



Modeling the evolution of electricity prices in a context of high penetration of renewables in electricity markets

Valentin Mahler

► To cite this version:

Valentin Mahler. Modeling the evolution of electricity prices in a context of high penetration of renewables in electricity markets. Chemical and Process Engineering. Université Paris sciences et lettres, 2022. English. NNT : 2022UPSLM018 . tel-03870568

HAL Id: tel-03870568

<https://pastel.hal.science/tel-03870568>

Submitted on 24 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à MINES ParisTech

**Modélisation de l'évolution des prix sur le marché de
l'électricité dans un contexte de forte pénétration des
énergies renouvelables**

Modeling the evolution of electricity prices in a context of
high penetration of renewables in electricity markets

Soutenue par

Valentin MAHLER

Le 4 avril 2022

École doctorale n°621

ISMME:

**Ingénierie des Systèmes,
Matériaux, Mécanique et
Énergétique**

Spécialité

**Energétique et
Génie des Procédés**

Composition du jury :

Angelica GIANFREDA Prof., Univ. of Modena and Reggio Emilia	<i>Présidente Examinatrice</i>
Yannick PEREZ Prof., CentraleSupélec	<i>Rapporteur</i>
Lennart SÖDER Prof., KTH - Royal Institute of Technology	<i>Rapporteur</i>
Richard GREEN Prof., Imperial College London	<i>Examineur</i>
Michael POLLITT Prof., University of Cambridge	<i>Examineur</i>
Philippe VASSILOPOULOS Dir. of Product Development, EPEX SPOT	<i>Examineur</i>
Georges KARINIOTAKIS Directeur de recherche, MINES ParisTech	<i>Directeur de thèse</i>
Robin GIRARD Enseignant chercheur, MINES ParisTech	<i>Co-Directeur de thèse</i>

Acknowledgements

Remerciements

To begin, I would like to thank Georges Kariniotakis and Robin Girard, my thesis supervisors. Thank you Georges for allowing me to join your research team and for your good advice. Thank you Robin for proposing and supervising this very interesting thesis topic and thanks for all that I could have learned from you during these years; your dedication to proposing and promoting solutions that take into account environmental limits is admirable. I wish you both much success in your current and future projects.

I thank the president of the jury, Angelica Gianfreda, the rapporteurs, Yannick Perez and Lennart Söder, as well as the jury members, Richard Green, Michael Pollitt and Philippe Vassilopoulos, for your assessment of the work and our informative discussions.

I thank Alexis Bocquet, François-Pascal Neirac and Fabrizio Sossan for your suggestions as members of the follow-up committee.

I acknowledge ADEME and ARMINES for funding this PhD; thank you Brice Arnaud and Valérie Pineau for your availability and our friendly exchanges. I would also like to thank ENTSO-E, EPEX SPOT, the French Ministry for the Ecological Transition, Meteo France, Open Power System Data and RTE for the provision of data that permitted the development and validation of the models presented in this thesis.

I would like to thank MINES Paris - PSL and the PERSEE center for welcoming and supporting me for this thesis. I thank in particular Marie-Jeanne Condo, Patrick Gatt, Laurent Fulcheri, Brigitte Leprat, Isabelle Liotta, Andrea Michiorri, Sophie Pierini, Magalie Prudon, Arnaud Rigacci, and especially Laurent Schiatti de Monza for your help and our pleasant conversations.

This thesis would not have been the same without the many interns, PhD students and post-docs of the PERSEE center with whom I had the pleasure of working. Special thanks to Fian, Hongxin, Kévin and Stéphano; sharing the office with you has been a privilege and an honor. Thanks also to Akylas, Alexis, Anaëlle, Antoine, Biswarup, Konstantinos, Pedro, Romain, Simon, Sylvain, Thomas C. and Thomas H. for your help. I thank those mentioned and all the others for the good moments we shared.

Maman, Papa, je vous dis un très grand merci ; sans votre amour, soutien et éducation rien de tout cela n'aurait été possible. Mathilde, j'ai énormément de chance d'être à tes côtés, merci pour tout. Merci Antoine d'illuminer mes journées de ton sourire. Béatrice et Henri, merci pour votre aide infailible et ô combien précieuse. Merci Laure et Nico, vos relectures et écoutes attentives m'ont apporté un regard extérieur très apprécié. Jo, Rod et Tom, merci d'avoir été là dans les moments forts qui ont ponctué cette thèse.

Note to French-speaking readers

Note pour les lecteurs francophones

A summary in French of each chapter is provided in the appendices.

Un résumé en français de chaque chapitre est proposé dans les annexes.

Contents

1	Introduction	1
1.1	Context	2
1.2	Problem statement	3
1.2.1	Research question	3
1.2.2	Relevance	3
1.2.3	Scope	5
1.3	State of the art and Gap analysis	6
1.4	Objectives and Contributions	7
1.5	Document outline	9
1.6	List of scientific communications	10
2	Energy transition and European Electricity markets	13
2.1	The climate-energy nexus	14
2.1.1	Fossil fuels contribution to climate change	14
2.1.2	What are fossil fuels used for?	14
2.1.3	Levers to reduce fossil fuel emissions	15
2.1.4	Low-carbon electricity for climate change mitigation	15
2.1.5	The energy transition in Europe	16
2.2	The European power system	17
2.2.1	Demand	17
2.2.2	Supply	20
2.2.3	Storage	21
2.2.4	Electrical grid	22
2.2.5	Prospective scenarios	24
2.3	Electricity markets in Europe	25
2.3.1	Structure of the wholesale electricity market	25
2.3.2	Institutions and regulation	28
2.3.3	Market participants	29
2.4	Conclusions	31
3	Price formation on day-ahead electricity markets	33
3.1	Operation of day-ahead markets	34
3.1.1	Setting prices that maximize social welfare	34
3.1.2	Impact of marginal costs on the merit order	36

3.1.3	Types of market orders	38
3.2	Interconnection of day-ahead markets	39
3.2.1	Bidding zones	39
3.2.2	Trading between market zones	41
3.2.3	Liquidity and price convergence in Europe	42
3.3	Stylized facts of electricity prices	45
3.3.1	Predictability of prices on day-ahead markets	45
3.3.2	Time series analysis of prices	45
3.3.3	Analyzing the influence of price drivers	48
3.3.4	Statistical modeling of prices	50
3.4	Conclusions	54
4	Data-driven structural modeling of prices	57
4.1	Literature review	58
4.2	Problem statement	60
4.3	Methodology	61
4.3.1	Market clearing with technical constraints	61
4.3.2	Parametrization of marginal costs	63
4.3.3	Bias correction	64
4.3.4	Estimation algorithm	64
4.4	Validating the approach on a case study	66
4.4.1	Method implementation	66
4.4.2	Data	69
4.4.3	Results	70
4.4.4	Discussion	76
4.5	Conclusions	80
5	Extensions and applications of the model	81
5.1	Further studies of the estimation method	82
5.1.1	Introducing variability into the cost function	82
5.1.2	Leveraging information from observed supply curves	84
5.1.3	Sensitivity analysis of initial conditions	86
5.1.4	Discussion	87
5.2	Model validation for prospective studies	88
5.2.1	Non-regression testing	88
5.2.2	Ramp constraints	88
5.2.3	Active storage management	90
5.2.4	Multi-zone simulation	91
5.2.5	Discussion	94
5.3	Model applications	96
5.3.1	Scenarios for prospective studies	96
5.3.2	Prospective studies	98
5.3.3	Discussion	101
5.4	Conclusions	102

6	Concluding remarks	105
6.1	General conclusions	106
6.2	Perspectives and future work	107
A	Résumé du chapitre 1 :	
	Introduction	109
B	Résumé du chapitre 2 :	
	Transition énergétique et marchés européens de l'électricité	117
C	Résumé du chapitre 3 :	
	Formation des prix de l'électricité sur les marchés journaliers	125
D	Résumé du chapitre 4 :	
	Modélisation structurelle des prix basée sur les données	133
E	Résumé du chapitre 5 :	
	Extensions et applications du modèle	145
F	Résumé du chapitre 6 :	
	Conclusion générale	155

