UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

STUDY ON NEGATIVE ELECTRICITY PRICES AND ITS IMPACT ON WIND FARMS OPERATION

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

By

RUTH NATALIA ANGARITA TARAZONA
Norman, Oklahoma
2016

STUDY ON NEGATIVE ELECTRICITY PRICES AND ITS IMPACT ON WIND FARMS OPERATION

A THESIS APPROVED FOR THE
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

BY

____
Dr. Ning Jiang, Chair

____
Dr. Tian You Yu

____
Dr. Keith Brewster

The older I get, the more I understand that there is nothing cheesy in dedicating a thesis to your love ones, but it is certainty a shame that it took me so long to figure that out. Everything I am is because of my family and no one else in the world could care more deeply about me, celebrate my achievements and hold me together during tough times.

This thesis is dedicated to them, the people who would love me the same with or without a master's degree.

I would like to express my special appreciation and thanks to my advisor Professor Ning Jiang for his valuable and constructive support during the development of this thesis. His willingness to help me put things in perspective not only make this thesis possible but definitely made me a better engineer.

Also, I would like to express my profound gratitude to everyone in the office: Milad Javadi, Xin Li, Dr. Di Wu, and Diego Rodriguez, and everyone in the ECE department; their amazing attitude and willingness to help me succeed have taught me the right way to treat people, work ethics and how knowledge connects people.

Table of Contents

List of Tables

List of Figures

Abstract

The relation between Negative Electricity Prices (NEPs) and Wind Power Producers (WPPs) in the wholesale market, shows the complex paradox of the integration of wind power, and renewable energy in general, in the day to day power system operation. The occurrence of negative prices is an intended price signal to inform the generation to reduce its output level, or the cost to be charged to wind generation for congestion management. Thanks to the intermittency of the wind and the encouragement of the government in the form of a Production Tax Credit (PTC) to produce wherever there is wind; WPPs often found themselves bidding and generating at negative prices. However, negative electricity prices are not reflecting that there is something wrong with wind power or the system per se, they reflect that wind power and other renewable technologies have new and unforeseen characteristics that are challenging the status quo of the system [1]. There is no doubt that the power sales and distribution system must adapt, however, the complete restructuration of the power system would take decades, which sharply collides with the fast-growing government encourage integration and investment on renewable energy.

Wind farms in the US could possibility take advantage of this situation, given that new alternatives to the current bid and operational strategy of WPPs have emerged. The introduction of the Investment Tax Credit (ITC), a federal incentive that is not tied to the amount of power generated by the wind farms allows wind farms to explore different alternatives to maximize their sales revenue. Additionally, the fact that new wind plants have more control to limit ramp rates and its output, make it possible for wind farms to implement predictive control avoiding then to generate at negative

electricity prices. In this thesis, a number of studies are conducted to investigate the nature of negative electricity price drivers since NEPs are a fairly new phenomenon and this is crucial information to effectively forecast them and migrate them; and to examine the economic feasibility of predictive control for WPPs receiving ITC subsidy. The studies evaluate the current impact of negative electricity prices on WPPs' sales revenue to further back up predictive control (curtailment at negative electricity prices) as an economic sound alternative for WPPs receiving ITC subsidy. Moreover, a fairly simple three steps discrete approach is proposed. The main aim of the proposed approach is to capture the nature of the negative price drivers through its influence in the negative prices' behavior. Last of all, the economic value of appropriately represent the nature of negative electricity price drivers for a wind farm is studied.

The results of the study suggest that given the current market structure, the negative price drivers show indication of being of a discrete kind. Furthermore, this finding has a positive impact on the sales revenue of a wind farms, as it helps to better describe and predict negative prices. Summarizing, results are promising towards a predictive control as a feasible- economic sound alternative for WPPs receiving ITC subsidy. Potentially, these findings could lead to a proper forecast of negative electricity prices that considers important characteristics like the nature of their drivers; that can be used by WPPs receiving ITC subsidy to apply predictive control. Predictive control would directly economically benefit wind farms, and indirectly improve system stability by reducing the occurrence of negative electricity prices by decreasing the number of wind farms that bid and generate at negative electricity prices.

Chapter 1: Introduction

Most wholesale electric markets are designed as a two-settlement system with a day-ahead market (DAM) and balancing/real-time market (RTM); where generators submit a bid before the market closes and then, the Independent System Operator (ISO) uses a full network model to determine least- cost unit commitments and market clearing prices taking into consideration demand, supply and constraints for the next day (DAM) or hour ahead (RTM). These constraints can be physical (generation and transmission), reliability related, financial and even political. Generation constraints are due to ramp up/down times, minimum down/up time and schedule or unscheduled down times. Similarly, transmission constraints concern to the capacity of the transmission lines to transport energy between different nodes of the grid. Reliability constraints ensure the quality of the energy sold in the market (voltage and frequency levels) and the stability of the system. Financial constraints represent the constrains due to the generation mix and the variability of their fuel prices. Last of all, political constraints in the form of policies, development plans and incentives for certain generation technologies, transmission grid expansion, etc [2].

The aggressive integration of renewable energy on the last 20 years and the subsequent rise of negative electricity prices are just two examples of how those constraints shaped and drive the power market. The dependency on foreign fossil fuel, the variability of its price, its non-renewable nature and the environmental damage of burning fossil fuels has lead most countries around the world to create and are currently implement long term plans to obtain a significant percentage of its electricity from non-polluting sources i.e. renewable sources [3]. North America is committed to obtain 50 percent of its electricity from non-polluting sources by 2025 [4], the EU Renewables Energy Directive2009/28/EC dictates that by 2020, 20% of all energy consumed by the European Union must come from a renewable source of energy [5], and in Australia ,the Renewable Energy Target has been set at 23.5% for 2020 [6]. The main driver of these initiatives is the protection of the environment by reducing greenhouse gas emissions and the dependency on high contaminant-nonrenewable fossil fuel power plants while satisfying the energy needs of current and future generations. With increasing penetration of renewable energy, negative electricity prices emerge as a sign of temporary excess of generation. Furthermore, NEPs are a sign of the mismatch between the current market design and the characteristics of the renewable generation introduced to the generation mix.

Section 1.1. Negative Electricity Prices: Characteristics and Causes

Negative electricity prices can occur consecutively or intermittently during a period of hours or days. In the real price data used in this thesis, the intermittent characteristic works in two ways, within short periods (hours) there will be high concentration of negative electricity prices, and sporadically there would be positive prices among them, and within long periods (days) there will be “islands” of high concentration of negative electricity prices surrounded by long periods of positive prices. There are many conditions under which negative prices are more likely to occur, and many actions that lead to negative prices. One of the most observed and reported conditions under which negative prices are more likely to occur are high wind generation and low demand scenarios. This is something that was also observed in the real sample data used in this study (data description can be found on sections 2.1.2 and 4.2) and those “islands” of negative prices can be seen as reflection of these episodes.

High wind generation and low demand episodes imply the commitment of wind farms during times where historically only base load units are committed, further suggesting negative bidding by wind farms, and other conventional generation with high shut down cost. Wind power is known to have a near zero production cost, this allow WPPs to bid at near zero prices. Moreover, most of WPPs receive some type of subsidy, the most popular one is called Production Tax Credit or PTC and it is tied to the amount of generation produced. Therefore, even though WPPs have the option to sell at negative prices or curtail production in a sense they are encouraged to produce as long as there is wind with no regards of electricity prices (as long as the negative prices are not higher than the PTC minus operation and maintenance cost).

On the other hand, base load units have a lower production cost (and a high shut down cost) than more flexible dispatchable units. It may not be profitable for base load plants to generate during negative electricity price episodes, but higher losses may arise if they are shut down due to their high shut down cost, minimum down time and/or required spin-up time that may prevent them to make a profit when prices are higher. Subsequently, base load plants may sometimes bid at negative prices to be committed too. In addition to the modest incentives to apply curtailment and the ambitious greenhouse gas reduction goals posed by the governments, wind power is perceived as a type of generation with little or no control on its output, therefore, other types of generation are directed to apply curtailment first. This translates in less demand of base load plants with slow ramping capability like thermal or nuclear plants, more demand of flexible generating resources like gas and combined cycle plants; and more downward reserve. Downward reserve is the type of generation power reserve that can reduce its output when the generation is exceeding the demand or the type of load that can increase their consumption during those episodes [7]. This challenges the flexibility of the current system as the steady base load plants are still needed for stability reasons and transmission constraints may prevent the use of most flexible resources [8].

In addition, if the market is forced to take in all renewable generation or there is no penalty for deviations from previous commitments in the DAM, forecast errors in demand, wind and weather in general exacerbates this type of situation creating perfect scenarios for negative prices to happen. The following are some of the solutions that have been proposed to reduce the likeliness of negatives prices: incentives for wind farms who apply curtailment as new wind plants have more control to limit ramp rates and their output; price responsive demand that is ready to go on and off line, to reduce peak load or take excess of generation during high wind generation and low load events; energy storage units so the excess of generation can be stored and later used at peak hours; and more investment in flexible - quick start generation and reserves so the system can bear those type of events [9][10].

Section 1.2. Relation Between Negative Prices and Negative Bidding by Wind Farms

Negative electricity prices are a tool of the market to indicate excess of generation and subsequently, inflexibility on the system to effectively deal with it. Not long ago, many markets did not even allow negative bidding by power producers. The bid is meant to indicate the lowest cost at which a power plant is willing to generate, and it is assumed that at its lowest, it is near the marginal power plant production cost. Traditional power plants have a production cost related to the price paid for the non-renewable energy source (i.e., fossil fuel, gas, etc.) and therefore it was unthinkable that a market participant would make negative bids. However, in the case of renewably energy alternatives, the energy source is virtually free (i.e., they are willing to generate at near zero prices) and when receiving an incentive or subsidy tied to their generation, they are willing to generate (and consequently bid) at negative electricity prices.

In the US, the renewable energy alternative with highest capacity installed, who historically have received a subsidy tied to its production and have the biggest participation in the power markets at least in Midwest and Southwest is the wind power. Therefore, the occurrence of negative electricity prices is more linked to this technology than any other renewable energy technologies. The main disadvantage of wind power is its intermittency that is out of the control of the producer. WPPs can only generate when there is wind and unfortunately there is generally more wind at off peak hours, i.e., WPPs generate more at off peak demand hours than on peak hours. While traditionally on peak hours and high demand seasons like summer, the market operator is concerned about having enough generation to meet the demand and market participants are willing to pay higher prices for it. On off peak hours, there is more generation than demand and therefore, only the generation with lowest bid gets dispatch. Ideally, all types of generation want to generate at on peak hours and at off peak hours only if it is profitable. Then, it is not that WPPs want to bid at negative prices but rather they are monetarily motivated by the PTC to do it.

Section 1.2.1. Impact of Federal Incentives for Wind Power Producers (WPP)

As it was mentioned before most wind farms in the US receive a government subsidy called Production Tax Credit (PTC), in order to spur capital investment in wind power plants to meet non-polluting electric generation goals. This tax credit grants a reduction in the income tax per every MWh produced by the wind farm. Alternatively, starting in 2008, the federal government is offering to wind farms another type of subsidy called Investment Tax Credit (ITC), this tax credit is independent of generation of the wind farm. PTC is inflation-adjusted every year and it is granted to qualifying technologies as long as the energy produced is sold to an unrelated party. Historically (since 1992), this has been the only type of federal subsidy available to wind farm developers. The subsidy has a duration of 10 years (starting from the first day of operation) and for wind farms projects starting construction on 2016 is equal to $0.023/kWh [11]. On the other hand, the Investment Tax Credit (ITC), for qualifying technologies, including the wind farms represents a tax deduction of 30% (for wind farms projects starting construction in 2016) of the project's qualifying costs (new equipment) spread evenly around the first five years of operation [12].

While PTC has been found to be more profitable, ITC gives more flexibility and has less associated risk. A study from LBNL[1] and NREL[2] in 2009 showed that the higher the net capacity factor and the lower the installed cost per kWh the more profitable PTC is in comparison to ITC. Nevertheless, other factors like the associated higher risk of relying on the performance of the wind farm and the weather for PTC, the freedom of selling energy to related parties and the shorter time of return of ITC favor ITC [13].

[1] Lawrence Berkeley National Laboratory
[2] National Renewable Energy Laboratory

Particularly, from the occurrence of negative electricity prices point of view, ITC would give more freedom to wind farms and WPPs in general to incorporate different bidding strategies to maximize their profits in the wholesale market; given that the subsidy is not tied to the power production of the wind farm.

Section 1.3. Motivation of the Study

It is imperative to investigate the impact of negative electricity prices on the cash flow of a wind farm to further evaluate the economic feasibility of predictive control. Moreover, the predictive control would require the forecast of negative electricity prices. Then, it is necessity to first explore the nature of negative electricity price drivers since NEPs are a fairly new phenomenon and this is vital information to effectively forecast them and to migrate them. Based on the LMP pricing mechanism, it can be seen that occurrence of negative prices is an intended price signal to inform the generation to reduced its output level, or the cost to be charged to wind generation for congestion management. Moreover, the factors that may cause negative prices in practice can be summarize in three categories: lack of dispatchability of generation for power balancing, transmission congestion and forecasting errors. Because the diversity and variation of the factors in these categories, it is difficult to identify the factors that actually cause the negative prices for the given market condition and the given time. However, based on the understanding of the power system and power markets, most of these variables are affected by the fundamental demand and supply situation.

The relation between negative electricity prices and WPPs in the wholesale market, shows the complex paradox of the integration of wind power and renewable energy in general in the day to day power system operation. There is no doubt that the current system needs to plan and invest in transmission, energy storage, reserve and generation to adapt itself to the intermittency of renewable energy that is due to the intermittency of its energy source. However, the complete restructuration of the power system would take decades, which sharply collides with the governments' time lines for deployment and integration of renewable energy sources. Wind farms in the US could possibility take advantage of this situation, given that new alternatives to the current bid and operational strategy of WPPs have emerged. First, the introduction of ITC, a federal incentive that is not tied to the amount of power generated by the wind farms allows wind farms to explore different alternatives to maximize their net income. Second, new wind plants have more control to limit ramp rates and their output, further enabling wind farms to implement predictive control. These two conditions, potentially would allow wind farms to actively benefit from and in some levels, mitigate the occurrence of negative electricity prices.

