ERASMUS UNIVERSITEIT ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS

THE IMPACT OF WIND POWER GENERATION ON THE ELECTRICITY DAY-AHEAD SPOT PRICE IN GERMANY AND IN THE NETHERLANDS: A FORECAST ANALYSIS

Master thesis

Candidate: Alberto Ferriolo
Student number: 469240
Thesis supervisor: Sjoerd van den Hauwe
Second reader: Dr. Jan Lemmen

ABSTRACT

The present research investigates the impact of wind power generation on the electricity day-ahead spot price by applying different forecast techniques in The Netherlands and in Germany, in order to understand the extensions and magnitude of wind power on the electricity price formation's process. In explaining its contents, the research sheds light on all the mechanisms and implications through which wind power generation might affect the electricity spot price and its volatility.

This paper proposes a combination of an autoregressive moving average (ARMA) and a Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models based on daily time-series

Key words: Electricity price, wind power generation, Forecast, Volatility, ARMA model, GARCH model, The Netherlands, Germany, Diebold-Mariano, MSE.

ACKNOWLEDGEMENT

Firstly, I would like to express my sincere gratitude to my supervisor Sjoerd van den Hauwe for the continuous support, aspiring guidance and patience in realizing my master thesis.
I thank all my fellow classmate and close friends for the stimulating discussion on relevant topics and for all the fun we had in the last year.
Last but not least I would thank my family; my parents and my sister for supporting me spiritually throughout all this long and insidious year. They always have been source of inspiration and motivation in tackling obstacles and achieving goals.

TABLE OF CONTENTS

Section 1: Introduction

The following research investigates the impact of the wind power generation on the electricity day-ahead spot price through a comparative forecast analysis between The Netherlands and Germany. In explaining its contents, the research sheds light on all the mechanism and implications through which wind power generation might affect the electricity spot price and its volatility.

Electricity price forecasting (hereafter EPF), due to the increased degree of deregulation of the power industry and the development of new renewable energy sources, has surely become an important area in the aftermath of worldwide. (Zhiwei Shena, 2015). The growing market share of wind power has in fact highlighted concerns and uncertainties about the predictability of the energy price. However, although this revolution towards more sustainable energy sources is helping reduce the quantity of C02 released in the atmosphere, on the other hand, it is posing increasing apprehensions on forecasting the electricity prices. Indeed, the development of renewable energy sources (hereafter RES) have galvanized many researchers and scholars, over the years, to carry out studies on how renewables could mitigate existent problem and interact with the well-established electricity system. The diatribe mainly involves the trade-off between the level of electricity price and its volatility.

The wind production, indeed, equally to all intermittent sources, is not adjustable to peak loads and therefore the variability in production can be seen as a volatility in price. Past empirical works have in fact demonstrated that, although the low marginal costs of production[1] reduce the electricity price overall, as confirmed by the merit order theory[2], clean energy generation might increase its volatility, making the price more unpredictable. Among all energy sources, in fact, wind and solar energy have the lower operational costs and they are dispatched on a legal priority basis with respect to the generation of not clean energy sources.

---

[1] Hydropower, photovoltaic solar and wind energy plants have less than 0.01 cent as operating cost (U.S Energy Information Administration, 2018).

[2] The merit order ranks available sources of energy (in particular electricity generation) based on their marginal costs in the short-run in combination with the amount of electricity that can be generated. Hence, sources with the lowest marginal cost are the first to be supplied on the market. This manner of dispatching the generators reduces the cost of electricity production significantly (Georg Wolff, 2017)

In this scenario, countries like The Netherlands and Germany, are facing numerous challenges along the mushrooming market share of wind power as well as the deregulation of the market.

Motivated by this background, the first goal of this paper is to evaluate the direct impact of wind power on the electricity day-ahead spot price in The Netherlands and in Germany by using a different approach based on the Mean Square Error (hereafter MSE approach). By looking at the differences in MSE[3], in fact, the divergence in accuracy between the two forecast models can be inferred, capturing the extent by which wind generation affects the electricity spot price. Consequently, two different forecast models, one autoregressive and one conditional to the wind power, are elaborated for both The Netherlands and Germany. The MSE from the autoregressive benchmark model[4] (in which the forecasted values are predicted using past and observed data) are contrast and compare to those where the wind power generation factor is added as an explicative variable[5]. In this regard, the increase in accuracy from an autoregressive to a conditional forecast can be intent as the impact of the exogenous factor on the dependent variable[6]. In order to reach this out, the autoregressive moving average (ARMA) model and its extension (with regards of exogenous variable "ARMAX") are applied to the Dutch APX and German EPEX phelix spot market.

The second goal of the study involves, instead, a structural analysis. It aims to explore the linkage between the electricity price and wind power generation, providing a helicopter overview of how the electricity day-ahead spot market behaves in relationship to one of its most discussed driver, the wind power. In this regard, direct and delayed effect, structural breaks and volatility effect are studied.

---

[3] The mean square error reflects the standard deviation of the residual or prediction errors. Hereby it reflects how far are predicted values from the regression line. The approach consists of comparing MSE from a simple autoregressive forecast to the one in which wind power generation is added as exogenous. The difference in MSE should therefore reflect the impact of wind power generation on the electricity price formation's process.
[4] The benchmark model is based on an ARMA (p, q) in which, the only variable presented is the electricity day-ahead spot price. Future prices are therefore forecasted only using past values.
[5] The benchmark model is compared to the ARMAX (p, q) model in which, wind power generation is added as exogenous(X) factor. Future prices are therefore generated using both past prices and wind power values.
[6] The dependent variable is the daily electricity day-ahead spot price

Overall, the analysis sheds light on where and to what extent the "variable wind power generation" has a greater impact on forecasting the electricity day-ahead spot price as well as all its connections to the price and its volatility in both mentioned countries.
The research is therefore mainly divided in two sections, in which, respectively the two points of view are discussed and elaborated in all their extensions. The purpose of this paper is threefold. First of all, given the scarce literature on the electricity price behaviour in the Netherland, an investigation of the electricity price formation and behaviour in the mentioned country is performed. Secondly, differently from other scholars, a different approach is used to analyse and quantify the impact of wind power generation on the electricity price. Forecasting methods are in fact, widely used among commodity firms and energy providers but none of them have been applied to capture the behaviour of wind on the electricity market. Thirdly, all past researches focus on different frequency data[7], which could lead to differences in results.

The organization of the paper is the following. The second section introduces previous works, spanning from econometric models to considerations on the wind power generation and the changeover that the market is actually undergoing. In the third section, all the research hypotheses are developed, thus, the purpose of the research is described. The analysis therefore begins in the fourth section where the dataset and the time series are introduced. The fifth section presents the methodology followed by this paper, which opens to the unique characteristics of the electricity day-ahead spot market as well as what is needed to consider in modelling the electricity price before applying forecast techniques, such as stationarity, seasonality, serial correlation of residuals and consideration on outliers.

Additionally, all econometric models are introduced, providing features that differentiate them from the others. Consequently, the ARMA, ARMAX, GARCH and its extension AR-GARCH are discussed in all their characteristics.

[7] Electricity prices are usually dispatched to the energy markets with an hourly frequency. This research however aims to investigate the impact on a daily periodicity.

Thereafter, all the forecast techniques previously described are applied to the dataset in section six, where the best forecast model is detected in accordance with few selected "accuracy criteria", such as the AIC and BIC criterion[8]. The MSE approach is lately applied to detect the importance and the effect of wind power generation on both the EPEX-phelix and APX day-ahead spot price from a forecasting point of view. Sub-sequentially, it follows the structural analysis which encompasses, to recall, an investigation of the direct and delayed effect of the wind power generation on the electricity spot price, a study on structural breaks and a volatility analysis[9].

To conclude, findings and results will be evaluated and compared from country to country, shedding light on the differences and on the importance of wind power generation on the APX day-ahead and EPEX Phelix price formation's process as well as all directs and delayed connections that wind power has with the energy price and its implied volatility.

[8] Respectively the Akaike information criterion and the Bayesian Information criterion are criterions for model section among a finite set of models. It is generally accepted in the forecasting techniques literature that; the best model should have the lowest AIC and BIC parameters.

[9] The price's volatility is modelled by using a AR-GARCH (1,1) model in order to capture all the characteristics of its formation's process as well as its conditional dependence on past volatility values and relationship with the wind power generation.

Section 2: Literature review

Literature on renewable energy is shown very rich although different methodologies are applied. Although the forecast approach has not been studied by scholar to measure the effect of an exogenous variable on the electricity price formation[10], the intrinsic characteristics of electricity price have, over the past years, motivated many researchers to carry out studies on the linkage between renewable energy sources and the well-established electricity system. The following section is divided in two sub-sections: The impact of renewables on the energy price, more generally, and lately the models used by scholars to capture it.

2.1 The impact of renewables on the electricity price

Concerning the impact of renewables on the electricity price, it is argued that, despite the positive effects derived by an increase usage of RES on the energetic system such as reduction of C02 emission and the decline in the electricity price, clean energy could pose its own disadvantages. Indeed, owing to their intermittent nature[11], RES might increase the volatility of the electricity market questioning whether they are economically beneficial. Considering that, electricity so far is not economically feasible to be stored, different studies on their seasonality and flexibility in production have been carried out through the last 20 years.

The literature has shown that, in accordance with the merit order theory, wind power generation has a dampening effect on the electricity spot price. Notwithstanding, no one have explicitly modelled its impact on the price through a forecast approach. Various studies have demonstrated this dampening effect for wind electricity generation (Di Cosmo et al. 2012; Nicolosi, 2010). Nicholson et al. (2010) found that the merit order effect changes between day and night. Paraschiv et. al (2014), have studied the EEX day-ahead prices in Germany, particularly focused on renewable energy sources such as wind and photovoltaic. They argue that intermittent RES increases the price fluctuations, which means that the price sensitivity is higher. In a research conducted by Nicolosi et al. (2009) it is found that intermittent green energy increases volatility of the residual demand, which turns into an increase of volatility of

[10] In the literature, usually, the impact of wind power generation, and more generally of an exogenous variable, is calculated by using a linear regression.

[11] The intermittency of these energy is attributable to their dependence on unpredictable climate factors such as sunlight, wind speed and rain.

electricity price. Additionally, in Ketterer (2014) studies, the intermittent wind power generation has been found to do not only dampen the electricity spot price but also positively influence its volatility. Further studies on other RES, are provided by Huisman, et al. (2013) and Huisman et al. (2015) in which it is shown how an increase of renewable energy supply with a low marginal cost would impact the electricity price in the Nord Pool Market. Specifically, it has been demonstrated that the marginal cost of hydro production is directly related to the level of the reservoir[12]. Hydro plants are willing to sell electricity at any price instead of losing water, that would spill from an overflowing reservoir. Results show that, as the marginal cost of hydro production varies according to the reservoir levels, the water level in the reservoir is one of the price drivers.

It is concluded that, a higher level of water in the reservoirs would drive the electricity price down. However, although the hydropower sample help understanding the mechanism behind RES, wind energy possesses clear difference in margins and flexibility in production that might differentiate it. Hence, as hydropower depends on the level of reservoir, wind power generation depends on the wind speed, which cannot be controlled. Therefore, it emerges, that, while hydropower plants could be switched on and off whenever it is needed, wind turbines work intermittently based on the weather conditions. It appears clear now that, although all renewables are categorized under the same name, effects and implications of their generation might diverge from each other.

2.2 THE ECONOMETRIC MODELS

In the area of EPF, several studies have been applied to the energy market in order to understand the implication of the merit order effect. Researches have been conducted mainly using the ARMA models, and in some case a combination with a GARCH model to capture the volatility behaviour of the electricity price. Literature, however, is debating whether ARMA and ARMA-GARCH approaches have enough predictive power, being too restrictive to capture the non-linear behaviour of wind speed and electricity spot price. Notwithstanding, the ARMA's models are found to be preferred in forecasting the electricity spot price by many scholars. It is used in fact to present the stationary time series based on autoregressive and moving average process.

[12] Hydroelectric reservoirs are large hydric basins in which water is collected and stored to guarantee a continuous stream of production. Hydropower is considered to be more flexible than wind and solar production since, reservoirs act as batteries providing it the ability to be switch on/off at short notice.

The various models differ in the data frequency used (usually daily or hourly), time series transportation (usually logarithms of prices or log-returns), the treatment of seasonality and the techniques. A first understanding is provided by Jonsson (2008) where a two-step methodology for forecasting the electricity spot prices is introduced focusing on the impact of predicted system load and wind generation. The chosen model involves an ARMA-type process and Holt-Winters models to account for residual autocorrelation and seasonal dynamics. In Jakasa (2011), the importance and various techniques in forecasting day-ahead electricity prices are investigated. The author used European Energy Exchange (EEX) data as the reference power market. His analysis follows the Box-Jenkins method applied to the ARIMA model, defined as the best approach to deal with the intrinsic and unique characteristic of electricity. In Ketterer (2014), a GARCH model is used to model price and volatility of the energy German Market in an integrated approach.

The analysis is conducted using daily levels of German wind power generation as an explanatory variable in the mean and the variance equation of a GARCH model of the German day-ahead electricity price. Results showed that wind power generation decreases the price in Germany in the period from 2006 to 2011 but increases the price volatility. According to Ketterer (2014), the AR-GARCH is the most appropriate model which allow to explicitly test the effect of wind power generation on the mean and volatility of the electricity price in an integrated system. Results from her research poses that, intermittent energy sources such as wind generation increase the volatility of the electricity prices.

In Shen et al. (2015) different other volatility forecasting techniques, such as Markov regime-switching and GARCH extensions for wind power production are studied. In comparing different models, it has been found that, the MRS-GARCH model significantly outperforms traditional GARCH models in predicting the volatility of wind power. The generalized Autoregressive Conditional Heteroscedasticity process has been tested specially to simulate spikes and volatility. Electricity spot prices, however, presents as matter of fact severals empirical regularities which differentiate it from other financial assets and commodities such as seasonal cycles, mean reverson and price spikes. On the light of this, in Cuaresma (2004), performance of different univariate time series models in forecasting electricity spot prices are compared. Specifically, in analysing the Leipzig Power Exchange (LPX) market, eight different ARMA models are estimated and used for forecasting hourly-time series, in which process, the AR and MA terms are chosen according to the ACF and PACF functions.

