UNIVERSIDAD PONTIFICIA COMILLAS

ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA (ICAI)

OFFICIAL MASTER'S DEGREE IN THE ELECTRIC POWER INDUSTRY

Master's Thesis

ECONOMIC ANALYSIS OF CONTINUOUS INTRADAY MARKET IN SPAIN FROM AN INDUSTRIAL CONSUMER'S PERSPECTIVE.

Author: Pilar Alfaro Cerezo
Supervisor: Pablo Fernández-Aller Horrillo

Madrid, July 2018

Master's Thesis Presentation Authorization

THE STUDENT:

Pilar Alfaro Cerezo

THE SUPERVISOR:

Pablo Fernández-Aller Horrillo

Signed: .... Date: 20/ 7/ 2018

Authorization of the Master's Thesis Coordinator

Dr. Luis Olmos Camacho

Signed: .... Date: ....../ ....../ ......

ECONOMIC ANALYSIS OF CONTINOUS INTRADAY MARKET IN SPAIN FROM AN INDUSTRIAL CONSUMER'S PERSPECTIVE

Author: Alfaro Cerezo, Pilar
Supervisor: Fernández-Aller Horrillo, Pablo

ABSTRACT

European energy policy in the last decades has focused on developing an electricity internal market. In this context, the XBID Project is created by several European power exchanges and transmission system operators, in order to develop an integrated Intraday Market, so that market participants' bids in one country can be matched with bids from market participants in any other country, as long as transmission capacity is available. Intraday Market in the Iberian Electricity Market was previously structured in six centralized auctions, based on marginal pricing with uniform prices. OMIE, REE and REN proposed a hybrid intraday market design in which continuous intraday market sessions are combined with the existing auctions. The Iberian Hybrid Intraday market started operating on June 12th, 2018, for the energy delivered on June 13th, 2018.

Advantages and disadvantages of this new market design have been discussed. It is argued that, as continuous trading prices vary from trade to trade, a unique and clear reference price is harder to identify. Auctions, on the other hand, provide a more reliable reference price. A hybrid market design as the one proposed by Iberian operators can help overcome this disadvantage of continuous trading, providing flexibility to agents' as well as clear price signals resulting from intraday auctions. Furthermore, the hybrid mechanism provides a playing field for small market participants, who may not have the resources to implement a 24/7 trading desk which would be required in an all-continuous intraday market. Also, the Iberian hybrid market design only considers hourly products. However, distributed resources are able to vary consumption and generation in very short periods of time. Thus, it would be beneficial to introduce shorter products as distributed resources penetration increases, especially as gate closure moves closer to real-time, in line with other European intraday markets. However, Spanish System Operator REE has already stated the operational difficulties associated to moving gate-closure up to one hour before delivery, using only hourly products, so resistance to further changes in this line can be expected. Also, this new model reduces lead time before delivery to 1 hour, increasing market agent's opportunities to correct their imbalances. In the case of an electro-intensive industrial consumer, reducing the extra-costs associated to imbalances could yield a competitive advantage. However, expected prices in intraday continuous market are needed in order to evaluate the incentive that industrial market agents have to correct their programs in the intraday continuous market.

In order to predict intraday continuous market prices, predictive models based on linear regression and multilayer perceptron methods have been used, based on variables such as hourly electricity demand, wind production or traded volume in intraday auctions. Non-linear models trained (MLP) show a better performance than linear regression models as the prediction of continuous intraday electricity prices is probably a non-linear function. Obtained continuous intraday prices over 2017 are generally higher than day-ahead prices and, in some cases, even higher than downward imbalance prices. High continuous intraday prices offer a high incentive to correct positive imbalances but a reduced incentive to correct negative imbalances. As the industrial consumer under-study generally creates positive imbalances, it may be able to strongly reduce the cost associated to imbalances by trading in the continuous intraday market. In the scenario considered, it would have been able to reduce its imbalance cost by 113%.

However, these results should be taken with caution. First, the ability of an industrial consumer to actually reduce its imbalance cost in this scenario depends on its productive process, as it is only a business opportunity if imbalances are generally positive and can be forecasted at least an hour before gate-closure. Second, results are subject to the ability of the model to accurately predict continuous intraday prices. As the training dataset used is small and only takes into account data from a specific time of the year, prediction models should be updated as the amount of data available increases, in order to improve their performance and avoid overfitting.

Table of Contents

List of Figures

List of Equations

Chapter 1

Introduction

European energy policy in the last decades has focused on developing an electricity internal market comprising several European states, that provides the right market signals in order to improve efficiency and sustainability, prices for industrial and domestic consumers and security and quality of supply [1]. In order to achieve these objectives in an efficient and cost-effective way, cross-border integration of the different European electricity markets is fundamental.

According to the European Commission, the net economic profit from the fulfilment of the internal market would be in the range of 16 and 40 billion euros per year [2]. It could satisfy demand at minimum cost and lower system operation costs and energy prices as dispatching expensive generators in one country would be avoided whenever more efficient and inexpensive units were available in another country. Competition will be enhanced as the number of agents in the market will be higher and it would drive prices down. System security would be increased due to the presence of many producers of different technologies and allow the development of renewable energy sources as their unpredictability is reduced in larger areas [3].

In order to reach the goal of developing an Internal Electricity Market in Europe, the European Commission promoted a Target Model for electricity day-ahead markets. According to it, these should be energy-only regional markets in which the price received by generators is the marginal price in each hour. In addition, these regional markets should be linked so that the lowest-priced bids in the coupled market are matched, regardless of the area where they have been submitted and only limited by the available transfer capacity between the regions [4]. This way, market liquidity would be increased and electricity prices volatility would be reduced.

Market coupling in Europe started by the Day-Ahead market. In 2006, the Trilateral Market Coupling among France, Belgium and the Netherlands Day-Ahead markets became effective. In 2010, Central West Europe (CWE) markets, including Belgium, the Netherlands, Luxembourg, France and Germany were coupled. In 2014, coupled markets included CWE, Great Britain, the Nordics and the Baltics. This was called the Multi-Regional Coupling (MRC) project, which, by 2016, included 19 European countries. In addition, Hungary, Romany, Czech Republic and Slovakia were coupled among them as part of the 4MMC project [5]. Figure 1 shows the countries included in the Multi-Regional Coupling initiative.

In parallel, in 2009, the Price Coupling of Regions (PCR) project started as an initiative of seven European Power Exchanges (APX, Belpex, EPEX SPOT, GME, Nord Pool Spot, OMIE and OTE). One of its key achievements was the development of the EUPHEMIA algorithm, acronym for Pan-European Hybrid Electricity Market Integration Algorithm, now used in 23 European countries [6] for the calculation of prices and to allocate cross-border capacity on a Day-Ahead basis. As the existing interconnections were only intended in the first place to increase system's reliability and they were not designed to put up with the flows resulting of an efficient regional market [3], efficient allocation of scarce transmission capacity among states is crucial. However, it may be argued that, in order to take advantage of the benefits of the Internal Electricity Market, interconnection capacity between European countries must be increased. A lack of transmission capacity may prevent many of the benefits of market coupling [5].

All in all, the Day-Ahead Market Target Model has been progressively implemented and in 2017, it covers approximately 75% of the consumption in Europe. However, there are other physical markets for trading electricity where coupling and harmonization are still needed: forward markets and balancing and intraday markets [7].

Regarding the latter, in August 2015, the European Commission published the Commission Regulation (EU) 2015/1222 establishing a guideline on Capacity Allocation and Congestion Management [8]. In this document, it established the Target Model for Intraday markets' integration, based on a cross-border continuous intraday market which will allow agents in different regions to trade electricity continuously throughout the day. ENTSO-E requires that cross-border capacity is allocated based on implicit methods, that is, market agents bid in energy only markets, and capacity is allocated implicitly. On the other hand, under explicit capacity allocation, market agents bid for capacity first. Once they have obtained the required capacity, they can bid in energy markets. Explicit allocation of cross-border capacity may lead to a less efficient result [9].

Figure 1 - Countries included in the Multi-Regional Coupling, 4MMC and using PCR algorithm – Source: [9]

In order to put in place this Target Model for Intraday Market, the Cross-Border Intraday Initiative – XBID Project is created: power exchanges and transmission system operators from 11 European countries are working together in order to develop an integrated intraday market. The different regions that wish to incorporate to the XBID project should develop Local Implementation Projects to develop and prove the different processes required by the XBID model [7].

In this context, OMIE, market operator at the Iberian Electricity Market, together with REE and REN, Spanish and Portuguese system operators, respectively, have proposed a new model for the Iberian Electricity Intraday Market [10], in order to comply with the latest European Commission requirements.

a. Motivation

Intraday Market in the Iberian Electricity Market is currently structured in six centralized auctions or sessions. Each session has a different timetable and allows trading electricity for different time horizons. Table 1 shows the schedule of the different intraday sessions, where D refers to the day of electricity delivery and D-1 the day before. Intraday Market 1 for delivery day D and Intraday Market 7 for delivery day D-1 are held in conjunction on day D-1.

Intraday Market	**IM 1**	**IM 2**	**IM 3**	**IM 4**	**IM 5**	**IM 6**	**IM 7**
Timetable	17h – 18h45 D-1	21h – 21h45 D-1	01h-01h45 D	04h-04h45 D	08h-08h45 D	12h-12h45 D	17h-18h45 D
Products traded	1-24 D	1-24 D	5-24 D	8-24 D	12-24 D	16-24 D	22-24 D

Table 1 - Schedule of the Iberian Intraday Electricity Market. Source: [11]

The new Intraday Market proposed by OMIE, REE and REN is a hybrid system as it mixes continuous intraday market sessions with the existent auctions. Different models are under study and rules defining functioning are still being developed. A transition scheme may be in place for some time but, eventually, the distribution of sessions according to OMIE, REE and the CNMC will be as shown in Table 2.

With an increased number of intraday sessions, it seems probable that agents will be able to renegotiate their production/consumption closer to delivery time, reducing their imbalance costs.

Session Schedule (D)		Products traded	
Opening	**Close**	**Auction session**	**Continuous Session**
0	1		3-24 D
1	2	5-24 D	4-24 D
2	3		5-24 D
3	4		6-24 D
4	5	8-24 D	7-24 D
5	6		8-24 D
6	7		9-24 D
7	8		10-24 D
8	9	12-24 D	11-24 D
9	10		12-24 D
10	11		13-24 D
11	12		14-24 D
12	13	16-24 D	15-24 D
13	14		16-24 D
14	15		17-24 D
15	16		18-24 D
16	17		19-24 D
17	18	22-24 D, 1-24 D+1	20-24 D
18	19		21-24 D
19	20		22-24 D
20	21		23-24 D
21	22	1-24 D+1	24 D
22	23		1-24 D+1
23	0		2-24 D+1

Table 2 – Potential schedule of the Iberian Intraday Electricity Market under new hybrid model. Source: [10]

b. Objectives

In light of all of the above, the main objectives of the present Master Thesis are defined.

- To understand the current functioning of the Iberian Day-Ahead and Intraday markets and the new functioning of the Iberian Intraday market under the hybrid intraday model.
- To determine the impact of the increased intra-day schedule on the operation of an industrial consumer.
- To estimate the cost-reduction associated to a more efficient correction of their imbalances by estimating the continuous intraday market average price using a prediction model.

Chapter 2

The Spanish Electricity Market

a. Introduction

In October 1st, 2004, the Spanish and Portuguese electricity markets merged to create the Iberian Electricity Market (MIBEL). It involves a sequence of markets where generation and demand trade energy for different time periods, as shown in Table 3: day-ahead and intraday markets, followed by the balancing services markets – technical constraints management, frequency restoration reserves, replacement reserves, imbalances management and additional upwards reserve. Spanish agents can also trade electricity through physical and bilateral contracts. Market agents holding physical bilateral contracts must inform the System Operator before the day-ahead market is held.

Schedule	Market type	Market name	Market Operator	Product
Weeks, months, years before delivery day D	**Forward market**	Over the counter	None	Physical and financial forward contracts
		MIBEL	OMIP (Portuguese pole of MIBEL)	
D-1	**Day-ahead market**		OMIE (Spanish pole of MIBEL)	Hourly energy for each hour of day D
D	Short-term and balancing services markets	**Intraday markets**	OMIE	Hourly energy for different hours of day D

Table 3 - MIBEL Market sequence. Source: [12]

OMIE manages the Day-Ahead and Intraday markets, without considering the restrictions imposed by the network. Later on, each TSO manages technical constraints and redispatches if necessary to obtain a feasible output. If the TSO considers that low reserve margins are detected - the day-ahead market schedule does not guarantee enough reserves to face demand and generation unbalances in real time- additional reserves are procured trough the Additional Upwards Reserve market. Once it is closed, the TSO procures Frequency Restoration Reserves (FRR), Replacement Reserves (RR) and, if necessary, calls the Deviation Management market, which will be further explained in Chapter 3.

The objective of this chapter is to carry out a descriptive investigation about the functioning of the Spanish Electricity Day-Ahead and Intraday Markets, as well as a description of the main European projects regarding Day-Ahead and Intraday market coupling in Spain, in order to define a clear starting point for the subsequent economic study.

b. The Spanish Electricity Day-Ahead Market

In the Spanish Day-Ahead Electricity Market, operated by OMIE, agents buy and sell energy every day for each hour of the next day. Up to 85% of the Spanish energy is traded in the Day-Ahead market, which makes it the most liquid energy market in Spain and thus, the one with the highest influence on the final price of electricity [13].

Market clearing is based on a marginal pricing model, using EUPHEMIA algorithm. All generators receive the same price and all consumers pay the same price, which is determined by the point where supply and demand curves meet. These curves are determined by aggregating all agents' bids. A generator bid must include the amount of energy he is willing to produce and the minimum price at which he is willing to produce it. Market rules establish that, in order to achieve an efficient market clearing, price should be the opportunity cost associated to generating electricity, that is, the costs associated to producing and the income the generator will not receive because of it. Once all generators have presented their bids, OMIE aggregates and sorts them from low to high prices, as it can be seen in Figure 2.

All available production units not bound to a physical bilateral contract must present bids on the Day-Ahead market. Since bid prices are generator's opportunity costs, solar, wind or run-of-river hydro generators appear in the lower part of the supply curve while thermal generators appear in the upper part, as they may use their fuel to produce electricity but they could also resell this fuel in another market. Manageable hydro plants are also in the upper part of the curve as their opportunity cost includes the possibility of holding the water for higher-priced periods.

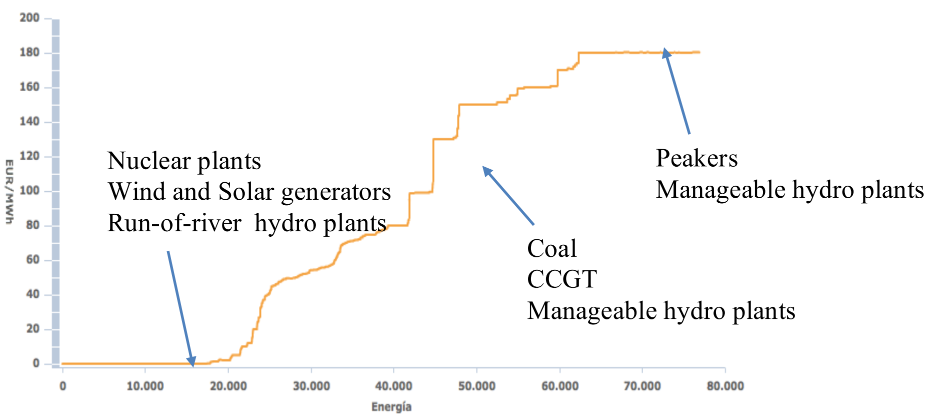

Figure 2 - Aggregated supply curve for hour 1, 28/12/2017. Source: [12]

On the other hand, direct consumers such as industrial consumers and electricity suppliers send demand bids stating the amount of energy they are willing to consume and the maximum price at which they are willing to buy it. Once all consumers have presented their bids, OMIE aggregates and sorts them from high to low prices, as it can be seen in Figure 3.

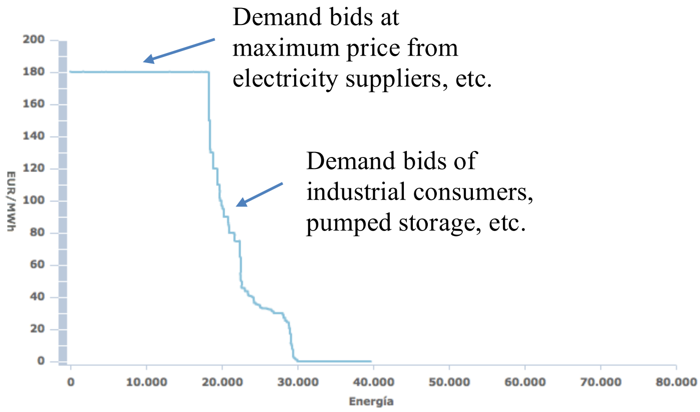

Figure 3 - Aggregated demand curve for hour 1, 28/12/2017. Source: [12]

Market price for hour h on day D and cleared supply and demand bids are determined by the cut-off point between supply and demand curves. In order to do this, complex bids must be taken into account.