Section 1.4. Brief Description of the Study

The study concentrates on evaluating the current impact of negative electricity prices on WPPs' sales revenue to further back up predictive control (curtailment at negative electricity prices) as an economic sound alternative for WPPs receiving ITC subsidy. Moreover, a fairly simple three steps approach is proposed. The main aim of the proposed approach is to capture the nature of the negative price drivers through its influence in the negative prices' behavior. It starts by identifying the groups or intervals of prices with high concentration of negative prices. And then, those groups are modeled as a Second Order Markov Process (SOMP), a discrete model that slightly accounts for the information carried by previous data points. Finally, the economic value of the appropriate representation on the nature of negative electricity price drivers for a wind farm is studied. All results are promising towards a predictive control as a feasible- economic sound alternative for WPPs receiving ITC subsidy. The proposed approach shows that under the current system, the nature of the negative electricity price drivers is appropriately represented as discrete. Furthermore, this finding has a positive impact on the sales revenue of a wind farms, as it helps to better describe negative prices.

This document is organized as follows: A sales revenue and cash flow analysis using real price data from a wind farm to sizing the economic impact of negative electricity prices on wind farm is depicted in Chapter 2. Chapter 3 describes the proposed approach to capture the nature of the negative price drivers through its influence in the negative prices' behavior and briefly overviews electricity price forecasting. Chapter 4 shows the validation study on the nature of drivers of NEP. Chapter 5 presents the economic value of the appropriate representation on the nature of negative electricity price drivers for a wind farm. Finally, the document culminates with the conclusions and future research.

Chapter 2: Negative Prices and Its Impact on Sales Revenue of WPPs

As the occurrence of Negative Electricity Prices (NEPs) will increase with the integration of more wind power generation, it is imperative to evaluate their current impact on WPPs' sales revenue to further back up economic sound alternatives to the current economic and operational strategies of most of the wind farms. To this end, different studies were performed. First, the impact of applying curtailment at negative prices was studied under three different scenarios and using real time market data. The first scenario represents the perfect forecast of negative prices and therefore, the maximum positive impact of applying curtailment at NEPs. The second scenario represents the influence of forecasting errors on the sales revenue when curtailment at NEPs is applied according to it; and the third and final scenario considers forecasting errors and minimum output level as representations of forecasting and physical limitations. As all scenarios suggest that it should be quite beneficial to implement curtailment at negative prices, the idea was further included in a cash flow analysis of a wind farm considering that it either opts for PTC and do not apply curtailment at negative prices or opts for ITC and apply curtailment at negative prices. The results show that an appropriate forecast of occurrence of negative price and coordinated control of wind generation may improve the sales revenue, making ITC more profitable than PTC; and highlight the necessity of a forecast of NEPs.

Section 2.1. Negative Prices and Its Impact on Sales Revenue: Curtailment at Negative Prices

The value of the impact of negative prices on sales revenue directly demonstrates the business opportunities for wind farms to apply curtailment at NEPs. The impact is quantified though the computation of the sale revenues with curtailment at NEPs and comparing that to the no curtailment strategy that most wind farms use nowadays as baseline. Three scenarios which represent curtailment at NEPs with perfect forecast, forecast limitations and forecast and physical limitations were studied. The economic benefits of applying curtailment at negative prices are long tied to frequency of occurrence and magnitude of NEPs. Likewise, the occurrence and magnitude are related to the specific characteristics of every market like generation mix, physical and geographical constraints, percentage of wind power integration and even less known percentage of wind farms who actively participate in the DAM and RTM markets. Using real data from a wind farm located in Northwest Oklahoma, it is aimed to at least partially portrait its impact for wind farms who participate in the Southwest Power Pool (SPP) market and similar markets.

Section 2.1.1. Study Framework

The study quantifies the impact of negative electricity prices on sales revenue though the computation of the sale revenues of a wind farm who participates in SPP spot market. Moreover, since it is desired to focus on the impact if negative electricity prices on sales revenue, only sales on the real time spot market are considered and the impact of the type of subsidy on the operation strategy is not considered. Additionally, different minimum down times and a constant generation of 1 MWh at any time when curtailment is not being applied are used to illustrated the possible impact. The sales revenue can be defined as:

$$Sales\ revenue = \sum_t^n price_t * curtialment * 1[MWh] \quad (2\text{-}1)$$

Where:

n: number of 5 minute prices in the period of time considered

$price_t$: electricity price at time t $[\frac{\$}{\mathrm{MWh}}]$

Curtailment: represent a binary signal to either generate or curtail

Finally, the illustration is down in a top to bottom approach with three scenarios of increasing realistic limitations to the improvement on sales revenue due to curtailment at NEPs. The first scenario is applying curtailment at NEPs in compliance to a perfect forecast and subsequently calculate the sales revenue. The second scenario is simple forecast where two different types of simple forecast algorithms are used to apply curtailment at NEPs accordingly; and subsequently calculate the sales revenue. Last of all, the third scenario considers the simple forecast algorithms and a minimum output of 5% of the rated generation to apply curtailment at NEPs and subsequently calculate the sales revenue. The results are shown as a percentage employing the no curtailment strategy that most wind farms use nowadays as baseline

Section 2.1.2. Data Description

The study was carried out using SPP 5-minute real time electricity prices for a wind farm located in Northwest Oklahoma and for the months of March, July, October and December of 2014. Due to limitations in data availability, each month used is intended to represented a different season, spring, summer, fall and winter, respectively, as it is known that wind and electrical consumption vary through seasons.

SPP is a power grid operator with members in 14 states: Arkansas, Iowa, Kansas, Louisiana, Minnesota, Missouri, Montana, Nebraska, New Mexico, North Dakota, Oklahoma, South Dakota, Texas and Wyoming. In 2014, it had a serving territory of 370,000 square miles 48,930 miles of transmission lines 4,103 substations 627 generating plants, a generating capacity of 77,366 megawatts and the following generation mix.

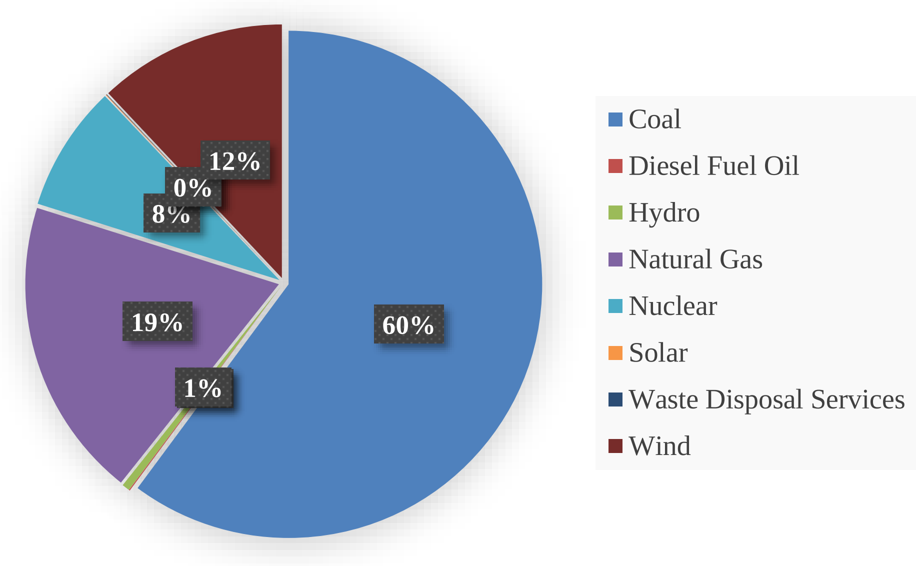

Figure 1. 2014 SPP Generation Mix (Capacity) [14]

As in 2015, the percentage of wind energy capacity increased to 14.86%. The amount of renewable generation, particularly those from wind in Northwest Oklahoma, will continue to increase rapidly. As the result of such a fast increase of wind generation, issues associated with slower growth of transmission capacity will become more serious. Consequently, transmission congestions, perhaps as well as negative prices at price-node of wind generation in Northwest Oklahoma, will be more frequently observed [15].

Currently SPP wholesale electricity market is a Locational Marginal Price (LMP) based integrated market of both energy and operating reserves using optimization techniques. This allows the system operator to allocate some cost of capacity (reserves) to the LMP prices, in both real-time market and day-ahead market. Wind generation is a dispatchable variable energy resources but it doesn't have to provide generation schedules even they are allowed to participate in the day-ahead market with benefit of credit and risk of charges. SPP performs the forecasts of wind and wind generation, and applies the forecast results in operations in both day-ahead market and real-time market.

Section 2.1.3. Perfect Forecast

In this scenario, it is assumed that the negative price can be perfectly forecasted, and wind generation can be reduced to zero promptly in response to negative price. Different minimum down times are considered related to different wind farms technologies 5, 10 and 15 minutes. The sales revenue applying curtailment at negative prices following a perfect forecast of NEPs is defined for every down time as following and the impact on sales revenue is shown in **Table 1** below.

For a 5 minute down time, the sales revenue is defined as:

$$Sales\ revenue = \sum_t^n Max(price_t * 1[MWh], 0) \tag{2-2}$$

Where:

n: number of 5 minute prices in the period of time considered

$price_t$: electricity price at time t $[\frac{\$}{\text{MWh}}]$

This is equivalent to applying curtailment every time $price_i$ is negative.

For a 10 minute down time, the sales revenue is defined as:

$$Sales\ revenue = \sum_{t}^{n} price_t * curtialment * 1[MWh]$$

Where:

n: number of 5 minute prices in the period of time considered

$price_t$: electricity price at time t $[\frac{\$}{\text{MWh}}]$

Curtailment: $(sign(price_t) = -\ OR\ sign(price_{t-1}) = -\ , then\ curtailment = 0)$

$$(sign(price_t) = +\ AND\ sign(price_{t-1}) = +\ , then\ curtailment = 1)$$

For a 15 minute down time, the sales revenue is defined as:

$$Sales\ revenue = \sum_{t}^{n} price_t * curtialment * 1[MWh]$$

Where:

n: number of 5 minute prices in the period of time considered

$price_t$: electricity price at time t $[\frac{\$}{\text{MWh}}]$

Curtailment:

$(sign(price_t) = -\ OR\ sign(price_{t-1}) = -\ OR\ sign(price_{t-2}) = -\ ,$

$then\ curtailment = 0)$

$(sign(price_t) = +\ AND\ sign(price_{t-1}) = +\ AND\ sign(price_{t-2}) = +,$

$then\ curtailment = 1)$

($, %)	March, 2014	$ 223,263	Improvement ($)	Improvement (%)
Revenue with Forecast and Generation Control	5 Minute Down Time	$ 245,361	22,098.16	9.90%
	10 Minute Down Time	$ 243,217	19,954.78	8.94%
	15 Minute Down Time	$ 242,159	18,896.58	8.46%
($, %)	July, 2014	$ 224,798	Improvement ($)	Improvement (%)
Revenue with Forecast and Generation Control	5 Minute Down Time	$ 245,246	20,447.22	9.10%
	10 Minute Down Time	$ 244,633	19,834.72	8.82%
	15 Minute Down Time	$ 244,181	19,382.24	8.62%
($, %)	October, 2014	$ 212,277	Improvement ($)	Improvement (%)
Revenue with Forecast and Generation Control	5 Minute Down Time	$ 233,706	21,428.45	10.09%
	10 Minute Down Time	$ 232,733	20,455.40	9.64%
	15 Minute Down Time	$ 231,713	19,435.65	9.16%
($, %)	December, 2014	$ 196,780	Improvement ($)	Improvement (%)
Revenue with Forecast and Generation Control	5 Minute Down Time	$ 213,061	16,280.75	8.27%
	10 Minute Down Time	$ 212,667	15,887.42	8.07%
	15 Minute Down Time	$ 211,937	15,156.77	7.70%

Table 1. Revenue Improvement Based on Perfect Forecast of Negative Prices

This ideal scenario can be used to calculate the maximum improvement of sales revenue, which serves as the benchmark upper bound for other studies when more constraints and limitations are taken into consideration. The results of this preliminary study suggest that there is quite an economical benefit in implementing an appropriate negative price forecasting algorithm, and active wind generation control strategy for wind farm operation. The percentage of improvement is similar in all cases with an average of 9.34%, 8.87% and 8.49% for a 5-minute start, 10-minute start and 15-minute start, respectively. The difference between increment in sales revenue considering the different minimum down times is less than 1% in all cases. Likewise, the percentage of sales revenue increment among different months is not significant with biggest difference months being 1.8%

As expected, the month of March and October present the highest percentage of increment in sales revenue from all months. March and October represent the season of Spring and Autumn, respectively, those months characteristically have a low to medium demand with a high to medium wind speed. Then, this translates in more high to medium wind power generation at medium to low demand, which is as discussed in the previous chapter the most frequent scenario where NEPs occur. The month of July which represent Summer is characterized as a high demand and low wind season, however their percentage of sales revenue improvement is the third highest. This season also has higher volatility prices as in hot summer days it is difficult to meet the demand and market participant are sometimes forced to pay maximum allowable prices. But as one of the main characteristics of electricity prices is mean reversals, high positive prices are followed by low (even negative) prices. Last, December shows the least increment in sales revenue from all the months. December is a stable month of low-medium wind and demand; the stability of both factors seems to contribute to either fewer NEPs and/or lower magnitude of NEPs.

Section 2.1.4. Simple Forecast

In this scenario, A simple moving average algorithm to forecast the occurrence of negative price for the months of July and October, and a simple exponential algorithm to forecast the occurrence of negative price for the months of July and December is employed. Two different algorithms were used to seek some generality. Different minimum down times are considered related to different wind farms technologies. The moving average, exponential algorithm, sales revenue applying curtailment at negative prices following the methods mentioned before are defined for every down time as following and the impact on sales revenue is shown in **Table 2** below.