In order to test how predicted values, diverge from observed value in the forecasting model, the author applied the Diebold-Mariano (DM) test for predictive accuracy. The DM tests the null hypothesis of equality of expected forecast accuracy[13]. The same methodology is followed by Jakasa (2011) in which, the ARIMA model is used to analysed and forecast the day ahead electricity spot price from EPEX power exchange. For the modelling purposes, the ARIMA method according to the Box and Jenkins is used. The autocorrelation and partial autocorrelation functions are used as based instruments to identify stationarity of the time series.

Although the variable renewable energy (hereafter VRE) with zero marginal costs decrease electricity prices, the literature is inconclusive about how the resulting shift in the supply curves impact price volatility. In Rintamäki (2013), the effect of exogenous variables such as wind and solar power on a depend variable such as electricity price volatility is studied using a seasonal autoregressive moving average model (SARMA (p, q, s)). The SARMA model, also implemented in Aiube (2013) studies, attempts to capture the linear relationship between actual and past values in time series, besides the seasonal pattern. The research sheds light on the fact that wind and solar power production have statistically and economically significant effect on day-ahead price in Denmark and German (Rintamäki, 2013), Results are therefore aligned with Ketterer (2014) and Jonsson et al (2010). In Cuaresma (2004), however, the seasonal part of the electricity spot price is tackled by adding *"s"* number of lags in the AR part of the ARMA model and creating dummy variables for the weekend days, which led to the same results.

However, the lack of literature in matter of "average-daily" electricity price behaviour has motivated me and stimulate my curiosity in carrying this research on a different frequency in order to understand whether electricity's behaviour changes according to the frequency analysed[14]. **Table 1** summarizes the main findings on the economic field of electricity price behaviour.

[13] Assuming two forecast's values h-steps ahead, the DM null hypothesis can be writes as: $= \mathrm{E}[\mathrm{g}(\mathrm{e}_t^{\mathrm{A}}) - \mathrm{g}(\mathrm{e}_t^{\mathrm{B}})] = 0$ where $\mathrm{e}_t^{\mathrm{i}}$ refers to the difference between the predicted values and the actual value of the model "i" (i =A, B) and g(.) is the corresponding loss function (The chosen loss function of this paper corresponds to the Mean root square error)

[14] The Energy market, is mainly divided in two sub-markets: The imbalance market in which electricity is exchange with a frequency of 15 minutes and the day-ahead market in which hourly contract for the day ahead are assigned by auctions.

2.3 THE ENERGY OUTLOOK

In order to acquire a better picture of this research and therefore of the impact of wind power on the electricity spot price, it is necessary to introduce the capacities and the electric mix supply of the aforementioned countries. The reason behind that, is that once information from the MSE approach are inferred, data can be evaluated, comparing them to the market share of wind power in the selected country, obtaining a better understanding about their impact. The supply in the Netherlands is composed of several different sources. It is characterised by a high share of total fossil fuels[15] (81.2% in 2016) and relatively small share of total renewable energy[16], namely 13% in 2016 (CBS Statistics, 2017). Wind energy generated the most amount of electricity **(7.1%),** biomass comprised of only 5%, Hydro power as well as solar photovoltaic accounted for the smallest shares, 0.1% and 1.4% respectively (CBS Statistics, 2017). Differently from the Netherlands, the energy outlook of Germany is more diversified among energy sources. Specifically, it is characterized by a higher share of wind power, which account for **20.3%** of the total generation, 38% coal, 7% gas, 32% nuclear and another 17% of other clean energy sources which include solar and hydropower (Statista, 2018).

[15] Total fossil fuel includes natural gas, hard coal, fuel oil and other fossil fuels (CBS Statistics, 2017).
[16] Total renewable energy includes solar photovoltaic, wind energy, hydro power as well as biomass (CBS Statistics, 2017).

Table 1: Relevant past contribution on electricity price's analysis

Paper	*Market*	*Frequency*	*Model*	*Seasonality*	*Findings*
(Tina Jakasa, 2011)	*European Energy Exchange*	*Hourly*	*ARIMA models*	*Increase the order of the order of the ARMA terms.*	*ARIMA (box Jenkins approach) model best fits for forecasting the electricity spot price*
(Rafał Weron, 2008)	*California power market*	*Hourly*	*ARMA and ARMAX (power plants data)*	*Dummies for weekly seasonality*	*Forecast performance of studied models.*
(Huisman, 2007)	*Dutch, German, French wholesale power market*	*Hourly*	*Stochastic component modelled as mean reverting process*	*Deterministic function*	*Hourly electricity prices do no behave as a time-series process. He proposes a panel data model*
Montero et al (2011)	*Spanish electricity market*	*daily*	*T-ARSV and GARCH (variations included)*	*Dummies for seasonal cycle*	*Comparative analysis of T-ARSV with four GARCH type specifications*
(Rintamäki, 2013)	*Danish, German electricity market*	*Hourly*	*SARMA*	*Seasonal factor in the ARMA model*	*Solar PV decreases the volatility in the German electricity market*
(Yan Dong, 2012)	*Nord-Pool electricity market*	*Hourly*	*ARMA, GARCH, E-GARCH*	*Dummies for seasonal cycle*	*Comparative analysis of different econometric model*
(Ketterer, 2014)	*German*	*daily*	*GARCH*	*Dummies for seasonal cycle*	*Wind power generation decreases the level of price but increases its volatility.*
(Fernando L. Aiube, 2013)	*Spanish, Austrian*	*Hourly*	*SARMA-GARCH model*	*Seasonal factor in the ARMA model*	*The SARMA model has enough explanatory power to forecast the electricity spot price*
(Zhiwei Shen, 2015)	*German market*	*Hourly*	*ARIMA/GARCH Markov regime-switching model*	*No seasonal treatment*	*Forecasting volatility of wind power production. MRS-GARCG better perform over the others.*
(Jesus Crespo Cuaresma, 2004)	*LPX market*	*Hourly*	*ARIMA's models*	*Dummies and variable lags*	*Hour by hour modelling strategy increases the forecast abilities of linear univariate time series*

***Note**: From the table, it can be observed how academics have carried research and studies on the electricity price using mostly hourly frequency and different treatments on seasonality. The table above summarizes the main academic contributions in matter of volatility modelling and electricity price forecasting. It shows different models adopted by scholars and researchers to tackle seasonality and all unique characteristics owned by the electricity spot price.*

Section 3: Hypotheses development

The paper develops two different research objectives:

- RO1: To investigate the impact of wind power generation on the energy price by applying different forecast techniques in The Netherlands and in Germany. The research's method embraces a forecast approach in which results from two different forecast models, are studied and compared to capture the wind generation effect. The increase in accuracy from an autoregressive to a conditional forecast can be intent as the impact of the exogenous variable on the dependent variable.
- RO2: To study and to model the electricity day-ahead spot price to explore its connection with the wind power generation. It aims to to understand its formation process and which connection links the volatility of the electricity price to the wind power generation in both Germany and The Netherlands.

The goal of the RO1 is to evaluate the statistical MSE differences between the simple autoregressive forecast model and the one in which the wind variable is embedded. In pursuance of the aforementioned topics, this paper develops few research questions necessary to decipher the dilemma behind the renewable energy sources. **Table 2**. encapsulates the main hypotheses for the research objective number 1 (RO1).

Table 2: Hypotheses developed for RO1.

Null hypothesis	Alternative Hypothesis
1) The variable wind generation does not increase the accuracy in forecasting the daily APX day-ahead spot price	1)The variable wind generation does increase the accuracy in forecasting the daily APX day-ahead spot price
2) The variable wind generation does not increase the accuracy in forecasting the daily EXPEX Phelix day-ahead spot price	**2**) The variable wind generation does increase the accuracy in forecasting the daily EPEX Phelix day-ahead spot price
3) The incremental accuracy in forecasting the electricity price is greater in the daily EPEX day-ahead spot price than in the daily EEX Phelix day-ahead spot price	**3**)The incremental accuracy in forecasting the electricity price is lesser in the daily APX day-ahead spot price than in the daily EPEX Phelix day-ahead spot price

Note: *The table describes the three hypotheses analysed with regard to the forecast analysis. The hypotheses .1) and 2) aim to demonstrate that wind power generation has a positive impact on EPF while 3) is based on a comparative analysis between Germany and The Netherlands.*

Moving on the RO2 (see **Table 3**), the paper intends to investigate, through a structural analysis, all characteristics and implications from the relationship between the electricity spot price and the wind power generation, and therefore understand the: Direct and delayed effect of wind power generation on the electricity spot price, changes in the relation price-wind generation over time and the Volatility's dependence.

Table 3: Hypotheses developed for the RO2.

Null hypothesis	Alternative hypothesis
4)Wind power Generation does not have a direct effect on the EPEX Phelix day-ahead price	4)Wind power Generation does have a affect the volatility of the EPEX Phelix day-ahead price
5)Wind power Generation does not have a direct effect on the APX day-ahead spot price	5)Wind power Generation does have a direct effect on the APX day-ahead spot price
7) The EPEX Phelix day-ahead spot price's volatility in time t-1 does not affect the day-ahead volatility in time t	6)The EPEX Phelix day-ahead spot price's volatility in time t-1 a does affect the day-ahead volatility in time t
8) The APX day-ahead spot price's volatility in time t-1 a does not affect the day-ahead volatility in time t	7)The APX day-ahead spot price's volatility in time t-1 a does affect the day-ahead volatility in time t
8)Wind power Generation does not affect the volatility of the EPEX Phelix day-ahead price	8)Wind power Generation does affect the volatility of the EPEX Phelix day-ahead price
9)Wind power Generation does not affect the volatility of the APX day-ahead price	9)Wind power Generation does affect the volatility of the APX day-ahead price

Note: *The table shows the hypotheses developed with respect to the structural analysis.*

From this research, I expect that, as stated by the merit order theory, wind power generation has a dampen effect on the electricity spot prices and a positive impact on its volatility. However, in comparing the results from The Netherlands to Germany these effects are supposed to be higher in the EPEX Phelix than in the APX market[17].

Furthermore, given the higher frequency (daily), it is not convinced to assume that the price will behave as same as explained in the literature. Consequently, the degree of connection between the two variables is expected to be lower.

[17] The assumption is based on the fact, that Germany's energy system relies more on wind energy than The Netherlands does.

Section 4: Dataset

The dataset employed in this study include four different time-series in a range period between 25May2015 to 20May2018 equally to a total of 1,090 daily observations (see **Table 5** for all statistical descriptions).

The first analysed time-series is the Dutch APX day-ahead electricity spot price. Established in 1999, under the name APX power NL, is an independent electronic exchange for trading electricity, in which hourly electricity contracts are exchanges among parties. The APX group however, merged its business with the EPEX spot market. It therefore operates under the EPEX spot brand name, but still in all informatics platform (data are released under this name). The acronym "APX" it is hold in this research to better identify the Dutch electricity spot price. All Data are transmitted to the market on hourly frequency, hence 24 different prices are presents every day. However, since research aims to understand how, daily wind power generation influence the daily electricity spot price, all 24-hours daily prices have been averaged in order to obtain an estimation of the daily price. Data are retrieved from Bloomberg.

With regard to the German electricity spot price, the EPEX spot market is used. It is defined as the Exchange for the power market in Europe, covering most of productive European countries such as Germany, France, UK, Belgium, Austria, Switzerland and Luxembourg. Specifically, as suggested by Ketterer (2014), the EPEX Phelix index is preferred as a benchmark for the electricity price in Germany. It corresponds to the arithmetic mean of the market clearing prices for the delivery periods between 0h00 CET and 24h00 for the EPEX spot German day ahead market (epex spot market, 2018). As mentioned before, the EPEX day-ahead market has a different regulation which differentiate dramatically its price formation from the APX day ahead spot market. The main reason relies on the fact that the price is allowed to be negative. As for the Dutch one, data are retrieved from Bloomberg.

Concerning the wind time-series, they have been retrieved manually from both the European Energy exchange transparency platform and from the ENTSOE transparency platform[18]. The time series refers to the forecast wind generation based on predicted wind speed.

[18] The ENTSOE transparency platform provides access to all European electricity market data for all users, across six main categories: Load, Generation, Transmission, Balancing, Outages and congestion Management. (Unicorn Systems a.s., 2018)

The creation of the Dutch wind time-series, however, has not been easy since the beginning, owing to the different frequency (15 minutes) through which data are transmitted to the system. Therefore, the dataset on the Dutch wind power generation was at the first stage, made of 104.640 different 15-minutes observations in 1090 different downloads.

Table 4: Description of the variables and their frequency.

Variable	Original frequency	Used Frequency	Source	Description
APX The Netherlands	Hourly	Daily	Bloomberg	Dutch electricity day-ahead price
EPEX Germany	Hourly	Daily	Bloomberg	German electricity day-ahead price
Dutch wind Generation	15 minutes	Daily	ENTSOE	Dutch wind power generation
German wind generation	15 minutes	Daily	Bloomberg (EEX)	German wind power generation

Note: *The table shows the original and new frequency for all variables as well as the respective database used.*

Successively the initial time series has been converted from 15 minutes frequency to hourly frequency and lately to daily frequency. The absence of adequate information on the renewables energy in the Netherlands is surely a symptom of the energy revolution that the country is currently undergoing.