List of Figures

2.1	Greenhouse gas emission intensity of electricity generation in the EU-27 .	16
2.2	Final energy consumption by fuel, EU, 1990-2019	18
2.3	Consumption of electricity by sector, EU-27, 2000-2018	19
2.4	Monthly electricity generation mix in the EU	20
2.5	Installed electrical capacity in the EU	21
2.6	Overview of flexibility solutions	23
2.7	Market participants on EPEX Spot by category	30
3.1	Simplified representation of the market clearing	36
3.2	The bidding zone configuration in Europe as of September 2020	40
3.3	Boxplot of daily traded volumes by month on day-ahead markets in the EU, 2019	43
3.4	Churn factors in major European day-ahead markets, 2015-2019	44
3.5	Price convergence on European day-ahead markets	44
3.6	Hourly electricity prices on the French day-ahead market	46
3.7	Boxplots of electricity prices on the French day-ahead market	47
3.8	Prices as a function of the demand forecast	48
3.9	Prices as a function of the share of weather-driven renewables	49
3.10	Prices as a function of the supply margin	50
3.11	Simulation of prices on the French day-ahead market: linear model . . .	52
3.12	Simulation of prices on the French day-ahead market: random forest . . .	53
3.13	The day-ahead process	54
4.1	Case study overview	67
4.2	Time series of prices on the French day-ahead market over four years (top), average prices by day of the week (bottom-left) and average prices by hour of the day (bottom-right)	71
4.3	Time series of prices on the French day-ahead market over one month . .	71
4.4	Density plot of prices on the French day-ahead market from 2015 to 2018	72
4.5	Mean of residuals by month from 2015 to 2018	73
4.6	Bias correction by hour of the day and day of the week	75
4.7	Performance comparison of benchmarked models	77
4.8	Unavailability of production units	78

4.9	Impact of forced outages on prediction errors. Hourly values (black dots), their corresponding densities (blue lines) and locally estimated scatterplot smoothing, i.e. Loess (the estimation is shown by the solid red line and the 95% confidence interval by the dashed red lines)	79
5.1	Out-of-sample performance for different values of the variability hyperparameter	83
5.2	Data augmentation principle to generate synthetic prices from supply curves	84
5.3	Out-of-sample performance for different augmented datasets (training set: 2015, testing set: 2016)	86
5.4	Best initial merit orders (first decile)	87
5.5	Comparison between results obtained in the previous case study with R (left) and the results obtained under similar conditions with Python (right)	89
5.6	Simulation of French day-ahead prices (reference case)	89
5.7	Multizone case: (A) French prices, (B) Belgian prices, (C) German prices	93
5.8	Boxplot: Comparison of the simulated cases	95
5.9	Metrics plots: Comparison of the simulated cases	96
5.10	Monte Carlo simulations of prices on the French day-ahead market (reference prospective scenario, first week of June)	99
5.11	Monte Carlo simulations of prices on the French day-ahead market (reference prospective scenario, first week of December)	99
5.12	Distribution of electricity prices on the French day-ahead market	100
5.13	Boxplot of French prices on day-ahead markets for the prospective scenarios	102

List of Tables

2.1	Consumption of electricity by sector, EU-27, 2018	19
2.2	Simplified representation of the value chain of power markets	26
4.1	Price drivers of marginal cost proxies considered by production class . . .	68
4.2	Data sources for the case study	70
4.3	Frequency of marginality by production class	74
4.4	Average contribution of each price driver to in the additive decomposition of order prices by production class (values in EUR/MWh)	74
5.1	Ramp limits by production class	90
5.2	Technical constraints for the aggregation of the French pumped hydro power plants in 2015	91
5.3	Computing time to generate and solve the optimization problem for dif- ferent test cases	94

Nomenclature

Variables

P	Production
L	Load
E	Energy transmitted between bidding zones
P^{min}	Minimum production
P^{max}	Maximum production
L^{min}	Minimum load
L^{max}	Maximum load
E^{min}	Minimum transmitted energy
E^{max}	Maximum transmitted energy
S^{hydro}	Hydro stock
Π	Observed market clearing price
$\hat{\Pi}$	Simulated market clearing price
$\hat{\pi}$	Simulated price of an order
α	Sell price parameter
β	Bias correction term
λ^{IC}	Contribution to prices from internal constraints
ϵ	Contribution to prices inferred from data
$\mathcal{F}$	Fuel price
$\mathcal{E}$	Emission price
$\mathcal{M}$	Supply margin
$\mathcal{R}$	Proxy variable for the rank of production unit
$\mathbf{X}$	Vector of price drivers
$\mathbf{Y}$	Vector of hours and weekdays (booleans)
ω	Weighting factor
Π^δ	Synthetic prices
$\mathcal{S}$	Supply curve function
V^δ	Synthetic transaction volumes
r	Ramp constraint
S^{level}	Storage level
S^{in}	Storage variation (charge)
S^{out}	Storage variation (discharge)
d	Storage dissipation
η	Storage efficiency

Subscripts

i	index for a production unit in set I
j	index for a consumption unit in set J
k	index for a storage unit in set K
t	index for a market time unit in set T
z	index for a bidding zone in set Z
$c(i)$	index for a production class in C of unit i

Chapter 1

Introduction

Contents

1.1	Context	2
1.2	Problem statement	3
1.2.1	Research question	3
1.2.2	Relevance	3
1.2.3	Scope	5
1.3	State of the art and Gap analysis	6
1.4	Objectives and Contributions	7
1.5	Document outline	9
1.6	List of scientific communications	10

1.1 Context

Electricity: a commodity in constant equilibrium

Electricity makes the modern world possible. Industrial societies are built on the assumption that power demand shall be met and that outages are rare occurrences[1]. Reliable techno-economic structures, guided by energy policies, are set up to ensure that production matches consumption at all times.

From a technical point of view, consumption units mainly get their electricity from production means connected through the electrical grid. These energy transfers must always be balanced at short time scales, which requires complex coordination.

2021 served as a good reminder that this balance can be fragile, even in developed countries. Issues can arise when the power system is not robust enough to face sudden natural events, such as a cold snap in Texas [2]. A rapid change in worldwide supply and demand for energy, such as the one observed during the economic recovery in mid-2021, can cause a simultaneous surge of coal, gas and oil prices [3], which also leads to higher power prices.

In addition, the long-term equilibrium depends on adequate investments. Power system assets typically have a multi-decade lifetime and usually take years to commission. Thus, anticipation is essential.

Equilibrium through wholesale electricity markets

In liberalized energy markets, the supply-demand equilibrium is met both from a physical point of view (i.e. consumption is equal to production) and from a financial point of view (i.e. energy is bought and sold).

Wholesale electricity markets allow producers and consumers to agree on electricity quantities to be exchanged and to set the corresponding prices. Market participants trade electricity on organized exchange markets or over-the-counter according to their preferences, which notably depend on temporal horizons and delivery points. The supply-demand dynamics are supplemented by fiscal policies (subsidies and taxes) as well as a wide range of regulations, which are defined by public authorities to guide the evolution of the power sector.

A constrained equilibrium

Finding an equilibrium in the short and long runs is a complex task that is directly linked to the three dimensions of the energy trilemma¹: (i) energy security, (ii) energy equity and (iii) environmental sustainability. Under the current societal paradigm, the power system has to be robust enough to ensure energy security, and not oversized to avoid wasting resources. This reliability comes at a cost, and so financial compensation can incentivize market participants to provide and deploy adequate technologies. Meanwhile, electricity is seen as a necessity, and fair access to all must be ensured; unaffordable energy can even spark social unrest. Nevertheless, technology-enabled energy consumption is also responsible for the damage we cause to the environment, which has escalated

¹Term defined by the World Energy Council

since the industrial revolution. Inconsiderate use of materials and energy is causing a biodiversity collapse that cannot be overlooked. Thus, supply-demand not only has to be balanced, but it also has to remain within acceptable physical limits.

The importance of electricity prices

The price associated with electricity is probably the most significant indicator of the value we attribute to this commodity. Considered as a signal[4], the evolution of prices indicates risks and opportunities. Understanding price formation mechanisms to anticipate future price developments helps us to define relevant investment strategies and public policies. This requires identifying key price drivers and their influence. In the European Union, the electricity sector is marked by two major trends: the energy transition, and the development of a fully integrated internal energy market. These two trends directly impact wholesale electricity prices and will be presented in this thesis.

1.2 Problem statement

1.2.1 Research question

Numerical models allow us to study prices in a quantitative fashion. Relying on economic theories and price formation hypotheses, empirical models can be built. This requires using real data about the power system and energy markets, made possible by open data, which are increasingly prevalent thanks to supportive regulation and data-sharing initiatives. Year after year, the pool of available data grows and the quality improves (higher granularity, standardization, etc.). Leveraging these data is key to developing relevant models of electricity prices.