The moving average algorithm is defined as:

$$forecast\ price_t = \sum_{i=1}^{N=3} \frac{price_{t-i}}{N} \tag{2-3}$$

Where:

$price_t$: electricity price at time $\mathrm{t-i}\ [\frac{\$}{\mathrm{MWh}}]$

The simple exponential algorithm is defined as:

$$forecast\ price_t = price_{t-1} + \alpha\varepsilon_{t-1} \tag{2-4}$$

Where:

$price_{t-1}$: electricity price at time $\mathrm{t-1}[\frac{\$}{\mathrm{MWh}}]$

$\alpha = Smoothing\ factor. It\ was\ selected\ as\ 0.3$

$\varepsilon_{t-1} = forecast\ error. It\ is\ defined\ as\ price_{t-1} - forecast\ price_{t-1}$

For a 5 minute down time, the sales revenue is defined as:

$$Sales\ revenue = \sum_{t}^{n} price_t * curtialment * 1[MWh]$$

Where:

n: number of 5 minute prices in the period of time considered

$price_t$: electricity price at time t $[\frac{\$}{\mathrm{MWh}}]$

Curtailment: $(sign(forecast\ price_t) = -\ , then\ curtailment = 0)$

$(sign(forecast\ price_t) = +\ , then\ curtailment = 1)$

For a 10 minute down time, the sales revenue is defined as:

$$Sales\ revenue = \sum_{t}^{n} price_t * curtialment * 1[MWh]$$

Where:

n: number of 5 minute prices in the period of time considered

$price_t$: electricity price at time t $[\frac{\$}{\mathrm{MWh}}]$

Curtailment:

$(sign(forecast\ price_t) = -\ OR\ sign(forecast\ price_{t-1}) = -,$

$then\ curtailment = 0)$

$(ssign(forecast\ price_t) = +\ OR\ sign(forecast\ price_{t-1}) = +,$

$then\ curtailment = 1)$

For a 15 minute down time, the sales revenue is defined as:

$$Sales\ revenue = \sum_{t}^{n} price_t * curtialment * 1[MWh]$$

Where:

n: number of 5 minute prices in the period of time considered

$price_t$: electricity price at time t $[\frac{\$}{\text{MWh}}]$

Curtailment:

$(sign(forecast\ price_t) = -\ OR\ sign(forecast\ price_{t-1}) = -$

$OR\ sign(forecast\ price_{t-2}) = -, then\ curtailment = 0)$

$(sign(forecast\ price_t) = +\ OR\ sign(forecast\ price_{t-1}) = +$

$OR\ sign(forecast\ price_{t-2}) = +, then\ curtailment = 1)$

($, %)	March, 2014	$ 223,263	Improvement ($)	Improvement (%)
Revenue with Forecast and Generation Control	5 Minute Down Time	$ 236,026	12,763.38	5.72%
	10 Minute Down Time	$ 236,424	13,161.26	5.89%
	15 Minute Down Time	$ 237,076	13,813.43	6.19%
($, %)	July, 2014	$ 224,798	Improvement ($)	Improvement (%)
Revenue with Forecast and Generation Control	5 Minute Down Time	$ 238,050	13,251.38	5.89%
	10 Minute Down Time	$ 238,715	13,916.24	6.19%
	15 Minute Down Time	$ 238,807	14,008.63	6.23%
($, %)	October, 2014	$ 212,277	Improvement ($)	Improvement (%)
Revenue with Forecast and Generation Control	5 Minute Down Time	$ 225,337	13,060.20	6.15%
	10 Minute Down Time	$ 226,225	13,947.90	6.57%
	15 Minute Down Time	$ 226,828	14,550.29	6.85%
($, %)	December, 2014	$ 196,780	Improvement ($)	Improvement (%)
Revenue with Forecast and Generation Control	5 Minute Down Time	$ 206,127	9,347.31	4.75%
	10 Minute Down Time	$ 206,852	10,072.15	5.12%
	15 Minute Down Time	$ 207,340	10,560.05	5.37%

Table 2. Revenue Improvement Based on Basic Forecasts of Negative Prices

These simple algorithms represent a scenario that can be used to calculate the minimum improvement of sales revenue, which serves as the benchmark lower bound for other studies. The results of this preliminary study suggest that there is quite an economical benefit in implementing even a simple price forecasting algorithm to actively apply wind generation control. The percentage of improvement is similar in all cases with an average of 5.63%, 5.94% and 6.16% for a 5-minute start, 10-minute start and 15-minute start, respectively. In comparison with **Table 1**, on average percentages of improvement are 3.71, 2.92 and 2.33 lower.

Section 2.1.5. Simple Forecast with Minimum Output Restriction

In this study, it is employed a simple moving average algorithm to forecast the occurrence of negative price for the months of July and October, and an exponential algorithm to forecast the occurrence of negative price for the months of July and December. In addition, it is taken into consideration the minimum output restriction on wind generation control: if negative price is forecasted, the output of wind generation will be reduced to 5% of its existing output level. Also, different minimum down times are considered related to different wind farms technologies. The sales revenue applying curtailment at negative prices following the methods mentioned before and considering a minimum output restriction is defined for every down time as following and the impact on sales revenue is shown in **Table 3** below.

For a 5 minute down time, the sales revenue is defined as:

$$Sales\ revenue = \sum_{t}^{n} price_t * curtialment * 1[MWh]$$

Where:

n: number of 5 minute prices in the period of time considered

$price_t$: electricity price at time t $[\frac{\$}{\text{MWh}}]$

Curtailment: $(sign(forecast\ price_t) = -\ , then\ curtailment = 0.05)$

$(sign(forecast\ price_t) = +\ , then\ curtailment = 1)$

For a 10 minute down time, the sales revenue is defined as:

$$Sales\ revenue = \sum_{t}^{n} price_t * curtialment * 1[MWh]$$

Where:

n: number of 5 minute prices in the period of time considered

$price_t$: electricity price at time t $[\frac{\$}{\text{MWh}}]$

Curtailment:

$(sign(forecast\ price_t) = -\ OR\ sign(forecast\ price_{t-1}) = -,$

$then\ curtailment = 0.05)$

$(ssign(forecast\ price_t) = +\ OR\ sign(forecast\ price_{t-1}) = +,$

$then\ curtailment = 1)$

For a 15 minute down time, the sales revenue is defined as:

$$Sales\ revenue = \sum_{t}^{n} price_t * curtialment * 1[MWh]$$

Where:

n: number of 5 minute prices in the period of time considered

$price_t$: electricity price at time t $[\frac{\$}{\text{MWh}}]$

Curtailment:

$(sign(forecast\ price_t) = -\ OR\ sign(forecast\ price_{t-1}) = -$
$OR\ sign(forecast\ price_{t-2}) = -, then\ curtailment = 0.05)$

$(sign(forecast\ price_t) = +\ OR\ sign(forecast\ price_{t-1}) = +$
$OR\ sign(forecast\ price_{t-2}) = +, then\ curtailment = 1)$

This scenario represents a more realistic scenario where forecast and a major physical limitation are considered. It can be used to calculate the minimum improvement of sales revenue, which serves as the benchmark lower bound for other studies under the same conditions. The results of this preliminary study suggest that there is quite an economical benefit in implementing even a simple price forecasting algorithm to actively apply wind generation control, and the minimum output restriction has a modest impact on the potential economic benefits. The percentage of improvement is similar in all cases with an average of 5.35%, 5.65% and 5.85% for a 5-minute start, 10-minute start and 15-minute start, respectively. Also in comparison with **Table 2**, on average percentages of improvement are 0.3 lower due to the minimum output restrictions.

($, %)	March, 2014	$ 223,263	Improvement ($)	Improvement (%)
Revenue with Forecast and Generation Control	5 Minute Down Time	$ 235,388	12,125.21	5.43%
	10 Minute Down Time	$ 235,766	12,503.20	5.60%
	15 Minute Down Time	$ 236,385	13,122.76	5.88%

($, %)	July, 2014	$ 224,798	Improvement ($)	Improvement (%)
Revenue with Forecast and Generation Control	5 Minute Down Time	$ 237,387	12,588.81	5.60%
	10 Minute Down Time	$ 238,019	13,220.43	5.88%
	15 Minute Down Time	$ 238,107	13,308.20	5.92%

($, %)	October, 2014	$ 212,277	Improvement ($)	Improvement (%)
Revenue with Forecast and Generation Control	5 Minute Down Time	$ 224,684	12,407.19	5.84%
	10 Minute Down Time	$ 225,528	13,250.50	6.24%
	15 Minute Down Time	$ 226,100	13,822.78	6.51%

($, %)	December, 2014	$ 196,780	Improvement ($)	Improvement (%)
Revenue with Forecast and Generation Control	5 Minute Down Time	$ 205,660	8,879.95	4.51%
	10 Minute Down Time	$ 206,348	9,568.54	4.86%
	15 Minute Down Time	$ 206,812	10,032.05	5.10%

Table 3. Revenue Improvement Based on Basic Forecasts of Negative Prices with Minimum Output Restriction

Section 2.1.6. Section Summary

Based on the results of the empirical study carried out on the negative price and its impact on sales revenue of wind generation, the necessity of developing an appropriate negative price forecast algorithm and predictive wind farm control can be summarized as follows. It is evident that an appropriate forecasting algorithm and predictive wind farm control will improve the sales revenue of wind generation. The results shown in Subsection 2.1.5 suggest that, even with a very simple forecasting algorithm, the improvement of sales revenue ranges from 5.35% to 5.85% according the historical record and depending on the minimum down time. These improvements appeared to be quite reliable ones since a number of major practical limitations have been applied in the study. Moreover, Subsection 2.1.3 suggest that the improvement of sales revenue could be as high as 8.49% to 9.34%, according the historical record and depending on the minimum down time. Finally, while there are some differences in revenues between the different seasons, they are considered small, contributing to conclude that there is a potential all year long increment in sales revenue.

Section 2.2. Illustration of The Necessity of Predictive Wind Farm Control

In the previous section, it was evidenced that an appropriate forecasting algorithm and predictive wind farm control will improve the sales revenue of wind farms. However, this finding needs to be further study in the context of the subsidy driven wind power industry. To this end, a cash flow analysis considering government subsidy, annual operating cost, and sales revenue of a hypothetical wind farm project is used to evaluate the impact of wind farm control at negative electricity prices under

PTC and ITC like tax credits. To simplify the analysis, we ignore the time value of money (interest rate) as it is not a significant factor in this qualitative analysis.

Section 2.2.1. Example Description

A wind farm with 250 MW installed capacity, initial investment of \$375 million (1,500 \$/kWh), 100 MW output capacity (capacity factor of 0.4) and annual operating cost such as amortization and maintenance costs estimated at \$2 million is considered. There are two major options of taking government subsidies for wind farm investor: option 1) taking an initial tax credit of one third of the initial investment up front, and option 2) taking the same amount of subsidy as that of option 1, but in the form of future production tax credit, assuming the production tax credit of ≈\$12.8 \$/MWh for the first 10 years of generation and as long that it generates according to the its wind capacity factor.

Note that as it was mentioned in Section 1.2.1, PTC (option 2) is usually more profitable than ITC (option 1), in [13] it was found that for the same \$/kWh of installed capacity and factor capacity, PTC is almost 15% more profitable than ITC considering other tax related particularities of each tax credit and the change of PTC through the years. Incorporating this to the cash flow analysis, would demand a more realistic evaluation of the power generation of the wind farm. Therefore, for simplicity, the options are considered equally profitable and any improvement of applying predictive control under option 1 must be substantial to consider the predictive control economically beneficial. Moreover, all these numbers were text book chosen for illustration purposes, but the author truly believes that they can be replaced with any other more realistic numbers and the conclusion will be similar.

Section 2.2.2. Results and Analysis

The cash flow analysis is computed on yearly basis and for the first 15 years of operation. Based on the real price data from the wind farm in Northwest Oklahoma, it was found that around 10% of all prices for 2014 were negative. In the cash flow analysis, this is equivalent to one year long of negative electricity prices for ten years of operation. To compensate, other electricity prices are little bit high in comparison with market average, since they are based on the average on the real data without considering NEPs.

First, the cash flows of two subsidies assuming there is no predictive wind farm control are found. Intentionally, the total cash flows of the two options were made exactly the same. This implies that the investor should be indifference between two options. Moreover, Table 4 and Table 5 illustrate that without predictive wind farm control, this is true whenever electricity prices are positive or negative.

The necessity of having predictive wind farm control is shown in Table 6. With predictive wind farm control, the investor who takes option 1 is able to respond to the market price signal by stopping production in presence of negative price, therefore, the investors taking option 1 will benefit from the sales revenue and increased total return by avoiding the impact of negative price. In this example, as long as the negative price is in a moderate range, the investor who takes option 2 would like to keep producing since the government subsidy depends on the production.

