UNIVERSIDAD PONTIFICIA COMILLAS DE MADRID
ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA (ICAI)
Instituto de Investigación Tecnológica (IIT)

SHORT-TERM FORECASTING OF ELECTRICITY PRICES: A HYBRID METHODOLOGY BASED ON FUNDAMENTAL AND STATISTICAL ANALYSIS

Predicción de precios del mercado eléctrico en el corto plazo: un procedimiento híbrido basado en análisis fundamentales y estadísticos

Tesis para la obtención del grado de Doctor

Directores: Dr. D. Javier Reneses Guillén

Dr. D. Antonio Bello Morales

Autor: MSc. D. Rodrigo Alejandro de Marcos Peirotén

COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI	ICADE	CIHS

Madrid
Febrero 2020

CONSTANCIA REGISTRAL DEL TRIBUNAL DEL ACTO DE LA DEFENSA DE TESIS DOCTORAL

TÍTULO: Short-term forecasting of electricity prices: a hybrid methodology based on fundamental and statistical analysis.

AUTOR: MSc. D. Rodrigo Alejandro de Marcos Peirotén

DIRECTOR: Dr. D. Javier Reneses Guillén y

Dr. D. Antonio Bello Morales

TUTOR-PONENTE: No aplica

DEPARTAMENTO: Instituto de investigación Tecnológica (IIT)

FACULTAD O ESCUELA: Escuela Técnica Superior de Ingeniería (ICAI)

Miembros del tribunal Calificador:

PRESIDENTE:	**Firma:**
VOCAL:	**Firma:**
VOCAL:	**Firma:**
VOCAL:	**Firma:**
SECRETARIO:	**Firma:**

Fecha de lectura:

Calificación:

To all the people that
have cared about me
and left their mark
on me

Acknowledgements

First of all, I would like to thank above all my thesis supervisors, Javier Reneses and Antonio Bello, for giving me the opportunity to work with them at the IIT, for aiding me in the development of my doctoral thesis and for including me in several collaboration projects with Endesa and broaden my professional career in the industry. Thank you so much for all your guidance, advice and teachings throughout the six years of my stay in your team at the IIT.

Special thanks to Álvaro Sánchez Miralles for taking me as an invited student in the IIT back in late 2011. If it were not for you, I may have missed the chance of becoming a PhD in the IIT, and everything that this implies.

For his friendship and the good moments we have spent at the IIT, I would like to thank Francisco Martín. You have shared with me your helpful experience that you gained throughout your PhD and I will be always thankful to you for that. I am so thankful to Diego and Aurora, who have recently finished their PhDs and are still supporting me so I can become a PhD as well. I have to thank them both for all the happy moments at the IIT. I would also like to thank the IIT in general, which has provided me with an excellent and very friendly environment for all these years. I cannot imagine a better place to develop a PhD. In addition, thanks a lot to my colleagues in the IIT, with special thanks to the spinning team, the breakfast team and all the friends I have made throughout my stay, and thus I must thank, among others, Alberto, Ángel, Geovanny, Luis, Santi, Nacho Herrero, Peyman, Javier Renedo, Eva Paz, Sonja, and Eduardo Alonso.

At this point, I would like to express my gratitude towards Derek Bunn from the London Business School. Thank you so much for hosting me at the LBS and for your insightful teachings and suggestions that have guided the latest stages of my thesis.

All my gratitude to Daniel Lestón, professional energy trader, and formerly my fellow Industrial Engineering classmate, for helping me with several topics and nuances of the Iberian power exchange, as well as for all the energy and support you have been giving me for over ten years.

Finally, my deepest gratitude goes to my family for their enormous moral support throughout my PhD. Thanks a lot to my brother for your advice, for all the days you have come by to have lunch and for being patient with my temper in the moments of stress. Thanks so much to my mother, for all your care and love since I was born. You have always been there in the moments of need and so I would like to dedicate this thesis to you.

Contents

Abstract

The ongoing evolutions that are exhibited in the electricity markets of today have brought up a new competitive environment in which traders and practitioners must adapt their strategies and look for support for the decision-making when operating in the market. This is a result of a combination of several factors, which are the increasing renewable penetration in power systems, ongoing regulatory changes, varying weather conditions, volatile fuel costs, and also the global financial instability. Due to the resulting uncertainty as well as the non-stationary and volatile behaviour that is present in electricity market prices, market agents and practitioners are resorting to electricity price forecasting models for several purposes, some of which are highly relevant in short-term contexts, such as risk management.

In order to consider all relevant drivers of electricity prices for said purposes, fundamental-statistical hybrid forecasting models may be utilised. On the one hand, fundamental models are responsible for simulating the market clearing while considering the operation of the market, physical features of the power system, regulatory effects, etc. On the other hand, statistical models take advantage of linear and non-linear trends that are assumed to be repetitive, such as intraday effects, seasonal patterns, etc. However, this particular hybrid model has been mostly utilised in longer-term applications, where they have proved beneficial. Therefore, employing this method in short-term applications would involve different and/or additional considerations that must be ascertained.

Keywords: electricity markets; fundamental models; hybrid approaches; short-term electricity price forecasting; statistical models

Resumen

Actualmente existe un panorama muy competitivo en el mercado eléctrico en el que los participantes deben ajustar sus estrategias y buscar apoyo a la hora de tomar decisiones al realizar transacciones. Esta situación es causa de los numerosos cambios que se dan en los mercados eléctricos que están relacionados con los siguientes factores: el auge de las tecnologías de generación renovable en el sistema eléctrico, los numerosos cambios regulatorios, las variaciones climatológicas, la volatilidad de los costes de los combustibles, y también la inestabilidad financiera global. Por lo tanto, el comportamiento de los precios del mercado eléctrico presenta una volatilidad e incertidumbre elevadas, además de ser no estacionario. Por ello, los agentes y profesionales del mercado recurren a modelos de predicción del precio eléctrico para obtener apoyo en sus transacciones, que en mayor parte son relevantes para el corto plazo, como la gestión de riesgos.

Una forma de tener en cuenta todos los factores que influyen en el precio del mercado eléctrico es por medio de la hibridación de modelos fundamentales y estadísticos. Por un lado, los modelos fundamentales llevan a cabo una simulación de la casación del mercado eléctrico en función de la dinámica del mismo, las propiedades físicas del sistema eléctrico, efectos regulatorios, etc. Por otro lado, los modelos estadísticos son capaces de detectar tendencias lineales y no lineales que ocurren de forma repetitiva, como los patrones intradiarios y otros efectos periódicos. Sin embargo, esta hibridación de modelos se ha llevado a cabo en su mayor parte en contextos de largo plazo, en los cuales se ha demostrado sus ventajas. En consecuencia, la aplicación de este método en el corto plazo conllevaría otras consideraciones y/o más hipótesis que deben determinarse.

Palabras clave: mercados de electricidad; modelos de fundamentales; procedimientos híbridos; predicción de precios de la electricidad en el corto plazo; modelos estadísticos

Dissertation

The thesis developments and original contributions have been presented in the following publications (de Marcos, Bello, et al., 2017, 2018, 2019a, 2019b; de Marcos, Bunn, et al., 2019):

Journal Articles:

de Marcos, R. A., Bello, A., & Reneses, J. (2019a). Electricity price forecasting in the short term hybridising fundamental and econometric modelling. *Electric Power Systems Research, 167,* 240–251. http://doi.org/10.1016/j.epsr.2018.10.034

de Marcos, R. A., Bello, A., & Reneses, J. (2019b). Short-Term Electricity Price Forecasting with a Composite Fundamental-Econometric Hybrid Methodology. *Energies*, *12*(6), 1067. http://doi.org/10.3390/en12061067

Conference Papers:

de Marcos, R. A., Bello, A., & Reneses, J. (2017). Short-term forecasting of electricity prices with a computationally efficient hybrid approach. In *14th International Conference on the European Energy Market - EEM17*

de Marcos, R. A., Bello, A., & Reneses, J. (2018). Hybridisation of Fundamental and Composite Econometric Modelling for Short-Term Electricity Price Forecasting. In *15th International Conference on the European Energy Market - EEM18*

Working Papers:

de Marcos, R. A., Bunn, D. W., Bello, A., & Reneses, J. (2019). Short-Term Electricity Price Forecasting with Recurrent Regimes and Structural Breaks. Working Paper

Chapter 1
Introduction

The ongoing evolutions that are exhibited in power exchanges have brought up a new competitive environment in which traders and practitioners must adapt their strategies and look for support for their decision-making when operating in the market. This is a result of a combination of factors, such as the increasing renewable penetration in power systems and ongoing regulatory changes. Furthermore, the current global instability heightens the volatility of power systems, which are becoming more complex than ever. Due to the resulting uncertainty, market agents and practitioners are resorting to electricity price forecasting models for several purposes, such as risk management and speculation. This thesis is thus aimed at the development of a suitable forecasting methodology that considers all relevant trends and behaviours that are exhibited in electricity market prices. Therefore, the core of the proposed methodologies of this thesis is based on the hybridisation of two forecasting approaches that have been widely used individually in the literature: fundamental and statistical. Both components are responsible of capturing the plethora of factors that drive electricity prices, whether they originate from market fundamentals or agent behaviours, and incorporate these effects into their electricity price forecasts. This combination has shown positive results for medium- and long-term applications. However, this approach has rarely been carried out in short-term contexts, in which different assumptions should be taken into consideration. This chapter introduces the context and motivation, defines the main objectives and describes the structure that has been followed in this thesis document.

1.1 Motivation

Throughout the previous decades, power exchanges worldwide have faced deregulation and liberalisation. As a result, market agents have a wide range of actions and investment options at their disposal, and thus the complexity of the market environment has considerably risen to the extent that it is imperative to invest significant resources in strategy and decision-making when it comes to participating and operating in the electricity market. Furthermore, there are several factors that are, as a whole, causing significant instabilities in the system, such as the growing renewable penetration, constant regulatory reforms and volatile fuel prices. Consequently, dealing with the underlying uncertainty and withstanding competition has become a highly challenging task for traders and practitioners. As such, electricity price forecasting models are at the centre of the spotlight for several purposes in this context, for instance, speculation and risk management. This also encourages interest in the academic environment, which is the primary source of these forecasting models.

Moreover, the relevance of the planning horizon is crucial when selecting the price prediction method, as the assumptions and considerations vary considerably depending upon the forecasting horizon. As stated in (Weron, 2014), there is no general consensus in the literature as to what time intervals pertain to each planning horizon in the electricity market price forecasting framework. The most popular uses of electricity price forecasting models fall in the short-term planning horizon category (horizons ranging from a few hours up to a week), which is closely related to decision-making in day-to-day market operations, such as bidding and unit commitment scheduling. Therefore, longer-term applications of electricity price forecasting models are less common, yet these are of vital importance for electricity producers when it comes to, for instance, trading via financial derivatives or making investment decisions. In any case, a proper forecasting methodology aids market participants in the ascertaining of favourable operations and transactions in advance.

As discussed in (Bello, Reneses, Muñoz, et al., 2016), statistical analysis[1] methods are prominent in the short-term electricity price forecasting literature, whereas medium- and long-term (horizons of generally a few months and a few years respectively) involve fundamental modelling of the market dynamics and the main features of power exchanges. Although there are other forecasting approaches in this context, both of these methods generally cover the most popular and common methods, as seen in (Weron, 2014).

In general, fundamental analyses focus on the operation and the behaviour of the electricity market in order to simulate the market clearing and thus obtain the electricity price forecasts. All physical elements (generation units, transmission lines, etc.) and law-related limits/issues (CO_2 emissions, renewable energy subsidies, taxes, etc.) are considered in fundamental models, which are usually put into practice by means of a market equilibrium optimisation model. Therefore, the resulting electricity price forecasts reflect the relevant physical and economic factors that are present in power exchanges. Nevertheless, as mentioned in the following chapters of this thesis, the literature is scarce of methodologies that propose a thorough representation of the power system with the hourly precision that is generally imposed by short-term contexts, which is

[1] It should be noted that several authors use both "technical" and "econometric" to refer to "statistical" analysis in forecasting modelling contexts. Consequently, these terms are used interchangeably throughout this thesis.

mainly due to the resulting large size and resolution times. Furthermore, an adequate modelling of thermal unit bids, mainly driven by their marginal or opportunity costs, is necessary in order to improve market clearing estimation accuracy in the short term.

As opposed to fundamental approaches, statistical analyses typically depend upon past information in order to determine functional relationships between the behaviour of electricity market prices and that of other factors or explanatory/exogenous variables. Consequently, statistical methods perform under the assumption that history repeats itself in the future, which is not suitable in most cases for today's power markets. Not all relevant events, such as the impending decommissioning of thermal units related to the decarbonisation of the electricity sector, can be treated as recurring phenomena. Therefore, punctual events are incompatible with this premise. Nevertheless, as mentioned earlier, these methods are considerably popular in short-term applications due to their linear and non-linear modelling capabilities. Furthermore, as explained in (Bello, Bunn, et al., 2017), statistical models are proficient at capturing the revealed behavioural aspects of market participants, such as strategic and speculative conducts.

However, certain kinds of events that lead to unstable market circumstances (e.g. high volatility) require a suitable adaptability from the forecasting models in order to avoid accuracy losses, which is of higher importance for short-term electricity price forecasting. Such changes in the market can be defined as structural breaks, in which case, the selection of the input data periods for the statistical models should not be disregarded. Otherwise, it would be counterproductive if a statistical model is trained with input data periods that include structural breaks, as stated in (Pesaran & Timmermann, 2007). These authors claim that adaptability is increased when utilising shorter calibration windows. By contrast, longer calibration windows result in a better estimation of the trained model's parameters. However, the length of these calibration windows may not be as decisive as the selection/removal of specific intervals in the historical dataset. In addition, there is no general consensus in the current literature regarding the appropriate training periods of statistical forecasting applications (regardless of the planning horizon), and thus this calls for a suitable method with robust criteria in order to determine it prior to the actual elaboration of the forecast (i.e. in an ex-ante manner).

Furthermore, certain circumstances that may occur in power systems cause sudden disruptions in electricity prices that last for a few hours. These occurrences are widely known as spikes and their volatility is considerably higher than those of other commodities. For instance, an immediate and sharp increase in renewable generation exercises a strong downward pressure on the price level as a result of the absence of fossil fuel generation in the market clearing. Moreover, upward spikes are fairly common in power systems with scarce overcapacity, typically caused by significant thermal outages, fuel price surges or network congestion. Therefore, this calls for a methodology that is able to capture the causes behind price spikes, which can be considered on a fundamental model when estimating the market clearing if relevant explanatory information to the price forecasts is available and can be considered in the market clearing, such as weather forecasts and planned generation unit maintenances.

1.2 Main remarks and identified gaps

The above facts suggest that statistical and fundamental methods provide different solutions to the issues and challenges present in electricity price forecasting applications. Moreover, combining both of these approaches[2] seems to be a growing trend, albeit mostly performed on medium- and long-term applications, in which positive results have been accomplished, as seen in (Bello, Bunn, et al., 2017). Therefore, the adequacy of these hybrid methods have not been sufficiently tested in the short term. Furthermore, it is most certain that these models will show poor performance if directly applied to the short term without any specific consideration just by taking into account the intraday effects that affect hourly forecasts as opposed to daily or monthly forecasts. Therefore, the question that arises in this regard and that this thesis attempts to answer is: Can fundamental-statistical hybrid models yield the same advantages as seen in medium-term applications?

Other issues involve the sudden evolution of market regimes that feature heightened levels of volatility. This is one of the major drawbacks of fundamental models, whose estimated prices are somewhat flat and thus fail to predict extremely low or high prices. Furthermore, another flaw lies in their inability to effectively capture short-term trends in their price forecasts, which can be corrected by means of statistical techniques. Moreover, a thorough representation in hourly basis of the power market, including all generation units and other physical elements, leads to a computationally cumbersome task when put into practice, although simplification methods may be applied in order to address this issue. By contrast, gradual market regime changes may be explained to some extent by the estimated prices of fundamental models if the market clearing is driven by said changes. Therefore, it would be interesting to verify if a hybrid fundamental-statistical forecasting approach is capable of responding to both sudden and gradual market regime changes and capture the varying features and behaviours of prices in such situations.

[2] This model combination or hybridisation, i.e. fundamental and statistical, will be referred to as hybrid models for the remainder of this document unless stated otherwise.

It is clear that the application of fundamental forecasting models to the short term is a challenging task. However, finding a suitable hybridisation approach that synergistically combines the strengths of both fundamental and statistical approaches is no easier. Such hybridisation approach should lessen the drawbacks and shortcomings of each methodology in order to enhance the overall predictive accuracy. Figure 1.1 displays a summary of the main aspects of fundamental and statistical methodologies, both separately and combined. The advantages and disadvantages are displayed with a white and black background respectively. Ideally, the hybridisation eliminates the drawbacks of each individual approach by merging their strengths. For instance, while fundamental approaches cannot reflect agent strategic behaviours whereas statistical approaches are able to extract such behaviours exhibited in the past. On the other hand, statistical models cannot deal appropriately with punctual events such as power unit decommissioning, which can be easily considered in a fundamental electricity market modelling approach. However, given that power systems are currently facing several structural changes, both modelling methods must be weighed accordingly, and thus these structural breaks must be taken into account in their corresponding forecasting techniques by, for instance, calibrating these models with training windows associated with similar market regime conditions than those of the forecasting period.

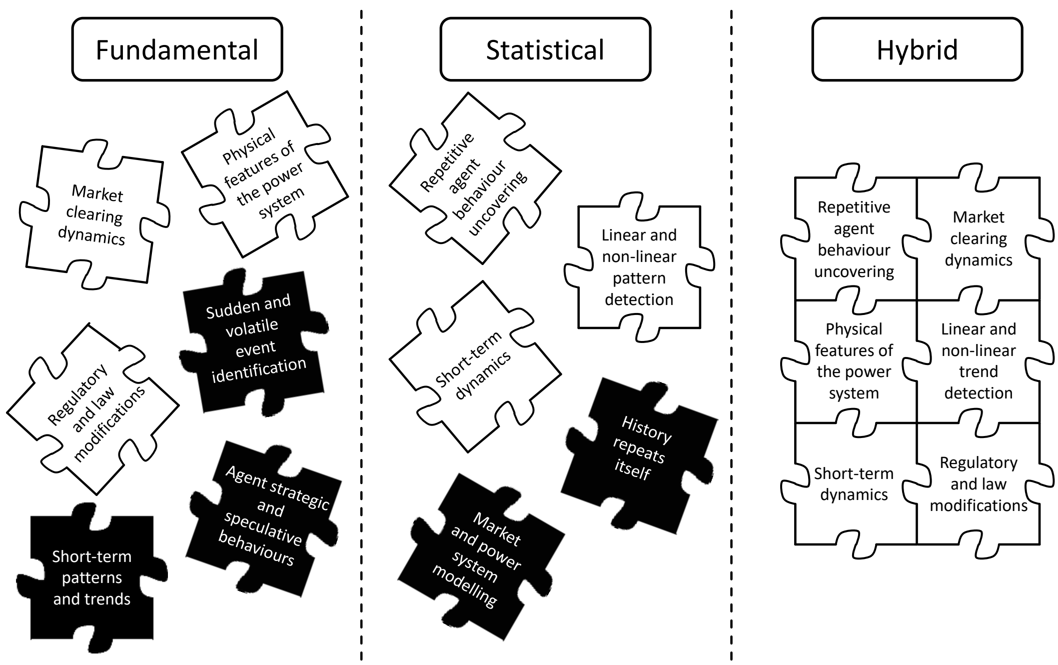

Figure 1.1 Advantages/shortcomings of fundamental and statistical methodologies vs. hybrid modelling

1.3 Objectives

This section contains the objectives to be achieved with this thesis, which are split into a main objective and a succession of specific objectives.

1.3.1 Main objective

The main objective of this thesis is to propose and develop a novel methodology that breaks new ground in the context of electricity market price forecasting models in the short term (short-term planning horizons are considered to be between a few hours and a week) with hourly precision. To this end, a fundamental approach and a statistical technique will be adequately combined so as to take advantage of both methods' strengths, as shown in Figure 1.1. In doing so, the key factors, idiosyncrasies and behaviours behind electricity prices will be identified depending on the state or regime of the power system.

1.3.2 Specific objectives

More specifically, the objectives that have been completed within the scope of the thesis are described below. Said completion has led to the provision of several findings and contributions in the context of short-term electricity market price forecasting.

1.3.2.1 Creation of a short-term fundamental forecasting approach

This objective's focus is centred on the elaboration a suitable fundamental model that is able to take the general fundamental factors of the electricity market into account, i.e. technical/physical and regulatory elements of the power system. Furthermore, an important part of this objective is related to the adaptation of the fundamental approach to short-term contexts where hourly arrangement of the data is the norm, and thus special considerations must be taken with regards to the computational efficiency, resolution times and hourly accuracy of the forecasts.

1.3.2.2 Development of short-term statistical methodologies

The purpose of this objective is twofold. On the one hand, given that an overabundance of training information may cause the statistical forecasting model to overfit, careful attention must be paid to the variables and periods that are used in order to calibrate the statistical models by means of data pre-processing approaches. This also includes the uncovering of patterns and trends in the training dataset that contribute or hamper the forecasting performance. On the other hand, an adequate short-term statistical forecasting model that is able to benefit from the fundamental model's results must be ascertained.

1.3.2.3 Hybridisation of short-term fundamental and statistical modelling

The purpose of this objective is to acquire a suitable fundamental-statistical hybrid methodology and apply it to the short term; and, by doing so, contribute to the short-term electricity market price forecasting literature. The estimated prices of this model should be able to reflect the most relevant drivers of electricity market prices via both fundamental and statistical approaches. On the one hand, fundamental models simulate the market clearing and thus the resulting electricity prices should follow the operation and behaviours of the electricity market, such as law-related limits and unit technical features (e.g. outages, maximum generation capacity, etc.). On the other hand, statistical methods should be capable of lowering the errors of the estimated market clearing prices and incorporate complex short-term and intraday dynamics with the aid of explanatory variables such as expected wind generation and demand as well as relevant behaviours exhibited in the past.

1.3.2.4 Proposal of suitable hybridisation techniques

This objective is focused on the interaction between both ends of the forecasting methodologies of this thesis. That is, the study and analysis of the combination schemes between the fundamental and the statistical models. For instance, the primary hybridisation scheme in the relatively-scarce fundamental-statistical modelling literature involves using the market clearing prices as an additional input to the statistical model. Therefore, this objective considers other ways of linking both methodologies in order to obtain more synergistic combinations and thus increase predictive performance, such as using market clearing prices as a calibration period selection criterion or employing other results from the fundamental model as further input data to the statistical model.

1.3.2.5 Detection and evaluation of electricity market regimes

In the electricity markets of today, sudden and gradual market regime changes are becoming the norm. Moreover, statistical forecasting performance is strongly affected by the presence of market regime switches in the calibration datasets. In addition, a proper prediction of the forecasting period's market conditions proves beneficial when selecting the corresponding training data. These aspects have not been sufficiently touched upon in the current short-term electricity forecasting literature, and therefore this gap must be covered.

1.3.2.6 Empirical evaluation of the proposed approaches

The completion of this objective involves the application of the methodologies that have been proposed and developed throughout this thesis to a recent and full-scale electricity market case study that presents several circumstances and scenarios of the behaviour that electricity market prices may exhibit. Furthermore, other forecasting models that are well established in the electricity price forecasting literature will be used as benchmarks as a way of validating the forecasting performance of the proposed forecasting methodologies of this thesis. More specifically, the case of the Iberian wholesale day-ahead electricity market, is mostly used as the case study throughout the entirety of the thesis. This power exchange is a representative case study given its increasing renewable penetration and the recurrent regulatory changes. Therefore, this case study constitutes a challenge when it comes to electricity price forecasting, and thus it has received considerable attention in the literature.

1.4 Outline and contents of the thesis

This thesis document contains, apart from this introductory chapter, four other chapters that attempt to address the aforementioned objectives. As a summary, chapters two, three and four present the original contributions and findings that have been achieved within the scope of this thesis. Chapter five contains the main conclusions, contributions and proposals of lines for future research. A general overview of the elements addressed in every chapter is shown in Figure 1.2 at the end of this subsection. The remainder of the thesis document consists of the following four chapters.

1.4.1 Chapter 2

This chapter proposes a fundamental-statistical hybrid model that is composed of a cost-production optimisation model (fundamental) and an artificial neural network model (statistical). The proposed fundamental model considers the full technical and physical structure of the Iberian power system, as well as its regulatory constraints. However, in order to reduce computational burden, generation units that shared similar cost functions and technical features were aggregated into larger units. By adequately representing the Iberian electricity market, this fundamental model is able to estimate market clearing prices by computing the dual variable of the demand vs. generation balance equation.

The employed artificial neural network model configuration follows the standard design that is established in the electricity price forecasting literature, which is a single hidden layer and an output layer. The usual predictors were used as input data to the neural network model, such as expected demand and expected wind generation. The hybridisation approach involves using the estimated market clearing prices from the fundamental model as an additional input variable to the neural network model. This hybrid forecasting methodology demonstrated a general superior performance compared to that of well-established methods in the literature, as well as its individual components due to the effect of incorporating market clearing prices to the neural network model which provided forecasts that are more centred in the daily average price level. Although overall accuracy was increased, this also yielded slightly flat predictions with lower volatility when compared to the pure neural network model. Therefore, this calls for an alternative hybridisation method that diminishes this effect on days of higher price volatility.

1.4.2 Chapter 3

The methodologies that are detailed in this chapter attempt to address the shortcomings of the hybrid procedure that is presented in the previous chapter by means of several improvements to the individual forecasting models, as well as new hybridisation schemes. One of the main findings of the previously presented methodology suggested an adjusted synergy between the fundamental and the neural network model due to the possibility of cancelling each other’s drawbacks in order to provide a higher forecasting accuracy.

Therefore, the procedures presented in this chapter constitute several forecast combination schemes between both components of the hybrid methodology. Additionally, other enhancements are proposed regarding the fundamental and statistical techniques, such as an increased level of detail in the fundamental model without leading to a considerable increase in its runtime, the improved estimation of the thermal units' variable costs and the inclusion of a similar-day method[3] prior to the neural network model forecast. The resulting enhanced hybrid methodologies are capable of outperforming the hybrid technique presented in Chapter 2, as well as providing insights as to how the models behave across the year.

1.4.3 Chapter 4

The focus of this chapter, motivated by the conclusions drawn in Chapter 3, is centred on the effect of the evolving market regimes on the predictive performance of the forecasting models. Given that one of the conclusions of Chapter 3 suggests that the performance of the neural network models in days characterised by unstable market conditions may be enhanced if attention is paid to the training dataset window. Following the rationale stated in (Pesaran & Timmermann, 2007), the concept of structural breaks was applied to the context of short-term forecasting of electricity market prices. Although the work presented in (Marcjasz, Serafin, et al., 2018) also proposes the implementation of the theory of structural breaks to this context, there is no approach in the electricity forecasting literature that provides specific guidance criteria regarding the selection of calibration dataset windows in statistical forecasting approaches. Moreover, said selection should be performed in an *ex ante* manner, that is, the guidance criteria must not be based on actual price behaviours at the moment of the forecast.

These issues call for the development of a unique methodology that effectively links the calibration dataset window selection to the market expectations related to the forecasting period, while also considering the market structural breaks that are exhibited in the historical time series of electricity prices. The former is carried out by an hourly clustering procedure regarding exogenous variables whose value is known for the forecasting period and the latter is performed by means of the methodology presented in (Zeileis, Kleiber, et al., 2003). An oversized initial calibration dataset window is filtered via this robust calibration period selection and provided significantly larger calibration datasets than those utilised in Chapter 3. Finally, the advantages of this novel technique are discussed, such as a significantly superior forecasting accuracy across periods with unstable market circumstances and the removal of the need of predefining a calibration period in electricity market price forecasting applications.

[3] The terms "similar-day method" and "similar days method" refer to the same methodology and are used interchangeably throughout this thesis.

1.4.4 Chapter 5

This is the final chapter of the thesis dissertation, which contains a summary of the main conclusions and findings obtained throughout the entire development of the thesis along with the main original contributions. In addition, a proposal of lines for future research is presented.

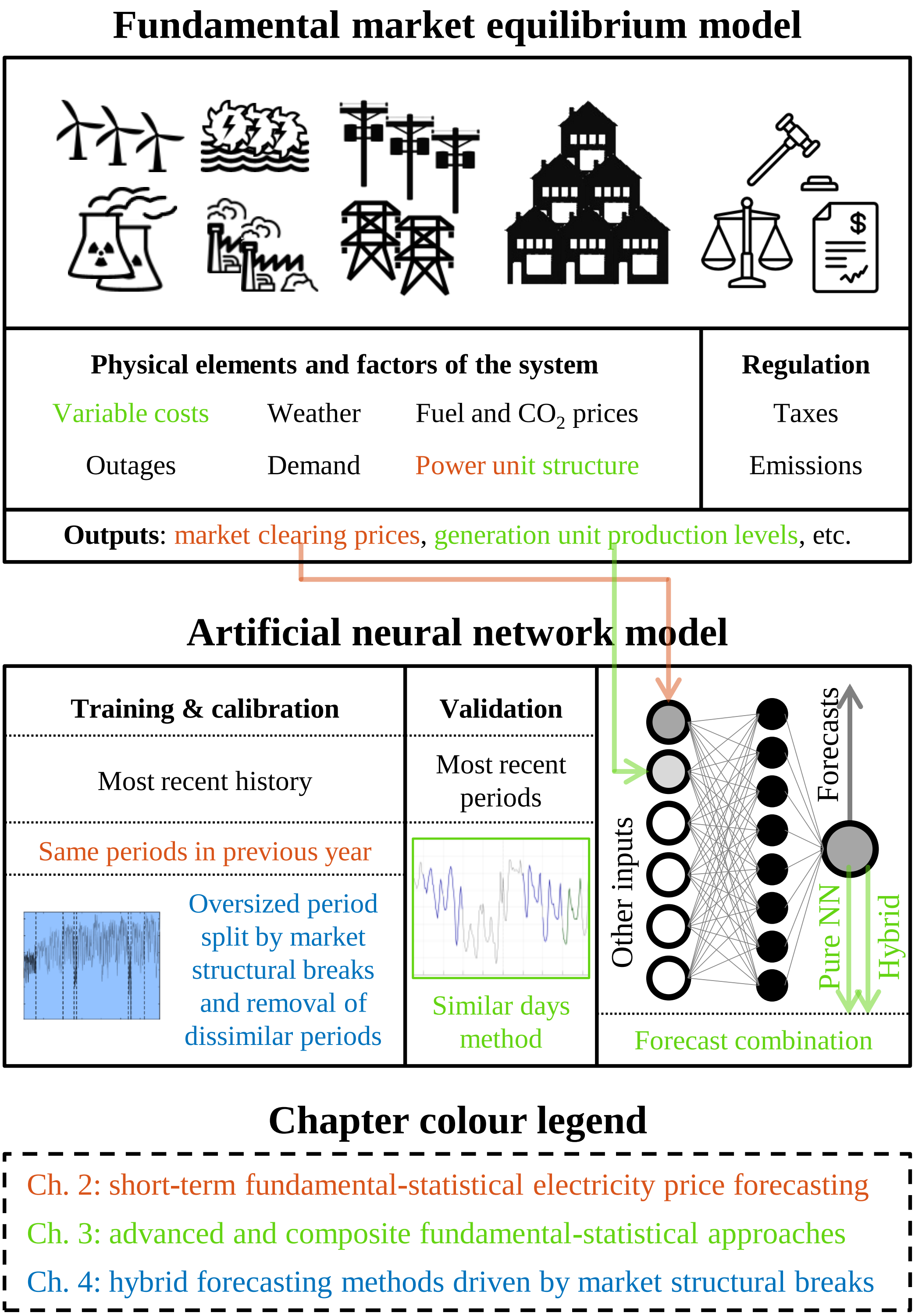

Figure 1.2 Overview of the hybrid electricity price forecasting model of the thesis as well as the targets and enhancements of the proposals of each chapter

Chapter 2

Hybridisation of Fundamental and Statistical Short-Term Electricity Price Forecasting Models

Traders and practitioners in diverse power exchanges are nowadays being most exposed to uncertainty than ever. The combination of several factors such as renewable generation and regulatory changes calls for suitable electricity price forecasting models that can deal with complex and unusual market conditions. Several authors have proposed combining fundamental approaches with statistical models in order to cover all relevant aspects for electricity price forecasting. This combination has shown positive results for medium-term horizons. However, this approach has rarely been carried out for short-term applications. Moreover, several day-to-day applications in electricity markets require fast responsiveness and accurate forecasts. All of these facts encourage the work presented in this chapter[4], which involves a novel short-term hybrid electricity price forecasting model that combines a cost-production optimisation (fundamental) model with an artificial neural network (statistical) model. In order to validate the advantages and contributions of the proposed model, it has been applied to a full-scale power system featuring complex price dynamics: the Iberian electricity market. Furthermore, the recent case studies of late 2016 and the entire year 2017 have been considered. Moreover, its forecasting performance has been compared with those of the two individual components of the hybrid model as well as other well-recognised methods. The results of this comparison prove that the proposed forecasting model of this chapter outperforms the benchmark models, especially in uncommon market circumstances.

[4] The proposed models, developments, findings and results are based on the journal paper (de Marcos, Bello, et al., 2019a)

2.1 Introduction

As mentioned in the previous chapter, traders and practitioners in power exchanges worldwide are currently facing very high competition, which is primarily due to the corresponding liberalisation of electricity systems. Furthermore, there are several sources of risk, such as the intermittent generation of renewable energy sources, as well as ongoing regulatory changes. These facts call for appropriate electricity price forecasting models that are able to deal with the resulting complex and uncommon market circumstances. Accurate price forecasts are highly valuable for decision-making support in day-to-day electricity market operations, which also require short runtimes and swift adaptability.

Consequently, in order to consider the aforementioned combination of factors, some works in the literature proposed combining statistical or econometric techniques with fundamental models. The former methods are highly utilised in short-term contexts, while the latter are more frequent in medium-term scopes. Medium-term applications have proven that one of the major advantages of the unification of these methods is the incorporation of structural and regulatory changes (e.g. decommissioning of thermal units, new subsidies or taxes, etc.) in a natural manner, which compensates for the misconception of the repetition of several price behaviours assumed in statistical models. In other words, the negative effects linked to the "history repeats itself" premise, which is linked to the calibration of statistical forecasting approaches, are, to some extent, nullified by fundamental models when physical and regulatory elements of the market are being altered. The ability to capture these alterations is appealing for market traders and practitioners, especially in today's power exchanges due to the ongoing evolutions that are exhibited in market conditions and circumstances. Nevertheless, these models are not able to represent the main features of short-term price dynamics on their own. This shortcoming is effectively covered by statistical procedures, and thus the hybridisation of both modelling methods is synergistic and advantageous. However, the hybridisation of these techniques has been rarely carried out in short-term applications, and it is most certain that the same assumptions and considerations will not yield a suitable short-term hybrid electricity market price forecasting model. Therefore, the medium-term hybridisation methods may not be completely extrapolated to the short term.

This chapter proposes a novel methodology for short-term electricity market price forecasting, which is based on the combination of a fundamental market equilibrium model and a neural network forecasting approach. The fundamental model, similar to the one proposed in (Bello, Bunn, et al., 2017) for a mid-term scope, considers a perfect competition environment in which total system costs are minimised. The neural network model uses the estimated market clearing prices yielded by the fundamental method as input data, alongside other widely-utilised predictors in statistical forecasting contexts. Every calculation and operation is carried out with hourly precision and forecasting horizons of one day and one week have been considered.