The context of uncertainty regarding the future of electricity prices and the possibilities offered by their modeling to gain a better understanding of price formation lead us to the following central question that motivates this thesis: "How can we best use available public data in order to model the relationship between price drivers and the dynamics of wholesale electricity prices while accounting for the actual price formation mechanism?"

1.2.2 Relevance

The energy transition implies the emergence of new usages and technologies. The profitability of the solutions deployed depends on the price of electricity on wholesale markets, but also on other revenue streams, such as ancillary services, capacity markets, and guarantees of origin. Electricity prices vary depending on time (e.g. prices are typically higher during the day than at night, notably due to demand variations) and location (e.g. bidding areas). Prices can be volatile and vary greatly from one hour to the next, so that high temporal granularity is needed in the modeling to make precise assessments. The following applications illustrate why wholesale electricity price modeling matters for market participants.

Renewable energy sources

Renewable energy sources (RES) are key elements of the energy transition, and their development has benefited from subsidies. One form of subsidy used in the past was feed-in tariffs, whereby renewable energy producers were not exposed to merchant risks, but benefited from a fixed price to sell their electricity. Now that the cost of installing RES has significantly decreased, policies are changing so that RES have to, or will have to, sell on markets as well. Still, in many cases², RES benefit from a feed-in-premium, a subsidy paid in addition to the wholesale price³. The rationale is that exposing RES to price signals should lead to a more effective and stable functioning of the power system (e.g. stop production when prices are negative). But these new rules mean that RES producers have to estimate their potential revenues from the market on the long-run based on price forecasts.

Flexibilities and storage

Once consequence of the growing share of intermittent energy resources is that power systems need to become more flexible in response to new uncertainties and increased variability. This is achieved through a combination of solutions. First, demand side response means that electricity needs are adjusted (i.e. reduced, deferred or advanced) based on price opportunities, which reflects a surplus in cheaper electricity produced (the cheapest electricity is also often the cleanest since wind, solar and rain are free, while carbon emissions induce additional costs due to carbon policies). Second, storage technologies with adequate responsiveness and capacities for different time scales (e.g. daily storage with batteries, seasonal storage with hydrogen) generate revenues from variation in prices, where energy is collected when prices are low and released when they are high (i.e. price arbitrage). Third, sector coupling (i.e. the interaction between electrical, gas and heat networks) helps to create new synergies and to use the most cost-effective energy vector for a given situation, which depends for example on the meteorological conditions. Finally, flexible production units, for which start-stop delays and ramping capabilities are competitive compared to the other technologies in the energy mix, can also benefit from price opportunities and avoid being exposed to prices below their short-run marginal production costs. The valuation of all the above-mentioned flexibilities and the incentives to develop them directly depend on the evolution of wholesale electricity prices.

Retail market

Wholesale prices are a major component of retail prices. Even though our primary focus is on the business-to-business market (B2B), our observations remain relevant for the business-to-consumer market (B2C). Expectation of future prices can guide individuals in their decisions regarding electricity production (whether or not to install PV and batteries, with auto-consumption or connected to the main grid), and also regarding the heating technology they install (e.g. electricity or gas). Smart electricity meters can

²<http://www.res-legal.eu/>

³<https://www.emissions-euets.com/internal-electricity-market-glossary/1811-feed-in-premium>

mean greater exposure to market prices for retail consumers depending on the contract (e.g. regulated or market-based). At the scale of a district or a city, initiatives can also arise regarding the development of local energy communities or, in the future, the development of peer-to-peer (P2P) energy-trading solutions in local energy markets. All of these cases reveal that, to some extent, retail consumers can also choose their level of risk exposure to the potential evolution of prices on wholesale electricity markets.

1.2.3 Scope

Modeling wholesale electricity markets is a broad subject matter. The specific scope of this thesis is defined in this subsection.

Day-ahead markets

Day-ahead markets occupy a central position in the trading of electricity on exchanges. Typically, auctions are organized a day before physical delivery, and a single price (also called uniform price or market clearing price) is set for all traded volume for a given time and location. This time is called market time unit (MTU) and is generally equal to one hour. The location is a node (e.g. a delivery point on the transmission grid) or a zone (e.g. a country) depending on the market design. Since the day-ahead market price is transparent, it gives a significant signal and is also suitable for modeling purposes. In addition, the high temporal granularity (one price per hour) is an essential feature to properly capture the value of energy production and flexibility solutions, which vary throughout the day.

The main other electricity markets are briefly presented in chapter 2, but will not be the focus of this thesis. Day-ahead markets are presented in detail in chapter 3.

The EU internal energy market

Through directives and regulations, legally binding frameworks for electricity markets are established at the European Union level.

First of all, the EU has set an objective to be climate-neutral by 2050, which has implications in terms of the technologies to be adopted in the coming decades (energy efficiency, sector coupling, carbon capture utilization and storage, etc.). The technical constraints imposed by technologies play a role in the choice of market design and in price formation. Moreover, a coordinated effort from the EU helps to combat carbon leakage (e.g. reduce imports of CO₂-intensive electricity from Western Balkan countries) by imposing a Carbon Border Adjustment Mechanism, thus reducing the risk of importing cheap but highly polluting electricity.

The Clean Energy for all Europeans package, published in 2015, was a major political milestone, which laid out the energy union strategy. Four years later, the directive on common rules for the internal market for electricity (EU) 2019/944 completed this package by defining how the market must develop in the coming years. Among the key elements of this directive, we note that the focus is on harmonization between countries to foster competition, while also boosting collaboration and synergies between energy

systems. Single Day-ahead Coupling (SDAC) will create a unique pan-European, cross-zonal, day-ahead electricity market, in which social welfare will be maximized thanks to a common algorithm named EUPHEMIA. Interconnections between areas are constantly being developed, and at least 70% of the capacity must be available to market participants. The European Union Agency for the Cooperation of Energy Regulators (ACER) ensures that long-term remedial actions are taken to limit grid congestion, which prevents the optimal use of energy resources. Similar to production and consumption units, storage assets must participate in the electricity market.

All of these elements indicate that the internal electricity market is set to be developed in a relatively predictable fashion for the coming decade. Major unsuspected structural changes in the coming years are unlikely to occur for the following reasons: (i) political decisions have to be taken by consensus at the EU level involving long processes, (ii) developing the interoperability of IT and governance structures requires years of work, (iii) national power systems have become interdependent, which implies that a sudden withdrawal from cooperation is virtually impossible.

Multiannual time series of hourly prices

We noted that, in the long-run, opportunities and risk arise for investors depending on the evolution of prices. Considering a longer time scale is also necessary for policy makers to establish an adequate market design and fiscal regulations. For these reasons, we will focus on multiannual time series of prices.

As mentioned previously, the value of electricity depends notably on the dynamics of the power system, that is to say, the constraints imposed on supply and demand. In order not to limit the study to global trends, which do not accurately reveal the potential of new investments, we will focus on time series of prices with an hourly resolution, which corresponds to a high temporal granularity.

1.3 State of the art and Gap analysis

The supply-demand equilibrium is typically represented by a bottom-up approach, where the price results from the optimal use of power system assets while accounting for techno-economic constraints. Such models, where the merit order plays a key role in the formation of prices, are called structural models in the scientific literature; a survey of structural models is proposed in [5]. This approach is especially suitable for long-term studies, since the method can account for an evolution in the power system and the market design. These models are widely used, but they require formulating numerous hypotheses in order to identify short-run marginal costs. In addition, the simulations do not necessarily lead to price forecasts that reflect the observed market dynamics, since this class of models is not primarily designed to capture the strategic behavior of market participants. A detailed analysis of the scientific literature is presented at the beginning of chapter 4. Here we examine salient works from the literature and identify how we could build upon these previous endeavors.