					Option 1 (ITC)			Option 2 (PTC)		
Year	Amortization [$]	Output [MW]	Generation [MWh]	Market price [$/kWh]	Sales [$]	Subsidy [$]	Cash Flow [$]	Sales [$]	Subsidy [$]	Cash Flow [$]
1	(2,000,000)	100	876,000	31.60	27,685,947	112,500,000	138,185,947	27,685,947	11,250,000	36,935,947
2	(2,000,000)	100	876,000	30.35	26,582,240		24,582,240	26,582,240	11,250,000	35,832,240
3	(2,000,000)	100	876,000	29.50	25,837,744		23,837,744	25,837,744	11,250,000	35,087,744
4	(2,000,000)	100	876,000	26.54	23,245,461		21,245,461	23,245,461	11,250,000	32,495,461
5	(2,000,000)	100	876,000	30.00	26,280,000		24,280,000	26,280,000	11,250,000	35,530,000
6	(2,000,000)	100	876,000	29.50	25,837,848		23,837,848	25,837,848	11,250,000	35,087,848
7	(2,000,000)	100	876,000	30.00	26,280,000		24,280,000	26,280,000	11,250,000	35,530,000
8	(2,000,000)	100	876,000	27.00	23,652,000		21,652,000	23,652,000	11,250,000	32,902,000
9	(2,000,000)	100	876,000	30.70	26,893,200		24,893,200	26,893,200	11,250,000	36,143,200
10	(2,000,000)	100	876,000	30.50	26,718,000		24,718,000	26,718,000	11,250,000	35,968,000
11	(2,000,000)	100	876,000	29.70	26,017,200		24,017,200	26,017,200	-	24,017,200
12	(2,000,000)	100	876,000	31.00	27,156,000		25,156,000	27,156,000	-	25,156,000
13	(2,000,000)	100	876,000	26.70	23,389,200		21,389,200	23,389,200	-	21,389,200
14	(2,000,000)	100	876,000	30.80	26,980,800		24,980,800	26,980,800	-	24,980,800
15	(2,000,000)	100	876,000	29.90	26,192,400		24,192,400	26,192,400	-	24,192,400
						Total	**471,248,041**		**Total**	**471,248,041**

Table 4. Comparison of Cash Flow (Without Predictive Wind Farm Control and Negative Prices)

					Option 1 (ITC)			Option 2 (PTC)		
Year	Amortization [$]	Output [MW]	Generation [MWh]	Market price [$/kWh]	Sales [$]	Subsidy [$]	Cash Flow [$]	Sales [$]	Subsidy [$]	Cash Flow [$]
1	(2,000,000)	100	876,000	31.60	27,685,947	112,500,000	138,185,947	27,685,947	11,250,000	36,935,947
2	(2,000,000)	100	876,000	30.35	26,582,240		24,582,240	26,582,240	11,250,000	35,832,240
3	(2,000,000)	100	876,000	29.50	25,837,744		23,837,744	25,837,744	11,250,000	35,087,744
4	(2,000,000)	100	876,000	26.54	23,245,461		21,245,461	23,245,461	11,250,000	32,495,461
5	(2,000,000)	100	876,000	(10.50)	(9,198,000)		(11,198,000)	(9,198,000)	11,250,000	52,000
6	(2,000,000)	100	876,000	29.50	25,837,848		23,837,848	25,837,848	11,250,000	35,087,848
7	(2,000,000)	100	876,000	30.00	26,280,000		24,280,000	26,280,000	11,250,000	35,530,000
8	(2,000,000)	100	876,000	27.00	23,652,000		21,652,000	23,652,000	11,250,000	32,902,000
9	(2,000,000)	100	876,000	30.70	26,893,200		24,893,200	26,893,200	11,250,000	36,143,200
10	(2,000,000)	100	876,000	30.50	26,718,000		24,718,000	26,718,000	11,250,000	35,968,000
11	(2,000,000)	100	876,000	29.70	26,017,200		24,017,200	26,017,200	-	24,017,200
12	(2,000,000)	100	876,000	31.00	27,156,000		25,156,000	27,156,000	-	25,156,000
13	(2,000,000)	100	876,000	26.70	23,389,200		21,389,200	23,389,200	-	21,389,200
14	(2,000,000)	100	876,000	30.80	26,980,800		24,980,800	26,980,800	-	24,980,800
15	(2,000,000)	100	876,000	29.90	26,192,400		24,192,400	26,192,400	-	24,192,400
						Total	**435,770,041**		**Total**	**435,770,041**

Table 5. Comparison of Cash Flow (Without Predictive Wind Farm Control and with Negative Price)

					Option 1 (ITC)			Option 2 (PTC)		
Year	Amortization [$]	Output [MW]	Generation [MWh]	Market price [$/kWh]	Sales [$]	Subsidy [$]	Cash Flow [$]	Sales [$]	Subsidy [$]	Cash Flow [$]
1	(2,000,000)	100	876,000	31.60	27,685,947	112,500,000	138,185,947	27,685,947	11,250,000	36,935,947
2	(2,000,000)	100	876,000	30.35	26,582,240		24,582,240	26,582,240	11,250,000	35,832,240
3	(2,000,000)	100	876,000	29.50	25,837,744		23,837,744	25,837,744	11,250,000	35,087,744
4	(2,000,000)	100	876,000	26.54	23,245,461		21,245,461	23,245,461	11,250,000	32,495,461
5	(2,000,000)	100	876,000	(10.50)	-		-	(9,198,000)	11,250,000	52,000
6	(2,000,000)	100	876,000	29.50	25,837,848		23,837,848	25,837,848	11,250,000	35,087,848
7	(2,000,000)	100	876,000	30.00	26,280,000		24,280,000	26,280,000	11,250,000	35,530,000
8	(2,000,000)	100	876,000	27.00	23,652,000		21,652,000	23,652,000	11,250,000	32,902,000
9	(2,000,000)	100	876,000	30.70	26,893,200		24,893,200	26,893,200	11,250,000	36,143,200
10	(2,000,000)	100	876,000	30.50	26,718,000		24,718,000	26,718,000	11,250,000	35,968,000
11	(2,000,000)	100	876,000	29.70	26,017,200		24,017,200	26,017,200	-	24,017,200
12	(2,000,000)	100	876,000	31.00	27,156,000		25,156,000	27,156,000	-	25,156,000
13	(2,000,000)	100	876,000	26.70	23,389,200		21,389,200	23,389,200	-	21,389,200
14	(2,000,000)	100	876,000	30.80	26,980,800		24,980,800	26,980,800	-	24,980,800
15	(2,000,000)	100	876,000	29.90	26,192,400		24,192,400	26,192,400	-	24,192,400
						Total	**446,968,041**		**Total**	**435,770,041**

Table 6. Comparison of Cash Flow (With Predictive Wind Farm Control and Negative Price)

With this simple illustration, it was shown how the appropriate forecasting algorithm and predictive wind farm control could improve the sales revenue of a wind farm that is subsidized with ITC. This clearly is a business opportunity for a wind farm operator who wants to capitalize on the negative bidding strategy of wind farms subsidized with PTC and other conventional base load generation with high shut down cost. In a more realistic situation, the investor who takes the option 2 is taking the risks associated with market volatility and variation in wind generation. Thus, it is important for option 1 investor to avoid the impact of negative price to secure the government subsidy as it is already a risk-discounted value.

Section 2.3. Chapter Summary: Significance of Predictive Wind Farm Control

Through a series of studies, it was demonstrated the necessity of developing an appropriate negative price forecast algorithm and predictive wind farm control. An appropriate negative price forecast is a forecast that takes into consideration the special characteristics of negative electricity prices. The predictive wind farm control would not only significant increase sales revenues but also could benefit wind farms subsidized with ITC, as this economic sound alternative brings economic benefits similar to be subsidized with PTC without the risks associated with market volatility and variation in wind generation. Indirectly the system´s stability would also benefit from this alternative, as wind farms would voluntary apply curtailment at negative prices. Subsequently, this could decrease the occurrence of NEPs and smooth the transition to a new power system better design to deal with renewable energy and its particular characteristics. Hence, it is on the best interest of market participant generators and operators to develop an appropriate negative price forecast algorithm and predictive wind farm control.

Chapter 3: Discrete Model for Negative Prices Identification

In this chapter, a discrete approach to forecast negative electricity prices is proposed. The main aim of the proposed approach is to capture the nature of the negative price drivers through its influence in the negative prices' behavior. To this end, first an overview of electricity prices forecasting is done to understand the importance of the objective and time horizon in the selection of a model or technique to forecast negative prices. Current work in explicitly forecasting negative electricity prices is revised, along with the forecasting of jumps in electricity prices using Markov Regime Switching (MRS) model. Though since they do not particularly focus on NEPs, the author of this thesis does not consider them completely adequate to examine the thesis motivation, they set the bases to propose an approach that is. The proposed approach is divided in two major steps; first the price data is separated using mean shift or kernel density estimation (since it is one-dimensional data), this first step lightly resembles the purpose of using MRS for jumps in prices. And then, intervals with high concentration of negative electricity prices are modeled as second order Markov processes.

Section 3.1. Overview of Electricity Price Forecasting

Section 3.1.1. Electricity Price Forecasting: Methods, Time Horizon and Purpose

There are three main factors that influence the selection of a forecast model: time horizon, purpose and characteristics of the electricity price that want to be captured using the model. The motivation of the forecast used in this thesis is to assist wind farms in their day to day operation by empower them with a negative price forecasting, that can be used to apply predictive control accordingly. This motivation or purpose aligns better with a short-term time horizon. There is not a common agreement upon the length of the time horizons, but short term forecasting is usually understood as minutes to days ahead.

Additionally, electricity prices are characterized by seasonality (daily, weekly and per season), high volatility, volatility clustering, mean reversion and spikes or jumps. All these characteristics must be account for in any forecasting model to accurately describe electricity prices. However, most of the forecasting algorithms do not consider infrequent events like jumps and focus on the mean reversion characteristic and others, since at least for power plants those characteristics has the biggest impact on their sales revenues. Electricity prices volatility is associated with the volatility transmitted by electricity price drivers like load, generation capacity, fuel prices, generation mix, weather, bidding strategies, between others; and the volatility is exacerbated in the RTM due to its balancing nature of unforeseen events.

Recapitulating, a short term forecast for the day to day operation of a wind farm who do not focus on special characteristics of electricity prices like jumps but do acknowledge the high volatility of the electricity prices is sought. There are at least three types of models who comply with the requirements: reduced form methods like time series methods where mainly only the prices are considered; statistical, where statistical or econometric approaches are used and machine learning methods or computational intelligence.

Time series models are the most popular methods to forecast short term electricity prices based only on historical data. As a result, they are chosen to validate the discrete approach proposed used in this thesis. The major assumption to model a time series using an autoregressive model is that the time series is stationary or weak stationary. A time series is stationary if the mean of the time series is constant, and covariance and autocorrelation only depends on the time difference between the data points in the time series, that is the variations around the mean have a constant amplitude and magnitude. Unfortunately, many time series including electricity prices are not weak stationary. As a solution, Box and Jenkin proposed the Autoregressive Integrated Moving Average (ARIMA) model, where the first step is to transform the non-stationary time series to a weak stationary time series by differentiating the time series as many times as it is necessary to remove the non-stationary characteristics. Then, an ARMA model is fitted to the weak stationary time series [16], [17]. Consequently, ARIMA model is selected as it is a good balance between solid estimation of time-varying trends and relatively small number of parameters i.e., it does a great job with serially correlated data without being overly complex.

ARIMA model is defined by the three following parameters:

p: The order of the autoregressive model AR(p)

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t \tag{3- 1}$$

Where:

c is a constant

ε_t is white noise

φ_i is the weight associated to the prior value X_{t-i}

d: Number of times, the data is differentiated to reach weak stationary status

if d=1

$$X_t = x_t - x_{t-1}$$

If d=2

$$X_t = (x_t - x_{t-1}) - (x_{t-1} - x_{t-2})$$

q: The order of the moving average model MA(q)

$$X_t = \mu + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \tag{3- 2}$$

Where:

μ is the mean of the series

θ_i is the weight associated to the error ε_{t-i}

ε_{t-i} is the error between the forecasted value X_{t-i} and the actual value x_{t-i}.

Then the ARIMA model is denoted ARIMA(p,d,q) and is defined as [18]:

$$X_t = c + \varepsilon_t + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \tag{3- 3}$$

Section 3.1.2. Forecasting of Negative Electricity Prices

While there has not been an extensive research on forecasting negative electricity prices, a number of papers were found where negative prices were especially considered while developing the price forecasting algorithm. In [17], it is developed a time invariant three regimes switching algorithm using deseasonilized log prices (negative prices are replaced 0.01 €/MWh to apply this transformation). The three regimes are the following: base regime, upward jump regime and downward jump regime. The latter is defined as any value 3 standard deviation lower than the mean of the deseasonilized log prices. Within this regime, for each price a number is randomly picked from a uniform distribution and it is compared to adjusted relative frequency of negative prices within the series of downward jumps. If the random value is lower than the adjusted relative frequency, the downward jump is replaced by a negative value. The negative values are picked from a lognormal distribution or an exponential distribution based on a similar process where a random value is picked from a uniform distribution and compared to the ratio between the number of negative prices greater than −80€/MWh and the total number of negative prices (the distributions used and the selection of −80€/MWh to divide the negative prices is based on historical data). It was reported that the negative price model improved the result of the overall electric price model, however, it only can be used to generated single negative prices and the need for an autoregressive approach is recognized to captured autocorrelation and lag of negative prices.

Other approaches were focused on creating alternatives, so that widely used price forecasting methods could be applied to modeling positive and negative electricity prices. [19] discusses the possibility of shifting all prices so that they can be treated as positive and the log transformation (that is commonly used as a first step to create a time series based price forecast model) can be employed, since negative prices are rare. Unfortunately, some price characteristics are compromised during this transformation. [20]proposed a sine hyperbolic transformation with an offset and scale parameters instead of the historical used log transformation. This transformation imitates the log behavior but allows negative prices without adding complexity to the price modeling.

Section 3.1.3. Forecasting of Jumps in Electricity Prices Using MRS

Another feature of the electricity prices that is usually overlooked due to their infrequency are the price spikes or price jumps. Price jumps are the result of transmission congestions, contingencies (in generation and/or in the transmission network), market working close to capacity constraints; and other unexpected events. They are characterized by short duration; positive jumps are commonly follow by negative jumps (jump reversals) and do not have a long-term impact on electricity price levels.

Markov Regime-switching is technique that captures well the short duration of price jumps, allows multiple consecutive spikes and naturally separates what many researchers consider two distinct trends or underlying stochastic processes within the price time series. Markov Regime-Switching usually is consisted of 2 or 3 regimes: base regime, jump regime and drop regime. The base regime is where the prices spend most of the time in, and sporadically jump to the jump and drop regimes spend some time in one or between them, and the return to the base regime. Different solutions have been proposed to determine the probability of transition between levels, time invariant transition probabilities based on historical data and its standard deviation is the most common [21]–[23], but there is a growing trend to model the transition probability as time variant and include external parameters to estimate these probabilities like demand and supply [24], [25]. Remarkably, spike identification schemes perform better when prices are used instead of log prices.