A complex bid is a bid that includes not only a minimum price and quantity to be produced, but also other conditions that must be fulfilled for the bid to be cleared, such as load gradients or minimum income conditions. Load gradients define the maximum difference between energy produced in one hour and the next one, in order to guarantee that bids cleared are compatible with production units' ramps. The minimum income condition defines that a production unit is only cleared if the total production cleared for day D obtains at least a certain income. As seen in Figure 4, after complex bid conditions are taken into account, the aggregated supply curve is altered. Price and quantity cleared are then determined. Price can vary from 0 to 180,3 €/MWh

Both Portuguese and Spanish generators and consumers participate in OMIE's Day-Ahead market. Thus, generally, both countries will have the same reference Day-Ahead price. However, in case of network congestion, a "market splitting" takes place and each country is cleared independently. In these cases, prices in Portugal and Spain are different. In 2017, Portugal and Spain had the same reference price in 93% of the hours [14].

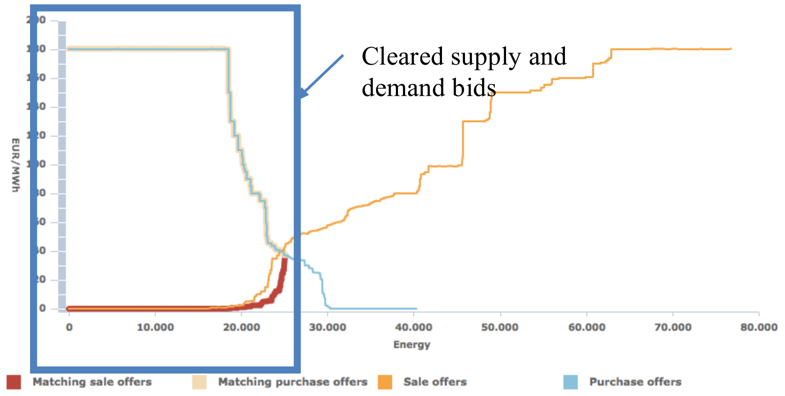

Figure 4 - Aggregated demand and supply curves for hour 1, 28/12/2017. Source: [12]

c. PCR Project: EUPHEMIA

In 2009, the Price Coupling of Regions (PCR) project started as an initiative of seven European Power Exchanges (APX, Belpex, EPEX SPOT, GME, Nord Pool Spot, OMIE and OTE) in order to advance in Europe's electricity market coupling by "developing a single price coupling solution to be used to calculate electricity prices across Europe respecting the capacity of the relevant network elements on a day-ahead basis" [15]. In the context of the PCR project, the EUPHEMIA algorithm was developed. EUPHEMIA is an acronym for Pan-European Hybrid Electricity Market Integration Algorithm, used for the calculation of prices and to allocate cross-border capacity on a Day-Ahead basis.

EUPHEMIA's objective is to determine which purchase and sale bids are cleared so that social welfare is maximized and power flows between countries resulting from clearing the market do not exceed network's capacity. Maximizing social welfare means maximizing consumer surplus, producer surplus and congestion rents across the regions [6].

Consumer surplus is defined as the difference between the total utility obtained by the consumer from the electricity purchased and the total price paid for it. Producer surplus, on the other hand, would be the difference between the total revenue obtained in the market and the total cost of producing the electricity sold [6]. Congestion rents are the product between the price spread between regions and the flow through the lines that connect them, in order to maximize flows from less to more expensive zones.

Euphemia returns a market price for each region - typically each country -, the quantities matched, the difference between the matched supply and the matched demand in that region - its "net position" - and the flow through the interconnectors between regions [6]. It has been used to clear the Iberian Day-Ahead market since 2014. In 2017, it is used to clear 23 European Day-Ahead markets [16]. EUPHEMIA development and implementation has been a milestone of vital importance in the process of attaining a fully-coupled European internal electricity market.

d. The Spanish Electricity Intraday-Market

Intraday Markets are adjustment markets where agents may modify their production or consumption schedules by buying or selling additional energy. Intraday Market in the Iberian Electricity Market is currently structured in six centralized auctions or sessions, organized by the power exchange OMIE and also based in marginal pricing with uniform prices, similar to the Day-Ahead Market. As shown in Table 1, each session has a different timetable and allows trading electricity for different time horizons. Both purchase and sale bids may include complex conditions.

Market clearing can be straightforward if only simple bids are considered, as different periods can be cleared separately. However, when complex bids are also considered, market clearing is an iterative process as different periods' outputs are linked by complex conditions. Likewise, for each period, price and cleared supply and demand bids are determined by the cut-off point between supply and demand curves.

Capacity in the Spain-France interconnection is allocated explicitly, that is, auctioned to the market separately from the marketplace where energy is auctioned, in the form of physical transmission rights [17].

e. XBID Project: Iberian Local Implementation Project

Unlike Day-Ahead markets, Intraday Markets in the European Union are still far from reaching the harmonization and coupling required for the achievement of an Internal European Market. The main Intraday Markets in Europe, besides the Iberian Intraday Market are:

- Central Europe Intraday Market, operated by EPEX Spot and including since 2016 France, Germany, the Netherlands, Belgium, Luxembourg, Switzerland, Austria and UK's Intraday Markets. It is based on a hybrid model including intraday auction sessions and continuous trading 24/7. Energy can be traded up to only 30 or 60 minutes before delivery [18].
- Nordic Countries Intraday Market, operated by NordPool Spot and Including Norway, Sweden, Finland, Denmark, Estonia, Latvia, Lithuania, UK and Germany's Intraday Markets. It is based on a continuous trading model open 24/7 where 15, 30 and 60 minute products can be traded. Energy can be traded up to only 60 minutes before delivery [19].
- Italian Intraday Market, operated by GME. It is similar to Iberian's Intraday Market as it is based on centralized auctions where hourly products are traded.

In particular, Italy has six price zones as interconnection between them is limited [20].

As seen in Figure 5, it is clear that traded products, gate closure times, bid formats, structure (continuous or auction-based) and other basic features need to be homogenized. As mentioned before, in August 2015, the European Commission published, as agreed by ACER and ENTSO-E, the Commission Regulation (EU) 2015/1222 establishing a guideline on capacity allocation and congestion management [8] in Intraday Markets. In this document, it established the Target Model for Intraday markets' integration, based on a continuous intraday market which will allow agents in different regions to trade electricity continuously throughout the day.

As a response to this, the Cross-Border Intraday Initiative – XBID Project is created by several European power exchanges and transmission system operators, in order to develop an integrated Intraday Market, so that market participants' bids in one country can be matched with bids from market participants in any other country, as long as transmission capacity is available. The different regions that wish to incorporate to the XBID project should develop Local Implementation Projects to develop and prove the different processes required by the XBID solution [7].

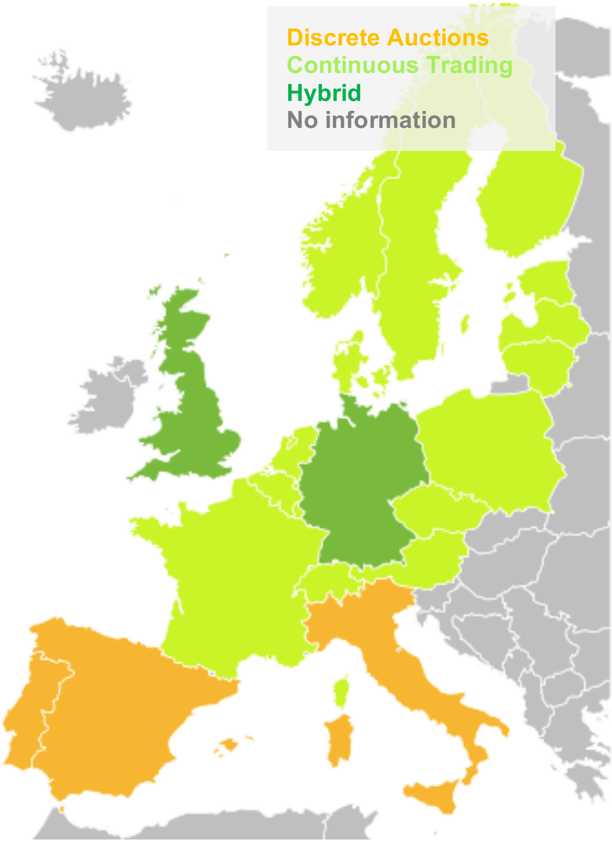

Figure 5 - Different designs of intraday markets in Europe in 2015. Source: [21]

According to this, OMIE, REE and REN proposed a hybrid system in which continuous intraday market sessions overlap with the existing auctions. ENTSO-E agrees with the possibility of a hybrid market mechanism combining both auctions and continuous trading. However, it should not impact liquidity at the European level or discriminate between adjacent regions and should allow market agents to trade closer to real time.

OMIE, REE and REN defined two different approaches, which were then submitted to public consultation [22]:

- Model A, preferred by the TSOs, in which continuous trading sessions would open after the closure of an intraday auction for those hourly periods that would not be negotiated in the next intraday auction. This model allows trading an hourly product after it has been negotiated in an Intraday auction up to an hour before delivery, increasing agent's flexibility in terms of deviations management.

- Model B, preferred by market agents, in which continuous trading sessions would overlap with intraday auctions and would allow negotiating all remaining hourly periods except the immediate next one, as can be seen in Table 2.

After public consultation, OMIE, REE and REN finally decided on implementing model A in the short-term, with a switch to model B after a 5-month period, in the context of the Iberian Local Implementation Project (LIP). **The Iberian Hybrid Intraday market started operating on June 12th 2018, for the energy delivered on June 13th 2018.**

Under this new Intraday Market, capacity in the Spain-France interconnection will be allocated implicitly, that is, included in the energy market, in order to define the most economic dispatch for the whole regional market, making sure that energy flows from low price areas towards high price areas.

In the new continuous intraday market, agents will present bids through the Local Trading Solution (LTS), a trading platform developed by OMIE. In it, each contract is associated to an hourly period with an Order Book, containing the bid and ask offers which, due to their prices, cannot be matched. When an order is introduced in the platform ("aggressor bid"), it is compared to the opposite nature bids in the Order Book for that specific contract ("passive bids"). If the buy bid price is higher or equal to the sell bid price, the aggressor bid can be matched with one or more passive bids, forming trades at the passive bid's price [9].

In [23] and [24], advantages and disadvantages of continuous and discrete auction-based intraday markets are discussed. In favor of continuous trading it is argued that it provides a higher flexibility to market agents than auctions do, as these are cleared at discrete times. This is an important feature in high renewable energy sources systems. However, continuous trading may reduce allocation efficiency as trades are formed in a first-come-first-served basis: trades increasing welfare might not take place while other trades decreasing welfare might. Also, in continuous trading it is not guaranteed that cross-border capacity is allocated to those agents willing to pay more for it [11] and liquidity tends to be lower than in auction-based markets, as auctions gather all agents at a specific time. Thus, the hybrid market design proposed by Iberian operators may combine both continuous trading and auctions advantages, providing both flexibility and a liquid marketplace.

However, it is also argued that the Spanish intraday market has high liquidity not only due to the market design, but also due to many other regulatory aspects which influence the behavior of market agents in intraday markets. For example, in Spain, a dual imbalance pricing mechanism is applied, as will be described in Chapter 3. This dual imbalance reduces significantly any incentives market agents may have to deviate from their programs, and provides a reason for agents to correct their programs in the intraday market. Also, since 2007, in Spain, renewable units with more than 1 MW of capacity are responsible for their imbalances. As intermittent RES-E can be subject to forecast errors, these agents have a very high incentive to participate in intraday markets. Finally, balancing markets and technical constraints solving procedures may require agents providing the service to vary their programs in the intraday market. For example, the provision of additional upwards reserve market is only done by thermal generators who have not been cleared in the Day-Ahead market. If they are required to provide additional upwards reserve, they will need to sell their minimum technical output in the intraday market and will probably do it at a price under their marginal price, as they also receive the marginal price of the upward reserve market, in €/MW [11].

As continuous sessions in the Iberian hybrid intraday market will allow agents to modify their programs up to one hour before delivery, when RES-E forecasts are more accurate, renewable generators may be able to further reduce their imbalances by trading in the continuous market sessions. Therefore, liquidity in intraday market might not be very much affected by the change in market design, although a future reduction in the number of intraday auctions might help promote liquidity in continuous trading.

It is also argued that, as continuous trading prices vary from trade to trade, a unique and clear reference price is harder to identify – price transparency is reduced. Auctions, on the other hand, provide a more reliable reference price. A hybrid market design as the one proposed by Iberian operators can help overcome this disadvantage of continuous trading, providing flexibility to agents' as well as clear price signals resulting from intraday auctions. Furthermore, the hybrid mechanism provides a playing field for small market participants, who may not have the resources to implement a 24/7 trading desk which would be required in an all-continuous intraday market.

Finally, the Iberian hybrid market design only considers hourly products. However, distributed resources are able to vary consumption and generation in very short periods of time. Thus, it would be beneficial to introduce shorter products as distributed resources penetration increases, especially as gate closure moves closer to real-time, in line with other European intraday markets. However, Spanish System Operator REE has already stated the operational difficulties associated to moving gate-closure up to one hour before delivery, using only hourly products, so resistance to further changes in this line can be expected.

To sum up, this new market design probably shows many advantages over the previous one, but further changes are required such as a reduction in the number of auctions and the introduction of shorter products. These changes may still take a long time to be implemented, due to the difficulties found by Iberian operators in the launching of the XBID hybrid intraday market. Also, the impact of the increased intra-day schedule on the operation of industrial consumers and the cost-reduction associated to a more efficient correction of their imbalances will depend on market liquidity and prices obtained in intraday continuous sessions.

f. Conclusions

The purpose of this chapter is to define a clear starting point for the following economic study. It can be concluded that:

- PCR project and the development of EUPHEMIA algorithm brought Day-Ahead Markets in Europe to the last stages of market coupling.

- However, harmonization of other markets such as forward or intraday markets is still needed.

- XBID project aims to develop an integrated Intraday Market based on continuous trading and implicit capacity allocation, a model far from the Iberian Intraday Market current design based on Intraday Auctions.

- OMIE, REE and REN have proposed a hybrid intraday market combining both centralized auctions and continuous trading, which makes possible to trade energy closer to delivery time.

- The impact of the increased intra-day schedule on the operation of industrial consumers will depend on market liquidity and prices obtained in the continuous market sessions.

- Market liquidity in the continuous market sessions will not only depend on market design but will also be influenced by other regulatory aspects characterizing the Spanish energy market such as imbalance pricing mechanisms.

- Further changes may be required such as a reduction in the number of auctions and the introduction of shorter products. These changes may still take a long time to be implemented, due to the difficulties found by Iberian operators in the launching of the XBID hybrid intraday market.

Chapter 3

Study of Imbalances in an Industrial Consumer

a. Introduction

Balancing supply and demand is one of the main responsibilities transferred to the System Operator after gate-closure. Real-time mismatches between supply and demand are mainly due to market agents' deviations from their expected consumption and production programs. These imbalances are managed by the System Operator through balancing markets, where market agents send their bids to balance the system at the lowest cost. The cost of resolving these real-time differences between generation and demand may be allocated to all users or just to the ones responsible for the imbalance [25] [20].

The objective of this chapter is to study the imbalance cost-allocation and pricing mechanisms in Spain, in order to quantify the extra-cost incurred by an industrial consumer due to deviations from its expected consumption program.

b. Imbalance pricing mechanisms. The Spanish case.

In most European countries, the cost of resolving the real-time system mismatches is allocated to those agents creating the imbalance. Imbalances are defined as the difference between the production or consumption measured at power station busbars and the final production or consumption program scheduled in the Day-Ahead and Intraday Markets. Imbalances can be negative (or *downwards*) or positive (or *upwards*). A negative imbalance creates a shortage of energy in the system: the production measured is lower or the consumption is higher than scheduled in the market. A positive imbalance creates a surplus of energy in the system: the production measured is higher or the consumption is lower than scheduled in the market. It should be noted that, in general, agents creating a positive imbalance *receive* the imbalance price, while agents creating a negative imbalance *pay* the imbalance price

One of the main features of an imbalance settlement mechanism is the imbalance pricing system, which can be based on dual pricing or single pricing. Under a single imbalance pricing system, positive and negative imbalances face the same imbalance price, related to the weighted average cost of balancing services. Under a dual imbalance pricing, positive and negative imbalances face different prices. Typically, imbalances contributing to the system imbalance receive a price that reflects the cost of balancing the system while imbalances helping reduce the system imbalance face the day-ahead market price. This system's typical functioning is detailed on Table 4 [11].