While electricity price forecasting is overall a well-developed field of research [6], few

models have been specifically designed for time horizons of one year or more, which corresponds to long-term studies [7]. Among this type of model, an approach for coupled electricity markets is proposed in [8], where the supply curve is constructed by stacking sell orders for which the price offers are a function of the supply margin (i.e. the difference between the total available capacity and the demand) and the production cost by technology. While the proposed approach reflects many aspects of the functioning of the market and the power system, it does not aim at accurately simulating the hourly dynamics of prices. Conversely, in [9], specific attention is given to the variability realism of simulations. Their model starts with a description of the supply curve per technology through the short-run marginal cost. Price offers are then reduced in the case of low demand (reflecting for example must-run constraints) and increased for high demand (inducing a scarcity rent). The authors make use of real data to simulate realistic prices, but their method accounts neither for technical constraints that are typically included in unit commitment problems, nor for the influence of other neighboring markets. In addition, complementary methodologies are proposed in [10] and [11] to model hourly prices using the merit order approach at the technology level to realistically model the price dynamics on the German day-ahead market. Both models make use of real data and linear regressions. The regressions are used in [10] to adjust the output of the merit order model to account for effects such as ramp constraints and strategic bidding, while [11] uses statistics to model the foreign trade balance. These models employ data to enhance the realism of simulations, but they do not propose to directly use statistical calibration within the optimization problem itself in a parametrized fashion, thus potentially limiting the modularity of the problem formulation.

1.4 Objectives and Contributions

The overarching objective of this thesis is to propose methods allowing to leverage the wealth of available data, especially electricity market data, in order to enhance the structural model approach used to simulate time series of electricity prices on day-ahead markets over multiannual horizons and with a fine (i.e. hourly) time step. The novelty of our model lies in its goal to bridge the gap between (i) optimization methods that estimate the minimal total cost of supplying electricity, and (ii) data-driven methods that aim at capturing the observed trends in how the market effectively solves this resource allocation problem. Bringing the two approaches together enables us to build upon robust theoretical foundations without neglecting the effects observed in practice. The model is developed with the intent of being suitable for the estimation of future prices on the long run based on scenarios of the evolution of electricity markets and power systems. Following the above defined overarching objective, complementary technical objectives are formulated:

- Highlight key trends in the EU power system and electricity markets, which will shape the future evolution of wholesale electricity prices
- Leverage open data in order to study the formation of electricity prices

- Propose innovative models that combine optimization and statistical approaches in order to simulate realistic time series of prices
- Formalize the quantitative link between price drivers and market prices through a convex optimization formulation
- Propose a parametrization of price offers in order to account for the techno-economic constraints and strategic behaviors
- Define an algorithm to estimate price parameters while accounting for real data and merit order effects
- Validate the proposed methods over multiple years on a real case study
- Use the models for prospective studies based on scenarios involving possible evolutions of the energy landscape

Key aspects of these objectives are further detailed thereafter.

Modeling the link between price drivers and price dynamics

Hourly prices on a day-ahead market depend on: (i) the state of the power system at the current, previous and upcoming hours, (ii) what happens on other European markets, and (iii) the positions taken by traders on other markets as well as the future positions that they intend to take. Numerous factors influence prices, but the modeling has to focus on those that are considered essential. This selection is non-trivial and partially guided by the availability of trustworthy information. Once the price drivers have been identified based on market analysis and the academic literature, the link between these factors and electricity prices has to be constructed through a model.

In reality, market prices spontaneously emerge from the decisions taken by uncoordinated market participants, and so modeling corresponds to explicitly revealing the underpinning price formation mechanisms. At the root of the market price formation, we find the orders submitted by buyers and sellers for energy volumes at a limited price. Modeling the formation of price offers in a parametrized fashion is thus a key aspect of our approach.

Estimating the model parameters from historical data

A proper formalism has to be proposed in order to exploit data so that the modeling matches the observations more closely. An estimation algorithm is required to learn from the wealth of available data. As the prices are implicitly and indirectly impacted by many drivers, modeling strategies have to be found to reveal these links in order to propose appropriate methods that quantify the contribution of each factor.

Generating time series of price from scenarios

As the evolution of future wholesale electricity prices is uncertain, prospective models have to rely on inputs built from scenarios. These scenarios reflect changes in technical

conditions and economic perspectives. Thus, one of the modeling challenges is to ensure that the simulated time series of prices are sensitive to these changes so that a comparative evaluation of multiple scenarios can be performed. The scenarios themselves have been built based on current trends in the public and private sectors.

1.5 Document outline

This section presents the structure of the remainder of the document.

In chapter 2, we begin with a global overview of the worldwide energy system, and, more specifically, we present the current role of fossil fuel. This introductory analysis led us to the conclusion that electricity will have to play an increasing role if we want to maintain modern lifestyles while reducing carbon emissions. From this worldwide perspective, we focus on Europe and on the EU energy policies that have been proposed to reach net zero emissions by 2050. To be implemented, the European power system will have to keep changing at a tremendous pace. We propose to give key figures and to analyze the current power system through different dimensions (demand, supply, storage and the electrical grid) in order to assess the current situation and reveal trends regarding the future of energy systems. Finally, we explain how wholesale electricity markets facilitate transactions to sustain the power system. The market is structured around power exchanges and other platforms that propose products covering a range of time frames. This structure, as well as the institutions, regulations, and market participants that constitute the European electricity market, are presented in order to understand the reality behind price formation.

In chapter 3, we take a look at day-ahead electricity markets, which represent a key element of the European internal energy market. We detail the operation of these markets, especially the mechanism underlying price formation. The notions of merit order and marginal pricing are explained, as they are a key element for our approach to modeling electricity prices. We also analyze the role of interconnections between bidding zones, notably how they lead to a more efficient use of assets, from both a technical and economic perspective. This understating of the functioning of electricity markets allows us to finally analyze time series of electricity prices. We start the analyses after having observed that wholesale electricity prices are predictable to some extent due to the inherent constraints imposed by the physics of power systems, the deterministic price formation mechanism of the market clearing, and the decisions of market participants based on sound economic reasoning. We progressively complexify the scope of these analyses by first presenting the main features of prices in a univariate fashion, then by highlighting the dependency on price drivers, and finally by showing how these prices can be modeled thanks to statistical and machine learning models. We discuss some limitations of these purely data-driven modeling methods and how they can be mitigated thanks to alternative modeling approaches.

In chapter 4, we propose to go beyond a purely data-driven model by explicitly accounting for our knowledge about the functioning of power systems and markets into our modeling approach. Similar methodologies proposed by academics and practitioners

are discussed in light of a literature review. Our methodology is innovative in the way we propose to take into account historical market and power system data within the structural model framework. We introduce a novel parametrization of market orders and a corresponding estimation algorithm used to statistically determine the value of these parameters. The methodology is presented in detail and then illustrated in a case study related to the French power system from 2015 to 2018. Using our model, multiannual time series of prices are simulated and analyzed. We discuss how the overall method is validated through the case study, but remark that some limitations must still be mitigated in order to apply this approach to prospective studies.

In chapter 5, we build upon the proposed model by allowing the estimation algorithm to account for additional relevant economic data (i.e. supply curves) and to better reflect the variability of time series of prices. These studies show the modularity of the method and how it can be extended depending on the datasets available to the modeler. We also perform a sensitivity analysis to draw attention to the importance of the initialization of the estimation method. We then extend the simulation model itself by introducing additional factors, such as the ramp constraints and active storage management, in order to better reflect the power system constraints that impact price formation. We also verify that the model is still effective when considering multiple bidding zones at once, since this is an essential feature to conduct prospective studies. Using the validated model, applications are finally proposed. A methodology used to build scenarios for the prospective studies is introduced. The results of the scenario-based studies are discussed in light of the strengths and limitations of the method.

Chapter 6 concludes this thesis. The main observations and contributions are summarized. Reflecting on the results achieved so far, perspectives are offered for future research.

1.6 List of scientific communications

This research has been presented to an audience of electricity market experts (peer-reviewed journal and international conference). The thesis has also been presented to a broader audience of scientists interested in various topics related to the energy transition during poster sessions.

In addition, winning the EEM20 Forecasting Competition with a team composed of members of our laboratory led to communications not directly related to the topic of this thesis. The goal of this competition was to propose hourly probabilistic forecasts of wind production aggregated at regional levels.

The list of scientific communications is presented thereafter.