In order to validate the advantages and contributions of this hybrid methodology, it has been applied to a full-scale electricity market featuring several market scenarios and circumstances: the Iberian electricity market. The forecasting performance of this proposal has been compared with those of its two individual components as well as with other well-recognised and established methods in the literature. The contents and results of this chapter are based on the journal publication of (de Marcos, Bello, et al., 2019a), which has been elaborated during the early stages of the development of this thesis.

It should be highlighted that the purposes of this thesis are aimed at the Iberian wholesale day-ahead electricity market, which is a representative case study in which a large number of features and idiosyncrasies of electricity prices can be observed. This is mainly due to the increasing renewable energy penetration and the ongoing and frequent regulatory changes, and thus forecasting the Iberian electricity market prices is highly challenging. Furthermore, its regulation enforces a lower and upper limit to price bids of 0 and 180 €/MWh respectively, and thus negative prices are not allowed. The day-ahead market clearing is performed for every hour of the following day and its price is set to the last or most expensive cleared supply bid. Supply bids pertaining to nuclear and intermittent renewable[5] energy generation technologies are typically cleared by the market, as their prices are significantly lower than the short-running costs of other thermal[6] and head-dependent hydro generation units.

In short, the main contributions of this chapter are described as follows:

1. A novel short-term electricity market price forecasting model is proposed and developed, which consists of a cost-production optimisation fundamental model and a neural network forecasting model. The superior performance of this proposal is justified by means of a comparison with five benchmarking models, including well-established methodologies in the electricity price forecasting literature.

2. The effects and synergies involved in the proposed modelling hybridisation were identified and thoroughly analysed.

3. Rearrangements and adjustments were performed to the input data of both components of the proposed hybrid model in order to decrease

[5] Intermittent renewable generation includes all renewable energy sources (RES) in the Iberian power system except the head-dependent component of hydro generation (i.e. not the river flow or the run-of-the-river hydro generation, but the production yielded by water reservoir turbines).

[6] These thermal generation technologies involve combined cycle gas turbine (CCGT) and coal power units, which are frequently the marginal generation technology in the Iberian electricity market clearing.

computational burden and thus increase efficiency, as well as reducing runtime and overfitting occurrences on the neural network model.

4. The forecasting capabilities of the proposed methodology and the benchmarks are closely examined by testing them on seven particular and insightful case studies of the Iberian electricity market in late 2016. Furthermore, a general analysis is carried out by testing these models on the entire 2017 case of the same market.

The chapter is organised in the following succession of subsections. Section 2.2 contains a state of the art review of the electricity market price forecasting context that encourages motivation for the development of the methodology that is described in Section 2.3. Section 2.4 presents the evaluation criteria and the benchmarking models that, along with the proposed methodology, are applied to the case studies described in Section 2.5. The corresponding results and discussions are contained in Section 2.6. In addition, Section 2.7 presents the conclusions and the main contributions of this chapter, as well as proposals of potential improvements to every aspect of the hybrid methodology introduced in this chapter, some of which are addressed on the remaining chapters of this thesis. Finally, Appendix A presents a preliminary analysis in which the probabilistic forecasting capabilities of the proposed methodologies are evaluated.

2.2 Electricity price forecasting

Going back three decades, the liberalisation of the electricity sector eliminated the traditional government monopolies and led to the introduction of competitive markets. Compared to the late 20th Century, power exchanges worldwide are facing a higher degree of competitiveness and a plethora of risk sources as a result of the increasing renewable penetration, ongoing regulatory changes and global financial instability. Consequently, participants and agents are encouraged to resort to electricity market price forecasting methods. Depending on the assumptions and purposes, the current literature offers a massive range of methodologies, as seen in (Weron, 2014). The most distinguishing element among forecasting techniques is the planning horizon, as short-term applications lean towards the use of statistical methods while longer-term horizons generally involve the fundamental modelling of the electricity market operations and dynamics, as stated in (Bello, Reneses, Muñoz, et al., 2016). Due to the higher frequency of day-to-day electricity market operations, statistical forecasting approaches are most prominent and thus a higher degree of research has been conducted when compared to medium- and long-term modelling procedures, as stated in (Yan & Chowdhury, 2013; Bello, Reneses, Muñoz, et al., 2016).

Within the category of short-term statistical techniques, the most established methods in the literature are based on time-series analysis, which involve linear regression models (Weron & Misiorek, 2006), ARIMA[7] models or transfer functions and its variants (Contreras, Espínola, et al., 2003; Nogales & Conejo, 2006; Cruz, Muñoz, et al., 2011; Sánchez De La Nieta, González, et al., 2016), GARCH[8] procedures (Garcia, Contreras, et al., 2005; Girish, 2016) and functional time series (Aneiros, Vilar, et al., 2013; Liebl, 2013; Portela, Muñoz, et al., 2017). Furthermore, given that several of these methods rely on the normality and stationarity assumptions in electricity price time series, which is becoming less appropriate in today's electricity markets, some authors have recently resorted to apply transformation approaches in electricity price forecasting contexts, such as the classic logarithmic transform, Box-Cox transform, probability integral transform, and mirror-log transform (Uniejewski, Weron, et al., 2017).

Another highly relevant category in short-term statistical forecasting is related to artificial intelligence (AI) techniques, which are mainly artificial neural networks or ANN (Keles, Scelle, et al., 2016; Monteiro, Ramirez-Rosado, et al., 2016; Panapakidis & Dagoumas, 2016) and support vector machine or SVM (Zhao, Dong, et al., 2008; Papadimitriou, Gogas, et al., 2014; Zahid, Ahmed, et al., 2019). Moreover, several authors have opted to combine several of these statistical forecasting models into pure statistical hybrid models such as ARIMA plus SVM (Amjady, 2006; Chaâbane, 2014; Angamuthu, Mukherjee, et al., 2018), as well as adding data pre-processing approaches to the mix such as wavelet decomposition (Conejo, Plazas, et al., 2005; Catalão, Pousinho, et al., 2011; Bento, Pombo, et al., 2018).

By contrast, the medium-term electricity price forecasting modelling literature is not entirely dominated by statistical approaches given that other forecasting methods are most suitable in longer planning horizons. A relevant category that falls in medium-term contexts are market agent models, which are aimed at the analysis of the interactions between market participants. This is mainly carried out by means of market equilibrium models, which derive from game theory principles such as Cournot[9] competition (Barquin & Vazquez, 2008; Weigt & Willems, 2012) or Bertrand[10] competition models (Federico & Rahman, 2003).

[7] ARIMA models involve a combination of three elements: autoregressive (dependency on past values), integrated (non-stationarities that may be corrected by differentiating the time-series) and moving-average (correction of current error with a weighted average of previous errors).

[8] GARCH stands for generalised autoregressive conditional heteroscedasticity, which considers a time series with a stochastic error variance that follows an ARMA model

[9] Cournot-Nash competition considers a market equilibrium in which agents, who are identical among each other, must decide their production quantities independently of each other, simultaneously and rationally.

[10] In contrast to Cournot models, market agents compete in prices instead of quantities under Bernard competition assumptions.

Furthermore, perfect competition may be assumed, as done in (Bello, Reneses, Muñoz, et al., 2016), although intermediate conditions between Cournot oligopoly and perfect competition can be assumed by the conjectural variations method (Barquín, Centeno, et al., 2005), which involves a parameter that measures the market power of the companies and represents the change in the electricity market price. This parameter may be estimated by using historical data, as proven in (Díaz, Villar, et al., 2010) or by means of other fundamental methods or in endogenous manners. Moreover, supply function equilibrium competition models consider a combination of Bertrand and Cournot models, given that prices and quantities are simultaneously ascertained by means of a set of differential equations (Hobbs, Metzler, et al., 2000; Ruiz & Conejo, 2009).

Other equilibrium models are focused on alternative aspects of the market. This is the case of the fundamental or structural models, which are targeted at a highly detailed representation of physical (e.g. transmission lines, generation units, etc.) and regulatory elements (e.g. taxes, CO_2 emissions, etc.) of the power exchange (Carmona & Coulon, 2014). In these models, the market clearing according to said elements is simulated in order to estimate the electricity prices. While these models are capable of considering market regimes in their forecasts, a great amount of information and computational burden is required so as to provide as much accuracy as possible. Several applications of fundamental models have considered trading accuracy for computation speed through simplification methods, for instance, hour aggregation according to load levels or system states (Wogrin, Dueñas, et al., 2014; Bello, Reneses, Muñoz, et al., 2016). Another way of reducing temporal information involves the representative periods method, which consists of selecting certain periods (e.g. days, weeks, etc.) that are representative of the different scenarios that may happen throughout the forecasting horizon, where every period is linked to their corresponding representative period (Tejada-Arango, Domeshek, et al., 2018). However, as mentioned in (Tejada-Arango, Domeshek, et al., 2018), these temporal representation methods do not fully preserve the original chronology, which poses no significant repercussions in medium- or long-term contexts. However, short-term applications demand an accurate chronology with the highest level of granularity.

Other means of reducing complexity in the model consist in aggregating generation units in order to consider a single power unit per generation technology (González, Contreras, et al., 2012), although this entails a perfect competition model in order to remove ownership and strategic constraints. Furthermore, as discussed in (Bello, Bunn, et al., 2017), the estimations yielded by these models are somewhat flat and lack the intraday effects that are exhibited in the day-ahead market. These facts have frequently encouraged the coupling of fundamental methods with statistical methods by using results from the former models as input to the latter procedures, as remarked in (Weron, 2014).

As mentioned previously, statistical models are more common in short-term application, although several works have attempted to apply these methods to the medium and long term. According to (García-Martos, Rodríguez, et al., 2011), accuracy tends to decay when lengthening horizons of the forecasting models that perform suitably in the short term. A common approach involves adapting data and calculations to daily, monthly and even yearly arrangement instead of hourly precision, as seen in (Torbaghan, Motamedi, et al., 2012; Azadeh, Moghaddam, et al., 2013; de Marcos, Reneses, et al., 2016). Alternatively, there are a few works in the literature that maintain hourly resolution in longer-term forecasting applications (Yan & Chowdhury, 2014; Maciejowska & Weron, 2016).

However, the adaptation of medium-term typical methods, such as fundamental models, to short-term contexts is even more unusual. The hybrid fundamental-statistical model presented in (González, Contreras, et al., 2012) is one of the first of these unique works in the literature, which has been applied to the 2008 case of the day-ahead APX electricity market of Great Britain. However, daily average prices are forecasted for a horizon of one-month and thus this work does not fully agree with the definition of short term in this thesis[11]. In order to overcome the computational challenges of fundamental modelling, all power units were stacked together according to their generation technology. The estimated electricity prices[12] are obtained by means of a simplified bidding curve that establishes a linear combination of technology variable costs weighted by their generation volumes. Specific fuel prices, such as NBP gas index and Brent crude oil index, were used so as to ascertain said costs. Furthermore, the interconnection with France was also considered, although its volume approximately represents a 4% share of the total U.K. generation profile in the considered case study. The estimated system costs were used as an additional explanatory variable to both a linear and a non-linear regression technique: an ARX[13] and an LSTR[14] model respectively. Although this work proves that these fundamental-statistical hybridisations yield superior day-ahead price forecasts, it would be interesting to extend the fundamental methodology by taking advantage of other elements in the market clearing process, such as the technical features of generation units, regulatory considerations (e.g. taxes and subsidies), etc. Furthermore, AI techniques such as neural networks are able to consider both linear and non-linear trends and may prove useful if the estimated market clearing prices are used as an extra input variable in these models.

[11] As mentioned in Chapter 1, this thesis considers short term horizons from a few hours up to a week ahead of time with hourly resolution.

[12] Also known as “system costs” in (González, Contreras, et al., 2012) and also “marginal costs” in other manuscripts pertaining to similar frameworks.

[13] ARX involves an autoregressive (AR) model with explanatory or exogenous variables.

[14] LSTR stands for logistic smooth transition regression, which considers gradual regime-switching by means of a nonlinear regression model, often applied to electricity prices distinguishing between normal and extreme price regimes (high and/or low prices).

Considering the above facts and issues, the methodology proposed in this chapter attempts to cover the insufficiencies and scarcities in the context of short-term electricity market price forecasting by means of a unique methodology that involves a fundamental-statistical approach. More specifically, this methodology is based on a cost-production optimisation model that is linked to a neural network model. The computational efficiency of each model is enhanced individually in order to adapt the resulting hybrid forecasting model to the short-term electricity market price forecasting contexts.

2.3 Methodology

This section contains a full description of the proposed methodology, whose objective is the creation of a novel short-term hybrid electricity market price forecasting model and put it to the test on a real, full-scale and complex electricity market, such as the Iberian (Spanish and Portuguese) power exchange. Forecasts have been considered for both one day and one week horizons so as to agree with the two ends of the short-term framework horizons. A diagram of the proposed fundamental-statistical model is depicted in Figure 2.1, which runs from left to right.

In short, electricity market prices are first estimated by means of a fundamental model (see subsection 2.3.1) that is applied to the case of the Iberian power exchange. The simulation of the market clearing considers thermal generation unit characteristics (maximum power output, outages, production costs, etc.), renewable generation (hydro reservoirs, estimated wind and solar generation, etc.), regulatory constraints (CO_2 emissions, generation taxes, etc.) and interconnections. Finally, these market clearing prices are used alongside other traditional predictors in a neural network forecasting approach (presented in subsection 2.3.2) in order to correct the errors from the fundamental model and

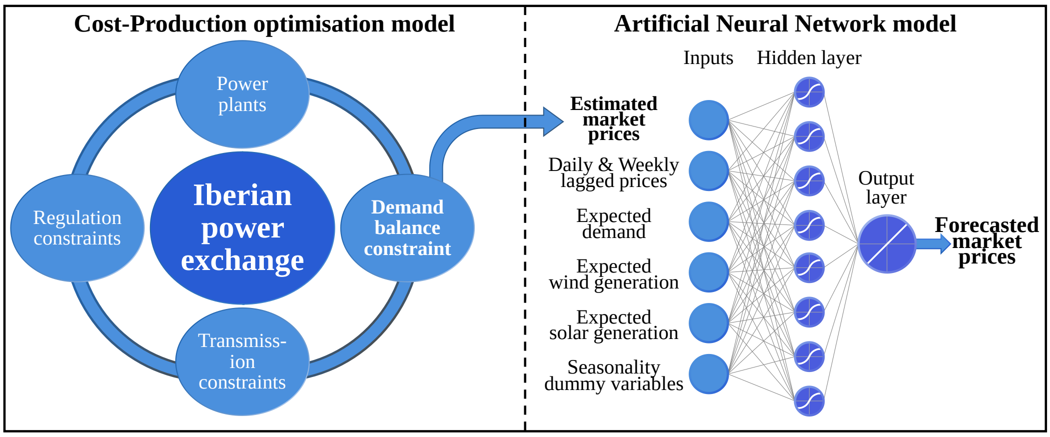

Figure 2.1: Overview of the proposed hybrid forecasting model

provide short-term and intraday dynamics in the final forecast. It is important to highlight that both approaches are run with the same frequency and forecasting horizon.

2.3.1 Fundamental modelling approach

The first phase of the model, displayed at the left part of Figure 2.1, is composed of a cost-production optimisation model whose purpose is the simulation of the market-clearing process by minimising the total system costs, which are constrained by generation unit technical attributes, regulatory rules, transmission and interconnection interactions and the energy equation which links generation volumes to the energy required by the demand. Therefore, in this model, the estimated electricity market price may be obtained as the dual variable of the demand balance constraint, as done in (Bello, Reneses, Muñoz, et al., 2016).

In order to effectively adapt the corresponding optimisation model to hourly precision without leading to high computational burden issues, similar power plants that share identical cost functions and other technical features were aggregated into larger ones. As a result, the optimisation problem size was significantly reduced and thus the estimation time of electricity market prices was reduced from a few minutes to a few seconds. More specifically, the simplified optimisation problem of the Iberian power system case for a forecasting period of one week consists of 12,440 equations and 71,024 variables with a total runtime of 0.91 seconds and a maximum RAM usage of 76 MB on a 64-bit Windows 7 PC with 16 GB installed RAM and the following processor: Intel® Core™ i7-3770 CPU@ 3.40 GHz of 4 cores and 8 logical processors.

The decision variables of this fundamental modelling approach involve production quantities of the non-intermittent generation technologies (i.e. thermal and head-dependent hydro production units). However, only the estimated market clearing prices were used for the second phase of the proposed hybrid methodology. Nevertheless, the possibility of incorporating additional output variables from the fundamental model to the statistical forecasting approach is left for future consideration in the remaining chapters of this thesis. Moreover, unlike the fundamental-statistical methodology of (Bello, Reneses, Muñoz, et al., 2016), only costs are considered in the optimisation objective function, thus ignoring the agent strategic behaviour term, defined as the "conjectured-price response". This term was discarded due to the suitable capability of short-term statistical approaches in detecting repetitive and complex agent strategic trends that are exhibited in the recent past.

The following subsections thoroughly present the main features of the Iberian power system and the optimisation model's equations.

2.3.1.1 Main characteristics of the Iberian power system

The Iberian electricity market, also known as MIBEL (*Mercado Ibérico de Electricidad(e)*[15]), is located at the Iberian Peninsula (Spain and Portugal). On the one hand, the spot electricity market operator is managed by the Spanish company *OMI-Polo Español S.A.* (OMIE), which involves a day-ahead market followed by six intraday auctions. On the other hand, the MIBEL derivatives exchange is supervised by the Portuguese company *Operador do Mercado Ibérico de Energia Pólo Português, S.G.M.R., S.A.* (OMIP). Although the market operator is the same in both countries (OMIE), the transmission system operator is different: REE (*Red Eléctrica de España*) in Spain and REN (*Redes Energéticas Nacionais*) in Portugal. Moreover, the Iberian electricity market[16] is interconnected to the French and the Moroccan power systems, which makes it somewhat islanded compared to other European markets that feature higher degrees of interconnection, such as Germany and the Netherlands.

The regulatory elements of the Iberian electricity market that are included[17] in the day-ahead fundamental model of this thesis are the following: a 7% generation tax that is imposed to generation units located in Spain[18], the lower and upper limits to bids of 0 €/MWh and 180 €/MWh respectively, a forced generation of nuclear power plants and the hydrocarbon or green tax[19]. The Iberian energy mix during 2017 is shown in Table 2.1 with a total installed capacity of 104 GW. However, this electricity system features a considerable amount of overcapacity, as the 2017 hourly demand never reached values over 50 GW. Therefore, the occurrences of extremely high prices are considerably less common than those of extremely low prices.

2.3.1.2 Model constraints and attributes

The fundamental model that has been utilised in this chapter's proposed hybrid methodology consists of a traditional market equilibrium model that is solved by

[15] The word for electricity in Spanish and Portuguese are *electricidad* and *electricidade* respectively.

[16] The Balearic Islands are also considered part of the Iberian electricity market, although special rules apply to them that are not considered in the fundamental model.

[17] Other regulatory aspects that are not included due to the modelling nature of the fundamental model (e.g. perfect competition, short-term forecasting horizon, etc.) involve a canon on hydro generation and a yearly CO_2 emission limit per Generation Company.

[18] In recent history, this tax is set to 7% of production-related income, although in some months of 2019 this tax was temporarily removed. Due to the fact that production incomes or revenues are unknown, this effect is incorporated by multiplying Spanish thermal fuel costs by 1 over 0.93.

[19] Known in Spanish as *Impuesto Especial sobre Hidrocarburos* or *Céntimo Verde*.

Generation technology	Installed capacity (MW)	Percentage of total capacity
Hydro generation	17,032	16.4%
Hydro pumping	3,329	3.2%
Nuclear energy	7,117	6.8%
Coal energy	10,004	9.6%
Fuel/gas energy	2,490	2.4%
Combined cycle gas turbine (CCGT)	26,670	25.6%
Wind energy	23,132	22.2%
Photovoltaic solar	4,687	4.5%
Thermosolar	2,304	2.2%
Other renewables	858	0.8%
Cogeneration	5,828	5.6%
Non-renewable waste	497	0.5%
Renewable waste	162	0.2%

Table 2.1: Installed capacity of the Iberian electricity system in the year 2017

minimising the costs of each market agent. In order to carry out the generation unit aggregation regardless of their ownership, perfect competition has been assumed, and thus total system costs are simultaneously minimised in its objective function. Taking into account that the main decision variables are the production quantities of each generation unit q_i, the following equation represents the model's objective function:

$$\min_{q_i} \sum_i c_i + ic \tag{2.1}$$

According to the generation technology of power unit i, the cost function c_i involves the following costs: fuel, start-up, CO_2 emission and maintenance costs. Furthermore, interconnections are considered in the variable ic. Apart from the corresponding technical and physical constraints of the system's generation units (e.g. maximum/minimum power output, planned maintenance periods, etc.), one of the relevant elements of this optimisation problem is the demand vs. generation balance equation (for every hour t, including interconnection interactions and energy not served, iq_t and ens_t respectively):

$$\sum_i q_{i,t} + iq_t = D_t - ens_t \qquad : \lambda_t \quad \forall t \tag{2.2}$$

Each unit's production quantities, i.e. $q_{i,t}$, are mainly limited by this constraint, and therefore its dual variable can be considered as the system's marginal cost λ_t, which represents the market clearing price under perfect competition assumptions that will be utilised in the remaining steps of the proposed hybrid forecasting methodology. As mentioned earlier, production quantities are the main decision variables of this model. Due to the fact that thermal power units are unable to produce under a certain threshold, additional decision variables must be taken into account, which are related to the commitment of the power unit. Given the nature of these decision variables (i.e. production levels and commitment), mixed integer programming[20] optimisation is typically used to solve this kind of problems. However, the resulting market clearing prices would only reflect the variable costs of the committed units. In order to account for all the generation units' costs when simulating the market clearing, a relaxed mixed integer programming[21] is chosen as the solving method of the optimisation problem. Even if the resulting generation unit schedule may not be fully feasible in practice, this poses no repercussions to the objectives of this work, which are targeted at the estimation of electricity market prices. The optimisation model has been run with the CPLEX solver (version 12.5.1.0).

The parameters, i.e. input data, of this optimisation model are listed below:

- $A_{i,t}$: expected availability (or reversed outage) for (thermal) power unit i and hour t
- CT_i: tonnes of CO_2 emitted per unit of volume produced for (thermal) power unit i
- CP_t: penalty per tonne of CO_2 emitted at hour t (i.e. CO_2 emission allowance price)
- D_t: expected system demand at hour t
- EC_t: expected interconnection export capacity at hour t
- EP_t: expected negotiated interconnection export price at hour t
- $ES_{i,t}$: expected solar generation for (solar) power unit i at hour t
- $EW_{i,t}$: expected wind generation for (wind) power unit i at hour t
- $FP_{i,t}$: fuel cost per unit of volume produced for (thermal) power unit i and hour t
- $GT_{i,t}$: generation tax[22] of (thermal) power unit i at hour t
- HT_i: hydrocarbon tax for (CCGT or coal) power unit i

[20] Mixed integer programming, also known as MIP or MILP, considers discrete variables, such as integer or binary variables

[21] Relaxed MIP or RMIP ignores the corresponding integrality constraints that ensure discrete values on integer and binary variables. Therefore, the RMIP solving time is generally lower than that of MIP models.

[22] This generation tax only applies to generation units located in Spain.

- $I_{i,t}$: expected hydro inflow for (hydro) power unit i and hour t
- IC_t: expected interconnection import capacity at hour t
- IP_t: expected negotiated interconnection import price at hour t
- M_i: maintenance costs per unit of volume produced for (thermal or wind) power unit i
- $\underline{P_i}$: minimum power output of (thermal) power unit i
- $\overline{P_i}$: maximum power output of (thermal) power unit i
- $R_{i,0}$: initial hydro reservoir level of (hydro) power unit i
- $\underline{R_{i,t}}$: minimum hydro reservoir level of (hydro) power unit i at hour t
- $\overline{R_{i,t}}$: maximum hydro reservoir level of (hydro) power unit i at hour t
- S_i: cost per start-up operation for (thermal) power unit i

Many of these parameters are effortlessly obtainable from the transparency platforms of the Spanish System Operator (e·sios, *Sistema de Información del Operador del Sistema*: https://www.esios.ree.es/en) and of the ENTSO-E (European Network of Transmission System Operators for Electricity: https://transparency.entsoe.eu). Commodity futures contracts were used as proxies for coal and CCGT fuel costs (API2 and NBP indexes, respectively). These are available at the financial platform of the International Commodity Exchange (ICE: https://www.theice.com). Moreover, CO_2 emission allowances futures are available at the European Energy Exchange platform (EEX, https://www.eex.com/en).

All generation units whose production quantities are being calculated may belong to the thermal power unit set (T: nuclear, CCGT and coal), the hydro power unit set (H) or the wind power unit set (W). Other renewable energy generation units such as solar, thermosolar and cogeneration are not physically considered and their expected production is a known parameter that is added to the term on the left-hand side of Equation (2.2).

The decision variables involved in this optimisation model are as follows:

- c_i: total costs of power unit i
- cc_i: CO_2 emission costs of (thermal) power unit i
- ch_i: costs derived from the hydrocarbon tax applied to (CCGT or coal) power unit i
- cf_i: fuel costs of (thermal) power unit i

- cm_i: maintenance costs of (thermal or wind) power unit i
- cs_i: start-up costs of (thermal) power unit i
- ens: costs associated with unsupplied energy or "energy not served"
- ic: total interconnection costs (or income if negative)
- ii_t: interconnection import volume at hour t
- ie_t: interconnection export volume at hour t
- iq_t: total interconnection incoming volume at hour t
- $p_{i,t}$: pumped energy of (hydro) power unit i at hour t
- $q_{i,t}$: production of power unit i at hour t
- $r_{i,t}$: reservoir energy level of (hydro) power unit i at hour t
- $s_{i,t}$: energy spillages of (hydro or wind) power unit i at hour t
- $u_{i,t}$: commitment of (thermal) power unit i at hour t (active or inactive)
- $y_{i,t}$: start-up transient state of (thermal) power unit i at hour t
- $z_{i,t}$: shut-down transient state of (thermal) power unit i at hour t

The following equations and constraints are implemented in the optimisation model:

- Equation (2.1): objective function which minimises system costs in general
- Equation (2.2): the balance of production plus imports equals demand plus exports
- Equation (2.3)-Equation (2.8): wind and thermal unit variable costs
- Equation (2.9)-Equation (2.12): interconnection interactions and costs
- Equation (2.13) & Equation (2.14): thermal unit production limits and availability bound
- Equation (2.14): enforced nuclear power plant production by law
- Equation (2.15): wind production and spillages
- Equation (2.16) & Equation (2.17): hydro unit energy balance and production/pumping

- Equation (2.18) & Equation (2.19): thermal unit commitment logic

The following equation formulations, aside from Equation (2.2), represent the operation and dynamics of the power system:

$$c_i = cc_i + cf_i + ch_i + cm_i + cs_i \qquad \forall i \in \{T, W\} \tag{2.3}$$

$$cc_i = \sum_t q_{i,t} \cdot CT_i \cdot CP_t \qquad \forall i \in T \tag{2.4}$$

$$cf_i = \sum_t q_{i,t} \cdot \frac{FP_{i,t}}{1 - GT_{i,t}} \qquad \forall i \in T \tag{2.5}$$

$$ch_i = \sum_t q_{i,t} \cdot HT_i \qquad \forall i \in T(coal, CCGT) \tag{2.6}$$

$$cm_i = \sum_t y_{i,t} \cdot M_i \qquad \forall i \in \{T, W\} \tag{2.7}$$

$$cs_i = \sum_t y_{i,t} \cdot S_i \qquad \forall i \in T \tag{2.8}$$

$$ic = ie_t \cdot EP_t - ii_t \cdot IC_t \qquad \forall t \tag{2.9}$$

$$0 \le ie_t \le EC_t \qquad \forall t \tag{2.10}$$

$$0 \le ii_t \le IC_t \qquad \forall t \tag{2.11}$$

$$iq_t = ii_t - ie_t \qquad \forall t \tag{2.12}$$

$$\underline{P_i} \cdot A_{i,t} \le q_{i,t} \le \overline{P_i} \cdot A_{i,t} \qquad \forall i \in T, t \tag{2.13}$$

$$q_{i,t} = \overline{P_i} \cdot A_{i,t} \qquad \forall i \in T(nuclear), t \tag{2.14}$$

$$q_{i,t} = EW_{i,t} - s_{i,t} \qquad \forall i \in W, t \tag{2.15}$$

$$r_{i,t} = r_{i,t-1} + I_{i,t} - q_{i,t} - s_{i,t} + p_{i,t} \qquad \forall i \in H, t > 1 \tag{2.16}$$

$$r_{i,t} = R_{i,0} + I_{i,t} - q_{i,t} - s_{i,t} + p_{i,t} \qquad \forall i \in H, t = 1 \tag{2.17}$$

$$y_{i,t} - z_{i,t} = u_{i,t} - u_{i,t-1} \qquad \forall i \in T, t > 1 \tag{2.18}$$

$$y_{i,t} - z_{i,t} = u_{i,t} - U_{i,0} \qquad \forall i \in T, t = 1 \tag{2.19}$$

It is worth noting that, thanks to Equation (2.16)-Equation (2.19), the estimated price from the fundamental model may reflect, to some extent, the information of daily lagged prices. These chronological or inter-temporal constraints link current values to others in the past according to unit commitment and hydro reserve balance, so it would be natural to assume that the current estimated price is affected by what has happened in the past.

In general, the set of equations that are presented in this fundamental model represent the main relationships that are present in the market clearing. Therefore, any kind of event that affects the physical or regulatory elements of the power system will be appropriately observed in its price forecasts. For instance, generation unit outages, tax modifications (i.e. generation and/or hydrocarbon taxes) and fuel price variations. Such occurrences have often taken place during the past few years in the Iberian power system and thus in the future they are expected to be as much or more relevant than before.

2.3.2 Statistical modelling approach

As shown in Figure 2.1, the estimated market price from the fundamental model is used as an additional input variable to the neural network model, which is trained alongside other inputs in order to produce short-term electricity price forecasts. The other input variables that were utilised are as follows:

- Expected values of demand, wind and solar generation in the Iberian electricity system.
- Lagged electricity market prices. The following lags were employed:
    - One day
    - Two days
    - One week
    - Two weeks

- Dummy variables indicating if a day belongs to one of the following day types[23]:
    - Saturdays
    - Sundays and holidays

These variables are available at the transparency platform of the Spanish System Operator (https://www.esios.ree.es/en), and their relationships with actual electricity market prices are adequately handled by neural network models (Cruz, Muñoz, et al., 2011) and are highly relevant in the Iberian electricity system, as recently observed in (Aineto, Iranzo-Sánchez, et al., 2019). In order to consider several levels of autoregression and seasonalities, four lagged prices were taken into account with the following delays: one day, two days, one week and two weeks. It is worth mentioning that no rolling-window forecasts are performed with the neural network models, and thus the minimum lag that can be considered is of one day given that the shortest considered horizon is one day.

It is also important to highlight that the same forecasts of demand, wind and solar generation in the Iberian power system have also been used in the fundamental model, and these affect the values of other unit production volumes (e.g. coal and CCGT) via the demand balance constraint. However, the modelling approaches of neural networks are entirely different, as these take into account linear and non-linear trends as well as statistical patterns between electricity prices and the previously mentioned forecasts. Therefore, even if the same data were used in both models, these were not handled and treated alike.

The neural network model structure and configuration are shown in the right-hand side of Figure 2.1, which involves the aforementioned inputs (including the estimated electricity market clearing prices from the fundamental model) that are used as input data in order to provide the final forecast. The following subsections describe in detail the data pre-processing methods that have been applied to this input data and the neural network forecasting approach.

2.3.2.1 Data arrangement and pre-processing methods

Given that statistical approaches are trained upon historical data, it is no surprise that selection and pre-processing methods are applied to said data before employing them in the calibration procedures of the forecasting models. In this case, the training data, which consists of ten variables, were modified with the intention of increasing efficiency as well as reducing the usual overfitting occurrences in neural networks. The training, validation and forecasting periods have been organised according to the timeline that is depicted in Figure 2.2.

[23] Regular Mondays to Fridays (not holidays) are considered when both of these dummy variables are false.

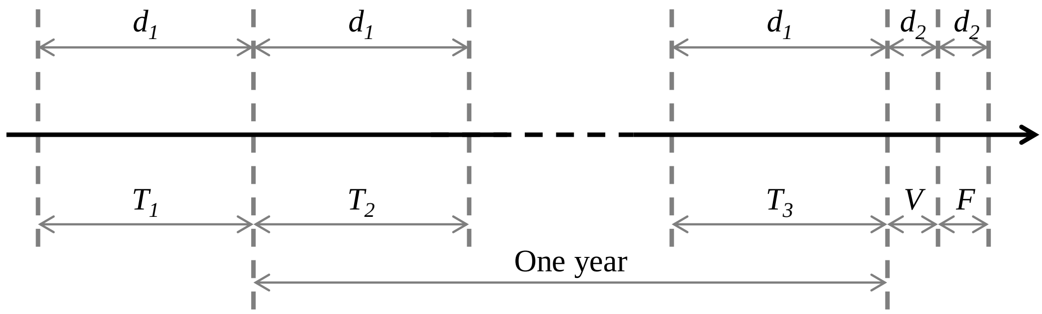

Figure 2.2: Training, validation and test/forecast periods arrangement

The training set consists of three time spans of d_1 days each: T_1, T_2 and T_3. The last one, T_3, is placed immediately prior to the validation period V and contains the most updated information for the neural network to train on. This information is reinforced by the data pertaining to T_1, which is located a year before T_3, and thus presents similar conditions provided that yearly seasonal behaviour is not altered. In the same sense, the data belonging to T_2 may reflect a possible evolution of electricity prices and all related factors, which is what happened a year before the forecasting period. This allows the neural network training algorithm to incorporate seasonality dynamics on its forecasts and thus enhance adaptiveness. Finally, as shown in Figure 2.2, the validation period V is placed right before the forecasting period F and both are set to be of equal length, which is d_2 days (i.e. the forecasting horizon).

This training set arrangement is more efficient than one entirely placed before the forecasting period. For instance, forecasting winter prices when including summer data in the calibration dataset may prove counterproductive. Furthermore, by splitting the training dataset as per Figure 2.2, these data contain more trends that better reflect electricity price behaviours in the forecasting period F.

Not only the arrangement of the input data was carefully analysed and modified, but also an elaborate test was carried out in order to assess variable importance. Additionally, it would be useful to increase the parsimoniousness[24] of the model so as to enhance efficiency and reduce potential overfitting occurrences in the neural network forecasting procedure.

To this end, the ten previously listed input variables were tested in a backward-elimination manner, i.e. evaluating all ten factors at once and discarding one by one the most noisy and redundant. Although this may effectively reduce any redundancy and irrelevancy in the input dataset, this procedure is more straightforward in linear regression applications by means of least squares methods such as LASSO (Tibshirani, 1996) and ridge regression (Hoerl & Kennard, 1970), which were applied to electricity price forecasting contexts in

[24] A parsimonious methodology presents a reduced complexity when compared to others, and parsimoniousness is usually sought with the intention to enhance its overall performance.