If an agent deviates upwards whenever the total system imbalance is downwards, he receives the Day-Ahead market price for the extra energy produced or the non-consumed energy. If an agent deviates upwards whenever the total system imbalance is upwards, he receives for the extra energy produced or the non-consumed energy the minimum between the Day-Ahead market price and the weighted cost of the balancing actions adopted by the System Operator to reduce the imbalance. If an agent deviates downwards whenever the total system imbalance is upwards, he pays the Day-Ahead market price for the extra energy consumed or the non-produced energy. If an agent deviates downwards whenever the total system imbalance is downwards, he pays for the extra energy produced or the non-consumed energy the maximum between the Day-Ahead market price and the weighted cost of the balancing actions adopted by the System Operator to reduce the imbalance.

	System negative imbalance	**System positive imbalance**
Agent positive imbalance Agent receives…	Day-Ahead market price	Min (Day-Ahead market price, weighted average price of Balancing Services)
Agent negative imbalance Agent pays…	Max (Day-Ahead market price, weighted average price of Balancing Services)	Day-Ahead market price

Table 4 - Dual imbalance pricing mechanism. Source: [20] [26]

Therefore, under a dual imbalance pricing mechanism, imbalances helping the system reduce its total imbalance are nor penalized nor incentivized while those imbalances against the system are penalized. This dual-pricing mechanism, as opposite to the single-pricing mechanism, aims at reducing every incentive to deviate, as it is never a business opportunity. In the single-pricing mechanism, deviating in the opposite direction than the system is incentivized as agents contributing to reduce the system imbalance share the cost-savings associated.

In Spain, imbalance pricing is based on a dual-pricing mechanism that allows allocating balancing costs among the responsible parties. It should be noted that the European Commission Network Code on Electricity Balancing, published November 2017, requires that each TSO develops by November 2018 a proposal in order to harmonize, among others, the use of single imbalance pricing mechanism, where both positive and negative imbalances face the same price. As these Network Codes are published as European Commission's Regulations, Spain may eventually have to switch to a single pricing mechanism [27] [28].

The balancing actions considered in Spain when computing the cost of balancing the system include Deviation Management, Tertiary Regulation and Secondary Regulation balancing energy services.

Deviation Management

The Deviation Management market mechanism - developed in REE's Operational Procedure 3.3 – is used by the System Operator to solve the supply and demand unbalances that may occur between each intraday market and the real-time operation. The SO asks generators for bids in the opposite direction of the expected deviation, and the cleared bids receive the marginal price (*pay-as-cleared*). The market only takes place if the expected imbalance in an hour surpasses 300 MW [29] [11].

Tertiary Regulation

The Tertiary Regulation market - Manual Frequency Restoration Reserve (mFRR) under the European nomenclature – is developed in REE's Operational Procedure 7.3. In this market, qualified generators must declare their whole upward and downward available capacity and present price bids for tertiary regulation energy. In real-time, if necessary, the System Operator will activate the cheapest energy bids, which will receive the marginal price (*pay-as-cleared*). Tertiary regulation capacity is not remunerated [29] [11].

Secondary Regulation

The Secondary Regulation market - Automatic Restoration Reserve (aFRR) market under the European nomenclature - is developed in REE's Operational Procedure 7.2. In this market, qualified generators present bids for a single product to be bought by the System Operator: a regulation band (in MW) which should comply with the ratio 7/5, that is, for each 7 MW of upwards capacity, the band should comprise 5 MW of downwards capacity. Cleared bids receive the marginal price (*pay-as-cleared*). During real-time operation, secondary regulation energy is effectively used proportionally to the capacity assigned to each agent. Each MWh receives the marginal price of the tertiary regulation energy that would have been required if secondary regulation had not been activated. An example is shown in Figure 6 [11] [29].

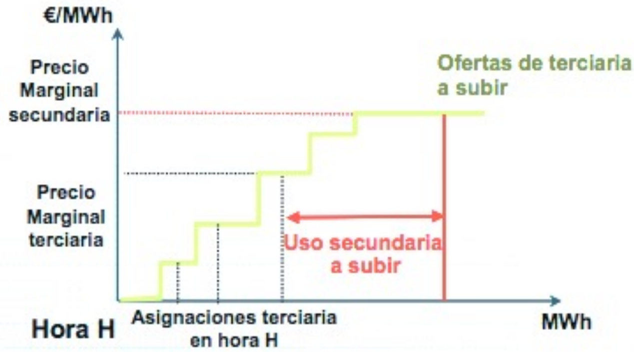

Figure 6 - Tertiary and secondary regulation energy pricing. Source: [25]

The cost of these services is used by the System Operator to compute the price of positive and negative imbalances in €/MWh to be paid or received by imbalanced agents in each hour. Figure 7 shows the evolution of positive and negative imbalance prices in Spain throughout 2017.

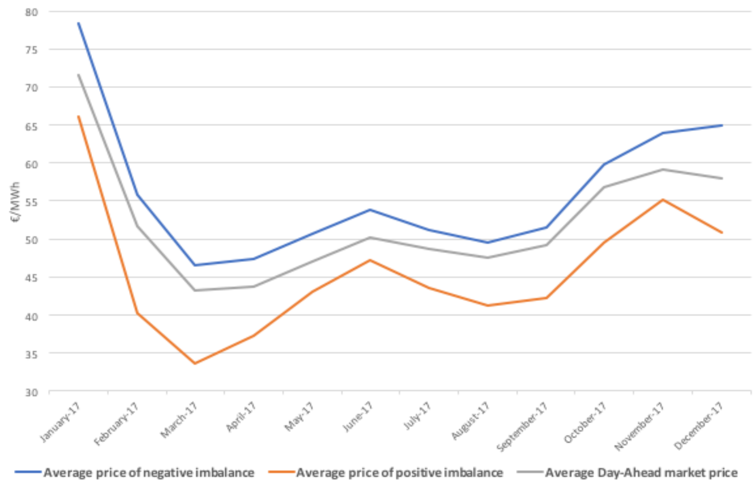

Figure 7 - Evolution of positive and negative imbalance prices, compared to Day-Ahead market price in 2017. Source: [13]

The average price of negative imbalance is always above the average day-ahead market price while the average price of positive imbalance is always below. This means that agents with a positive imbalance will always receive a lower or equal price for that energy than the one they would have received if they had sold that energy in the Day-Ahead market. Agents with a negative imbalance will always pay for the extra energy a higher or equal price than the one they would have paid in the Day-Ahead market. Therefore, under the dual pricing mechanism established in Spain, there is no incentive to deviate, even if that imbalance contributes to reducing the overall system's imbalance.

However, negative and positive imbalances do not imply the same cost, that is, the same difference with the day-ahead market. Figure 8 shows the average cost of both positive and negative imbalances in Spain throughout 2017, with respect to Day-Ahead market price in each hour. It can be seen that a positive imbalance tends to be more expensive than a negative imbalance. In the case of an industrial consumer, this means that it might have the incentive to under-buy in the market, in order to avoid consuming less than what was programmed and at the risk of creating a negative imbalance, which was significantly less expensive throughout 2017.

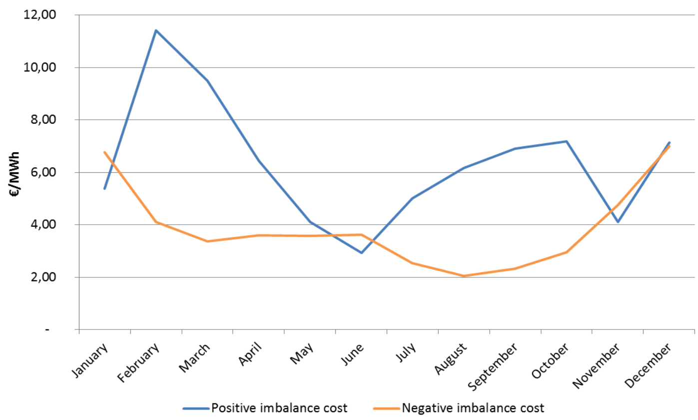

Figure 8 - Evolution of positive and negative unitary imbalance (€/MWh) in Spain throughout 2017. Source: [13]

c. Imbalance costs incurred by an industrial consumer

The industrial consumer under-study belongs to the electrometallurgy sector. Thus, electricity procurement is one of the major costs-components of the production process. Adapting its production to expected electricity prices is crucial in order to reduce this cost, which makes it a very price sensitive consumer. Figure 9 shows the high correlation existing between the industrial consumer's energy consumption and the average day-ahead market price by hour of the day during 2017.

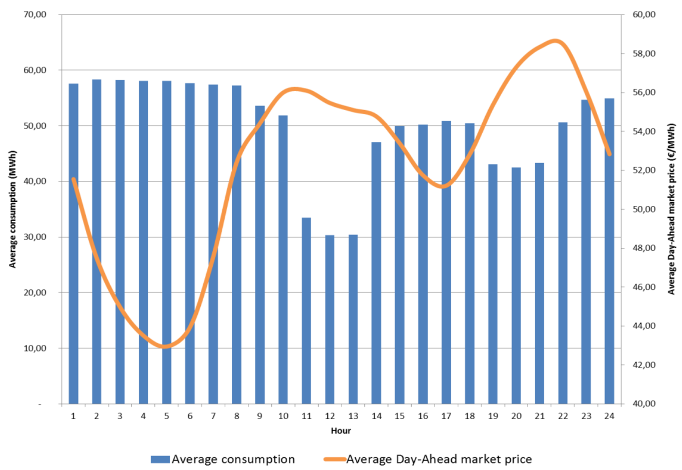

Figure 9 - Correlation between the industrial consumer's consumption and the average day-ahead market price in each hour, throughout 2017.

In case of changes in the production process or unexpected outages, correcting the market program in the next available market will allow the industrial consumer to avoid the extra-cost of imbalances. However, the intraday market mechanism existing in Spain, based on auctions, does not allow market agents to modify their program in the 3 or more hours before delivery time.

Table 5 shows every hourly product last trading time and the duration of the non-trading period between the last trading time and the delivery time, also known as lead-time. During the lead-time, the consumer may know it will deviate from the market program but it will not have the possibility of correcting it before delivery, generating an imbalance. It can be seen that consumption in some hours cannot be modified in the 5, 6 or even 7 hours before delivery time.

The new hybrid intraday market model proposed by OMIE, REE and REN reduces this lead time to 1 hour, increasing agent's capability of adjusting their programs in intraday markets in order to avoid costly imbalances.

Hour	Delivery time	Last trading time	Lead-time
1	00:00	21:45 D-1	2h15
2	01:00	21:45 D-1	3h15
3	02:00	21:45 D-1	4h15
4	03:00	21:45 D-1	5h15
5	04:00	01:45 D	2h15
6	05:00	01:45 D	3h15
7	06:00	01:45 D	4h15
8	07:00	04:45 D	2h15
9	08:00	04:45 D	3h15
10	09:00	04:45 D	4h15
11	10:00	04:45 D	5h15
12	11:00	08:45 D	2h15
13	12:00	08:45 D	3h15
14	13:00	08:45 D	4h15
15	14:00	08:45 D	5h15
16	15:00	12:45 D	2h15
17	16:00	12:45 D	3h15
18	17:00	12:45 D	4h15
19	18:00	12:45 D	5h15
20	19:00	12:45 D	6h15
21	20:00	12:45 D	7h15
22	21:00	18:45 D	2h15
23	22:00	18:45 D	3h15
24	23:00	18:45 D	4h15

Table 5 - Delivery time, last-trading time and non-trading period duration before delivery of each hourly product. Source: [10]

During 2017, both positive and negative imbalances generated an extra-cost for the industrial consumer under study. Energy bought at the day-ahead market price was not consumed in delivery time. Thus, the industrial consumer received a lower price than what was initially paid. On the other hand, energy consumed at delivery time was not bought at the day-ahead market. Thus, the industrial consumer paid for this energy a higher price than the day-ahead market price.

Figure 10 shows the estimated monthly extra-costs incurred by an electrometallurgy consumer due to positive and negative imbalances. It can be seen that extra-costs due to negative imbalances are higher, even though during 2017 negative imbalances were clearly less expensive than positive imbalances. This may be a consequence of the industrial consumer's strategy to reduce imbalance costs: in order to reduce the total cost, it is preferable to under-buy in the market in order to make sure that in case of deviating, the resulting imbalance will be negative and therefore, less expensive.

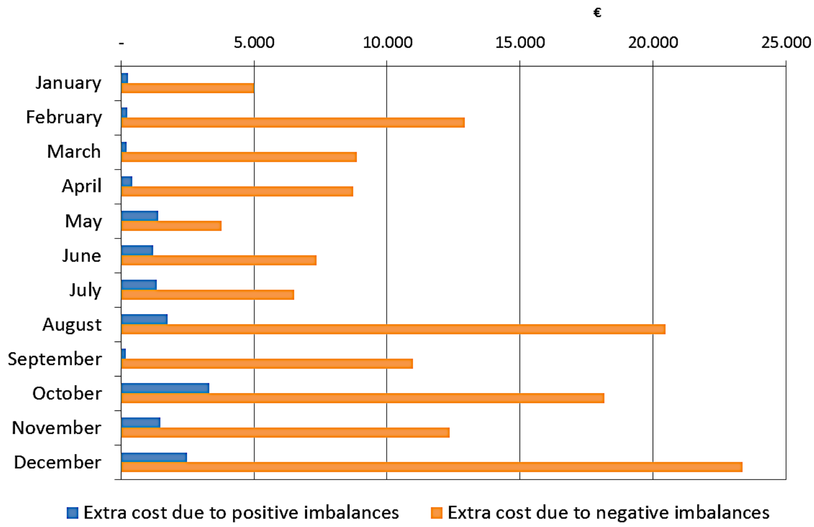

Figure 10 - Monthly extra-costs due to positive and negative imbalances throughout 2017.

Total imbalance extra-costs in 2017 are summarized on Table 6. In the case of an electrometallurgy industry, electricity costs can account from 30% to 50% of the total production cost, thus, an average value of 40% has been assumed. It can be seen that reducing the extra-costs associated to imbalances could yield a 0,65% reduction in the total electricity cost, and a 0,3% reduction in the total production cost, which could be used by the producer as a competitive advantage in the electrometallurgy market.

2017	MWh	€
Energy bought in DA/ID markets	456.687 MWh	23,4 M€
Total imbalance	-18.456 MWh	1,2 M€
Extra-cost associated	152.741 € **0,65 % of total electricity cost** **~ 0,3 % of total production cost**	

Table 6 - Imbalance extra-costs as compared to total energy costs in 2017.

However, the industrial consumer under study would only be able to reduce the extra-costs associated to imbalances in continuous intraday sessions if these allow buying electricity at a market price below the corresponding imbalance price. Expected volumes and prices in Spanish continuous intraday market sessions will be analyzed in Chapter 4.

d. Conclusions

The purpose of this chapter is to study the imbalance cost-allocation and pricing mechanisms in Spain, in order to quantify the extra-cost incurred by an industrial consumer due to deviations from its expected consumption program. It can be concluded that:

- Imbalances are defined as the difference between the production or consumption measured at power station busbars and the final production or consumption program scheduled in the Day-Ahead and Intraday Markets. Imbalances can be negative (or downwards) or positive (or upwards).

- In Spain, imbalance pricing is based on a dual-pricing mechanism, which aims at reducing every incentive to deviate, as it is never a business opportunity: agents with a positive/negative imbalance will always receive/pay a lower/higher or equal price for that energy than the one they would have received/paid if they had sold that energy in the Day-Ahead market.

- The new hybrid intraday market model proposed by OMIE, REE and REN reduces lead time before delivery to 1 hour, increasing the chances of adjusting agents' programs and reducing imbalances.

- Reducing the extra-costs associated to imbalances could yield a competitive advantage in the ferroalloy market, where competition is highly fragmented.

- Consumers will only be able to reduce the extra-costs associated to imbalances in continuous intraday sessions if these allow buying electricity at a market price below the corresponding imbalance price, which will depend on session's liquidity and price convergence.

Chapter 4

Economic Analysis of a Hybrid Intraday Electricity Market

a. Introduction

Market liquidity has been defined as the easiness of trading an asset, without any transaction having an impact on the asset's value. Even though it is not directly observable, this easiness of trade seems to be correlated with the number of market agents and the number of trades. Thus, the volumes traded in a market are a typical indicator for liquidity. Market liquidity can really impact market outcomes: without enough liquidity in the market, agents may fear that their transactions move the market price, making him pay a worse price than the unaffected market price.

As explained in [30], there are different types of costs associated to illiquidity in a market: bid-ask spread, market impact costs and delay and search costs. The bid-ask spread would be the difference between the best bid price and the best offer price in a market: due to this spread, market agents have to pay an extra cost in each transaction. Market-impact costs occur when a market agent drives the price up or down when buying or selling, respectively. Finally, delay and search costs are incurred when a market agent finds difficulties in closing a transaction or delays it in order to obtain a better price, due to a lack of liquidity in the market.