Table 5: Summary statistics of the dataset

Variable	Observations	Mean	Std. Dev	Min.	Max.
APX The Netherlands	1.090	37.46	8.69	15.37	88.98
EPEX Germany	1.090	32.09	11.58	-52.11	101.92
Dutch wind Generation	1.090	1108.468	819.76	42.7	3688.4
German wind generation	1.090	10154.19	7452.85	741.8	38701.2

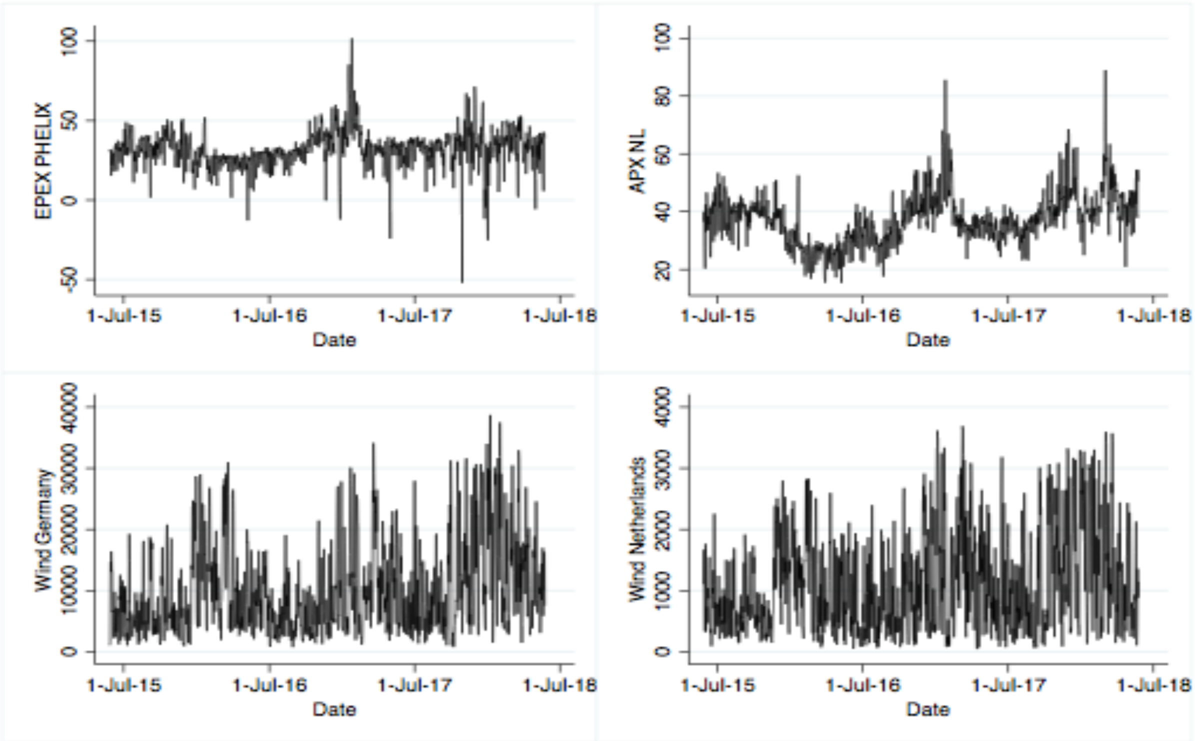
Note: *The table provides a statistical summary of the dataset. As can be noticed, APX and EPEX spot market prices performs tangible difference in the Min. and Max. values. This is mainly due to the different regulation applied to the respective markets.*

Surely, electricity displays unique characteristics which differentiate it from financial assets as well as commodities. Most importantly, electricity is not storable, features that causes paramount consequences: Demand and supply directly determine the market price resulting in a strong prices fluctuation, moreover they need to be balanced any moment of time causing difficulties for grid operators. Besides the intrinsic characteristic of the commodity itself, other features can be observed, summarized under the concept of seasonality, spikes, mean reversion (Geninasca, 2012).

The first distinctive characteristic is given by periodicity in different lengths and consequently seasonal fluctuations. Electricity in fact displays, regardless distinctions in term of frequency, various seasonality over days, weeks and months (Dong, 2012).

From the **figure 1**, it can be observed how electricity spot prices display significant seasonal patterns with respect to the frequency.

Figure 1: Graphic development of the studied time series



Note: *From the graphs, the peculiar seasonal characteristics can be observed for both electricity price time series. The upper graphs display weekly seasonality and numerous outliers that must be eliminate in order to applied the forecast methodology. From the lower graphs, the wind power generation for both Germany and The Netherlands can be observed. As for electricity, seasonality is present given the difference during the 4 seasons.*

Generally, the seasonal structure is mostly determined by the demand, which vary from months to months and from day and night. Secondly, a distinctive peculiarity is given by presence of spikes, mainly observable in the intraday market, which are related to instantaneous supply and demand. Thirdly, within certain interval, stationarity is noticed. It means that electricity is mean reversion over shorter periods.

Looking down at the two bottom charts, wind power generation can be observed for both Germany and The Netherlands. In both cases, the time series presents spikes and trace of monthly and weekly seasonality.

Section 5: Methodology

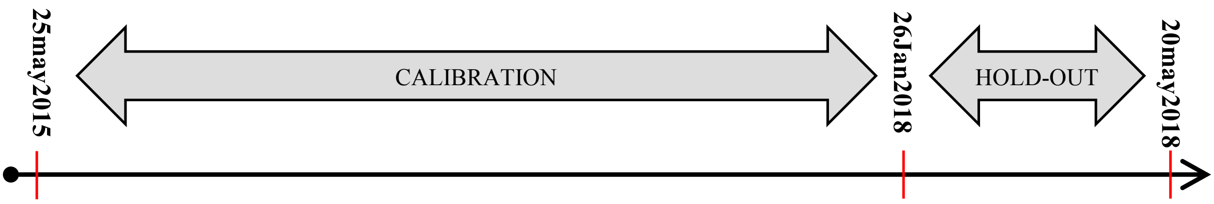
5.1 The forecast analysis

The analysis involves two different forecasts, respectively a static and a dynamic one. The methodology followed by this research is based on the Box-Jenkins approach. Specifically, it encompasses the ARMA model, and its extension which account for the exogenous variable (ARMAX). The approach consists of extracting predictable movements from observed data through a series of iterations. The applied box-Jenkins method is entirely considered a forecasting tool which follows two steps:

- **Model Identification:** Through a graphic and statistical analysis, autocorrelations, partial autocorrelations, seasonal patterns, trend and possible lags are detected. Most commonly, all parameters of the models are calculated in this phase.
- **Model Estimation and verification:** Once the best model is detected, predicted values are estimated and verify through accuracy measures.

However, before proceeding with the analysis, the first step involves the so-called data splitting of the time series (Elisabeth Woschnagg, 2004). The dataset is therefore divide in two different parts:

1. The construction sample or calibration (ranging from 25may2015 to 26jan2018) which is used to provide the ground-floor for the estimation of the model.
2. The validation or hold-out sample (from 27jan2018 till 20may2018), which is used to back-test the forecasting power (test period) of the model estimated during the first step of the work[19].



The research's model begins with an autoregressive moving average (ARMA) forecast model in which predict values are generated using only past data.

[19] Analysist usually hold back 10% of the sample, however there is not a theoretical meaning or mathematic rule. (Elisabeth Woschnagg, 2004).

However, before that, all time series need to be checked in all their properties, meeting the forecast's requirement, hence stationarity, seasonality and distribution of residuals are studied. Along these lines, correlations (AC) and partial correlations (PAC) functions are performed to detect correlation among residuals as well as the orders of the ARMA's models. Performed all statistical tests, the forecast techniques are applied to the best model, chosen in accordance to Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC)[20]. Predicted values are therefore evaluate in terms of accuracy through a MSE analysis. It evaluates how much the predicted values diverge from actual value (hold-out sample).

In order to detect the wind impact on electricity price formation, the procedure is repeated twice. Firstly, in an autoregressive model (ARMA) and lately in a conditional one (ARMAX) in which electricity prices not only depends on past values but also on the wind power generation. This model is consequently an essential part of this analysis, since the predicted values generated by the ARMA process with the exogenous variables (ARMAX) are used to create a comparison with those values before generated without the "wind power generation" variable, capturing the wind effect in electricity price formation. It follows hereby that, in order to gauge statically differences between ARMA and ARMAX forecast's accuracy, the Diebold-Mariano (DM) test is performed on the two MSE loss functions. Results from the test will then shed light on whether, wind power generation has a positive effect on the electricity spot price.

5.1.1 Model identification

In pursuing the forecasts, the ARMA family models are introduced, which, however are applied in the next section. According to Jakasa (2011) "all models are good enough to forecast day-ahead electricity spot prices".

Most time series can be described by Autoregressive Moving Average (ARMA) model. The stationary series with white noise is said to be ARMA (*p, q*) if: (Kruangpradit, 2013).

$$\left(1-\sum_{k=1}^{p}\alpha_k L^k\right)Y_t=\left(1-\sum_{n=1}^{q}\beta_n L^n\right)\varepsilon_t \qquad [1]$$

[20] BIC usually estimate the quality of the model and more informative if the number of observations is very high (verbeek,2004).

Where ε_t is the white noise, L is the lag operator, α_k are the coefficients for the autoregressive part of the model (referring to the lags of the dependent variable), while β_n are the coefficients for the moving average part of the model (referring to the lags of the noise).
The ARMA (p, q) model is therefore constitutes by two different parameters, the autoregressive (p) and moving average (q) parts:

- AR (p). The parameter "p" refers to the number of autoregressive orders in the ARMA model, accounting for its Autoregressive(AR) part. The p orders describe which previous values from the series are used to predict current values.
- MA (q). The last part involves the moving average aspect of the ARMA model, specifically q embeds the number of moving average orders in the model. It specifies how deviations from the series mean for previous values are used to predict current values.

The basic ARMA model can be additionally extended by adding an exogenous variable (In [2] is presented as W to recall the WIND dependent variable) forming the so-called ARMAX model. It displays the same autoregressive and moving average parameters but, given the presence of an additional variable, it is presented as follow:

$$\left(1-\sum_{k=1}^{p}\alpha_k L^k\right)(Y_t - W_t) = \left(1-\sum_{n=1}^{q}\beta_n L^n\right)\varepsilon_t \qquad [2]$$

Where $\mathrm{W_t}$ is the exogenous variable of the model which reflects the independent variable, namely wind power generation and α_k and β_n represents the order of the parameters p and q respectively.

Model [2] is based on the assumption that wind energy, as all renewables, is not storable, meaning that, whenever the production is greater than the forecasted and predetermined agreed quantity by the grid operator, the surplus must be wasted in order to do not overcharge the power grid. In the energy finance field, indeed, demand and supply match immediately according to the frequency of the market. In the particularity of the day-ahead spot market, the electricity price in time t+1 is computed on the forecasted production for the period t+1 in time t. The price formation, therefore, is highly influenced by the merit order theory, meaning that an excessive RES production will shift the curve dampening the price and vice-versa.

During the last decade, the entire energy system can be considered as renewable energy driven, since all hard producers (coil/gas/nuclear) adapt their production on those who have priority in the merit order theory. The no-storability issue helps us understanding how, although it might appear that wind power it is heavily auto-correlated, the energetic market system does not allow to the past production to influence the future one. Differently from the financial market, indeed, energy is a tangible and durable goods which, once is generated it is either used or wasted.

The connection could arise exclusively from the wind forecasts, which however as the intuition suggests, depends on meteorology, science highly variable and susceptible on different factors out of the competences of this paper. It follows that, delaying the independent variable, namely wind power, by 1,2,3 lags in the forecast model would bias the forecast, rowing against the theory and the intuition behind the energy market. By delaying the wind by 1,2,3 lags in the forecast, the model will forecast the actual price using respectively the wind production at time t-1, t-2, t-3. Further lags of the wind variable are therefore not used in order to detect the direct effect in forecasting the electricity spot price.

The electricity price is therefore modelled based on an ARMA model, which to recall encompasses an autoregressive and moving average part. Whenever the order of the AR part increase, past values of the electricity prices are used. It follows that the wind power generation would be lagged by the extension of the AR parameters in the ARMA forecast model in order to discover the direct effect of the two variables.

In pursuing the objective of this research, which follow a comparison of different forecasts, the random walk model surely plays an important role. The random walk model can be considered as the simplest and most important model in time series forecasting. The model in fact assumes that in each period the variable takes a random step away from its previous value. These steps are independently and identically distributed.

The model might have drift or no, depending on whether the steps size allows a non-zero mean or a zero one.

The random walk model without drift, for the variable PRICE[21] is:

$$Price_t = Price_{t-1} + \varepsilon_t \qquad [3]$$

[21] The variable Price refers to the electricity day-ahead spot price

It therefore predicts that; all future values will equal to the last observed value. It means that all predicted values, are expected to be close to the observed value, although they are equally likely to be higher and lower. The random walk model is elaborated in order to offer a further mean of comparison between the ARMA forecast model.

Many time series, however, possess in practise a non-stationary behaviour. Usually due to a trend, a change in the local mean or seasonality. Since the box-Jenkins forecast method is for stationarity models only, the time series must be adjusted before testing the model. An analysis to detect stationarity, seasonality and peaks or outliers is hereby performed.

5.1.2 Stationarity

The Augmented Dickey Fuller test for unit root is firstly used to detect whether a variable is stationary or not. In case of a failure to reject the null hypothesis of a unit root[22], the variable is differenced (returns) and the Augmented Dickey Fuller test is then calculated on the differenced variable. The procedure is repeated until the null hypothesis is rejected, meaning that the p-values is lower that 5% confidence level.

5.1.3 Seasonality

Tackling seasonality is paramount important in forecasting the price, since omitting it in the calculation could bias the results. If the series is detected stationary, seasonality is studied by looking at the partial autocorrelation (PAC) functions. A partial autocorrelation can be explained as the amount of correlation between a variable and a lag of itself. Hence, the correlation of a time series Y at lag 1 is the coefficient of correlation between Y_t and Y_{t-1} which is presumably the correlation between Y_{t-1} and Y_{t-2} .Seasonal pattern will arise in the eventuality in which correlations is recursive for determined numbers of lags.

However, it can be "squeezed-out" by creating a dummy variable for those days in which values follow a specific pattern (such as weekend days in which the price is generally lower than the weekly average). Consequently, in order to remove eventual seasonal pattern, weekly dummy variables and upper/lower constraints are applied to the time series before modelling. The seasonal treatment is therefore applied before modelling the time series.

[22] The null hypothesis is that of non-stationarity (or existence of a unit root).

5.1.4 Order of ARMA model

Having checked both stationarity and seasonality, the next step in fitting an ARMA model is to determine the orders of AR (p) and MA (q) terms. The approach involves the analysis of residuals, specifically of the autocorrelation functions (AC) and partial autocorrelation (PAC) functions. In the analysis, the partial autocorrelation at lag p estimates the AR(p) coefficient, in an autoregressive model with p terms. Thus, by exploring the properties of the PACF the AR term can be detected (Nau, 2018). The ACF plays the same role as PACF but instead it explains the MA(q) term, by displaying how many moving average (MA) terms are likely to be needed to remove the remaining autocorrelation from the differenced series.

5.1.5 AIC criterion

Having accomplished all the pre-requirement required by the Box-Jenkins approach in forecasting the electricity spot price, the model can be generated and the prices forecasted. Last but not least, all models need to be ranked by AIC criterion. As suggested by Yan (2012), the order selection is quite important in using an ARMA process, especially for forecast. Although ACF and PACF suggest determined number of p and q terms, it is not surely true that they will generate the best forecast model, hence multiple forecast models need to be generated and compared by the AIC criterion. As a matter of fact, when the model is applied for forecasting, the mean square errors (MSE) might be large/small which depends on errors from predicted parameters of the fitted model used. Many criteria are used to detect whether a model is desirable or not. The most widely used encompass the AIC criterion.