(Uniejewski, Nowotarski, et al., 2016). Moreover, the work of (Muñoz & Czernichow, 1998) proposes a sensitivity analysis that involves calculating the derivative of the output with respect to every input after performing the neural network training algorithm. If the values of the neuron weights and their activation functions are known, it is possible to link the output with every input with a function that is possible to differentiate. However, this method has not been utilised in this application due to the fact that it does not provide a way to consider interaction[25] between the neural network inputs. Therefore, the validation set mean-square error (MSE) was used as variable selection criterion, which is available at the end of the neural network training procedure. The MSE is calculated as per the following formula with the conventional notation ($\widehat{Y}_i$ represent the forecasted values for a certain period of N hours, whereas Y_i are the real values pertaining to the same period):

$$MSE = \frac{1}{N} \cdot \sum_{i=1}^{N} \left(\widehat{Y}_i - Y_i \right)^2 \tag{2.20}$$

The backward-elimination procedure is displayed in Figure 2.3. First of all, the neural network validation forecast MSE was calculated considering all 10 input variables. Then, the process was repeated for every combination of 9 input variables and the resulting MSE was compared to the one obtained with 10 variables. If discarding any variable yielded an MSE reduction, then the variable that led to the highest error reduction was discarded and the process was repeated

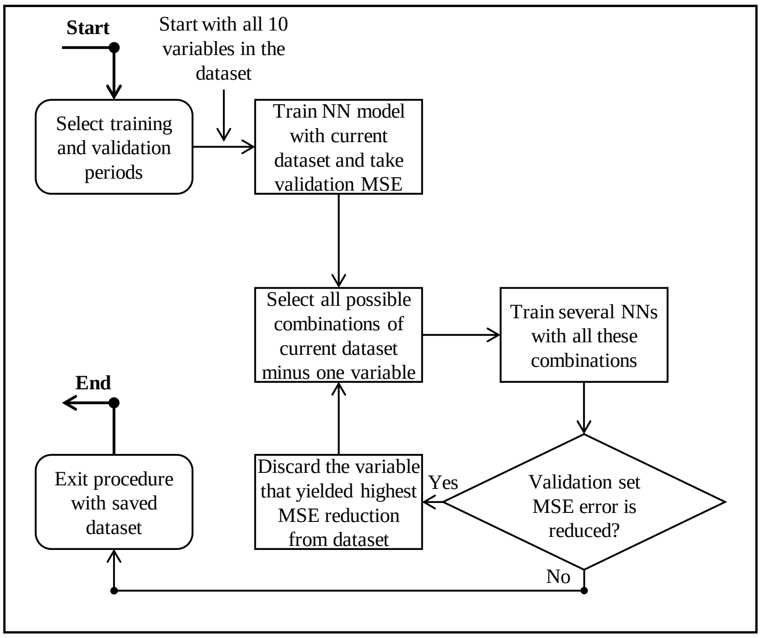

Figure 2.3: Flow diagram of the variable selection method

[25] Interaction between variables in this context is linked to the redundancy of the information yielded by them. If several variables are highly correlated to electricity prices, it does not necessarily imply that using all variables as input data is not redundant.

for every combination of 8 input variables ignoring the discarded variable. If the MSE could not be further reduced by removing one additional variable, the test reached its end and thus all remaining variables were considered for the NN forecast. This procedure was carried out for several days prior to the case studies and the most frequent dataset did not contain the Saturday dummy variable and the two-day lagged electricity prices.

The estimated market clearing prices from the fundamental model thus proved more useful than these discarded variables, which suggests that the underlying information within these prices is useful to the hybrid model, such as coal unit operation costs, CO_2 emission allowances, maintenance schedules, etc. Moreover, elastic nets, which is a combination of LASSO and Ridge Regression proposed in (Zou & Hastie, 2005), were also employed with the same intention, although this method is more appropriate in linear regression applications, and the results were similar for a certain tolerance level (i.e. elastic net parameters). However the backward-elimination procedure of Figure 2.3 is a non-parametric approach and thus does not require further studies in order to ascertain additional information.

The employed data pre-processing methods reduced the total volume of data that is used as input data to the NN model, reducing the NN overfitting occurrences and thus increasing efficiency.

2.3.2.2 Neural network forecasting model

The neural network model that is depicted in Figure 2.4 shows its configuration, which is comprised of a hidden layer and an output layer. According to (Bello, Reneses, & Muñoz, 2016), experience shows that one hidden layer is suitable for most applications in electricity price forecasting contexts. The hyperbolic tangent sigmoid was utilised as the activation function of the hidden layer's neurons,

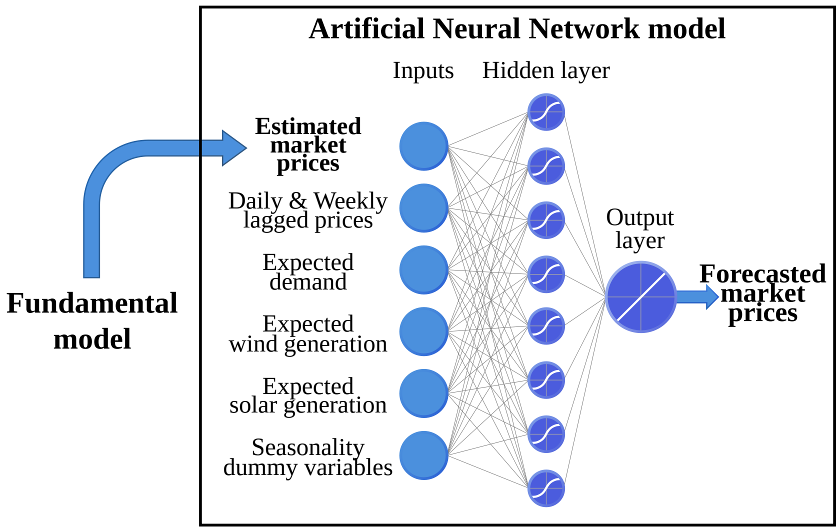

Figure 2.4: Structure of the NN model

whereas a pure linear transfer function was selected for the output layer.

The neural network was trained according to the standard Levenberg-Marquardt algorithm, explained in (Moré, 1978), which is one of the most popular neural network training methods employed in electricity market price forecasting applications, such as (Amjady, 2007; Catalão, Mariano, et al., 2007). Regarding the number of neurons in the hidden layer, some works, such as (Monteiro, Fernandez-Jimenez, et al., 2015; Dudek, 2016; Sandhu, Fang, et al., 2016) propose using $2 \cdot n + 1$ neurons on the hidden layer, where n is the number of input variables. Authors using this criterion claim that this configuration enables the neural network model to fit any finite input-output problems. Other works, such as (Cipriano, Lira, et al., 2009; Bento, Pombo, et al., 2018), ignore this criterion and attempt to ascertain the optimal number of hidden layer neurons. However, given that neural networks are used in a massive number of contexts, there is no specific criterion in the literature as to how many neurons should be used in the hidden layer and thus the number of neurons should not be predefined. Furthermore, according to some authors, such as those proposing the works of (Cruz, Muñoz, et al., 2011; Bello, Reneses, & Muñoz, 2016), a fixed number of neurons in the hidden layer should not be chosen due to the inherent randomness of the neural network training procedures and algorithms.

As explained earlier, the neural network datasets are divided into three subsets: training, validation and test or forecast. These data are standardised[26] in order to avoid issues that may arise when several orders of magnitude are present. The standard optimisation problem behind neural network training algorithms attempts to minimise the validation set MSE by modifying the weights pertaining to each neuron and the bias values within the hidden layer, whose values are computed according to the training set data. However, said values are initialised randomly, and thus the neural network training algorithms may reach their optimality conditions upon reaching local, and not global, minima.

In order to solve these issues, several numbers of neurons were tested, as done in (Bello, Reneses, & Muñoz, 2016), and the neural network with the hidden layer number of neurons that yielded the lowest validation set MSE was utilised for the final forecast. More specifically, the neuron number sweep consisted in the following: 10 to 60 with a step of 5, which results in 11 different neural networks. In addition, in order to deal with the resulting volatility in the neural network forecasts, this procedure was carried out a high number of replications (100 for this work's case) and the mean of all the replications is considered the final forecast of the proposed hybrid methodology.

Furthermore, the neural network training procedures have been paired with an early stopping technique in order to improve generalisation and avoid overfitting in the resulting forecasts. As it is normal with any regular optimisation problems,

[26] Standardisation in this context refers to the process of transforming a variable to one with a mean of 0 and a standard deviation of 1.

a high number of iterations is involved. Typically, given a random starting point, the validation set MSE is reduced throughout the first iterations. However, in the event that this error rises with each successive iteration, there is a strong possibility that the neural network is overfitting the training set data. In this case, the early stopping method finalises the optimisation algorithm and takes the weight and bias values pertaining to the moment prior to the error rise. All the specific optimality conditions and parameters that were applied to the neural network training algorithm utilised throughout this thesis can be checked on https://www.mathworks.com/help/deeplearning/ref/trainlm.html[27].

2.4 Evaluation criteria and benchmarking models

This section presents the studies and analyses that have been carried out so as to evaluate the proposed forecasting approach. In addition, comparisons have been performed with other forecasting models with the intention of validating the predictive accuracy of the proposed methodology.

2.4.1 Evaluation criteria and performance metrics

Some of the most used error metrics in the literature, e.g. (Sandhu, Fang, et al., 2016), have been chosen for this purpose, which are: mean absolute percentage error (MAPE), mean absolute error (MAE) and root-mean-square error (RMSE). These error measures are computed according to the equations below, following the same notation as the one used in Equation (2.20):

$$MAPE = \frac{100}{N} \cdot \sum_{i=1}^{N} \left| \frac{\hat{Y}_i - Y_i}{Y_i} \right| \tag{2.21}$$

$$MAE = \frac{1}{N} \cdot \sum_{i=1}^{N} \left| \hat{Y}_i - Y_i \right| \tag{2.22}$$

$$RMSE = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} \left(\hat{Y}_i - Y_i \right)^2} \tag{2.23}$$

[27] The neural network models of this thesis have been implemented in MATLAB.

As mentioned previously, it is important to highlight that prices in the Iberian power exchange may go to zero, and thus may result in infinite MAPE. Nevertheless, the lowest price that has occurred in the considered case studies is 2.3 €/MWh. Furthermore, it is also common in the literature, such as (Uniejewski, Nowotarski, et al., 2016; Uniejewski, Weron, et al., 2018), to provide statistically significant conclusions regarding forecasting performance comparisons via the Diebold-Mariano (DM) test, introduced in (Diebold & Mariano, 1995), and thus it has also been applied to this work's case studies. A 10% significance level has been considered, an absolute error difference as the loss differential series, and a two-sided perspective, i.e. testing for both out- and underperformance.

2.4.2 Benchmarking and competing models

The above evaluation criteria will be used for the proposed hybrid forecasting methodology as well as other models for benchmarking purposes. Two of these benchmarks represent the split versions of the hybrid model or its individual components. Benchmark 1 (BM_1) is a pure NN approach that is essentially the NN model used in the proposed hybrid scheme without the market clearing prices from the fundamental model. Said prices represent the second benchmark (BM_2).

The third benchmark (BM_3) is a slight modification of a linear regression model that was introduced in (Misiorek, Trueck, et al., 2006) and most recently applied to electricity price forecasting in (Marcjasz, Serafin, et al., 2018). This linear regression model can be represented as per the following equations in order to calculate the log-price, $p_{d,h}$, at day d and hour h:

$$\begin{aligned} p_{d,h} = \beta_{h,1} \cdot p_{d-1,h} + \beta_{h,2} \cdot p_{d-2,h} + \beta_{h,3} \cdot p_{d-7,h} + \beta_{h,4} \cdot p_{d-1}^{min} + \beta_{h,5} \cdot z_{d,h} \\ + \beta_{h,6} \cdot D_{Sat} + \beta_{h,7} \cdot D_{Sun} + \beta_{h,8} \cdot D_{Mon} + \varepsilon_{d,h} \end{aligned} \tag{2.24}$$

$$p_{d,h} = log(P_{d,h}) - \frac{1}{24 \cdot D} \cdot \sum_{d=1}^{D} \sum_{h=1}^{24} log(P_{d,h}) \tag{2.25}$$

The betas in Equation (2.24) are the regressor coefficients that are determined in the linear regression model, which respectively represent lagged log-prices (one, two and seven days), the minimum log-price of the 24 hours in day d minus one, the expected demand and three dummy variables indicating if day d is Saturday, Sunday or Monday.

Moreover, as mentioned earlier, a slight modification was carried out, which pertains to the logarithmic transform of Equation (2.25), which includes an hourly average of the log prices, $P_{d,h}$. Due to the possibility that prices may go to zero in the Iberian electricity market, the logarithmic transform that was applied is the mirror-log transform, which has been recently applied to electricity market price forecasting in (Uniejewski, Weron, et al., 2017):

$$n_{d,h} = \frac{(P_{d,h} - \mu_T)}{\sigma_T} \tag{2.26}$$

$$p_{d,h} = sgn(n_{d,h}) \left[log\left(n_{d,h} + \frac{1}{c} \right) + log(c) \right] \tag{2.27}$$

First of all, as per Equation (2.26), the prices were normalised by subtracting their mean in the training period (μ_T) and dividing by their standard deviation in the training period (σ_T). The parameter c in Equation (2.27) was set to one third as done in (Uniejewski, Weron, et al., 2017).

The fourth benchmark (BM_4) is based on ARIMA models, which are well recognised and more established in the literature than the aforementioned methodologies. These models have been widely used in electricity price forecasting, including the Iberian electricity market (Contreras, Espínola, et al., 2003). In this case, the model consists of a transfer function with SARIMA noise, which has been developed according to the methodologies presented in the works of (Box & Jenkins, 1970; Pankratz, 2012). Electricity price variance was stabilised by means of the Box-Cox transformation (Box & Cox, 1964). Furthermore, two seasonalities were considered that are most suitable in electricity price forecasting: 24 hours (one day) and 168 hours (one week). The Bayesian information criterion (BIC) value of the fitted models was used as model selection rule. The obtained SARIMA noise presents the following parameters with the standard notation: SARIMA$(1,0,0)_{168}(1,0,2)_{24}(1,0,0)_1$. The expected demand was used as an exogenous variable in this model, and thus this methodology can be also referred to as a SARIMAX model.

The last benchmark (BM_5) is a simple naïve approach in which the real electricity prices from the previous week are directly taken as the forecast. This benchmark is also chosen as a means to quantify the difference between weeks and thus the degree of evolution of electricity prices. The naïve model is represented as per the following formula, using the same notation as in BM_3's equations:

$$P_{d,h} = P_{d-7,h} \tag{2.28}$$

2.5 Case study description

In order to fully demonstrate the capabilities of the forecasting methodologies, these were tested in seven specific and instructive case studies on late 2016 of the

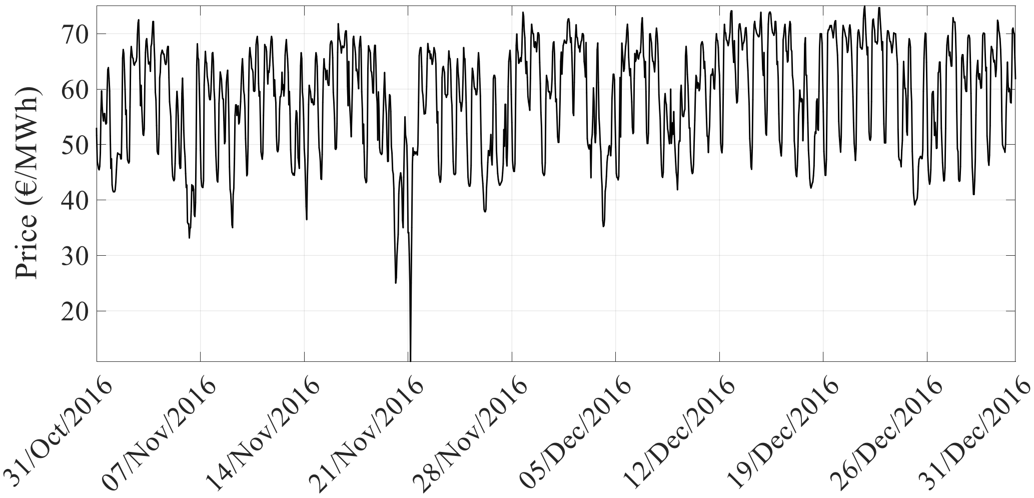

Figure 2.5: Iberian electricity market prices for late 2016

Iberian electricity market, as well as a general study for the entire year 2017 of the same power exchange. The cases of late 2016 are listed in Table 2.2, which are characterised by different types of complexities and pertain to the period depicted in Figure 2.5.

The most outstanding period is 21/Nov/2016, when prices collapsed due to unusually high levels of wind generation in the Iberia power system, reaching 10.88 €/MWh at early morning hours. Moreover, this day presents a price range of 56.61 €/MWh and a standard deviation of 15.07 €/MWh. These values are significantly higher than those of its adjacent days, which present a price range of 30.12 €/MWh and a maximum standard deviation of 9.48 €/MWh. Therefore, it would be interesting to analyse the models' forecasting performance on such a

Case study	Training periods		Validation period	Forecasting period
	$T_1 \cup T_2$	T_3	V	F
C_1	15/Oct/2015 to 13/Dec/2015	15/Oct/2016 to 13/Nov/2016	14/Nov/2016	15/Nov/2016
C_2	21/Oct/2015 to 19/Dec/2015	21/Oct/2016 to 19/Nov/2016	20/Nov/2016	21/Nov/2016
C_3	10/Nov/2015 to 09/Jan/2016	10/Nov/2016 to 10/Dec/2016	11/Dec/2016	12/Dec/2016
C_4	24/Nov/2015 to 23/Jan/2016	24/Nov/2016 to 24/Dec/2016	25/Dec/2016	26/Dec/2016
C_5	08/Oct/2015 to 06/Dec/2015	08/Oct/2016 to 06/Nov/2016	07/Nov/2016 to 13/Nov/2016	14/Nov/2016 to 20/Nov/2016
C_6	29/Oct/2015 to 26/Dec/2015	29/Oct/2016 to 26/Nov/2016	27/Nov/2016 to 04/Dec/2016	05/Dec/2016 to 11/Dec/2016
C_7	05/Nov/2015 to 03/Jan/2016	05/Nov/2016 to 04/Dec/2016	05/Dec/2016 to 11/Dec/2016	12/Dec/2016 to 18/Dec/2016

Table 2.2: Training, validation and forecast periods of case studies in late 2016

day (case C_2).

Furthermore, there is an apparent price level difference between November and December, which implies a slight market structural change. Moreover, coal plants were slightly less available in December than in November, whereas demand levels also increased in December. For these reasons, a day and a week that do not present uncommon behaviours and relatively stable price patterns have been selected for both November and December in order to compare both market circumstances (cases C_1, C_3, C_5 and C_7).

The most erratic weekly period in Figure 2.5 takes place between 05/Dec/2016 and 11/Dec/2016, which coincides with two Spanish National holidays that happen on the 6th and 8th of December (case C_6). Due to these holidays, electricity prices are lower than on their adjacent days. Additionally, another uncommon day included in Figure 2.5 is 26/Dec/2016 (case C_4), which is a Monday. However, the usual Christmas Day holiday happened on Sunday and thus the studied day was considered a holiday on most areas of Spain, so it cannot be considered a normal business days. These seven case studies put the proposed forecasting methodology and its benchmarks to the test under diverse scenarios and challenges, all of which are analysed and discussed in the following subsection.

Given that these specific cases may provide evidence as to how the considered forecasting models may perform under specific circumstances, the corresponding results cannot be appropriately generalised. Therefore, in order to provide more statistically significant results and further proof as to how these models perform, the entire year 2017 of the Iberian electricity market was also used as a case study. This case study is recent, and presents several kinds of market circumstances. For example, early 2017 was characterised by an uncommon mix of factors: very cold weather, high natural gas prices, low renewable generation and external alterations mainly caused by France's decommissioning of nuclear power plants. Consequently, prices rose to 101.99 €/MWh in the Iberian electricity system. Summer 2017 was relatively stable, although it is the starting point of the recent and ongoing increase in CO_2 emission allowance prices. Specifically, these prices rose by approximately 25% throughout the second half of the year 2017. Therefore, the Iberian electricity market of 2017 presents a plethora of situations and market regimes that can be used in order to analyse several capabilities of the forecasting methods.

2.6 Results and discussion

First of all, the cost-production optimisation model has been run for the training and forecasting periods that were mentioned in the previous subsection with hourly precision. In this test, regular circumstances in the Iberian power system were assumed regarding weather conditions (wind generation, hydro inflows, etc.). As a result, the estimated electricity market price was determined and used as an additional input to the neural network model.

Regarding the neural network runs, forecasting horizons of one day and one week with hourly precision were utilised. Once the day or week to forecast was selected, all periods according to the timeline of Figure 2.2 can be set. In all cases, a training set arranged as per Figure 2.2 was used with d_1 equal to 30 days, thus a total of 90 days were used as training data. Neural networks with this training set were trained considering different numbers of neurons (see subsection 2.3.2.2). The neural network with the lowest MSE on the validation set was later used to forecast the electricity market price on the test/forecast period. This procedure (from neural network training to forecasting) was carried out 300 times, whose mean was used as the resulting forecast of the proposed electricity market price forecasting model.

For the weekly cases, the author previously tested this methodology by both using a standard 168-hour horizon and a rolling window of a 24-hour horizon, i.e. forecasting day by day up to one week. However, by using a rolling-window method on a neural network model, the forecasts given in every simulation had a considerably higher volatility than those yielded by the neural network model with a 168-hour horizon, as well as a generally higher MAPE. Therefore, the rolling-window method was not utilised in these cases. However, one-day lagged prices were not included in the one-week forecasts, because, in reality, the price of the previous day becomes unknown when forecasting further than one day. Nevertheless, the estimated price from the fundamental model may contain, to some extent, the information of one-day lagged prices due to the chronological constraints (e.g. unit commitment and hydro reserve balance) that set links between current prices and other factors in the past.

2.6.1 Specific cases in late 2016

This subsection is arranged in two parts. First of all, a specific analysis of the proposed hybrid methodology is presented. The second part contains a general study and a comparison with the benchmarking models that were described in previous subsections.

The first case study (C_1) is displayed in Figure 2.6, which shows the forecast for 15/Nov/2016 (Tuesday). The daily pattern that electricity prices usually exhibit is successfully mimicked and the model yields a 2.179% MAPE, which mostly corresponds to the early morning hours, whereas the other hours, especially afternoon hours, are considerably accurate.

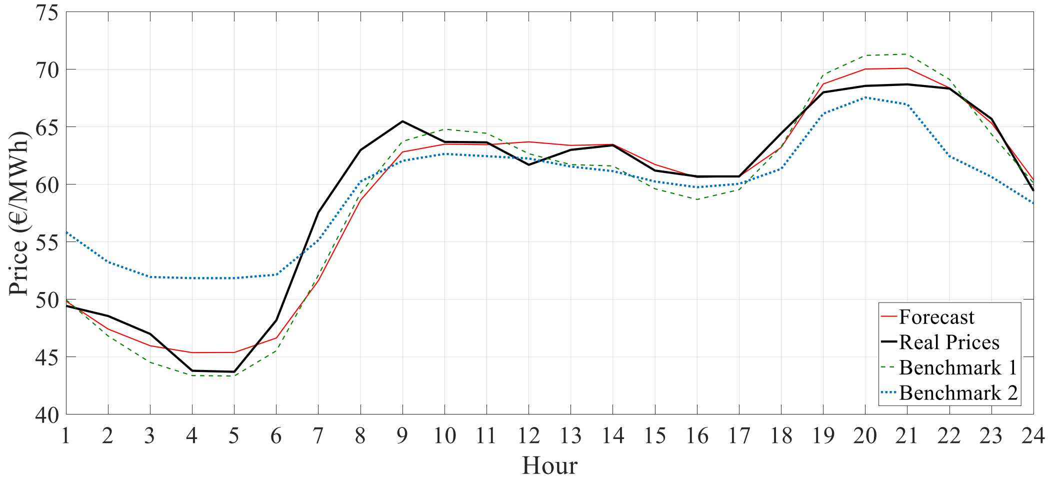

Figure 2.6: Electricity price forecast for 15/Nov/2016 (C_1)

The dashed and dotted lines of Figure 2.6 represent the forecasts of both components of the proposed hybrid methodology on their own, i.e. only the neural network model without the estimated market clearing prices from the cost-production optimisation model (Benchmark 1 or BM_1), as well as said estimated prices on their own (BM_2). By analysing and comparing these results, the benefits of the hybridisation of both methods can be checked and verified.

The forecasted market clearing price of the fundamental model (BM_2) is considerably flat and clearly lacks intraday dynamics, and thus yields a higher error (5.675% MAPE). Nevertheless, the estimation is centred on the daily average price level, which proves useful for the proposed hybrid methodology. The daily behaviour exhibited by the pure neural network forecast (BM_1) better resembles intraday electricity price patterns mainly thanks to its adaptability for capturing non-linear trends and agent strategic behaviours, although its accuracy is significantly lower on the afternoon and the evening (2.856% MAPE).

The combination of the daily equilibrium price level and the intraday patterns of the fundamental and the neural network model respectively proves advantageous

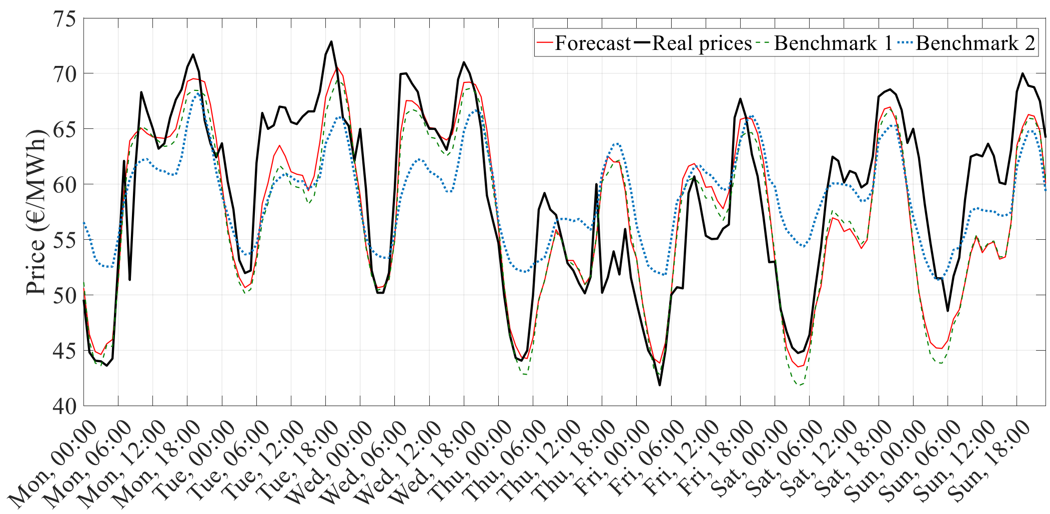

Figure 2.7: Electricity price forecast from Mon, 05/Dec/2016 to Sun, 11/Dec/2016 (C_6)

in this case. Regarding the one-week case studies, the forecast of case C_6 is depicted in Figure 2.7. In this case study, the resulting MAPE is 5.878%. Once again, it can be seen in Figure 2.7 that the estimated price from the fundamental model (BM_2) fails to follow the intraday patterns (8.709% MAPE). The neural network model on its own (BM_1) shows a more appropriate performance (6.136% MAPE), although its seems to yield considerably lower values on the early hours of Thursday to Sunday, which may be caused by a slight underperformance of the neural network model due to the degrading of the forecasts as the forecasting horizon increases.

This performance reduction is somewhat dampened by the estimated market clearing prices of BM_2, which, as seen in Figure 2.6, provides the daily equilibrium price level even at longer horizons. Therefore, the resulting hybrid forecast becomes more accurate and the benefits are experienced yet again, which strongly supports the statement that the combination of both models' advantages is highly valuable. Furthermore, in this case study, the fundamental model's ability to incorporate the effects of the reduction of availability in the system coal power plants provided a slight upward pressure on the hybrid model's price forecasts, and thus the contribution of the fundamental model proved relevant and useful.

The MAPE, MAE and RMSE results of the rest of the cases (C_2, C_3, C_4, C_5 and C_7) are shown in Table 2.3, Table 2.4 and Table 2.5 respectively, along with those of the other benchmark models. First of all, it is worth noting that the proposed model outperforms the others in all cases except C_2, in which the pure neural network model provides a significantly higher accuracy. This is mostly due to the fact that, in early morning hours, BM_1's forecast is closer to the real value than that of the proposed model. Furthermore, in a case in which prices

Model	C_1	C_2	C_3	C_4	C_5	C_6	C_7
Proposed – Hybrid BM_1/BM_2	**2.179**	15.96	**3.726**	**3.242**	**6.146**	**5.878**	**4.119**
BM_1 – Pure NN model	2.856	**13.79**	4.075	3.825	6.172	6.136	4.166
BM_2 – Fundamental model	5.675	26.73	6.812	10.83	11.23	8.709	8.135
BM_3 – ARX model	7.522	27.84	5.421	13.63	11.95	10.84	7.746
BM_4 – SARIMAX model	9.280	22.48	6.238	6.657	9.894	13.67	7.224
BM_5 – Naïve model	7.595	31.91	8.544	19.57	12.84	14.11	10.85

Table 2.3: Comparison of the proposed model with five benchmarks in terms of MAPE (%)

Model	C_1	C_2	C_3	C_4	C_5	C_6	C_7
Proposed – Hybrid BM_1/BM_2	**1.250**	5.294	**2.414**	**1.767**	**3.062**	**3.462**	2.609
BM_1 – Pure NN model	1.690	**4.444**	2.715	2.045	3.123	3.634	**2.542**
BM_2 – Fundamental model	3.062	7.575	3.414	3.374	5.496	4.893	4.890
BM_3 – ARX model	4.646	7.935	3.422	7.226	5.811	6.173	4.750
BM_4 – SARIMAX model	5.695	7.365	4.068	3.397	5.043	8.367	4.641
BM_5 – Naïve model	4.727	9.658	5.378	10.51	6.456	7.870	6.623

Table 2.4: Comparison of the proposed model with five benchmarks in terms of MAE (€/MWh)

Model	C_1	C_2	C_3	C_4	C_5	C_6	C_7
Proposed – Hybrid BM_1/BM_2	**1.861**	7.166	**2.622**	**2.474**	**3.897**	**4.466**	3.213
BM_1 – Pure NN model	2.035	**5.893**	3.063	2.718	3.930	4.621	**3.149**
BM_2 – Fundamental model	3.817	10.85	4.018	3.771	6.986	5.619	5.534
BM_3 – ARX model	5.249	9.938	4.697	8.241	7.151	7.617	5.655
BM_4 – SARIMAX model	6.202	8.105	4.407	3.743	6.163	9.704	5.552
BM_5 – Naïve model	6.043	11.24	6.876	12.41	8.333	9.780	8.453

Table 2.5: Comparison of the proposed model with five benchmarks in terms of RMSE (€/MWh)

collapse to such a low value (10.88 €/MWh), the difference is more apparent and noticeable.

This trend also happens in the common days pertaining to cases C_1 (see Figure 2.6) and C_3, although the proposed hybrid methodology's forecast in the rest of the hours of the day makes up for it more than enough, offering a higher overall accuracy. This fact also confirms that the fundamental model contribution enhances forecasting performance on late morning hours up to midnight, whereas on early morning hours it yields a reduced accuracy. Furthermore, said reduction is heightened on uncommon price collapsing situations such as C_2. This calls for a procedure that adequately weighs both models in order to obtain an enhanced predictive performance.

Moreover, regarding the case studies with a one-week forecasting horizon (C_5, C_6 and C_7), the proposed methodology and BM_1 show similar accuracies. However, on early December (case C_6), when the overall price levels are beginning to increase (see Figure 2.5), there is a more notable accuracy increase from the proposed hybrid model. This may suggest that whenever such a structural market evolution is underway, fundamental-related information should be taken into consideration.

Furthermore, case C_6 includes two Spanish National holidays, and thus it may imply that the proposed model is also the most proficient at forecasting prices on non-business days. The same conclusion can be reached from the results of case C_4, which is also a holiday. This may also indicate that by considering the estimated market clearing prices from the fundamental model, the bias effect from the previous week is lessened.

Model comparison*	C_1	C_2	C_3	C_4	C_5	C_6	C_7
Proposed vs BM_1	1	-1	0	1	0	1	0
Proposed vs BM_2	1	0	1	1	1	1	1
Proposed vs BM_3	1	1	1	1	1	1	1
Proposed vs BM_4	1	1	1	1	1	1	1
Proposed vs BM_5	1	1	1	1	1	1	1

*A value of 1 indicates significant outperformance of the proposed model forecasts
*A value of -1 indicates significant underperformance of the proposed model forecasts
*Otherwise no significant difference between model forecasts

Table 2.6: Results of the Diebold-Mariano test across the seven case studies of late 2016

The results shown by the MAPE, MAE and RMSE values have shown some differences between the proposed model and the five benchmarks. However, in order to appropriately analyse the performance in a statistically significant manner, a DM test with a 10% significance level was carried out, whose results are displayed in Table 2.6.

It can be seen that the proposed model's forecasts generally yield statistically superior performances than those of its competitors. The only exception in which its forecasts are statistically underperforming is in case C_2, where BM_1 has proved to be more accurate according to the error metrics of the previous tables. By contrast, three cases, C_1, C_4 and C_6, show significant differences in favour of the proposed hybrid forecasting methodology when compared to BM_1.

2.6.2 General case study of the entire 2017

The proposed forecasting methodology, as well as the five benchmarks, have also been tested for the entire year 2017 for both one-day and one-week forecasting horizons. The MAPE, MAE and RMSE forecasting error results are shown in Table 2.7, Table 2.8 and Table 2.9 respectively.

It is important to highlight that the difference in the fundamental model's results for each horizon are not exactly identical. This is due to how the one-week horizon case was carried out, where the weeks did not overlap each other and

Model	One-day forecasting horizon					One-week forecasting horizon				
	Winter	Spring	Summer	Autumn	Average	Winter	Spring	Summer	Autumn	Average
Prop.	**12.83**	**8.840**	5.016	**6.764**	**8.341**	**17.08**	8.330	**5.120**	8.000	**9.633**
BM_1	12.97	9.018	**4.977**	6.815	8.424	17.31	**8.202**	5.305	**7.919**	9.684
BM_2	26.67	21.54	12.39	19.01	19.89	26.38	21.94	12.13	19.18	19.91
BM_3	16.79	13.58	7.153	10.51	11.99	17.69	13.24	7.050	10.56	12.13
BM_4	15.06	9.293	5.097	7.654	9.248	17.64	10.36	5.915	8.662	10.65
BM_5	25.93	17.55	9.343	12.82	16.37	26.17	17.02	9.462	12.91	16.39

Table 2.7: MAPE comparison of the proposed model with five benchmarks for 2017 (%)

Model	One-day forecasting horizon					One-week forecasting horizon				
	Winter	Spring	Summer	Autumn	Average	Winter	Spring	Summer	Autumn	Average
Prop.	**5.137**	3.068	2.359	**3.331**	**3.465**	7.369	**3.247**	**2.407**	4.044	**4.267**
BM_1	5.178	**3.059**	**2.335**	3.383	3.480	7.486	3.257	2.491	**4.028**	4.315
BM_2	14.96	9.413	5.949	11.03	10.31	14.68	9.667	5.797	11.12	10.32
BM_3	6.838	4.765	3.262	5.066	4.972	**7.213**	4.652	3.198	5.116	5.045
BM_4	8.113	4.150	2.473	4.454	4.780	9.382	4.694	2.844	5.051	5.493
BM_5	10.53	6.225	4.266	6.387	6.828	10.54	6.043	4.315	6.434	6.833

Table 2.8: MAE comparison of the proposed model with five benchmarks for 2017 (€/MWh)

Model	One-day forecasting horizon					One-week forecasting horizon				
	Winter	Spring	Summer	Autumn	Average	Winter	Spring	Summer	Autumn	Average
Prop.	**5.921**	3.658	2.840	**4.003**	**4.096**	**8.927**	**4.074**	**3.065**	5.113	**5.295**
BM_1	5.953	**3.638**	**2.822**	4.089	4.115	9.046	4.115	3.164	**5.083**	5.352
BM_2	16.59	10.97	6.985	12.39	11.70	16.84	11.87	7.114	12.87	12.17
BM_3	7.809	5.552	3.885	6.055	5.814	9.010	5.944	4.105	6.612	6.418
BM_4	10.84	5.585	4.531	4.959	6.460	12.76	6.414	5.341	5.639	7.537
BM_5	11.48	7.092	5.030	7.567	7.773	12.32	7.775	5.458	8.298	8.464

Table 2.9: RMSE comparison of the proposed model with five benchmarks for 2017 (€/MWh)

seasons were not considered changed until the end of a week.