As explained in Chapter 2, continuous markets tend to present lower liquidity levels than auctions do. If continuous intraday market sessions under the hybrid intraday model in Spain happen to be illiquid, the extra-costs associated to this illiquidity will increase the risk associated to trading and could noticeably worsen the price that can be obtained in the intraday market.

In this section, Spanish intraday current market outcomes will be analyzed, in order to provide a clear starting point from the subsequent analysis. Afterwards, in order to evaluate the impact of low liquidity in the Iberian hybrid intraday market, a forecasting model will be developed in order to predict hourly average continuous intraday market prices, based on the data gathered since its launch on June 13th 2018. This model will be used to generate hourly average continuous intraday market prices for 2017, in order to estimate the potential cost reduction that the industrial consumer under study would have been able to obtain by correcting its program in the intraday continuous market.

Two forecasting approaches will be compared: linear regression (LR) and multilayer perceptron (MLP). The analysis will be carried out using RStudio, a free and open-source development environment for R, a programming language for statistical analysis.

b. Spanish intraday market outcomes: intraday prices and volumes.

In order to analyze price outcomes from intraday market sessions, the day-ahead premium has been computed, that is, the difference between the hourly intraday price and the hourly day-ahead price, considering it positive when intraday market prices are lower than day-ahead prices and negative when intraday market prices are higher than day-ahead prices. Results are shown in Figure 11 and 14.

		ID1	ID2	ID3	ID4	ID5	ID6	ID7
2016	Premium	0,27	-0,01	-0,05	-0,11	-0,12	-0,05	1,03
	Standard deviation	2,19	2,20	2,30	2,00	2,44	3,10	4,30
2017	Premium	0,21	0,09	0,03	0,02	-0,02	-0,65	0,33
	Standard deviation	2,37	2,17	2,04	1,79	2,40	2,95	3,90

Figure 11 – Day-ahead premium and standard deviation of each intraday auction in 2016 and 2017. Source: [13]

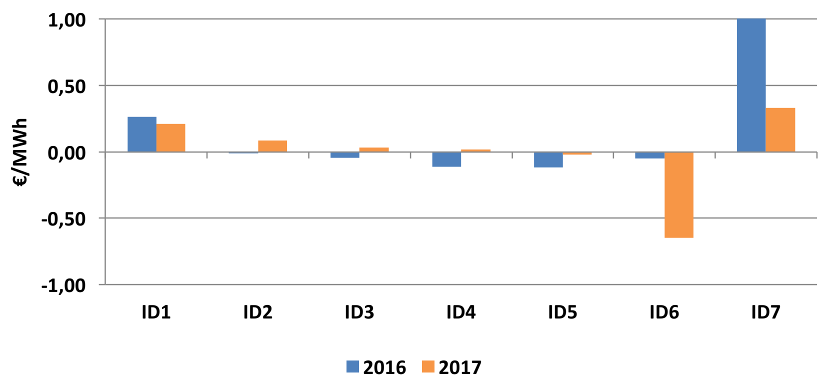

Figure 12- Day-ahead premium of each intraday auction in 2016 and 2017. Source: [13]

Figure 12 shows that, in the last 2 years, intraday sessions 1 and 7 have been less expensive that those same hours in the day-ahead market. Sessions 2,3,4 and 6 have not shown such a clear behavior and seem to show a better price convergence, except for intraday session 6 in 2017, which was clearly more expensive. Overall, this shows that intraday market prices are approximately the same or even better than day-ahead prices for an industrial consumer. Nevertheless, in order for an industrial consumer to actually have an incentive to correct its program in an intraday market session, its price should be more attractive than the imbalance price. In this case, day-ahead premium represents the gains that the industrial consumer would get from not deviating at all from its day-ahead program.

However, in order to analyze the incentive that an agent has to correct its program in an intraday market, instead of deviating, it is more interesting to look at the difference between the imbalance price and the price of the last intraday auction where a certain hourly product was traded. Figure 13 shows the extra-cost associated to deviating from the program, instead of correcting it in the last intraday session available.

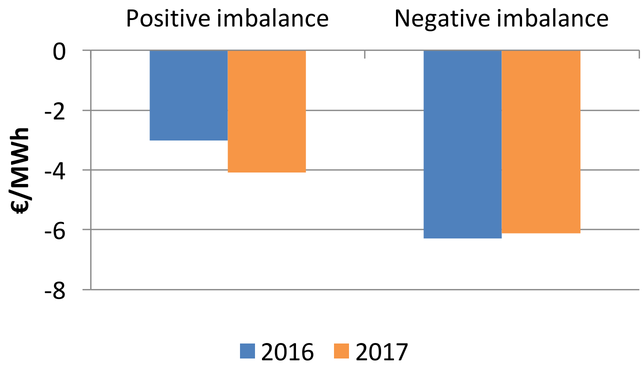

Figure 13 – Extra-cost of deviating from the program instead of correcting it in the last intraday session. Source: [13]

It can be seen that, on average, it is less costly for an agent to modify its program in an intraday market auction than to deviate. This premise may or may not be true in the case of intraday continuous market, depending on the price difference between the price at which the deal was closed in the continuous intraday market and the imbalance price. In order to estimate this difference, the average intraday continuous market price will be predicted.

In 2017, 245.650 GWh were traded in the Day-Ahead market, and 35.601 GWh in the intraday market, which represents a 12,6% over the total energy traded [11]. Figure 14 shows hourly volumes traded by hour and intraday auction in Spain, as a percentage of the volume traded in the day-ahead market, throughout 2017.

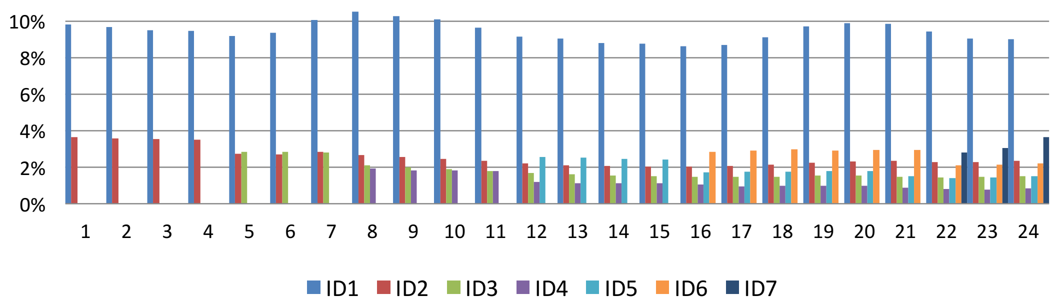

Figure 14 – Hourly volumes traded by hour and intraday auction, as a percentage of the volume traded in the day-ahead market in 2017. Source: [13]

It is observed that the first intraday auction is the most liquid one, as technical constraints solving procedures may require agents providing the service to vary their programs in the first intraday auction. For example, the provision of additional upwards reserve market is only done by thermal generators who have not been cleared in the Day-Ahead market. If they are required to provide additional upwards reserve, they will need to sell their minimum technical output in the first intraday market. In Figure 15, it is shown how traded volumes increase in the different market sessions when additional upwards reserve is required. Volumes in intraday session 1 increase by more than 66%, while other sessions' liquidity is not very much affected. Figure 14 also shows that those hourly periods that will not be negotiated in the next intraday auction are more liquid, in all intraday auctions.

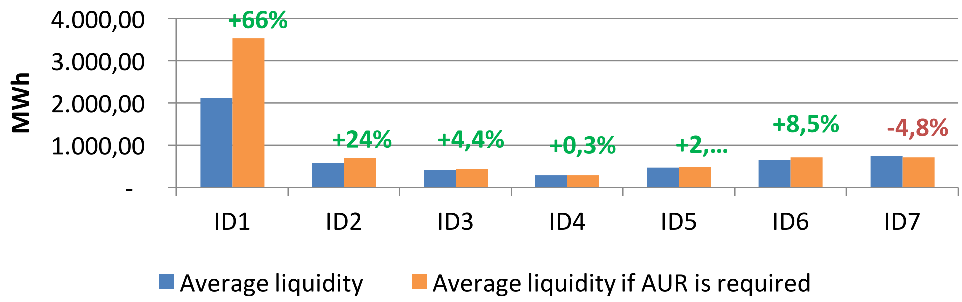

Figure 15 – Variations in intraday auctions' liquidity when Additional Upwards Reserved is required by the System Operator. Source: [13]

Figure 16 shows the hourly average volume traded in the continuous intraday market from June 13th to June 25st 2018, compared to auction intraday market volumes in the same period, as percentage of the total volume traded. It can be observed that traded volumes in the intraday continuous market are small, but in line with intraday market sessions 3, 4 and 5. Figure 17 shows the average hourly imbalance that the industrial consumer under study would have to trade in the continuous intraday market.

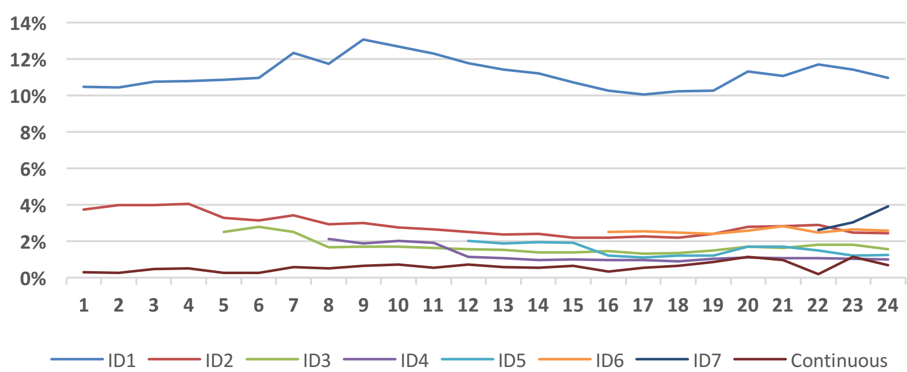

Figure 16 - Hourly average traded volume in each intraday market as a % of day-ahead volume (13-25th June). [13]

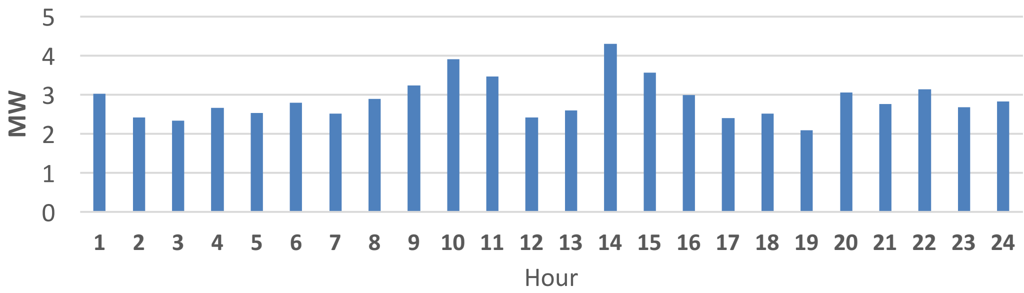

Figure 17 – Hourly average absolute imbalance of the industrial consumer under study

The average absolute imbalance of the industrial consumer under study over 2017 was 5,88 MW, with a standard deviation of 2,88 MW. It can be seen that average absolute deviations of the industrial consumer under study are relatively small, and it could be reasonable to suppose that it will not be difficult to buy those volumes in the continuous intraday market. Unusual deviations will be observed when estimating the potential cost reduction that the industrial consumer under study would have been able to obtain by correcting its program in the intraday continuous market.

c. Expected prices in the Hybrid Intraday Electricity Market: Forecasting Models.

In this section, two forecasting approaches will be used to predict continuous intraday market prices: linear regression (LR) and multilayer perceptron (MLP). The analysis will be carried out using RStudio, a free and open-source development environment for R, a programming language for statistical analysis.

First, an exploratory analysis of the data to be used to predict intraday continuous prices will be carried out, and a proper metric to evaluate model's performance will be chosen. Afterwards, both LR and MLP models will be fit, and their performance assessed. Then, the best model will be chosen in order to estimate the continuous intraday market prices that would have taken place during 2017, in order to compute the potential cost reduction that the industrial consumer under study would have been able to obtain by correcting its program in the intraday continuous market.

Data exploratory analysis

A database including all variables that may have an impact on continuous intraday prices has been built, including data from the launch of the intraday continuous market on June 13th 2018, until June 30th 2018. In order to do so, ESIOS, the information platform of Spanish System Operator REE was very useful. The initial input variable set included in the model and their reference names are described hereinafter:

- Hourly day-ahead price (€/MWh) - PMD
- Auction intraday average hourly price: average of the price of an hourly product in each of the intraday markets in which it has been traded (€/MWh) - PMIS
- Hourly upwards imbalance price (€/MWh) - PDS
- Hourly downwards imbalance price (€/MWh) - PDB
- Hourly demand (GWh) - DEM
- Hourly wind production (GWh) – PRODEOL
- Hourly CCGT production (GWh) – PRODCCGT
- Hourly coal production (GWh) – PRODCARBON
- Hourly solar production (GWh) – PRODSOLAR
- Hourly hydro production (GW)h - PRODHIDRO
- Hourly energy traded in the day-ahead market (GWh) – VOLMID
- Hourly energy traded in the auction intraday market (GWh) – VOLMIS
- Continuous intraday average hourly price (€/MWh) – PMIC

Afterwards, the initial dataset has been explored in order to see whether there were missing values, outliers or strong correlations which could provide misleading information. NA values have been removed from the initial data set and boxplots and correlation graphs for all variables have been run. Results are shown in Figure 18 and Figure 19.

In Figure 18, no point is clearly far enough to be seen as an outlier. Therefore, no data was removed. Relevant information may be extracted from Figure 19: the output, PMIC, may have some linear relationship with PMD, PMIS and DEM, as can be extracted from the scatter diagrams. These variables also show high correlation levels with the output, as well as PRODCARBON AND PRODCCGT. No variables present a particularly small correlation value with the output variables.

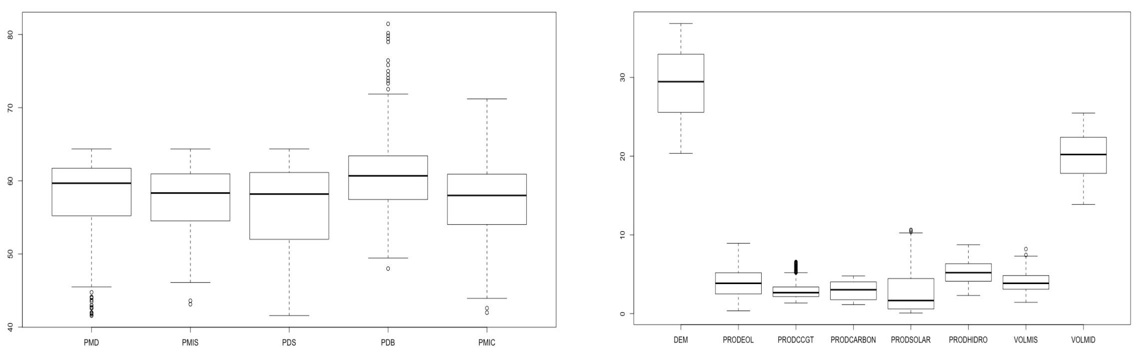

Figure 18 - Box plot of all variables in the initial set

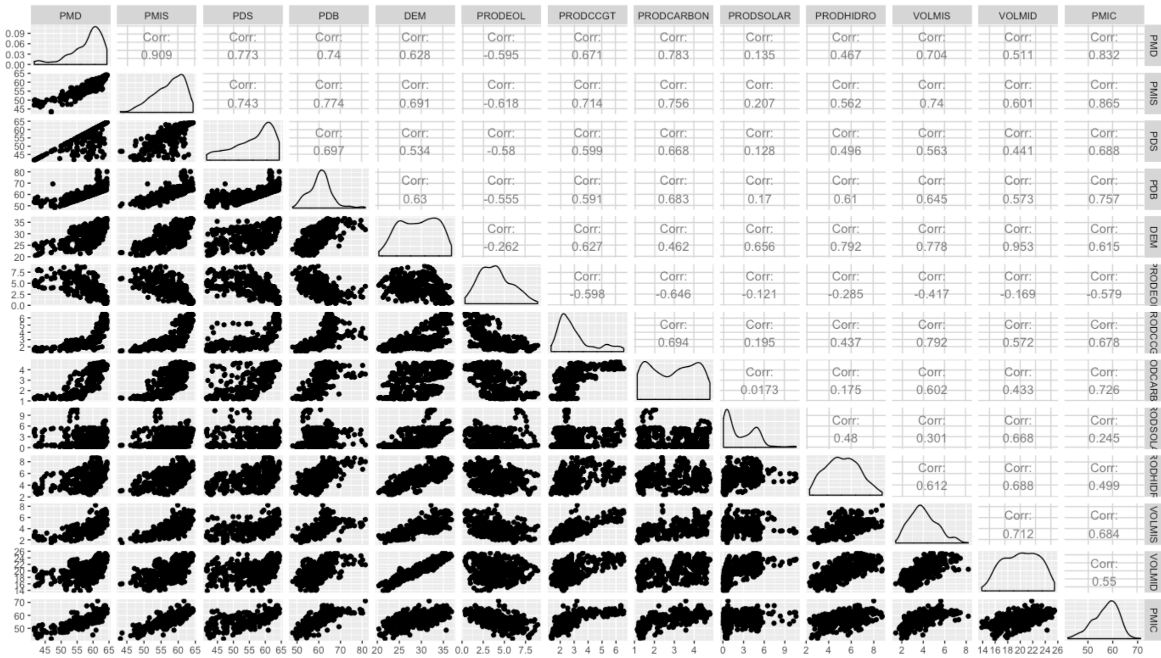

Figure 19 - Correlation plot between all variables in the initial dataset

Before any model was set, the input dataset has been split randomly into two smaller sets: the training set, including 80% of the data, and the validation set, including 20% of the data. The R script used for the exploratory analysis is shown in Figure 20. All R libraries required were installed beforehand.

```
#Load data from txt files
fdata =read.table("TRAINING13a30.txt",header=TRUE, fill=TRUE) #Input data
fdata = na.omit(fdata)  #Eliminate NA

#Exploratory analysis
ggpairs(fdata,columns=2:14, cardinality_threshold = 150)
boxplot(fdata[,c(2,3,4,5,14)])
boxplot(fdata[,c(6,7,8,9,10,11,12,13)])

#Divide the data into training and validation sets
set.seed(150) #For replication
ratioTR = 0.8 #Percentage for training: 80%
trainIndex <- createDataPartition(
                              fdata$PMIC,     #output variable.
                               p = ratioTR,    #split probability
                               list = FALSE,   #do not show output as list
                               times = 1)      #only one partition

fdata_train = fdata[trainIndex,] #Training set
fdata_val = fdata[-trainIndex,] #Validation set
fdataTR_eval =fdata_train #Set equal to validation set, to evaluate model
fdataTV_eval =fdata_val #Set equal to validation set, to evaluate model
```