5.1.6 Evaluate the forecast accuracy

As specified above, the second step of the Box-Jenkins approach involves model estimation and verification, thus once all parameters have been obtained and understood, the valuation and verification of the model begins. The residuals of models are in fact, checked for any remaining patterns or normality. To test the existence of a white noise process of the residuals the diagnostic Portmanteau test is used. It is based on the asymptotic distribution of the residual autocorrelations, $r_1, r_2, . . ., r_m$, where m<n-1 is the largest selected lag. Additionally, the Shapiro test is used to detect whether the residuals of the model are normally distributed although graphical tools such as histogram and auto-correlogram are investigated as well.

The final step of the approach comprises an evaluation of accuracy of the predicted values. The three most common measures of predicted and forecast accuracy are considered to be the mean squared errors (MSE), mean absolute error (MAE) and Theil's inequality coefficient (Fair, 1986). These measures have been used to evaluate prediction of ex-post as well ex-ante forecast (Elisabeth Woschnagg, 2004). Ex-post forecast is defined as the one in which the actual values of the exogenous variables (wind power generation) are used while the ex-ante is the one in which guessed values of the exogenous variable are used. In this analysis, the forecast involves an ex-ante approach since all predicted values are compared to observable values (see data splitting)[23].

However, Chen et al. (2004) separated forecast accuracy measures in stand-alone and relative accuracy measure, defining two different categories. In this analysis only the first categories, is used. The Stand-Alone measures can be defined as those that can be obtained without additional reference forecast, for instance measures associated with a certain loss function or based on quadratic functions.

To this category belong the MSE. In order to achieve the paper 's objective a first analysis of MSE is therefore assessed. The MSE depends on the scale of the dependent variable. Hence it can be used as a relative measure to compare forecasts using same time series among different econometric models. It is defined as "how much predicted values deviate from actual value", hence smaller is error, better is forecasting power of the model. Since the effect of each error on MSE is proportional to the size of the squared error, larger error has non-linear effect on the result. On the light of this, this measure is highly sensitive to outliers. The MSE is described in the form of:

$$MSE = \frac{\sum_{t=1}^{\tau}(y_t^* - y_t)^2}{\tau} \quad [4]$$

The MSE of predicted values y^* for time t of a regression's dependent variable y_t with variables observed over time T, is computed for T different predictions as the mean of the deviations.

[23] In order to achieve an ex-post forecast, the multivariate vector autoregressive (VAR) model best fit the need.

5.2 STRUCTURAL ANALYSIS

Moving on the second research objective, the following sub-section describes the methodology followed by the structural analysis. It investigates and explores all characteristics and implications from the relationship between the electricity spot price and the wind power generation, and therefore understand the:

- Direct and delayed effect of wind power generation on the electricity spot price
- Changes in relation price-wind generation over time
- The electricity spot price's volatility

5.2.1 DIRECT AND DELAYED EFFECT OF WIND POWER ON THE ELECTRICITY SPOT PRICE

Literature argues that renewable energy sources, given the lower marginal cost of production, dampen the electricity spot price. Following the analysis conducted by Ketterer (2014) this research elaborates an ARMAX model in which the exogenous variable (X), namely wind power generation, is added up to 4 lags, to detect whether wind has a direct and/or a delayed effect on the electricity spot price. The analysis therefore encompasses a valuation between the results from the no-delayed model and the one in which the variable is lagged. Moreover, from the parameters, the effect of wind power can be inferred.

5.2.2 STRUCTURAL BREAK DUE TO CHANGES IN THE WIND POWER GENERATION FEED-IN

Based on Alaa Abi Morshed (2015), most of economic time series are characterized by multiple or single structural changes in their parameters of their mean as well as in their volatility. A structural break can be defined as an unexpected shift in a time series which might lead to a huge forecasting errors and more generally to an increase imprecision of the model. The paper therefore investigates whether structural changes have affected the relationship between the electricity spot price and its driver over-time.

The method encompasses a Wald test[24] over a linear regression of wind power generation on the electricity day ahead spot price without fixing a known break-date but combining the test statistic for each possible break-date during the sample.

[24] The wald test studies whether the coefficients in the time-series regression has changed over time, defining a break date. The null hypothesis is that there is no structural break during the sample.

5.2.3 THE ELECTRICITY SPOT PRICE'S VOLATILITY

In order to investigate how volatility changes over time, the generalized autoregressive heteroscedasticity (GARCH) model is introduced. According to Erni (2012), the electricity spot price's volatility is defined as not constant and clustered.

Volatility clustering can be defined as the phenomena in which high period of changes are followed by period of high volatility resulting in persistence of the amplitudes of price changes, while periods of small changes are followed by periods of low volatility (Mandelbrot, 1963). As electricity is not storable, the literature states that the price tends to spike and then revert as soon as the difference in supply and demand is resolved.

The GARCH is a time series model that allows explicitly to test the effect of daily wind power generation on the mean and volatility of the electricity price in an integrated approach across different time periods, (Ketterer, 2014). It is consequently used to test whether the volatility of the electricity price is conditional to its past values and whether it is conditional to the wind power generation variable. The reason relies on the fact that it has the advantages of determining the effect of short and long-term price volatility (Wirdemo, 2017).

The ARCH, GARCH models were firstly proposed by Engle (1982) and Bollerslev (1986) respectively, in order to capture simultaneously volatility clustering and leptokurtosis (AL-Najjar, 2016).

Engle (1982) introduces the use of the ARCH (m) process which can be intend as model in which the variance at time t is conditional to an observation at the previous m times. However, the model showed few limitations, overcome subsequently through the introduction of the GARCH model. The new proposed process, not only share the main assumption of the ARCH model, regarding the conditional variance, but also allows the lagged conditional variances to enter in the model. It in fact uses values of the past squared observation and past variance to model the variance at time t. Given the purpose of the research and the time series taken into consideration, the literature suggests the GARCH (1,1) specification. It is considered as the simplest and most robust among volatility models. On the light of this, the model is applied. Higher orders are usually applied when the dataset displays a long span of data.

The main goal of the model is explaining the causes of changes in volatility. By analysing graphically both time series, it can be noticed the presence of volatility clustering, meaning that the present volatility depends on past observations and past volatilities, hence volatility h_t

can be explained by past volatility h_{t-1} (Yan, 2012). Analogous to the ARMA models, the GARCH model use autoregressive terms (Ketterer, 2014), which make it suitable for analysing seasonal effects.

The GARCH (*p, q*) model is strictly stationary with finite variance whenever the condition $\omega > 0$ and $\sum_{i=1}^{q} \alpha + \sum_{j=1}^{p} \beta_j < 1$ are required. As above described, electricity price displays unique characteristics such as not storability. Price therefore tends to spike and then revert as soon as demand and supply converge again. Suggested and motivated by the work of Ketterer on the German electricity price (2014), this paper apply an AR-GARCH and ARX-GARCHX model, including a mean reversion parameter. Therefore, it follows an interested addition which include the mean reverting characteristic of electricity price. Following Ketterer (2014), it can be captured by a Gaussian AR (1) process $y_t = \mu + \phi y_{t-1} + \varepsilon_t$ where $\phi = 1 - k \; and \; \varepsilon_t \sim iidN(0, \sigma^2)$. The speed of the mean reversion is consequently calculated from the coefficient for the AR parameters.

The mean reversion AR-GARCH model consequently become:

$$PRICE_t = \mu + \sum_{i=1}^{l} \phi PRICE_{t-1} + \varepsilon_t \quad [5]$$

$$PRICE\ VOLATILITY_t = \omega + \sum_{i=1}^{q} \alpha PRICE\ SHOCKS_{t-1} + \sum_{j=1}^{p} \beta_j PRICE\ VOLATILITY_{t-j} \quad [6]$$

The Equation [5], namely mean equation, explains how past prices (y_{t-1}) influence the actual electricity price (Y_t). The equation [6], namely, variance equation, instead describe how the volatility at time t, is conditional to how past price changes (h_{t-j}) and current price shocks (ϵ_{t-1}^2)(Wirdemo, 2017). In the equation ω, is defined as the long-run variance. Following the interpretation provided by Campbell et al. (1997), α measure the extent to which a today volatility shock feeds throught the next period volatility, while $\alpha + \beta$ measures the rate at which this effect dies over time. The two equation are therefore connected through the term ε_t Since $\sum_{i=1}^{q} \alpha PRICE\ SHOCKS_{t-1} = \sum_{i=1}^{q} \alpha \varepsilon_t{}_{t-1}^{2}$.

Moreover, in order to test the hypotheses of RO2, the wind power generation variable is introduced as an explanatory variable for the volatility. It follows that a new model which account for this feature and an additional exogenous variable can be elaborated. The daily data for wind generation w are therefore included in the mean and variance equations of the new, so called ARX-GARCHX model.

$$PRICE_t = \mu + \sum_{i=1}^{l} \phi PRICE_{t-1} \sum_{j=1}^{m} \theta_j WIND_{t-1} + \varepsilon_t \quad [7]$$

$$PRICE\ VOLATILITY_t = \omega + \sum_{i=1}^{q} \alpha PRICE\ SHOCKS_{t-1}^2 + \sum_{j=1}^{p} \beta_j PRICE\ VOLATILITY_{t-j} + \sum_{k=1}^{s} \gamma_k\, WIND_{\ t-k} \quad [8]$$

Where, as the equation [5], [7] explains how past values and the additional chosen parameter WIND influence the price. The equation [8] instead shows how past price changes and the current price shocks as well as the chosen parameter $WIND_{t-k}$ affect the price at time t (Wirdemo, 2017). In a normal GARCH model, the coefficients in the variance equation, including γ_k should be non-negative to ensure that the variance is always positive (Ketterer, 2014).

In order to estimate the parameters of the conditional GARCH models, the maximum likelihood estimation is used[25]. However, before that, some preparatory works are necessary to check whether the variables satisfy the model's requirements. The dataset in fact, needs to be slightly modelled: identify and remove both seasonality and extreme values and test whether ARCH effect[26] and white noise process is present. For this purpose, the Lagrange multiplier tests is used on squared residuals. Stationarity and seasonality check follow the same procedure described in 5.1.2 and 5.1.3

Lately, in order to verify whether the model has effectively a power explanation, the residuals of the models are checked for any remaining patterns or normality. To test the existence of a white noise process of the residuals, the diagnostic Portmanteau test is used. Additionally, the Shapiro test is used to detect whether the residuals of the model are normally distributed although graphical tools such as histogram and auto-correlogram are investigated as well.

[25] The algorithmic is automated by STATA Inc.

[26] The squared residuals of the time series model exhibit autocorrelation

Section 6: Empirical results

6.1 THE FORECAST MODEL

Before assessing the forecast model, all time-series are analysed in all their characteristics as suggested by the methodology and Box-Jenkins approach.

6.1.1 TIME SERIES ANALYSIS

Stationarity is firstly studied. From **Figure 1** it can be observed that the prices hover around a mean level, non-displaying any trend (Dong, 2012). However, in order to confirm the hypothesis of stationarity, the augment dickey fuller test is employed. Results from the test are displayed in **Table 6** below. It can be noticed that all variables are detected to be stationary at 1% confidence level in the sample period, meaning that first different or integration are not necessary to forecast the electricity spot price.

Table 6: The table displays Dickey-Fuller test's result for all four-time series.

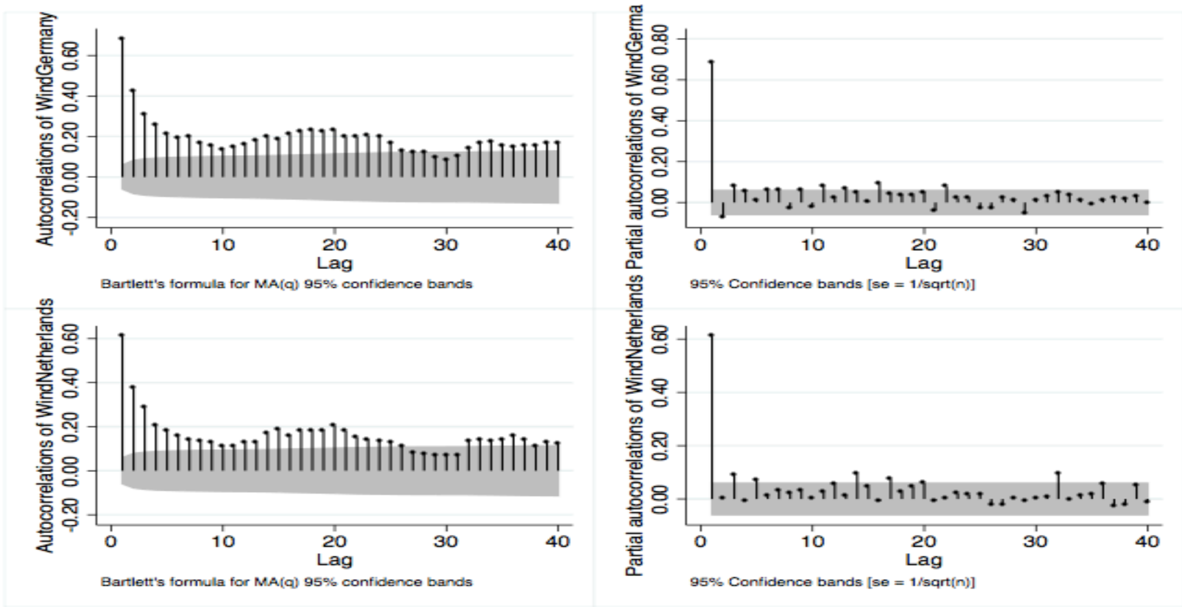
Variable	Test statistic	1% critical value	P-value
Wind Generation in Germany	-14.378	-3.430	<0.0001
Wind Generation in The Netherlands	-16.231	-3.430	<0.0001
APX spot price	-11.547	-3.430	<0.0001
EPEX Phelix spot price	-15.979	-3.430	<0.0001

Note: *As can be observed from the last column the P-value is found to be lower that 1%, meaning that all variables are stationary and no first difference are required. The ADF statistics, used in the test is a negative number, meaning that the left tail of the distribution is investigated, hence all terms are intended in absolute values. As the number increase, stronger rejection of the hypothesis that there is a unit root occurs.*

It follows the seasonal study which encompasses the autocorrelation (AC) and partial-autocorrelation(PAC) analysis. This seasonal treatment is considered as an ex-ante procedure since all adjustment are computed before the mere phase of modelling.