Moreover, it is worth noting that the forecasting errors of all models are significantly higher during winter than in the rest of the year mainly because of the extreme conditions that were present in the Iberian power system during January 2017. Furthermore, January's monthly average price was 71.49 €/MWh (18.6% and 38.2% higher than that of the previous and the following month respectively) and its standard deviation was 14.26 €/MWh (52.4% and 23.9% higher than that of the previous and the following month respectively).

Although the overall mean of the proposed model's forecasting errors is the lowest during the entire year 2017, it slightly underperforms during some seasons. For example, for the one-day forecasting horizon cases, the proposed model's performance during the early morning hours of the day in summer is lower, which coincides with the aforementioned trend described in the previous subsection.

In order to check if there are any statistically significant differences in predictive performances, a DM test has also been carried out for each model throughout the year 2017, whose results are displayed in Table 2.10 (with the same notation as in Table 2.6). According to the DM test results, the proposed model is not significantly outperformed in any of the cases and there seems to be a general outperformance in the entire year 2017 for the one-week forecasting horizon cases.

Taking into account that neural network forecast accuracy is reduced for longer

Model	One-day forecasting horizon					One-week forecasting horizon				
	Win.	Spr.	Sum.	Aut.	Avg.	Win.	Spr.	Sum.	Aut.	Avg.
Proposed vs. BM_1	0	0	0	1	0	1	0	1	0	1
Proposed vs. BM_2	1	1	1	1	1	1	1	1	1	1
Proposed vs. BM_3	1	1	1	1	1	0	1	1	1	1
Proposed vs. BM_4	1	1	0	1	1	1	1	1	1	1
Proposed vs. BM_5	1	1	1	1	1	1	1	1	1	1

Table 2.10: Results of the Diebold-Mariano test for the entire year 2017

Price	Hours	BM_1 MAE*	BM_2 MAE*	Total gen.	CCGT gen.†	Coal gen.†
<= 40 €/MWh	1038	5.253 €/MWh	9.030 €/MWh	24351 MWh	4.22%	7.02%
<= 35 €/MWh	415	8.220 €/MWh	10.22 €/MWh	23519 MWh	2.38%	4.07%
<= 30 €/MWh	203	11.20 €/MWh	10.57 €/MWh	22981 MWh	1.83%	2.81%
<= 25 €/MWh	117	16.10 €/MWh	7.870 €/MWh	22845 MWh	1.35%	1.92%
<= 20 €/MWh	83	18.24 €/MWh	6.283 €/MWh	22522 MWh	1.14%	1.45%
<= 15 €/MWh	51	18.43 €/MWh	4.612 €/MWh	22427 MWh	1.02%	1.54%
<= 10 €/MWh	22	20.16 €/MWh	3.075 €/MWh	22258 MWh	0.94%	1.56%

Table 2.11: Analysis of BM_1 and BM_2's performance on low price real cases in 2017

*Forecasting errors for one-day horizons

†As a percentage of the total generation

horizons, it can be deduced from this result that the contribution of the equilibrium price level that is provided by the estimated fundamental price is more notable, which coincides with case C_6's results that are explained in the previous subsection. This is also why, in several works in the recent literature, such as (Bello, Bunn, et al., 2016, 2017), this effect has proven useful for extended forecasting horizons windows (i.e. medium-term horizons).

Furthermore, extremely low prices that have occurred throughout 2017 are, on average, better forecasted by the market clearing prices when compared to the pure neural network's performance. These occurrences are mainly caused by the significant absence of the Iberian power exchange's most expensive generation technologies in the market clearing, i.e. coal and CCGT power units. This trend can be seen on Table 2.11, which shows the performances of the individual components of the proposed hybrid methodology (i.e. BM_1 and BM_2) for several price levels in the Iberian electricity market of 2017. As lower prices are selected throughout the year 2017, BM_2's forecasting accuracy beats that of BM_1, which may suggest that the price drivers in the power exchange are better captured by means of fundamental procedures on such occasions. Although it was repeatedly observed that the proposed fundamental-statistical hybrid model's performance needs to be improved in the early morning hours of the day that feature significantly lower prices when compared to the daily average, the events displayed on Table 2.11 do not involve abrupt or immediate price downfalls, i.e., the daily average prices are also lower than those of normal days.

In addition to these low-price occurrences, another remarkable event that can be observed and captured in the fundamental model throughout the year 2017 is related to the downfall in nuclear power plant production and overall temperature in France that caused its electricity market prices to rise, and, at the same time, the Iberian's electricity market prices saw an increase due to the interconnection between Spain and France. Another example of relevant occurrences that cannot be captured by a historical-based model has to do with regulatory changes that have taken place recently, albeit not during the case study, which is when the Spanish government temporarily suppressed the generation tax from October 2018 to March 2019 (both inclusive). This would represent an indirect reduction to the power system costs and thus a decrease in market clearing prices.

Another distinctive factor outside of the year 2017 that is currently present in European power exchanges is the CO_2 emission allowance price, which has almost reached a value of 30 €/tonne in early 2020, almost quintupling the average value in 2017 (5.84 €/tonne). This has a direct impact on fossil-fuel generation technologies, and thus causes an upward pressure on market clearing prices. Furthermore, the decarbonisation of the Spanish power system has led to the decommissioning of coal power plants, and thus market clearing prices are mostly set by CCGT power units unless left out of the market clearing by renewable power units. If any of these structural changes or punctual occurrences that may be only captured by fundamental models were to take place in the chosen case study, the fundamental model that has been developed in this chapter will be able to capture all of these events in a natural way and thus is expected to respond adequately and provide advantageous information for the forecasts.

2.7 Conclusions

The novel methodology that is presented in this chapter attempts to contribute to the short-term fundamental-statistical hybrid electricity market price forecasting literature, which is significantly scarce. This hybrid model consists of a cost-production optimisation fundamental model and a neural network model. The hybridisation of both models involves using the estimated market clearing prices of the cost-production optimisation model as an additional input data to the neural network model. Furthermore, the input data on both components of this novel hybrid approach were rearranged and modified in order to decrease computational burden and therefore increase efficiency, as well as reducing runtime and overfitting occurrences on the neural network model, which are of high value in short-term contexts.

The proposed hybrid model has shown adequate performance in seven particular and instructive case studies of the Iberian power system of late 2016, all of which have presented diverse circumstances and challenges. The five benchmark models, including some well-established methodologies in the literature, were outperformed by the proposed model in most case studies.

Furthermore, one of the main findings of this chapter proves that the proposed hybrid forecasting model's accuracy is generally increased by the daily equilibrium price level provided by the fundamental model. The combination with the neural network modelling capabilities for non-linear patterns in electricity prices makes for higher forecasting performance. In other words, the synergy behind the longer-term price level yielded by the fundamental model and the intraday pattern given by the statistical model has unquestionably proven to be advantageous, including uncommon market situations such as holidays, increased unit outages and extremely low price occurrences that do not feature sudden variations.

Additionally, the results of the case study of the entire year 2017 have demonstrated a general outperformance of the proposed hybrid forecasting methodology. This advantage is more notable on the one-week forecasting horizon results, which supports the conclusion that the equilibrium price level of the fundamental model enhances predictive accuracy even if the forecasting horizon is longer.

However, the results suggest that extremely low prices compared to the daily average level, such as those of early morning hours, a combination technique with another pure statistical model, or a regime-switching model within a hybrid framework, may enhance the resulting model's accuracy. Furthermore, other variables from the fundamental model may provide interesting results and an alternative and unique hybridisation approach, such as the thermal unit generation technologies or hydro power dispatch. Nevertheless, considering more input variables in this application may call for a more suitable sensitivity analysis or variable selection procedure than the proposed backward-elimination procedure, although developing a computationally efficient method may result in a highly challenging task due to the high level of complexity of the neural network training algorithms.

Moreover, the results of the case studies involving highly unstable periods, such as winter 2017, suggest that further adaptability enhancements should be determined in order to reduce forecasting error, which is considerably higher than that pertaining to case studies of relative stability, such as summer 2017. For instance, the calibration dataset periods of the neural network model may be modified according to market circumstances. Most of the proposed extensions of the hybrid methodology are addressed on the following chapters.

The figure on the next page, which is identical to Figure 1.2, can be referred to as a way of visualising not only the achievements and contributions provided in this chapter, but also the focus and the targets of the following chapter that arise as a result of these conclusions. As a brief anticipation to Chapter 3, the following enhancements are carried out in order to address the shortcomings that have been identified in this Chapter:

1. The fundamental model is improved by increasing the level of detail in the power unit structure regarding coal and natural gas power units, whose variable costs are estimated by means of an auxiliary forecasting procedure.

2. In order to consider market conditions in the forecast, a similar days method is used as guidance for the neural network's validation period selection. Moreover, more variables from the fundamental model that may

contain further information related to circumstances of the market are transferred to the neural network model

3. Given that the hybrid and pure neural network forecasting methodologies outperform each other in different scenarios, a forecast combination method is performed so as to take advantage of both method's accuracies in any scenario.

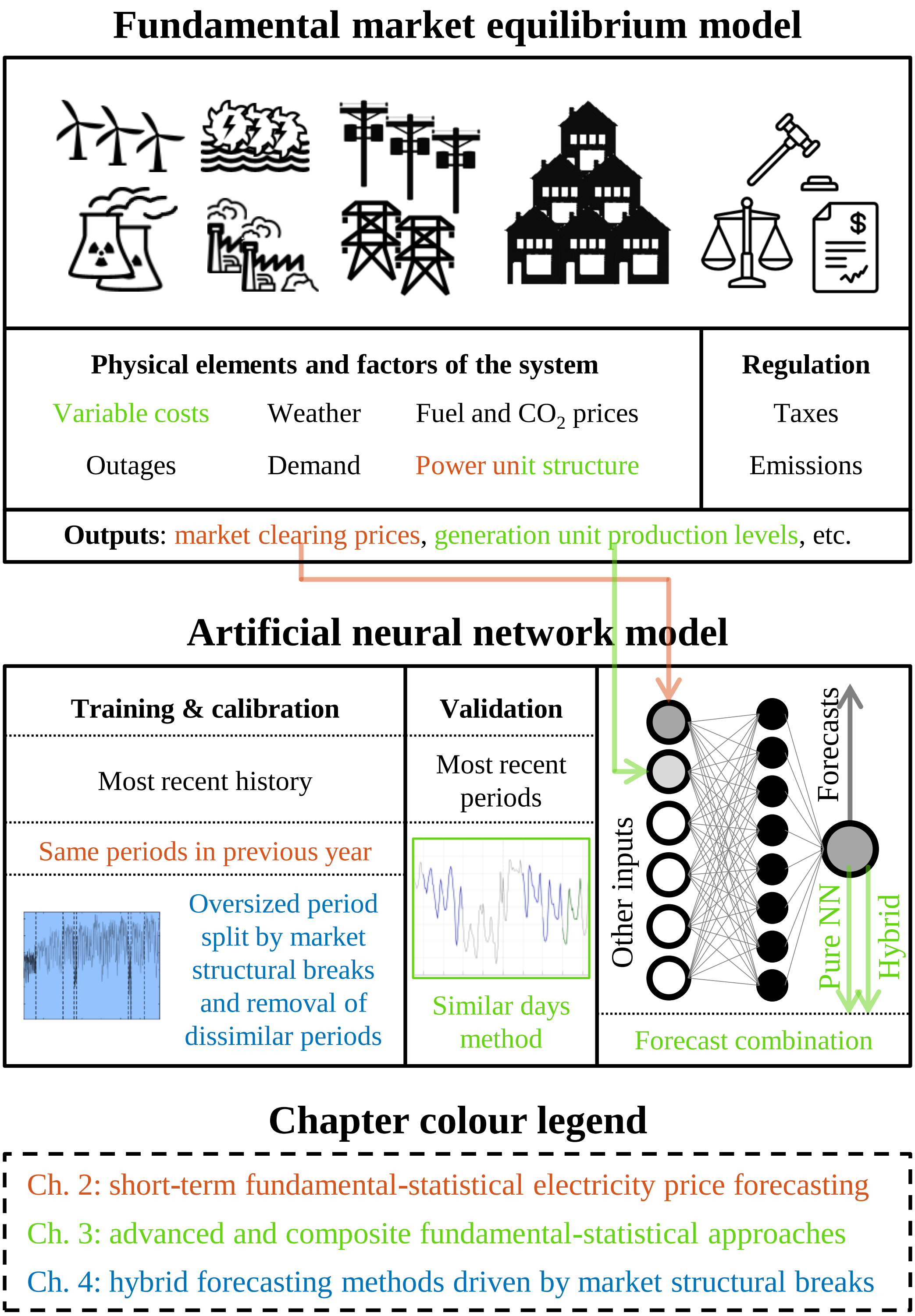

Figure 2.8 Overview of the hybrid electricity price forecasting model of the thesis as well as the targets and enhancements of the proposals of each chapter (identical to Figure 1.2)

Appendix A: A preliminary approach to probabilistic forecasting

Throughout the previous years, the field of probabilistic forecasting is gradually increasing in popularity, albeit mostly for longer-term horizons, as mentioned in (Ziel & Steinert, 2018). Point forecasts are still the most preferred option in short-term contexts, and thus it would be interesting to carry out a preliminary probabilistic analysis for the proposed models of this chapter, although the main objective of this work is focused on reducing the error of the mean of the forecasts. Moreover, probabilistic studies may provide some insight related to risk analysis and worst case scenario evaluation.

Some works in the literature have applied fundamental-statistical hybrid models in probabilistic forecasting contexts, as seen in (Bello, Bunn, et al., 2017; Nowotarski & Weron, 2017). A great part of these forecasting methods yield prediction intervals (Weron & Misiorek, 2008; Zhao, Dong, et al., 2008; Wan, Xu, et al., 2014), which are similar to confidence intervals, but associated with a random variable, for instance, electricity prices, instead of a model parameter. By contrast, the significantly less popular density forecasting methods have been rarely touched upon, which consider the full probability density function of electricity prices in order to provide the full picture of the underlying uncertainties in the probability distribution (Serinaldi, 2011; Jónsson, Pinson, et al., 2013).

The probabilistic analysis that is presented in this appendix involves two different methods related to the percentiles of the forecasts of the proposed model, denoted by $\hat{Y}_{ia}$, with $a = \{1, 5, 95, 99\}$. Firstly, the percentage of times the actual values of the electricity price (i.e. Y_i) in the are above the different percentiles of the predicted cumulative distribution has been measured. This measure will be referred to as exceedance rate. Ideally, given a percentile forecast $\hat{Y}_{ia}$, its exceedance rate should be of $(100 - a)$%. Secondly, the proposed model was tested via the pinball loss function (PLF) as done in several works in the current literature that focus on probabilistic analyses, such as (Liu, Nowotarski, et al., 2017). Lower values of the PLF score indicate that the forecasts are statistically superior and better reflect the probability of the occurrence of the price value associated with the target percentile, a. The PLF for a certain hour is calculated as per the following equation:

$$PLF(a)_i = \begin{cases} \left(1 - \frac{a}{100}\right) \cdot \left(\hat{Y}_{ia} - Y_i\right) & if\ \hat{Y}_{ia} > Y_i \\ \left(\frac{a}{100}\right) \cdot \left(Y_i - \hat{Y}_{ia}\right) & if\ \hat{Y}_{ia} < Y_i \end{cases} \tag{2.29}$$

Percentile	Model	One-day forecasting horizon					One-week forecasting horizon				
		Win.	Spr.	Sum.	Aut.	Avg.	Win.	Spr.	Sum.	Aut.	Avg.
P1	Proposed	**95.60**	**95.92**	**94.02**	94.05	**94.90**	93.77	**95.51**	**91.44**	**89.42**	**92.54**
	BM_1	95.42	93.98	92.57	**94.19**	94.03	**93.82**	94.96	89.88	88.14	91.70
P5	Proposed	**89.49**	**92.07**	**88.99**	87.87	**89.61**	**87.27**	**91.62**	**84.89**	**82.46**	**85.56**
	BM_1	88.94	90.85	87.45	**88.19**	88.86	86.26	90.57	84.16	82.01	85.75
P95	Proposed	11.06	**14.27**	**14.49**	**18.13**	**14.50**	**20.65**	**19.92**	**24.95**	**30.40**	**23.98**
	BM_1	**10.97**	15.08	15.35	18.96	15.10	22.34	19.09	28.39	30.49	25.08
P99	Proposed	5.324	6.612	**6.205**	**7.692**	**6.461**	**12.82**	10.12	**15.38**	19.55	**14.47**
	BM_1	**4.491**	**5.661**	8.605	8.929	6.929	13.51	**9.570**	19.09	**16.53**	14.67

Table 2.12: Exceedance rate of the percentile forecasts for the entire year 2017 (%)

These methods can be applied as a means to verify the statistical features of the forecasts that have been given by the proposed model. The exceedance rate of the percentile forecasts of the proposed model and BM_1 is displayed in Table 2.12. As mentioned earlier, the ideal exceedance rate for the percentile 1, 5, 95 and 99 forecasts are of 99%, 95%, 5% and 1% respectively.

First of all, it is natural that the longer horizon forecasts deviate more from these ideal values. The proposed model is generally closer to the ideal values than BM_1, especially the percentile 1 and percentile 5 values. This means that the proposed model's forecasts are generally more suitable for risk analyses such as extreme or worst case scenario evaluation. However, no notable improvement was achieved for the percentile 95 and percentile 99 cases. Furthermore, for both forecasting horizons, the percentile 95 and percentile 99 exceedance rates seem to be farther from their ideal values, which may be an indication that the probability distribution of the forecasts presents a positive skew and therefore is not symmetrical given that it somewhat fails in the upper tail. This can be verified by calculating the PLF according to Equation (2.29) of the proposed model's forecasts, whose results can be seen in Table 2.13.

The PLF results suggest that the proposed hybrid model yields overall superior probabilistic forecasts for percentiles 1 and 5 as opposed to percentiles 99 and 95 respectively, which indicates that the forecasts do not capture the probability of the occurrence of extremely high prices as well as the occurrence of extremely low prices. Therefore, this may call for a peak or extreme value detection procedure if this imbalance is to be solved.

Percentile	One-day forecasting horizon					One-week forecasting horizon				
	Win.	Spr.	Sum.	Aut.	Avg.	Win.	Spr.	Sum.	Aut.	Avg.
P1	**0.259**	**0.199**	0.180	0.239	**0.219**	**0.415**	**0.162**	**0.162**	**0.341**	**0.270**
P99	0.324	0.212	**0.172**	**0.217**	0.231	0.711	0.287	0.399	0.516	0.478
P5	**0.813**	**0.503**	**0.408**	0.590	**0.577**	**1.141**	**0.455**	**0.391**	**0.748**	**0.684**
P95	0.816	0.515	0.428	**0.588**	0.586	1.410	0.656	0.672	1.031	0.942

Table 2.13: Pinball loss function score of the proposed methodology for the entire year 2017

As mentioned previously, probabilistic forecasting in short-term applications is frequently ignored as much more importance is given to point forecasting. This preliminary analysis in the probabilistic spectrum has shown that the proposed forecasting model yields overall superior probabilistic forecasts. In other words, the market clearing prices from the fundamental model contribute to the improvement of the statistical performance of the neural network model. However, there seems to be room for improvement in order to obtain percentile forecasts that are closer to the ideal values regarding exceedance rates. This will be considered a future line of research of the thesis, given that it is more appealing in the short-term electricity price forecasting context to address and enhance the predictive accuracy of the expected value before approaching probabilistic forecasts.

Chapter 3

Adaptive Combination Methods for Hybridising Enhanced Fundamental and Statistical Models

As seen in Chapter 2, fundamental-statistical hybrid approaches can provide a suitable price forecasting methodology in the light of the plethora of disruption sources in the current electricity sector. In short, the pairing of the ability to incorporate regulatory and physical variations in the electricity system from the fundamental model with the linear as well as non-linear modelling capabilities of statistical techniques provide a significantly robust forecast that is suitable for the current electricity market contexts. However, these hybrid methodologies have rarely been applied in short-term contexts, where other considerations and issues must be borne in mind, such as the work that has been proposed throughout Chapter 2. For instance, the hourly resolution that is required in short-term contexts grants significant relevance to hourly trends such as intraday patterns, which are not adequately addressed by fundamental models alone. Therefore, one of the most crucial elements is the combination or hybridisation of the fundamental and the statistical model, which must constitute an adequate synergy between both modelling approaches. According to the findings of Chapter 2, this synergy may be heightened by weighing the contributions of both techniques in a dynamic manner. This is one of the main motivations of the forecasting methodology of this chapter[28], which features a unique hybridisation approach, including forecast combination methods. Moreover, several methods have been utilised in this work in order to modify the input datasets for enhanced predictive accuracy. The performance of this proposal has been analysed in the real and recent case study of the Iberian power exchange of the year 2017 and has outperformed other well-recognised and traditional methods.

[28] The studies, analyses and results presented in this chapter are based on the journal paper (de Marcos, Bello, et al., 2019b)

3.1 Introduction and literature review

Diverse power exchanges in the entire world have undergone considerable changes since their corresponding deregulation and liberalisation events, as mentioned in Chapter 2. Therefore, traders and practitioners were given much more investment options than in the previous monopolistic market environment and, as a result, these markets have grown significantly competitive and their participants are thus forced to adjust their strategies in order to withstand competition. Furthermore, there are many other disruptive forces that are involved in the electricity market price formation, such as a significant amount of intermittent generation and constant regulatory reforms.

As was stated in Chapter 2, an interesting trend albeit uncommon in the literature proposes combining statistical models with fundamental models in order to consider all of the aforementioned forces in the electricity market price forecast, such as (González, Contreras, et al., 2012; Bello, Reneses, Muñoz, et al., 2016). These hybridisation procedures are aimed at the elimination of both methods' shortcomings in order to provide an accurate and robust forecast. More specifically, this synergy consists of two important effects:

1. The estimated market clearing prices from fundamental models are somewhat devoid of short-term dynamics (such as intraday patterns) and statistical models are proficient at mimicking these high-frequency trends.

2. Including punctual events, such as generation unit decommissioning as well as subsidies and taxes, in statistical models is a highly challenging task due to the inherent assumptions in their training algorithms (e.g. history repeats itself). Resorting to fundamental modelling allows for the incorporation of such events in the estimated market clearing prices in a natural way.

This hybridisation is highly valuable in the current context of electricity market price forecasting and has shown positive results in medium-term applications (Bello, Reneses, Muñoz, et al., 2016). However, the short-term literature of these hybrid forecasting models is considerably scarce and, as seen in Chapter 2, there are significant areas of improvement, despite the adequate performance of the proposed hybrid forecasting approach.

The most important trend that was identified in the model hybridisation of Chapter 2 is the combination of the daily equilibrium price level from the cost-production optimisation model (fundamental approach) with the intraday effects provided by the neural network model (statistical method) as well as other non-linear effects that are not driven by market fundamentals. In other words, a neural network forecast alone is dragged to the daily average price thanks to the incorporation of the market clearing prices from the fundamental model, which not only reduces daily forecasting error, but also hourly error. However, it was observed that the hybrid forecast was somewhat flat and thus underperformed on hours of extreme prices when compared to a pure neural network model. Therefore, this calls for another hybridisation method that better weights the effects from both modelling approaches, for instance, reducing the effect of the fundamental model on hours of extreme prices while relying more on the statistical model.

This can be done by means of forecast combination techniques, although very few works have resorted to forecast combination methods in electricity price forecasting contexts, such as (Bordignon, Bunn, et al., 2013; Nowotarski, Raviv, et al., 2014; Alonso, Bastos, et al., 2016). The authors of these works are also aware of the scarcity of forecast combination in the electricity market price forecasting literature. Regarding the more general statistics literature, many authors refer to one of the first works that propose combination of multiple forecasting methodologies as an alternative method to using a single one, which is presented in (Bates & Granger, 1969). In this work, an adaptive combination method is proposed that assigns weights to the forecasting models depending on the inverse of their mean squared errors in a specific in-sample period. The weights assigned in this method are also known as time-varying coefficients, given that most forecasting applications involve distinct in-sample windows when forecasting different periods.

Since then, several extensions of this kind of forecast combination methodologies in the general statistical context have been proposed, such as the AFTER algorithm introduced in (Yang, 2004). Other extensions have been proposed in (Sánchez, 2008) for wind power forecasting and (Andrawis, Atiya, et al., 2011) for inbound tourism forecasting. Another kind of forecast combination in the literature is based on the Bayes' theorem: the Bayesian model averaging method (Hoeting, Madigan, et al., 1999). Some of these works were referenced and used for electricity price forecasting in (Bordignon, Bunn, et al., 2013; Nowotarski, Raviv, et al., 2014; Alonso, Bastos, et al., 2016), although none of them consider the application of forecast combination with fundamental and statistical modelling approaches.

There are other hybridisation methods outside of the model combination field that were pointed out in (de Marcos, Bello, et al., 2019a), such as using other variables computed in the cost-production optimisation model as input data to the statistical model. The main decision variables of this fundamental model involve production volumes, which also carry relevant information regarding the market clearing. For instance, if the thermal generation technology in the Iberian market clearing is entirely related to nuclear technology, prices are bound to fall. Therefore, production levels of coal and CCGT units, which constitute almost the remaining thermal power units in the Iberian system, can be considered as additional input variables to the neural network model of the fundamental-statistical hybrid modelling approach. This may be useful as a means of indicating specific market circumstances with extremely low or high prices, aside from the well-known regime-switching models. One of the most prominent regime-switching approaches is the Markov regime-switching model, typically aimed at capturing price spikes as well as other abrupt and erratic trends that may be exhibited in electricity prices (Bordignon, Bunn, et al., 2013). Once calibrated, a Markov regime-switching estimates regime probabilities and employs them as weights for a linear combination of a set of forecasts belonging to each regime, which is essentially another forecast combination scheme.

Furthermore, if several market regime changes are present in recent history, the predictive performance of statistical models may be considerably hindered due to their calibration and training methods, whose primal assumption is that the trends that are present in their input datasets will repeat in the future. For instance, if weather conditions have significantly varied within the most recent month, one should carefully consider including months prior to the previous month in the training dataset, as these months would most likely belong to a distinct market regime with regards to the current market conditions. Therefore, selecting the most appropriate training period in electricity market price forecasting applications is crucial. However, in spite of this clear relationship, calibration period selection is frequently disregarded in electricity market price forecasting applications. The authors of (Marcjasz, Serafin, et al., 2018) have recently pointed out this issue and claim that forecasting models with shorter calibration windows adapt better to changes, whereas longer calibration windows result in a better estimation of the trained model's parameters, which is also mentioned in (Pesaran & Timmermann, 2007). Nevertheless, ARX forecasting models are the main focus of the work presented in (Marcjasz, Serafin, et al., 2018) and thus this calls for a suitable procedure that can be applied to artificial intelligence models such as neural networks in order to verify if a more accurate forecast can be obtained.

Therefore, the motivations that encourage the work of this chapter are aimed at the betterment of the hybrid methodology of Chapter 2 while considering the issues and scarcities pointed out in the previously mentioned statements and facts. It should be noted that the entirety of this chapter is based on the journal paper of (de Marcos, Bello, et al., 2019b), which was written throughout the middle stages of the development of this thesis. The main contributions of this chapter are summarised as follows:

1. An improved version of Chapter 2's fundamental-statistical hybrid model is proposed. On the one hand, the fundamental component, a cost-production optimisation model, was enhanced by increasing the level of detail in its thermal generation unit structure and by incorporating an auxiliary model that estimates thermal unit variable costs based on their past bids and relevant fuel commodities. On the other hand, more useful predictors were used in the statistical component, a neural network model,

as well as an advanced validation period selection via similar day methods and a shortening calibration procedure in order to enhance the resulting forecast.

2. Several hybridisation schemes were studied and tested. Aside from market clearing prices from the fundamental model's results, thermal and hydro generation levels were also used as additional inputs to the neural network model. Furthermore, an additional hybridisation stage was incorporated at the end of the neural network forecast that involves an adaptive forecast combination with a pure neural network model.

3. The proposed models were put to the test on the real-size market case of the Iberian power exchange of the year 2017. Every stage and component of the proposed composite hybrid methodology is evaluated separately and a comparison with other well-recognised methods in the literature is provided by means of error metrics and a robust statistical test.

This chapter is organised as follows. Section 3.2 presents a thorough description of the proposed methodology of this chapter. Section 3.3 contains the evaluation criteria and the benchmarking models. The results and analyses are shown in Section 3.4, as well as the corresponding discussions. Finally, Section 3.5 presents the main findings, conclusions and contributions of this chapter, as well as potential areas of improvement that are mostly covered in Chapter 4.

3.2 Methodology

The main objective of this chapter is to propose and develop a novel short-term hybrid forecasting model and verify its performance on a real, full-scale and

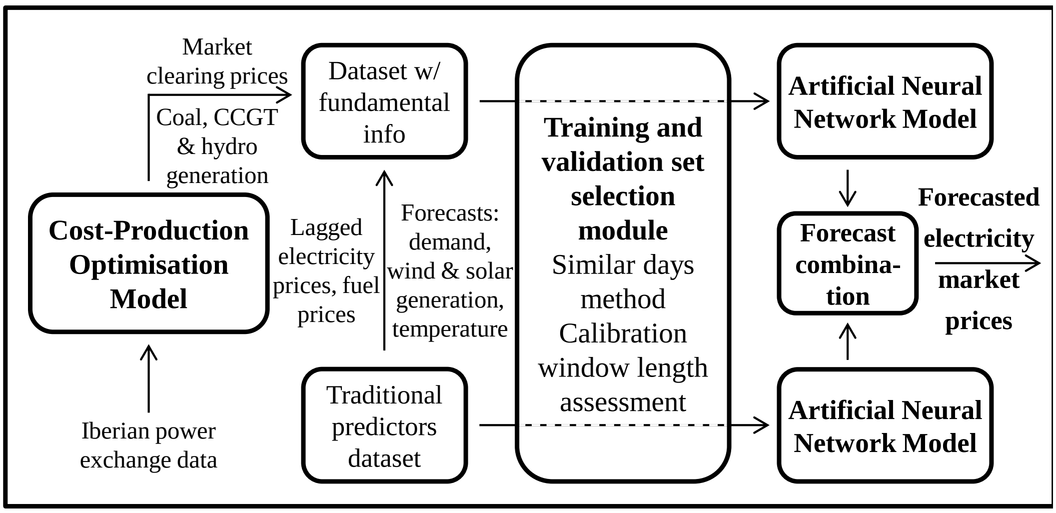

Figure 3.1: Overview of the proposed composite fundamental-statistical approach

complex case study, such as the Iberian electricity market. A diagram that summarises the workflow of the proposed hybrid methodology is shown in Figure 3.1.

First of all, the fundamental methodology, a cost-production optimisation model, is run to obtain its corresponding output variables, which are later utilised as additional variables in the statistical model's input dataset. After applying the necessary data pre-processing approaches to the input datasets, two neural network forecasting models are run with and without said additional variables to finally unite the resulting sets of forecasts by means of a forecast combination method.

The following subsections present the specific details of each component of this chapter's proposed methodology.

3.2.1 Fundamental model improvements

The fundamental component, displayed at the left-hand side of Figure 3.1 is based on the fundamental model that has been utilised throughout Chapter 2. A full description of this market equilibrium model is located in Section 2.3.1. Said modifications were encouraged by the computational difficulties of thoroughly modelling the Iberian power system with hourly granularity as per short-term standards, which translate to relatively high problem sizes and resolution times. In short, thermal units with similar cost functions and technical properties were aggregated into larger ones in order to reduce the problem size.

However, it was observed that the hourly accuracy of the resulting market clearing price estimations was somewhat low, which is mainly a result of its flatness that becomes more apparent on hours of extreme prices. Therefore, it is worth verifying if increasing the level of detail in the fundamental model, thus increasing problem size and resolution time, may have a positive effect on the forecasting accuracy. More specifically, given that CCGT and coal generation technologies are the most expensive in the Iberian electricity system, the generation unit aggregation that was carried out in Chapter 2 regarding these technologies was reversed and thus these power units were considered individually. Moreover, the variable costs of these thermal power units (e.g. fuel, CO_2 emissions, etc.) were unified into a single variable cost. This removes the need for ascertaining all specific costs (e.g. CO_2 emissions, fuel commodity costs, etc.) of each of the 30 coal and 63 CCGT programming units[29]. This also

[29] All programming units in the Iberian electricity system can be found at http://www.mercado.ren.pt/EN/Electr/MarketInfo/StructuralInfo/MarketUnits/Pages/ProgrUnits.aspx for those located in Portugal and at https://www.esios.ree.es/en/programming-units for those located in Spain.

significantly reduces the problem size. This variable cost is estimated by means of an hourly auxiliary linear regression procedure that is based on the following regressors:

1. Day ahead CO_2 emission allowance futures price
2. NBP natural gas (CCGT units) or API2 coal (coal units) commodity futures[30] price
3. USD to EUR currency exchange value
4. Daily variable cost with a 90-day lag[31]
5. Most expensive hourly bid price with a 90-day lag

A special bid is used as proxy for the variable cost, which is the fourth regressor in the above list. In the Iberian electricity system, all programming units are able to establish a minimum revenue or income[32] condition[33] for each market clearing (i.e. the whole 24-hour period), which is represented by the following inequality:

$$\sum_{t=1}^{24} cv_t \cdot cp_t \geq FT + \sum_{t=1}^{24} cv_t \cdot VT \tag{3.1}$$

Where:

- FT is the fixed term in euros (no decimal figures allowed) of this special bid
- VT is the variable term in €/MWh (three decimal figures allowed) of this special bid
- cv_t is the cleared volume in MWh after a certain iteration of the market clearing at hour t
- cp_t is the market clearing price in €/MWh after a certain iteration of the market clearing at hour t

[30] More specifically, month-ahead NBP natural gas futures and month-ahead API2 coal futures.

[31] A 90-day delay is enforced according to the Iberian market confidentiality rules, so in practice no one can obtain updated information regarding the bids of all the programming units in the Iberian power system. This is done according to the 18th clause of the Official State Gazette BOE-A-2018-6925 (content not available in English): https://www.boe.es/diario_boe/txt.php?id=BOE-A-2018-6295.

