-0.105484	0.025954	-4.06431	0.0000
ϕ_9	-0.069618	0.018058	-3.85527	0.0001
ϕ_{10}	0.084468	0.014688	5.750996	0.0000
θ_1	-0.111232	0.014123	-7.87606	0.0000
θ_2	-0.91004	0.009399	-96.8207	0.0000
θ_3	-0.274217	0.011847	-23.1466	0.0000
θ_4	-0.215398	0.002377	-90.6257	0.0000
θ_5	0.68586	0.000448	1530.617	0.0000
θ_6	0.333586	0.007515	44.39117	0.0000
θ_7	-0.130202	0.001048	-124.263	0.0000
α_0	2.154021	0.15540	13.86117	0.0000
α_1	0.311638	0.021765	14.31841	0.0000
β	0.582584	0.01558	37.39314	0.0000

Table 10 indicates that ARMA lags are statistically significant with an even higher significance than in the pure ARMA(10,7) process, except at lag 4. The coefficients of the variance equation are also highly statistically significant, while the unconditional variance is constant as the sum of the two lagged coefficients in GARCH is smaller than one. Recall that if this sum would be larger than one, it would mean that the model is non-stationary in variance, and there would be no theoretical motivation for its existence. The following figures plot the residuals of ARMA(10,7)-GARCH(1,1) and its sample acf and pacf.

Figure 15. Residuals of ARMA(10,7)-GARCH(1,1) process

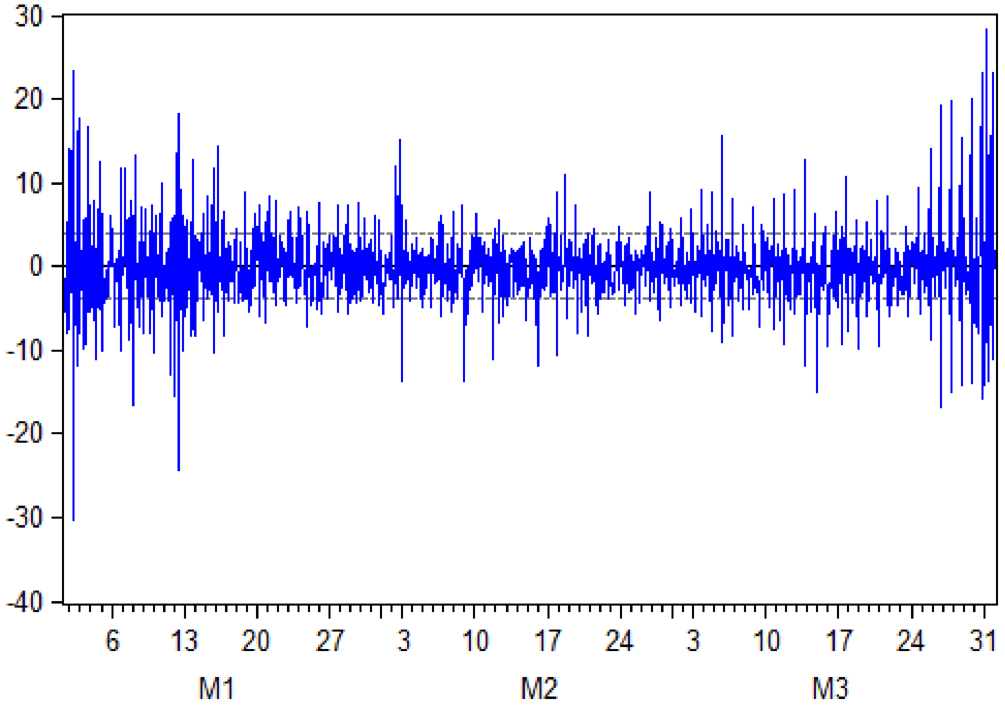


Figure 16. ACF for residuals of ARMA(10,7)-GARCH(1,1) process with confidence intervals

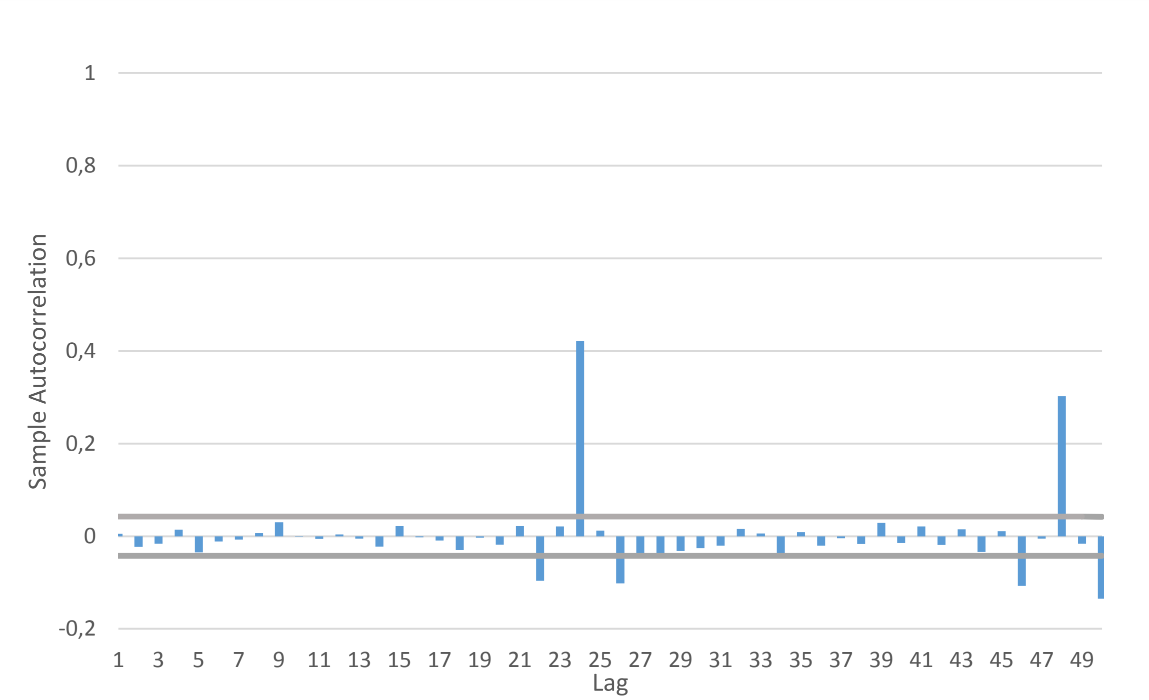


Figure 17. PACF for residuals of ARMA(10,7)-GARCH(1,1) process with confidence intervals

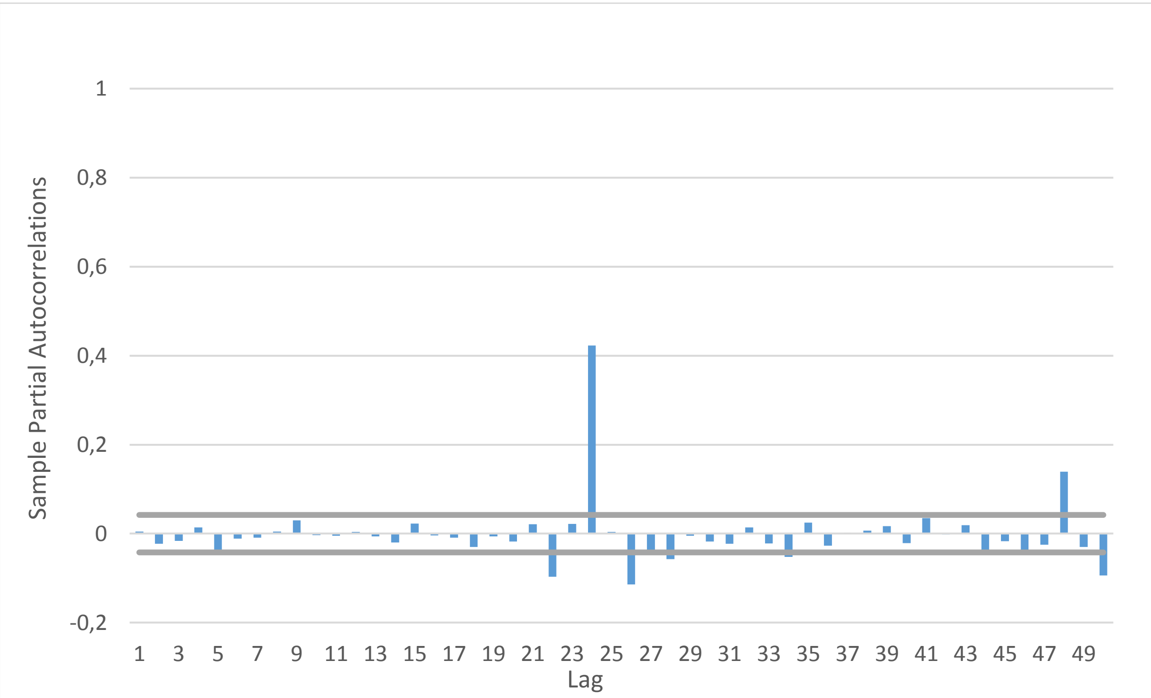
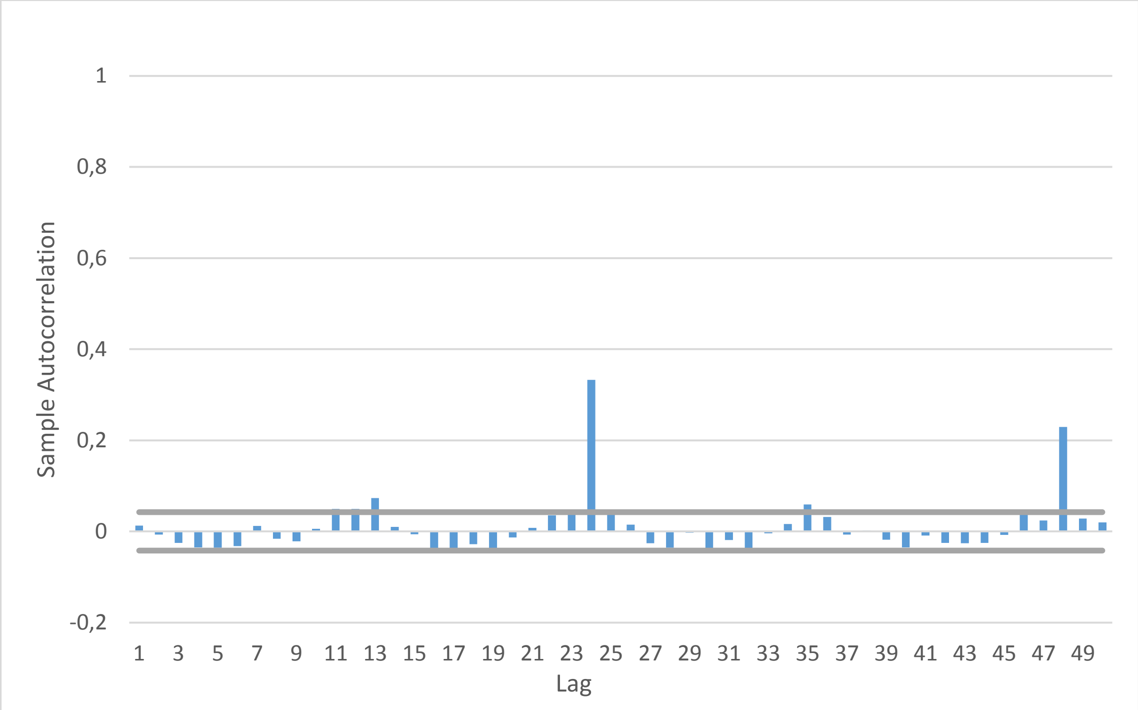


Figure 18. acf for squared residuals of ARMA(10,7)-GARCH(1,1) process with confidence intervals



From the above plots of autocorrelations, we can observe that most of the values for residuals are within the 95% confidence interval, specifically at the early lags of correlations. However, the 24-hour periodicity is still significantly present in the series. We further formally check the presence of autocorrealtions in the model with a Ljung Box test for residuals and squared residuals.