Base and jump regimes are commonly modeled by the Ornstein-Uhlenbeck process while most of the proposed models for the jump regime are Ornstein-Uhlenbeck processes and heavy-tailed random variables including external parameters like outages, demand and supply (some of those may forecast negative prices [26]) . Using the same random noise process helps to keep the mean reversion characteristic of electricity prices while using independent regimes allows more flexibility to model the jump regime [27], [28].

Section 3.1.4. Section Summary

In this section, forecast and modeling methods for short time forecasting and operational assistance purpose were explored. Although, none of the methods align perfectly with the purpose of studying the nature of negative price drivers, they were useful to understand how unconventional electricity price features like jumps are negative prices are treated and modeled. Moreover, the importance of the purpose of a forecasting model is recognized through the differences between whether or not negative electricity prices or jump in the electricity prices are considered; and the logic behind the selection of each model to emphasize specific features of the electricity prices. Also, the necessity of studying the negative electricity price drivers is emphasized as some authors mainly identified the necessity of including negative electricity prices in modeling electricity prices without considering if they are driven by different forces and how those affect their occurrence. In the next section, the proposed approach to study the nature negative price drivers is presented. It is influenced by the studied methods in the section, but fairly simple to reflect the nature of the drivers rather that the superiority of the approach.

Section 3.2. Proposed Approach

To study the nature of negative electricity price drivers, a fairly simple three steps approach is proposed. It starts by identifying the groups or intervals of prices with high concentration of negative prices. And then, those groups are modeled as a Second Order Markov Process (SOMP), a discrete model that slightly accounts for the information carried by previous data points. The separation of the electricity prices into groups or intervals is motivated by the MSR method used in the forecasting of jumps in electricity prices; and the recognition that at least for the motivation of this thesis, the study of negative electricity prices must be done separately from the rest of the electricity prices. Then, it was decided that electricity prices without any type of preprocessing will be used based on the indication that identification techniques perform better using prices instead of log prices and important characteristics like negative prices can be preserved in exchange of higher volatility. Additionally, to avoid seasonality, the method is meant to be apply on season basis i.e., model's parameters were found for every season. After the brief overview of the approach and some of the assumption for its implementation, the three steps can be defined as follow:

1. Data segmentation

Mean shift algorithm was applied to find "clusters" or intervals in the electricity price data. Since, the only input of the model are historical electricity price data, the mean shift procedure further simplifies into finding the Kernel Density Estimation (KDE) of the electricity prices and use its local minimums as interval borders and its local maximums as the mean of each interval.

While this algorithm gives the advantage of dividing data according to what it might be different market scenarios or conditions, it would also result in losing important information about the current state of the market; information that is transmitted by immediately previous prices in the time series. To compensate, for every electricity price, the two immediately previous prices were considered, if those were part of the same interval then the electricity price was also considered part of that interval. This would help to quantify the likelihood of negative or positive electricity prices given that the previous electricity prices were part of a certain interval or high density zone. Also, given the restrictions that arise with this condition, a new group or interval was created to account for consecutive prices that belong to different intervals. It would be denoted as "Mixed" interval o group. This group constitutes on average 12% of the training price data. Therefore, these interactions between prices of different intervals are a reduced percentage of the total number of prices, and most consecutive data points are part of the same cluster or interval.

2. Assessment of the distribution of negative prices within the intervals

For each interval, the conditional probability of occurrence of negative or positive prices given that the two previous prices were negative or positive is computed. Along, with the total number of data points in each interval, it helps to determine the intervals with the higher concentration of negative electricity prices. For the intervals with low concentration of negative electricity prices or low number of data points, this conditional probability is further used to forecast future prices; as either the prices in those intervals are not the main focus of the proposed approach; or the conditional probability overwhelmingly shows that future prices are highly likely to be negative if the previous prices were negative.

3. Negative electricity price modeling

The intervals with high concentration of negative electricity prices are modeled as a Second Order Markov Process (SOMP), i.e., it is assumed that the processes are independent from prior states. Then, their transition probability matrices are used to forecast future prices.

A detail explanation of the methods used in steps one and three, further reasoning behind their selection and assumptions are discussed in the next subsections.

Section 3.2.1. Segmentation of prices

For a continuous random variable, the probability density function (pdf) describes its relative likelihood to take on a given value. This important information helps to define and characterize a random variable, but it is often unknown. While it is possible to make assumptions about the pdf of a random variable, a widely popular and well-studied non parametric approach called Kernel Density Estimation does a great job estimating the pdf of a random variable based only on data samples and bandwidth *h* [29].

Supposing that random variable $\{x_i\}\ i = 1, \ldots, n, x_i \epsilon \mathbb{R}$ are i.i.d data drawn from an unknown density *f(x)*, *f(x)* can be estimated using Kernel Density Function defined as:

$$\boldsymbol{f}(\boldsymbol{x}) \approx \hat{\boldsymbol{f}}(\boldsymbol{x}) = \frac{\boldsymbol{1}}{\boldsymbol{nh}} \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{K}\left(\frac{\boldsymbol{x}-\boldsymbol{x_i}}{\boldsymbol{h}}\right) \quad (3\text{- }4)$$

Where:

Kernel function has the following properties:

a) non-negative
b) non-increasing: $K(a) \geq K(b)$ if $a<b$
c) piecewise continuous $\int_0^\infty K(r)dr < \infty$
d) symmetric
e) $\int_{-\infty}^{\infty} K(r)dr = 1$

Some of most commonly used Kernel functions include: Unit Flat Kernel, Epanechnikov Kernel and Gaussian Kernel. Since it is known that the Kernel function does not play a key role in the mean shift performance, the Gaussian Kernel was picked out based on its convenient mathematical properties, and it is defined as following:

$$\boldsymbol{K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}} \tag{3- 5}$$

Contrasting, the bandwidth h is the most important value of the equation (3-4), and it also can be seen as a smooth parameter. A large h will over smooth the data and therefore, important density information could be lost; while a small h would give too much importance to noise in the data. The goal is to pick a bandwidth that minimizes the error between the estimated density and the actual density, which it is difficult since the actual density is unknown. While this bias-variance tradeoff is still a debate topic, there is a “rule of thumb bandwidth estimator” when Gaussian kernel is employed and there is some indication that the actual density is Gaussian (This is the formula used to computed h by most statistical programs), according to this rule h can be estimated as [30]:

$$\boldsymbol{h = \left(\frac{4\hat{\sigma}^5}{3n}\right)^{\frac{1}{5}}} \tag{3- 6}$$

Where:

$\hat{\sigma}$ is the standard deviation

n is the number of data points

As electricity prices, have been historical known to be heavy-tailed, calculating the bandwidth using (3-6) would introduce some error. This is acceptable, since the study don't strive to perfectly model the negative electricity prices but rather capture the essential dynamics in them.

The mean shift algorithm finds the local maximums in the KDE by using ascending gradients for each value in the density until the "means" or "modes" that are the local maximums converge; and while doing so, it associates every value to a local maximum to create data clusters. In other words, the mean shift algorithm calculates a mean shift vector at every data point using the bandwidth h as the size of the radio sphere and subsequent influence of adjacent data points for this calculation, and the Gaussian Kernel as the weight that is given to each value (i.e., closer values are given more importance that far away values within the radio of influence; values outside the radio of influence are not included in this calculation) [31], [32]. However, since one-dimensional data (time was not directly taken into consideration) is only being considered in this study, there is no need to apply the mean shift algorithm to find local maximums and the clusters associated to them. A simple procedure to find local minimums and maximums for one-dimensional data would find the local maximums, which are the modes or means of each "cluster" or interval; and the local minimums, which are the borders between "clusters" or intervals.

Section 3.2.2. Modeling of Negative Electricity Prices

If it is assumed that the processes are independent from prior states or values the processes has taken, and considering as it was mentioned before that no direct connection between the occurrence of negative prices and time of the day was found. Then, negative electricity prices can be modelled as a second order time homogenous Markov Chain.

Let *X(t)* be a stochastic process, which can take any value from a discrete finite set *S=(1,2,…,k)*. Then *X(t)* is a second order Markov Chain (SOMC) if [33]

$$Pr\,\{X(t_n) = i_n \mid X(t_1 = i_1, \ldots, X(t_{n-1}) = i_{n-1}\} =$$

$$Pr\{X(t_n) = i_n \mid X(t_{n-1}) = i_{n-1}, X(t_{n-2}) = i_{n-2}\} \tag{3- 7}$$

Where:

$$t_1 < t_2 < \ldots < t_{n-1} < t_n$$

That is, the probability of transitioning between the finite set of values or states depends only on the two prior states. Moreover, if the probability does not depend of time, the Markov Chain is known as time homogenous Markov Chain and the SOMC can be parameterized by empirically estimating the transition probability matrix that describes the probability of transitioning between the different states. The transition probability *Pr{X(t)=k|X(t-1)=j,X(t-2)=i}=Pij,k* denotes the conditional probability of transitioning to state *k* given that current state is *j* and previous state was *i*. For *k* states, the transition probability matrix of a SOMC takes the form:

$$M = \begin{bmatrix} p_{1.1,1} & p_{1.1,2} & \cdots & p_{1.1,k} \\ p_{1.2,1} & p_{1.2,2} & \cdots & p_{1.2,k} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1.k,1} & p_{1.k,2} & \cdots & p_{1.k,k} \\ p_{2.1,1} & p_{2.1,2} & \cdots & p_{2.1,k} \\ p_{2.2,1} & p_{2.2,2} & \cdots & p_{2.2,k} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ p_{k.k,1} & p_{k.k,2} & \cdots & p_{k.k,k} \end{bmatrix} \tag{3- 8}$$

The transition probability matrix becomes the basis to predict the likelihood of new states. Every transition probability can take a value between 0 and 1, and the summation of any row of the transition probability matrix is 1 i.e., it is certain that the matrix will move from the current state to one state in *S* [34], [35].

For the Markov processes describe in this study, 4 by 2 probability transition matrices will illustrate the transition between states. The probabilities elements can be interpreted as following:

- $P_{11,1}$ is the probability of transition to a positive price given that the current and previous prices were positive, it also can be denoted as $P_{++,+}$
- $P_{11,2}$ is the probability of transition to a negative price given that the current and previous prices were positive, it also can be denoted as $P_{++,-}$
- $P_{21,1}$ is the probability of transition to a positive price given that the current price is positive and the previous price was negative, it also can be denoted as $P_{-+,+}$
- $P_{21,2}$ is the probability of transition to a negative price given that the current price is positive and the previous price was negative, it also can be denoted as $P_{-+,-}$
- $P_{12,1}$ is the probability of transition to a positive price given that the current price is negative and the previous price was positive, it also can be denoted as $P_{+-,+}$
- $P_{12,2}$ is the probability of transition to a negative price given that the current price is negative and the previous price was positive, it also can be denoted as $P_{+-,-}$
- $P_{22,1}$ is the probability of transition to a positive price given that the current and previous prices were negative, it also can be denoted as $P_{--,+}$
- $P_{22,2}$ is the probability of transition to a negative price given that the current and previous prices were negative, it also can be denoted as $P_{--,-}$

In the next chapter, the implementation and the results of the proposed approach using real price data are presented in a comparative study.

Chapter 4: Empirical Study on The Nature of Drivers of NEP

In this chapter the results of implementing the approach proposed in Chapter 3 are shown and discussed. First, the framework of the forecasting, and techniques used to show the significance of the forecasting results are discussed. Then, the real price data used in the thesis is further described. The implementation of the proposed approach is detailed step by step to highlight the potential useful insights the approach provides. Finally, the cross-validation results are compared against the results from the ARIMA models, which represent the classic continuous type of forecasting of electricity prices. The results of the study suggest that given the current market structure, the negative price drivers show indication of being of a discrete kind

Section 4.1. Proposed Forecasting Framework

First of all, for all methods described in the prior section electricity prices were used instead of log prices. Hence, negative prices can be taken into consideration (in expense of higher volatility), and a better estimation of the intervals limits can be expected. The forecast framework is the following:

1. The Kernel Density Estimation is found for the training sample and the local minimums are used as limits of the high-density intervals. The same limits are used for the testing data. Therefore, based on the current and previous values the conditional probability or transition probability matrix of a certain interval is used to forecast the future value.
2. The conditional probabilities for every interval are found. They are used along with the distribution of prices in each interval to determine which intervals have the

highest concentration of negative electricity prices. For intervals with very small or very high concentration of negative electricity prices, the same probabilities are used to predict future prices given that the current and previous prices belong to such intervals.

3. The transition probability matrices for the intervals with high concentration of negative electricity prices are computed for the training sample data. The same transition probability matrices are used to predict future prices given that the current and previous prices belong to such intervals. There is a special consideration when at least half of the prices in the previous hour were negative, in those cases, prices that show a negative slope (i.e. previous price value was greater than current price) were assigned a higher probability for future negative electricity prices.

Section 4.2. Data Description

Expanding the Description of Section 2.1.2, in this study the data is divided so that conventional validation (70% of the data is used for training and 30% for testing) and cross-validation (data was divided into two equally sized groups. The groups were not randomly sampled. In one case, first two weeks were set as the training set and the last two weeks as the testing set; and in the other, the first and the third week were set as the training set and the second and fourth weeks as the testing set) is used to quantify the generality of the results.

Negative prices account for 12.9% of electricity prices from the month of March with a range between -0.0001 to -223.787 $/MWh, 9.45% of electricity prices from the month of July with a range between -0.0215 to -250.303 \$/MWh, 11.12% of electricity prices from the month of October with a range between -0.0085 to -187.938 \$/MWh, and 10.04% electricity prices from the month of December with a range between -0.0001 to -87.737 \$/MWh. As it was expected fewer negative prices occurred in summer, however contrary to what had been reported; negative prices were not more likely to occur in winter. In general, the number of negative prices did not extremely fluctuate within seasons. Also, negative jumps or spikes are less extreme (see the lowest price for every month above) than positive jumps or spikes which highest values are 2051.32, 1121.246, 1140.25 and 1654.433 \$/MWh for the months of March, July, October and December, respectively. If the occurrence of downward jumps in electricity prices are a response of the system trying to revert to normal values after an upward jump, they seem to be less driven by the extreme conditions of the system than upward jumps.