Peer-reviewed journal:

- *Data-driven Structural Modeling of Electricity Price Dynamics* - Valentin MAHLER, Robin GIRARD, Georges KARINIOTAKIS - Energy Economics, Elsevier, 2022, 107, pp.105811, doi:10.1016/j.eneco.2022.105811 (<https://hal.archives-ouvertes.fr/hal-03542564>)

International conferences:

- *Simulation of day-ahead electricity market prices using a statistically calibrated structural model* - Valentin MAHLER, Robin GIRARD, Sébastien BILLEAU, Georges KARINIOTAKIS - 16th International Conference on the European Energy Market (EEM19) - IEEE PES - September 2019 - Ljubljana, Slovenia.
(<https://hal.archives-ouvertes.fr/hal-02351932/document>)
- *Probabilistic Forecasting of Regional Wind Power Generation for the EEM20 Competition: a Physics-oriented Machine Learning Approach* - Kévin BELLINGUER, Valentin MAHLER, Simon CAMAL, Georges KARINIOTAKIS - 17th International Conference on the European Energy Market (EEM20) - KTH, IEEE - September 2020 - Stockholm, Sweden (by visio)
(<https://hal.archives-ouvertes.fr/hal-02952589/document>)
- *Forecasting regional wind production based on weather similarity and site clustering for the EEM20 Competition*, Kévin BELLINGUER, Valentin MAHLER, Simon CAMAL - 40th International Symposium on Forecasting (ISF20) - International Institute of Forecasters - October 2020 - Rio de Janeiro, Brazil (by visio)
(<https://hal.archives-ouvertes.fr/hal-03157849>)

Webinar:

- *EEM20 Forecasting Competition: Probabilistic forecast of aggregated wind production with hourly resolution* - Kévin BELLINGUER, Valentin MAHLER, Simon CAMAL - International Energy Agency Wind Task 36
(<https://youtu.be/n0m3Sl8Zwtk>)

Poster presentations:

- *Modélisation de l'évolution des prix sur le marché de l'électricité dans un contexte de forte pénétration des énergies renouvelables* - PhD student seminar - French Agency for Ecological Transition (ADEME) - March 2019 - Angers, France
- *Simulation of day-ahead electricity market prices using a statistically calibrated structural model* - Summer School on Data-Driven Analytics and Optimization for Energy Systems - Technical University of Denmark, Center for Electric power and Energy (DTU CEE) - June 2019 - Copenhagen, Denmark
(<https://hal.archives-ouvertes.fr/hal-03445344/document>)
- *Modélisation de l'évolution des prix de l'électricité dans un contexte de forte pénétration des énergies renouvelables* - PhD students seminar - Doctoral school ISMME (Ingénierie des Systèmes, Matériaux, Mécanique, Énergétique) - September 2020 - Online

Chapter 2

Energy transition and European Electricity markets

Contents

2.1	The climate-energy nexus	14
2.1.1	Fossil fuels contribution to climate change	14
2.1.2	What are fossil fuels used for?	14
2.1.3	Levers to reduce fossil fuel emissions	15
2.1.4	Low-carbon electricity for climate change mitigation	15
2.1.5	The energy transition in Europe	16
2.2	The European power system	17
2.2.1	Demand	17
2.2.2	Supply	20
2.2.3	Storage	21
2.2.4	Electrical grid	22
2.2.5	Prospective scenarios	24
2.3	Electricity markets in Europe	25
2.3.1	Structure of the wholesale electricity market	25
2.3.2	Institutions and regulation	28
2.3.3	Market participants	29
2.4	Conclusions	31

A prerequisite to properly study electricity markets is to gain a good understanding of supply, demand and how trades are made possible. This is what we propose to do in this chapter. The goal is not only to present the system as it is today, but also how it is likely to evolve. To do so, a multidisciplinary approach is adopted. Key figures and studies are presented in order to base analyses on quantitative data.

We begin by presenting the worldwide energy consumption and its link to the Anthropocene. The specific place of electricity among all energy sources is highlighted and its specific role with regards to the energy transition is explained. After this global overview, we focus on the European power system and how electricity is produced, consumed and supplied to end-users. Finally, we analyze how wholesale markets facilitate the trade of electricity across Europe.

2.1 The climate-energy nexus

2.1.1 Fossil fuels contribution to climate change

In order to significantly reduce the risks and impacts of climate change, the Paris agreement adopted in December 2015, set the goal to hold "the increase in the global average temperature to well below 2°C above pre-industrial levels". The increase in global average temperatures is notably due to fossil CO₂ emissions, which are the main sources of emitted greenhouse gas (GHG) [12]. A carbon budget, i.e. the maximum amount of carbon dioxide (CO₂) that can still be released in the atmosphere before compromising a climate objective, can be estimated thanks to climate models. The 2°C carbon budget is incompatible with the coal, oil and natural gas consumption trends that have been observed since the beginning of the industrial revolution [13, 14]. This budget not only impose to reduce fossil fuel consumption, but also to consider a high share of the carbon reserves as unburnable [15]. Nevertheless, we are absolutely not on track to bridging the emissions gap [12]. The magnitude of climate change will depend on how quickly we reduce these emissions and to which extent we manage to capture emitted greenhouse gases. According to projections, implementing no climate policies should lead to a global average temperature rise relative to pre-industrial levels of around 4.5°C by 2100. The current policies, if implemented properly, should lead to an increase of approximately 3°C. Changes in the energy sector are essentials since energy consumption is responsible for three-quarters of the global greenhouse gas emissions [16].

2.1.2 What are fossil fuels used for?

Fossil fuels, which store energy in potential form (chemical energy), are mainly used to generate kinetic energy (motion energy, electrical energy and thermal energy). Globally, the motion energy used in the transport sector comes quasi exclusively from oil. According to the International Energy Agency [17], around two-third of the electrical energy comes from fossil fuels (mostly from coal and natural gas). Besides being used for electricity generation, coal and natural gas are also used to produce heat for two main

usages: (i) in industrial processes and (ii) to warm air and water for commerce, public services and the residential sector. In addition, fossil fuels are also used as feedstock to produce goods.

2.1.3 Levers to reduce fossil fuel emissions

Reducing emissions imply to decrease the product between the final energy consumption and the emission intensity. Ideally, emission intensity should be estimated thanks to whole life-cycle assessment in order to account for upstream emissions, such as conversion losses from primary to final energy.

Final energy consumption depends on the world population and the consumption per capita (keeping in mind that the variability between individuals is considerable: "the emissions of the richest 1% the global population account for more than twice the combined share of the poorest 50%" [12]). The fair level of consumption per capita relates to the concept of energy sufficiency [18], which is hard to quantify and to transpose into acceptable public policies. Energy sufficiency is also unlikely to be spontaneously adopted by most individuals [19].

Emission intensity can be reduced by using electricity produced from low-carbon sources (mainly hydropower, nuclear, wind and solar). Large scale electrification of the heating & cooling sector and the transport sector is thus an ongoing industrial challenge and a key component of the energy transition [20].

2.1.4 Low-carbon electricity for climate change mitigation

A set of solutions complementary to low-carbon electricity have been proposed to foster the energy transition:

- Develop storage technologies, such as batteries and hydrogen, which enable to use low-carbon energy that would have been curtailed otherwise. Storage devices can replace fossil fuels reservoirs for many applications.
- Expand transmission and distribution infrastructures to ensure that green energy can be used where and when it is needed.
- Reduce energy consumption through energy efficiency measures, while ensuring that economy-wide rebound effects are managed [21].
- Negate emissions thanks to the capture of carbon in flue gas or in the atmosphere, using nature-based solutions or dedicated technologies [22] (the potential of these solutions is limited and they should not be seen as a viable alternative to drastic emission reductions [23]).

In spite of these innovations, it remain unclear whether green growth is feasible [24, 25] and if it is a suitable solution to the problems we face [26], knowing that Anthropocene challenges are not limited to climate change [27]. The progressive increase in primary

energy consumption from low-carbon sources over the last decades, both in absolute numbers with regards to the share of the total primary energy consumption, did not coincide with a decrease in fossil fuel consumption. As a consequence, effective methods to globally and durably reduce emissions from fossil fuels still have to be empirically demonstrated.

2.1.5 The energy transition in Europe

As the cradle of the industrial revolution, Europe is directly and indirectly responsible for a significant share of cumulative emissions. The continent is now dedicated towards net-zero targets and is at the forefront of energy transition policies (as of October 2021, outside Europe, only Japan and New-Zealand have passed net-zero targets in law¹). Since the European Union is responsible for 7% of the annual greenhouse gas emissions¹, a rapid and successful transition that do not rely on emission outsourcing would demonstrate a feasible transition path on a large scale. So far, the decarbonization efforts are encouraging: as result of the decreasing share of fossil fuels in the energy mix, the greenhouse gas emission intensity of electricity generation has decreased by a half since 1990, as illustrated in Figure 2.1.