Table 11. Ljung Box test for residuals of ARMA(10,7)-GARCH(1,1) model

Lag	AC coefficient	Q-statistics	Critical value (at
5	-0.035	4.8590	11.070
10	-0.001	7.2649	18.307
15	0.022	9.5067	24.996
20	-0.018	12.314	31.410
23	0.021	34.431	35.172
24	0.422	422.62	36.415

Table 12. Ljung Box test for squared residuals of ARMA(10,7)-GARCH(1,1) model

Lag	AC coefficient	Q-statistics	Critical value (at 5%)
5	-0.036	7.1097	11.070

10	0.006	11.351	18.307
12	0.049	21.720	21.026

The Ljung Box test confirms that the series is not autocorrelated till lag 23 for residuals and lag 11 for squared residuals, indicating that the mixed ARMA-GARCH model deals much better with heteroscedasticity than the pure ARMA model does.

Before we begin forecasting with the ARMA(10,7)-GARCH(1,1) model, we examine the ARMA(10,7)-GJR(1,1) and ARMA(10,7)-EGARCH(1,1) models.

4.6 Mixed ARMA-GJR-GARCH model

Using the extension to account for non-negativity constraints that may be violated by the basic GARCH model, we now propose a mixed GJR-GARCH model with an additional term in the variance equation to account for possible asymmetries. Estimated coefficients of the mixed ARMA(10,7)-GJR-GARCH(1,1) process are presented in Table 13.

Estimated values of the ARMA(10,7)-GJR-GARCH(1,1) process indicate that all coefficients except ϕ_9 are highly significant in the mean equation, while in the conditional variance equation, both GARCH terms and constant α_0 are statistically significant, but this does not hold for the additional asymmetry term γ. Consequently, we cannot make safe conclusions about the leverage effect from ARMA(10,7)-GJR-GARCH(1,1) and hopefully the ARMA(10,7)-EGARCH(1,1) specification will give better results regarding the movements after negative and positive shocks within the model. Nevertheless, we check the plot of residuals in the model, the structure of autocorrelations and perform Ljung Box test for stationarity, as the next chapter compares forecasts obtained from the ARMA-GJR-GARCH model to forecasts of the ARMA-GARCH and ARMA-EGARCH models.

Table 13. Estimated parameters for ARMA(10,7)-GJR-GARCH(1,1) process

Parameter	Estimated value	Standard error	t-Statistic	p-values
μ	40.14858	1.757081	22.84959	0.0000
ϕ_1	0.930169	0.003615	257.3283	0.0000
ϕ_2	0.825624	0.008014	103.0251	0.0000
ϕ_3	-0.334890	0.003233	-103.5850	0.0000
ϕ_4	0.100585	0.019253	5.224275	0.0000
ϕ_5	-1.329264	0.021089	-63.03203	0.0000
ϕ_6	0.317025	0.002998	105.7294	0.0000

ϕ_7	0.719204	0.010402	69.13857	0.0000
ϕ_8	-0.310001	0.007799	-39.75125	0.0000
ϕ_9	0.028900	0.022082	1.308742	0.1906
ϕ_{10}	0.035794	0.014683	2.437785	0.0148
θ_1	0.054053	0.001098	49.22262	0.0000
θ_2	-0.930841	0.001302	-714.9833	0.0000
θ_3	-0.464889	0.001155	-402.4625	0.0000
θ_4	-0.458388	0.001363	-336.2986	0.0000
θ_5	0.929887	0.000890	1045.393	0.0000
θ_6	0.608079	0.000536	1134.744	0.0000
θ_7	-0.334809	0.000173	-1931.167	0.0000
α_0	1.810502	0.133997	13.51152	0.0000
α_1	0.266554	0.023139	11.51961	0.0000
β	-0.03196	0.033791	-0.945662	0.0000
γ	0.639017	0.014916	42.84055	0.3443

Figure 19. Residuals of ARMA(10,7)-GJR-GARCH(1,1) process

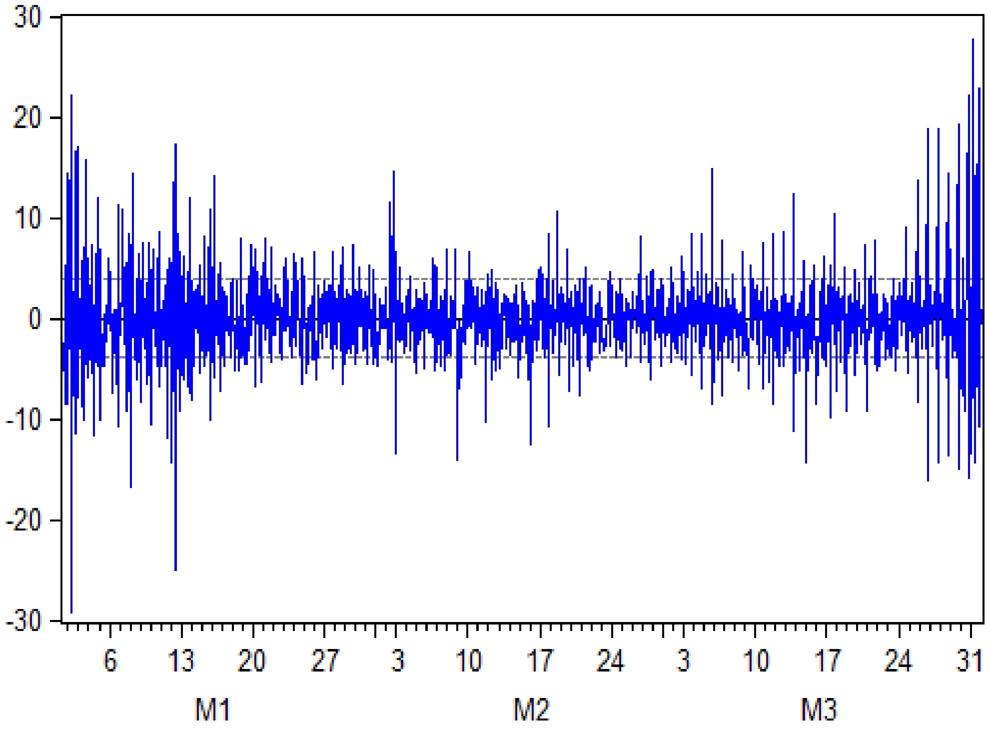


Figure 20. acf for residuals of ARMA(10,7)-GJR-GARCH(1,1) process with confidence intervals

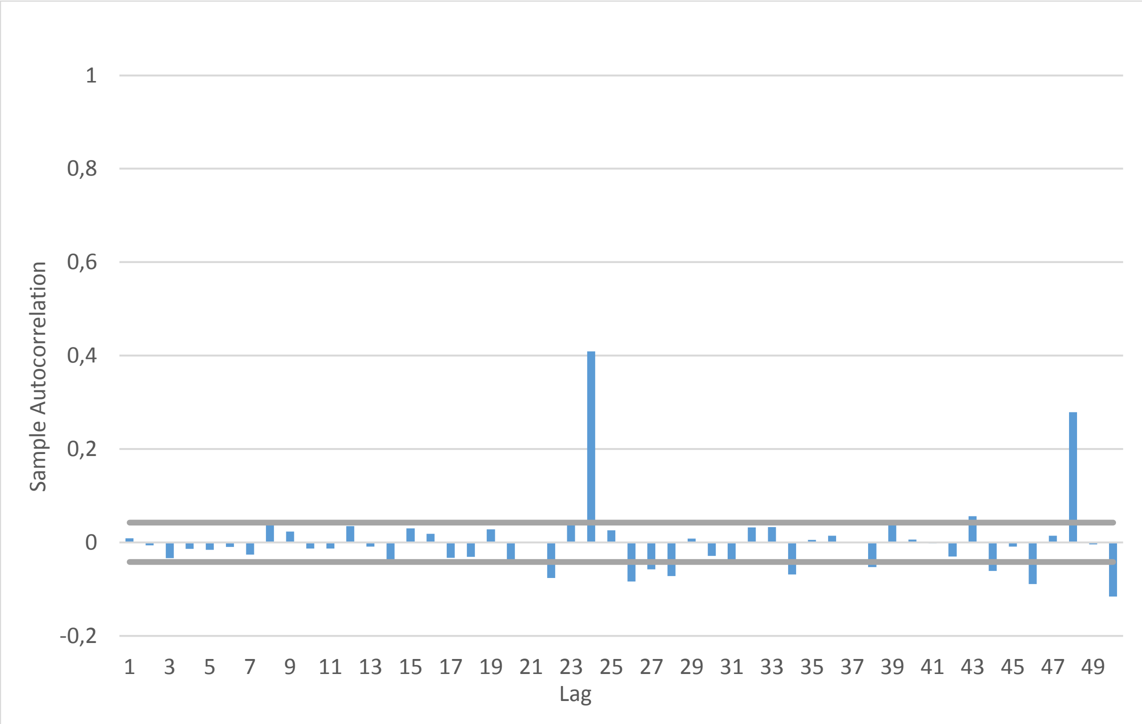


Figure 21. pacf for residuals of ARMA(10,7)-GJR-GARCH(1,1) process with confidence intervals