Fig. 2 shows the complexity of forecasting negative prices under a unique underlying assumption for their distribution. Most of the months' exhibit two density peaks, one near zero and one more negative that varies in distance (from the first peak) and magnitude for each one of the months. Then, their KDEs advocate for similar day like forecast since unique features for each month are clearly displayed. Additionally, the modeling of the negative prices as lognormal distribution for values below certain limit and as exponential distribution for values near zero use in [17] appear appropriate as long as the parameters for the distributions are estimated from similar days.

Section 4.3. Results and Analysis

As it was mentioned on Section 4.2, multiple cross validation techniques were used. However, to display the characteristics captured by the different methods, the training data corresponding to the first two weeks of every month are displayed here; and the general forecasting results would display the weight average of the results (they are weighted since all methods do not hold the same number of data points for the training (testing) data sets).

Fig. 3 shows the probability density function of the training data for each month, they were estimated using Kernel density estimation and the identification of the intervals or "clusters" was made based on the local minimums. While no data points were filtered for the estimation, all graphs are cropped to the span where all high-density points are located. It is worth mentioning that all the high densities points shown in Fig. 2 are still visible and identifiable in the density estimation of the training price data. That is, the distribution of NEPs does not change significantly between the NEPs data for the whole month and the training data. This could have two explanations, nor most negative prices occurred within the first two weeks of every month or the first two weeks are a great representation of the distribution of negative prices for the month. Negative prices of the training electricity price data account for 43.65% of all negative electricity prices for the month of March, 65.88% of all negative electricity prices for the month of July, 69.69% of all negative electricity prices for the month of October, and 63.73% of all negative electricity prices for the month of December. While most of the negative prices occurred within the first two weeks for the month of October, it is fair to say that negative prices seem to be relatively homogenously distributed for the months of March, July and December.

Then, it can be inferred from this that the drivers behind negative prices could be linked to slow changing conditions like the weather pattern (it is directly related to the wind generation and load) for the month or season, with fluctuations related to unexpected changes like forecast errors or abnormal weather for the season.

To illustrate the clustering or interval identification using Kernel Density Estimation (i.e. mean shifting for one-dimensional data), in Fig. 3, for the first two weeks of March the following clusters or intervals were identified:

1. $-\infty <= X(t\text{-}2)$ *AND* $X(t\text{-}1) < -60.304$
2. $-60.304 <= X(t\text{-}2)$ *AND* $X(t\text{-}1) < -18.735$
3. $-18.735 <= X(t\text{-}2)$ *AND* $X(t\text{-}1) < 11.301$
4. $11.301 <= X(t\text{-}2)$ *AND* $X(t\text{-}1) < 66.97$
5. $66.97 <= X(t\text{-}2)$ *AND* $X(t\text{-}1) < =\infty$

And a sixth interval or “mixed” group that include any consecutive electricity price values that are not part of the same interval. It can be seen that an advantage of using this method is that the number of intervals and their limits don’t have to be specified. However, given the limitations of the method used to select the bandwidth and the number of not significant local minimums and maximums, the forecasting had to be carried out taking into consideration different numbers of interval to evaluate the influence. In Fig. 3, final chosen limits or borders are represented by a vertical dotted line.

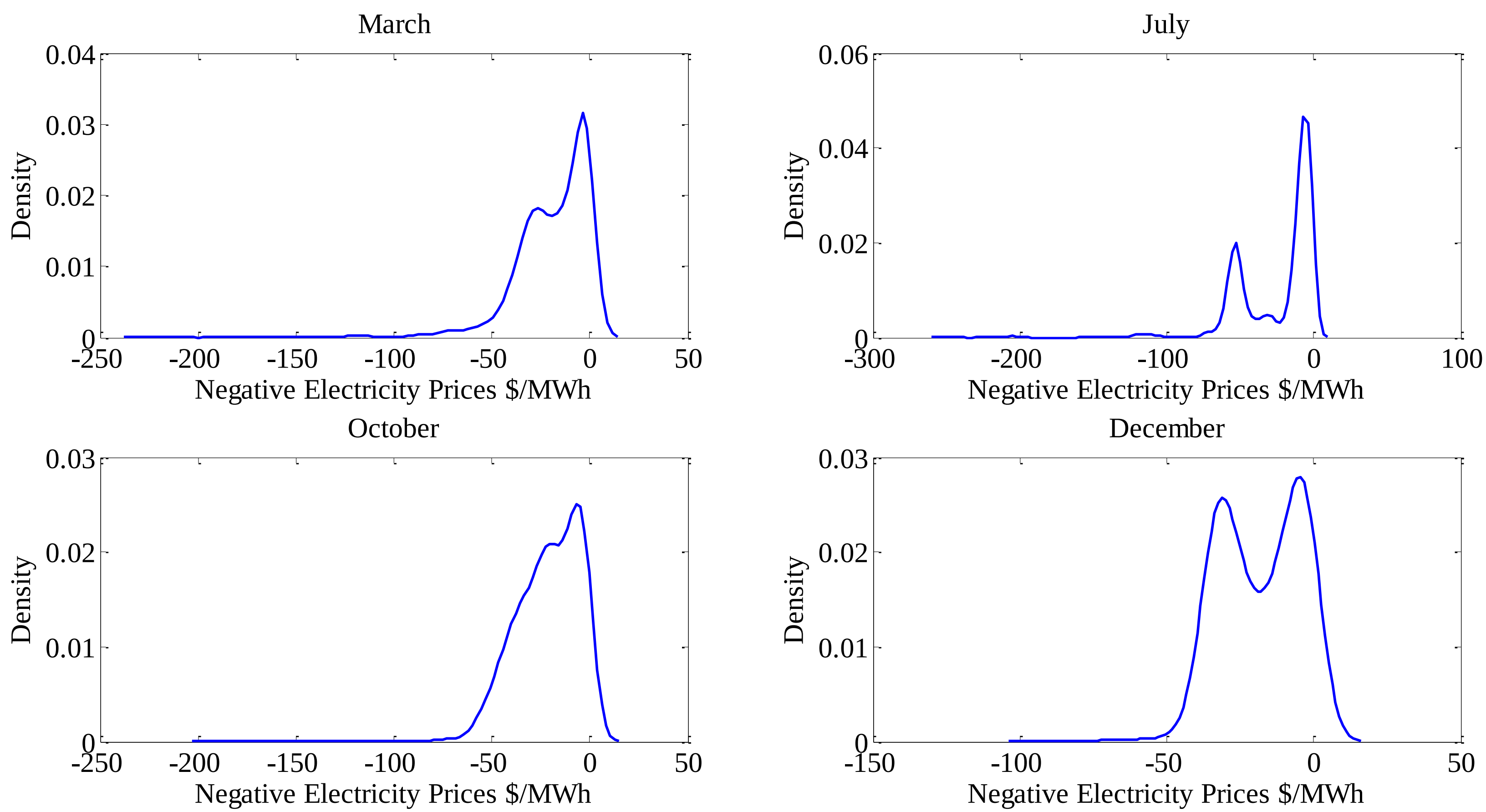

Figure 2. Probability Density Function of negative prices estimated for the months of March, July, October and December using Kernel Density Estimation

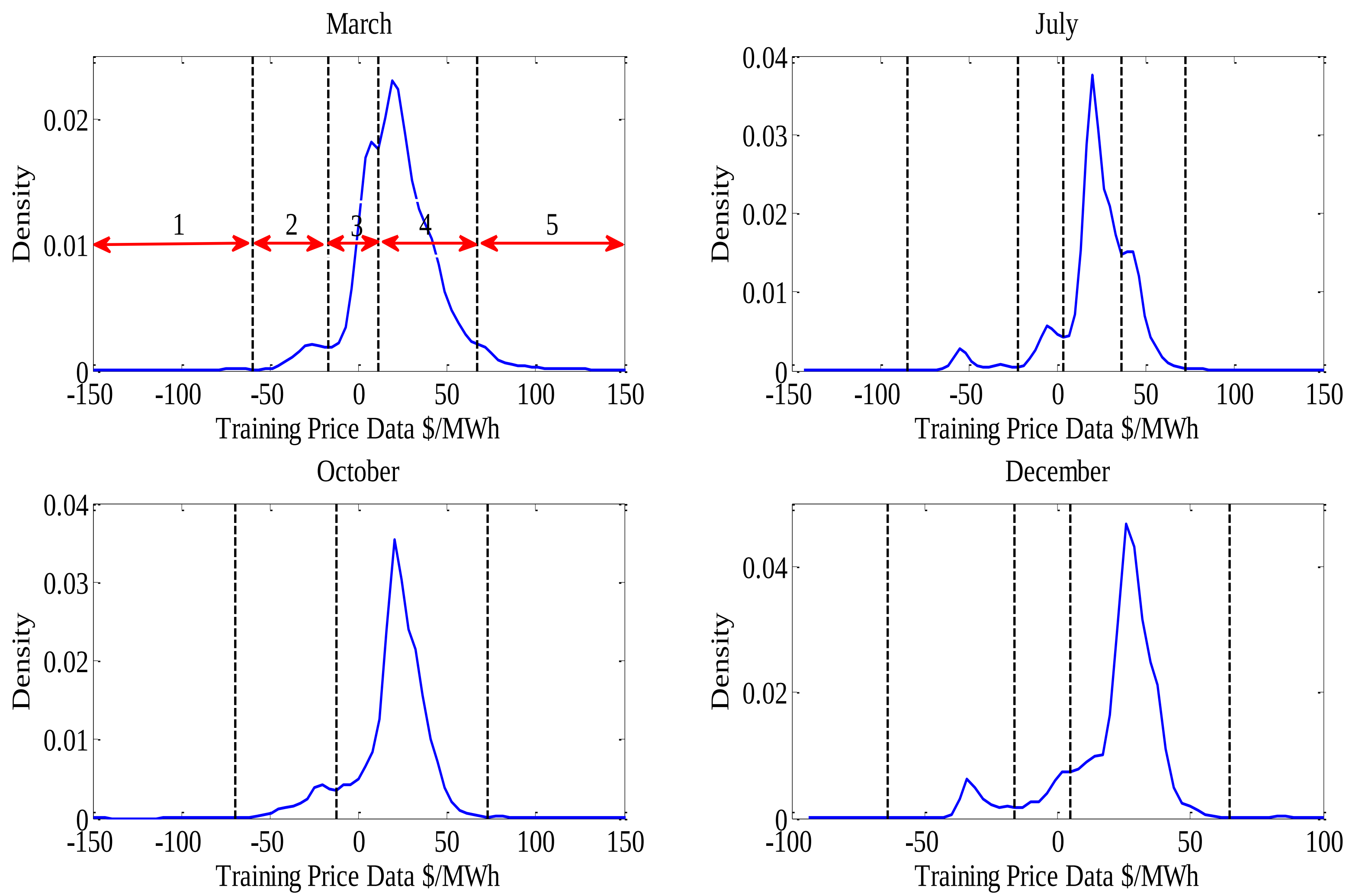

Figure 3. Probability Density Function of the first two weeks of March, July, October and December using Kernel Density Estimation

Table 7 shows the percentage of electricity prices on each interval, as it can be observed, most of the data points are allocated in the positive intervals since negative prices are still a rare event, and consecutive prices most often belong to the same interval. Yet the percentage of preceding consecutive prices that belong to different intervals does illustrate the sudden changes that electricity prices are well known for. Additionally, it shows that the highest percentage of negative prices are not generally associated with the highest density peak or “base regime”, instead they belong to interval 3 or the mixed interval. Interval 3 for in all months has one negative border or limit and one positive border or limit, within where the highest negative density peak shown in Fig. 2 is captured. Alongside, the mixed interval captures jumps and more subtle changes in prices.

Table 8 shows the conditional probability of a negative or positive electricity price given that the current and previous price values were part of certain interval. In association with Table 7, it can be inferred that the occurrence of positive prices when previous prices were negative beyond its first high density peak is almost non-existent (Intervals 1 and 2). Except for the first interval of October, where there is only two data point, one positive and one negative, that explains the sharply divided probability or either a negative or a positive price. Likewise, the occurrence of negative prices on intervals with positive borders or limits is very small. Therefore, with the mean shifting like separation and conditional probability, it is possible to identified intervals where the occurrence of negative electricity prices is almost sure or very unlikely, since there is not much price diversity in those price intervals. However, the complex dynamics of negative prices that belong to interval 3 and the mixed group is not reflected in the conditional probability where the negative prices are often undermine by the number of positive prices, especially for October where the high density positive and negative peak were too close.

To investigate further about the dynamics and occurrence of negative prices on those groups, and following the assumption that those are Markov processes; the electricity prices were modeled as a second order Markov Chain and the elements of probability transition matrices are shown in Table 9.

While modeling those intervals as Markov processes further help to identify the conditions under which negative prices occur, some especial cases were also highlight during the process. When current and immediately previous prices had been positive and negative, respectively, the future price has roughly the same probability to be either negative or positive. In those cases, it was considered that further price information was needed to forecast future prices; and therefore, future prices under this condition were forecasted using ARIMA(1,1,1).