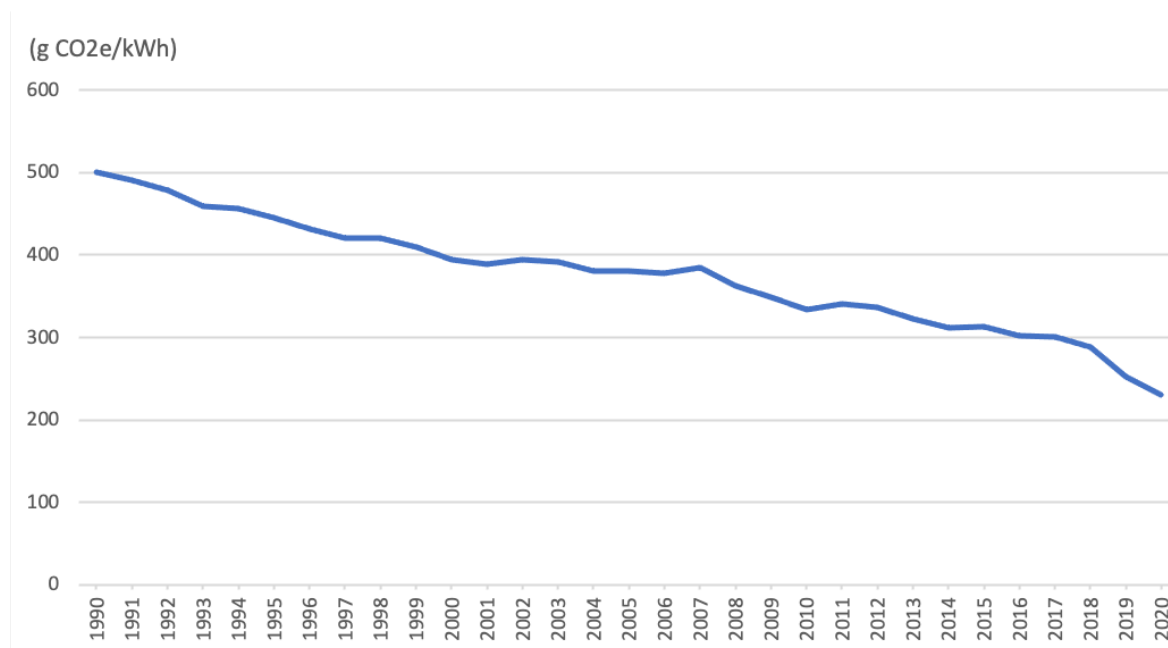


Figure 2.1: Greenhouse gas emission intensity of electricity generation in the EU-27 [28]

To reach climate neutrality by 2050, the EU commission set the target to reduce greenhouse gas emission to at least 40% in 2030 compared to 1990. This 2030 target should be achieved thanks to share of renewable energy reaching at least 32% and energy

¹<https://www.climatewatchdata.org/net-zero-tracker>

efficiency of 32.5% compared to 1990 [29]. This ambition is likely to be revised upwards with the upcoming Fit for 55 package, which is discussed among member states. To reach the 2030 objectives, each country develops his own national strategy. The chosen approach is detailed in National energy and climate plans (NECPs) [30], which have to be in line with the national long term strategies meant to reach the Paris agreement at the horizon 2050 [31]. The International Energy Agency conducted a comprehensive review of the EU energy policy, leading to the publication of a detailed report on energy system transformation and energy security [32].

Carbon trading is one of the key instruments of transition policies. A limited amount of greenhouse gases emissions by the power, industry and aviation sectors is allowed in the EU. The market price of emissions evolves according to supply and demand within the EU Emission Trading System (EU ETS). The cap on the number of emission allowances decreases every year on a regulatory basis. For stationary installations such as power plants, the EU-wide cap was reduced by 1.74% per year for the period 2013-2020 and is reduced by an annual factor of 2.2% for the period 2021-2030. For the year 2021, emissions corresponding the global warming potential of around 1.6 billion tons of CO₂ is allowed [33].

2.2 The European power system

2.2.1 Demand

As for the rest of the world, electricity represents a minor share of the final energy consumption in Europe (around 10% of the total consumption), while fossil-based energies are dominant, as shown in Figure 2.2. This means that the potential for electrification is still consequent in Europe. Overall, electricity consumption is quite stable for the period 2000-2018. Electricity is mainly consumed by the industry sector (around 38%), services (around 29%) and households (around 28%), as shown in Figure 2.3. The main industrial contributors are (i) chemical & petrochemical, (ii) machinery, (iii) iron & steel, (iv) food, beverage & tobacco, and (v) paper, pulp & printing (details are given in Table 2.1). More specific data about energy consumption are published by the European Commission [34], including energy balance sheets, which contain country-specific information [35].

The composition of electricity demand will evolve due to electrification of end-uses and the development of new usages. The main drivers for electrification are heating, ventilation and air conditioning (heat-pumps), industrial heat (boilers, furnaces), e-mobility (electric vehicles) and hydrogen production. In the future, carbon capture should represent a new electricity-intensive industry. We also note that computer systems play an increasingly important role and lead to new usages (internet of things, industry 4.0, high-definition video, virtual reality, cryptocurrencies, smart contracts, ...). Energy efficiency measures, such as the use of combined heat and power systems, can help to counter-balance some of the consumption increase.

Demand side management is likely to influence the future temporal consumption

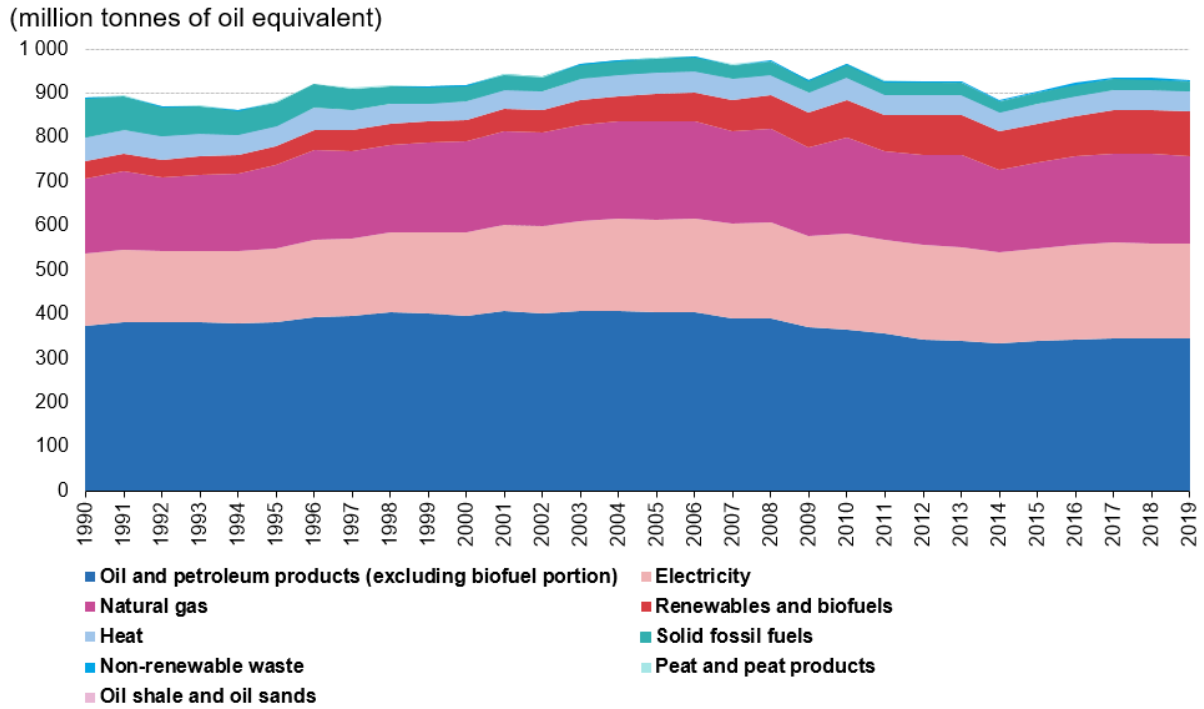


Figure 2.2: Final energy consumption by fuel, EU, 1990-2019 [36]

patterns of electricity. Costs and CO₂ emissions reductions can be achieved by taking into account the supply constraints when using electricity. As hourly emissions by country resulting from electricity generation can be estimated from open data and retrieved in real time [37], some consumers can reduce their carbon footprint by balancing loads across time and space (e.g. non-urgent computation on a server).