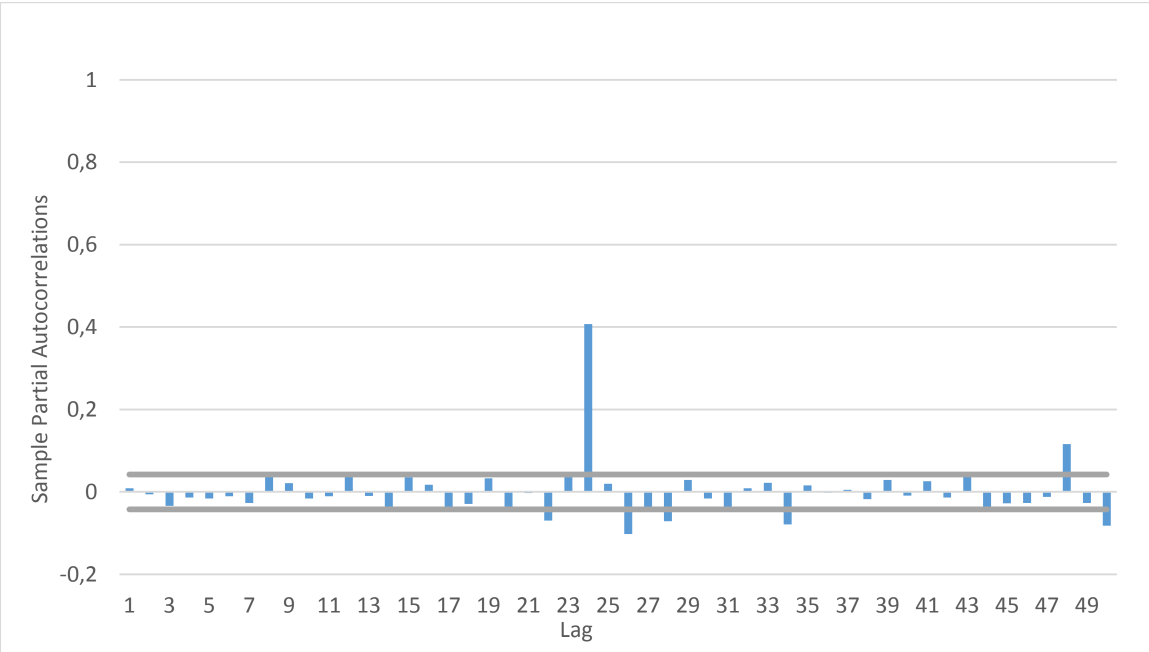
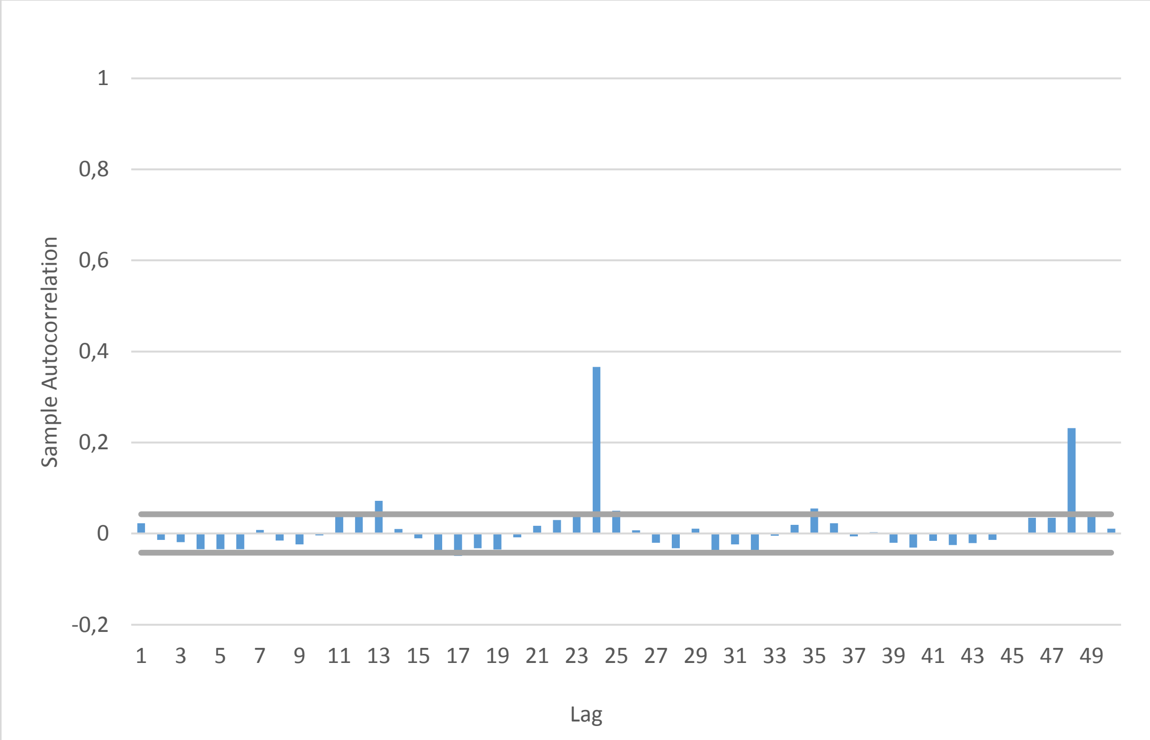


Figure 22. acf for squared residuals of ARMA(10,7)-GJR-GARCH(1,1) process with confidence intervals



Similarly to the ARMA-GARCH specification, the 24-hour seasonal effect is still strongly present in the series, while the autocorrelations at the early lags of correlations are mostly within bounds. We test this with the Ljung Box test for residuals and squared residuals.

Table 14. Ljung Box test for residuals of ARMA(10,7)-GJR-GARCH(1,1) model

Lag	AC coefficient	Q-statistics	Critical value (at 5%)
5	-0.016	3.6622	11.070
10	-0.013	9.8303	18.307
15	0.030	19.087	24.996
21	0.002	29.905	32.671
22	-0.076	42.377	36.781

Table 15. Ljung Box test for squared residuals of ARMA(10,7)-GJR-GARCH(1,1) model

Lag	AC coefficient	Q-statistics	Critical value (at 5%)
5	-0.034	7.3574	11.070
12	0.040	19.519	21.026
13	0.072	30.589	22.362

The Ljung Box test shows that the series is not autocorrelated until lag 21 for residuals and lag 12 for squared residuals, indicating that the mixed ARMA-GJR-GARCH process behaves very similarly to the ARMA-GARCH model. This suggests that the models will give similar forecasts and that different extensions to the GARCH model, apart from the 'corrections' for the restrictions and limitations in the basic GARCH model, do not fit the data much better than the original GARCH specification.

4.7 ARMA-EGARCH model

We now estimate another extension of GARCH that also accounts for asymmetric responses, i.e. the exponential GARCH model. Estimated coefficients of the mixed ARMA(10,7)-EGARCH(1,1) process are presented in Table 16.

The results of the ARMA(10,7)-EGARCH(1,1) process show that all of the coefficients in both equations are highly statistically significant. Furthermore, the positive estimate of the asymmetry term γ in the conditional variance equation suggests that a positive shock implies a lower next-period conditional variance than negative shocks: in other words, that an increasing electricity price leads to a lower next-period volatility than when the electricity price decreases by the same amount. The following figures plot the residuals of ARMA(10,7)-EGARCH(1,1) and its sample acf and pacf.

Table 16. Estimated parameters for ARMA(10,7)-GJR-GARCH(1,1) process

Parameter	Estimated value	Standard error	t-Statistic	p-values
μ	39.32399	2.143386	18.34667	0.0000
ϕ_1	0.620480	0.018406	33.71069	0.0000
ϕ_2	-0.362532	0.020585	-17.61185	0.0000
ϕ_3	0.408554	0.015724	25.98318	0.0000
ϕ_4	0.876752	0.015160	57.83302	0.0000
ϕ_5	-0.470137	0.019688	-23.87914	0.0000
ϕ_6	0.617195	0.022704	27.18394	0.0000
ϕ_7	-0.558247	0.016729	-33.37071	0.0000
ϕ_8	-0.307416	0.016327	-18.82826	0.0000
ϕ_9	0.046578	0.014812	3.144661	0.0017

ϕ_{10}	0.107896	0.015628	6.904030	0.0000
θ_1	0.508747	0.001244	408.8592	0.0000
θ_2	0.828208	0.002088	396.5658	0.0000
θ_3	0.444119	0.000678	654.5801	0.0000
θ_4	-0.561020	0.003247	-172.7951	0.0000
θ_5	-0.246137	0.001696	-145.1643	0.0000
θ_6	-0.893253	0.002109	-423.5668	0.0000
θ_7	-0.433196	0.001648	-262.8495	0.0000
ω	0.600119	0.070767	8.480177	0.0000
β	0.574302	0.028872	19.89165	0.0000
γ	0.131857	0.025791	5.112610	0.0000
α	0.659645	0.039072	16.88299	0.0000

Figure 23. Residuals of ARMA(10,7)-EGARCH(1,1) process

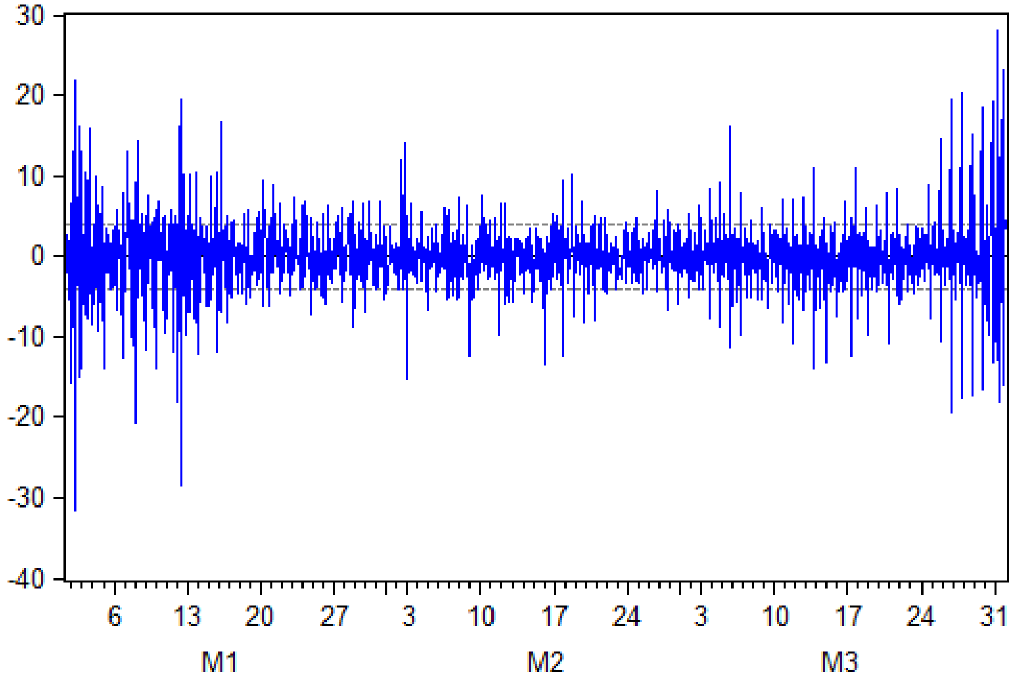


Figure 24. *acf for residuals of ARMA(10,7)-EGARCH(1,1) process with confidence intervals*

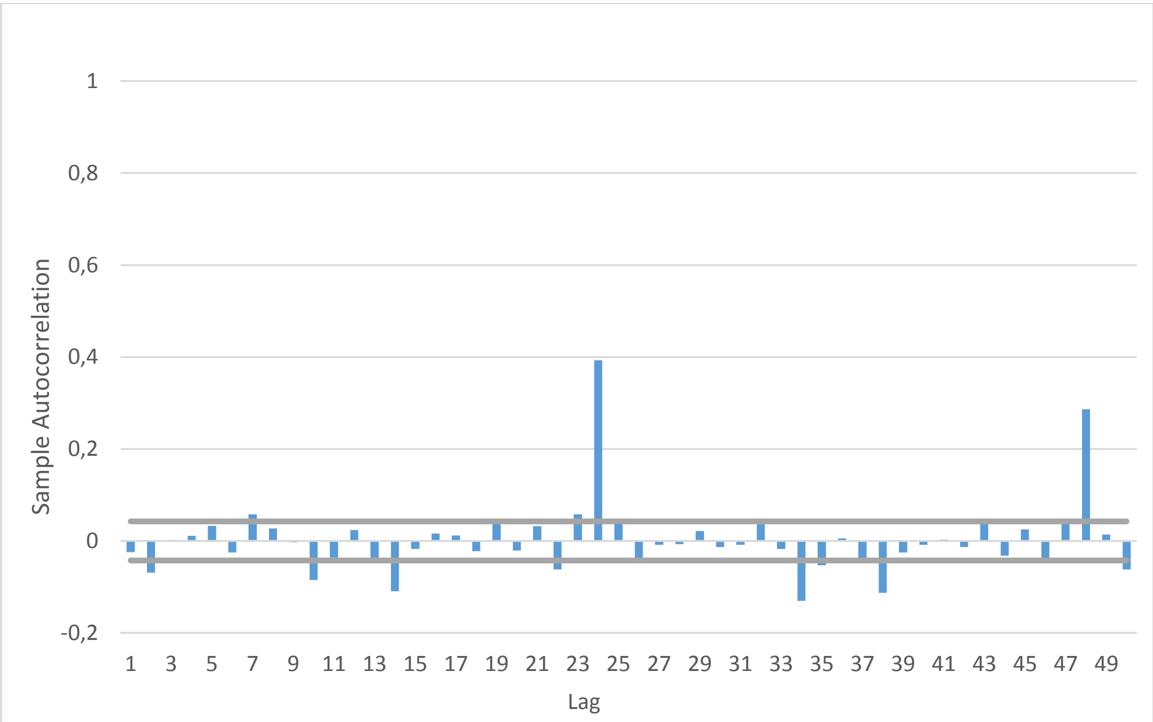


Figure 25. pacf for residuals of ARMA(10,7)-EGARCH(1,1) process with confidence intervals