Figure 20 - R script used for exploratory analysis of the initial dataset

Assessing the accuracy of predictive models

Being able to correctly assess the accuracy of predictive model's performance is critical in order to be able to evaluate if output information can be trusted and used for decision-making. Predictive model's performance should be measured based on the difference between real, observed values and predicted values [31]. The most common measures used for evaluating model's accuracy and performance are described hereinafter [32]:

- Mean Absolute Error: measures the average magnitude of the residuals, that is, the difference between the observed values and the predicted values. It is a positive value, measured on the same units as the data, and smaller values mean that the model is better fit [33].

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y[i] - \hat{y}[i]|$$

Equation 1 - Mean Absolute Error (MAE) calculation

- Root Mean Squared Error: measures the square root of the average of squared residuals. It is a positive value, measured on the same units as the data, and smaller values mean that the model is better fit. As the errors are squared before they are averaged, RMSE penalizes especially large errors [33].

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y[i] - \hat{y}[i])^2}$$

Equation 2 - Root Mean Square Error (RMSE) calculation

- Coefficient of determination: measures the proportion of variance that is explained by the model. It can take values from 0 to 1, and higher values mean that the model is better fit. One disadvantage of the coefficient of determination is that there are no universal acceptable values for it, but instead, what is considered a low or high value depends very much on the research area [34].

$$R_i^2 = 1 - \frac{\sum_{i=1}^{N}(y[i] - \hat{y}[i])^2}{\sum_{i=1}^{N}(y[i] - \bar{y})^2}$$

Equation 3 - Coefficient of determination (R-squared) calculation

According to these definitions, the main metric chosen for model comparison in this case was RMSE: it penalizes large errors and it is in the same units as the output variable, that is, in €/MWh. Therefore, using RMSE makes it easier to assess when the model performance is acceptable for predicting continuous intraday prices.

Linear regression with backwards feature selection and cross-validation

The objective of multiple linear regression techniques is to model the behavior of a dependent variable as a linear function of a given set of independent variables and a random error term, which represents the behavior of the dependent variable not explained by the regression function. Equation 4 shows the general form of a multiple linear regression model with n independent variables. β coefficients measure the marginal contribution of each input variable x_n to the output variable y [35].

$$y_t = \beta_0 + \beta_1 \cdot x_{1t} + \cdots + \beta_n \cdot x_{nt} + \varepsilon_t$$
$$\hat{y}_t = \beta_0 + \beta_1 \cdot x_{1t} + \cdots + \beta_n \cdot x_{nt}$$

Equation 4 - General form of a multiple regression model

The multiple linear regression method is based on a series of assumptions [36]:

- The expected value of the output variable is a linear function of the n input variables, which implies:
  - Linearity: if an input variable x_n changes an amount Δx, the output variable y should change proportionally to β_n
  - Additivity: the total effect of the input variables n_n in the output variable y is equal to the sum of their separate effects.
- The random terms ε_t are independent random variables.
- The random terms ε_t are independent are homoscedastic (they all have the same variance) and are normally distributed.

If input variables are not independent but instead they show a strong correlation among them, estimation of the model coefficients can be arbitrary and behave badly with new data. This effect is called multicollinearity and can be detected with the Variance Inflation Factor (VIF), as described in Equation 5. It is computed from the coefficient of determination R_i^2, which measures the proportion of the variance of the output variable y that is explained by the input variables [32].

$$VIF(\beta_i) = \frac{1}{(1 - R_i^2)}$$

Equation 5 - Variable Inflation Factor (VIF)

Also, not all input variables may be useful in order to predict the output variable. In order to find a subset of the original input variables that does a good job of predicting the dependent variable, subset selection techniques may be used. In this case, linear regression with backward feature selection was applied by using the recursive feature elimination technique through the rfe() function available in *R*'s library *caret*. First, the algorithm fits the model to all variables, and ranks them according to their importance to the model. Then, iteratively, the least useful variable is removed, obtaining the optimal number of input variables to be used [32].

As the training sample is relatively small, a technique called *k-fold cross-validation* is used to avoid losing important information. This method consists on randomly splitting the original data set in k equal sized subsets. Of the k subsets, one is retained as the validation set and the remaining k-1 are used as training sets. The cross-validation process is then repeated k times, with each of the k datasets used exactly once as the validation data. The k results can then be averaged to produce a single prediction [37].

Model

The R code used to fit a linear regression model with recursive feature elimination is shown in Figure 21. The metric that will be minimized is RMSE.

```
ctrl_none <- trainControl(method = "none",
       summaryFunction = defaultSummary,
       returnResamp = "final",
       savePredictions = TRUE)

## Specifies the cross validation method used for selecting the
 optimum number of variables
ctrl_rfe <- rfeControl(method = "cv",
                       number = 10,
                       verbose = TRUE,
                       functions = caretFuncs)

set.seed(150)
subsets <- 1:20 #The number of features that should be retained
lm.RFE <- rfe(form = PMIC~.,
              data = fdata_train[,c(2:14)],
              method = "lm",
              preProcess = c("center","scale"),
              trControl = ctrl_none,          # Arguments for rfe
              sizes = subsets,
              metric = "RMSE",
              rfeControl = ctrl_rfe)

lm.RFE
ggplot(lm.RFE,metric = "RMSE")
lm.RFE$fit
summary(lm.RFE$fit)
lm.RFE$fit$finalModel

vif(lm.RFE$fit$finalModel)
```

Figure 21 - R script for linear regression with recursive feature elimination

Information given by the fitted model is shown in Figure 22. According to the information provided by the trained model, RMSE is minimized with 9 out of 12 input variables, removing the following predictors: PDS, VOLMIS and VOLMID. The trained model explains 81,9% of the variance (Rsquared) with an error of 2,153 (RMSE).

```
Recursive feature selection
Outer resampling method: Cross-Validated (10 fold)
Resampling performance over subset size:
 Variables  RMSE Rsquared   MAE RMSESD RsquaredSD  MAESD Selected
         1 3.449   0.5065 2.724 1.1349    0.32776 1.1108
         2 2.442   0.7688 1.773 0.3626    0.06985 0.2558
         3 2.436   0.7708 1.769 0.3558    0.06992 0.2583
         4 2.422   0.7752 1.751 0.3184    0.05925 0.2302
         5 2.337   0.7931 1.660 0.3082    0.04681 0.1773
         6 2.327   0.7927 1.651 0.3333    0.05458 0.2004
         7 2.269   0.8033 1.627 0.2990    0.05622 0.1727
         8 2.241   0.8064 1.619 0.2970    0.05476 0.1591
         9 2.153   0.8190 1.586 0.3391    0.05946 0.1864        *
        10 2.186   0.8147 1.609 0.3212    0.06050 0.1822
        11 2.180   0.8158 1.603 0.3239    0.05988 0.1876
        12 2.171   0.8170 1.592 0.3186    0.06036 0.1857
The top 5 variables (out of 9):
   PMIS, PRODSOLAR, DEM, PDB, PMD

Residuals:
     Min       1Q   Median       3Q      Max
-10.3464  -1.3212   0.1622   1.0959   9.2841

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  57.0940     0.1161 491.571  < 2e-16 ***
PMIS          2.6113     0.3351   7.792 8.73e-14 ***
PRODSOLAR     1.7310     0.2386   7.255 2.91e-12 ***
DEM          -2.9972     0.4704  -6.372 6.29e-10 ***
PDB           1.1640     0.2290   5.083 6.26e-07 ***
PMD           1.3191     0.3032   4.350 1.82e-05 ***
PRODEOL       1.0006     0.2347   4.263 2.63e-05 ***
PRODCARBON    1.0799     0.2983   3.621  0.00034 ***
PRODHIDRO     0.9201     0.3049   3.018  0.00275 **
PRODCCGT      1.0052     0.2162   4.648 4.85e-06 ***
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Residual standard error: 2.138 on 329 degrees of freedom
Multiple R-squared:  0.8282, Adjusted R-squared:  0.8235
F-statistic: 176.2 on 9 and 329 DF, p-value: < 2.2e-16

VIF (Variance Inflation Factor)
      PMIS  PRODSOLAR        DEM        PDB        PMD    PRODEOL PRODCARBON
  8.300981   4.207702  16.352611   3.876506   6.796712   4.070918   6.575026
 PRODHIDRO   PRODCCGT
  6.870381   3.456220
```

Figure 22 - Linear regression with recursive feature elimination: information given by the model using 12 predictors.

However, predictor DEM shows a variance inflation factor over 10, showing multicollinearity (high correlation with other input variables), which could also be seen in the correlation plot in Figure 19. When multicollinearity is present, estimation of the model coefficients can be arbitrary and behave badly with new data. Therefore, a new model was run removing predictor DEM. Information given by the model is shown in Figure 23.

```
Recursive feature selection
Outer resampling method: Cross-Validated (10 fold)
Resampling performance over subset size:
 Variables  RMSE Rsquared   MAE RMSESD RsquaredSD  MAESD Selected
         1 2.439   0.7709 1.756 0.4043    0.07095 0.3070
         2 2.358   0.7908 1.688 0.3302    0.05433 0.2453
         3 2.366   0.7875 1.716 0.2823    0.05421 0.2050
         4 2.356   0.7894 1.702 0.2723    0.05447 0.1847
         5 2.337   0.7912 1.695 0.3067    0.05689 0.2110
         6 2.318   0.7942 1.684 0.3122    0.05768 0.2289
         7 2.290   0.7987 1.658 0.3276    0.05740 0.2238
         8 2.234   0.8066 1.623 0.3484    0.05993 0.2107        *
         9 2.246   0.8052 1.628 0.3406    0.05964 0.2058
        10 2.251   0.8040 1.632 0.3402    0.06033 0.2071
        11 2.254   0.8037 1.635 0.3408    0.06023 0.2081
The top 5 variables (out of 8):
   PMIS, PDB, PRODSOLAR, VOLMID, PRODCCGT

Residuals:
     Min       1Q   Median       3Q      Max
-10.4833  -1.1660   0.0599   1.0736   9.3171

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  57.0940     0.1207 473.177  < 2e-16 ***
PMIS          2.5643     0.3479   7.370 1.38e-12 ***
PDB           1.2932     0.2154   6.004 5.07e-09 ***
PRODSOLAR     1.1264     0.2281   4.938 1.26e-06 ***
VOLMID       -1.2653     0.3162  -4.002 7.76e-05 ***
PRODCCGT      0.6620     0.2165   3.057  0.00242 **
PMD           0.7490     0.3136   2.389  0.01747 *
PRODCARBON    0.7450     0.2359   3.158  0.00173 **
PRODEOL       0.5154     0.2273   2.268  0.02400 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Residual standard error: 2.222 on 330 degrees of freedom
Multiple R-squared:  0.814,   Adjusted R-squared:  0.8095
F-statistic: 180.5 on 8 and 330 DF,  p-value: < 2.2e-16

> vif(lm.RFE$fit$finalModel)
      PMIS        PDB  PRODSOLAR     VOLMID   PRODCCGT        PMD PRODCARBON
  8.289797   3.176778   3.562912   6.846035   3.211006   6.733121   3.810525
   PRODEOL
  3.537871
```

Figure 23 - Linear regression with recursive feature elimination: information given by the model using 11 predictors.

According to the information provided by the trained model, RMSE is minimized with 8 out of 12 input variables, removing the following predictors: PDS, VOLMIS and VOLMID. The trained model explains 80,66% of the variance (Rsquared) with an error of 2,234 (RMSE). According to the variable inflation factors obtained, the remaining variables show no significant correlations (VIF<10) and multicollinearity can be discarded, although new error is slightly higher than in the previous model considering all predictors.

Artificial neural network: multilayer perceptron (MLP) with cross-validation

Artificial neural networks are based on the functioning of the human brain: it imitates the human brain structure, making use of a large number of highly interconnected processing elements, called neurons. Artificial neural networks (ANNs) are able of recognizing patterns and relationships in the data, learning from examples and generalizing that knowledge to new data [38]. ANNs are structured in a series of layers, as shown in Figure 24.

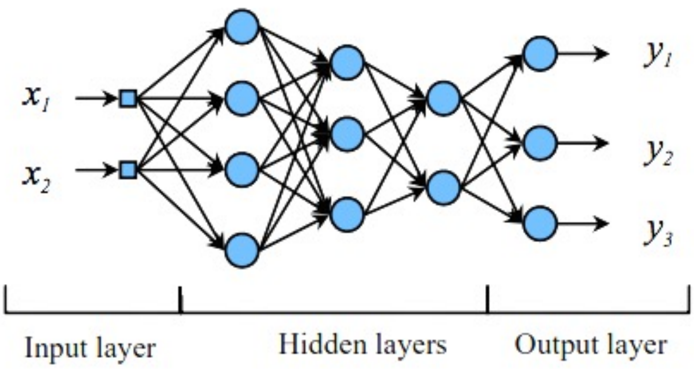

Figure 24 - General form of a neural network. Source: [32]

Typically, ANNs are composed of three types of layers of neurons: input layers, hidden layers and output layers. Each neuron is characterized by its weight, bias and activation function. First, input data is fed into the input layer neurons, which do a linear transformation of the input data using the synaptic weights, and a non-linear transformation using the activation function, generating an activation signal which progressively moves from the input layers, through the hidden layers up to the output layer [39]. This process is schematically shown in Figure 25 and Equation 6.

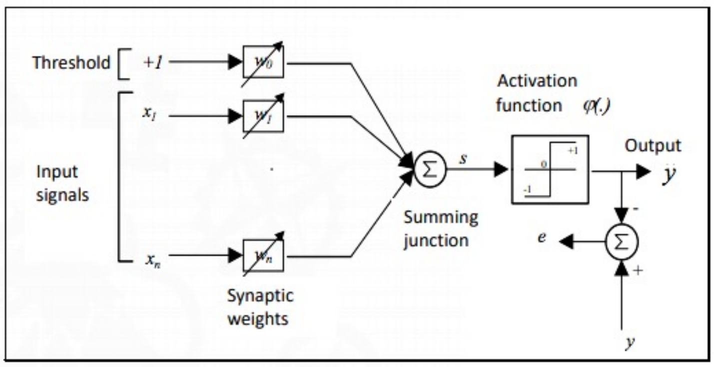

Figure 25 - Scheme of information flow in a neuron. Source: [32]

$$\hat{y} = \varphi\left(w_o + \sum_{i=1}^{n} w_i x_i\right)$$

Equation 6 - Synaptic weights, intercept and activation function. Source: [32]

The activation function may be linear, where the output is proportional to the weighed input; a threshold, where the output is set at one of two levels, depending on whether the total input is greater than or less than a threshold value; or non-linear functions, where the output varies continuously, but not linearly, as the input changes [39].