Starting from the wind variables, the ACF and PACF plots of the Dutch and German wind power generation show a slow decay and a slight recursive seasonal pattern in all their sample. However, although the autocorrelation of residuals is significant for a large number of lags, it is perhaps merely caused by the propagation of the autocorrelation at lag 1. This is confirmed by the PACF plot which has a significant spike only at lag 1, meaning that all autocorrelation shown in the ACF are explained by the autocorrelation the lag 1.

Figure 2: The ACF and PACF plots for WINDNETHERLANDS, and WINDGERMANY



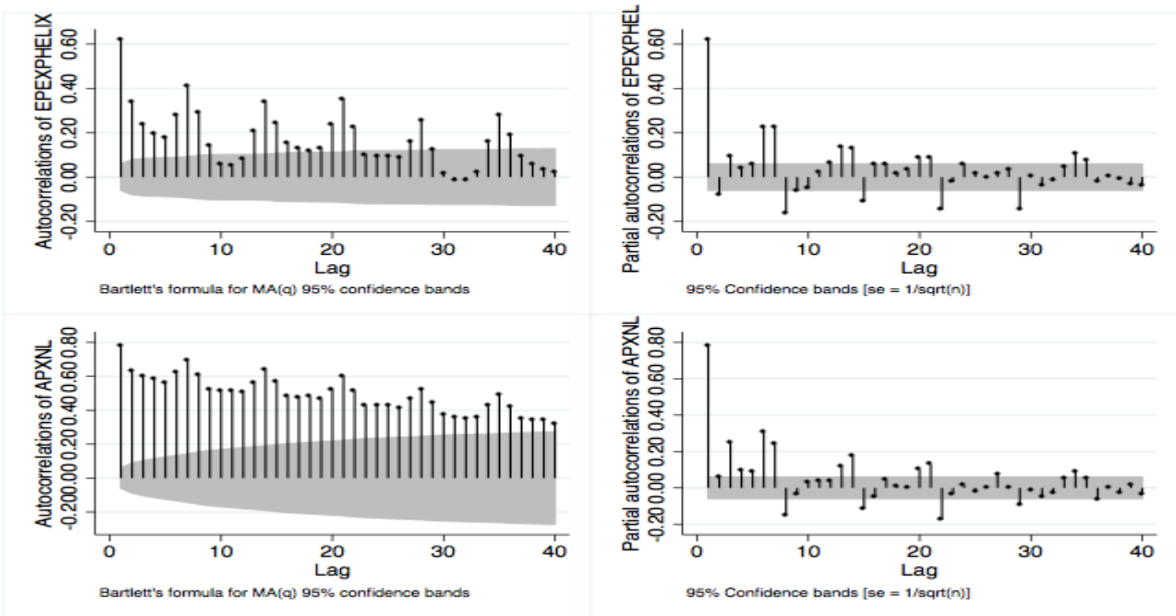
Note: *The present graphs show the ACF and PACF for both wind time series. They reveal that both time series present a stationary process and trace of seasonal patterns. Specifically, both ACF shows sine waves which reflect the monthly seasonal scheme of the wind generation. Furthermore, the slow decay in the upper left charts symbolise the presence of a slow trend, which however does not persist in the DF test. With regard to the order of the ARMA, the AC and PACF suggest an ARMA (1,1)*

Most importantly however, the seasonal patterns displayed in the ACFs are pronounced in the period between the lag 13 and 25. As suggested by the Box-Jenkins approach, a time series can be de-smoothed by differencing or using logarithm/exponential functions. In any case, this seasonal effect, does not seem enough pronounced in the sample period to justify either a logarithmic/exponential transformation or introduction of the dummy variables, since it the mere characteristic of wind power generation. The application of a dummy variable in this case would false the result and therefore the goal of the research. The time series is consequently not adjusted.

Differently from that, the AC and PAC functions of the electricity spot prices shed light on different features that must be cancelled off from the time series before assessing a forecast model. Specifically, from **figure 4**, it can be noticed how, the seasonal pattern clearly appears in the evolution of the time series. The ACF plots for APX and EPEX-phelix present the same slow decay and a seasonal pattern in all its sample. The degree of autocorrelation is significant for a large number of the lags but it is presumed that the autocorrelation from lags 2 on, is merely caused by the propagation of the autocorrelation at lag 1. Furthermore, the AC plots suggest a seasonal term of 7, which must be corrected before modelling.

It therefore brings me to the conclusion that, during the weekend-days the price is on average lower than in the week-days for both electricity prices time-series, which is confirmed by looking at the values of the time-series. The seasonal analysis suggests an ARMA (1,1) or ultimately an ARMA (2,1) model.

Figure 3: AC and PAC functions for the APX and EPEX phelix spot price with their returns.



Note: *The present graphs show the ACF and PACF for both wind time series. They reveal that both time series present a stationary process and trace of seasonal patterns. Specifically, both ACF shows peaks every 7 lags which refer to the weekend days. Furthermore, the slow decay in the upper left charts symbolise the presence of a slow trend, which however does not persist in the DF test. With regard to the order of the ARMA, the AC and PACF suggest an ARMA (1,1).*

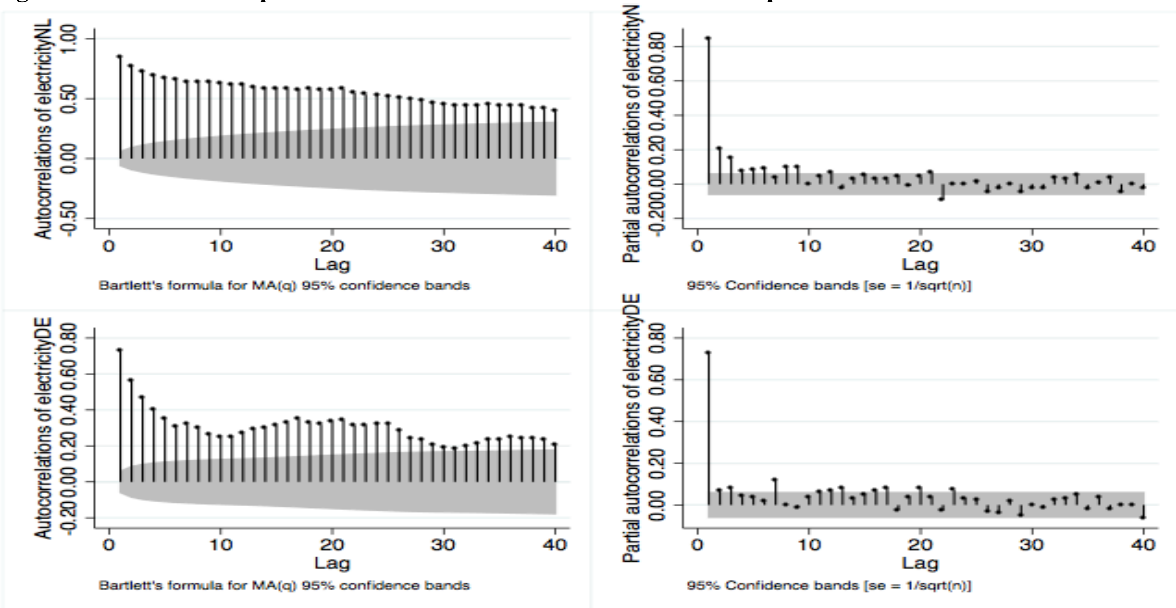
6.1.2 Seasonal treatments

The paper hereby elaborates a strategy aiming to eliminate the seasonal patterns displayed in the price time series. It involves a preliminary adjustment based on four different dummy (0/1) variables, in order to delete the outliers (both values higher than 2 times the average and lower than half time the average[27]) and weekend values presents in the ACF and PACF analysis. The dummy variables therefore substitute all Saturday/Sunday values in which the price is statistically lower than the weekly average. Additionally, as suggested by Ketterer (2014) all outliers are deleted. From **Figure 3,** all improvements in the ACF and PACF functions can be observed. The new graphs, although still display a slow decay in autocorrelation, do not present any form of seasonal patterns able to bias the models. The new summary statistics is presented

[27] The proposed values are chosen according to Ketterer (2014) in order to make the time series hover around the mean value. (e.g. 37.46 with regards to the APX electricity spot price).

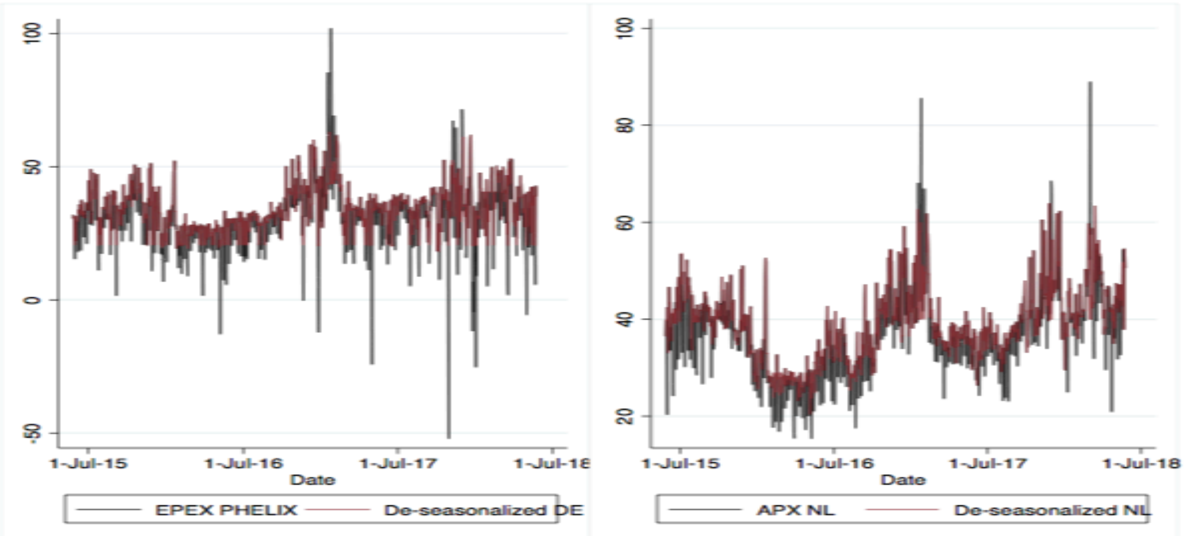
in the **Table 7** below. The graphic price development also exhibits improvements (see **Figure 4**).

Figure 3: AC and PAC plots of the de-seasonalized APX and EPEX-phelix time series.



Note: *The new ACF and PACF for the EPEX phelix and APX show clear difference in terms of seasonality. Spikes are no longer present in the ACF graph.*

Figure 4: DE-Seasonalized APX and EPEX day-ahead spot price development



Note: *The figure shows the development of both time series over time, highlighting the difference between the original and the one in which seasonal treatment is applied. The new time series, displays less outliers. Moreover, all week-end values have been replaced using a dummy variable.*

Table 7: Summary statistics describing the changes before and after the seasonal treatment

Variable	Observations	Mean	Std. Dev	Min.	Max.
Seasonal EPEX Phelix	1.090	32.09	11.58	-52.11	101.92
De-seasonalized EPEX PHELIX	1.090	34.18	8.41	18.03	62.88
Seasonal APX	1.090	37.46	8.69	15.37	88.98
De-seasonalized APX	1.090	38.90	7.68	20.39	71.53

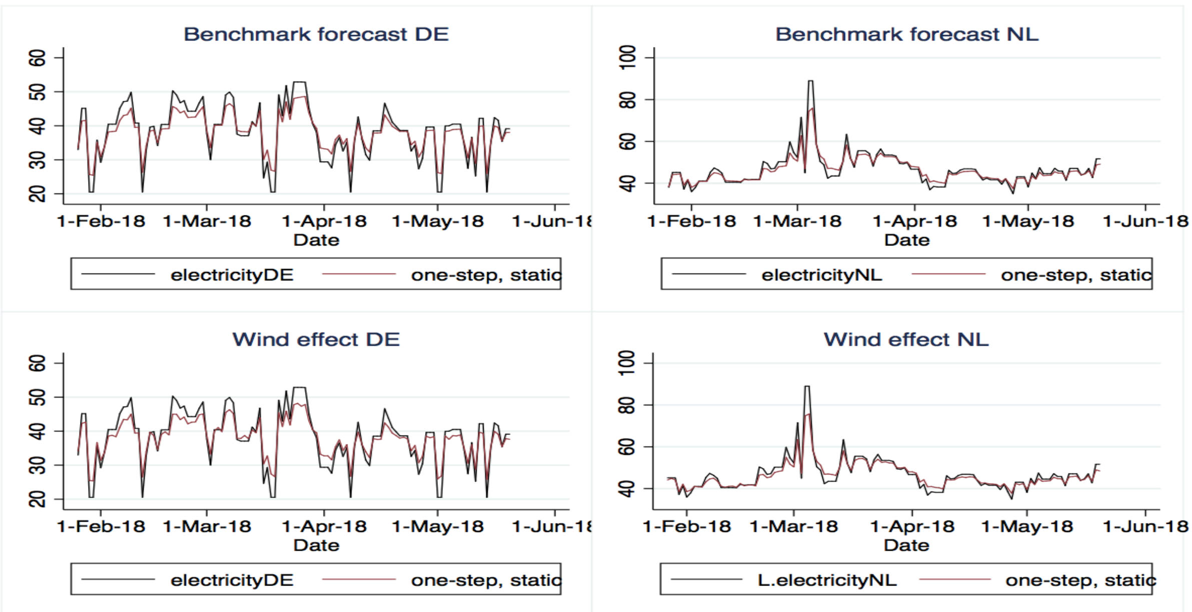
Note: *This table presents summary statistics for all-time series before and after the seasonal treatment. The introduction of the dummy variables for the weekend days and upper/lower restriction (non-negative price in the EPEX Phelix) make the time series more suitable for a forecast model.*

6.1.3 The static forecasts

Once all adjustments in terms of seasonality and stationarity are executed and the orders of ARMA (*p, q*) are detected, both static and dynamic forecast models can be elaborated.

Starting from a static one, it uses the actual values of the dependent and explanatory variable in predicting the forecast. It means that, predicted values will never deviate from the hold-out sample's value. Results from the German analysis show that, in accordance with the AIC criterion, the ARMA (2,1) model is detected as the more appropriate (see appendix). The model, moreover displays significance at 5% confidence level of all AR and MA parameters as well as the constant of the model. The ARMA (p, q) is elaborate in the calibration period in order to predict values in the hold-out sample.