[32] Known in Spanish as *Condición de ingresos mínimos*, also accessible at the 28th clause of the Official State Gazette BOE-A-2018-6925.

[33] There are other conditions that programming units must fulfil, such as the maximum number of offers, which is 25 per hour.

FT and VT must be set by the programming unit owner before the market clearing occurs, and these are fixed for the whole 24-hour period of the following day, as seen in Equation (3.1). When the market clearing is underway, these inequalities are verified and, if not satisfied, the programming unit bids are removed from the order book and replaced by other, more expensive bids. Given that the owner is the one who adjusts this condition, it is safe to assume that VT reflects the variable costs of the programming unit (including taxes), and thus it is used as a proxy. However, the real and current VT should be estimated given the 90-day confidentiality rule. This information with the 90-day lag is available on the historical datasets of the Iberian market operator website (OMIE): http://www.omie.es/en/.

The computational differences as a result of these modifications to the fundamental model are shown in Table 3.1 for a forecasting period of one week and a comparison of the mean absolute error (MAE) of the obtained market clearing prices throughout the year 2017. As expected, an increase in the number of generation units in the system led to a significantly higher problem size, a quadruple maximum RAM (random-access memory) usage and a longer resolution time. However, the MAE was reduced by approximately one third with respect to that of Chapter 2 throughout the entire year 2017, which makes this increase in the level of detail and computational burden a worthwhile exchange. These results were obtained under identical conditions and the same computer[34] as the one indicated in Chapter 2.

3.2.2 Enhancements with respect to statistical approaches

The statistical component of the proposed hybrid methodology of this chapter has also been improved. More specifically, the following enhancements were carried out: an extended dataset with more predictors was used, a sophisticated validation period selection approach was applied for the neural network training algorithm, and several forecast combination methods were proposed in order to increase overall predictive accuracy. The following subsections present these extensions in detail.

Fundamental model	Equations	Variables	Runtime	Max. RAM usage	2017 MAE
This chapter	50745	118905	7.40 sec.	278 MB	6.84 €/MWh
Chapter 2	12440	71024	3.91 sec.	76 MB	10.31 €/MWh

Table 3.1: Comparison of computational statistics and MAE of the market clearing prices

[34] More specifically, a 64-bit Windows 7 PC with 16 GB installed RAM and the following processor: Intel® Core™ i7-3770 CPU@ 3.40 GHz of 4 cores and 8 logical processors.

3.2.2.1 Extended predictors dataset

As shown in Figure 3.1, not only the market clearing prices were taken as the linking output of the fundamental model, but also the generation levels of the coal, CCGT and hydro units. These were merged into a specific dataset, as well as more common predictors, which are:

- Expected values of demand, wind and solar generation.
- Expected mean temperature[35] in the Iberian Peninsula.
- Two dummy variables indicating if it is a working day or a Sunday/holiday, thus leaving the case of Saturday for when both of these dummies are false.
- Month-ahead futures prices of API2 coal, month-ahead futures prices of NBP natural gas and month-ahead futures prices of European CO_2 emission allowances.
- Lagged electricity market prices: specifically: one day, two days, one week and two weeks.

The sources for these variables were indicated throughout Section 2.3, except for the temperature forecasts, which can be found at https://rp5.ru, an open-source platform of weather forecasts around the globe. According to the above list, 17 input variables are considered in the dataset, including the four outputs of the fundamental model. This set of input variables includes several kinds of explanatory factors that influence the Iberian electricity market according to the authors of (Monteiro, Fernandez-Jimenez, et al., 2015).

3.2.2.2 Calibration and validation period selection

Before running the neural network model with these variables, a calibration period selection procedure must be carried out so as to reduce overtraining issues in the neural network forecasts. The calibration set variables of the neural networks were arranged as per the timeline displayed in Figure 3.2, where the bottom labels indicate the interval names and the top labels the interval length in days. Given a certain forecasting day *F*, the neural network training set is split into three periods with their corresponding intervals in days (“1y” represents one year):

- $T_1 = [F - 1\text{y} - d_1, F - 1\text{y})$

[35] In actuality, a weighted mean of the temperature forecasts for the main cities throughout the Iberian Peninsula was performed according to their population.

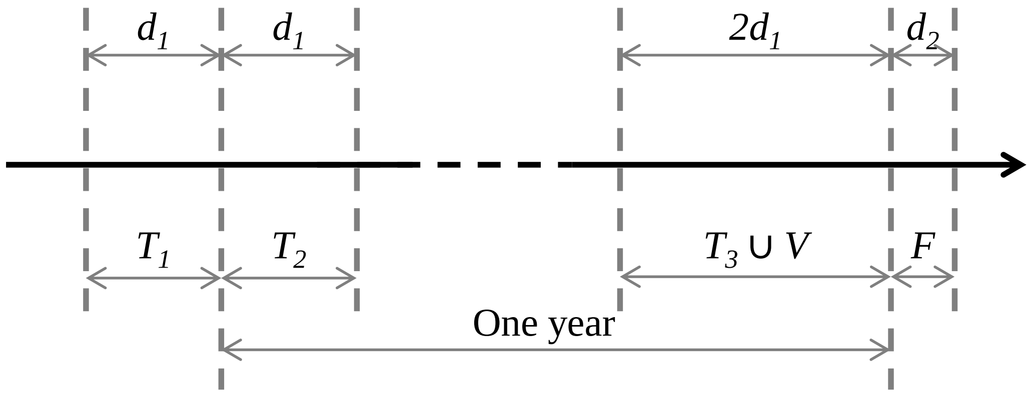

Figure 3.2: Training, validation and test/forecast periods arrangement

- $T_2 = [F - 1y - d_1, F - 1y + d_1)$
- $T_3 = [F - 2 \cdot d_1, F)$

The length of these three intervals depends on the value of one parameter, d_1. The first two periods, T_1 and T_2, present relevant information related to similar conditions regarding calendar effects (e.g. season) in the previous year, whereas the third period contains the most recent information. However, contrary to what is usually performed on neural network forecasting applications that do not consider cross-validation[36], the validation period V is not placed immediately prior to the forecasting period F. In this case, the validation period V is placed anywhere within T_3's interval according to a similar-day method, which has been recently considered in (Bento, Pombo, et al., 2018), although their guidance criteria[37] are significantly different to this work's application.

The similarity test that has been carried out is a modified version to the method proposed in (Mandal, Senjyu, et al., 2006), which selects days in the past as calibration dataset according to its similarity in terms of variables that are available at the moment of the forecast. In this case, the following similarity criteria were taken so as to select days within T_3's range as validation set: expected demand (ED), expected demand deviation (EDD), expected temperature (ET) and expected wind generation (EW). A Euclidean norm for every hour i is used in order to evaluate the similarity between the forecasting period F and sub-periods of equal length (i.e. of d_2 days) contained in training period T_3. Furthermore, given that all the similarity criteria variables do not coincide regarding their orders of magnitude, weighted factors w were applied, as shown in the following equations:

[36] Cross-validation was considered in the neural network models of the thesis. However, given the neural network modelling approach (number of neurons sweep and replications, see Section 2.3.2.2), applying cross-validation would yield significant computation issues.

[37] More specifically, the authors of (Bento, Pombo, et al., 2018) select three days in recent history that present similar price patters in price behaviours to those exhibited in the day prior to the forecast period.

$$\|D_i^{sub}\| = \sqrt{D_{i,1}^{sub} + D_{i,2}^{sub} + D_{i,3}^{sub} + D_{i,4}^{sub}} \tag{3.2}$$

$$D_{i,1}^{sub} = \widehat{w}_1 \cdot \left(ED_i^F - ED_i^{T_3^{sub}}\right)^2 \tag{3.3}$$

$$D_{i,2}^{sub} = \widehat{w}_2 \left(EDD_i^F - EDD_i^{T_3^{sub}}\right)^2 = \widehat{w}_2 \left[\left(ED_i^F - ED_{i-1}^F\right) - \left(ED_i^{T_3^{sub}} - ED_{i-1}^{T_3^{sub}}\right)\right]^2 \tag{3.4}$$

$$D_{i,3}^{sub} = \widehat{w}_3 \left(ET_i^F - ET_i^{T_3^{sub}}\right)^2 \tag{3.5}$$

$$D_{i,4}^{sub} = \widehat{w}_4 \left(EW_i^F - EW_i^{T_3^{sub}}\right)^2 \tag{3.6}$$

The estimated weights, $\widehat{w}_i$, are calculated by means of a linear regression model across every hour i that belongs to every sub-period within T_3 (i.e. every T_3^{sub}). This regression model is represented in the following equation:

$$ED_{i+1} = \widehat{w}_1 \cdot ED_i + \widehat{w}_2 \cdot EDD_i + \widehat{w}_3 \cdot ET_i + \widehat{w}_4 \cdot EW_i \tag{3.7}$$

The sub-period that presents the lowest average value of $\|D\|$ is thus the most similar to the forecasting period *F*. In this case, the top 20% most similar sub-periods were chosen as the validation period *V*. This will enable the neural network to perform its optimisation algorithm of the errors on a period which is more akin to the forecasting period, and thus the resulting neural network's capabilities of suitably forecasting period *F* should be superior. This data arrangement is more efficient and reduces redundancy in the neural network training set as well as overfitting occurrences. Furthermore, the proposed similar-day method provides a robust control that takes calendar effects into account, such as ignoring non-business days as validation data when forecasting business-day prices and vice versa.

However, the length of the training set or the value of d_1 should be chosen carefully. As mentioned earlier, shortened calibration windows have demonstrated to be an effective means of increasing adaptability and responsiveness to sudden changes. As seen in Chapter 2, the Iberian electricity system on early 2017 was affected by a highly uncommon combination of factors: very cold weather, low renewable generation (especially hydro and wind generation), high natural gas prices and disrupted interconnection with France due to its decommissioning of nuclear power plants. Therefore, shortening calibration periods may prove a suitable solution during such unstable periods so as to increase adaptability and thus provide a higher forecasting accuracy.

Using smaller calibration windows was recently considered in the work of (Marcjasz, Serafin, et al., 2018), which proposes averaging forecasts provided by the same forecasting methodology (in their case, they utilised ARX-type models) but using distinct calibration windows. Given that the proposed methodology of this chapter employs neural network forecasting approaches, the application of the forecast averaging as per (Marcjasz, Serafin, et al., 2018) would significantly increase computational burden. Instead, the number of days d_1 is set according to a preliminary test based of validation set mean-square error (MSE) during unstable periods that is explained throughout the next subsection. Otherwise, it is set to 30 days as done in Chapter 2.

3.2.2.3 Neural network modelling approach

The neural network forecasting model configuration is similar to the one utilised throughout Chapter 2, which can be summarised as follows (refer to Section 2.3.2.2 for a more detailed description):

- A single hidden and output layer structure with hyperbolic tangent sigmoid and linear activation functions respectively.
- A Levenberg-Marquardt training algorithm.
- A sweep of several numbers of neurons (more specifically, 10 to 60 with a step of 5) in the hidden layer.
- Several replications of the training procedure so as to compute the mean of each replication's forecast as the last step.

However, given that the input dataset is considerably larger than that of Chapter 2's neural network model, a short analysis was conducted in order to observe the variation of the mean error of the forecasts when performing an additional replication of the neural network training process. The result of this analysis is shown in Figure 3.3. The average effect of adding an additional replication to the forecast is within the ±0.2% MAE variation range starting from the 40th replication approximately. Therefore, for the purpose of safety, 50 replications are chosen by default.

However, for unstable periods, several values of d_1 were tested considering several dynamics (e.g. idiosyncratic features of prices, seasonal behaviours, etc.), as displayed in Table 3.2. Furthermore, shortened calibration windows mean that the neural network has less information to train on, and therefore its resulting forecasts will exhibit a higher volatility, and thus the number of replications is set to be inversely proportional to d_1.

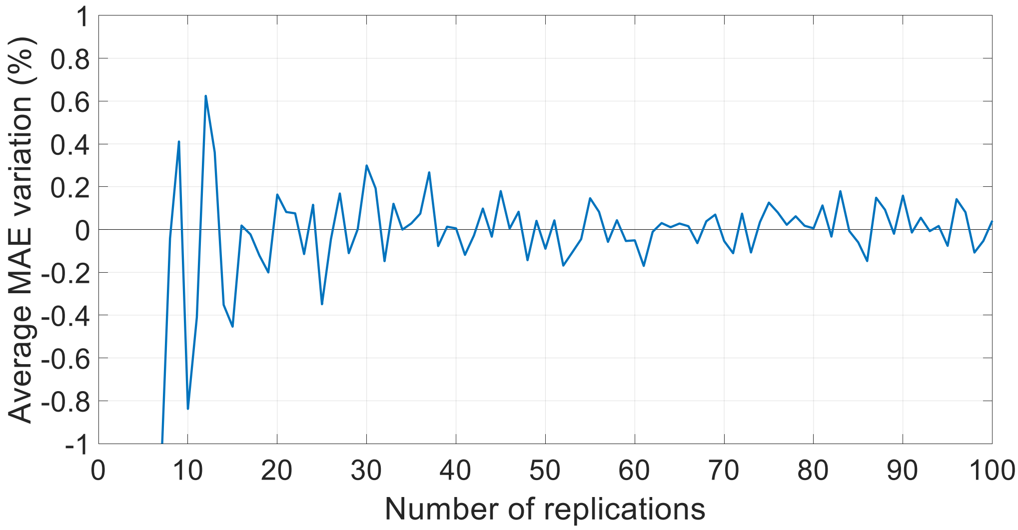

Figure 3.3: MAE variation as a function of performing an additional replication

d_1 length	$T_1 \cup T_2 \cup T_3 \cup V$ length	V length	Number of replications
30 days	120 days	12 days	50
15 days	60 days	6 days	75
10 days	40 days	4 days	85
5 days	20 days	2 days	100

Table 3.2: Training and validation sets for neural network forecasting on unstable periods

The first few replications of the neural network training procedure are run for each value of d_1, whose resulting average validation set MSE is taken as the criterion for the selection of the value of d_1. The default value of 30 days is set on periods of more relative stability, due to the fact that longer calibration windows entail a better estimation when no sudden changes are present, as mentioned in (Marcjasz, Serafin, et al., 2018). As mentioned earlier, the mean of the forecasted prices of each replication is taken as the final forecast of the neural network model.

3.2.3 Forecast combination techniques

According to Figure 3.1, the ultimate stage of the proposed hybrid methodology involves a forecast combination procedure, where the neural network forecasts with and without fundamental information on their training datasets are combined. This forecast combination is encouraged by the effects seen on the hybrid forecasting model of Chapter 2, which yielded superior forecasts on relatively stable periods (i.e. with few abrupt changes and spikes, such as summer) than the pure neural network model, whereas the pure neural network model outperforms on hours of extremely high and/or low prices. Therefore, it is essential to combine both of these positive effects and obtain a synergy in order

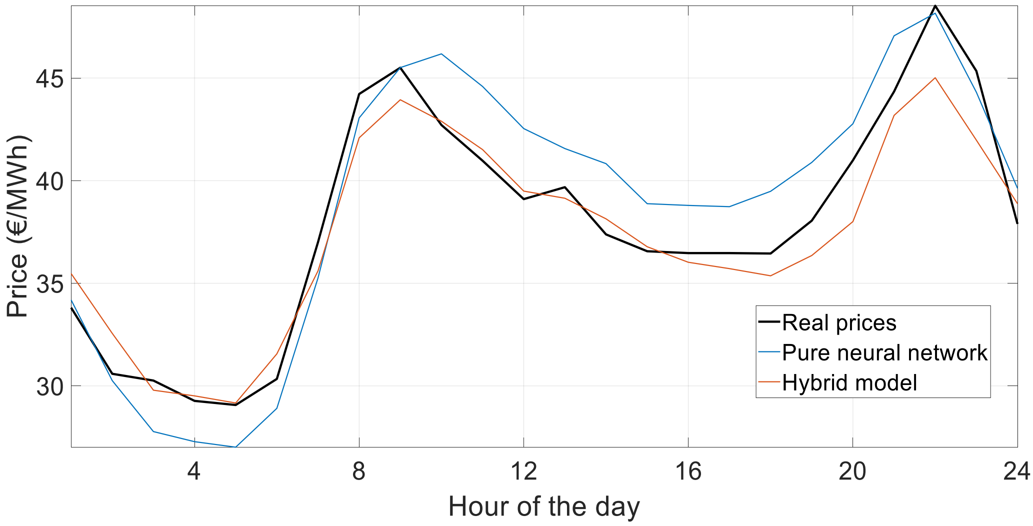

Figure 3.4: Forecasts of the hybrid and pure NN methodologies of Chapter 2 on 06/Apr/2017

to minimise the intraday adaptability reduction of the hybrid model while taking advantage of its better estimation of the equilibrium price levels.

One of the most representative and illustrative examples of the aforementioned effects is displayed in Figure 3.4, where the hybrid model shows a lower performance on the hours of higher prices, although in this case this is the reverse trend than that of the first hours in the early morning. In this example, a suitable combination procedure between these two forecasts should assign more weight to the hybrid model on the first 18 hours of the day, while assigning more weight to the neural network forecast for the remaining hours of the day.

However, given that the literature regarding forecast combination in electricity price forecasting contexts does not clearly favour a specific combination method, the following techniques have been tested: simple averaging, inverse validation error weighting, and Bayesian model averaging. Given this work's purpose, hourly combinations may prove useful so as to assign weights to the most outperforming model, carried out in an *ex-ante* manner. All combination methods can be represented by the following equation, where $\hat{Y}_{i,m}$ represents the estimated values provided by the forecasting model m for a specific hour i:

$$\hat{Y}_i^c = \sum_{m=1}^{M} w_{i,m} \cdot \hat{Y}_{i,m} \qquad \forall i \tag{3.8}$$

The hourly weights for each model, $w_{i,m}$, differ among combination methods, but they all satisfy the usual conditions assumed in these applications, which are as follows:

$$\sum_{m=1}^{M} w_{i,m} = 1 \qquad \forall i \tag{3.9}$$

$$w_{i,m} \geq 0 \qquad \forall i,m \qquad (3.10)$$

These conditions ensure that the value of the combined forecasts lies within the range delimited by the smallest and highest values among the individual forecasts. Two models are considered for these forecast combination methods, which are the fundamental-statistical model and the pure statistical model. Therefore, M is equal to two. When considering simple averaging, all the weights, $w_{i,m}$, are set to $1/M$, which is one half in this case.

The other two combination procedures are carried out for every hour of the day. Ideally, as mentioned earlier, the resulting hourly weights should favour the hybrid model when real prices are closer to their daily average value, whereas the pure statistical model's weights should be higher on hours of highest and lowest prices in the intraday pattern. This would be in accordance with the behaviours exhibited by market agents, given that peak and valley hours present higher chances of carrying out favourable transactions due to its higher volatility with respect to the remaining hours of the day.

The first hourly combination approach assigns weights inversely proportional to the square value of the forecast error, as proposed in (Bates & Granger, 1969). This can be therefore applied to every hour of the days pertaining to the validation set V (represented by V_i) as follows:

$$w_{i,m} = \frac{\left(\sum_{j\in V_i}\left(\hat{Y}_{j,m} - Y_j\right)^2\right)^{-1}}{\sum_{m'=1}^{M}\left(\sum_{j\in V_i}\left(\hat{Y}_{j,m'} - Y_j\right)^2\right)^{-1}} \qquad \forall i,m \qquad (3.11)$$

The forecast error in the above equation is simply the squared difference between the forecasted and real values in the validation set ($\hat{Y}_{j,m}$ and Y_j respectively). The numerator contains the error pertaining to one of the two models, while the denominator contains the sum of both model errors, and the negative exponential value is applied to both elements in the division so as to set the weight inversely proportional to the error. For a given forecasting period F, the corresponding weights pertaining to its associated validation set V are calculated. Therefore, these weights are different for every forecasting period, which provides a certain adaptability for the combined forecasts.

The last combination method is a Bayesian model averaging (BMA) method, which is carried out in a similar hourly manner as the previous combination technique in the sense that the validation dataset performance of the models is used in order to obtain hourly weights. It is worth noting that there are only two neural network models to combine. However, these forecasts are the result of the average of several individual replications of the neural network training procedure, as mentioned in the previous subsection.

Let *K* denote the number of individual forecasts (or replications, as per Table 3.2) carried out in the neural network models of both modelling procedures (hybrid and pure statistical). In order to ensure feasibility in terms of resolution time, the *K* forecasts of each model are divided into five subsets and their mean is later computed. As a result of a limited optimisation study, ten forecasts are used for combination in the BMA method in order to establish a suitable compromise between resolution time and accuracy.

In the BMA technique, the model space M_b is considered, which contains the previously mentioned ten forecasts: $M_b(b=1,2,\ldots,B),\ B=10$. The BMA method calculates the model weights for every considered combination option among the B model forecasts as the posterior probability in the same hour i of the days in the validation period (i.e. V_i). This posterior probability is represented in the following equation:

$$w_{i,b} = p(M_b|Vi) \qquad \forall i, b \qquad (3.12)$$

Therefore, by using the Bayes theorem, the probability distribution of the BMA forecast is computed as a weighted average of the posterior distributions:

$$p(\hat{Y}_i^c|V_i) = \sum_{b=1}^{B} w_{i,b} \cdot p(\hat{Y}_i{}^c|M_b, V_i) \qquad \forall i, b \qquad (3.13)$$

Finally, the posterior mean of the BMA forecast is represented by the following equation:

$$E[\hat{Y}_i^c|V_i] = \sum_{b=1}^{B} p(M_b|V_i) \cdot E[\hat{Y}_i{}^c|M_b, V_i] = \sum_{b=1}^{B} w_{i,b} \cdot \hat{Y}_{i,b} \qquad \forall i \qquad (3.14)$$

This method was applied in R and is based on the BMA package in R, which is developed by the authors of the work presented in (Raftery, Painter, et al., 2005).

3.3 Evaluation criteria and benchmarking models

As done in Chapter 2, several evaluation criteria and benchmarking models have been used in order to suitably validate the proposed hybrid methodology of this chapter. Forecasting performances have been compared as per the MAPE, MAE and RMSE error metrics. All errors have been computed with hourly precision. Furthermore, the Diebold-Mariano (DM) test was also used considering a 5% significance level.

Moreover, given the hybridisation methods that have been proposed throughout this chapter, several versions of the proposed hybrid model have been set as benchmarks. Without taking the combination procedures into account, there are two hybridisation methods that have been performed:

- PM_1: proposed model 1, the fundamental-statistical hybrid model with only the market clearing prices from the fundamental model.
- PM_2: proposed model 2, the fundamental-statistical hybrid with all four output variables of the fundamental model: market clearing prices as well as CCGT, coal and hydro output levels.

A comparison of these two versions of the fundamental-statistical hybrid model should provide useful information regarding the effect of the three additional outputs. The two models are then combined as per the three forecast combination methods with the pure neural network model, leading to six different models:

- $CMSA_1$: combined model with simple averaging 1, which involves model PM_1 and the pure neural network model.
- $CMSA_2$: combined model with simple averaging 2, which involves model PM_2 and the pure neural network model.
- $CMIEW_1$: combined model with inverse error weighting 1, which involves model PM_1 and the pure neural network model.
- $CMIEW_2$: combined model with inverse error weighting 2, which involves model PM_2 and the pure neural network model.
- $CMBMA_1$: combined model with Bayesian model averaging 1, which involves model PM_1 and the pure neural network model.
- $CMBMA_2$: combined model with Bayesian model averaging 2, which involves model PM_2 and the pure neural network model.

This makes 8 total versions of the proposed hybrid methodology of this chapter, which will be compared with the fundamental-statistical forecasting method of Chapter 2: the first benchmark (BM_1). The second benchmark (BM_2) is the pure neural network model forecasting method without any information provided by the fundamental model, as shown in the lower-left part of Figure 3.1. It should be highlighted that this pure neural network model includes all improvements explained within Section 3.2.2 and generally outperforms the pure neural network model of Chapter 2. The third benchmark (BM_3) is the pure fundamental model, which, as shown in Table 3.1, is also superior to that of Chapter 2 according to the enhancements described in Section 3.2.1.

The following and remaining three benchmarks, BM_4, BM_5 and BM_6, coincide with the last three benchmarks utilised in Chapter 2, which are, respectively: the slightly modified ARX model of (Misiorek, Trueck, et al., 2006); a double seasonal ARIMA model with the expected demand as exogenous variable; and a simple naïve approach that sets the price forecast to the actual value that occurred one week (or 168 hours) ago (see Section 2.4 for further and specific details).

3.4 Case study, results and discussion

As explained throughout Chapter 2, the case of the Iberian electricity market of the entire year 2017 presents several circumstances and situations in which the forecasting models may be put to the test. For instance, winter 2017 presents the highest standard deviation in prices ever experienced in the Iberian power exchange's recent history, whereas summer 2017 presented relatively stable market conditions. Therefore, providing suitable performance in all of these circumstances is a highly challenging task. Furthermore, given that this case study has been used in Chapter 2, it may be used again so as to compare its proposed hybrid methodology with the models developed within this chapter.

According to Figure 3.1, the fundamental model or cost-production optimisation model is first run in order to obtain the following outputs: market clearing prices as well as coal, CCGT and hydro unit generation outputs. This has been done for the considered training, validation and forecasting periods as per the timeline of Figure 3.2. Regarding the neural network forecasts, forecast horizons of one day in hourly resolution have been considered, i.e. d_2 is considered to be of one day.

Consequently, in order to perform this work's neural network forecast for January 1st of 2017, calibration data will be needed related to the months of December 2015, January 2016 and December 2016. Therefore, the cost-production optimisation model must be run for the months between December 2015 and December 2017 so as to have the necessary data to perform the neural network forecasts. As it is common in the literature, e.g. (Bento, Pombo, et al., 2018), the forecasting models have been evaluated for every season of the year, as well as a general assessment for the whole year 2017.

All of the proposed forecasting models and selected benchmarks have been tested for every day of the year 2017 under the conditions of the Iberian electricity market. Their MAPE, MAE and RMSE errors are displayed in Table 3.3, Table 3.4 and Table 3.5 respectively, including the combinations between both variants of the proposed hybrid methodology of this chapter (PM_1 and PM_2) and the pure neural network model (BM_2) with the simple average, inverse error weighting and Bayesian model averaging (CMSA, CMIEW and CMBMA respectively).

The bold values of these tables indicate the lowest forecasting error measures for every considered period of the year 2017 (i.e. the four seasons and the entire year). According to these results, the most accurate forecasting models seem to be

Model	Winter	Spring	Summer	Autumn	Average
PM_1 – Proposed hybrid model 1	11.40	**7.377**	4.746	6.689	7.534
PM_2 – Proposed hybrid model 2	11.68	8.106	**4.450**	6.812	7.744
BM_1 – Chapter 2's proposed hybrid model	12.83	8.840	5.016	6.764	8.341
BM_2 – Pure neural network model	**11.12**	7.804	4.605	6.834	7.575
BM_3 – Pure fundamental model	20.47	13.60	10.99	10.58	13.88
BM_4 – ARX of (Misiorek, Trueck, et al., 2006)	16.79	13.58	7.153	10.51	11.99
BM_5 – SARIMAX model	15.06	9.293	5.097	7.654	9.248
BM_6 – Simple Naïve approach	25.93	17.55	9.343	12.82	16.37
$CMSA_1$: PM_1+BM_2 simple average	11.21	7.488	4.584	**6.645**	**7.464**
$CMIEW_1$: PM_1+BM_2 inverse error weighting	11.21	7.490	4.584	**6.645**	7.465
$CMBMA_1$: PM_1+BM_2 Bayesian model averaging	11.20	7.475	4.586	6.722	7.478
$CMSA_2$: PM_2+BM_2 simple average	11.30	7.902	4.477	6.756	7.591
$CMIEW_2$: PM_2+BM_2 inverse error weighting	11.33	7.905	4.475	6.751	7.597
$CMBMA_2$: PM_2+BM_2 Bayesian model averaging	11.36	7.924	4.470	6.774	7.615

Table 3.3: Forecasting error in terms of MAPE (%)

the pure neural network model (BM_2) on winter, the neural network model with the market clearing prices from the fundamental model (PM_1) on spring, the neural network model with the four outputs from the fundamental model (PM_2) on summer and the simple average combination of PM_1 and BM_2 ($CMSA_1$) on autumn and generally during the entire year 2017. Nevertheless, the performance of the inverse validation error weighting combination of PM_1 and BM_2 ($CMIEW_1$) is almost equal to that of the simple average between the same models ($CMSA_1$).

The pure neural network model (i.e. BM_2) is capable of outperforming all other models in the most unstable period of 2017 thanks to its innate adaptability, as well as the similar-day and the calibration period shortening procedures that have been paired with it. This may suggest that electricity prices were less driven by market fundamentals than by agent strategic behaviours. However, adding fundamental-related information to this neural network model proves useful in all the other seasons of 2017 and throughout the year 2017 in general.

The differences between the predictive accuracies of PM_1 and PM_2 indicate the benefits and drawbacks of incorporating additional variables from the fundamental model (i.e. market clearing prices alone, PM_1, or also hydro as well as coal and CCGT unit generation levels, PM_2). The highest error differences can be seen between spring and summer. Additionally, model PM_1 seems to outperform on the other two seasons and on the entire year 2017. This suggests that the price formation in summer is more characterised by market fundamentals and thus the contribution provided by the fundamental model is more advantageous.

Finally, the hybridisation approach of PM_1 between the fundamental model (BM_3) and the neural network model (BM_2) reduces overall forecasting error as a result of the synergy between the neural network model's adaptability and the

Model	Winter	Spring	Summer	Autumn	Average
PM_1 – Proposed hybrid model 1	4.641	**2.641**	2.197	3.350	3.199
PM_2 – Proposed hybrid model 2	4.756	2.882	**2.070**	3.453	3.282
BM_1 – Chapter 2’s proposed hybrid model	5.137	3.068	2.359	3.331	3.465
BM_2 – Pure neural network model	**4.562**	2.826	2.136	3.440	3.233
BM_3 – Pure fundamental model	10.81	5.696	5.066	5.872	6.842
BM_4 – ARX of (Misiorek, Trueck, et al., 2006)	6.838	4.765	3.262	5.066	4.972
BM_5 – SARIMAX model	8.113	4.150	2.473	4.454	4.780
BM_6 – Simple Naïve approach	10.53	6.225	4.266	6.387	6.828
$CMSA_1$: PM_1+BM_2 simple average	4.577	2.690	2.123	**3.329**	**3.172**
$CMIEW_1$: PM_1+BM_2 inverse error weighting	4.577	2.691	2.123	3.330	**3.172**
$CMBMA_1$: PM_1+BM_2 Bayesian model averaging	4.583	2.693	2.125	3.376	3.186
$CMSA_2$: PM_2+BM_2 simple average	4.610	2.832	2.079	3.409	3.224
$CMIEW_2$: PM_2+BM_2 inverse error weighting	4.624	2.831	2.078	3.406	3.227
$CMBMA_2$: PM_2+BM_2 Bayesian model averaging	4.618	2.831	2.077	3.418	3.228

Table 3.4: Forecasting error in terms of MAE (€/MWh)

Model	Winter	Spring	Summer	Autumn	Average
PM_1 – Proposed hybrid model 1	5.415	**3.134**	2.651	4.022	3.796
PM_2 – Proposed hybrid model 2	5.479	3.407	**2.517**	4.129	3.874
BM_1 – Chapter 2’s proposed hybrid model	5.921	3.658	2.840	4.003	4.096
BM_2 – Pure neural network model	**5.308**	3.342	2.588	4.089	3.823
BM_3 – Pure fundamental model	12.19	6.685	5.759	7.158	7.927
BM_4 – ARX of (Misiorek, Trueck, et al., 2006)	7.809	5.552	3.885	6.055	5.814
BM_5 – SARIMAX model	10.84	5.585	4.531	4.959	6.460
BM_6 – Simple Naïve approach	11.48	7.092	5.030	7.567	7.773
$CMSA_1$: PM_1+BM_2 simple average	5.337	3.194	2.575	**3.986**	**3.764**
$CMIEW_1$: PM_1+BM_2 inverse error weighting	5.338	3.194	2.575	**3.986**	**3.764**
$CMBMA_1$: PM_1+BM_2 Bayesian model averaging	5.349	3.207	2.583	4.032	3.783
$CMSA_2$: PM_2+BM_2 simple average	5.334	3.349	2.525	4.069	3.810
$CMIEW_2$: PM_2+BM_2 inverse error weighting	5.350	3.349	2.524	4.066	3.813
$CMBMA_2$: PM_2+BM_2 Bayesian model averaging	5.359	3.356	2.523	4.092	3.823

Table 3.5: Forecasting error in terms of RMSE (€/MWh)

equilibrium price level provided by the cost-production optimisation model. Regarding the forecast combinations between PM_1 or PM_2 with BM_2, they do not seem to provide lower errors on some specific periods of the year 2017 when compared to both individual models prior to the combination, but they do when considering the entire year 2017, which is mainly due to the results in autumn. Therefore, the accuracy improvement as a result of the forecast combination methods coincide with the hypotheses provided in Chapter 2 and the previously mentioned synergy.

Furthermore, the results for every combination method seem to indicate that the simple average is most beneficial, albeit closely followed by the inverse validation error weighting procedure. As in other works in the same forecasting

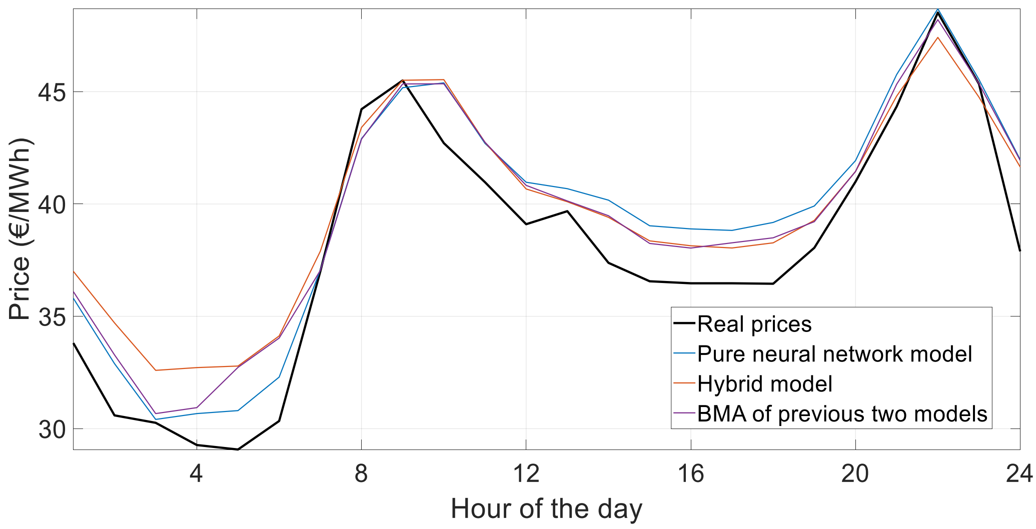

Figure 3.5: Forecasts of the proposed hybrid and pure NN models as well as the BMA for 06/Apr/2017

context, e.g. (Bordignon, Bunn, et al., 2013), the simple average technique seems to be challenging to outperform, even with more sophisticated methods.