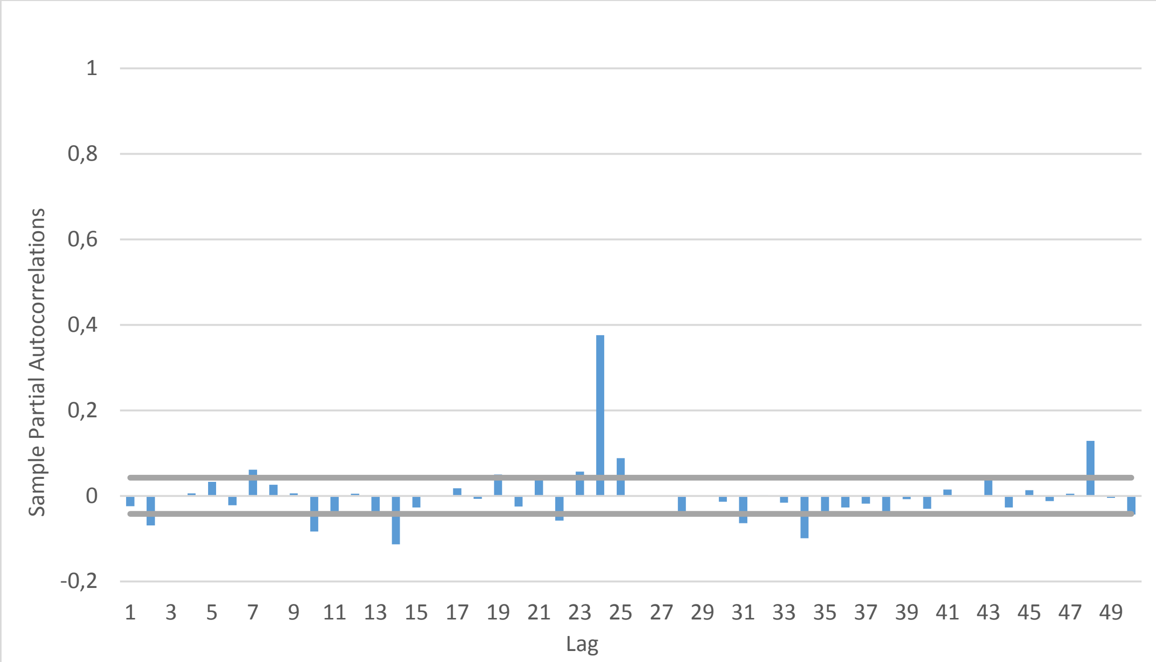
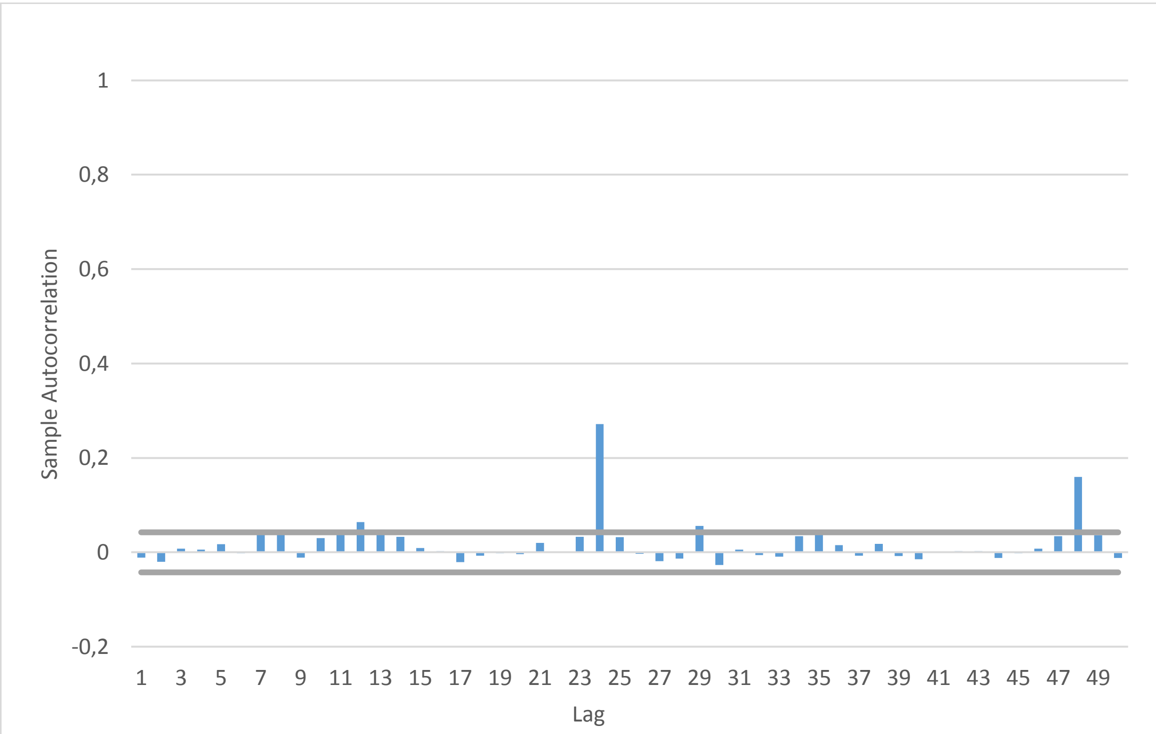


Figure 26. acf for squared residuals of ARMA(10,7)-EGARCH(1,1) process with confidence intervals



Unlike in the ARMA-GARCH and ARMA-GJR-GARCH models, the sample acf and pacf for residuals in the ARMA-EGARCH model show that the residuals at their early lags are not within the 95% confidence interval, indicating that the series is autocorrelated in its levels. However, the autocorrelations of the squared residuals seem not to be autocorrelated until the 11th lag. We further formally check the presence of autocorrelations of the model with a Ljung Box test for residuals and squared residuals.

Table 17. Ljung Box test for residuals of ARMA(10,7)-EGARCH(1,1) model

Lag	AC coefficient	Q-statistics	Critical value (at 5%)
1	-0.024	1.2389	3.841
2	-0.069	11.455	5.991

Table 18. Ljung Box test for squared residuals of ARMA(10,7)-EGARCH(1,1) model

Lag	AC coeffcient	Q-statistics	Critical value (at 5%)
5	0.017	1.9487	11.070
11	0.037	15.134	19.675
12	0.064	23.991	21.026

The Ljung Box test confirms observations from the plots that the residuals of the series are autocorrelated in its levels, and that variances are not autocorrelated within 11 lags.

In the next section, we forecast the proposed ARMA-GARCH, ARMA-GJR-GARCH and ARMA-EGARCH models and compare their results.

4.8 Forecasting

In our empirical analysis, we forecast the proposed ARMA-GARCH model and its two extensions for electricity prices using two different methods, i.e. the dynamic forecast and static forecast. The dynamic forecast uses rolling step-ahead forecasts for the lagged dependent variables, whereas the static forecast uses the actual values for the lagged dependent variables. With the dynamic forecast, we produce a series of rolling step-ahead forecasts to predict day-ahead electricity prices for the interval period of the 24-hour observations for three weekdays of 17th, 21st and 25th March 2014. With the static forecast, we produce a sequence of one-step-ahead forecasts, which roll the sample forward one observation after each forecast and use the actual rather than forecasted values for lagged dependent variables. Using this method, we can produce a longer out-of-sample forecast as the dynamic forecast quickly converges upon the long-term unconditional mean value. Thus, using the static method, we forecast the interval period of 408 hourly observations from March 15th to March 31st. These forecasts are generated using the sample of 1752 estimated observations from January 1st to March 14th in order to produce the forecasts from the estimated parameters for the remaining 408 sample observations. Following this, we evaluate the accuracy of the generated out of sample forecasts by comparing their values with the actual values in the series.

The following figures are graphical representation of performance for the dynamic forecasts of the deseasonalised electricity prices produced by ARMA(10-7)-GARCH(1,1), ARMA(10-7)-GJR-GARCH(1,1) and ARMA(10-7)-EGARCH(1,1) models for March 17th, 21st and 25th.

Figure 27. Forecast of ARMA(10,7)-GARCH(1,1) vs. the deseasonalised series for March 17th 2014

Figure 28. Forecast of ARMA(10,7)-GJR-GARCH(1,1) vs. the deseasonalised series for March 17th 2014

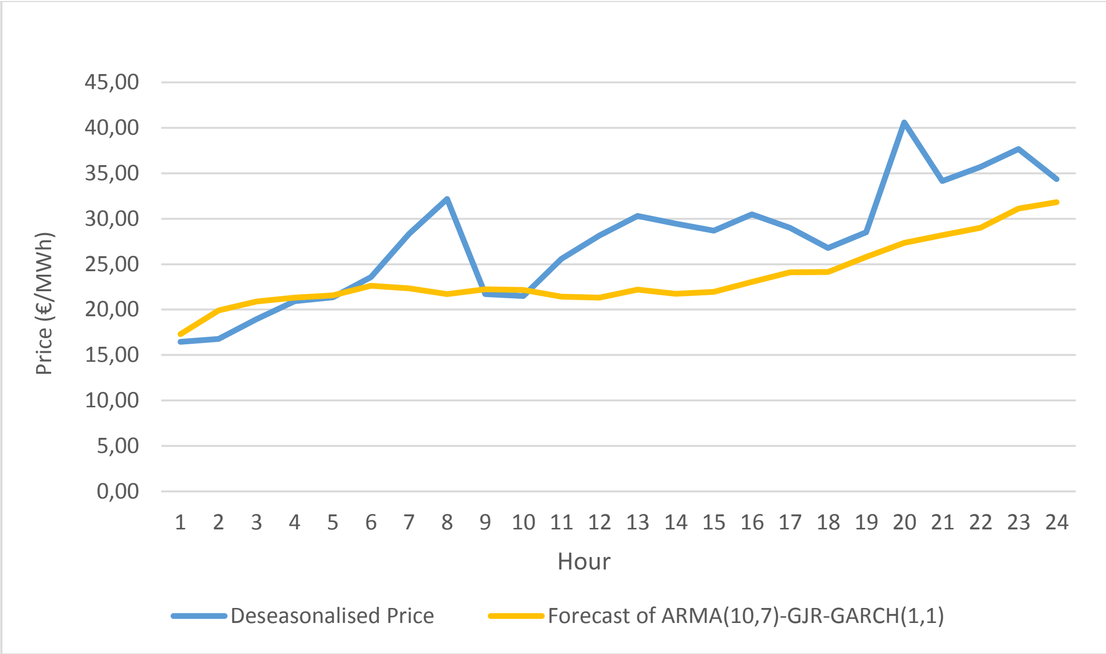


Figure 29. Forecast of ARMA(10,7)-EGARCH(1,1) vs. the deseasonalised series for March 17th 2014