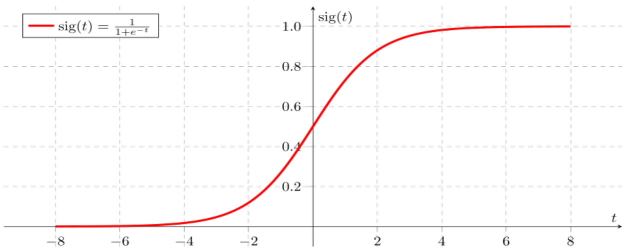

Figure 26 - Sigmoid function: logistic function. Source: [37]

Neural network architecture may be feed-forward or feedback. If activation signals are only allowed to travel from input from output, and thus, the output of a layer does not affect that same layers, the network is said to be a feed-forward network. If signals can travel through the network both from input to outputs and from outputs to inputs, in loops, then network is said to be a feedback network. Feedback networks can be extremely complicated [39]. Also, the learning process of a neural network can be classified as supervised learning or unsupervised learning. In supervised learning, the neural network infers information about the relationship between inputs and outputs from a training dataset, which includes input data and output values, and applies that knowledge to new examples. In unsupervised learning, input data without labeled responses is provided to the model, which tries to learn about the inferences and structure of the provided data [40].

Multilayer perceptron neural networks (MLP) are supervised learning, feedforward neural networks consisting of at least three layers of nodes. MLP networks are fully connected, that is, each node in one layer connects, with a certain weight, to every node in the following layer. MLPs use a non-linear activation function, which typically is a sigmoid function: a hyperbolic tangent, which ranges from -1 to 1, or a logistic function, which ranges from 0 to 1 [41]. The latter is depicted in Figure 26. The architecture of a multilayer perceptron neural network is shown in Figure 27.

To train the neural network, the weights must be adjusted so that the error between the predicted output and the observed values is reduced. The network must compute how the error changes as each weight is slightly increased or decreased, which is called the error derivative of the weights (EW). To compute the EW, MLP uses a supervised learning algorithm called backpropagation.

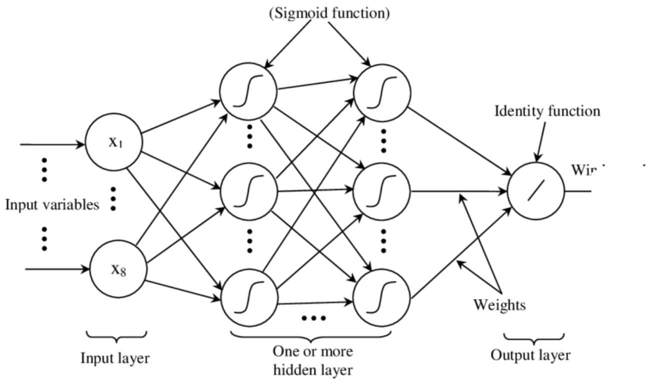

Figure 27 - Architecture of a multilayer perceptron neural network. Source: [41]

In order to fit the model, the train() function available in R's library caret has been used. The method used is *nnet*, which fits a single-hidden-layer neural network, with a sigmoid activation function in the hidden layer. The tuning parameters used in order to optimize the model are *size* and *decay*. *Size* is the number of neurons in the hidden layer, and *decay* is a parameter that that limits the growth of the weights in the training process.

Model

Figure 28 shows the R code used to fit an MLP neural network with all variables as input, using a linear activation function in the output, centering and scaling variables and minimizing RMSE.

```
ctrl_tune <- trainControl(method = "cv",           #K-fold with 10 folds
                          number = 10,
                          summaryFunction = defaultSummary,
                          returnResamp = "final",
                          savePredictions = TRUE)

set.seed(150)
mlp.fit = train(form = PMIC~.,
                data = fdata_train[,c(2:14)],
                method = "nnet",
                linout = TRUE,
                maxit = 400,
                tuneGrid = expand.grid(size=70, decay=2.227778),
                preProcess = c("center","scale"),
                trControl = ctrl_tune,
                metric = "RMSE")

mlp.fit #information about the resampling settings
ggplot(mlp.fit)
SensAnalysisMLP(mlp.fit) #Statistical sensitivity analysis
```

Figure 28 – R script for MLP training with cross validation

The model was provided several arrays of values for size and decay, until optimal size (70) and decay (2.22778) values were found. To do so, graphs relating RMSE with size and decay values were used, as can be seen in Figure 29.

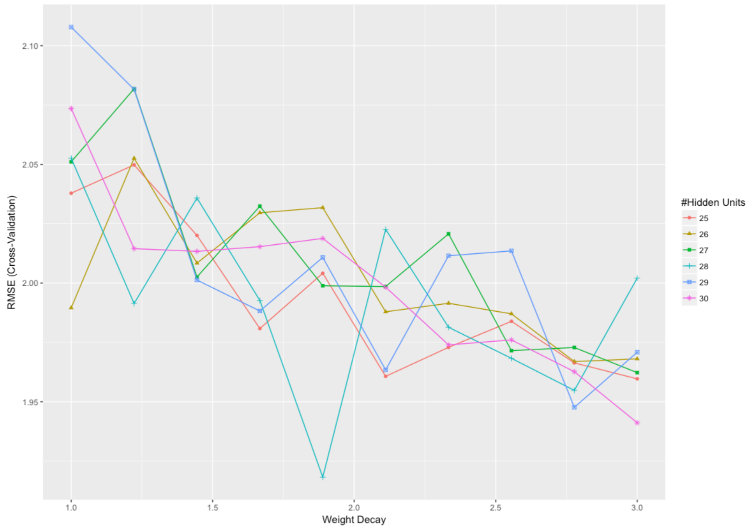

Figure 29 - RMSE relationship with size and decay in MLP models

The trained model explains 84.96% of the variance (Rsquared) with an error of 1,9426 (RMSE). A sensitivity analysis is then performed to analyze which variables strongly affect the output and which ones can be pruned. Results are shown in Figure 30.

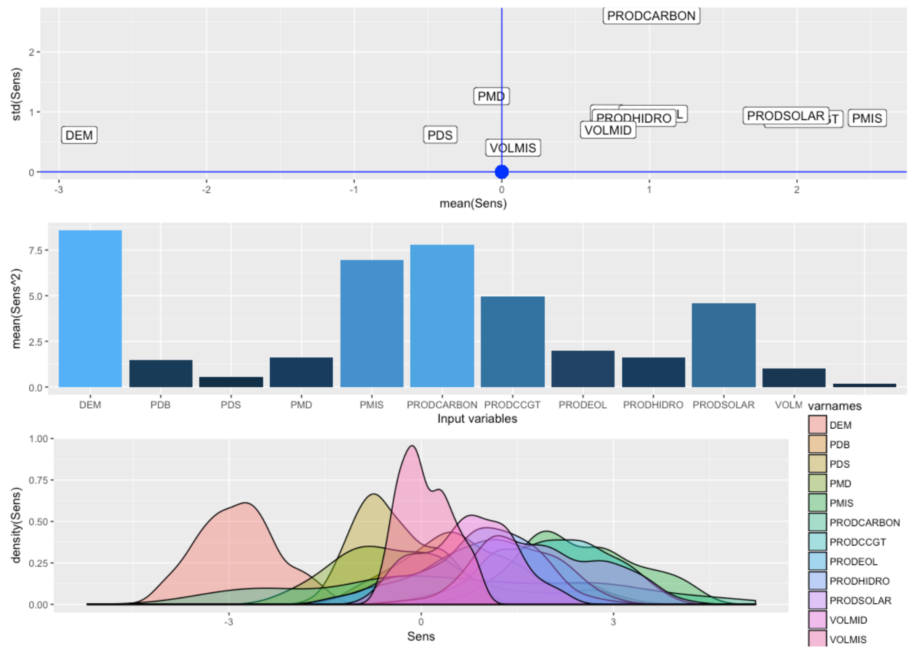

Figure 30 - Sensitivity analysis of MLP with all predictors (mlp)

Based on the results shown in Figure 30, variable VOLMIS is removed, as its sensibility's mean and standard deviation are very close to zero. Optimal size (47) and decay (1.93333) values are found iteratively. Results are improved, as the trained model explains 85,8% of the variance (Rsquared) with an error of 1,8944 (RMSE). A new sensitivity analysis is performed to analyze which variables strongly affect the output and which ones can be pruned. Results are shown in Figure 31.

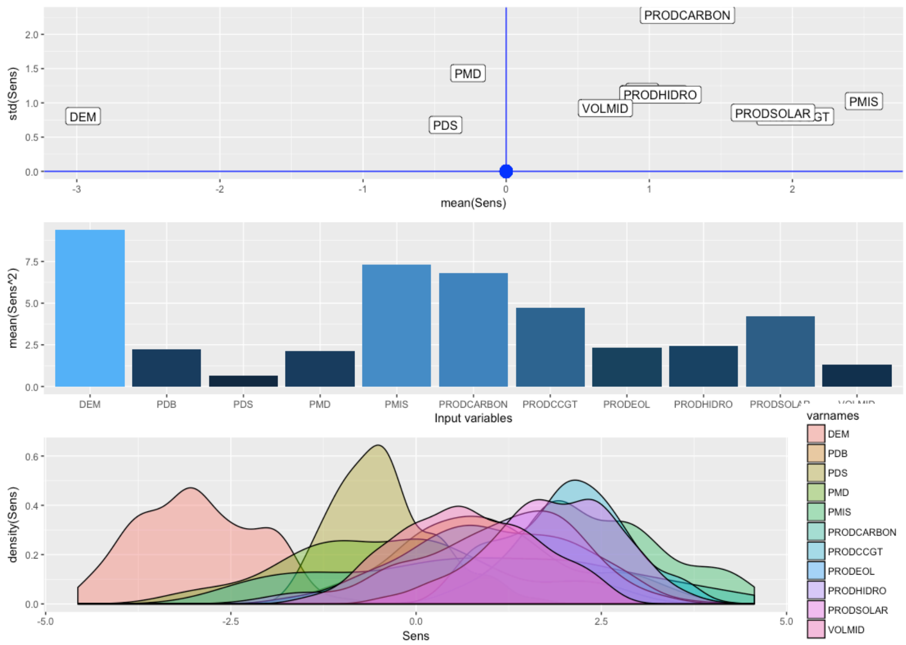

Figure 31 - Sensitivity analysis of MLP, without predictor VOLMIS (mlp2)

Based on the results shown in Figure 31, variable PDS is also removed, as its sensibility's mean and standard deviation are close to zero, and smaller compared with other predictors' sensibilities. Optimal size (30) and decay (2.3333) values are found iteratively. Results are improved, as the trained model explains 86,57% of the variance (Rsquared) with an error of 1,847 (RMSE). A new sensitivity analysis is performed to analyze which variables strongly affect the output and which ones can be pruned. Results are shown in Figure 32.

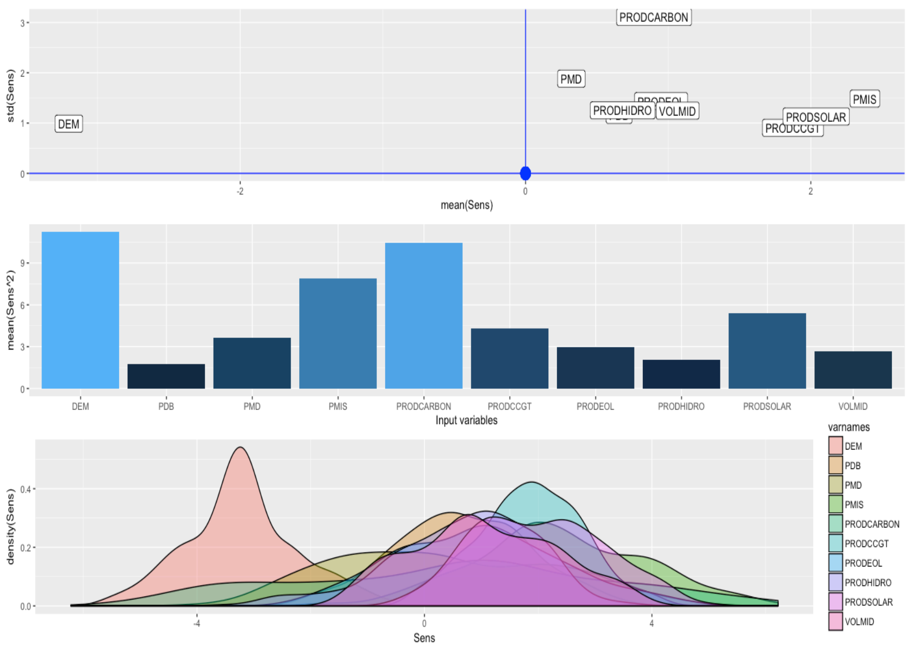

Figure 32 - Sensitivity analysis of MLP, without predictors VOLMIS and PDS (mlp3)

Variable PDB is also removed, in order to ensure consistency and because, based on the results shown in Figure 31, its sensibility's mean is close to zero, and smaller compared with other predictors' sensibilities. Optimal size (16) and decay (2.1111) values are found iteratively. Results are improved, as the trained model explains 86,86% of the variance (Rsquared) with an error of 1,835 (RMSE). A new sensitivity analysis is performed to analyze which variables strongly affect the output and which ones can be pruned. Results are shown in Figure 33.

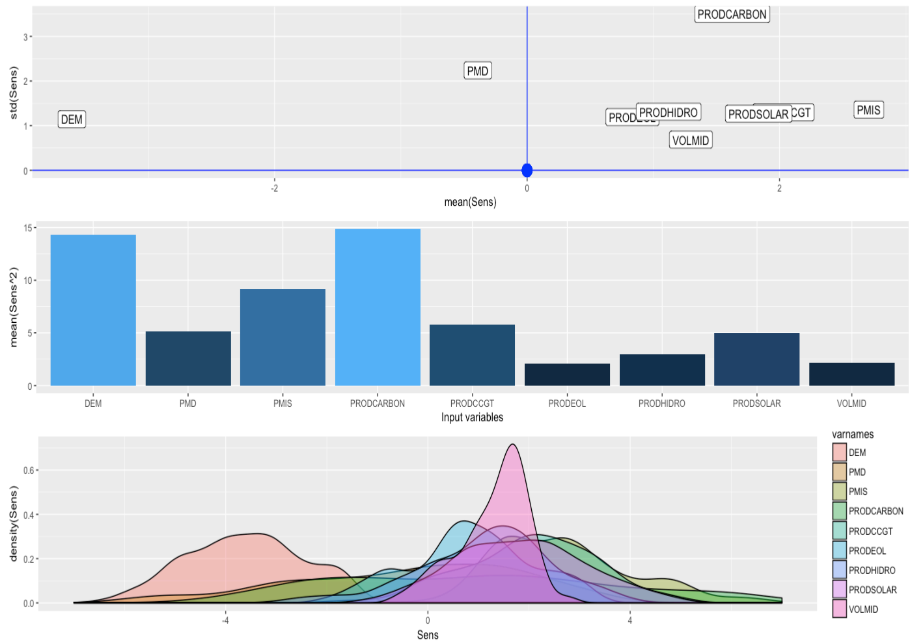

Figure 33 - Sensitivity analysis of MLP, without predictors VOLMIS, PDS and PDB (mlp4)

The new sensitivity analysis also shows that now all variables have a density distribution along the sensitivity not centered in zero neither they have a mean value too close to zero meaning that all variables have an impact over the output.

Model comparison

Figure 34 the performance of the different methods used. Cross-validation, training and validation errors are shown. Also, new data (belonging to July 1st to July 3rd, 2018 market outcomes) has been used to evaluate model's overfitting to the initial dataset. As it can be seen, non-linear models (MLP) show a better performance than linear regression models both in cross-validation, training, and validation errors, as the prediction of continuous intraday electricity prices is probably a non-linear function.

Among the non-linear models trained, mlp3 and mlp4 show a better performance in cross-validation and training errors. However, their performance is worse in validation and new data errors, which reflects overfitting in model training: the aim of the model is to be able to predict the output when fed with validation data which it has not encountered during its training. If the model is too closely fit to the initial data set, the quality of its predictions could be reduced. Thus, the chosen model for predicting continuous intraday prices is *mlp,* using all predictors, as cross-validations and training errors are acceptable, and validation and new data errors are improved.