Furthermore, the forecast combination methods performed well in the situations that were identified in Chapter 2 and encourage said combination methods. One of the most illustrative examples is the previously presented case of the 6th of April 2017 that is depicted in Figure 3.4. In this case, the most beneficial model combination is $CMABMA_2$ (i.e. Bayesian model averaging combination of the neural network model with PM_2). This is shown in Figure 3.5, which shows a combination that favours the neural network model on the first four hours and the last three, while using the hybrid forecast on the middle hours of the day. The MAE errors are 1.891, 1.665 and 1.614 €/MWh for the hybrid, neural network and BMA combination respectively.

Finally, a Diebold-Mariano (DM) test has been carried out so as to verify the statistical significance of the error measures shown in Table 3.3, Table 3.4 and Table 3.5. Figure 3.6 shows the DM test results for the most outperforming models. Its colour-bar[38] indicates the value of the DM test statistic that assesses if the model on the top header significantly outperforms the model on the left header. Given that this test is run with a 5% significance level, the corresponding critical value is 1.96. Therefore:

- DM test statistic < -1.96 implies significant outperformance
- DM test statistic > 1.96 implies significant underperformance
- Otherwise no significant out- or underperformance

[38] According to the colour-bar, green and red colours mean significant out- and underperformance, respectively.

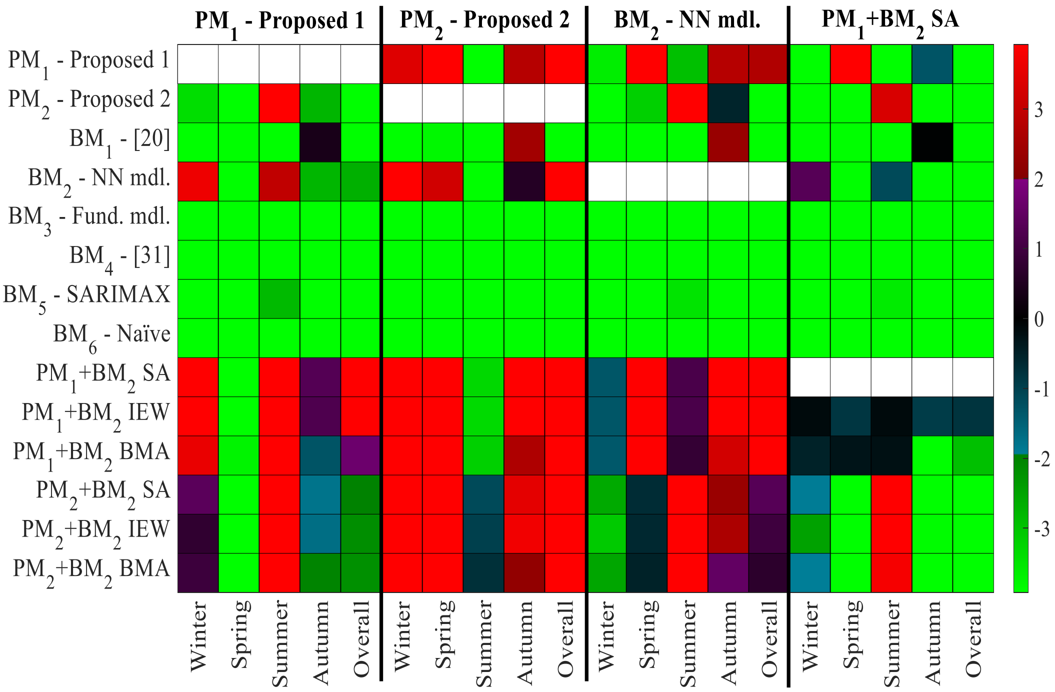

Figure 3.6: DM test for PM_1, PM_2, BM_2 and the simple average of PM_1 and BM_2 ($CMSA_1$)

According to Figure 3.6, the hybrid model of PM_1 is the most outperforming model during spring and shows suitable overall performance when considering the entire year 2017. It is also one of the few models not significantly bested by the hybrid model of Chapter 2 (i.e. BM_1) during autumn. Regarding model PM_2, it is not significantly outperformed by any other model during summer, but shows otherwise a slight underachievement when tested against PM_1.

The most remarkable model on the winter is the pure neural network model, although its combinations with PM_1 are almost significantly bested by it. Furthermore, the simple average method between PM_1 and BM_2 seems to generally outperform every other model when considering the whole year 2017, which is coherent with several authors in the literature that claim that the simple average offers a suitable diversification in which model misspecification is diminished.

3.5 Conclusions

In this chapter, a novel methodological approach that is composed of a fundamental-statistical hybrid electricity market price forecasting model has been presented. The components of this hybrid methodology were individually enhanced and demonstrated a superior forecasting performance than the individual components of Chapter 2's proposed forecasting model. More specifically, the cost-production optimisation model was improved by considering CCGT and coal units separately and by estimating their variable costs according to their past bids and other relevant factors, such as commodity prices. On the other hand, the neural network model was improved by incorporating a validation period selection via a customised similar-day method on its training procedure, and its calibration period was shortened based on validation set error on unstable periods.

Another contribution of this chapter is the proposal of several hybridisation schemes, which involve using different variables from the cost-production optimisation model as input data to the neural network model, as well as performing distinct forecast combination procedures with a pure neural network model. The following observations and findings summarise the conclusions drawn in this chapter.

First of all, the proposed hybrid model of this chapter is capable of simultaneously benefitting from the neural network's adjustability for volatile prices and from the equilibrium price level provided by fundamental-related information. Furthermore, highly unstable periods, such as early 2017, can be dealt with shortened calibration windows in order to further increase the neural network model forecasts' adaptability.

In addition, on periods of more relative stability, such as summer 2017, electricity market price behaviours are responding more to market fundamentals, and thus incorporating additional variables to the hybrid model, such as thermal/hydro generation unit levels, proves advantageous. Furthermore, on the other periods and generally throughout the entire year 2017, a simple average combination method between the fundamental-statistical hybrid model and the pure neural network model further increases forecasting accuracy, providing a heightened and better balanced synergy between the considered fundamental and statistical approaches.

In short, the unique set and combination of methodologies that constitute this chapter's proposal has demonstrated a suitable performance for short-term electricity market price forecasting in the recent case of the Iberian electricity power exchange throughout the year 2017, while also outperforming other benchmark models.

However, some of the methodologies employed in this work may be modified or extended in order to explore some potential improvements. For example, given that the results suggest that a simple averaging procedure outperforms the other forecast combination methods, there may be some useful evidence regarding specific trends and behaviours in electricity prices that can lead to a more optimal forecast combination method.

Furthermore, the similar-day method can be altered in order to enhance the similarity assessment based on other explanatory variables (i.e. apart from expected demand, expected wind power and expected temperature). Moreover, the calibration window length shortening procedure may also be done by analysing market stability, which is correlated to some of the considered explanatory variables, such as fuel prices (e.g. unstable market regimes with high volatility seem to be frequently related to high natural gas prices, as observed in early 2017).

Moreover, the calibration period selection of the neural network may also be carried out by establishing links to market regimes or statistical criteria, such as pattern detection or structural break tests. This would result in an as-of-now unseen calibration period selection method driven by robust criteria, at least in the electricity price forecasting context. Additionally, including a suitable variable selection procedure for non-linear contexts may also prove beneficial so as to further analyse the effect of incorporating more variables of the fundamental model to the statistical forecasting method.

Chapter 4

Short-Term Electricity Price Forecasting Considering Market Structural Breaks

As seen on previous chapters, fundamental-statistical hybrid electricity market price forecasting models have demonstrated a suitable performance by providing a forecast in accordance with most of the factors that drive the electricity system's prices. However, the results and findings of Chapter 3 suggest that, in order to provide a more robust forecast, guidance criteria should be set based on past and expected price behaviours and trends. Moreover, the degree of market structural breaks in recent history calls for a versatile methodology that is able to adapt to any market regime. This encourages the need for an advanced hybridisation method that is able to provide an accurate forecast driven by trends and behaviours provided by both the fundamental and the statistical component of the forecasting methodology. As a means of dealing with this issue, the main focus of this chapter[39] is the selection of the calibration data periods that are used to train the neural network model that screens out periods related to dissimilar market circumstances. This aims to enhance the synergy related to the transfer of fundamental market regime indicators to the statistical forecasting model. Furthermore, calibration period selection is frequently disregarded, as most of the forecasting applications in this context lack robust criteria related to the training period selection. By adding a robust calibration period selection to one of the hybrid fundamental-statistical models that have been developed throughout Chapter 3, this chapter's proposed forecasting methodology is able to outperform other benchmark models in the real case of the Iberian electricity market of 2017, which presents a considerable number of market structural breaks and circumstances.

[39] The contents of this chapter are based on the working paper (de Marcos, Bunn, et al., 2019)

4.1 Introduction and literature review

In the current context of electricity markets, several factors are interfering with the stability of the system, thus increasing complexity and instability, which traders and practitioners have to confront. During the monopolistic era that ended in the late 20th Century for most power exchanges, these system alterations were almost entirely due to yearly seasonal effects. Therefore, the trends that one had to be concerned about when it came to developing price forecasting models, which mainly consisted of mean-reversion techniques, were only a few. The most typical example is not relying on past summer prices when forecasting winter prices, whose price levels were significantly different from each other. However, price levels have lately experienced much more sudden variations than those of the past, and thus market regimes may vary more often than seasons do throughout an entire year. For instance, an immediate and sharp increase in wind power generation exerts a sudden and strong downward pressure on electricity market prices[40]. In such situations, electricity price forecasting models will fail if no specific considerations are made. Considering that these spikes are not punctual but recurring phenomena, this issue should not be understated.

One of the first forecasting methodologies that consider price spikes in electricity market contexts is the work presented in (Huisman & Mahieu, 2003), which is based on previous works mainly applied to stock markets. The authors propose a regime switching model that separates price spikes from normal mean-reverting prices. Further regime switching methodologies were developed later, such as threshold autoregressive models (Rambharat, Brockwell, et al., 2005), Markov regime switching models (Kosater & Mosler, 2006; Cruz, 2013) and smooth transition regime switching models (Chen & Bunn, 2010). In addition, some authors have proposed establishing an additional regime in which price spikes are split into upward and downward spikes, as performed in (Paraschiv, Fleten, et al., 2015). Moreover, some works, such as (González, Contreras, et al., 2012), consider reserve margin[41] as regime distinctive feature instead of price levels, given that the relationship between reserve margin and electricity prices is significant, especially on markets with scarce overcapacity.

Considerable changes in market regimes can also be perceived as structural breaks in the general econometrics and statistics context. A structural break in a time-series splits it into two segments, where the coefficients of a linear regression model should be significantly different between said segments. One of

[40] This holds true as long as no significant changes are exhibited regarding other important aspects of the electricity system. For instance, if demand levels are significantly high, prices may not fall as much as expected when wind generation increases.

[41] An electricity system's reserve margin is the difference between its generation capacity and the demand, which is inversely proportional to prices in markets where the most expensive cleared bid is set as the market price.

the first works in this regard is the Chow test, proposed in (Chow, 1960), which is a statistical test that assesses if a certain moment in a time series is a significant structural break. However, if several structural breaks are actually present in a time-series, applying the Chow test as a means of identifying these structural breaks is a highly challenging task. Furthermore, if the moment of a structural break is not known *a priori*, the Chow test is also ineffective. In these situations, there are other structural break tests that are able to consider a multiple but unknown number of structural breaks, such as the methodologies presented in the works of (Page, 1954; Andrews, 1993; Bai & Perron, 2003; Zeileis, Kleiber, et al., 2003). Furthermore, the concept of structural breaks has been rarely applied to electricity price forecasting contexts (Apergis & Lau, 2015). However, the idea of market structural changes in the same sense has been mentioned in (Bello, Reneses, Muñoz, et al., 2016), which is related to the physical and regulatory alterations of the electricity system that can be considered by means of market equilibrium models, as mentioned on previous chapters of this thesis. Furthermore, as seen in Chapter 3, there may be ways of considering market regimes in order to obtain a more optimal forecast combination and calibration window selection procedures.

Calibration period selection is relevant when market regimes are constantly varying, as mentioned in Chapter 3. The authors of the work presented in (Pesaran & Timmermann, 2007) point out the issues of structural breaks in calibration datasets for econometric or statistical models and propose a forecast combination method based on averaging forecasts of the same model trained with different calibration window sizes. This idea was taken to the electricity market price forecasting context in the work of (Marcjasz, Serafin, et al., 2018), which proved to be useful in linear regression models such as ARX. This usefulness was also demonstrated in the probabilistic forecasting field in (Serafin, Uniejewski, et al., 2019). These authors claim that calibration period selection in electricity price forecasting contexts is frequently disregarded. However, their work's methodology does not consider structural breaks or any guidance criteria as to the calibration window selection for the regression models in an *ex-ante* manner. A recent work in the literature that considers calibration period selection in electricity price forecasting contexts with neural network procedures is presented in (Bento, Pombo, et al., 2018), as well as the methodology proposed in Chapter 3 or (de Marcos, Bello, et al., 2019b). The former uses a training set involving the seven days prior to the forecasting day adds three extra days based on the similarity with respect to the day immediately prior to the forecasting day in terms of daily price patterns. The latter utilises a modified version of the similar days method proposed in (Mandal, Senjyu, et al., 2006) in order to select the 12 most similar days in a predefined 4-month calibration period as validation set according to exogenous variables available at the moment of the forecast such as expected demand and temperature. However, both of these methods are more encouraged by neural network overfitting issues rather than by market regime alterations.

This subject is central to the purpose of the work and is also a highly critical aspect of electricity market price forecasting in the light of the significant instability issues that power systems are currently facing. To the best knowledge of the author, there is currently no work in the electricity price forecasting literature that considers a dynamic calibration data period that is supported by robust criteria according to explanatory variables associated with market circumstances or regimes. Moreover, this should be carried out in an *ex-ante* manner, thus removing the need for establishing or predefining said period based on empirical analyses of out-of-sample forecasts. For instance, as remarked in (Chen & Bunn, 2014), although Markov regime switching's model fitting provides suitable in-sample results, their out-of-sample generalisation capabilities may not be acceptable due to overfitting issues. This matter is also prominent in neural networks, which may arise when using data pertaining to dissimilar market regimes in the training dataset and thus the neural network may elaborate biased neuron weights that cause it to underperform on the forecasting period.

Considering all the previously mentioned facts and issues, the work presented in this chapter attempts to provide a unique forecasting technique that properly addresses market structural breaks by both utilising a fundamental-statistical hybrid model and an enhanced calibration period selection. By doing so, the market-related information provided by the fundamental model is exploited to a higher extent. This hybrid model is identical to one of the models proposed in Chapter 3 which involves a short-term fundamental-statistical electricity market price forecasting model that is composed of an hourly cost-production optimisation model whose outputs provide market-related information to a neural network model. It should be noted that the work of this chapter is based on the working paper (de Marcos, Bunn, et al., 2019). The methodologies proposed in this chapter that contribute to the betterment of this base model are listed below:

1. Prior to the neural network forecast, the neural network period, which is initially set to a very large window, is filtered by means of a structural break analysis method. More specifically, the behaviours exhibited by actual electricity prices in this window are examined and periods whose prices significantly differ from those immediately prior to the forecasting period (i.e. most recent prices) are discarded.

2. Furthermore, the hourly trends in the actual forecasting period according to market regime related variables are evaluated via a K-means clustering procedure. The hours of the initial neural network calibration period whose assigned cluster coincides with that of the hours in the forecasting period are included in the previously filtered calibration period by the structural break analysis method. This combination of training window

selection techniques is thus carried out *ex-ante* and therefore provides a dynamic calibration dataset regardless of any particular predefinitions.

3. The proposed set of methodologies is put to the test on the real and full-scale case of the Iberian electricity market of 2017 and the usefulness of the calibration period selection techniques are assessed separately. Moreover, the performance of this work's proposed model is compared with that of other well-recognised models in the literature, as well as recently proposed forecasting techniques.

The remainder of this chapter is organised as follows. The proposed methodology is fully described in Section 4.2. Section 4.3 presents the utilised evaluation criteria and the benchmarking models that have been used in order to validate the proposed methodology of this chapter. Section 4.4 contains the case studies in which all the forecasting methods were tested as well as the corresponding results and comments. Section 4.5 presents the conclusions that were drawn in this work, including the suggestions for potential extensions of this chapter's proposed methodology. Finally, Appendix B contains several analysis that attempt to detect any potential overfitting issue in the final neural network model that is proposed in this chapter.

4.2 Methodology

Essentially, this work's proposed methodology is comprised of the methods displayed in Figure 4.1, all of which have been tested on a real-size power exchange with complex price dynamics: the Iberian electricity market. The first phase of the methodology is associated with its fundamental component, the cost-production optimisation model. The next stage involves several approaches that aim to enhance the final step of the methodology: an artificial neural network model. More specifically, a robust calibration period selection for the neural network model has been proposed and developed, which is driven by past

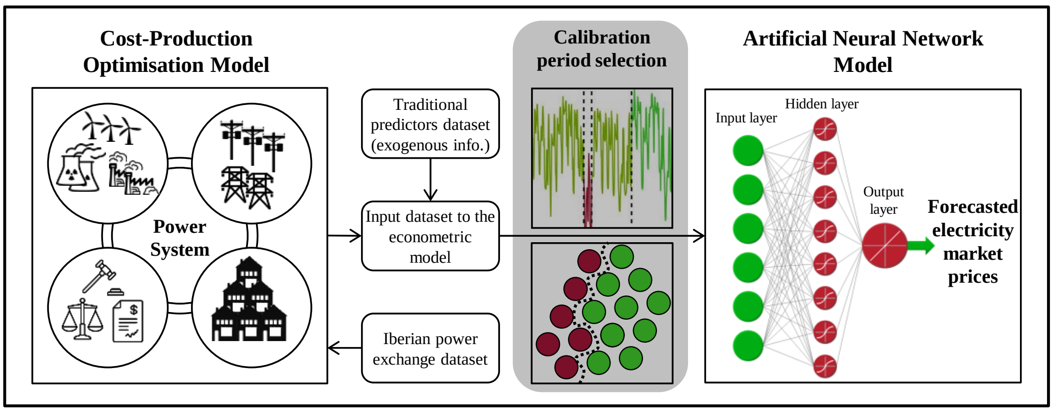

Figure 4.1: Overview of the proposed hybrid forecasting methodology

behaviours in prices and the values of predictors that are available at the moment of the forecast. Each element of the proposed methodology is thoroughly explained on the following subsections.

4.2.1 Cost-production optimisation model

In order to consider physical elements, regulatory limits and the operation of the market in this work's final electricity price forecast, a cost-production optimisation model, which is based on the Iberian power exchange, is run. As mentioned in previous chapters, the pertinent information is obtainable at the transparency platforms of the Spanish System Operator (https://www.esios.ree.es/en) and of the ENTSO-E (https://transparency.entsoe.eu). This optimisation problem, which is similar to the one presented throughout Chapter 2, is solved so as to estimate the electricity market price as a result of the market clearing (see Section 2.3.1 for a detailed description of this model). These prices are also known as system marginal prices, and they can be ascertained from the dual variable of the demand and generation balance constraint. Furthermore, it was observed in Chapter 3 that considering thermal units separately allows for a worthy exchange between resolution speed (lower) and accuracy (higher). More specifically, a week is solved in 7.4 seconds (up from 3.91) and the forecasting error is reduced by approximately 33%.

Moreover, the optimisation problem is solved via relaxed mixed-integer programming (RMIP) in order to consider all units' variable costs and not only those of the committed ones. Although this may not be appropriate for other purposes (e.g. the resulting unit uptimes may prove infeasible in reality), this work's objectives do not consider generation unit scheduling, but mainly the market clearing prices as well as the CCGT, coal and hydro production levels. These variables, among others, are later used as input data for neural network forecasting models, as explained in the following subsections.

4.2.2 Period selection

As mentioned previously, evaluating the input data periods is an often overlooked topic. In the current electricity price forecasting contexts, the importance of the ongoing market structural breaks should not be understated. This work's calibration period selection methodology provides a suitable and unique solution to this issue, allowing the neural network model to handle only the necessary data by paying attention to the relevant circumstances or regimes present in the power system at the moment of the forecast. This methodology is split into three steps, as explained in the following three subsections. Firstly, the initial dataset window is split into several segments according to the structural breaks revealed in the historical price time series. Secondly, this reduced dataset is extended by an hourly clustering algorithm that links past hours with those of the forecasting period according to explanatory information that is based on market fundamentals. Although the Iberian electricity market has been used as case study for this methodology, its versatility makes it applicable to any power system as well as any forecasting technique that involves calibration or training procedures.

4.2.2.1 Structural breaks

Before applying any filtering method, the initial dataset period arrangement for the neural network model should be oversized. In this case, 13 months prior to the forecasting period are taken (i.e. a 13-month rolling window dataset), which may be perceived as a large calibration dataset for neural network models if hourly precision is considered. The fact that structural patterns change in electricity prices throughout this 13-month period is not in question, not only due to several seasonal effects in the system, but also depending on abrupt market condition fluctuations or other structural breaks. An example can be seen in Figure 4.2, which shows the evolution of the Iberian electricity market prices during the autumn of 2016. It can be observed that early autumn is significantly different than late autumn. When it comes to forecasting on late autumn (e.g. shortly after December 6th), one should consider discarding the periods with the lowest prices, as they clearly correspond to other market circumstances.

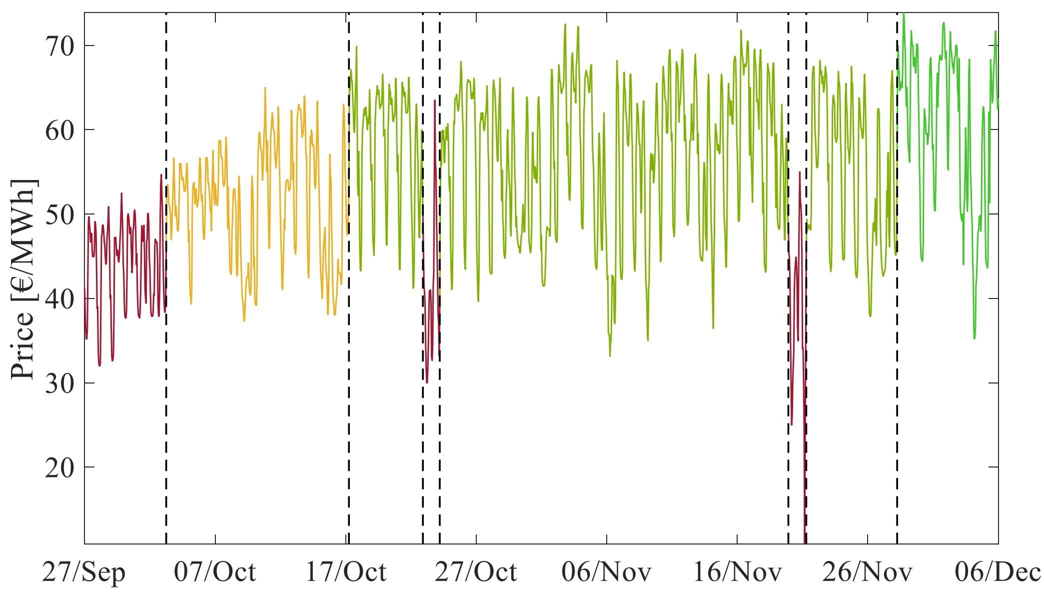

Figure 4.2: Iberian electricity market prices during autumn 2016

The different market circumstances are separated by the vertical lines, which correspond to the structural breaks that were previously mentioned. These structural breaks have been computed based on the “strucchange” package in R (available on https://cran.r-project.org/web/packages/strucchange/index.html) whose theoretical foundations are explained in the work presented in (Zeileis, Kleiber, et al., 2003). This structural breaks package involves a function that calculates the optimum number of breakpoints in a time-series. As mentioned earlier, structural breaks split a time-series into several segments which feature significantly different coefficients and independent term for a linear regression model that estimates said time-series[42]. The goal of this optimisation algorithm is the minimisation of the total residual sum of squares of the regression models associated with each segment. This optimisation model is constrained by two parameters: minimum segment size and maximum number of breakpoints.

It is worth noting that evaluating a 13-month hourly dataset is greatly cumbersome if high precision is desired. Therefore, the number of breaks must be limited. In order to capture most of the structural breaks in the 13-month price time-series, the breakpoints were computed in two sequential runs in which daily and intraday trends are assessed in each run. Therefore, the first run involves a daily arrangement of the 13-month dataset with a minimum breakpoint distance of one week. On the other hand, the second run involves an hourly arrangement of the remaining days as a result of the first run. When it comes to selecting or discarding segments, the following assumption is considered: all the hours within the period between the most recent structural break and the forecasting period (inclusive) belong to the same market regime. Therefore, the most recent segment is never discarded.

After computing the breakpoints in a run, the input dataset is divided into segments, which are compared to the most recent segment in terms of price average. Therefore, a certain persistence is considered in which the price average of the most recent segment will not be significantly different to that of the forecasting period. In order to discard sufficiently dissimilar periods that belong to other market circumstances, the periods whose price average falls outside the interval $\mu \pm \sigma$, where μ and σ represent the most recent period’s price average and price volatility (i.e. standard deviation) respectively, are discarded. As a result, this unique manner of performing the methodology of (Zeileis, Kleiber, et al., 2003) provides an efficient way of detecting structural breaks in a 13-month dataset with hourly precision, as well as discarding significantly different periods as per price behaviours.

Figure 4.3 depicts the resulting calibration period selection according to the structural breaks method for a 13-month initial dataset window. On the one hand, the left y-axis is related to the curve, which represents the real Iberian electricity

[42] In this case, the linear regression model is merely “price equals a constant”, and thus the constant is the varying element in the segments that are separated by the structural breakpoints.

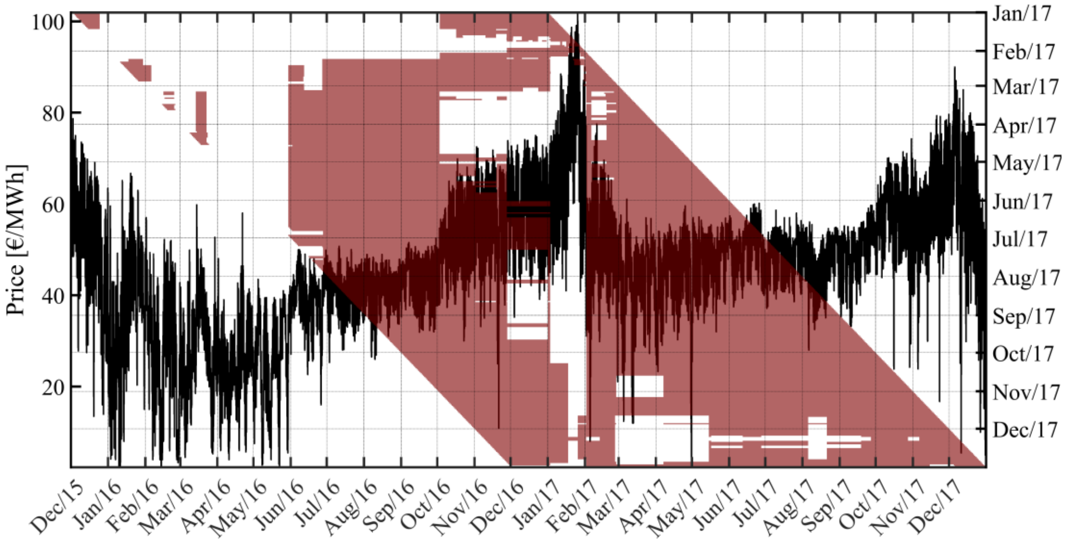

Figure 4.3: Structural breaks algorithm period selection

market price from December 2015 up to December 2017. On the other hand, the shape indicates the calibration periods (x-axis) selected for a certain forecasting day (right y-axis). Several patterns can be observed in Figure 4.3, for instance, given that early 2017 was characterised by uncommonly high prices, the selected calibration periods were much shorter than those of later 2017. Furthermore, January's peak is generally discarded from calibration datasets when forecasting days within 2017's remaining months. Moreover, summer 2016 is considered while forecasting summer 2017. Therefore, this algorithm is able to detect periods in the past that are highly dissimilar to the forecasting period.

A step-by-step example is shown in Figure 4.4 and Figure 4.5, which represents the structural breaks calibration selection procedure that results in one of the most segmented windows: the case of 19th of March 2017. The periods coloured in red are those that are discarded according to the $\mu \pm \sigma$ criterion. Most periods are discarded on the first step (i.e. daily precision and daily averages), which correspond to extremely low or high prices when compared to the average price

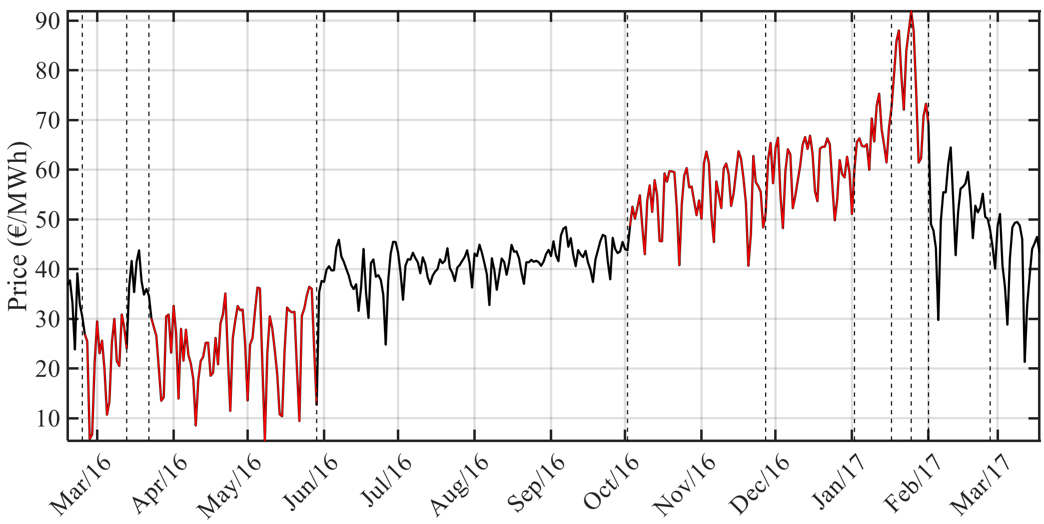

Figure 4.4: Example of the structural breaks period selection for 19/Mar/2017 - Step 1

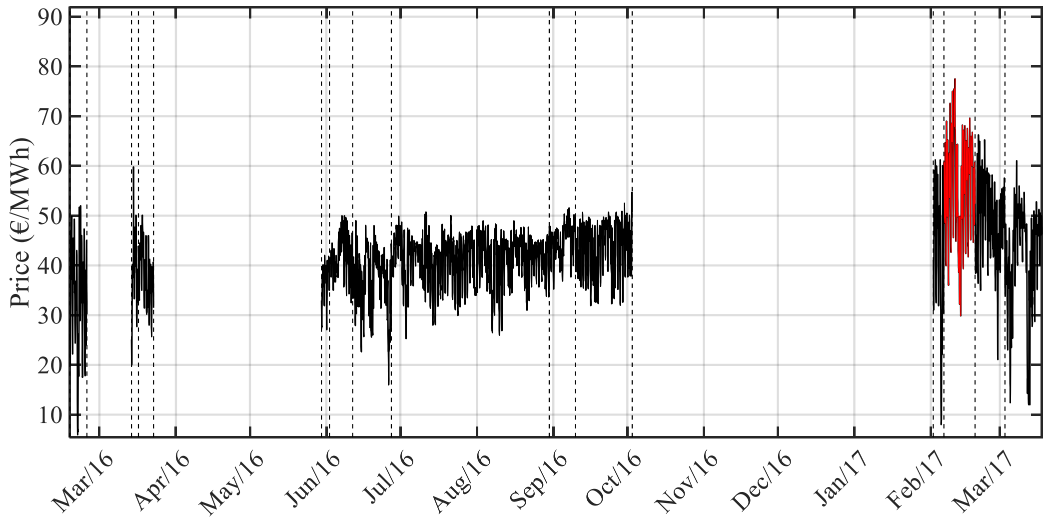

Figure 4.5: Example of the structural breaks period selection for 19/Mar/2017 - Step 2

level of the most recent segment. On the hourly step, the red segment's price mean is above the most recent segment's mean plus its standard deviation, and thus it is discarded. As a result, the training dataset period for the neural network is split into five windows.

However, a significantly higher level of accuracy that leads to an unfeasible level of computational burden is needed in order to spot much more sudden price changes, such as those depicted in Figure 4.2 and Figure 4.5. These occurrences should be discarded from training algorithms in forecasting models in order to reduce noise and overfitting. However, if a reliable wind forecast is available in advance, perhaps it is worth considering these outstanding periods if the forecasting period presents high wind power generation, as these occurrences usually exert a strong downward pressure on prices, which has been recently confirmed in (Aineto, Iranzo-Sánchez, et al., 2019) for the case of the Iberian power system. Furthermore, this structural breaks algorithm is based on the actual electricity prices in the historical data series up to immediately prior to the forecasting period, so this yields insufficient guidance as to what characterises said forecasting period.

These facts call for additional guidance regarding calibration period selection, and such guidance should be based on exogenous variables whose value is known at the forecasting period, for instance, wind production and demand forecasts that are available at the transparency platform of the Spanish System Operator (https://www.esios.ree.es/en).

4.2.2.2 Hourly clustering

First of all, a study was conducted so as to determine the most relevant factors regarding market conditions during the forecasting period. The variable with most predictive content is the estimated or expected price from the fundamental model, which reflects several aspects of the operation and the dynamics of the market. Furthermore, another piece of information that is often resorted to, not only in electricity markets, but also in other commodity markets, is future prices, given that many works in the literature suggest that market futures are linked to upcoming spot prices, such as (Steinert & Ziel, 2019). However, the day-ahead electricity futures[43] of the Iberian power exchange proved less useful than market clearing prices given that they lack intraday trends and are unsusceptible to abrupt price changes. Another variable that responds well to sudden market condition disruptions is the expected thermal gap[44], which represents the difference between the expected demand and the expected renewable generation corresponding to wind and solar energy. Prices are bound to fall if the gap is low. Although the expected market clearing prices also capture this effect, the expected thermal gap contains a higher level of short-term dynamics information and thus provides intraday trends with a higher definition. Moreover, the expected temperature may also be used as a means to remove periods with significantly different temperature (e.g. discarding winter when forecasting summer).

A K-means clustering method may be applied so as to take these three exogenous variables into account (estimated market clearing prices from the fundamental model, expected thermal gap and expected temperature) and relate the hours in the forecasting period to those of the training period. The K-means clustering application involves the ascertaining of centroids as per the values of the three aforementioned variables throughout the 13-month initial dataset. Consequently, each hour in the dataset belongs to the closest centroid in terms of squared Euclidean distances in the 3D plane formed by said three variables. Depending on the predefined number of clusters, the centroids are placed so as to minimise the total quantisation error or the sum of squared Euclidean distances, and thus a greater number of clusters lead to lower quantisation errors and higher complexity levels. In order to appropriately set the number of clusters, the K-means algorithm is computed for several numbers of clusters and, by means of a Pareto optimal frontier procedure, a suitable compromise between complexity level and total quantisation error is obtained[45]. Finally, the clusters that include

[43] These day-ahead electricity futures refers to the futures that are traded with an expiry or delivery date set to the following day. The reference to “day-ahead” in these futures should not be confused with the day-ahead wholesale electricity market.