Figure 5: Static forecast for both APX and EPEX phelix day-ahead spot price



Note: *The figure shows the price evolution for both know and static predicted values of the autoregressive forecast in Germany. As can be observed, the prediction is based on one-step ahead forecast.*

Results from the analysis in Germany shows a MSE of 54.77 in the period between 27Jan2018 to 20may2018. This value is hold and considered as a benchmark, lately compared to the those from the ARMAX forecast model in which the wind is embedded. As for Germany, results from the analysis in The Netherlands are based on ARMA (2,1) which minimize the AIC at 53.752, showing significance at 5% confidence level of both AR and MA parameters. The deviation from actual values is of 16.82. In comparing the different results from the Autoregressive and Conditional forecast, the Diebold-Mariano (DM[28]) test is used. It is a test for the predictive accuracy measures which determines whether forecast accuracies of the models are significantly different. Having defined the two-loss function MSE, respectively for Germany and the Netherlands, the paper aims to compare the forecasts in terms of mean square errors. The test is repeated twice, for both countries with respect of the autoregressive and conditional forecast. The test is based on the loss function differential between the two models, ARIMA and ARIMAX for both countries.

Table 8: Diebold-Mariano test for electricity in Germany

Criterion: MSE over 114 observations	
SERIES	MSE
Arima (2,0,1)	54.77
Arimax (2,0,1)	54.32
Difference: 0.3851	
By this criterion, arimax (2,0,1) is the better forecast	
H0: Forecast accuracy is equal.	
S(1) =0.868 p-value = 0.3851	

Note: *The table summarizes results from the DM test for ARMA models in Germany, The MSE is found to be lower in the conditional forecasting, revealing that the prediction is more accurate when wind power is included in the model. The significance at 1% level confirm the findings.*

Table 9: Diebold-Mariano test for electricity in The Netherlands

Criterion: MSE over 114 observations	
SERIES	MSE
Arima (2,0,1)	16.82
Arimax (2,0,1)	16.63
Difference: 0.1903	
By this criterion, arimax (2,0,1) is the better forecast	
H0: Forecast accuracy is equal.	
S(1) =1.951 p-value = 0.0511	

Note: *The table summarizes results from the DM test for ARMA models in Germany, The MSE is found to be lower in the conditional forecasting, revealing that the prediction is more accurate when wind power is included in the model. The test rejects the null hypothesis at 5% confidence level.*

[28] The null hypothesis of the DM test is that there is no difference in the accuracy of the two competing forecasts.

Evidences from the tables above show that, once the wind variable is added in the ARMAX model, the MSE decreases. It follows that wind power generation increase the accuracy in predicting the EPEX phelix and APX spot price. However, by looking at the differences in MSE between the two countries it can be inferred that the model works better for Germany than for the Netherlands. Specifically, the MSE reduction is found to be greater in Germany, meaning that it has a greater effect on the merit order theory.

Findings from the analysis, however show that the reduction in MSE is not as large as expected[29] (see **Table 10**). The table below shows the result from the ARMAX model. As can be seen the MSE it is observed to be lower, meaning that with energy has a positive effect in forecasting the electricity spot price. The incremental accuracy of the forecasted values is positive, highlighting the direct relationship between wind and price and the merit order theory but, given the chosen average daily frequency, wind power generation does increase significantly the accuracy in forecasting the electricity spot price. However, the result is in line with the literature, specifically it confirms that as stated by Ketterer (2014), wind power generation has a positive effect on electricity price formation.

Table 10: Differences in MSE between the static ARMA and ARMAX forecast model.

Model	*Country*	*ΔMSE*
ARMA/ ARIMAX (2,1)	*Germany*	0.4546
ARMA/ ARIMAX (2,1)	*The Netherlands*	.0.1903

***Note**: Findings reveal the impact of wind power generation under a forecast point of view. From the table, it is shown that the wind has a greater effect in Germany than in The Netherlands, given the delta MSE of 5%. These values are successively confirmed through the use of DM test. These numbers can be interpreted as follow: Wind power generation has a positive effect of 0.8% on the German electricity spot price and only 0.16% on the Dutch electricity spot price.*

In both scenarios, the DM test reject the null hypothesis in which forecast accuracy are equal, stating that whenever exogenous variable, namely WIND power generation, is included in the model, the forecast better performs in terms of precise.

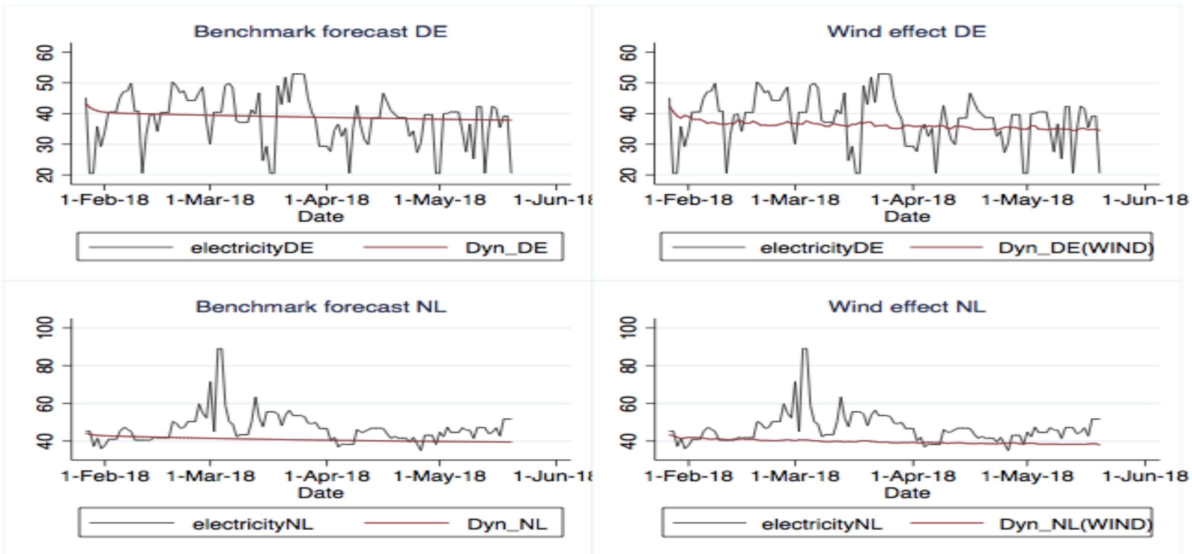
[29] The research expected that the reduction in MSE was as big as the wind market share in the energy market.

6.1.4 THE DYNAMIC FORECAST

Results from the dynamic analysis are similar to the static one although the clear difference in modelling. The main distinctions emerge because of their estimation procedure.

The dynamic forecast uses in fact the value of the previous predicted value of the dependent variable to elaborate and compute the next one. It means that it pretends to do not have any information about the dependent variable during the hold-out period. Hence all values are the product of different forecasts. On the other hand, static forecast uses the actual value for each subsequent forecast. It follows that although the one-step-ahead forecasts never deviate far from the observed values, the dynamic forecasts have larger errors over time. To understand the reason behind, the model need to be rewritten as the forecasted value of PRICE at time t depends on the value of PRICE at time t – 1 or on the PRICE and WIND in the case of the ARMAX model. When making the one-step-ahead forecast for period t, we know the actual value of PRICE at time t−1.

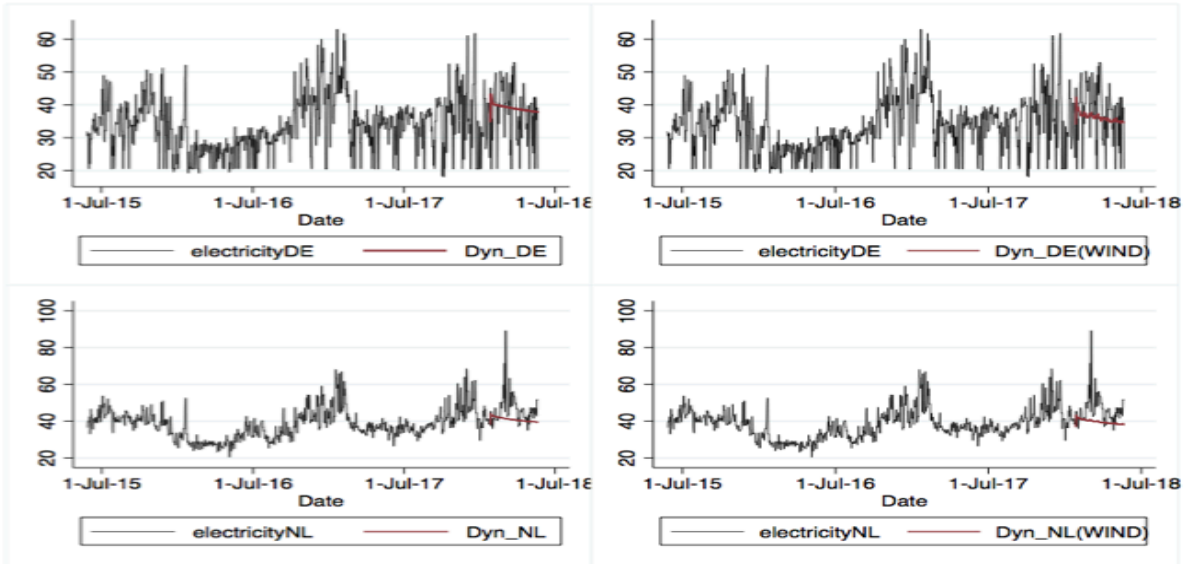
Figure 6: Dynamic forecast for both the APX and EPEX phelix day-ahead spot price.



Note: *The upper figure shows the price evolution for both known and predicted values of the autoregressive forecast in Germany and The Netherlands for the period between from 27jan2018 till 20may2018. As can be observed, the prediction is based on a dynamic forecast, rather than a one-step ahead forecast to simulate the development of the electricity price day by day. The dynamic forecast, predicted values for the day* n_{t+k}*, based on the all values before* n_{t+k-1}.

On the other hand, with the dynamic(td(27jan2018)) option, the forecasted value of electricity price for the day 27jan2018 is based on the observed value of electricity price in the day 26jan2018, but the forecast for 28jan2018 is based on the forecasted value of 27jan2018, the forecast for 28jan2018 is based on the forecasted value for 27jan2018, and so on. Thus, with dynamic forecasts, the error accumulates over time. The following graph illustrates this effect.

Figure 7: Dynamic forecast for both the APX and EPEX phelix day-ahead spot price.



Note: *The dynamic forecast shows therefore a downward curve, since it cannot predict shocks and spikes. The lower figure displays the entire time series highlighting the differences between the new forecasted value and the observed one.*

As for the static prediction, the ARMA (2,1) model is detected as the more appropriate for both countries given the lower AIC and BIC statistics. Evidences from analysis shows that, the dynamic ARMA (2,1) model is less accurate than the static one. It in fact, has not enough power to predict sudden spikes and mean reverse movements, differently from the static one which used current and past values to prompt the forecast values.

Table 11: Diebold-Mariano test for dynamic prediction in Germany

Criterion: MSE over 114 observations	
SERIES	MSE
Arima (2,0,1)	72.88
Arimax (2,0,1)	60.53
Difference: 12.350	
By this criterion, arimax (2,0,1) is the better forecast	
H0: Forecast accuracy is equal.	
S(1) =2.284 p-value = 0.001	

Note: *The table summarizes results from the DM test for ARMA models in Germany, The MSE is found to be lower in the conditional forecasting, revealing that the prediction is more accurate when wind power is included in the model. The significance at 1% level confirm the findings.*

However, in practice terms, the dynamic model is considered to be more helpful to predict and address future trends of time series.

Analysing the different results from the Autoregressive and Conditional forecast, as for the static one, the Diebold-Mariano (DM) test is additionally employed to test whether predictive values from the two different models statistically diverge in terms of accuracy.

Results show how, in Germany, the DM test reject the null hypothesis in which forecast accuracy are equal, stating that in the dynamic forecast, as the previous one, whenever exogenous variable, namely WIND power generation, is included in the model, the forecast better performs in terms of accuracy. While, on the other side, wind has a negative effect on the Dutch electricity spot price, reducing the accuracy in forecasting the price.

Table 12: Diebold-Mariano test for dynamic prediction in The Netherlands

Criterion: MSE over 114 observations	
SERIES	MSE
Arima (2,0,1)	52.89
Arimax (2,0,1)	66.9
Difference: -14.01	
By this criterion, arima (2,0,1) is the better forecast	
H0: Forecast accuracy is equal.	
S(1) =-3.955 p-value = 0.001	

Note: *The table summarizes results from the DM test for ARMA models in Germany, The MSE is found to be lower in the conditional forecasting, revealing that the prediction is more accurate when wind power is included in the model. The significance at 1% level confirm the findings.*

Table 13: Differences in MSE between the dynamic ARMA and ARMAX forecast model.

Model	*Country*	*ΔMSE*
ARMA/ ARIMAX (2,1)	*Germany*	12.350
ARMA/ ARIMAX (2,1)	*The Netherlands*	-14.01

Note*: Findings reveal the impact of wind power generation under a forecast point of view. From the table, it is shown that the wind has a greater effect in Germany than in The Netherlands, given the delta RMSE of 5%. These values are successively confirmed through the use of DM test.*

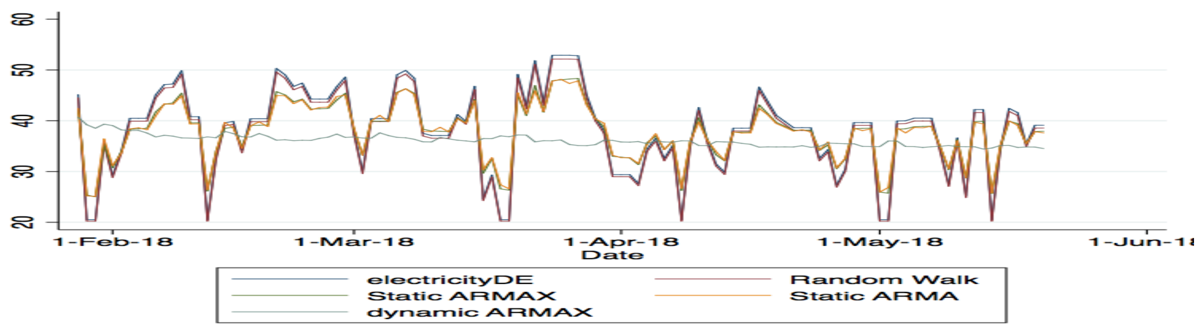
In conclusion, the evidences support the hypothesis in which, according to the merit order theory, wind power generation does influences both country in the electricity price formation, although **Table 10 and Table 13** contradict each other. The static prediction in fact, have shown how the wind generation has a higher effect in Germany than in The Netherlands. This is surely attributable to the higher number of wind turbines present in the German territory which, with an increased supply of wind power, affect more the merit order curve, ergo the merit order effect is more pronounced. However, in the dynamic prediction, although the model is considered less accurate, it displays a reverse result in terms of accuracy. In any case, what emerge from the MSE approach is that wind generation does have a positive effect on the

electricity price formation's process. It, however is not as big as one might expect from a qualitative analysis based on the energy supply mix.
Recalling the data on the Wind market share from Section 3, it can be observed in fact that, although 20% of the total German's electricity come from wind turbines, in terms of forecast, wind affect only the 0.8% in the electricity price's formation process.