2.2. The European power system

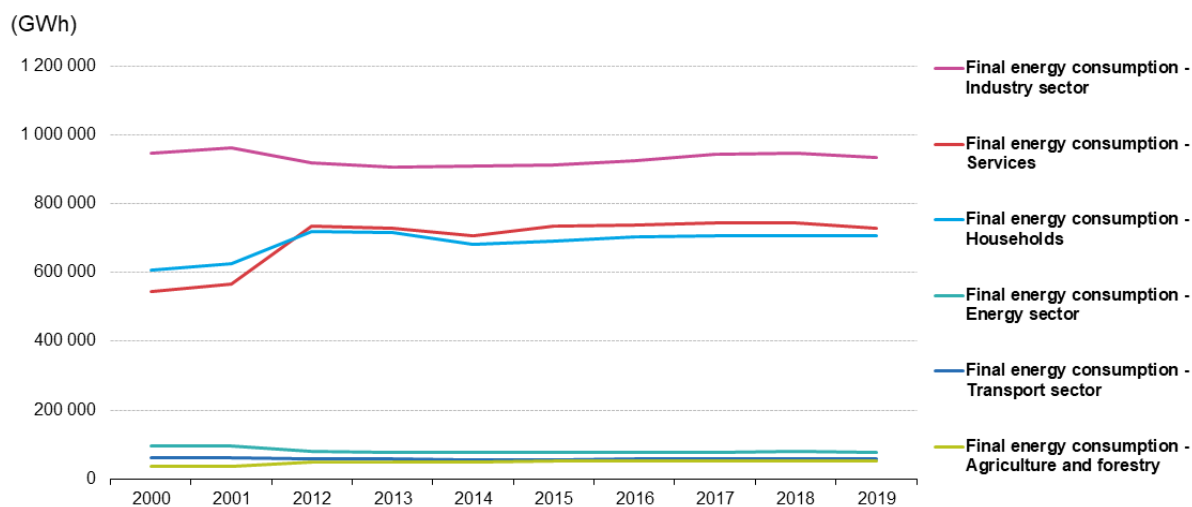


Figure 2.3: Consumption of electricity by sector, EU-27, 2000-2018 [38]

Table 2.1: Consumption of electricity by sector, EU-27, 2018 [38]

(GWh)	Electricity
Final energy consumption	2511766
INDUSTRY SECTOR	956463
Iron & steel	112171
Chemical & petrochemical	169042
Non-ferrous metals	63616
Non-metallic minerals	64895
Transport equipment	50451
Machinery	113102
Mining & quarrying	20364
Food, beverages & tobacco	109985
Paper, pulp & printing	105077
Wood & wood products	25993
Construction	25260
Textile & leather	18550
Not elsewhere specified (industry)	77956
TRANSPORT SECTOR	58929
Rail	48590
Road	1675
Domestic aviation	0
Domestic navigation	0
Pipeline transport	2059
Not elsewhere specified (transport)	6606
OTHER SECTORS	1496374
Commercial & public services	737063
Households	705522
Agriculture & forestry	50846
Fishing	675
Not elsewhere specified (other)	2268

2.2.2 Supply

The generation mix evolves according to demand, installed capacity, weather and technico-economic constraints. The fossil fuel share in electricity production is progressively decreasing in Europe and is replaced by renewable energy sources, as shown in Figure 2.4.

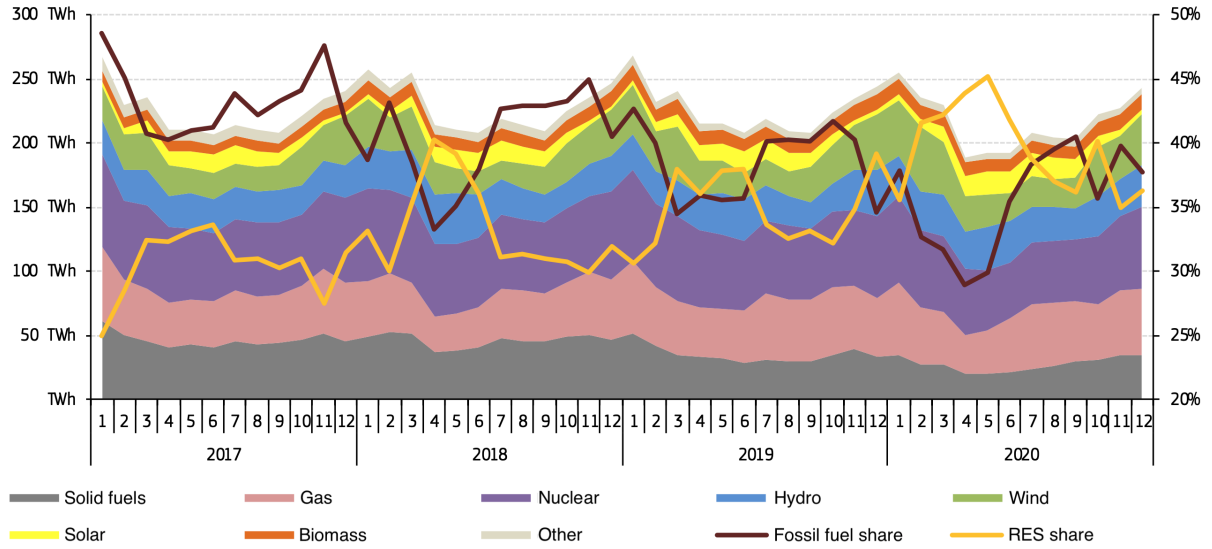


Figure 2.4: Monthly electricity generation mix in the EU [39]

Over the past two decades, significant increase in global installed capacity has been observed, especially for wind, solar photovoltaic (PV) and gas power plants, as shown in Figure 2.5. This evolution can be understood in light of the change of unsubsidized levelized cost of energy (LCOE) for utility-scale generation. According to Lazard’s analysis [40], the mean LCOE in 2020 for the key technologies are by \$175/MWh for gas peaker (-36% compared to 2009), \$163/MWh for nuclear (+33%), \$141/MWh for solar thermal tower (-16%), \$112/MWh for coal (+1%), \$80/MWh for geothermal (+6%), \$59/MWh for gas in combined cycle power plants (-29%), \$40/MWh for wind (-70%) and \$37/MWh for solar using crystalline silicon photovoltaics (-90%). The cost reduction for renewables is driven by decreasing capital costs, improving technologies and increased competition. An acceleration of the shift in energy mix would have significant consequences in term of revenues for the owners of stranded assets (e.g. coal power plants).

Power plant databases are available online and a tool has been proposed to match and compare them [41] (European power plants by capacity and fuel type also can be visualized on an interactive map [42]). In addition to large power plants, distributed generation is gaining traction for renewables (mainly PV), but also for nuclear technologies (small modular reactors[43], which represent a largely unproven solution for a fast and effective energy transition).

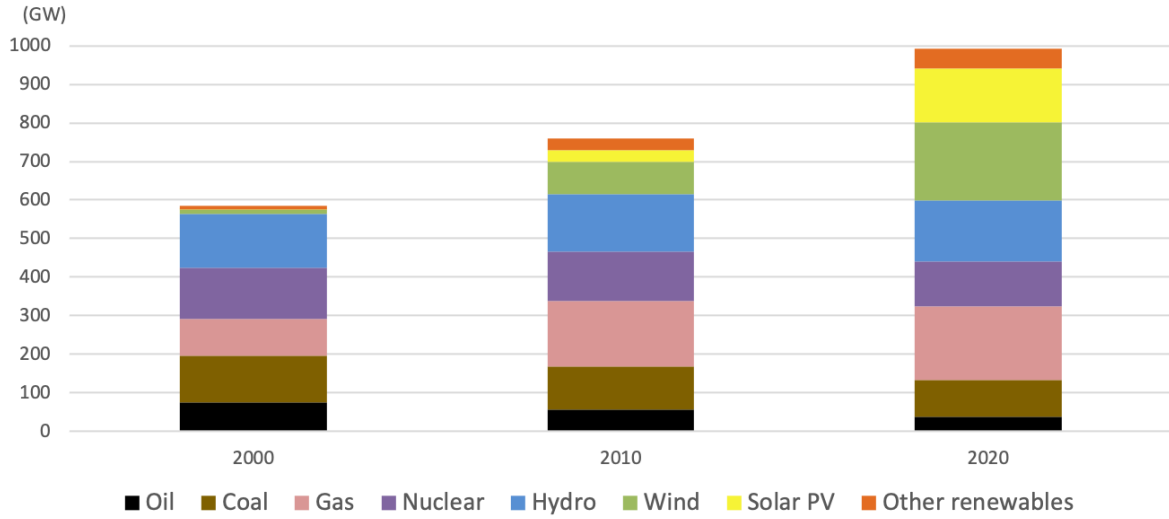


Figure 2.5: Installed electrical capacity in the European Union [44]