[44] This is the case for electricity markets such as the Iberian power exchange, although reserve margin levels are more relevant in power systems of significantly lower overcapacity.

[45] In most cases, 4 clusters were selected. This agrees with the “n+1” criterion that is frequently seen in the literature, given that 3 variables are involved.

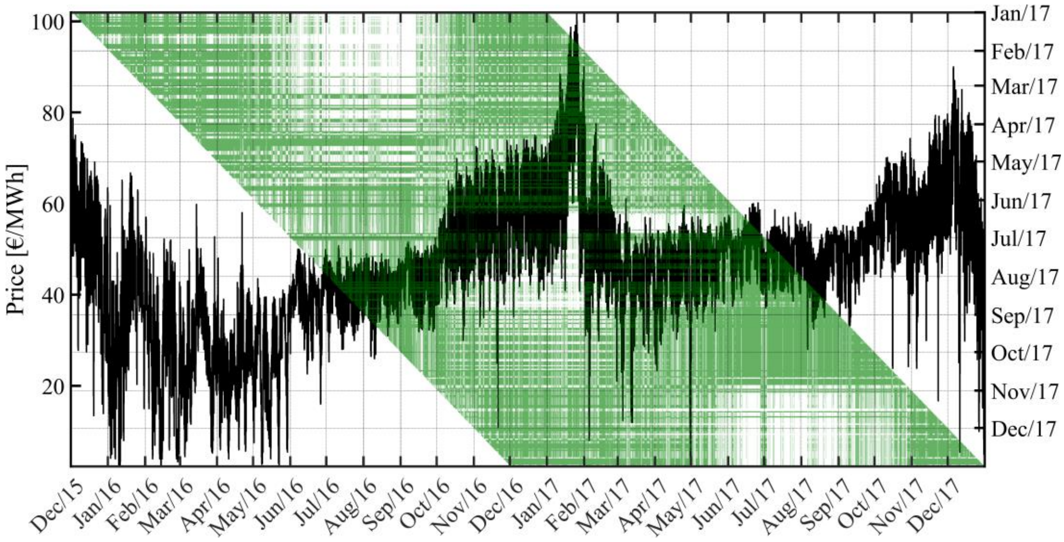

Figure 4.6: Hourly clustering method period selection

the hours of the forecasting period are deemed relevant and thus the hours of the input dataset that do not belong to said clusters are discarded.

The hours selected by the K-means clustering method are shown in Figure 4.6. This new shaded shape is somewhat hollow given that the clustering has been performed hourly. This provides useful information as to what intraday patterns in the past are the most similar to that of the forecasting period. This filtered dataset may be used as calibration data for any model, whether it is a hybrid or a pure statistical forecasting methodology. However, some hours on the most recent periods are generally discarded which most likely belong to the same market regime. Therefore, it would be beneficial if the union between the structural break and the hourly clustering period selection methods is performed, as depicted in Figure 4.7.

The combination of these period selection algorithms aims to provide a calibration dataset that contains two sets of information: the recent dynamics such

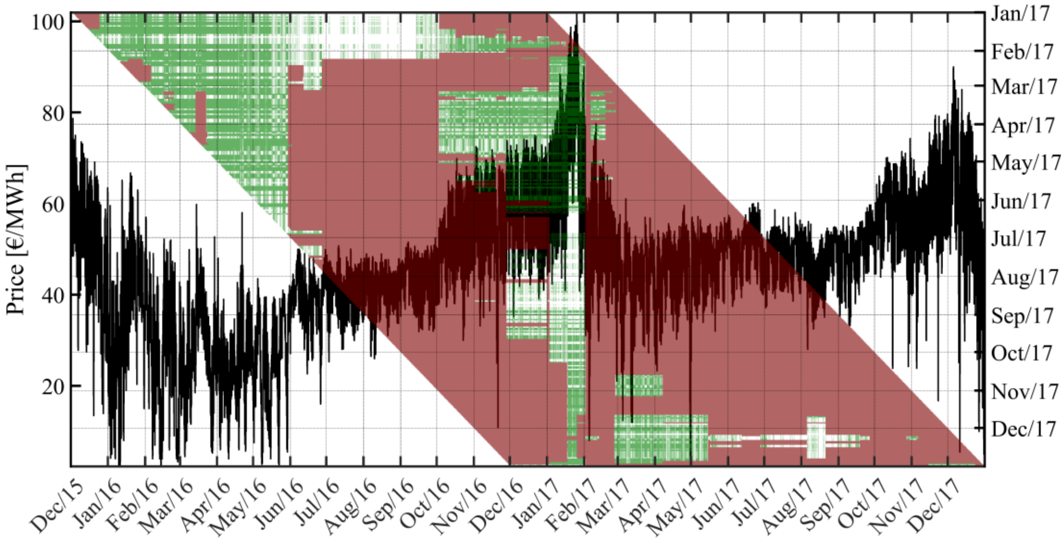

Figure 4.7: Union of the hourly clustering and the structural breaks period selection methods

as agent strategic behaviours provided by the structural breaks method and the patterns that are driven by market fundamentals yielded by the hourly clustering technique. All in all, this combined dataset discards the information pertaining to dissimilar market regimes according to recent price behaviours and forecasted market regime indicators. As seen in both Figure 4.3 and Figure 4.6, there are some hours which were selected by both algorithms. However, for the sake of simplicity, those hours have been coloured in the same colour as the hours selected by the structural breaks period selection algorithm in Figure 4.3.

4.2.2.3 Neural network validation set

Considering the length of the filtered dataset, a validation set is obtained as per the similar-day method performed in Chapter 3, which selects days in the historical dataset as per their similarity with respect to the forecasting period in terms of daily patterns regarding exogenous variables such as expected demand, as explained in Section 3.2.2.2. For this work's case, the top 20% similar days of the most recent segment (i.e. between the most recent structural break and immediately prior to the forecasting period) are selected as the neural network validation set.

If the structural breaks period selection algorithm is not carried out (i.e. only the hourly clustering as displayed in Figure 4.6), an alternative comparative is performed, which selects the top 2.5% similar hours, after the hourly clustering selection method is performed, as validation dataset. This is approximately 3.2 days on average throughout the entire year 2017, while the previously explained method selects 3 days on average.

4.2.3 Artificial neural network model

As explained on Section 4.2.1, four outputs of the fundamental model are taken as additional inputs to the neural network model: market clearing prices as well as the CCGT, coal and hydro production levels. These are combined with common predictors to form the set of input variables for the neural network model. This set of common predictors consists of the following factors:

- Expected values of demand, wind and solar generation.
- Expected mean temperature in the Iberian Peninsula.
- Two dummy variables corresponding to working days or a Sunday/holiday, thus Saturdays would correspond to both dummy variables being false.

- Actual electricity market prices with the following lags: one day, two days, one week and two weeks.
- Commodity related futures prices: month-ahead API2 coal, month-ahead NBP natural gas and month-ahead European CO_2 emission allowances.
- Fundamental model output variables: market clearing prices; and coal, CCGT and hydro production levels.
- Day-ahead Iberian electricity market futures.

Essentially, these are the same variables chosen in Chapter 3's proposed methodology except for the day-ahead Iberian electricity market futures, which provide more information as to the average level of spot prices throughout the next day, as stated on the recent work of (Steinert & Ziel, 2019). Once the previously mentioned period filtering methods have been carried out, the remaining data are used as training inputs to a neural network forecasting method. See Section 2.3.2.2 for detailed descriptions regarding the neural network forecasting model features, configuration and training methods. Moreover, Appendix B presents several experiments targeted at the avoidance of overfitting in the neural network forecasts applied to the final model of this chapter.

4.3 Evaluation criteria and benchmarking models

In order to fully demonstrate the forecasting performance of the proposed forecasting models of this chapter, distinct evaluation criteria and benchmarking models have been utilised, as done in previous chapters. The forecasting accuracy of the involved models has been measured by means of the MAPE, MAE and RMSE error metrics with hourly precision. Moreover, in order to assess performance comparisons in a statistically significant manner, the Diebold-Mariano (DM) test has been carried out with a 5% significance level with the absolute error difference as the loss differential series. This test has been carried out considering both out- and underperformance.

In order to quantify the contributions of the proposed calibration period selection methods for the neural network forecasting technique, the forecasting methodology is split into several stages, where each stage involves none, one or both calibration window selection procedures:

- Stage 0: this starting point is a base hybrid fundamental-statistical model which coincides with the Proposed Model 2 that was presented in Chapter 3. This model uses 120 days of calibration data, although a filtering procedure[46] on winter 2017 reduced this data length by approximately 70% on average.

- Stage 1a: 13 months of calibration data are used and these are filtered via the structural breaks technique.
- Stage 1b: same as stage 1a, but the hourly clustering method is used instead.
- Stage 2: both the structural breaks and the hourly clustering approaches are combined.

These models will be referred to as PMS_i, which stands for Proposed Model at its stage *i*. As done in several works in the literature, such as (Bento, Pombo, et al., 2018), the performance of these models has been analysed for every season of the year and compared with that of nine other electricity price forecasting models, some of which correspond to well-established methodologies in the literature and others pertain to specific methodologies presented throughout Chapter 3:

- The first chosen benchmark model (Benchmark 1 or BM_1) consists of one of the most accurate models in Chapter 3, which involves a simple average between the forecasts of the fundamental-statistical hybrid model (PMS_0) and a pure neural network model.
- Benchmark two (BM_2) only involves this pure neural network model that utilises the same input variables as BM_1 and PMS_0 (except those pertaining to the fundamental model) and the same calibration window. This 120-day window includes four months within the 13-month window established in this work, more specifically, the 13th, 12th, 2nd and 1st month prior to the forecasting day, as established in Chapter 3.
- The third benchmark model (BM_3) has been utilised in previous chapters, which is the linear regression model presented in (Misiorek, Trueck, et al., 2006). Given that in Chapters 2 and 3, a 3-month calibration dataset window was chosen and placed immediately prior to the forecasting period, this benchmark has also been utilised with 3 months of calibration data in this chapter.
- The next benchmark (BM_4) is the extension of BM_3 as per the proposed methodology of (Marcjasz, Serafin, et al., 2018), which involves a weighted average of this regression model across the following calibration windows (in terms of days prior to the forecasting period): 56, 84, 112, 714, 721 and 728 days. The weights of these six forecasts are computed by means of an inverse MAE weighting procedure, which is similar to that

[46] This involves computing the validation set MSE of the NN for several calibration window lengths. The window length that yielded the lowest MSE was chosen for the forecast.

of (Bates & Granger, 1969), when testing the linear regression models at the day prior to the forecasting period. Furthermore, the linear regression model of BM_3 has also been extended with one or both of the calibration period selection methods that have been proposed in this chapter:

  - BM_5: BM_3 with PMS_{1a}'s calibration dataset window.
  - BM_6: BM_3 with PMS_{1b}'s calibration dataset window.
  - BM_7: BM_3 with PMS_2's calibration dataset window.

- Benchmarks eight and nine (BM_8 and BM_9) coincide with the SARIMAX and the simple naïve approach that have been used in both Chapter 2 and Chapter 3, which are thoroughly explained in Section 2.4.

4.4 Case study, results and discussion

As done in previous chapters, the case study for the forecasting methodologies of this chapter are tested on the recent case study of the Iberian electricity market of the entire year 2017, whose distinct features and circumstances have previously been explained in Sections 2.5 and 3.4. First of all and as per Figure 4.1, the cost production model is run so as to obtain the necessary information to carry out the remaining components of the proposed methodology. Said information translates to the previously mentioned fundamental model output variables: market clearing prices as well as CCGT, hydro and coal unit generation output levels. Furthermore, given that the aim of this work is to provide forecasts for the entire year 2017 and that the neural network model's initial training dataset is 13 months, all input variables must be made available from December 2015 up to December 2017. Once these 13 months are filtered according to the methodologies presented in the previous section, the neural network model is run for every day of the year 2017 and the final forecast is obtained. Therefore, the chosen forecasting horizon is one day.

The proposed model in all of its stages as well as the nine benchmark models have been put to the test for every day of the entire year 2017 and their error measures across the four seasons of 2017 are shown in Table 4.1, Table 4.2 and Table 4.3. Furthermore, the average calibration dataset windows for each of the models involving a neural network forecasting technique are displayed in Table 4.4.

As seen in previous chapters, the forecasting methodologies that do not involve neural networks are underperforming. This is mainly due to the modelling capabilities of neural networks, which are highly versatile when compared to the classic statistical models. Although they are able to outperform linear regression

Model	Winter	Spring	Summer	Autumn	Average
PMS_0 – Base model (PM_2 in Chapter 3)	11.68	8.106	**4.450**	6.812	7.744
PMS_{1a} – Structural breaks period selection method	11.02	7.706	4.501	**6.237**	7.348
PMS_{1b} – Hourly clustering period selection method	11.76	8.048	4.711	6.802	7.812
PMS_2 – Union of both period selection methods	**10.05**	**7.303**	4.467	6.284	**7.012**
BM_1 – Simple avg. of PMS_0 + BM_2 ($CMSA_2$ in Ch. 3)	11.30	7.902	4.477	6.756	7.591
BM_2 – Pure NN model (BM_2 in Chapter 3)	11.12	7.804	4.605	6.834	7.575
BM_3 – ARX of (Misiorek, Trueck, et al., 2006)	16.79	13.58	7.153	10.51	11.99
BM_4 – W. ARX of (Marcjasz, Serafin, et al., 2018)	16.27	13.21	7.015	10.14	11.64
BM_5 – ARX of BM_3 with PMS_{1a}'s period selection	17.34	13.12	7.041	10.39	11.95
BM_6 – ARX of BM_3 with PMS_{1b}'s period selection	15.84	12.04	6.801	9.581	11.04
BM_7 – ARX of BM_3 with PMS_2's period selection	16.06	12.73	7.023	9.984	11.43
BM_8 – SARIMAX model	15.06	9.293	5.097	7.654	9.248
BM_9 – Simple Naïve approach	25.93	17.55	9.343	12.82	16.37

Table 4.1: Forecasting error in terms of MAPE (%)

Model	Winter	Spring	Summer	Autumn	Average
PMS_0 – Base model (PM_2 in Chapter 3)	4.756	2.882	2.070	3.453	3.282
PMS_{1a} – Structural breaks period selection method	4.266	2.487	2.063	3.150	2.984
PMS_{1b} – Hourly clustering period selection method	4.761	2.626	2.175	3.476	3.251
PMS_2 – Union of both period selection methods	**4.133**	**2.433**	**2.045**	**3.139**	**2.930**
BM_1 – Simple avg. of PMS_0 + BM_2 ($CMSA_2$ in Ch. 3)	4.610	2.832	2.079	3.409	3.224
BM_2 – Pure NN model (BM_2 in Chapter 3)	4.562	2.826	2.136	3.440	3.233
BM_3 – ARX of (Misiorek, Trueck, et al., 2006)	6.838	4.765	3.262	5.066	4.972
BM_4 – W. ARX of (Marcjasz, Serafin, et al., 2018)	6.390	4.500	3.211	4.880	4.736
BM_5 – ARX of BM_3 with PMS_{1a}'s period selection	6.773	4.506	3.238	5.007	4.870
BM_6 – ARX of BM_3 with PMS_{1b}'s period selection	6.134	4.224	3.093	4.561	4.494
BM_7 – ARX of BM_3 with PMS_2's period selection	6.276	4.307	3.221	4.798	4.641
BM_8 – SARIMAX model	8.113	4.150	2.473	4.454	4.780
BM_9 – Simple Naïve approach	10.53	6.225	4.266	6.387	6.828

Table 4.2: Forecasting error in terms of MAE (€/MWh)

and time-series approaches, neural networks require much more computational power and runtime. The weighted average of linear regression forecasts (BM_4) has proven to be superior to the individual one (BM_3), albeit not as accurate as all neural network related models. The proposed extensions of the individual linear regression model (BM_5, BM_6 and BM_7) indicate that the hourly clustering period selection method alone is most beneficial. This suggests that linear relationships are better captured if explanatory variables are taken into account in the moment of the forecast while disregarding past price behaviours, which most likely feature non-linear patterns such as agent strategic behaviours. Therefore, BM_6 is the most accurate among the considered ARX models, outperforming the base ARX model represented by BM_3 and the forecast window averaging method of BM_4.

Compared with the base model of PMS_0, the implementation of the structural breaks technique increased (PMS_{1a}) the neural network training set by well beyond the predefined number of 120 days that was established in Chapter 3. The reason behind the reduced dataset during 2017's winter is due to its high instability, and it was observed in Chapter 3 that a reduction of the 120-day

Model	Winter	Spring	Summer	Autumn	Average
PMS_0 – Base model (PM_2 in Chapter 3)	5.479	3.407	2.517	4.129	3.874
PMS_{1a} – Structural breaks period selection method	4.917	2.960	2.509	3.820	3.543
PMS_{1b} – Hourly clustering period selection method	5.461	3.110	2.667	4.143	3.835
PMS_2 – Union of both period selection methods	**4.785**	**2.892**	**2.504**	**3.818**	**3.492**
BM_1 – Simple avg. of PMS_0 + BM_2 ($CMSA_2$ in Ch. 3)	5.334	3.349	2.525	4.069	3.810
BM_2 – Pure NN model (BM_2 in Chapter 3)	5.308	3.342	2.588	4.089	3.823
BM_3 – ARX of (Misiorek, Trueck, et al., 2006)	7.809	5.552	3.885	6.055	5.814
BM_4 – W. ARX of (Marcjasz, Serafin, et al., 2018)	7.314	5.268	3.857	5.874	5.568
BM_5 – ARX of BM_3 with PMS_{1a}'s period selection	7.683	5.174	3.863	6.012	5.671
BM_6 – ARX of BM_3 with PMS_{1b}'s period selection	7.006	4.952	3.723	5.567	5.302
BM_7 – ARX of BM_3 with PMS_2's period selection	7.162	5.032	3.863	5.807	5.456
BM_8 – SARIMAX model	10.84	5.585	4.531	4.959	6.460
BM_9 – Simple Naïve approach	11.48	7.092	5.030	7.567	7.773

Table 4.3: Forecasting error in terms of RMSE (€/MWh)

dataset provided useful results. This agrees with the rationale that consists of increasing adaptability on unstable periods by reducing the calibration window in order to remove structural breaks from the input dataset. However, in this work, an average dataset of 152.9 days yields lower forecasting errors. Furthermore, PMS_1 discards most of the previous winter, which is considerably different from 2017's winter as depicted in Figure 4.3. This also seems to be the case for spring, as 2016's spring yielded approximately twice as much hydro generation as 2017's spring. In general, the structural breaks algorithm provides a generally lower error throughout 2017. However, summer 2017 seems to be the exception, where prices are relatively stable and thus it lacks room for improvement, as proven by the generally low errors yielded by most models.

The implementation of the hourly clustering period selection method alone (PMS_{1b}) with K-means seems to discard important information pertaining to the recent periods or moments that coincide with the market regime of the forecasting periods, which translates to the accuracy loss that can be observed on the previous tables. Furthermore, it seems to be somewhat generally restrictive, to a lesser or greater extent, throughout the entire year 2017, given the overall hollowness that can be observed in Figure 4.6's shape, reducing the initial 13-month dataset by approximately 35% throughout the 365 days of the case study. Although linear relationships may be better captured with this filtered dataset, as previously observed from BM_6's performance, neural networks involve a significantly different algorithm that also considers non-linear trends and patterns.

Model	Winter	Spring	Summer	Autumn	Average
PMS_0 – Base model (PM_2 in Chapter 3)	36.67	120	120	120	99.17
PMS_{1a} – Structural breaks period selection method	152.9	237.0	324.7	300.5	254.2
PMS_{1b} – Hourly clustering period selection method	241.8	253.6	234.2	271.7	250.3
PMS_2 – Union of both period selections methods	288.8	324.5	344.7	348.1	326.7

Table 4.4: Average calibration window length of the neural network models (days)

Furthermore, given that the expected values of the variables that are used for the hourly clustering method may be inaccurate themselves, their misspecifications may be transferred to the neural network, thus hindering its forecasting performance. Although said variables, namely the expected market clearing prices from the fundamental model as well as the expected thermal gap and expected temperature, respond well to market fundamentals and are relevant to the electricity price formation, it seems that the ascertainment of the market regime by analysing past price behaviours is more valuable for neural network forecasting methodologies with regards to calibration period selection.

The union of PMS_{1a}'s calibration dataset selection method with that of PMS_{1b}'s (PMS_2) further reduces the overall forecasting error when compared to that of PMS_{1a} alone. This is most notable during winter, whose average calibration dataset is greatly increased to 288.8 days. As for the other seasons, a calibration dataset of approximately one year proves to be beneficial for electricity price forecasting with neural network models even with hourly arrangement and seems to increase their generalisation capabilities. The results on summer suggest yet again that it is characterised by relatively stable conditions, and thus PMS_2 does not fully outperform PMS_0.

Although PMS_2 yields a lower error overall, the statistical significance of these error measures must be verified in order to confirm its superiority against the highest ranked models according to Table 4.1, Table 4.2 and Table 4.3. Therefore, a DM test was carried out for PMS_2 paired with PMS_0, PMS_{1a}, PMS_{1b}, BM_1 and BM_2. The DM test statistic is evaluated with a 5% significance level, which involves a critical value of 1.96. Therefore:

- DM test statistic < -1.96 implies significant outperformance
- DM test statistic > 1.96 implies significant underperformance
- Otherwise no significant out- or underperformance.

The results of the DM test statistic are shown in Table 4.5. The three values in bold indicate the three occasions that PMS_2 was unable to significantly outperform. The comparison with PMS_0 suggests that the increase in calibration data window lengths does not significantly contribute to summer forecasts, albeit not detrimental to the accuracy. This may also imply that a robust calibration

Model comparison	Winter	Spring	Summer	Autumn	Average
PMS_2 vs. PMS_0	-8.834	-12.31	**-0.975**	-6.787	-14.75
PMS_2 vs. PMS_{1a}	-2.903	-3.199	**-1.833**	**-0.436**	-3.917
PMS_2 vs. PMS_{1b}	-12.15	-7.429	-5.254	-8.535	-17.24
PMS_2 vs. BM_1	-6.528	-8.042	-2.883	-5.726	-11.29
PMS_2 vs. BM_2	-6.316	-11.36	-3.262	-6.877	-13.18

Table 4.5: DM statistic values of PMS_2 against highest ranked models

period selection is not highly crucial on such a stable market regime. Therefore, the same conclusion can be drawn from the summer comparison with PMS_1. Furthermore, the DM statistic value on autumn when compared with PMS_1 may indicate that the information provided by the hourly clustering method is not significantly different than that provided by the structural breaks technique. However, these values indicate that PMS_2 is significantly outperforming all other models throughout the entire year 2017, especially during winter, which is the most unstable season of the year.

4.5 Conclusions

This chapter presents a novel short-term hybrid electricity price forecasting methodology which is comprised of three main elements: a cost-production optimisation model, a sophisticated period filtering approach and a neural network model. These three elements were utilised sequentially with the calibration selection procedure as the main focus of this work. Given a forecasting day, the structural patterns in actual prices corresponding to the 13 months prior to said day are analysed and those deemed unimportant were discarded from the neural network training data. Moreover, a K-means clustering method was also applied so as to relate moments in said 13 months to the forecasting day in terms of the estimated fundamental market clearing prices, expected thermal gap and expected mean temperatures in the Iberian Peninsula.

The results and analyses presented in this chapter translate to the following findings and observations. The combination of structural break analysis and hourly clustering provides a dynamic calibration period selection that removes the need for predefining said period due to the involved guidance criteria based on past price behaviours and useful market regime explanatory variables. In addition, this sophisticated training window selection for the neural network model yields appealing results in every market circumstance present in the relatively recent and challenging case study of the Iberian electricity market of 2017. The period selection technique is more selective in volatile market conditions, such as early 2017, albeit providing a considerably longer training window length than other works that claim that employing much shorter calibration windows is most suitable in these situations, such as the work presented in Chapter 3. This also may be due to the segmentation of the calibration dataset, as most electricity price forecasting models in the literature do not consider this kind of training windows. In addition, the proposed methodology proves most useful on volatile periods, although accuracy is marginally increased in stable market regimes, such as summer 2017.

Taking everything into consideration, this work's proposed short-term fundamental-statistical electricity price forecasting model, which features a unique hybridisation approach, has yielded appropriate results when applied to a real-size electricity system with complex price dynamics, such as the Iberian power exchange of 2017. Furthermore, the performance of this proposal has been proven superior to that of well-recognised benchmark models. However, there seems to be room for improvement regarding the utilised structural breaks period selection algorithm, as it is highly challenging to ascertain a convenient compromise between accuracy and computational burden. It would most certainly be beneficial if sudden price changes (i.e. price spikes) are adequately considered in a computationally feasible manner.

In addition, the fact that the hourly clustering period selection method only proves beneficial when combined with the structural breaks method requires further investigation. It has been observed that linear regression models benefit most from the hourly clustering method alone, which has the opposite effect on the considered fundamental-statistical approaches. It would seem that recent periods should not be segmented as much by the hourly clustering method. Perhaps this is due to the possibility that, throughout the training process of the neural network models, these may need as much information as possible regarding the most recent moments up to the latest structural break in order to enhance their predictive accuracy in the forecasting period. Then, the least recent information may be more restrictively selected and added to the neural network dataset, as performed in this chapter.

Appendix B: Neural network model overtraining and sensitivity analyses

It is of common knowledge that neural network forecasting models are prone to overfitting if no attention is paid to their calibration datasets. A suitable balance between model complexity and generalisation capability is vital for a successful forecast, albeit challenging to perfect it and it also depends on the context. Throughout the entirety of this thesis, two dimensions of the neural network training datasets were modified: input variables and periods. This chapter has presented a suitable procedure regarding the latter. However, a suitable variable selection method is yet to be determined.

Given that neural networks handle non-linear relationships, some variable selection methods that are usually applied in linear contexts may not prove useful, such as LASSO and Ridge regression, recently utilised for electricity price forecasting with linear regression models in (Uniejewski, Nowotarski, et al., 2016). Outside the electricity price forecasting context, mutual information and partial mutual information (MI and PMI respectively) are utilised for feature or variable selection in order to consider both the relevance with respect to the forecasted variable and the redundancy among other variables in the input dataset (Frenzel & Pompe, 2007; Zheng, Yang, et al., 2017). This method may be applied in order to check the usefulness of the 18 input variables that were used in this chapter's proposed neural network model and the open source code available at http://www.cs.tut.fi/~timhome/tim/tim.htm has been implemented in order to calculate the values of MI and PMI. However, given the discrete and discontinuous nature of the dummy variables, these were discarded from the MI and PMI experiments.

Table 4.6 and Table 4.7 show the average rankings of the variables (except the dummy ones) as per their MI and PMI values throughout the year 2017. Although it seems that the demand is less relevant than others, it would seem that it is the least redundant variable amongst all of them. On the other hand, the MI values suggest that the one-day lagged electricity price and the estimated market clearing prices are the most relevant, which is expected given that they represent electricity prices. The implementation of a variable selection method based on both MI and PMI would require an iterative procedure with a stopping criterion. However, the final forecasting model of this chapter did not yield better results when paired with this variable selection method, apart from increasing model runtime.

Input variable	Winter	Spring	Summer	Autumn	Average
Expected demand	12	13	11	11	12
Expected wind generation	15	15	16	16	16
Expected solar generation	16	14	15	15	15
Forecasted market clearing prices (fundamental)	2	6	2	2	2
One-day lagged electricity prices	1	1	1	1	1
Two-day lagged electricity prices	9	9	9	9	9
One-week lagged electricity prices	6	5	3	4	3
Two-week lagged electricity prices	7	7	6	7	7
Expected Iberian Peninsula temperature	13	12	14	14	13
Natural gas futures prices	4	4	4	5	4
Coal futures prices	5	2	7	6	5
CO_2 emission allowances futures prices	11	11	12	12	11
Forecasted CCGT production (fundamental)	10	10	10	10	10
Forecasted coal production (fundamental)	8	8	5	3	8
Forecasted hydro production (fundamental)	14	16	13	13	14
Electricity market futures prices	3	3	8	8	6

Table 4.6: Sixteen variables ranked from most to least relevant according to MI values

Input variable	Winter	Spring	Summer	Autumn	Average
Expected demand	1	1	1	1	1
Expected wind generation	2	2	16	16	5
Expected solar generation	16	15	2	15	14
Forecasted market clearing prices (fundamental)	5	5	9	5	4
One-day lagged electricity prices	6	6	8	8	6
Two-day lagged electricity prices	11	12	6	7	9
One-week lagged electricity prices	9	9	7	6	7
Two-week lagged electricity prices	10	4	5	4	3
Expected Iberian Peninsula temperature	14	14	15	14	16
Natural gas futures prices	7	10	10	9	10
Coal futures prices	8	7	12	11	12
CO_2 emission allowances futures prices	13	13	14	13	15
Forecasted CCGT production (fundamental)	3	3	3	12	2
Forecasted coal production (fundamental)	12	11	11	2	11
Forecasted hydro production (fundamental)	15	16	4	3	13
Electricity market futures prices	4	8	13	10	8

Table 4.7: Sixteen variables ranked from least to most redundant according to PMI values

Another way of calculating a direct sensitivity analysis is by extracting the weights of the trained neural networks that were used in the proposed methodologies of this chapter. Given a certain number of neurons in the hidden layer, *n*, the neural network presents *n* weights per each input variable. These weights are averaged across the number of neurons and replications every day of 2017, and the resulting weights are used in order to rank the variables from most to least relevant on Table 4.8. Although the expected demand always presents the lowest weight, the NN considers similar information that can be found on other variables such as production outputs from the fundamental model or the expected wind generation. This does not necessarily mean that the expected demand is significantly redundant, which may go against several traditional approaches that

Input variable	Winter	Spring	Summer	Autumn	Average
Working-day dummy variable	6	7	5	2	4
Sunday/holiday dummy variable	3	3	3	3	2
Expected demand	18	18	18	18	18
Expected wind generation	8	9	7	6	9
Expected solar generation	5	6	8	9	8
Forecasted market clearing prices (fundamental)	15	15	14	14	14
One-day lagged electricity prices	10	5	4	8	6
Two-day lagged electricity prices	12	13	13	13	13
One-week lagged electricity prices	13	12	12	12	12
Two-week lagged electricity prices	4	4	11	4	5
Expected Iberian Peninsula temperature	11	8	9	7	10
Natural gas futures prices	2	2	2	10	3
Coal futures prices	9	11	10	11	11
CO_2 emission allowances futures prices	7	10	6	5	7
Forecasted CCGT production (fundamental)	17	17	17	17	17
Forecasted coal production (fundamental)	1	1	1	1	1
Forecasted hydro production (fundamental)	14	14	15	15	15
Electricity market futures prices	16	16	16	16	16

Table 4.8: Eighteen variables ranked from most to least relevant according to NN weights

use the expected demand as the first explanatory variable in electricity price forecasting models.

Moreover, given the results of Table 4.8, one may reach the conclusion that the error may be reduced if the weights on certain variables are reduced. This can be directly applied to the objective function of the neural network training algorithm with a regularisation technique that penalises high weight values. This is done by incorporating an additional term apart from the sum of squared errors on the validation set as seen on the following objective function:

$$\min\left[(1-\gamma)\cdot\frac{1}{N}\cdot\sum_{i=1}^{N}(e_i)^2+\gamma\cdot\frac{1}{N}\cdot\sum_{i=1}^{N}(w_i)^2\right] \tag{4.1}$$

Where N represents the number of hours in the validation set and γ is the regularisation ratio that must range from zero to one. Originally, γ is set to zero, and it would be interesting to see if modifying its value could provide a higher forecasting accuracy. However, it is challenging to ascertain the optimum value for this parameter. Therefore, different values are tested on the final model of this chapter and its results are shown on Table 4.9. Unfortunately, no value of γ has provided a higher frequency, which may indicate that the weights displayed on

Value of regularisation ratio γ	Winter	Spring	Summer	Autumn	Average
0	4.133	2.433	2.045	3.139	2.930
0.25	4.170	2.465	2.162	3.156	2.981
0.5	4.236	2.499	2.178	3.141	3.007
0.75	4.435	2.622	2.177	3.152	3.089

Table 4.9: Forecasting accuracy as a function of γ in terms of MAE (€/MWh)

Table 4.8 do not seem to be linked to any overfitting or overtraining issue.

In conclusion, all of the methodologies presented in this appendix, which are targeted at the reduction of overfitting occurrences and the improvement of neural network generalisation, do not seem to reduce forecasting error in this application. Nevertheless, the number of neurons sweep, the early stopping methodology and the replication of the neural network training algorithm (refer to Section 2.3.2.2 for specific details) are suitable procedures that have been utilised throughout this thesis that seem to be sufficient for avoiding overtraining and improving neural network forecasting performance.

Chapter 5
Conclusions, Contributions and Future Research

This final chapter of the thesis document contains the main conclusions and findings that derive from the research that has been presented throughout the document. This chapter is composed of three parts. Firstly, the summary of the studies and analyses of the thesis is presented. Secondly, the contributions that have been achieved while elaborating all the involved forecasting methodologies of the thesis are explained. Finally, several lines for future research are identified and proposed that are derived from the results and developments that have been carried out in this thesis.

5.1 Summary and conclusions

Due to the constant evolutions and alterations that the electricity markets of today are facing, stakeholders are strongly encouraged to adapt their strategies and seek assistance when operating in the market. Traders and practitioners in power exchanges are experiencing a very high degree of uncertainty and thus resort to electricity price forecasts for several purposes, such as unit scheduling, risk management and fuel trading. However, electricity cannot be treated as a regular commodity, given its non-storable signature feature, which causes its price to reach considerable levels of volatility. In addition, the ongoing regulatory changes, strong weather variations and increasing deployment of intermittent generation are heightening the complexity of the system.

The possibility of combining fundamental and statistical approaches has been thoroughly discussed throughout this thesis, which is a suitable way to consider all sources of disruption that are present in electricity markets. However, merging these models has been mostly carried out in medium-term contexts, where they have proved beneficial due to their capabilities of capturing several aspects into account, such as market dynamics and structural/regulatory market changes (fundamentals), as well as strategic/speculative behaviour and linear/non-linear modelling capabilities (statistics). However, the literature regarding this kind of hybrid models in short-term contexts is relatively scarce and thus it encourages motivation to determine if the same advantages can be attained for the short term. Therefore, a unique methodology to apply fundamental-statistical hybrid models in the short term has been proposed and developed in this thesis.

First of all, in Chapter 2, a novel methodology that is based on a hybridised fundamental model and a statistical model has been presented. On the one hand, the fundamental model consists of a cost-production optimisation model that simulates the market clearing considering all physical and regulatory elements of the power system. On the other hand, a neural network forecasting model is employed so as to correct any bias of the fundamental model's estimated prices alongside other input variables, such as wind power and demand forecasts. Furthermore, the input data on both components of the proposed hybrid model were rearranged and modified in order to decrease computational burden and therefore increase efficiency, as well as reduce runtime and overfitting occurrences on the neural network model. For instance, in the fundamental model, the generation technologies that shared similar cost functions were aggregated into larger ones in order to reduce resolution time.