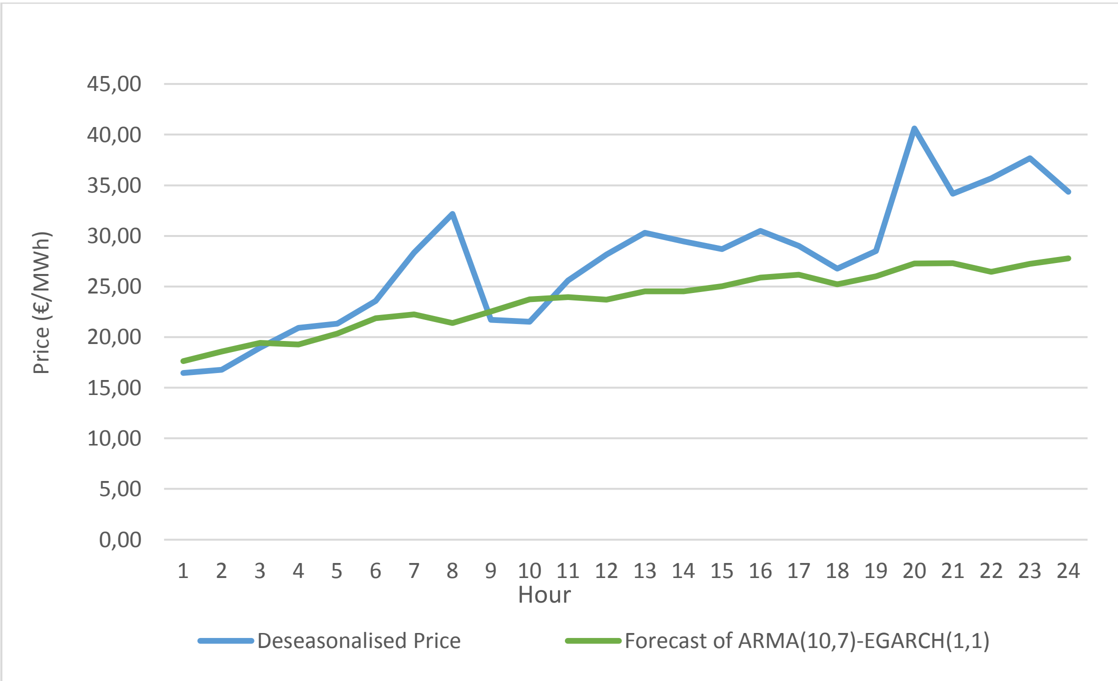


Figure 30. Forecast of ARMA(10,7)-GARCH(1,1) vs. the deseasonalised series for March 21st 2014

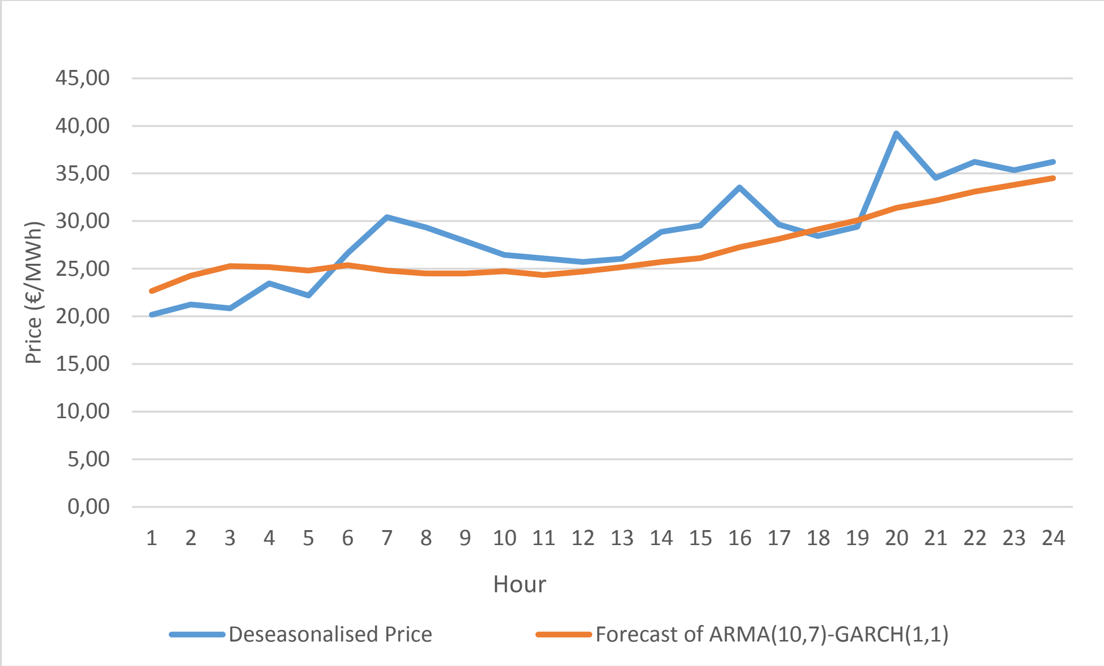


Figure 31. Forecast of ARMA(10,7)-GJR-GARCH(1,1) vs. the deseasonalised series for March 21st 2014

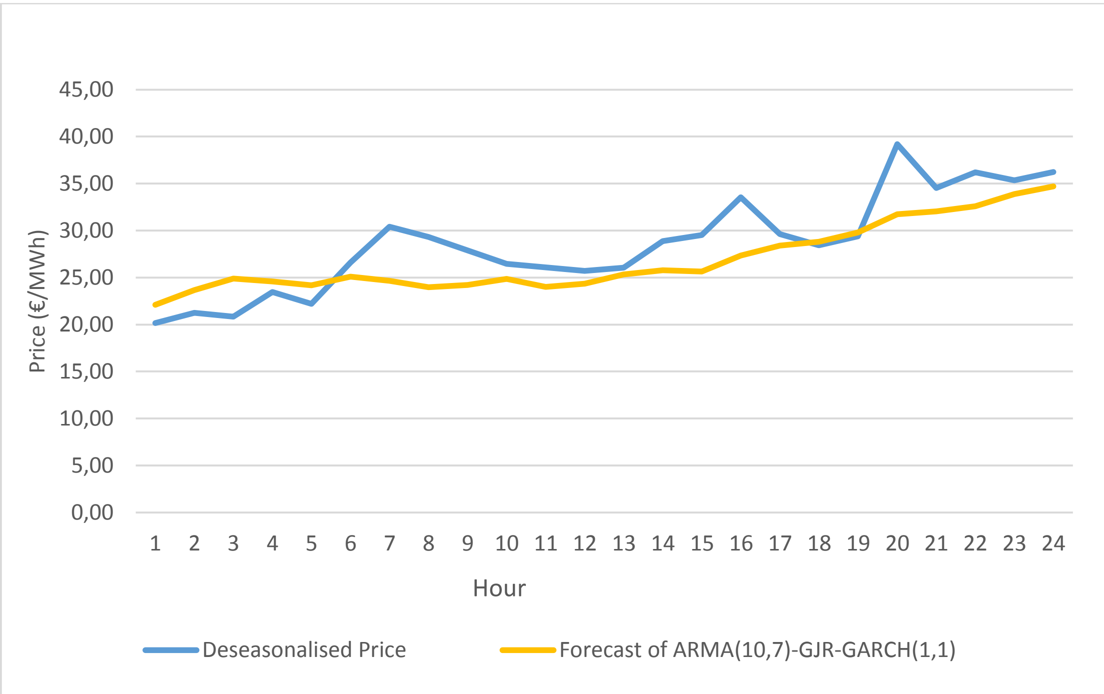


Figure 32. Forecast of ARMA(10,7)-EGARCH(1,1) vs. the deseasonalised series for March 21st 2014

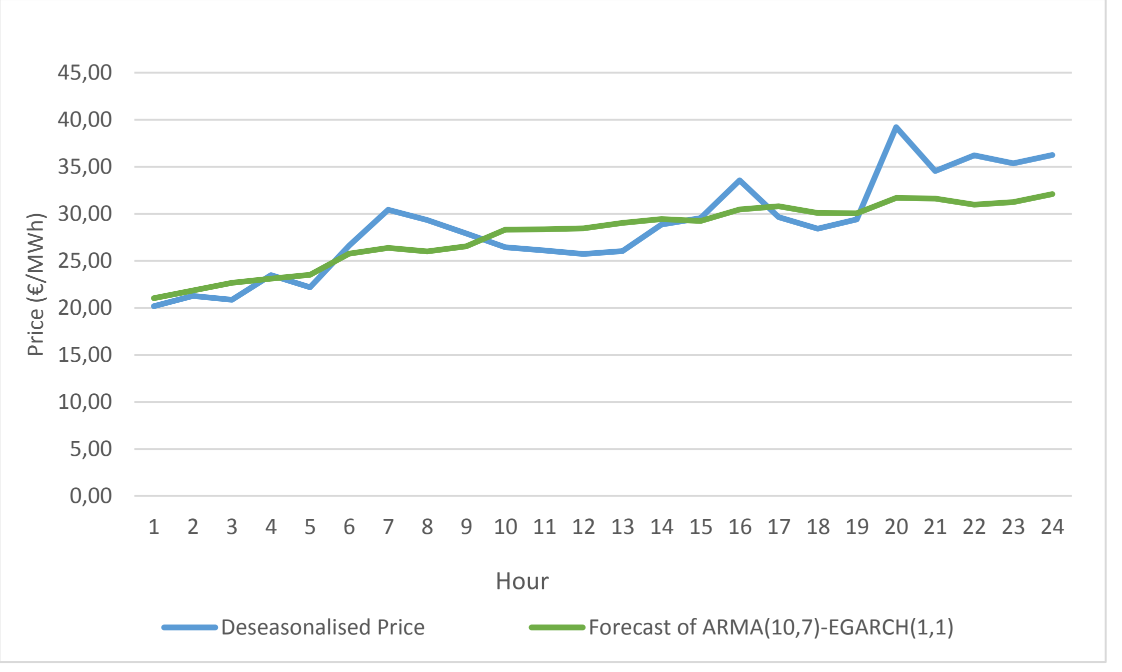


Figure 33. Forecast of ARMA(10,7)-GARCH(1,1) vs. the deseasonalised series for March 25th 2014

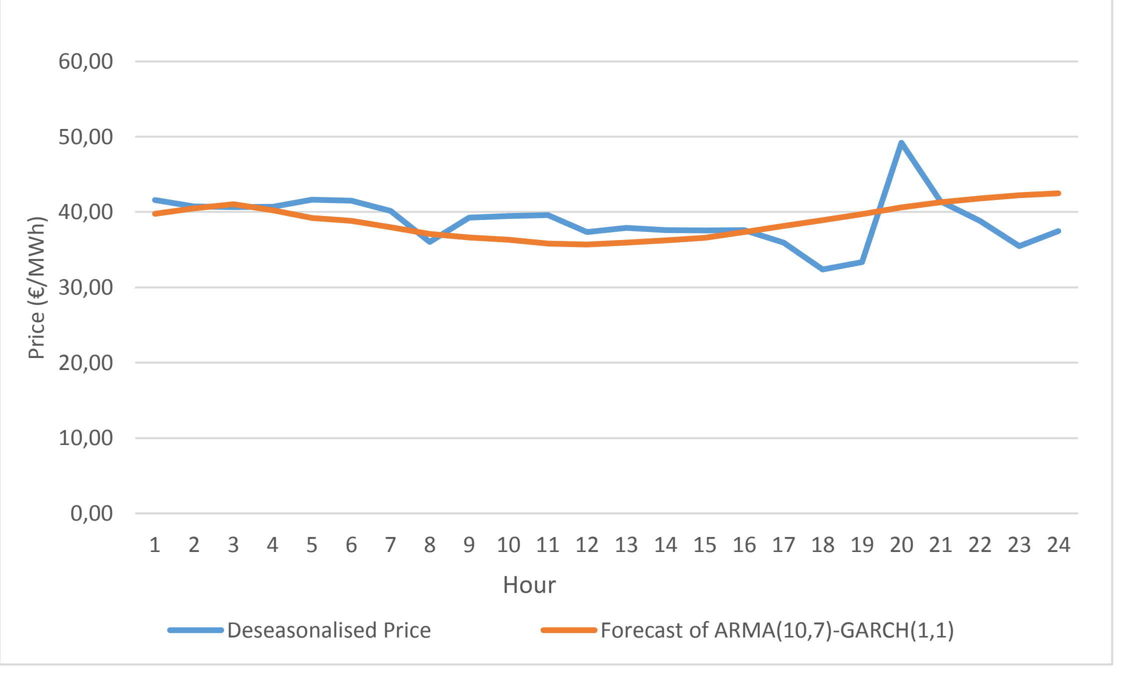


Figure 34. Forecast of ARMA(10,7)-GJR-GARCH(1,1) vs. the deseasonalised series for March 25th 2014