Interval	March		July		October		December	
	% prices	*% negative prices*	*% prices*	*% negative prices*	*% prices*	*% negative prices*	*% prices*	*% negative prices*
1	0.37	3.1	0.09	0.72	0.04	0.14	0.04	0.35
2	2.09	15.74	2.43	17.48	7.06	43.06	4.30	31.93
3	20.42	**48.01**	4.80	**25.95**	85.13	**33.82**	5.36	**23.51**
4	58.38	4.58	56.71	14.77	0.56	0.00	78.96	10.53

5	5.03	0.00	18.67	0.18	7.21	**22.98**	0.78	0.0
6	13.71	**28.49**	0.56%	0.00			10.55	**33.68**
7			16.74	**40.90**				

Table 7. Percentage of Electricity Prices on Each Interval

Interval	March		July		October		December	
	Positive Price	*NEP*	*Positive Price*	*NEP*	*Positive Price*	*NEP*	*Positive Price*	*NEP*
1	0.06	0.94	0.00	1.00	0.50	0.50	0.00	1.00
2	0.18	0.82	0.13	0.87	0.08	0.92	0.08	0.92
3	**0.74**	**0.26**	**0.35**	**0.65**	**0.94**	**0.06**	**0.46**	**0.54**
4	0.99	0.01	0.97	0.03	1.00	0.00	0.98	0.02
5	1.00	0.00	1.00	0.00	**0.52**	**0.48**	1.00	0.00
6	**0.77**	**0.23**	1.00	0.00			**0.60**	**0.40**
7			**0.71**	**0.29**				

Table 8. Conditional Probability of Negative and Positive Electricity Prices for Each Interval

Interval 3								
	$P_{11,1}$	$P_{11,2}$	$P_{21,1}$	$P_{21,2}$	$P_{12,1}$	$P_{12,2}$	$P_{22,1}$	$P_{22,2}$
March	0.885	0.115	0.727	0.273	0.5	0.5	0.237	0.763
July	0.571	0.429	0.640	0.360	0.250	0.750	0.292	0.708
October	0.970	0.030	0.821	0.179	0.446	0.554	0.365	0.635
December	0.674	0.326	0.615	0.385	0.405	0.595	0.316	0.684

Mixed Interval								
March	0.949	0.051	0.761	0.239	0.569	0.431	0.233	0.767
July	0.958	0.042	0.703	0.297	0.462	0.538	0.190	0.810
October	0.957	0.043	0.660	0.340	0.227	0.773	0.283	0.717
December	0.847	0.153	0.787	0.213	0.375	0.625	0.296	0.704

Table 9. Elements of the Transition Probability Matrices

For the ARIMA model, two models were considered ARIMA(0,1,1) and ARIMA(1,1,1), and the parameters in (3-3) were estimated using maximum likelihood, and they are shown in Table 10.

Table 11 compares the forecasting results of the proposed method to the ARIMA methods. The results are evaluated by the percentage of correctly forecast negative prices, false negatives, correctly forecast positive prices and false positive. While the author acknowledges that the use percentages to compare results is restricting, the use of a simple method to forecast the value of the prices and then apply a commonly use forecast errors technique like MSE, MAE or MAPE would then undermine the method to identity the nature of negative prices drivers as discrete. The proposed approach forecasts correctly a higher number of negative prices. It consistently correctly forecast at least 5% more negative prices than the ARIMA models, except for March where the difference is 1.79%. However, as it was expected, the ARIMA models did a better job at correctly forecasting positive prices, but the proposed model was only behind for less than 1% in every case.

Parameter	March		July		October		December	
	ARIMA (0,1,1)	*ARIMA (1,1,1)*	*ARIMA (0,1,1)*	*ARIMA (1,1,1)*	*ARIMA (0,1,1)*	*ARIMA (1,1,1)*	*ARIMA (0,1,1)*	*ARIMA (1,1,1)*
Constant	0.000	0.009	0.000	0.003	0.000	-0.023	0.000	0.005
φ_1	-0.474	0.282	-0.690	0.308	-0.709	0.284	-0.524	-0.258
θ_1	0.000	-0.707	0.000	-0.924	0.000	-0.887	0.000	-0.297

Table 10. Estimated Model Parameters of ARIMA(0,1,1) and ARIMA(1,1,1)

Forecast		Negative	False Positive	Positive	False Negative
March	Proposed	71.89%	28.11%	91.87%	8.13%
	ARIMA (0,1,1)	68.45%	31.55%	92.76%	7.19%
	ARIMA (1,1,1)	70.17%	29.83%	92.91%	7.64%
July	Proposed	68.52%	31.48%	98.18%	1.82%
	ARIMA (0,1,1)	58.86%	41.14%	98.21%	1.79%
	ARIMA (1,1,1)	54.82%	45.18%	98.42%	1.58%
October	Proposed	81.83%	18.12%	95.56%	4.45%
	ARIMA (0,1,1)	75.34%	25.19%	95.35%	4.55%
	ARIMA (1,1,1)	74.31%	26.22%	95.62%	4.27%
December	Proposed	61.81%	38.28%	95.25%	4.75%
	ARIMA (0,1,1)	55.84%	44.26%	95.58%	4.41%
	ARIMA (1,1,1)	56.49%	43.61%	95.51%	4.49%

Table 11. Forecasting Results

Section 4.4. Chapter Summary

Through a comparative empirical study and using cross validation to quantify the generality of the results. It was found that the discrete approach used to forecast negative electricity prices outperformed the results of its continuous counterpart, and by proxy it showed that the nature of negative electricity price drivers is more appropriate represented as discrete. It consistently correctly forecast at least 5% more negative prices than the ARIMA models for all months, except for March where the difference is 1.79%; without excessively increasing the percentage of false negatives (less than 1%, when it is compared to the continuous counterparts).

Chapter 5: Demonstration of Economic Value of The Finding: Impact on The Economic Return of a Wind Farm

In this chapter, the significance of the appropriate representation on the nature of negative electricity price drivers on the sales revenue of a wind farm is empirically studied. In chapter 2, it was found that an appropriate forecasting algorithm and predictive wind farm control will improve the sales revenue of wind generation. In this case, an appropriate forecast is one who models the important characteristics that affect the economic return of a wind farm. To maximize the economic return of a wind farm, the forecast method should maximize the number of correctly forecast negative prices and minimize the number of false negatives (this is equivalent to maximize the number of correctly forecast positive prices).

As it was observed in chapter 4, neither the continuous nor the discrete models satisfy both requirements to maximize sales revenue. While, the simple method used to explore the nature of the negative electricity price drivers forecasts correctly more negative prices than its continuous counterpart without excessively generating false negatives; the continuous model correctly forecasts more positive prices but does a mediocre job of forecasting negative prices. Therefore, a hybrid model between the models would reflect how each model captures specific characteristics of the prices and together, they form a model that appropriately respond to the necessities of a prices forecast to improve sales revenue.

Section 5.1. Hybrid Model

The hybrid model is a slight modification of the approach proposed in Chapter 3. In the intervals of interest, when the present and previous prices were positive, instead of using the transition probability matrix to find the next state, ARIMA(1,1,1) was employed to forecast the future prices. The reasoning behind it is that either the probability of the next state to be positive was overwhelming or the probability of the next state to be either positive or negative is roughly the same. Either way, the scenarios are dominated by positive prices or marginally explained by the probabilities, hence the continuous model could be advantageous in those cases.

Section 5.2. Study Framework

This study is similar to the one made on section 2.1.4 but considering more composite and appropriate forecast algorithms. Moreover, the comparative approach is the same as in Chapter 4. This is, the impact on sales revenue of applying predictive control according of the proposed approach in Chapter 3, ARIMA(1,1,1) and the proposed hybrid model is computed for the testing data of every validation scenario, then the percentage of improvement with respect to the no curtailment scenario is calculated; finally, the results are the weighted average from all validation scenarios.

Section 5.3. Results and Analysis

As it was mentioned on Section 4.2, conventional validation (70% of the data is used for training and 30% for testing) and cross-validation (data was divided into two equally sized groups. The groups were not randomly sampled. In one case, first two weeks were set as the training set and the last two weeks as the testing set; and in the other, the first and the third week were set as the training set and the second and fourth weeks as the testing set) were used to quantify the generality of the results.

In Figure 4, the average percentage of increment in sales revenue for the testing price data of the month of March is shown. For all down times, the hybrid model has the highest percentage of increment in sales revenue, with difference of under 2% with the proposed approach and around 1% with ARIMA(1,1,1). The highest percentage of increment on sales revenue is 15.24%. The month of March presents the highest percentage of increment in sales revenue from the all four months. Alongside the magnitude of the prices, the number of NEPs in the testing data is a key factor in the increment of sales revenue. On average, the number of NEPs in the testing data is 624 out of 3833 electricity prices, this means that NEPs represent on average 16.7% of the price data which is quite higher than the average of the month (12.9%).

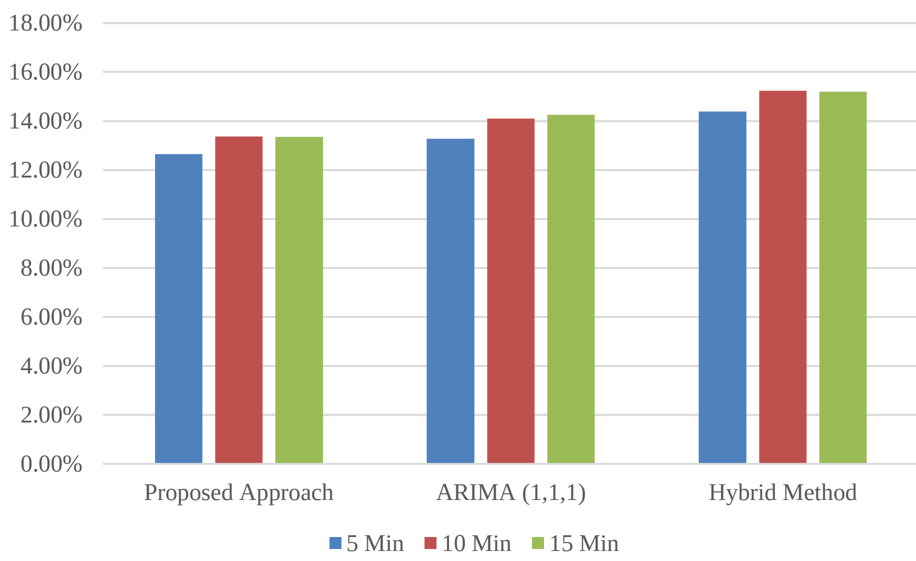

Figure 4. Average percentage of increment in sales revenue for the testing price data of the month of March

In Figure 5, the average percentage of increment in sales revenue for the testing price data of the month of July is shown. For all down times, the hybrid model has the highest percentage of increment in sales revenue, with difference of under 0.02% with the proposed approach and around 0.16% with ARIMA(1,1,1). The highest percentage of increment on sales revenue is 2.15%. The month of March presents the lowest percentage of increment in sales revenue from the all four months. Alongside the magnitude of the prices, the number of NEPs in the testing data is a key factor in the increment of sales revenue. On average, the number of NEPs in the testing data is 211 out of 3837 electricity prices, this means that NEPs represent on average 5.51% of the price data which is quite lower than the average of the month (9.45%). This low number of negative electricity prices would limit the impact of the forecasting of negative electricity prices on the sales revenue.

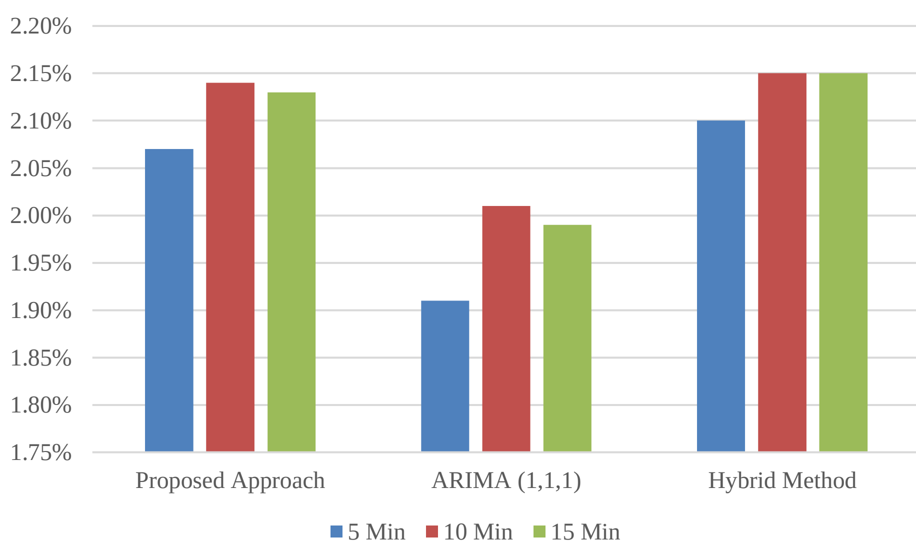

Figure 5. Average percentage of increment in sales revenue for the testing price data of the month of July

In Figure 6, the average percentage of increment in sales revenue for the testing price data of the month of October is shown. For all down times, the hybrid model has the highest percentage of increment in sales revenue, with difference of under 0.21% with the proposed approach and around 0.46% with ARIMA(1,1,1). The highest percentage of increment on sales revenue is 7.96%. On average, the number of NEPs in the testing data is 436 out of 3837 electricity prices, this means that NEPs represent on average 11.35% of the price data which is pretty close to the average of the month (11.12%). There are slight differences between the increment in sales revenue when different down times are considered, the difference between 5-minute minimum down and 10-minute minimum down time being the most perceptible. Since the differences are not significant, they can be attributed to nearby negative prices that were not initially captured by the approaches, but were close to negative prices that were captured and therefore, they are enclosed under the down time.

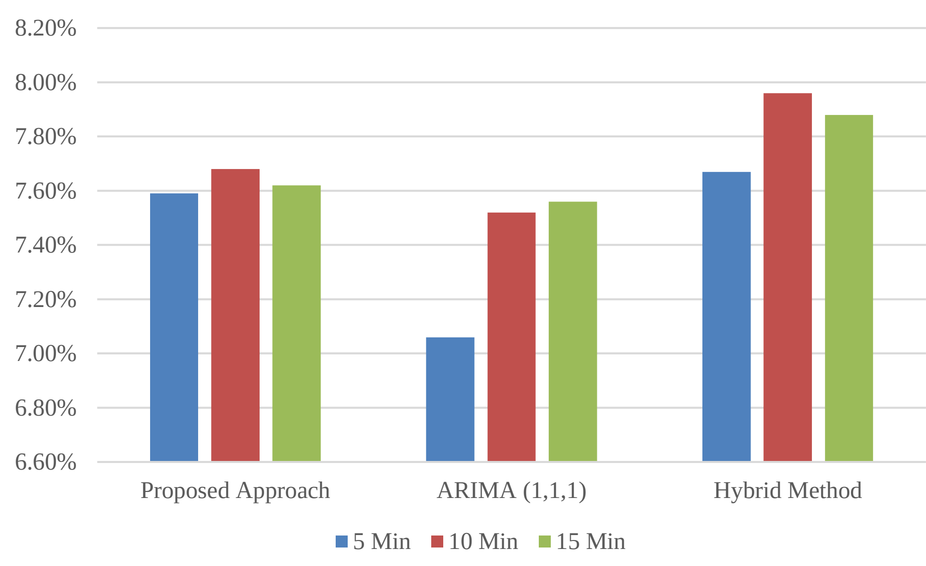

Figure 6. Average percentage of increment in sales revenue for the testing price data of the month of October

In Figure 7, the average percentage of increment in sales revenue for the testing price data of the month of December is shown. For all down times, the hybrid model has the highest percentage of increment in sales revenue, with difference of under 0.88% with the proposed approach and around 0.96% with ARIMA(1,1,1). The highest percentage of increment on sales revenue is 4.42%. On average, the number of NEPs in the testing data is 377 out of 3837 electricity prices, this means that NEPs represent on average 9.82% of the price data which is relatively similar to the average of the month (10.02%).