In the light of the results and analyses of Chapter 2, the proposed hybrid short-term electricity market price forecasting model has shown adequate performance in seven specific and representative case studies of the Iberian electricity market of late 2016, and also in the more general case study of the entire year 2017 of the same power exchange. Not only traditional methods have been outperformed, such a SARIMAX model and a linear regression model, but also the effects yielded by the incorporation of market clearing prices to the neural network dataset has proven to be beneficial. It has been thus concluded in Chapter 2 that the hybrid model's accuracy is generally increased by the estimated price from the fundamental model. The main concluding remark is related to the combination of the longer-term price level yielded by the fundamental model and the intraday pattern provided by the statistical model on the forecasts of the proposed hybrid methodology. This advantage is more notable for longer forecasting horizons, which further indicates that the equilibrium price level of the fundamental model enhances predictive accuracy. In addition, a preliminary probabilistic analysis was conducted whose results are also in favour of the hybrid model, although there seems to be room for improvement regarding the exceedance rates of its percentile forecasts.

Moreover, it was repeatedly observed that the forecasting error was increased during hours of extremely low or high prices. The effect of dragging the neural network forecast to the daily average level by the market clearing prices is the main cause of this issue. This suggested that a combination procedure that discards the effect of the fundamental on hours of extremely high/low prices should be used in order to further reduce the error. Another modification of the hybridisation approach that was proposed is targeted at the output variables that are obtained from the cost-production optimisation model, such as head-dependent hydro, coal or CCGT unit production levels. Furthermore, the variable selection procedure that was carried out iteratively calls for a more suitable sensitivity analysis or a suitable variable selection procedure for neural network models, and making it more efficient could be highly challenging due to the inherent complexity in neural network training algorithms.

Several of these issues and limitations of this forecasting model encouraged the work proposed and developed in Chapter 3, which addresses forecast combinations with improved versions of the fundamental and the statistical models. The cost-production optimisation model was extended by considering coal and CCGT power units individually in order to increase accuracy albeit doubling resolution time. Furthermore, the bids of these units were estimated based on past bids and relevant commodity prices. As a result, the errors of the estimated market clearing prices throughout the year 2017 were reduced by approximately a third, which makes it a worthwhile tradeoff. The statistical component was paired with an improved calibration period selection and a similar days method. More specifically, the calibration period was shortened for unstable periods, such as winter 2019, and the similar days methods selects as the neural network validation procedure the days in the past that are similar to the forecast period in terms of expected demand, expected wind power and expected temperature. Moreover, multiple hybridisation methods were employed throughout Chapter 3 in the form of additional outputs from the fundamental model passed on to the statistical model as well as several forecast combination procedures.

The concluding remarks that have arisen from the proposed hybrid methodologies of Chapter 3 are mainly related to the behaviours exhibited in the distinct market circumstances that are present throughout the Iberian electricity market of 2017. For instance, during periods of relative stability, the electricity market price behaviours seem more driven by market fundamentals and thus using additional variables from the fundamental model in the statistical model proved beneficial. On highly volatile periods, the neural network shortened calibration window (approximately 70% of its normal length) further increased its adaptability and accuracy. Moreover, generally throughout the entire case study, the simple average between the hybrid and the pure neural network forecasts provided the lowest error and thus outperformed the other forecast combination methods. This calls for a more sophisticated and optimal forecast combination method, given that the analyses of Chapter 2 suggested a less balanced combination procedure. Furthermore, a more generalised calibration selection procedure should be utilised instead of a specific one, which was used only throughout the winter season.

These remarks and issues encouraged the studies that have been conducted throughout Chapter 4, which contains a proposal of a robust calibration period selection driven by explanatory information prior to the elaboration of the forecast, which has been directly applied to one of the models proposed in Chapter 3. First of all, the initial dataset was increased from four months to 13 months. Two techniques were implemented in this methodology, which are responsible for identifying distinct trends and patterns that select periods in the past 13 months that are sufficiently similar to the market regime that corresponds to the forecasting period. The first procedure identifies structural breaks in this 13-month dataset and splits it into several segments. The most recent segment is assumed to pertain to the same market regime as that of the forecasting period, and thus the other segments are discarded if they are dissimilar to the recent segment. The second method is a clustering method that utilises variables that are available at the moment of the forecast, which involve the expected thermal gap, the expected temperature, and the estimated market clearing prices from the fundamental model. Consequently, the hours in the 13-month dataset that do not belong to the same clusters as the hours in the forecasting period are discarded. This hourly clustering method constitutes a unique hybridisation procedure that aims to heighten the synergies between the fundamental and the statistical components of the hybrid methodology. The union of both period selection methods yielded a significantly lower error when compared to the proposed forecasting models of Chapter 3, especially on volatile periods such as early 2017. Although applying the hourly clustering method alone did not prove useful, the structural breaks period selection reduced error by approximately 10% throughout the entire 2017. Furthermore, the union with the clustering method increased this error reduction to 11.5%. An hourly MAPE value of 7% throughout the entire year of 2017 is quite appealing for traders and practitioners in the Iberian power system.

In general, the employed calibration periods were of at least 10 months, which contradicts the conclusions of Chapter 3 and of other works related to the advantage of shortened calibration periods in volatile periods. For instance, (Pesaran & Timmermann, 2007) recommend discarding every period prior to the most recent structural break, while the results of Chapter 4 show that using segmented and much larger datasets yields superior predictive performances when compared to those of Chapter 3. Moreover, the period selection methodology of Chapter 4 selects 12 of the 13 months in stable periods, where there does not seem to be any room for improvement given that most models perform well in such circumstances. Therefore, the proposed calibration period selection is also more selective on unstable periods. However, this methodology may be extended by means of price spike detection and forecasting, as the proposed model does not fare adequately in the event of abrupt price changes, which are much less gradual than structural breaks.

Taking everything into consideration and in the light of the results of the studies and analyses of this thesis, it can be concluded that short-term electricity price forecasting, which is usually carried out by means of statistical approaches, can be enhanced by pairing these with fundamental approaches in a hybrid complementary framework. However, in short-term applications, an efficient and optimal hybridisation is highly challenging to obtain. This provides the forecast with a robust adaptability and thus can yield accurate forecasts in several market regimes and circumstances. Fundamental models internalise physical and regulatory alterations of the power system in their forecasts, while statistical approaches are able to detect key trends and repetitive patterns such as the strategic and speculative behaviours of traders. However, it is worth noting that day-ahead forecasts require a daily data acquisition as well as an ongoing monitoring of the electricity system. Nevertheless, the proposed forecasting methodologies and analyses of this thesis can prove useful for electricity market participants, ranging from generating companies and retailers to regulators and authorities.

5.2 Original contributions

Throughout the development and elaboration of this thesis, multiple original contributions have been provided, which are listed below:

1. The first contribution is related to the rationales, ideas and reasoning processes that have guided the developments and breakthroughs of the thesis. The research gaps in the short-term electricity price forecast context have been pointed out in the literature reviews that have been presented in the previous chapters. In the light of the current circumstances and challenges pertaining to electricity markets, the encouragement behind the proposed hybrid methodologies of the thesis has been justified from both academic and practical perspectives.

2. The successful application of a fundamental market equilibrium model to the short term that provides price forecasts driven by physical and regulatory elements of the power system with hourly precision. Several enhancements have been proposed with the intention to reduce computational burden and increase forecasting accuracy, such as the simplification of the production structure and the estimation of the generation programming units' bids. This model has proven useful in estimating the long-run equilibrium level of electricity prices and adequately responds to several market events.

3. Another important contribution of this thesis lies in the proposal of multiple combination and hybridisation schemes between the fundamental model and statistical techniques. These are targeted at the ascertainment of an optimal synergy between fundamental and statistical approaches in order to consider as much relevant factors and forces that drive electricity prices as possible. By doing so, robust forecasts are provided, which

outperform individual fundamental and statistical methods. These hybridisation schemes have been developed in the form of passing variables from the fundamental to the statistical forecasting approaches, forecast combination methods and period selection designs.

4. The idea of calibration period selection in the context of electricity price forecasting has been rarely addressed. Given that it is more relevant in other fields, resorting to calibration period selection is not essentially a novelty. However, advanced statistical methods, such as neural networks, can greatly benefit from carefully selected input data, as demonstrated on this thesis. The similar days methodology of this thesis that has been paired with the neural network models provides a suitable validation set for their training algorithms, which allows the resulting neurons to better adapt to the forecasting period conditions when carrying out the actual forecast.

5. Furthermore, in this thesis, the novel idea of utilising calibration period selection methods driven by robust criteria related to market regimes is presented. It is based on past price behaviours and explanatory variables available at the moment of the forecast. By starting off with a considerably oversized calibration window, this robust procedure removes the need of specific predefinitions such as discarding summer data when forecasting winter prices, as well as selecting the most relevant information for the statistical forecasting models in order to enhance predictive performance in the presence of any kind of market events and under any level of volatility.

6. All the forecasting models and hybrid methodologies that have been proposed throughout this thesis have been put to the test and thoroughly verified in a real-size electricity system featured by complex price dynamics: the recent case of the Iberian electricity market of late 2016 and the entire year 2017. Furthermore, a detailed comparison with some of the most well-established forecasting methodologies in the short-term electricity price forecasting literature is provided, which reveals interesting results that can also be recognised as a contribution.

5.3 Future research

As a result of the methodologies that have been developed throughout this thesis, several future lines of research have been identified and proposed that are relevant in the context of electricity price forecasting. These are described below:

1. A notable trend in the electricity market price forecasting literature is that short-term applications are mainly being carried out from a point-forecasting perspective, and thus there is a considerable amount of work and research to carry out in the field of probabilistic forecasting, which, according to (Ziel & Steinert, 2018), is becoming increasingly popular since 2016, although mostly in medium- and long-term contexts. While point forecasting focuses on the mean and sometimes on the variance, probabilistic forecasts cover the full probability density of electricity market prices and thus provide more information regarding risk levels. Although the importance and usefulness of probabilistic analyses has been justified by the authors of (Bello, Bunn, et al., 2016) in the context of medium-term hybrid electricity market price forecasting models, little work has been carried out in the short term, such as the percentile forecasts presented at the end of Chapter 2, whose results suggested that the hybrid methodology yielded percentile forecasts that are slightly closer to the ideal values. However, the occurrence of extremely low or high prices are not accurately estimated by the neural network forecasts, and thus further investigation should be conducted in this regard.

2. As mentioned earlier, several novel hybridisation procedures have been presented throughout this thesis in order to combine the fundamental and statistical approaches. However, there are multiple combination possibilities that have not been considered throughout this thesis that may prove beneficial. For instance, a novel application could involve a regime-switching model within a similar hybrid framework that may be guided by the outputs of the fundamental model. Another possible extension can be found if a deeper analysis is performed in the field of forecast combination, given that it was concluded that a simple averaging procedure between the hybrid and the pure statistical models is challenging to outperform.

3. A potential extension of the final electricity price forecasting methodology of this thesis may be centred in the field of price spikes. In the light of the ongoing expansion of renewable sources in most power systems and the recent weather instabilities, the occurrence of price spikes is expected to increase over the next few years. Therefore, sudden structural breaks may become more common than the gradual ones, and thus the structural breaks calibration period selection method presented in Chapter 4 must be readjusted in order to capture the sudden price changes in the initial dataset. Careful attention must be paid in this oversized dataset in order to maintain computational feasibility.

4. Although feature selection or variable assessment was addressed throughout Chapter 2, Chapter 3 and on Chapter 4's Appendix B, no solid contribution could be provided. However, depending on the market

circumstances, some of the variables considered in the neural network forecasts may become irrelevant. For instance, if the auxiliary forecasts indicate a sharp increase in wind production and thermal generation is excluded from the market clearing, fuel prices will lack any influence over the resulting electricity prices. Therefore, if the dataset is filtered correctly, only similar circumstances should remain that exhibit the key behaviours and relationships with electricity prices that would allow a variable selection procedure to screen out the least relevant and most redundant information for the forecast period. Furthermore, the variable selection procedure must be suitable for non-linear contexts, and appropriate criteria should be chosen so as to stop any iterative variable elimination procedures, such as the one followed in Chapter 2.

Bibliography

Aineto, D., Iranzo-Sánchez, J., Lemus-Zúñiga, L. G., Onaindia, E., & Urchueguía, J. F. (2019). On the influence of renewable energy sources in electricity price forecasting in the Iberian market. *Energies*, *12*(11), 2082. https://doi.org/10.3390/en12112082

Alonso, A. M., Bastos, G., & García-Martos, C. (2016). Electricity price forecasting by averaging dynamic factor models. *Energies*. https://doi.org/10.3390/en9080600

Amjady, N. (2006). Day-Ahead Price Forecasting of Electricity Markets by a New Fuzzy Neural Network. *IEEE Transactions on Power Systems*, *21*(2), 887–896. https://doi.org/10.1109/TPWRS.2006.873409

Amjady, N. (2007). Short-Term Bus Load Forecasting of Power Systems by a New Hybrid Method. *IEEE Transactions on Power Systems*, *22*(1), 333–341. https://doi.org/10.1109/TPWRS.2006.889130

Andrawis, R. R., Atiya, A. F., & El-Shishiny, H. (2011). Combination of long term and short term forecasts, with application to tourism demand forecasting. *International Journal of Forecasting*, *27*(3), 870–886. https://doi.org/10.1016/j.ijforecast.2010.05.019

Andrews, D. W. K. (1993). Tests for Parameter Instability and Structural Change With Unknown Change Point. *Econometrica*, 821–856.

Aneiros, G., Vilar, J. M., Cao, R., & Munõz San Roque, A. (2013). Functional prediction for the residual demand in electricity spot markets. *IEEE Transactions on Power Systems*, *28*(4), 4201–4209. https://doi.org/10.1109/TPWRS.2013.2258690

Angamuthu, R., Mukherjee, A., Campion, M., Salehfar, H., Hansen, T., Lin, J., & Ranganathan, P. (2018). A Multi-Stage Price Forecasting Model for Day-Ahead Electricity Markets. *Forecasting*, *1*(1), 3. https://doi.org/10.3390/forecast1010003

Apergis, N., & Lau, M. C. K. (2015). Structural breaks and electricity prices: Further evidence on the role of climate policy uncertainties in the Australian electricity market. *Energy Economics*, *52*(A), 176–182. https://doi.org/10.1016/j.eneco.2015.10.014

Azadeh, A., Moghaddam, M., Mahdi, M., & Seyedmahmoudi, S. H. (2013). Optimum long-term electricity price forecasting in noisy and complex environments. *Energy Sources, Part B: Economics, Planning and Policy*. https://doi.org/10.1080/15567249.2012.678559

Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*. https://doi.org/10.1002/jae.659

Barquín, J., Centeno, E., & Reneses, J. (2005). Stochastic market equilibrium model for generation planning. *Probability in the Engineering and Informational Sciences*, *19*(4), 533–546. https://doi.org/10.1017/S0269964805050357

Barquin, J., & Vazquez, M. (2008). Cournot equilibrium calculation in power networks: An optimization approach with price response computation. *IEEE Transactions on Power Systems*. https://doi.org/10.1109/TPWRS.2008.919198

Bates, J. M., & Granger, C. W. J. (1969). The Combination of Forecasts. *Journal of the Operational Research*, *20*(4), 451–468. https://doi.org/10.2307/3008764

Bello, A., Bunn, D. W., Reneses, J., & Muñoz, A. (2016). Parametric Density Recalibration of a Fundamental Market Model to Forecast Electricity Prices. *Energies*, *9*(11), 959. https://doi.org/10.3390/en9110959

Bello, A., Bunn, D. W., Reneses, J., & Muñoz, A. (2017). Medium-Term Probabilistic Forecasting of Electricity Prices: A Hybrid Approach. *IEEE Transactions on Power Systems*, *32*(1), 334–343. https://doi.org/10.1109/TPWRS.2016.2552983

Bello, A., Reneses, J., & Muñoz, A. (2016). Medium-term probabilistic forecasting of extremely low prices in electricity markets: Application to the Spanish case. *Energies*, *9*(3). https://doi.org/10.3390/en9030193

Bello, A., Reneses, J., Muñoz, A., & Delgadillo, A. (2016). Probabilistic forecasting of hourly electricity prices in the medium-term using spatial interpolation techniques. *International Journal of Forecasting*, *32*(3), 966–980. https://doi.org/10.1016/j.ijforecast.2015.06.002

Bento, P. M. R., Pombo, J. A. N., Calado, M. R. A., & Mariano, S. J. P. S. (2018). A bat optimized neural network and wavelet transform approach for short-term price forecasting. *Applied Energy*, *210*(January 2018), 88–97. https://doi.org/10.1016/j.apenergy.2017.10.058

Bordignon, S., Bunn, D. W., Lisi, F., & Nan, F. (2013). Combining day-ahead forecasts for British electricity prices. *Energy Economics*, *35*, 88–103. https://doi.org/10.1016/j.eneco.2011.12.001

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, *26*(2), 211–252. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x

Box, G. E. P., & Jenkins, G. (1970). Time Series Analysis - Forecasting and Control. https://doi.org/10.1111/j.1467-9892.2009.00643.x

Carmona, R., & Coulon, M. (2014). A Survey of Commodity Markets and Structural Models. In *Quantitative Energy Finance: Modeling, Pricing, and Hedging in Energy and Commodity Markets* (pp. 41–83). https://doi.org/10.1007/978-1-4614-7248-3

Catalão, J. P. S., Mariano, S. J. P. S., Mendes, V. M. F., & Ferreira, L. A. F. M. (2007). Short-term electricity prices forecasting in a competitive market: A neural network approach. *Electric Power Systems Research*, *77*(10), 1297–1304. https://doi.org/10.1016/j.epsr.2006.09.022

Catalão, J. P. S., Pousinho, H. M. I., & Mendes, V. M. F. (2011). Hybrid wavelet-PSO-ANFIS approach for short-term electricity prices forecasting. *IEEE Transactions on Power Systems*, *26*(1), 137–144. https://doi.org/10.1109/TPWRS.2010.2049385

Chaâbane, N. (2014). A hybrid ARFIMA and neural network model for electricity price prediction. *International Journal of Electrical Power & Energy Systems*, *55*, 187–194. https://doi.org/10.1016/j.ijepes.2013.09.004

Chen, D., & Bunn, D. (2014). The forecasting performance of a finite mixture regime-switching model for daily electricity prices. *Journal of Forecasting*, *33*(5), 364–375. https://doi.org/10.1002/for.2297

Chen, D., & Bunn, D. W. (2010). Analysis of the nonlinear response of electricity prices to fundamental and strategic factors. *IEEE Transactions on Power Systems*, *25*(2), 959–606. https://doi.org/10.1109/TPWRS.2009.2035129

Chow, G. C. (1960). Tests of Equality Between Sets of Coefficients in Two Linear Regressions. *Econometrica*, *28*(3), 591–605. https://doi.org/10.2307/1910133

Cipriano, A., Lira, F., Núñez, F., & Muñoz, C. (2009). Short-term forecasting of electricity prices in the Colombian electricity market. *IET Generation, Transmission & Distribution*, *3*(11), 980–986. https://doi.org/10.1049/iet-gtd.2009.0218

Conejo, A. J., Plazas, M. A., Espinola, R., & Molina, A. B. (2005). Day-Ahead Electricity Price Forecasting Using the Wavelet Transform and ARIMA Models. *IEEE Transactions on Power Systems*, *20*(2), 1035–1042. https://doi.org/10.1109/TPWRS.2005.846054

Contreras, J., Espínola, R., Nogales, F. J., & Conejo, A. J. (2003). ARIMA models to predict next-day electricity prices. *IEEE Transactions on Power Systems*, *18*(3), 1014–1020. https://doi.org/10.1109/TPWRS.2002.804943

Cruz, A. (2013). Contribuciones a la predicción a corto plazo del precio de la electricidad mediante modelos de series temporales (Doctoral dissertation). Universidad Pontificia Comillas. Madrid (Spain).

Cruz, A., Muñoz, A., Zamora, J. L., & Espínola, R. (2011). The effect of wind generation and weekday on Spanish electricity spot price forecasting. *Electric Power Systems Research*, *81*(10), 1924–1935. https://doi.org/10.1016/j.epsr.2011.06.002

de Marcos, R. A., Bello, A., & Reneses, J. (2017). Short-term forecasting of electricity prices with a computationally efficient hybrid approach. *14th International Conference on the European Energy Market - EEM17*.

de Marcos, R. A., Bello, A., & Reneses, J. (2018). Hybridisation of Fundamental and Composite Econometric Modelling for Short-Term Electricity Price Forecasting. *15th International Conference on the European Energy Market - EEM18*.

de Marcos, R. A., Bello, A., & Reneses, J. (2019a). Electricity price forecasting in the short term hybridising fundamental and econometric modelling. *Electric Power Systems Research*, *167*, 240–251. https://doi.org/10.1016/j.epsr.2018.10.034

de Marcos, R. A., Bello, A., & Reneses, J. (2019b). Short-Term Electricity Price Forecasting with a Composite Fundamental-Econometric Hybrid Methodology. *Energies*, *12*(6), 1067. https://doi.org/10.3390/en12061067

de Marcos, R. A., Bunn, D. W., Bello, A., & Reneses, J. (2019). Short-Term Electricity Price Forecasting with Recurrent Regimes and Structural Breaks. Working Paper.

de Marcos, R. A., Reneses, J., & Bello, A. (2016). Long-Term Spanish Electricity Market Price Forecasting with Cointegration and VEC Models. *The 2016 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS 2016)*. https://doi.org/10.1109/PMAPS.2016.7764158

Díaz, C. A., Villar, J., Campos, F. A., & Reneses, J. (2010). Electricity market equilibrium based on conjectural variations. *Electric Power Systems Research*. https://doi.org/10.1016/j.epsr.2010.07.012

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, *13*(3), 253–263. https://doi.org/10.1080/07350015.1995.10524599

Dudek, G. (2016). Multilayer perceptron for GEFCom2014 probabilistic electricity price forecasting. *International Journal of Forecasting*. https://doi.org/10.1016/j.ijforecast.2015.11.009

Federico, G., & Rahman, D. (2003). Bidding in an Electricity Pay-as-Bid Auction. *Journal of Regulatory Economics*. https://doi.org/10.1023/A:1024738128115

Frenzel, S., & Pompe, B. (2007). Partial mutual information for coupling analysis of multivariate time series. *Physical Review Letters*, *99*(20), 1–4. https://doi.org/10.1103/PhysRevLett.99.204101

García-Martos, C., Rodríguez, J., & Sánchez, M. J. (2011). Forecasting electricity prices and their volatilities using Unobserved Components. *Energy Economics*. https://doi.org/10.1016/j.eneco.2011.07.005

Garcia, R. C., Contreras, J., van Akkeren, M., & Garcia, J. B. C. (2005). A GARCH forecasting model to predict day-ahead electricity prices. *IEEE Transactions on Power Systems*, *20*(2), 867–874. https://doi.org/10.1109/TPWRS.2005.846044

Girish, G. P. (2016). Spot electricity price forecasting in Indian electricity market using autoregressive-GARCH models. *Energy Strategy Reviews*, *11–12*, 52–57. https://doi.org/10.1016/j.esr.2016.06.005

González, V., Contreras, J., & Bunn, D. W. (2012). Forecasting power prices using a hybrid fundamental-econometric model. *IEEE Transactions on Power Systems*, *27*(1), 363–372. https://doi.org/10.1109/TPWRS.2011.2167689

Hobbs, B. F., Metzler, C. B., & Pang, J. S. (2000). Strategic gaming analysis for electric power systems: An MPEC approach. *IEEE Transactions on Power Systems*. https://doi.org/10.1109/59.867153

Hoerl, A. E., & Kennard, R. W. (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, *12*(1), 55. https://doi.org/10.2307/1267351

Hoeting, J. A., Madigan, D., Raftery, A. E., & Volinsky, C. T. (1999). Bayesian model averaging: A tutorial. *Statistical Science*. https://doi.org/10.1214/ss/1009212519

Huisman, R., & Mahieu, R. (2003). Regime jumps in electricity prices. *Energy Economics*. https://doi.org/10.1016/S0140-9883(03)00041-0

Jónsson, T., Pinson, P., Nielsen, H. A., Madsen, H., & Nielsen, T. S. (2013). Forecasting electricity spot prices accounting for wind power predictions. *IEEE Transactions on Sustainable Energy*. https://doi.org/10.1109/TSTE.2012.2212731

Keles, D., Scelle, J., Paraschiv, F., & Fichtner, W. (2016). Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks. *Applied Energy*, *162*, 218–230. https://doi.org/10.1016/j.apenergy.2015.09.087

Kosater, P., & Mosler, K. (2006). Can Markov regime-switching models improve power-price forecasts? Evidence from German daily power prices. *Applied Energy*, *83*, 943–958. https://doi.org/10.1016/j.apenergy.2005.10.007

Liebl, D. (2013). Modeling and forecasting electricity spot prices: A functional data perspective. *Annals of Applied Statistics*. https://doi.org/10.1214/13-AOAS652

Liu, B., Nowotarski, J., Hong, T., & Weron, R. (2017). Probabilistic Load Forecasting via Quantile Regression Averaging on Sister Forecasts. *IEEE Transactions on Smart Grid*, *8*(2), 730–737. https://doi.org/10.1109/TSG.2015.2437877

Maciejowska, K., & Weron, R. (2016). Short- and Mid-Term Forecasting of Baseload Electricity Prices in the U.K.: The Impact of Intra-Day Price Relationships and Market Fundamentals. *IEEE Transactions on Power Systems*. https://doi.org/10.1109/TPWRS.2015.2416433

Mandal, P., Senjyu, T., & Funabashi, T. (2006). Neural networks approach to forecast several hour ahead electricity prices and loads in deregulated market. *Energy Conversion and Management*, *47*(15–16), 2128–2142. https://doi.org/10.1016/j.enconman.2005.12.008

Marcjasz, G., Serafin, T., & Weron, R. (2018). Selection of calibration windows for day-ahead electricity price forecasting. *Energies*, *11*(9), 1–20. https://doi.org/10.3390/en11092364

Misiorek, A., Trueck, S., & Weron, R. (2006). Point and interval forecasting of spot electricity prices: Linear vs. non-linear time series models. *Studies in Nonlinear Dynamics & Econometrics, 10*(3). https://doi.org/10.2202/1558-3708.1362

Monteiro, C., Fernandez-Jimenez, L. A., & Ramirez-Rosado, I. J. (2015). Explanatory information analysis for day-ahead price forecasting in the Iberian electricity market. *Energies*, *8*(9), 10464–10486. https://doi.org/10.3390/en80910464

Monteiro, C., Ramirez-Rosado, I. J., Fernandez-Jimenez, L. A., & Conde, P. (2016). Short-term price forecasting models based on artificial neural networks for intraday sessions in the Iberian electricity market. *Energies*, *9*(9). https://doi.org/10.3390/en9090721

Moré, J. J. (1978). The Levenberg-Marquardt algorithm: Implementation and theory. In *Numerical analysis* (pp. 105–116). https://doi.org/10.1007/bfb0067700

Muñoz, A., & Czernichow, T. (1998). Variable selection using feedforward and recurrent neural networks. *Engineering Intelligent Systems for Electrical Engineering and Communications*, *6*(2), 91–102.

Nogales, F. J., & Conejo, A. J. (2006). Electricity price forecasting through transfer function models. *Journal of the Operational Research Society*, *57*(4), 350–356. https://doi.org/10.1057/palgrave.jors.2601995

Nowotarski, J., Raviv, E., Trück, S., & Weron, R. (2014). An empirical comparison of alternative schemes for combining electricity spot price forecasts. *Energy Economics*, *46*, 395–412. https://doi.org/10.1016/j.eneco.2014.07.014

Nowotarski, J., & Weron, R. (2017). Recent advances in electricity price forecasting: A review of probabilistic forecasting. *Renewable and Sustainable Energy Reviews*. https://doi.org/10.1016/j.rser.2017.05.234

Page, E. S. (1954). Continuous Inspection Schemes. *Biometrika, 41*(1/2), 100–115.

Panapakidis, I. P., & Dagoumas, A. S. (2016). Day-ahead electricity price forecasting via the application of artificial neural network based models. *Applied Energy*, *172*, 132–151. https://doi.org/10.1016/j.apenergy.2016.03.089

Pankratz, A. (2012). Building Dynamic Regression Models: Model Identification. In *Forecasting with dynamic regression models (Vol. 935)* (pp. 167–201). https://doi.org/10.1002/9781118150528.ch5

Papadimitriou, T., Gogas, P., & Stathakis, E. (2014). Forecasting energy markets using support vector machines. *Energy Economics*. https://doi.org/10.1016/j.eneco.2014.03.017

Paraschiv, F., Fleten, S. E., & Schürle, M. (2015). A spot-forward model for electricity prices with regime shifts. *Energy Economics*, *47*, 142–153. https://doi.org/10.1016/j.eneco.2014.11.003

Pesaran, M. H., & Timmermann, A. (2007). Selection of estimation window in the presence of breaks. *Journal of Econometrics*, *137*(1), 134–161. https://doi.org/10.1016/j.jeconom.2006.03.010

Portela, J., Muñoz, A., & Alonso, E. (2017). Forecasting functional time series with a new Hilbertian ARMAX model: Application to electricity price forecasting. *IEEE Transactions on Power Systems*. https://doi.org/10.1109/TPWRS.2017.2700287

Raftery, A. E., Painter, I. S., & Volinsky, C. T. (2005). BMA: An R package for Bayesian Model Averaging. *R News*, *5*(2), 2–8. https://doi.org/10.1159/000323281

Rambharat, B. R., Brockwell, A. E., & Seppi, D. J. (2005). A threshold autoregressive model for wholesale electricity prices. *Applied Statistics*, *54*(2), 287–299. https://doi.org/10.1111/j.1467-9876.2005.00484.x

Ruiz, C., & Conejo, A. J. (2009). Pool strategy of a producer with endogenous formation of locational marginal prices. *IEEE Transactions on Power Systems*. https://doi.org/10.1109/TPWRS.2009.2030378

Sánchez De La Nieta, A. A., González, V., & Contreras, J. (2016). Portfolio decision of short-term electricity forecasted prices through stochastic programming. *Energies*, *9*(12), 1–19. https://doi.org/10.3390/en9121069

Sánchez, I. (2008). Adaptive combination of forecasts with application to wind energy. *International Journal of Forecasting*, *24*(4), 679–693. https://doi.org/10.1016/j.ijforecast.2008.08.008

Sandhu, H. S., Fang, L., & Guan, L. (2016). Forecasting day-ahead price spikes for the Ontario electricity market. *Electric Power Systems Research*, *141*, 450–459. https://doi.org/10.1016/j.epsr.2016.08.005

Serafin, T., Uniejewski, B., & Weron, R. (2019). Averaging predictive distributions across calibration windows for day-ahead electricity price forecasting. *Energies*, *12*(13), 1–12. https://doi.org/10.3390/en12132561

Serinaldi, F. (2011). Distributional modeling and short-term forecasting of electricity prices by Generalized Additive Models for Location, Scale and Shape. *Energy Economics*. https://doi.org/10.1016/j.eneco.2011.05.001

Steinert, R., & Ziel, F. (2019). Short- to mid-term day-ahead electricity price forecasting using futures. *Energy Journal*, *40*, 105–127. https://doi.org/10.5547/01956574.40.1.rste

Tejada-Arango, D. A., Domeshek, M., Wogrin, S., & Centeno, E. (2018). Enhanced representative days and system states modeling for energy storage investment analysis. *IEEE Transactions on Power Systems*, *33*(6), 6534–6544. https://doi.org/10.1109/TPWRS.2018.2819578

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B*, 267–288. https://doi.org/10.1111/j.1467-9868.2011.00771.x

Torbaghan, S. S., Motamedi, A., Zareipour, H., & Tuan, L. A. (2012). Medium-term electricity price forecasting. *2012 North American Power Symposium, NAPS 2012*. https://doi.org/10.1109/NAPS.2012.6336424

Uniejewski, B., Nowotarski, J., & Weron, R. (2016). Automated variable selection and shrinkage for day-ahead electricity price forecasting. *Energies*, *9*(8). https://doi.org/10.3390/en9080621

Uniejewski, B., Weron, R., & Ziel, F. (2017). Variance Stabilizing Transformations for Electricity Spot Price Forecasting. *IEEE Transactions on Power Systems*, *PP*(99). https://doi.org/10.1109/TPWRS.2017.2734563

Uniejewski, B., Weron, R., & Ziel, F. (2018). Variance stabilizing transformations for electricity spot price forecasting. *IEEE Transactions on Power Systems*, *33*(2), 2219–2229. https://doi.org/10.1109/TPWRS.2017.2734563

Wan, C., Xu, Z., Wang, Y., Dong, Z. Y., & Wong, K. P. (2014). A hybrid approach for probabilistic forecasting of electricity price. *IEEE Transactions on Smart Grid*. https://doi.org/10.1109/TSG.2013.2274465

Weigt, H., & Willems, B. (2012). The Effect of Divestitures in the German Electricity Market. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1791583

Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, *30*(4), 1030–1081. https://doi.org/10.1016/j.ijforecast.2014.08.008

Weron, R., & Misiorek, A. (2006). Short-term electricity price forecasting with time series models: A review and evaluation. In *Complex Electricity Markets* (pp. 231–254). IEPŁ & SEP.

Weron, R., & Misiorek, A. (2008). Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models. *International Journal of Forecasting*. https://doi.org/10.1016/j.ijforecast.2008.08.004

Wogrin, S., Dueñas, P., Delgadillo, A., & Reneses, J. (2014). A new approach to model load levels in electric power systems with high renewable penetration. *IEEE Transactions on Power Systems*, *29*(5), 2210–2218. https://doi.org/10.1109/TPWRS.2014.2300697

Yan, X., & Chowdhury, N. A. (2013). Mid-term electricity market clearing price forecasting: A hybrid LSSVM and ARMAX approach. *International Journal of Electrical Power and Energy Systems*, *53*(1), 20–26. https://doi.org/10.1016/j.ijepes.2013.04.006

Yan, X., & Chowdhury, N. A. (2014). Mid-term electricity market clearing price forecasting using multiple support vector machine. *IET Generation, Transmission & Distribution*, *8*(9), 1572–1582. https://doi.org/10.1109/ISGTEurope.2013.6695249

Yang, Y. (2004). Combining Forecasting Procedures: Some Theoretical Results. *Econometric Theory*, *20*(1), 176–222. https://doi.org/10.1017/S0266466604201086

Zahid, M., Ahmed, F., Javaid, N., Abbasi, R., Zainab Kazmi, H., Javaid, A., … Ilahi, M. (2019). Electricity Price and Load Forecasting using Enhanced Convolutional Neural Network and Enhanced Support Vector Regression in Smart Grids. *Electronics*. https://doi.org/10.3390/electronics8020122

Zeileis, A., Kleiber, C., Walter, K., & Hornik, K. (2003). Testing and dating of structural changes in practice. *Computational Statistics and Data Analysis*, *44*(1–2), 109–123. https://doi.org/10.1016/S0167-9473(03)00030-6

Zhao, J. H., Dong, Z. Y., Xu, Z., & Wong, K. P. (2008). A statistical approach for interval forecasting of the electricity price. *IEEE Transactions on Power Systems*. https://doi.org/10.1109/TPWRS.2008.919309

Zheng, J., Yang, W., & Li, X. (2017). Training data reduction in deep neural networks with partial mutual information based feature selection and correlation matching based active learning. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2017*, 2362–2366.

Ziel, F., & Steinert, R. (2018). Probabilistic mid- and long-term electricity price forecasting. *Renewable and Sustainable Energy Reviews*, *94*, 251–266. https://doi.org/10.1016/j.rser.2018.05.038

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, *67*(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x