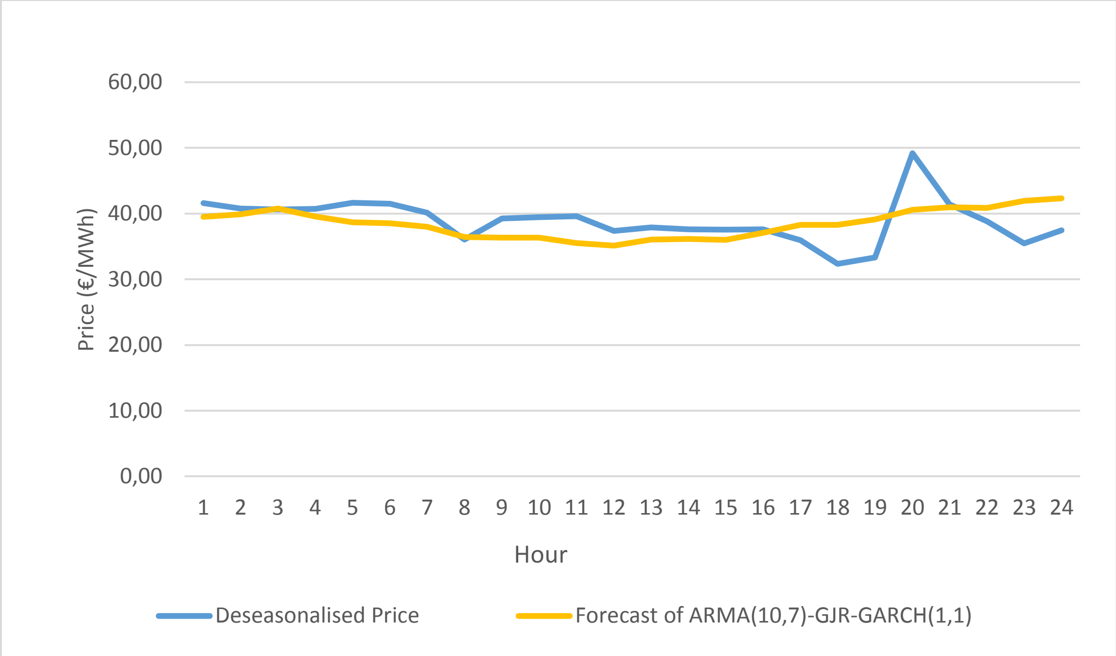
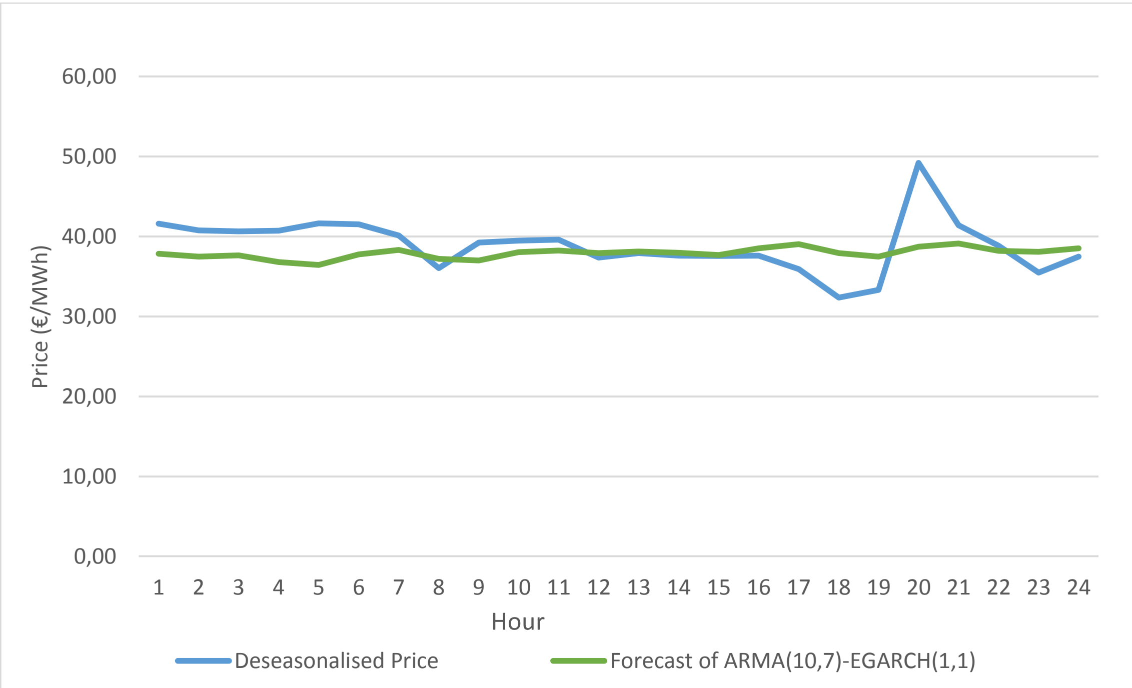


Figure 35. Forecast of ARMA(10,7)-EGARCH(1,1) vs. the deseasonalised series for March 25th 2014



Examining the above forecasts for three different days, we can see that the dynamic forecasts of all three models closely follow the deseasonalised prices; however, the forecasts cannot explain all of the price jumps, particularly at the final parts of the forecasted horizon, perhaps indicating that the forecasts are more accurate for shorter forecasting periods. In the following tables, we compare the average forecasted errors of the proposed models.

Table 19. Errors of the forecasted models for March 17th, 21st and 25th

Date	ARMA(10,7)-GARCH(1,1)	ARMA(10,7)-GJR-GARCH(1,1)	ARMA(10,7)-EGARCH(1,1)
March 17th	15.98%	15.49%	14.38%
March 21st	9.78%	9.43%	7.63%
March 25th	7.17%	7.27%	6.62%

The above table indicates that the ARMA(10,7)-EGARCH(1,1) process outperforms both ARMA(10,7)-GARCH(1,1) and ARMA(10,7)-GJR-GARCH(1,1) models. It also shows that all models perform better when there is less volatility in the series. Tables 20, 21 and 22 show the estimated forecasts and their errors for each of the forecasted observations during March 17th, 21st and 25th.

Table 20. Forecasted values and errors of the forecasted models for March 17th.

Hour	Deseasonalised price	ARMA(10-7)-GARCH(1,1)		ARMA(10-7)-GJR-GARCH(1,1)		ARMA(10-7)-EGARCH(1,1)	
		Forecast	Error (%)	Forecast	Error (%)	Forecast	Error (%)
00–01	41.59	39.77	4.37	39.48	5.07	37.85	8.99
01–02	40.76	40.50	0.62	39.89	2.12	37.47	8.06
02–03	40.64	41.03	0.95	40.76	0.29	37.64	7.38
03–04	40.72	40.21	1.24	39.55	2.87	36.81	9.60
04–05	41.64	39.23	5.78	38.67	7.13	36.43	12.50
05–06	41.50	38.84	6.42	38.51	7.21	37.75	9.05
06–07	40.13	37.98	5.37	37.98	5.36	38.32	4.52
07–08	36.02	37.07	2.91	36.46	1.20	37.20	3.26
08–09	39.26	36.60	6.77	36.33	7.45	36.99	5.77
09–10	39.48	36.31	8.02	36.35	7.94	38.05	3.61
10–11	39.61	35.80	9.62	35.54	10.29	38.26	3.42

(table continues)

(continue)

11–12	37.35	35.69	4.45	35.11	5.98	37.92	1.53
12–13	37.91	35.94	5.20	36.05	4.91	38.12	0.56
13–14	37.60	36.23	3.66	36.13	3.93	37.96	0.95
14–15	37.57	36.58	2.63	36.01	4.15	37.68	0.29
15–16	37.60	37.35	0.65	37.05	1.45	38.53	2.47
16–17	35.94	38.15	6.16	38.28	6.53	39.04	8.64
17–18	32.35	38.90	20.25	38.30	18.39	37.91	17.18
18–19	33.34	39.72	19.15	39.12	17.35	37.48	12.44
19–20	49.19	40.63	17.41	40.55	17.57	38.73	21.28
20–21	41.40	41.28	0.31	40.98	1.03	39.10	5.56
21–22	38.83	41.81	7.68	40.89	5.30	38.19	1.65
22–23	35.48	42.23	19.05	41.93	18.20	38.09	7.36
23–24	37.48	42.47	13.32	42.31	12.89	38.53	2.80

Table 21. Forecasted values and errors of the forecasted models for March 21st.

Hour	Deseasonalised price	ARMA(10-7)-GARCH(1,1)		ARMA(10-7)-GJR-GARCH(1,1)		ARMA(10-7)-EGARCH(1,1)	
		Forecast	Error (%)	Forecast	Error (%)	Forecast	Error (%)
00–01	20.16	22.67	12.43	22.09	9.56	21.03	4.32
01–02	21.25	24.26	14.13	23.65	11.28	21.84	2.78
02–03	20.84	25.28	21.32	24.90	19.47	22.65	8.69
03–04	23.47	25.18	7.29	24.58	4.73	23.09	1.61
04–05	22.19	24.80	11.75	24.19	9.01	23.53	6.01
05–06	26.61	25.38	4.61	25.09	5.71	25.76	3.20
06–07	30.42	24.81	18.41	24.66	18.93	26.37	13.29
07–08	29.34	24.50	16.48	23.97	18.29	26.00	11.38
08–09	27.90	24.52	12.13	24.21	13.25	26.56	4.82
09–10	26.45	24.74	6.46	24.86	6.01	28.33	7.11
10–11	26.10	24.35	6.70	23.99	8.06	28.36	8.65
11–12	25.71	24.71	3.90	24.34	5.33	28.47	10.71
12–13	26.04	25.17	3.36	25.34	2.70	29.03	11.46
13–14	28.87	25.71	10.93	25.77	10.73	29.44	1.98
14–15	29.53	26.13	11.53	25.65	13.15	29.25	0.95

(table continues)

(continue)

15–16	33.55	27.27	18.73	27.34	18.50	30.46	9.23

16–17	29.64	28.12	5.14	28.40	4.19	30.79	3.86
17–18	28.42	29.13	2.50	28.82	1.38	30.07	5.79
18–19	29.40	30.09	2.34	29.79	1.30	30.05	2.21
19–20	39.20	31.38	19.94	31.76	18.99	31.69	19.16
20–21	34.55	32.16	6.92	32.04	7.27	31.63	8.45
21–22	36.21	33.11	8.55	32.59	9.99	30.97	14.48
22–23	35.37	33.81	4.41	33.87	4.23	31.24	11.67
23–24	36.24	34.51	4.75	34.71	4.21	32.10	11.42

Table 22. Forecasted values and errors of the forecasted models for March 25th.