	Predictors	Tuning parameters	Cross-validation		Training		Validation		New data	
			RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
lm.rfe	All	-	2,153	0,819	2,1066	0,8281	2,4195	0,7322	2,5101	0,7468
lm2.rfe	2,3,4,5,7,8,9,10,11,12,13	-	2,234	0,8066	2,1919	0,8139	2,5024	0,7148	2,3687	0,7424
mlp	All	Size=70 Decay=2.22778	1,9378	0,85	1,4266	0,9217	2,0318	0,8085	2,3421	0,7335
mlp2	2,3,4,5,6,7,8,9,10,11,13	Size=47 Decay=1.93333	1,894	0,8584	1,3266	0,9324	2,0814	0,8004	2,5908	0,7352
mlp3	2,3,5,6,7,8,9,10,11,13	Size=30 Decay=2.33333	1,847	0,8657	1,3607	0,9288	2,0684	0,8052	3,0193	0,5998
mlp4	2,3,6,7,8,9,10,11,13	Size=16 Decay=2.11111	1,835	0,8686	1,4291	0,9214	2,0548	0,8049	2,7852	0,696

*Predictors: [2] PMD, [3] PMIS, [4] PDS, [5] PDB, [6] DEM, [7] PRODEOL, [8] PRODCCGT, [9] PRODCARBON, [10] PRODSOLAR, [11] PRODHIDRO, [12] VOLMIS, [13] VOLMID

Figure 34 - Summary of models' performance

Results

Figure 35 shows day-ahead, upwards and downwards imbalance monthly average prices, compared with MLP monthly average prediction of continuous intraday market prices, over 2017. It can be seen that, generally, continuous intraday prices are higher than day-ahead prices. Furthermore, according to predicted values, continuous intraday prices would have been higher than upwards imbalance prices in 82% of the hours.

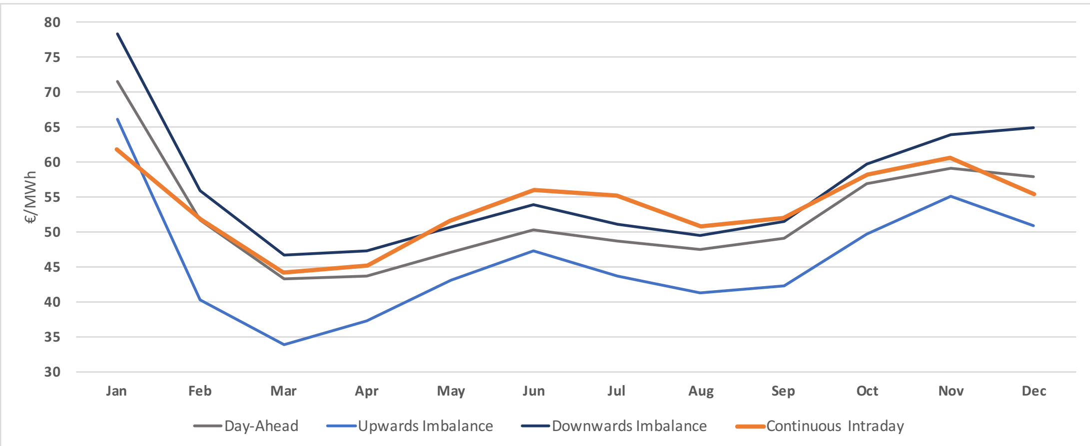

Figure 35 – Monthly price comparison of day-ahead, upwards imbalance, downwards imbalance and predicted continuous intraday monthly prices over 2017.

As a consequence, intraday continuous market may not offer economic incentives to agents to correct downward imbalances: if a consumer foresees that it will be consuming more than what he has bought in previous markets, it will probably prefer to create a downward imbalance than to buy the extra-energy required in the continuous intraday market, as prices tend to be higher than downward imbalance prices.

On the contrary, as the average continuous market price is generally higher than upwards imbalance prices, continuous intraday prices seem to be very attractive for correcting positive imbalances. This is the case of the industrial consumer under study: as can be seen in Figure 36, it generally consumes less than what he had bought in the different markets, creating a positive imbalance. Therefore, it will have a very high incentive to sell the extra-energy in the intraday continuous market, in order to take advantage of the high intraday continuous market prices and avoid lower positive imbalance prices.

It should be noted that the industrial consumer strives to accurately predict their energy consumption in order to buy it in the day-ahead or intraday market. Therefore, it could be assumed that these imbalances could not be reduced by consistently under-buying in the day-ahead market.

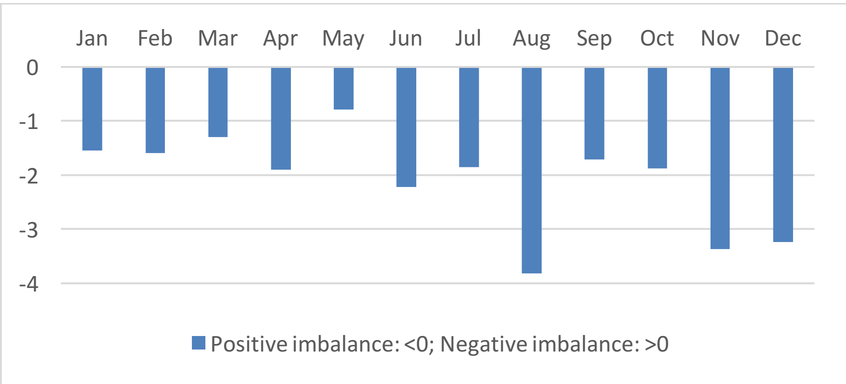

Figure 36 – Average monthly imbalance of the industrial consumer under study over 2017.

Specifically, Figure 37 shows the cost-reduction that the industrial consumer analyzed would have obtained if he had been able to correct every imbalance over 2017 in the intraday continuous market. If he had been able to accurately buy its program in the day-ahead market, income arising from the settlement of the imbalanced energy would have been 991.549,28 €. However, as the consumer was not able to do so, income arising from the settlement of the imbalanced energy was 839.483,13 €, which implies an imbalance cost of 152.066,15 €. Finally, if the consumer had been able to correct every imbalance over 2017 in the intraday continuous market, income arising from the settlement of the imbalanced energy would have been 1.011.347,49 €, which implies an extra income as compared to the day-ahead settlement of 19.798,21 €. Therefore, its imbalance cost would have been reduced by 113%. According to these results, the industrial consumer under study may be interested in investing on the technology and staff required to operate in the intraday continuous market. In case it is represented by a supplier, it could renegotiate contract terms to be able to take advantage of this new mechanism.

	Day-Ahead	Imbalance	Continuous-market	Difference
Settlement (€)	-991.549,28	-839.483,13	-1.011.347,49	-171.864,36
Cost (€) (compared to DA settlement)	0	152.066,15	-19.798,21	-171.864,36

Figure 37 - Settlement and cost of creating an imbalance compared to correcting the program in intraday continuous market over 2017.

However, these results clearly depend on the industrial process of this specific consumer - as it generally creates positive imbalances - and cannot be extrapolated to others. Also, these results show the theoretical cost-reduction that the industrial consumer would obtain in case it had been able to predict every imbalance over 2017 at least one hour before delivery. If imbalances are unpredictable or occur closer to real-time, gate closure would have already taken place and the imbalance could not be corrected in any market.

Also, as mentioned before, high continuous intraday prices offer a high incentive to correct positive imbalances, but a reduced incentive to correct negative imbalances. This may result in a lack of buying interest in the intraday market and an increase in the bid-ask spread caused by the reduced liquidity. Although observed traded volumes in the intraday continuous market are generally higher than the volumes that the industrial consumer under-study would be interested in trading, the lack of liquidity can complicate finding an interested counterpart. However, as intraday continuous market operates in a European scale, agents in other countries are subject to different imbalance pricing mechanisms and they may face different incentives than agents in Spain. Increased interconnection capacity could probably help improve liquidity in intraday continuous market and reduce the market costs associated to the lack of buying interest in Spain.

Finally, presented results are subject to the ability of the MLP model of accurately predicting intraday continuous prices. Every model trained for this study predicts an imbalance cost-reduction over 2017 for the industrial consumer under-study. However, as the training dataset used is small and only takes into account data from a specific time of the year, prediction models should be updated as the amount of data available increases, in order to improve their performance and avoid overfitting.

d. Conclusions

In this section, Spanish intraday current market outcomes were analyzed and a forecasting model was developed in order to predict hourly average continuous intraday market prices. This model was used to generate hourly average continuous intraday market prices for 2017, in order to estimate the potential cost reduction that the industrial consumer under study would have been able to obtain by correcting its program in the intraday continuous market. It can be concluded that:

- Market liquidity can really impact market outcomes. Costs associated to illiquidity in a market are bid-ask spread, market impact costs and search costs. Continuous markets tend to present lower liquidity levels than auctions do, increasing the risk associated to trading.

- In Spain, the first intraday auction is the most liquid one. Intraday continuous market volumes traded are small, but in line with intraday market session 3,4 and 5. It can be assumed that an industrial market agent will be able to correct small deviations in the intraday continuous market.

- In the last 2 years, auction intraday market prices have generally been lower or equal than day-ahead market prices. Also, auction intraday market prices have been more than 3 €/MWh less expensive than imbalance prices. Therefore, agents have an incentive to modify its program in an intraday market auction than to deviate.

- Expected prices in intraday continuous market are needed in order to evaluate the incentive that market agents have to correct their programs in the intraday continuous market. Predictive models can be used in order to estimate prices in intraday continuous markets over 2017.

- Linear regression and multilayer perceptron methods have been used to predict intraday continuous market prices, looking to minimize RMSE error. Non-linear models (MLP) show a better performance than linear regression models both in cross-validation, training, and validation errors, as the prediction of continuous intraday electricity prices is probably a non-linear function. MLP model using all predictors is chosen, as cross-validations and training errors are acceptable, and validation and new data errors are smaller than other models'.

- Continuous intraday prices obtained over 2017 are generally higher than day-ahead prices and, in some cases, even higher than downward imbalance prices. High continuous intraday prices offer a high incentive to correct positive imbalances but a reduced incentive to correct negative imbalances. As the industrial consumer under-study generally creates positive imbalances, it may be able to reduce the cost associated to imbalances by trading in the continuous intraday market.

- However, these results should be taken with caution. First, the ability of an industrial consumer to actually reduce its imbalance cost in this scenario depends on its productive process, as it is only a business opportunity if imbalances are generally positive and can be forecasted at least an hour before gate-closure. Second, they are subject to the ability of the model to accurately predict continuous intraday prices. As the dataset used in this study is small, model performance could be progressively improved by including new, different data to train the model.

Chapter 5

Conclusions

European energy policy in the last decades has focused on developing an electricity internal market, comprising several European states, that provides the right market signals in order to improve efficiency and sustainability, prices and security and quality of supply. Price Coupling of Regions project and the development of EUPHEMIA algorithm brought Day-Ahead Markets in Europe to the last stages of market coupling. However, there are other physical markets for trading electricity where coupling and harmonization are still needed: forward markets and balancing and intraday markets. In August 2015, the European Commission established the Target Model for Intraday markets' integration, based on a cross-border continuous intraday market which will allow agents in different regions to trade electricity continuously throughout the day, in which cross- border capacity is allocated based on implicit methods. In this context, the Cross-Border Intraday Initiative – XBID Project is created by several European PXs and TSOs, in order to develop an integrated Intraday Market, so that market participants' bids in one country can be matched with bids from market participants in any other country, as long as transmission capacity is available.

Intraday Market in the Iberian Electricity Market was previously structured in six centralized auctions, based on marginal pricing with uniform prices. In the context of the XBID project, OMIE, REE and REN proposed a hybrid intraday market design in which continuous intraday market sessions are combined with the existing auctions. This new model reduces lead time before delivery to 1 hour, increasing the chances of adjusting agents' programs and reducing imbalances. The Iberian Hybrid Intraday market started operating on June 12th 2018, for the energy delivered on June 13th 2018.

Advantages and disadvantages of this new market design have been discussed. It argued that, as continuous trading prices vary from trade to trade, a unique and clear reference price is harder to identify. Auctions, on the other hand, provide a more reliable reference price. A hybrid market design as the one proposed by Iberian operators can help overcome this disadvantage of continuous trading, providing flexibility to agents' as well as clear price signals resulting from intraday auctions. Furthermore, the hybrid mechanism provides a playing field for small market participants, who may not have the resources to implement a 24/7 trading desk which would be required in an all-continuous intraday market.

Also, the Iberian hybrid market design only considers hourly products. However, distributed resources are able to vary consumption and generation in very short periods of time. Thus, it would be beneficial to introduce shorter products as distributed resources penetration increases, especially as gate closure moves closer to real-time, in line with other European intraday markets. However, Spanish System Operator REE has already stated the operational difficulties associated to moving gate-closure up to one hour before delivery, using only hourly products, so resistance to further changes in this line can be expected.

Finally, market liquidity can really impact market outcomes. Costs associated to illiquidity in a market are bid-ask spread, market impact costs and search costs. Continuous markets tend to present lower liquidity levels than auctions do, as auctions gather all agents at a specific time, increasing the risk associated to trading. However, market liquidity in the Spanish continuous market sessions will not only depend on market design but will also be influenced by other regulatory aspects characterizing the Spanish energy market such as renewable units being responsible for their imbalances or dual imbalance pricing.

In the case of the industrial consumer analyzed, reducing the extra-costs associated to imbalances could yield a competitive advantage in the electrometallurgy industry. However, consumers will only be able to reduce the extra-costs associated to imbalances in continuous intraday sessions if these allow buying/selling electricity at a market price below/over the corresponding imbalance price, which will depend on session's liquidity and price convergence. Expected prices in intraday continuous market are needed in order to evaluate the incentive that industrial market agents have to correct their programs in the intraday continuous market. Predictive models can be used in order to estimate prices in intraday continuous markets over 2017.

Linear regression and multilayer perceptron methods have been used to predict intraday continuous market prices, looking to minimize RMSE error, based on a dataset comprising market outcomes from 13th to 30th June 2018. Non-linear models (MLP) show a better performance than linear regression models both in cross-validation, training, and validation errors, as the prediction of continuous intraday electricity prices is probably a non-linear function. MLP model using all predictors is chosen, as cross-validations and training errors are acceptable, and validation and new data errors are smaller than other models'.

Continuous intraday prices obtained over 2017 are generally higher than day-ahead prices and, in some cases, even higher than downward imbalance prices. High continuous intraday prices offer a high incentive to correct positive imbalances but a reduced incentive to correct negative imbalances. As the industrial consumer under-study generally creates positive imbalances, it may be able to reduce the cost associated to imbalances by trading in the continuous intraday market.

However, these results should be taken with caution. First, the ability of an industrial consumer to actually reduce its imbalance cost in this scenario depends on its productive process, as it is only a business opportunity if imbalances are generally positive and can be forecasted at least an hour before gate-closure. Therefore, these results should not be extrapolated. Instead, the study should be carried out individually for each industrial productive process. Second, results are subject to the ability of the model to accurately predict continuous intraday prices. As the dataset used in this study is small and only considers summertime behavior, model performance should be improved in the future by increasing the training dataset with new, different data.

Chapter 6

References

[1] «Directive 2009/72/EC of the European Parliament and of the Council of 13 July 2009 concerning common rules for the internal market in electricity and repealing Directive 2003/54/EC.,» Official Journal of the European Union, 2009.

[2] *Progress towards completing the Internal Energy Market,* European Comission, 2014.

[3] L. Olmos y I. Pérez-Arriaga, «Regional Markets,» de *Regulation of the Power Sector*, Springer, 2013.

[4] M. Key, *The EU Target Model for electricity markets: fit for purpose?,* The Oxford Institute for Energy Studies, 2013.

[5] P. J. Mejía, *Hacia un nuevo diseño del mercado eléctrico en Europa,* Cuadernos de la Energía nº47, 2016.

[6] PCR Price Coupling of Regions, «EUPHEMIA Public Description,» 2016.

[7] *Cross-border intraday: questions and answers,* XBID Cross-border intraday market project, 2017.

[8] *Comission Regulation (EU) 2015/1222 of 24 July 2015 establishing a guideline on capacity allocation and congestion management,* Official Journal of the European Union, 2015.

[9] OMIE, «Nuevo Funcionamiento del Mercado Intradiario MIBEL,» [En línea]. Available: http://www.nuevomercadointradiario.omie.es/. [Último acceso: 29 October 2017].

[10] OMIE, [En línea]. Available: www.omie.es. [Último acceso: Diciembre 2017].

[11] C. Á. José Pablo, «European short-term electricity market designs under high penetration of wind power.,» Universidad Pontificia de Comillas, 2014.

[12] P. Alfaro Cerezo y J. Montes Ponce de Leon, «Estudio crítico del coste y del precio de la energía eléctrica en España. Aplicación a la economía de las redes inteligentes.,» Universidad Pontificia de Comillas, 2016.

[13] REE, «Esios - Sistema de Información del Operador del Sistema,» Febrero 2018. [En línea].

[14] M. Ventosa, P. Linares y I. Pérez-Arriaga, «Power System Economics,» de *Regulation of the Power Sector*, Springer, 2013.

[15] PCR Price Coupling of Regions, *PCR Main Features,* 2013.

[16] PCR Price Coupling of Regions, «Los mercados diarios de electricidad del suroeste y del noroeste de Europa se acoplan con éxito.,» Mayo 2014.

[17] RTE, REE, «Capacity Allocation Rules for the France-Spain Interconnection for the Intradat timeframe».

[18] EPEX Spot, «Operational Rules,» 2018.