6.1.5 Random walk comparison

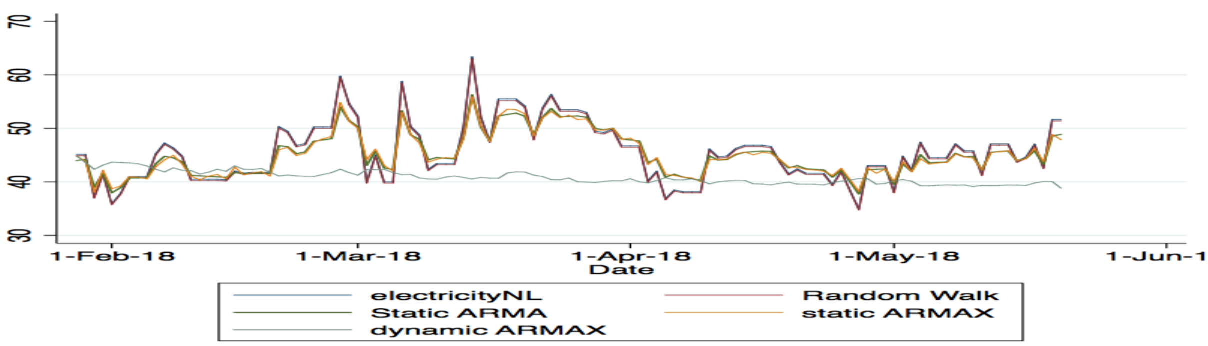
The random walk model is further elaborated in order to gauge statistical differences with the other forecast models such as ARMA and ARMAX.

Figure 8: Model comparison in Germany



Note: *The graph summarizes results from the different forecast models in Germany. Specifically, it is shown the graphical development of predicted values from the original time series, ARMA model, ARMAX model, Dynamic ARMAX and random walk.*

Figure 9: Model Comparison in The Netherlands



Note: *The graph summarizes results from the different forecast models in The Netherlands. Specifically, it is shown the graphical development of predicted values from the original time series, ARMA model, ARMAX model, Dynamic ARMAX and random walk.*

From figure **8** and **9**, it can be graphically observed that the random walk better performs the previous model in matter of accuracy, given this specific frequency of data. It can be concluded from the random walk analysis that; the electricity price is better described by a random walk than an ARMA or ARMAX process.

6.2 RELATION BETWEEN WIND POWER AND ELECTRICITY SPOT PRICE

6.2.1 DIRECT AND DELAYED EFFECT

Moving on the structural analysis, the direct and delayed effect of wind power generation on the electricity spot price is firstly analysed. Results are shown in **Table 14.** It clearly appears that wind power generation has a both a direct and delayed effect on the electricity spot price. Specifically, an increase of 1 unit of Mw/h in wind power generation can be seen as a reduction of .0006€ in the Dutch electricity spot price and .00067€ in the German electricity spot price. The following findings, therefore, although differently from other scholar who studied the merit order effect using an hourly frequency, confirm the hypothesis that a lower marginal cost in production (and hence renewable energy source) have a dampening effect on the electricity spot price, showing a negative relationship.

Table 14: The impact of wind power generation on the electricity spot price in The Netherlands and in Germany.

Parameter	**(A)** ARMAX (2,1) NL	**(B)** ARMAX (2,1) DE
wind	-.00060692*	-.00006751*
	(0.027)	(0.041)
ϕ_1	0.00038*	0.0006*
	(0.015)	(0.022)
ϕ_2	-0.00013	0.0001***
	(0.487)	(0.000)
ϕ_3	0.00016	00001
	(0.367)	(0.647)
ϕ_4	0.00031	0.0002
	(0.868)	(0.460)

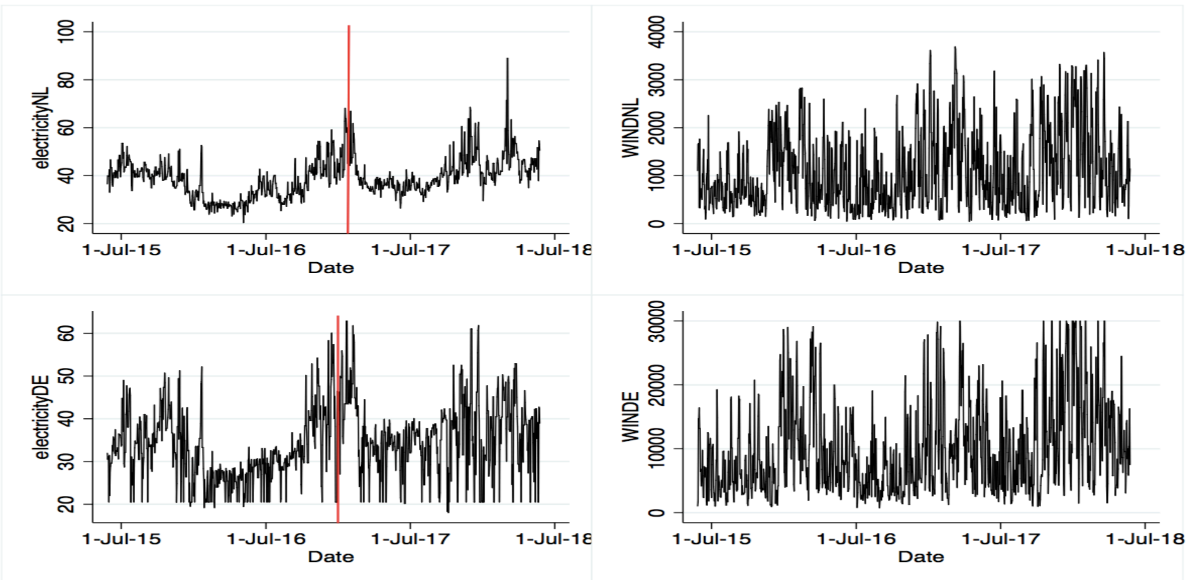
Note: *The table above shows the results from an ARMAX model in which the wind power generation variable is added as exogenous factor. The ϕ parameter indicate the lagged variable, respectively until lag 4. The model is computed for both The Netherlands and Germany. From the table, it can be inferred that, while "wind" has a negative effect on the electricity spot price, delayed "wind" behave oppositely.*

However, from the table is noticeable how, once wind power generation is delayed, it positively affects the electricity spot price. As the number of lags increase, however, the relationship become weaker and not statistically significant. Findings, therefore highlight how the movement between the electricity price and its driver, diverge as soon as the exogenous variable is lagged. The intuition behind is that while the electricity spot price is determined by direct and indirect factors, as such its power production or temperature, Wind turbines produces electricity only based on the wind speed. Hence, wind power generation in t_{k-1} should not influence the electricity spot price in time t_k, since they are not directly correlated.

6.2.2 Structural breaks

Findings from the structural breaks analysis have instead shown how both APX electricity spot price and EPEX-PHELIX spot price experienced a structural change over time in their relation to the wind power generation. Specifically, with regard to the electricity spot price in The Netherlands, results from the post-regression wald test, force to reject the null hypothesis of no structural breaks during the sample period and detects a break in the first week of November 2017. As for The Netherlands, the EPEX Phelix performs, according to the test, another structural break during the day 8 October 2016, changing therefore is co-movement with the wind generation.

Figure 10: Structural breaks present within the electricity spot price for both The Netherlands and Germany.



Note: *the graph displays structural breaks in the APX and EPEX-phelix spot price development. The right graphs show the wind power generation development for both countries.*

The identified structural changes might have been driven by different factors such as the increasing production of other power plants or the undergoing de-regulation in the market and in whole energy system. It could be argued however, that, since wind power generation and electricity spot price manifest the relationship "cause-consequences", structural break could be attributed to all factors that might affect the wind turbines production, hence climate and technology factor. In order to deeply understand the cause of these structural break, a comparative analysis, focused on wind production and electricity spot price development should be studied. In any case it is important to keep in mind that, whenever these structural breaks were taken into consideration in the analysis, results could be different for many aspects. This paper, however, since aims to understand the real and observable effect during the years of wind power on the electricity price, have limited the structural analysis as a merely add-on in understanding the wind-energy price relationship.

6.2.3 VOLATILITY EFFECT

Given the pronounced volatility in the liberalised markets, the conditional heteroscedasticity model helps explaining the price performance. Based on Huisman (2008), which recognised the need to enrich the price model with fundamentals as temperature variables to detect changes in price behaviour, this research estimates a GARCH model for both the Netherlands and Germany adding the variable “wind power generation” to the model. Recalling the previous methodology’s chapter, the GARCH model helps describe time series in which volatility can change over time, becoming more or less pronounced depending on different period of time. The term “Heteroscedasticity”, describe in fact the irregular patters of variation of the error term, in other words the inconsistency of linear patterns on the time series. Volatility therefore could tend to be clustered.

The purpose of this sub-section is analysing whether changes in volatility of the electricity day-ahead spot price can be attributable to changes of the wind power generations. Hence whether the volatility of the electricity spot price is conditioned by past values of itself, model errors and successively whether the volatility is conditioned by the wind. The models, however, as noticeable, differently from other scholars are not used in matter of prediction and forecast, but solely to estimate whether the electricity is conditional to other variables. The analysis will therefore stop once all GARCH parameters are estimated, focusing on the significance or insignificance of them. Reminiscing Patton et al. (2000) “no-one believes that financial assets prices and most generally financial time series evolve independently of the market around”, the paper expect that, given the described merit order theory in matter of electricity price formation, other variables such as wind may contain relevant information for the volatility of the series. (Patton, 2000). The paper, on the stream of the literature, applies a AR-GARCH (1,1) model.

Results from the estimations for the German and Dutch day ahead electricity spot price can be found in **Table 15**. The first columns, namely column (A) and (B) shows the benchmark specification for models solely determined by using the electricity spot price.

All coefficients are highly significant, the variance parameters are all positive, and their sum is smaller than one. Findings show that the volatility of the time series is quite persistent, given the sum of the key statistic of the two main parameters α and β of 0.97 (Ketterer, 2014). The sum of them is significantly less than one meaning that, the volatility process does return to the mean. Volatility clustering in fact, implies that volatility hovers through a mean level.

Mean reversion in volatility is therefore generally interpreted as the fact the current information has no influence on the long run forecast (Patton, 2000).The size of the GARCH term β with 0.92 indicates that the autoregressive persistence β is particularly strong for the electricity price. The GARCH term α on the other side reflects the impact of new shocks the conditional variance h_t, transmitted though the error term ε_t from [4]. Looking at the mean equation The AR term depicts a specificity of the power market. The coefficient of 0.88 in (A) shows that the price reverts back to its long-run mean. But the speed of reversion, given by $1 - \varphi_1$, is low. (Ketterer, 2014).

Table 15: Estimation results of traditional AR-GARCH and ARX-GARCHX models

Parameter	(A) GARCH (1,1) NL	(A*) ARX-GARCHX (1,1) NL	(B) GARCH (1,1) DE	(B*) GARCH-X (1,1) DE
		Mean equation		
μ	36.858***	37.116***	31.865***	31.912***
	(0.000)	(0.000)	(0.000)	(0.000)
AR (1)	0.886***	0.880***	0.7430***	0.631***
	(0.000)	(0.000)	(0.000)	(0.000)
w		-0.00027		-0.0001
		(0.053)		(0.553)
		Conditional variance equation		
ω	0.665***	0.680**	0.478***	0.597
	(0.000)	(0.001)	(0.002)	(0.166)
α	0.125***	0.139***	0.060***	0.055***
	(0.000)	(0.000)	(0.000)	(0.000)
β	0.843***	0.791***	0.927***	0.935***
	(0.000)	(0.000)	(0.000)	(0.000)
γ		0.006**		0.0004
		(0.000)		(0.531)

Note: The table shows that all parameters are highly significance. The mean equation, built up through an AR process shows how the variable WIND influence the electricity spot price, while the AR parameters can be intended as the speed of the mean reversion. The conditional variance equation [7], instead shows how past values and the WIND variable (γ) influence price volatility.

In column **A***, the exogenous wind variable is included in the mean as well as the variance equation of the GARCH (1,1). The negative coefficient for the wind variable shows that the day-ahead price decrease when high wind electricity generation is forecasted which is in line with findings by Jonsson et al. (2010), Woo et al. (2011) and the merit order effect.

However, the coefficient less than 0,002% tell us how weak the positive relationship is displayed with this frequency of data and with this particular model. In the present specification, the coefficients can be interpreted as elasticities. When the wind electricity in-feed (MWh per day) increases by 1 per cent, the price volatility changes accordingly between 0.004 and 0.006. In the variance equation, the wind variable is significant different from zero and positive. Hence, the fluctuating wind in-feed increases the volatility of the electricity price. To make sure that these results are not driven by the outliers that remain an outlier dummy is included in all mean equations. A similar and parallel picture arises in column (B) and (B*) with respect for the Germany. If the volatility clustering is adequately explained by the model, three indicators need to be confirmed.

- The standardized residuals from the GARCH model should follow a normal distribution
- The standardized squared residuals from the GARCH model should not be auto-correlated
- No remaining ARCH effect on the residuals

In checking the followed criteria, the Shapiro-Wilk and Jarque-Bera tests are used for normality, Ljung-Box Q statistic for correlation and ARCH LM test for the remaining effect. From the diagnostic check, the volatility clustering remains for 2 out 4 models, specifically from the auto-correlogram and Ljung-Box Q statistics, it is showed that standardized residuals, although all variables seem to be significant, are serial correlated from lag 6. Once the mean equation is correctly specified, the residuals diagnostic of the GARCH model should indicate the absence of serial correlation as well as no ARCH effect in the residuals.