Hour	Deseasonalised price	ARMA(10-7)-GARCH(1,1)		ARMA(10-7)-GJR-GARCH(1,1)		ARMA(10-7)-EGARCH(1,1)	
		Forecast	Error (%)	Forecast	Error (%)	Forecast	Error (%)
00–01	41.59	39.77	4.37	39.48	5.07	37.85	8.99
01–02	40.76	40.50	0.62	39.89	2.12	37.47	8.06
02–03	40.64	41.03	0.95	40.76	0.29	37.64	7.38
03–04	40.72	40.21	1.24	39.55	2.87	36.81	9.60
04–05	41.64	39.23	5.78	38.67	7.13	36.43	12.50
05–06	41.50	38.84	6.42	38.51	7.21	37.75	9.05
06–07	40.13	37.98	5.37	37.98	5.36	38.32	4.52
07–08	36.02	37.07	2.91	36.46	1.20	37.20	3.26
08–09	39.26	36.60	6.77	36.33	7.45	36.99	5.77
09–10	39.48	36.31	8.02	36.35	7.94	38.05	3.61
10–11	39.61	35.80	9.62	35.54	10.29	38.26	3.42
11–12	37.35	35.69	4.45	35.11	5.98	37.92	1.53
12–13	37.91	35.94	5.20	36.05	4.91	38.12	0.56
13–14	37.60	36.23	3.66	36.13	3.93	37.96	0.95
14–15	37.57	36.58	2.63	36.01	4.15	37.68	0.29
15–16	37.60	37.35	0.65	37.05	1.45	38.53	2.47
16–17	35.94	38.15	6.16	38.28	6.53	39.04	8.64
17–18	32.35	38.90	20.25	38.30	18.39	37.91	17.18
18–19	33.34	39.72	19.15	39.12	17.35	37.48	12.44
19–20	49.19	40.63	17.41	40.55	17.57	38.73	21.28

(table continues)

(continue)

20–21	41.40	41.28	0.31	40.98	1.03	39.10	5.56

21–22	38.83	41.81	7.68	40.89	5.30	38.19	1.65
22–23	35.48	42.23	19.05	41.93	18.20	38.09	7.36
23–24	37.48	42.47	13.32	42.31	12.89	38.53	2.80

We now model the static forecasts for all of the proposed models for the period of March 15th to March 31st. The following plots show the dynamics of the static forecasts.

Figure 36. Forecast of ARMA(10-7)-GARCH(1,1) process vs. the deseasonalised price series for the period of March 15–31.

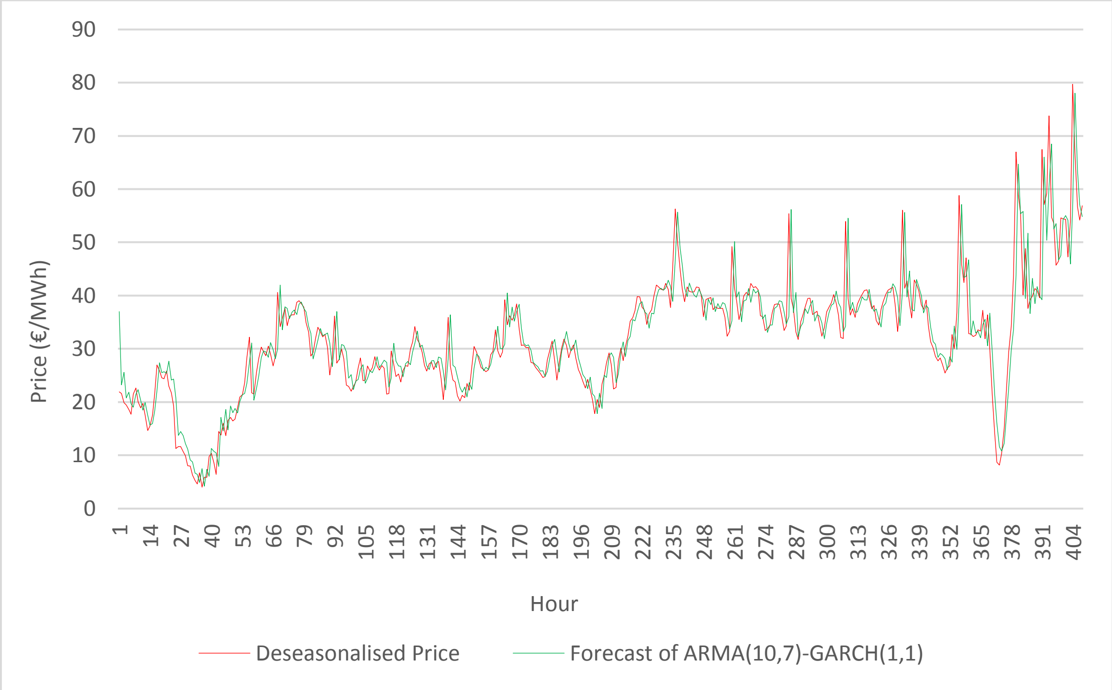
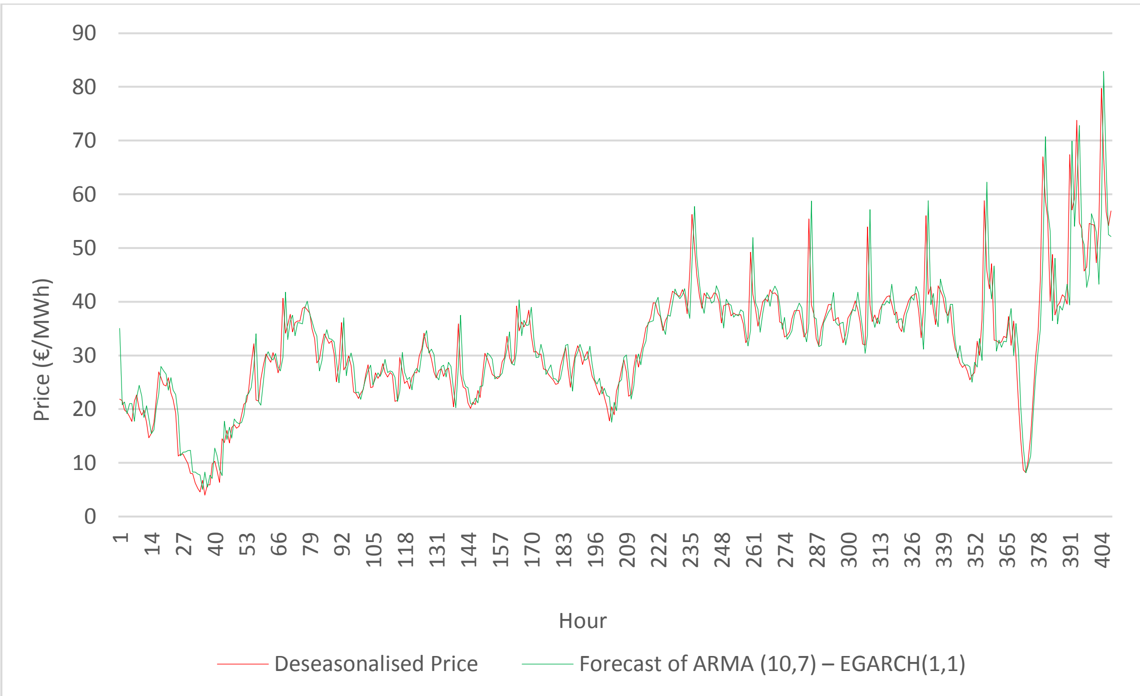


Figure 37. Forecast of ARMA(10-7)-GJR-GARCH(1,1) process vs. the deseasonalised price series for the period of March 15–31.

Figure 38. Forecast of ARMA(10-7)-EGARCH(1,1) process vs. the deseasonalised price series for the period of March 15–31.



Observing the above figures, the plots of forecasts indicate that the generated forecasts are quite close to the actual deseasonalised values. Since these are a series of multiple rolling step-ahead forecasts for the conditional variance, they show much more volatility than for the dynamic procedure. The forecasts lag a little behind their actual values; however, all the models seem to accurately follow volatility peaks. In Table 19, we compare the forecasted errors of the three models for the whole forecasted period.

Table 23. Errors of the forecasted models for the period of March 15–31.

	ARMA(10-7)-GARCH(1,1)	ARMA(10-7)-GJR-GARCH(1,1)	ARMA(10-7)-EGARCH(1,1)
MSE	4.886387	4.807610	5.005760
MAE	3.070867	3.089950	3.100570
MAPE	10.67626	10.81107	10.40616

Table 23 indicates a reasonably good performance of all models with a mean average percentage error of around 10% for all forecasts. In the following section, we examine the errors during the shorter 24-hour static forecast periods and compare forecasted prices to the actual deseasonalised prices hour-by-hour during a period of lower prices and volatility (March 19th) and a period of higher volatility (March 24th). Tables 24 and 25 show the errors during these two periods.

Table 24. Errors of the forecasted models during the period of lower volatility on March 19

Hour	Deseasonalised price	ARMA(10-7)-GARCH(1,1)		ARMA(10-7)-GJR-GARCH(1,1)		ARMA(10-7)-EGARCH(1,1)	
		Forecast	Error (%)	Forecast	Error (%)	Forecast	Error (%)
00–01	23.19	29.44	26.98	28.86	24.46	28.24	21.80
01–02	22.93	24.46	6.70	24.61	7.32	23.06	0.59
02–03	22.00	25.17	14.44	25.15	14.32	23.16	5.27
03–04	22.94	22.31	2.72	22.08	3.75	21.83	4.84
04–05	23.68	23.86	0.77	23.80	0.50	23.86	0.74
05–06	26.49	24.26	8.44	24.38	8.00	25.81	2.58
06–07	28.24	26.24	7.09	26.18	7.31	27.14	3.91
07–08	24.04	27.02	12.42	26.81	11.56	28.20	17.33
08–09	24.14	23.51	2.63	23.98	0.69	24.50	1.49

(table continues)

(continue)

09–10	26.79	24.42	8.85	24.50	8.54	25.69	4.09
10–11	25.76	25.89	0.51	25.70	0.21	26.73	3.80
11–12	26.44	25.52	3.49	25.63	3.06	26.11	1.25
12–13	28.51	26.58	6.76	26.98	5.36	27.29	4.26
13–14	26.80	28.54	6.49	28.26	5.44	29.25	9.16
14–15	26.00	26.44	1.69	26.32	1.23	26.80	3.07
15–16	26.82	27.11	1.08	27.68	3.18	27.13	1.15
16–17	26.38	27.86	5.61	27.74	5.18	26.97	2.24
17–18	21.46	27.46	27.93	27.04	25.99	25.90	20.67
18–19	21.61	22.78	5.44	22.95	6.22	21.45	0.75
19–20	29.62	24.47	17.38	24.76	16.42	23.29	21.38
20–21	27.04	31.04	14.79	30.48	12.70	30.55	12.98
21–22	24.82	27.65	11.40	27.60	11.22	26.88	8.30
22–23	25.25	26.75	5.93	26.87	6.41	25.52	1.09
23–24	23.78	26.69	12.26	26.71	12.34	25.95	9.14

Analysing the forecasts for the two 24-hour periods of March 19th and March 24th reveals that the proposed models using the static forecasting method perform better during the hours of higher volatility (March 24th) with an average error value of 6.04%, 5.95% and 5.32% for ARMA(10-7)-GARCH(1,1), ARMA(10-7)-GJR-GARCH(1,1) and ARMA(10-7)-EGARCH(1,1), respectively, while the average error values for the March 19th forecasts are a bit higher at 8.82%. 8.39% and 6.75% for ARMA(10-7)-GARCH(1,1), ARMA(10-7)-GJR-GARCH(1,1) and ARMA(10-7)-EGARCH(1,1). Therefore, we can conclude that the ARMA-EGARCH model, like in the case of the dynamic forecasts, again outperforms basic ARMA-GARCH and ARMA-GJR-GARCH models by a small margin, during both estimated periods of low and high volatility, as well as during the entire out-of-sample forecasted period.