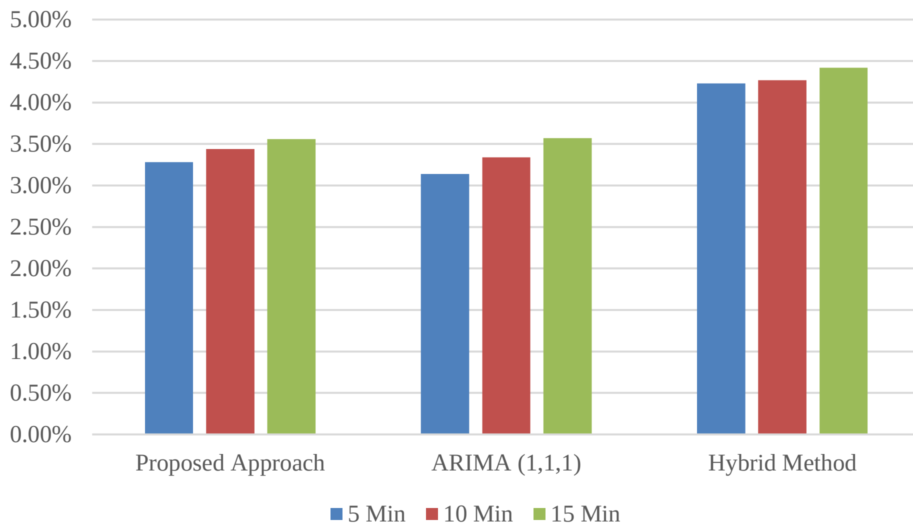

Figure 7. Average percentage of increment in sales revenue for the testing price data of the month of December

Section 5.4. Chapter Summary

Through the results, it was shown the positive influence of including the simple approach that captures the nature of negative electricity price drivers on sales revenue of a wind farm. The results showed that the hybrid model consistently has the highest increase on sales revenue of all methods. The hybrid method reflects how each model captures specific characteristics of the prices and together bring higher sales revenue to the wind farm. Inevitably, the revenue is tied and limited by the number of negative prices in the sample and their magnitude, then the increase in sales revenue varies greatly from one month to another depending on those factors. This is illustrated by the months of March and July. For March, the percentage of increase in revenue is as high as 15.19% and the testing price data has on average 624 negative electricity prices. Meanwhile for July, the percentage of increase in revenue is as high as 2.15% and the testing price data has on average 211 negative electricity prices.

Conclusions

1. The results of the study suggest that given the current market structure, the negative price drivers show indication of being of a discrete kind. It has been discussed in the literature that specific events prompt negative prices like negative bids and forecasting errors in demand, load and weather. A better understanding of the nature of those drivers is needed to effectively forecast and model negative electricity prices, since they would be reflected in the negative electricity prices' behavior. Negative electricity prices were therefore modeled as discrete events with the aim of showing the discrete nature of those drivers. Then the simple discrete approach used to forecast negative electricity prices outperformed the results of its continuous counterpart, and by proxy it showed that the nature of negative electricity price drivers is more appropriate represented as discrete.
2. Predictive wind farm control can be economically beneficial to wind farms who sign up for ITC. Before, wind energy generation was known to be driven by intermittency of their energy source, government's production incentives (PTC) and lack of control of its output. Nowadays, wind turbines and related technology have improved to a point where it is possible to control its output; and the federal government offers an alternative subsidy alternative that is not related to the amount of generation (ITC). These conditions make it possible for wind farms to explore curtailment alternatives to increase their sales revenue. For wind farms who signed up for ITC, curtailment at negative prices translates into higher sales revenues since no energy generated was paid at negative prices

while the decrease in generation do not affect the amount of subsidy received. In contrast, this is not the case for wind farms who opted for PTC since reductions in generation do affect the amount of subsidy received, and it would only be beneficial where the negative prices are greater than PTC plus shut down cost of the wind farm.

3. The benefit of such an understanding on the investment return on wind farm is demonstrated with a more general example. Particularly, it was shown how under certain conditions, wind farms who apply smart curtailment and receive ITC can be economically more attractive than a wind farm who receives PTC and do not apply curtailment. If it is considered the special case where the amount of money received by either PTC or ITC is the same, it is observable how for a wind farm who signed up for ITC, curtailment at negative prices translates into higher sales revenues without compromising the amount of subsidy granted by the government. Additionally, wind farms who apply smart curtailment and receive ITC are avoiding part of the risks associated with market volatility and variation in wind generation.

4. The economic value of integrating the negative prices forecast within the forecast of electricity prices and the impact of using the adequate type of forecast to apply curtailment on wind farms' revenue was shown. The forecasting of electricity prices was used to apply curtailment at negative forecasted prices, and the revenues considering the curtailment and different down times after shut downs were computed. The hybrid method between the proposed method to forecast negative prices and a continuous type forecasting

of electricity prices, produced higher sales revenues than both methods when they were considered individually. Therefore, the discrete forecast of negative electricity prices and the continuous forecast of electricity prices capture the different price characteristics, helping wind farms achieve greater sales revenues. This finding potentially encourages smart curtailment by wind farms, which indirectly benefits market participants and the for grid´s stability, and highlight the necessity of predictive wind farm control.

Future Work

1. Given that the negative price drivers show indication of being of a discrete kind under the current market structure, and the potential economic value of forecasting negative electricity prices for wind farms, future work on the negative price modeling could be done considering and exploring more complex models that further considers the dynamics of the time series and important additional parameters like wind speed and demand.
2. Some characteristics observed in the price data like slow changing and longer time duration than other "out of ordinary" price events like jumps could be further study to determine if negative electricity prices must be forecasted separately of other electricity prices as they may indicate that they are a different stochastic process within the electricity prices.
3. Future work would include a more comprehensive study on the economic value of predictive wind farm control, considering the generation curve and other technical and tax related dispositions to accurately compared the economic return of a wind farm signed up for ITC and applying predictive control against a wind farm signed up for PTC and that does not apply predictive control.

References

[1] M. Sathyajith, *Wind Energy:Fundamentals, Resource Analysis and Economics*, First. Springer-Verlag Berlin Heidelberg, 2006.

[2] D. Kirschen and G. Strbac, *Fundamentals of Power System Economics*. 2004.

[3] U. Aswathanarayana, "Renewable energy policies," in *Green Energy:Technology, Economics and Policy*, CRC Press, 2010, pp. 269–279.

[4] J. Eilperin and B. Dennis, "U.S., Canada and Mexico vow to get half their electricity from clean power by 2025," *The Washington Post*, 2016. [Online]. Available: https://www.washingtonpost.com/news/energy-environment/wp/2016/06/27/u-s-canada-and-mexico-to-pledge-to-source-half-their-overall-electricity-with-clean-power-by-2025/?utm_term=.d26e3d95d59a. [Accessed: 18-Aug-2016].

[5] A. H. Miranda and C. M. Villar, "Impacts of large penetration of wind in electricity systems. Probabilistic analysis on Spanish market.," *Energy Mark. (EEM), 2010 7th Int. Conf. Eur.*, pp. 1–7, 2010.

[6] Australian Goverment: Deparment of the Enviroment and Energy, "The Renewable Energy Target (RET) scheme." [Online]. Available: https://www.environment.gov.au/climate-change/renewable-energy-target-scheme. [Accessed: 08-Oct-2016].

[7] M. Milligan, P. Donohoo, D. Lew, E. Ela, B. Kirby, E. Lannoye, D. Flynn, M. O. Malley, N. Miller, P. B. Eriksen, A. Gøttig, B. Rawn, and M. Gibescu, "Operating Reserves and Wind Power Integration : An International Comparison Preprint," 2010.

[8] L. Deng, B. F. Hobbs, and P. Renson, "What is the Cost of Negative Bidding by Wind ? A Unit Commitment Analysis of Cost and Emissions," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1805–1814, 2015.

[9] R. Baldick, "Wind and Energy Markets : A Case Study of Texas," *IEEE Syst.*, vol. 6, no. 1, pp. 27–34, 2012.

[10] J. C. Smith, S. Member, S. Beuning, H. Durrwachter, S. Member, D. Hawkins, L. Member, B. Kirby, W. Lasher, K. Porter, K. Schuyler, and P. Sotkiewicz, "Impact of Variable Renewable Energy on US Electricity Markets," 2010.

[11] U.S. Internal Revenue Service, "RENEWABLE ELECTRICITY PRODUCTION TAX CREDIT (PTC)," *U.S. Deparment of Energy*, 2016. [Online]. Available: http://energy.gov/savings/renewable-electricity-production-tax-credit-ptc.

[12] U.S. Internal Revenue Service, "BUSINESS ENERGY INVESTMENT TAX CREDIT (ITC)," *U.S. Deparment of Energy*, 2016. [Online]. Available:

http://energy.gov/savings/business-energy-investment-tax-credit-itc.

[13] C. Grant, M. Bolinger, R. Wiser, K. Cory, and T. James, “PTC, ITC, or Cash Grant? An Analysis of the Choice Facing Renewable Power Projects in the United States,” no. March, 2009.

[14] Southwest Power Pool, “SPP: Generation Mix Historical,” *Southwest Power Pool Website*, 2016. [Online]. Available: https://marketplace.spp.org/generation-mix-historical.

[15] Southwest Power Pool, “2014 SPP Strategic Plan,” Little Rock, AR, 2014.

[16] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*. New York, NY, USA: Springer-Verlag New York, Inc., 1991.

[17] D. Keles, M. Genoese, D. Möst, and W. Fichtner, “Comparison of extended mean-reversion and time series models for electricity spot price simulation considering negative prices,” *Energy Econ.*, vol. 34, no. 4, pp. 1012–1032, 2012.

[18] J. Mei, D. He, R. Harley, and T. Habetler, “A Random Forest Method for Real-Time Price Forecasting in New York Electricity Market,” 2003.

[19] M. SEWALT and C. DE JONG, “Negative Prices in Electricity Markets,” no. June, pp. 74–77, 2003.

[20] S. Schneider, “Power Spot Price Models with Negative Prices,” *J. Energy Mark.*, 2011.

[21] J. Janczura, S. Trück, and R. C. Wolff, “Identifying spikes and seasonal components in electricity spot price data : A guide to robust modeling,” *Energy Econ.*, vol. 38, pp. 96–110, 2013.

[22] A. M. González, A. Muñoz, S. Roque, and J. García-gonzález, “Modeling and Forecasting Electricity Prices with Input / Output Hidden Markov Models,” *IEEE Trans. Power Syst.*, vol. 20, no. 1, pp. 13–24, 2005.

[23] O. Wu, T. Liu, B. Huang, and F. Forbes, “Predicting Electricity Pool Prices Using Hidden Markov Models,” *Int. Fed. Autom. Control*, vol. 48, no. 8, pp. 343–348, 2015.

[24] T. Kanamura and Ō. Kazuhiko, “On transition probabilities of regime switching in electricity prices,” *Energy Econ.*, vol. 30, pp. 1158–1172, 2008.

[25] T. D. Mount, Y. Ning, and X. Cai, “Predicting price spikes in electricity markets using a regime-switching model with time-varying parameters,” *Energy Econ.*, vol. 28, pp. 62–80, 2006.

[26] F. Paraschiv, S. Fleten, and M. Schürle, “A spot-forward model for electricity

prices with regime shifts," *Energy Econ.*, vol. 47, pp. 142–153, 2015.

[27] M. Coulon, W. B. Powell, and R. Sircar, "A MODEL FOR HEDGING LOAD AND PRICE RISK IN THE TEXAS," *Energy Econ.*, pp. 0–22, 2012.

[28] J. Janczura and R. Weron, "An empirical comparison of alternate regime-switching models for electricity spot prices," *Energy Econ.*, vol. 32, no. 5, pp. 1059–1073, 2010.

[29] M. S. Uddin and A. Kuh, "Online Bad Data Detection Using Kernel Density Estimation," pp. 0–4, 2015.

[30] R. Du, Q. Wu, X. He, and J. Yang, "MIL-SKDE : Multiple-instance learning with supervised kernel density estimation," *Signal Processing*, vol. 93, no. 6, pp. 1471–1484, 2013.

[31] C. C. Aggarwal and C. K. Reddy, *Data Clustering: Algorithms and Applications*. Chapman & Hall/CRC, 2013.

[32] Y. Cheng, "Mean Shift, Mode Seeking, and Clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 790–799, 1995.

[33] A. Shamshad, M. A. Bawadi, W. M. A. Wan Hussin, T. A. Majid, and S. A. M. Sanusi, "First and second order Markov chain models for synthetic generation of wind speed time series," *Energy*, vol. 30, no. 5, pp. 693–708, 2005.

[34] A. Carpinone, M. Giorgio, R. Langella, and A. Testa, "Markov chain modeling for very-short-term wind power forecasting," *Electr. Power Syst. Res.*, vol. 122, pp. 152–158, 2015.

[35] J. R. Norris, *Markov Chains*. Cambridge University Press, 1998.