The analysis however, sheds lights on the fact that, given the studied dataset and its consequent frequency of data, models in which wind power generation is added in the variance equation are not serially supported by diagnostic test. On the contrary, the GARCH models in which only the dependent variable "electricity day-ahead price" is used, accept all three described indicators of validity and hence it can be confirmed that values at t-1 statistically influence the volatility at the time t. What emerges, is however that, every GARCH model rejects the model's requirements, regardless changes in the mean equation or transformation in the time series.

Looking back the ACF and PACF functions of residuals, it clearly emerges that problems occur in correspondence with the 7 lags, which coincide with the above describe seasonal pattern.

Additionally, dummy variables are implemented yet, although standard residuals become no serial correlated the ω parameter become negative symbolizing a negative variance.

To conclude, findings from findings from the mean equation show that, as other scholars and past researcher, the AR-GARCH model confirm the hypothesis that wind power generation as negative effect on the electricity price. In contrast, ARX-GARCHX shows how the electricity's volatility is influenced by its past values in time t-1 and by the exogenous wind power generation variables, although the model lack in accuracy.

Section 7: Discussion and conclusion

In conclusion, the present paper tried to decipher few dilemmas with regard to the impact and the extent by which renewable energy sources affect the electricity price formation. Modelling the electricity spot price has presented numerous pitfalls, mainly driven by the intrinsic and unmistakable characteristic that only the electricity presents. The econometric excursus of the electricity price formation faced through the employment of forecasting and volatility models shed light on results which to some extent could diverge from the existing literature in matter of electricity's behaviour in the market.

The impact of wind generation has been tracked following a distinctive and new approach based on forecast techniques to show how electricity and wind generation can be controlled and manipulated in all their characteristics. Supported by the literature and by the merit order theory, wind power generation, in both The Netherlands and Germany increase the forecast accuracy in predicting the day-ahead electricity spot price.

The observed incremental accuracy is displayed, greater in Germany than in the Netherlands, which is in line with the higher wind energy disposal presented in the German country. On the other side, with regard to the RO2, evidences show that the relationship between the electricity spot price and wind power has changed over time. However, it has been found that, wind power generation has a direct negative effect on the electricity spot price, dampening his price and a positive one on the electricity day-ahead price's volatility.

7.1 Limitation of the research

Although the present research aimed to improve the existent literature in matter of renewable energies, few limitations are observable. In this paper, the effect of the wind power generation is treated as the effect of an exogenous variable in a forecast model. Specifically, the impact of wind power is captured by looking at the incremental accuracy between an autoregressive and a conditional forecast.

The electricity price, however, is determined by numerous factor in the market, and although forecasts can be considered as sophisticated techniques, they leave back few aspects which a multiple regression might capture. Due to the lack of available public data, the research could not have gathered all information necessary to elaborate a sophisticated wind model. It might be interesting in fact, collect updated data on actual and forecasted wind speed, wind generation and electricity spot price in order to understand how much clean electricity is wasted and how much the wind forecast errors influences the imbalance and the spot market.

Moreover, the time series are elaborated over average data, which might result in a dispersion of data throughout the modelling process. The research therefore is considered limited by the dataset and consequently by the lack of available public data. Moreover, this research has showed few constraints related to the volatility model, highlighting the lack of power of the GARCH model. Further advanced volatility model would therefore improve the result capturing few aspects which are not fully explained by the proposed model.

7.2 Further studies

In pursuing the aim of reducing carbon emission, the increasing market share acquired by renewable energy is changing the general outlook of the whole energy market and sub-sequentially everything that surround it, including our day-life routine. Findings, overall, reveals that wind power generation does not only influence the electricity price formation but also affect its volatility. Retracing the hypothesis developed in Section 3, with regards to the RO1, the paper highlighted the fact that wind power generation has a positive impact on the electricity day-ahead price formation

The research sheds light on possible future studies in understanding how wind power generation and more generically renewable energy sources affect the electricity market. Given the undergoing energy revolution, further studies on all renewable energy should be conducted, more on the consumer level. Specifically, how much, this dampening effect might increase the consumer utility and how much its inconstant volatility might danger it. The trade-off between volatility effect and reduction in price is therefore, still a dilemma and considered as both a limitation of this paper and a possibility of new studies.

The higher volatility however, might have consequences in terms innovation and "green" investments. It could be argued in fact that, higher is the volatility, higher is the uncertainty around RES, which according to Pindyck (1994) might lead to a delay and decay of investments. Surely, in order to achieve perfect synchrony between the electricity price and renewable energy sources, devices such as batteries need to be developed and improved to offset the uncertainty given by these intermittent energy sources such as the wind power.

Appendix

The following tables refer to the static and dynamic forecast. All forecasts are based on an ARIMA (2,0,1) given the lower AIC criterion.

TABLE A1: STATIC FORECAST IN GERMANY

Variable	ARIMA (1,1,1)	ARIMA (2,1,1)	ARIMA (1,0,2)	ARIMA (1,0,1)	**ARIMA (2,0,1)**
PRICE	.00362857	.00360431	33.756450***	33.757267***	33.759725***
α1	.62369492***	.62102138***	-.3261758***	.80059227***	1.6080284***
α2		.0054723			-.61220957***
β1	-.9580303***	-.95851739***	-.17446981***	-.119544***	-.94884256***
β2			-.10157775***		
σ	5.4090654***	5.4089609***	5.4063001***	5.4089343***	5.3951105***
Statistics					
AIC	6072.852	6069.255	6070.831	6097.249.	6071.012***
BIC	6087.3823	6094.377	6100.126	6121.667	6095.43
LL	-3029.9263	-3029.912	-3029.41	-3043.624	-3030.506

Table A2: STATIC CONDITIONAL FORECAST IN GERMANY

Variable	ARIMA (1,1,1)	ARIMA (2,1,1)	ARIMA (1,0,2)	ARIMA (1,0,1)	**ARIMA (2,0,1)**
PRICE	.0029082	.00287954	33.076921***	.33.164623***	33.078905***
WIND	-.00006632*	-.00006656*	-.00007016*	-.00006132*	-.00006751*
α1	.61667641***	.62102138***	.86031755***	.80314038***	1.6003392***
α2		.0054723			-.60470707***
β1	-.95645335***	-.95851739***	-.19083161***	-.12934762**	-.94689404***
β2			-.10901773**		
σ	5.3987904***	5.4089609***	5.4574361***	5.4741253***	5.3844245***
Statistics					
AIC	6069.1353	6068.0778	6095.1032	6099.0424	6066.1494***
BIC	6090.5474	6097.3725	6124.404	6123.4597	6098.4501
LL	-3028.0676	-3028.0389	-3041.5516	-3044.5212	-3028.5747

TABLE B1: STATIC FORECAST IN THE NETHERLANDS

Variable	ARIMA (1,1,1)	ARIMA (2,1,1)	ARIMA (1,0,2)	ARIMA (1,0,1)	**ARIMA (2,0,1)**
PRICE	.00196901	.001911095	37.98242***	37.91167***	38.25725***
α1	.62369492***	.62102138***	-.971758***	.9289227***	1.4080284***
α2		.0054723			-.41220957***
β1	-.9580303***	-.95851739***	-.4026981***	-.339544***	-.84884256***
β2			-.19157775***		
σ	3.4090654***	3.4089609***	3.7063001***	3.8089343***	3.3951105***
Statistics					
AIC	5345.8437	5341.4927	5374.1092	5393.2752	5341.1634 ***
BIC	5361.3734	5365.9049	5398.5265	5412.8095	5369.5807
LL	-2666.9218	-2665.7463	-2682.0546	-2692.6378	-2667.5817

TABLE B2: CONDITIONAL FORECAST IN THE NETHERLANDS

Variable	ARIMA (1,1,1)	ARIMA (2,1,1)	ARIMA (1,0,2)	ARIMA (1,0,1)	**ARIMA (2,0,1)**
PRICE	.00157866	.00149163	37.257313***	37.211766***	37.529185***
WIND	-.00059397*	-.00058648*	-.00065678*	-.00064041*	-.00060692*
α1	.50112434***	.49457504***	.97129037***	..9286571***	1.4782355***
α2		.05293256			-.48407324***
β1	-.8929928***	-.90782807***	-.40480393***	-.33521244***	-.8781053***
β2			-.19247019***		
σ	3.7112811***	3.7070924***	3.7530007***	3.8089343***	3.7008177***
Statistics					
AIC	5337.817	5334.6301	5364.916	5384.5637	5334.7452***
BIC	5359.2292	5363.9247	5394.2167	5408.981	5367.0459
LL	-2662.4085	-2661.3151	-2676.458	-2687.2818	-2662.8726

BIBLIOGRAPHY

Alaa Abi Morshed, E. A. (2015). *Structural break tests robust to regression misspecification.* Tilburg University,University of Cyprus,.

AL-Najjar, D. (2016). *Modelling and Estimation of Volatility Using ARCH/GARCH Models in Jordan's Stock Market.* Asian Journal of Finance & accounting, Applied Science University , Dept. of Finance and Banking.

Bellona. (2012). *Renewable energy policies in Europe .* The bellona Foundation.

John Y. Campbell, A. W. (1997). *The Econometrics of Financial Markets.*

Chaleampong Kongcharoen, T. K. (2013). *Autoregressive Integrated Moving Average with Explanatory Variable (ARIMAX) Model for Thailand Export.*

Chen, Z., & Yang, Y. (2004). *Assessing forecast accuracy Measures.* Iowa State University, Department of Economics;Department of Statistics.

Christoph Böhringer, T. F. (2009). *THE EU 20/20/2020 targets: An overview of the EMF22 assessment.*

Di Cosmo, M. V. (2012). *Relation between wind and electricity prices in a deregulated market: the case of Ireland.*

Dong, Y. (2012). *ARMA and GARCH-type modeling: Electricity prices.* Gotenborg University, Department of mathematical Sciences.

Elisabeth Woschnagg, J. C. (2004). *Evaluating forecast accuracy.*

ENTSOE. (n.d.). Retrieved from https://transparency.entsoe.eu/

epex spot market. (2018). *Description of the epex spot markets indices.*

Erni, D. (2012). *Day-ahead electricity spot prices- Fundamental modelling and the role of Expected wind electricity infeed at the european energy.* Dissertation, University of St.Gallen.

European Parliament, Council of the European Union. (2009). *DIRECTIVE 2009/28/EC.* Official Journal of the European Union.

Fair, R. C. (1986). *Evaluating the predictive accuracy of models.*

Fernando L. Aiube, T. K. (2013). *Modeling Hourly European Electricity Spot Prices via a SARMA-GARCH Approach.*

Geninasca, D. (2012). *Wind power Feed-In- Effects on prices.* University of St.Gallen(HSG).

Georg Wolff, S. F. (2017). *Short-term dynamics of day-ahead and intraday electricity prices.* National Institute of Informatics, Tokyo.

Jesus Crespo Cuaresmaa, J. H. (2004). *Forecasting electricity spot-prices using linear univariate time-series models.* University of Vienna, Institute for Advanced Studies, International Institute for Applied Systems Analysis, Department of Economics, Department of Economics and Finance, Department of Finance. Elsevier.

Jonsson, T. (2008). *Forecasting electricity spot prices accounting fro wind power predictions.*

Ketterer, j. C. (2014). *The impact of wind power generation on the electricity price in Germany.*

Kruangpradit, C. K. (2013). *Autoregressive Integrated Moving Average with Explanatory Variable (ARIMAX) Model for Thailand Export.* Thammasat University, Faculty of Economics.

Nicolosi, M. (2010). *Wind power integration, negative prices and power system flexibility — an empirical analysis of extreme events in Germany.* Institute of Energy Economics EWI, Cologne.

Makridakis, S. (1997). *ARMA models and the Box jenkins.* Neapolis University.

Mandelbrot, B. B. (1963). *The Variation of Certain Speculative Prices.* The Journal of Business 36, No. 4.

Moreno, B., & Lopez, A. J. (2011). *The impact of renewable energies and electric market liberalization on Electrical prices in the Europan Union. An econometric Panel Data Model.*

Nau, R. (2018). *Fuqua School of Business Duke University*. Retrieved from https://people.duke.edu/~rnau/411arim3.htm

Nicholson, R. P. (2010). *Relationship between wind generation and balancing energy market prices in ERCOT: 2007–2009.* . National Renewable Energy Laboratory (NREL).

Partanen, S. V. (2013). *Price Forecasting in the Day-Ahead Energy Market by an Iterative Method with Separate Normal Price and Price Spike Frameworks.* Lappeenranta University of Technology.

Patton, R. F. (2000). *What good is a volatility model?* NYU Stern School of Business, University of California, Department of Finance, Department of Economics.

Pindyck, A. K. (1994). *Investment Under Uncertainty.*

Rafał Weron, A. M. (2008). *Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models.*

Rintamäki, T. (2013). *Impact of renewable energy on electricity prices – comparative analysis of Denmark and Germany.* Aalto University.

Seán Collins, B. Ó. (2015). *Quantifying the "merit order" effect in the european electricity market.*

Statista. (2018). Retrieved from www.statista.com

Stefano clò, A. C. (2014). *The merit order effect in the italian power market: The impact of solar and wind generation on the national wholesale electricity prices.* Department of treasury of the italian ministry of economy and Finance, Italy.

Tina Jakasa, I. A. (2011). *Electricity price forecastin- ARIMA model approach.*

Tryggvi Jo ́nsson, P. P. (2008). *Forecasting Electricity Spot Prices Accounting for Wind Power Predictions.*

U.S Energy Information Administration. (2018). *Levelized Cost and Levelized Avoided Cost of New Generation Resources in the Annual Energy Outlook 2018.*

Unicorn Systems a.s. (2018). Retrieved from www.entose.eu: https://transparency.entsoe.eu/

Weron, R. (2014). *Electricity price forecasting: A review of the state-of-the-art with a look into the future.*

Wirdemo, A. (2017). *The impact of wind power production on electricity price volatility: A time series analysis.*

Wolff, G. (2017). *Short-term dynamics of day-ahead and intraday electricity prices.*

Yan, D. (2012). *ARMA and GARCH-type modelling electricity prices.* Master thesis, Gotenborg University, Department of Mathematical Sciences, Gotenborg.

Zhiwei Shen, M. R. (2015). *Forecasting volatility of wind power production.* Humboldt-Universität zu Berlin, Germany.