Table 25. Errors of the forecasted models during the period of higher volatility on March 24

Hour	Deseasonalised price	ARMA(10-7)-GARCH(1,1)		ARMA(10-7)-GJR-GARCH(1,1)		ARMA(10-7)-EGARCH(1,1)	
		Forecast	Error (%)	Forecast	Error (%)	Forecast	Error (%)
00–01	35.16	32.31	8.10	31.77	9.65	32.63	7.20
01–02	35.82	35.46	1.03	35.73	0.26	36.13	0.84
02–03	37.14	35.27	5.03	35.66	3.98	36.31	2.23
03–04	39.78	36.86	7.35	36.36	8.60	36.47	8.33
04–05	39.77	38.46	3.30	38.46	3.28	39.68	0.24
05–06	37.83	38.85	2.69	39.53	4.50	40.84	7.97
06–07	37.24	36.91	0.90	36.77	1.25	37.55	0.82
07–08	34.60	36.56	5.66	36.06	4.22	35.31	2.04
08–09	36.55	33.86	7.36	34.45	5.73	33.88	7.30
09–10	37.32	36.66	1.79	36.94	1.03	37.57	0.66
10–11	39.89	36.58	8.29	36.04	9.65	37.28	6.55
11–12	41.98	39.63	5.59	39.46	6.00	39.75	5.31
12–13	41.60	41.05	1.32	41.92	0.76	42.38	1.86
13–14	41.18	41.39	0.52	41.02	0.38	41.19	0.03
14–15	41.03	40.99	0.09	40.60	1.04	40.57	1.11
15–16	42.25	41.43	1.95	41.90	0.83	41.25	2.36
16–17	41.14	42.88	4.23	43.18	4.97	42.44	3.17
17–18	37.78	41.74	10.48	40.90	8.25	39.38	4.23
18–19	45.54	38.88	14.63	38.91	14.57	36.91	18.97
19–20	56.25	46.85	16.71	47.27	15.96	47.28	15.93
20–21	50.06	55.69	11.25	55.52	10.90	57.74	15.35
21–22	44.91	48.95	8.99	48.25	7.43	48.34	7.65
22–23	41.15	45.63	10.88	46.12	12.08	43.43	5.55
23–24	38.82	41.50	6.91	41.73	7.49	39.57	1.92

CONCLUSION

Since the mid-1990s, electricity markets have been deregulated in order to reduce electricity prices and improve efficiency by determining prices based on the behaviour of supply and demand forces on the markets. Electricity is uniquely characterized by its non-storability, particularly reflected in its high volatility, which makes it significantly different from other commodities traded on financial exchanges. Other factors that strongly influence electricity prices are the fact that transmission system networks frequently exhibit irregularities, which are reflected in geographical price variations across different areas due to different maintenance and transmission costs. Consequently, the relationship between consumption in the market and electricity prices is extremely difficult to predict. Consumption, though less volatile in comparison to electricity spot prices, presents the same seasonal treatment, indicating that demand elasticity is low. Nonetheless, prices are significantly affected by the degree of consumption, which is seen in the critical cause of spikes in electricity price time series. Increases or sudden jumps in demand levels at certain moments force the producers and distributers of electricity to provide electricity from more expensive energy resources in the production of electricity. In other words, increases in electricity consumption are directly linked to the levels of electricity production, whose marginal costs rise depending on the use of energy inputs. Furthermore, deregulated electricity markets indicate various series of seasonality, due to external factors such as seasonal weather conditions, weekly seasonality due to differences in consumption during weekdays and weekends, as well as intra-daily periodic components due to the peak and off-peak phases of the day. Altogether, deregulation along with the abovementioned characteristics of electricity markets lead to a considerable growth of electricity price volatility, incomparable to any other financial assets or commodities, with an average intra-day volatility as high as 50% in our analysed time series of the Slovenian electricity prices for the first quarter of 2014.

Various studies have been developed to analyse and forecast electricity prices, reliant on the intended temporal horizons. For the purpose of this thesis, we focused on the determination of trends and forecasts of Slovenian electricity spot prices in the short run. As electricity spot prices showcase different forms of nonlinear dynamics, most notably the robust dependence of the variance of the series on its own previous values, we proposed three different stochastic volatility forecasting methods for the Slovenian hourly electricity prices traded on the BSP Southpool electricity exchange, based on the family of autoregressive conditional heteroscedastic models. To analyse the volatility in a price heteroscedasticity framework, we modelled short-term forecasts with a GARCH model, along with its two extensions that measure asymmetric effects: the EGARCH and GJR-GARCH models. Each of these models was processed simultaneously with an ARMA model specified for the conditional mean equation. Modelling and forecasting a pure ARMA model could represent a risk of higher forecasting errors as the volatility in our time series is not independent, though our estimated ARMA model can successfully decompose AR and MA processes, as well as obtain shortly stationary residuals. Therefore, in order to overcome this, we proposed mixed ARMA-GARCH, ARMA-GJR-GARCH and ARMA-EGARCH models that can address volatility clustering, and hence outperform a pure ARMA model in forecasting. Before identifying the most suitable ARMA orders, we first deseasonalised the series for its daily and weekly periodic seasonal components. Once the series had been deseasonalised, we choose an optimal order for an ARMA model with different information criteria functions. Following with the estimation for parameters of both conditional mean and conditional variance equations, we validated the robustness of parameters using various statistical significance and hypotheses tests, such as the Ljung Box test, to test the stationarity of residuals. Once the models were validated, we began forecasting the proposed models for day-ahead electricity prices with dynamic forecasting for March 17th, 21st and 25th as well as with the static forecasting for a time span of 17 days between March 15th and March 31st. In both cases, the ARMA-EGARCH model slightly outperformed both ARMA-GARCH and ARMA-GJR-GARCH. The output of the parameters of the ARMA-EGARCH model also reveals that a positive shock to the series implies a lower next-period conditional variance than negative shocks, or in other words, that increasing electricity prices lead to a lower next-period volatility than when electricity prices decrease by the same amount.

REFERENCE LIST

1. Bollerslev, T. (1986). Generalised autoregressive conditional heteroscedasticity. *Journal of Econometrics*, *31*(3), 307–327.

2. Bollerslev, T. (1987). A Conditionally Heteroscedastic Time Series Model for Speculative Prices and Rates of Return. *Review of Economics and Statistics*, *69*(3), 542-547.

3. Box, G. E. P., & Jenkins, G. M. (1994). *Time Series Analysis: Forecasting and Control* (3rd ed.). Englewood Cliffs: Prentice Hall.

4. Branger, N., Reichmann, O., & Wobben, M. (2010). Pricing electricity derivatives on an hourly basis. *The Journal of Energy Markets, (3)*3, 51-89.

5. Brooks, C. (2002). *Introductory Econometrics for Finance.* Cambridge: Cambridge University Press.

6. Bunn, D. W. (2004). *Modeling prices in competitive electricity markets*. Hoboken: John Wiley & Sons.

7. Burger, M., Klar, B., Mueller, A., & Schindlmayr, G. (2004). A spot market model for pricing derivatives in electricity markets. *Journal of Quantitative Finance, (4)*1, 109-122.

8. Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The Econometrics of Financial Markets.* Princeton: Princeton University press.

9. *Day-ahead trading results*. (2014). Retrived April 17th, 2014, from http://www.bsp-southpool.com/day-ahead-trading-results-si.html

10. Deng, S. J. (2000). *Stochastic models of energy commodity prices and their applications: Mean reversion with jumps and spikes* (Working paper). Atlanta: Georgia Institute of Technology.

11. Deng, S. J., & Oren, S. S. (2006). *Electricity derivatives and risk management* (Working paper). Atlanta: Georgia Institute of Technology.

12. Dong, Z., & Zhang, P. (2010). *Emerging Techniques in Power System Analysis.* Berlin: Springer.

13. Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, *50*(4), 987-1008.

14. Escribano, A., Peña, J. I., & Villaplana, P., (2002). "*Modelling Electricity Prices: International Evidence* (Working Paper). Madrid: Universidad Carlos III de Madrid.

15. Eydeland, A., & Geman, H. (1999). *Fundamentals of Electricity Derivatives in Energy Modelling and the Management of Uncertainty*. London: Risk books.

16. Eydeland, A., & Wolyniec, K. (2003). *Energy and Power Risk Management*. Hoboken: John Wiley & Sons.

17. Faruqui, A., & Eakin, K. (2000). *Pricing in competitive electricity markets.* Berlin: Springer.

18. Francq, C., & Zakoian, J. M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli, (10)*4, 605-637.

19. Garcia, R. C., Contreras, J., Akeren, M. & Garcia, J. B. C. (2005). A GARCH model to predict a day-ahead electricity prices. *IEEE Transactions on Power Systems, 20*(2), 867–874.

20. Geman, H. (2002). *Towards a European Market of Electricity: Spot and Derivatives Trading* (Working paper). Paris: University Paris IX Dauphine and ESSEC.

21. Gençay, R., Selçuk, F., & Ulugülyagci , A. (2001). EVIM: A Software Package for Extreme Value Analysis in MATLAB. *Studies in Nonlinear Dynamics and Econometrics, (5)*3, 1-29.

22. Greene, W. H. (2003). *Econometric Analysis.* Englewood Cliffs: Prentice Hall.

23. Gourieroux, C. (1997). *ARCH Models and Financial Applications*. Berlin: Springer.

24. Hull, J. C. (2010). *Risk Management and Financial Institutions* (2nd ed.). Englewood Cliffs: Prentice Hall.

25. Hull, J. C. (2012). *Options, Futures, And Other Derivatives* (8th ed.). Englewood Cliffs: Prentice Hall.

26. Janssen, M., & Wobben, M. (2009). Electricity pricing and market power: evidence from Germany. *European Transactions on Electrical Power, (19)*4, 591–611.

27. Kaminski, V. (1997). The Challenge of Pricing and Risk Managing Electricity Derivatives. In *The US Power Market* (pp. 149-171). London: Risk Publications.

28. Karakatsani, N. V., & Bunn, D. W. (2004). *Modeling the Volatility of Spot Electricity Prices.* London: London Business School.

29. Kluge, T. (2006). *Pricing Swing Options and other Electricity Derivatives.* Doctoral thesis, St. Hugh's College, University of Oxford, Oxford, United Kingdom.

30. Knittel, C., & Roberts, M. (2002). An Empirical Examination of Restructured Electricity Prices. *Energy Economics, (27)*5, 791-817.

31. Lucia, J., & Schwartz, E. (2002). Electricity prices and power derivatives: evidence from the Nordic power exchange. *Review of Derivatives Research, (5)*1, 5–50.

32. Naeem, M. (2010). *A comparison of electricity spot prices simulation using ARMA-GARCH and mean reverting models.* Master's thesis, Lappeenratanta University of technology, Department of Mathematics and Physics, Lappeenratanta, Finland.

33. Pilipovic, D. (1997). *Energy Risk: Valuing and Managing Energy Derivatives*. New York: McGraw-Hill.

34. Ptak, P., Jablonska, M., Habimana, D., & Kauranne, T. (2008). Reliability of ARMA and GARCH models of electricity spot market prices. *Proceedings of ESTSP 2008 European Symposium on Time Series Prediction.* Porvoo, Finland: Helsinki Institute of Technology.

35. Reider, R. (2009). *Volatility Forecastng I: GARCH models* (Working paper). New York: New York University.

36. Tsay, R. S. (1982). *Analysis of Financial Time Series.* Hoboken: John Wiley & Sons, Inc..

37.Woo, C., K., Karimov, R., & Horowitz I. (2004). Managing electricity procurement cost and risk by a local distribution company. *Energy Policy, (32)*5, 635-645.