[19] NordPool, «Rules and Regulations,» [En línea]. Available: https://www.nordpoolgroup.com/TAS/Rules-and-regulations/. [Último acceso: Enero 2018].

[20] C. Batlle y I. Perez-Arriaga, «Electricity Generation and Wholesale Markets,» de *Regulation of the Power Sector*, Springer, 2013.

[21] R. Scharff y M. Amelin, «Trading behaviour on the continuous intraday market ELBAS,» *Elsevier, Energy Policy 88, 544–557,* 2016.

[22] OMIE, REE, REN, «Informe sobre la consulta pública a la propuesta de funcionamiento para el mercado intradiario en el MIBEL,» 2017.

[23] C. Weber, « Adequate intraday market design to enable the integration of wind energy into the European power systems,» *Elsevier, Energy Policy 38, 3155-3163,* 2010.

[24] M. A. Pérez, «Economic analysis of the increase in number of intraday market sessions in the iberian market.,» 2016.

[25] Red Eléctrica de España, «Servicios de Ajuste del Sistema - Avance 2017».

[26] Red Eléctrica de España, «ESIOS - Sistema de información del Operador del Sistema,» [En línea]. Available: www.esios.ree.es. [Último acceso: 2017].

[27] European Commission, «Regulation (EU) 2017/2195 establishing a guideline on electricity balancing,» 2017.

[28] P. Willis, «Refining electricity imbalance and balancing prices,» 24 Noviembre 2016. [En línea]. Available: https://www.twobirds.com/en/news/articles/2016/global/refining-electricity-imbalance-and-balancing-prices.

[29] Red Eléctrica de España, «La operación del sistema eléctrico y la gestión de los servicios de ajuste».

[30] Y. Amihud y H. Mendelson, «Liquidity, Asset Prices and Financial Policy,» Financial Analysts Journal, 47, 1991.

[31] J. Li, «Assessing the accuracy of predictive models for numerical data: not r nor r2. Why not? Then what?,» PLoS ONE, 2017.

[32] A. Muñoz y J. Portela, «Machine Learning: course notes,» 2018.

[33] C. Willmott y K. Matsuura, «Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance,» *Climate Research,* vol. 30, pp. 78-82, 2005.

[34] Penn State - Eberly College of Science, «Stat 501 - Regression Methods: The Coefficient of Determination,» [En línea]. Available: https://onlinecourses.science.psu.edu/stat501/node/255/.

[35] Y. Xin y G. S. Xiao, Linear Regression Analysis. Theory and Computing, Singapore: World Scientific, 2009.

[36] R. Nau, «Notes on linear regression analysis».

[37] Wikipedia. [En línea]. Available: https://www.wikipedia.org/.

[38] Y.-w. Wang, «Neural Network Forecasting of the Production Level of Chinese Construction Industry».

[39] C. Stergiou y D. Siganos, «Imperial College London - Neural Networks,» [En línea]. Available:

https://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html#What%20is%20a%20Neural%20Network. [Último acceso: 2018].

[40] MathWorks Discovery , «Supervised and Unsupervised Learning,» [En línea]. Available: https://es.mathworks.com/discovery/unsupervised-learning.html?s_tid=srchtitle.

[41] H. Bouzgou, «Advanced methods for the processing and analysis of multidimensional signals: application to wind speed.,» University of Hadj Lakhdar, Batna.

[42] NordPool Spot, «Explicity and implicity capacity auction».

[43] C. F. Fernandes, «Adapting european market arrangements to facilitate the contribution of renewable generators to electricity balancing.,» 2017.

[44] R. v. d. Veen, A. Abassy y R. Hakvoort, «A comparison of imbalance settlement designs and results of Germany and the Netherlands,» 2010.

Annex

R Script

```
#Insert required libraries
library(caret)
library(ggplot2)
library(GGally)
library(gam)
library(splines)
library(corrplot)
library(lubridate)
library(pls)
library(car)
library(kernlab)
library(gridExtra)
library(ISLR)
library(readxl)
library(NeuralNetTools)
source("RegressionTools.R")

#Load data from txt file----------------------------------------------------
fdata =read.table("TRAINING13a30.txt",header=TRUE, fill=TRUE)
fdata_new =read.table("NEWDATA13A30.txt",header=TRUE, fill=TRUE)
fdata_2017 =read.table("Datos2017.txt",header=TRUE, fill=TRUE)

fdata_2017$PDS=as.numeric(as.character(fdata_2017$PDS))

fdata = na.omit(fdata) #Eliminate NA
fdata_new=na.omit(fdata_new)
fdata_2017=na.omit(fdata_2017)

#Exploratory analysis ------------------------------------------------------
ggpairs(fdata,columns=2:14, cardinality_threshold = 150)
boxplot(fdata[,c(2,3,4,5,12)])
boxplot(fdata[,c(6,10,11)])
boxplot(fdata[,c(7,8,9)])

#Divide the data into training and validation sets -----------------------
set.seed(150) #For replication
ratioTR = 0.8 #Percentage for training: 80%
#createDataPartition creates proportional partitions.
trainIndex <- createDataPartition(fdata$PMIC,     #output variable.
                                  p = ratioTR,    #split probability
                                  list = FALSE,   #no output as list
                                  times = 1)      #only one partition
fdata_train = fdata[trainIndex,]
fdata_val = fdata[-trainIndex,]
fdataTR_eval = fdata_train #Set equal to training set, for evaluation
fdataTV_eval = fdata_val #Set equal to validation set, for evaluation


#K-fold with 10 folds ---------------------------------------------------
ctrl_tune <- trainControl(method = "cv",
                          number = 10,
                          summaryFunction = defaultSummary,
                          returnResamp = "final",
                          savePredictions = TRUE)
```

```
#----------- LINEAR REGRESSION WITH RFE FOR VARIABLE SELECTION -----------
ctrl_none <- trainControl(method = "none",
                          summaryFunction = defaultSummary,
                          returnResamp = "final
                          savePredictions = TRUE
#Specifies the cross validation method used for selecting the optimum number of
 variables
ctrl_rfe <- rfeControl(method = "cv",
                       number = 10,
                       verbose = TRUE,
                       functions = caretFuncs)

#TRAIN MODEL WITH ALL VARIABLES
set.seed(150)
subsets <- 1:20 #The number of features that should be retained
lm.RFE <- rfe(form = PMIC~.,
              data = fdata_train[,c(2:14)], #Arguments passed to train()
              method = "lm",
              preProcess = c("center","scale"),
              trControl = ctrl_none,          # Arguments for rfe
              sizes = subsets,
              metric = "RMSE",
              rfeControl = ctrl_rfe)

lm.RFE                        #Cross validation results and variable selection
ggplot(lm.RFE,metric = "RMSE")
lm.RFE$fit
summary(lm.RFE$fit)
lm.RFE$fit$finalModel         #Final model trained
vif(lm.RFE$fit$finalModel)

fdataTR_eval$lmRFE_pred = predict(lm.RFE,  newdata = fdata_train)
fdataTV_eval$lmRFE_pred = predict(lm.RFE,  newdata = fdata_val)
fdata_new$lmRFE_pred = predict(lm.RFE,  newdata = fdata_new)


#TRAIN MODEL WITHOUT PREDICTOR DEM
set.seed(150)
subsets <- 1:20 #The number of features that should be retained
lm2.RFE <- rfe(form = PMIC~.,
              data = fdata_train[,c(2:5,7:14)],   # Arguments passed to train()
 function
              method = "lm",
              preProcess = c("center","scale"),
              trControl = ctrl_none,          # Arguments for rfe
              sizes = subsets,
              metric = "RMSE",
              rfeControl = ctrl_rfe)

lm2.RFE                           #Cross validation results and variable selection
ggplot(lm2.RFE,metric = "RMSE")
lm2.RFE$fit
summary(lm2.RFE$fit)
lm2.RFE$fit$finalModel            #Final model trained
vif(lm2.RFE$fit$finalModel)

fdataTR_eval$lm2RFE_pred = predict(lm2.RFE,  newdata = fdata_train)
fdataTV_eval$lm2RFE_pred = predict(lm2.RFE,  newdata = fdata_val)
fdata_new$lm2RFE_pred = predict(lm2.RFE,  newdata = fdata_new)
```

```r
#----------- MULTILAYER PERCEPTRON -----------
#MLP
set.seed(150) #For replication
mlp.fit = train(form = PMIC~.,
                data = fdata_train[,c(2:14)],
                method = "nnet",
                linout = TRUE,
                maxit = 400,
                tuneGrid = expand.grid(size=70, decay=2.22778),
                preProcess = c("center","scale"),
                trControl = ctrl_tune,
                metric = "RMSE")
mlp.fit                  #Information about the resampling settings
ggplot(mlp.fit)
SensAnalysisMLP(mlp.fit) #Statistical sensitivity analysis
plotModelDiagnosisAll(fdata_train[,c(2:14)], fdata_train$PMIC, fdataTR_eval$mlp_pred)

fdataTR_eval$mlp_pred = predict(mlp.fit,  newdata = fdata_train)
fdataTV_eval$mlp_pred = predict(mlp.fit,  newdata = fdata_val)
fdata_new$mlp_pred = predict(mlp.fit,  newdata = fdata_new)

#MLP2
set.seed(150) #For replication
mlp2.fit = train(form = PMIC~.,
                 data = fdata_train[,c(2:11,13:14)], #Sin VOLMIS
                 method = "nnet",
                 linout = TRUE,
                 maxit = 400,
                 tuneGrid = expand.grid(size=47, decay=1.93333),
                 preProcess = c("center","scale"),
                 trControl = ctrl_tune,
                 metric = "RMSE")
mlp2.fit                  #Information about the resampling settings
ggplot(mlp2.fit)
SensAnalysisMLP(mlp2.fit) #Statistical sensitivity analysis
plotModelDiagnosisAll(fdata_train[,c(2:11,13:14)], fdata_train$PMIC,
 fdataTR_eval$mlp2_pred)

fdataTR_eval$mlp2_pred = predict(mlp2.fit,  newdata = fdata_train)
fdataTV_eval$mlp2_pred = predict(mlp2.fit,  newdata = fdata_val)
fdata_new$mlp2_pred = predict(mlp2.fit,  newdata = fdata_new)

#MLP3
set.seed(150) #For replication
mlp3.fit = train(form = PMIC~., #sin VOLMIS, PDS
                 data = fdata_train[,c(2,3,5:11,13,14)],
                 method = "nnet",
                 linout = TRUE,
                 maxit = 400,
                 tuneGrid = expand.grid(size=30, decay=2.33333333),
                 preProcess = c("center","scale"),
                 trControl = ctrl_tune,
                 metric = "RMSE")
mlp3.fit                  #Information about the resampling settings
ggplot(mlp3.fit)
SensAnalysisMLP(mlp3.fit) #Statistical sensitivity analysis
plotModelDiagnosisAll(fdata_train[,c(2,3,5:11,13,14)], fdata_train$PMIC,
 fdataTR_eval$mlp3_pred)

fdataTR_eval$mlp3_pred = predict(mlp3.fit,  newdata = fdata_train)
fdataTV_eval$mlp3_pred = predict(mlp3.fit,  newdata = fdata_val)
fdata_new$mlp3_pred = predict(mlp3.fit,  newdata = fdata_new)
```

```r
#MLP4
set.seed(150) #For replication
mlp4.fit = train(form = PMIC~., #sin VOLMIS, PDS, PDB
                 data = fdata_train[,c(2,3,6:11,13,14)],
                 method = "nnet",
                 linout = TRUE,
                 maxit = 400,
                 tuneGrid = expand.grid(size=16, decay=2.1111111),
                 preProcess = c("center","scale"),
                 trControl = ctrl_tune,
                 metric = "RMSE")
mlp4.fit                  #Information about the resampling settings
ggplot(mlp4.fit)
SensAnalysisMLP(mlp4.fit) #Statistical sensitivity analysis
plotModelDiagnosisAll(fdata_train[,c(2,3,6:11,13,14)], fdata_train$PMIC,
 fdataTR_eval$mlp4_pred)

fdataTR_eval$mlp4_pred = predict(mlp4.fit,  newdata = fdata_train)
fdataTV_eval$mlp4_pred = predict(mlp4.fit,  newdata = fdata_val)
fdata_new$mlp4_pred = predict(mlp4.fit,  newdata = fdata_new)


# Evaluate training performance -----------------------------------------------------
transformResults <- resamples(list(lm.RFE=lm.RFE$fit, lm2.RFE=lm2.RFE$fit,
 mlp=mlp.fit, mlp2=mlp2.fit, mlp3=mlp3.fit, mlp4=mlp4.fit))
summary(transformResults)
dotplot(transformResults)

# Evaluate validation performance ---------------------------------------------------
caret::R2(fdataTR_eval$lmRFE_pred,fdataTR_eval$PMIC)
caret::R2(fdataTR_eval$lm2RFE_pred,fdataTR_eval$PMIC)
caret::R2(fdataTR_eval$mlp_pred,fdataTR_eval$PMIC)
caret::R2(fdataTR_eval$mlp2_pred,fdataTR_eval$PMIC)
caret::R2(fdataTR_eval$mlp3_pred,fdataTR_eval$PMIC)
caret::R2(fdataTR_eval$mlp4_pred,fdataTR_eval$PMIC)

caret::RMSE(fdataTR_eval$lmRFE_pred,fdataTR_eval$PMIC)
caret::RMSE(fdataTR_eval$lm2RFE_pred,fdataTR_eval$PMIC)
caret::RMSE(fdataTR_eval$mlp_pred,fdataTR_eval$PMIC)
caret::RMSE(fdataTR_eval$mlp2_pred,fdataTR_eval$PMIC)
caret::RMSE(fdataTR_eval$mlp3_pred,fdataTR_eval$PMIC)
caret::RMSE(fdataTR_eval$mlp4_pred,fdataTR_eval$PMIC)

caret::R2(fdataTV_eval$lmRFE_pred,fdataTV_eval$PMIC)
caret::R2(fdataTV_eval$lm2RFE_pred,fdataTV_eval$PMIC)
caret::R2(fdataTV_eval$mlp_pred,fdataTV_eval$PMIC)
caret::R2(fdataTV_eval$mlp2_pred,fdataTV_eval$PMIC)
caret::R2(fdataTV_eval$mlp3_pred,fdataTV_eval$PMIC)
caret::R2(fdataTV_eval$mlp4_pred,fdataTV_eval$PMIC)

caret::RMSE(fdataTV_eval$lmRFE_pred,fdataTV_eval$PMIC)
caret::RMSE(fdataTV_eval$lm2RFE_pred,fdataTV_eval$PMIC)
caret::RMSE(fdataTV_eval$mlp_pred,fdataTV_eval$PMIC)
caret::RMSE(fdataTV_eval$mlp2_pred,fdataTV_eval$PMIC)
caret::RMSE(fdataTV_eval$mlp3_pred,fdataTV_eval$PMIC)
caret::RMSE(fdataTV_eval$mlp4_pred,fdataTV_eval$PMIC)

caret::R2(fdata_new$lmRFE_pred,fdata_new$PMIC)
caret::R2(fdata_new$lm2RFE_pred,fdata_new$PMIC)
caret::R2(fdata_new$mlp_pred,fdata_new$PMIC)
caret::R2(fdata_new$mlp2_pred,fdata_new$PMIC)
caret::R2(fdata_new$mlp3_pred,fdata_new$PMIC)
caret::R2(fdata_new$mlp4_pred,fdata_new$PMIC)
```

```
caret::RMSE(fdata_new$lmRFE_pred,fdata_new$PMIC)
caret::RMSE(fdata_new$lm2RFE_pred,fdata_new$PMIC)
caret::RMSE(fdata_new$mlp_pred,fdata_new$PMIC)
caret::RMSE(fdata_new$mlp2_pred,fdata_new$PMIC)
caret::RMSE(fdata_new$mlp3_pred,fdata_new$PMIC)
caret::RMSE(fdata_new$mlp4_pred,fdata_new$PMIC)

fdata_2017$lmRFE_pred = predict(lm.RFE,  newdata = fdata_2017)
fdata_2017$lm2RFE_pred = predict(lm2.RFE,  newdata = fdata_2017)
fdata_2017$mlp_pred = predict(mlp.fit,  newdata = fdata_2017)
fdata_2017$mlp2_pred = predict(mlp2.fit,  newdata = fdata_2017)
fdata_2017$mlp3_pred = predict(mlp3.fit,  newdata = fdata_2017)
fdata_2017$mlp4_pred = predict(mlp4.fit,  newdata = fdata_2017)

write.table(fdataTR_eval,"Resultados_TR.txt",sep="\t")
write.table(fdataTV_eval,"Resultados_TV.txt",sep="\t")
write.table(fdataTV_eval,"Resultados_TV.txt",sep="\t")
write.table(fdata_new,"Resultados_New.txt",sep="\t")
write.table(fdata_2017, "Resultados2017.txt", sep="\t")
```