In the coming decades, new geographical areas will be used by the power sector. Renewable energy production will increasingly take place offshore [45, 46] (mainly thanks to wind turbines, since waves and tidal technologies are not expected to grow significantly). Moreover, because of drastic cost reduction of enabling technologies, renewable energy produced outside Europe can also be imported directly as electricity [47] or through a chemical conversion, such as the production of green hydrogen [48]. Bioenergy (biomass and biogas) will replace some of the uses of fossil fuels, including electricity generation, but its bio-feedstock would have to come mainly from wastes, agricultural residues and industrial by-products in order to not cause detrimental impacts in terms of land and ocean uses. Geothermal technologies, which enable to make use of energy stored underground [49], is not a key element of the European transition strategy. For the sake of completeness, we also note that energy production could theoretically take place at higher altitudes or in space, but it is at the research stage and should play no role in energy production in the foreseeable future. In addition to these considerations regarding the locations of power system assets, the energy transition will depend on adequate access to materials and components for all technologies. Finally, sufficient water resources near hydraulic power plants and nuclear power plants will be essential.

2.2.3 Storage

Storage technologies have different applications, which can be classified as (i) generation support services and bulk storage services, (ii) services to support transmission infrastructure, (iii) services to support distribution infrastructure, (iv) ancillary services, and (v) services providing customer energy management [50]. To fulfill this range of applications, different technologies have to be used depending on the required capacity

(i.e. energy), responsiveness (i.e. power) and the needed efficiency to be competitive. The technologies can be mechanical (e.g pumped hydro, compressed air, flywheels), electrochemical (e.g. batteries), electrical (e.g. supercapacitors, superconducting magnetic energy storage), chemical (e.g. power-to-gas) or thermal (e.g. heat storage, thermochemical storage) [51].

Driven by the conjunction of new needs and technological innovation, storage takes an increasingly important role in power systems. Pumped hydro storage is the main energy storage in the EU, by such a margin that it could be considered as the only significant one (it is the only storage technology reported in the transparency platform for the pan-European market [52]). Power-to-gas, especially green hydrogen (i.e. hydrogen produced thanks to low-carbon electricity), can be used for long-term energy storage; but it will have to compete with other usages, which might be more interesting in terms of costs and carbon reduction (chemical and industrial processes, transportation, ...)[53]. Batteries become competitive for short- to mid-term storage and are progressing rapidly both as bulk storage and as distributed solution "behind-the-meter" [54].

2.2.4 Electrical grid

The electrical grid is composed of a transmission network and a distribution network. For a given geographical area, these grids are respectively managed by a transmission system operator (TSO)² and a distribution grid operator (DSO).

Structure

Europe has the largest interconnected electrical grid in the world. As of 2015, around 310 000 km of transmission lines were managed by the European TSOs [55]. The European TSOs are represented by the European Network of Transmission System Operators (ENTSO-E). The legal mandate of ENTSO-E, as defined in regulation (EC) 714/2009 [56], is to “ensure optimal management of the electricity transmission network and to allow trading and supplying electricity across borders in the Community”. ENTSO-E missions include enhancing the creation of the Internal Electricity Market (IEM) and to perform pan-European resource adequacy assessment [57]. Maps of the EU transmission grid are published by ENTSO-E [58] and the corresponding data are available online [59]. A list of open data related to transmission grids is maintained by the Open Energy Modelling Initiative [60].

In the past decades, distribution grids were more seen as subordinated to the transmission grid and thus more passive regarding electricity markets. But, with the growing share of distributed generation and storage, combined with demand side response, DSOs play an increasingly active role in the coordination of supply and demand. The Electricity regulation (EU) 2019/934 [61] lead to the creation of the EU DSO Entity in order to "increase efficiencies in the electricity distribution networks in the Union and to ensure close cooperation with transmission system operators and the ENTSO for

²RTE is the unique operator of the whole French transmission system, but other countries such as Germany are split in geographical areas managed by different TSOs

Electricity". The French distribution network is mainly operated by Enedis, which is in charge of around 95% of all lines, while the remaining 5% are under the responsibility of 160 smaller companies [62]. The process of publishing open data regarding the French distribution network is ongoing [63].

Opportunities and risks

The electrical grid plays a key role in the energy transition as it enables to exploit the full potential of flexibility solutions. An increase in flexibility at different time scales would facilitate the integration of weather-driven renewable energy sources while ensuring the security of supply. Figure 2.6 summarizes the main actions to consider. In addition, the electrical grid is a central element of both end-use sector coupling (transport, industry, heating & cooling) and cross-vector integration (electricity, gas and heat networks), which are beneficial in terms of efficiency, flexibility, reliability and adequacy [64].

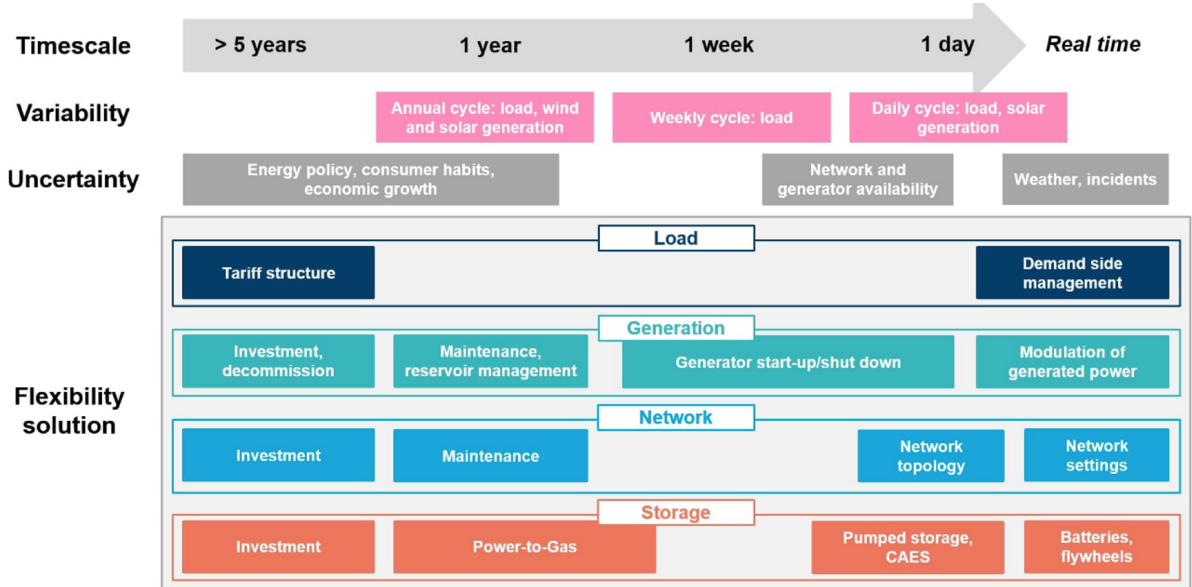


Figure 2.6: Overview of flexibility solutions [65]

Nevertheless, new challenges arise with this increasingly important role of the grid. Due to the importance of power electronics in inverter-based systems, notably led by the penetration of renewables [66], additional technical solutions have to be implemented to ensure the robustness and stability of the grid [67, 68]. Moreover, the energy sector is at the center of the critical infrastructure interdependencies [69], which means that ensuring reliable operation of the electrical grid is of utmost importance. Strategic assets have to be protected against physical threats (natural phenomenon and attacks); the risks of cyberattacks, which are prevalent because of digitalization, also have to be carefully mitigated.

2.2.5 Prospective scenarios

The path towards net-zero emissions in 2050 is uncertain and scenarios at different geographical scales have been proposed.

Worldwide

The International Energy Agency (IEA) proposes a pathway at the world scale. Power system data relative to electricity demand, the total capacity and generation by technology have been published but they are only presented at the global scale (no data for Europe or France are available).

Europe

Scenarios until 2050 at the European scale with detailed power system data at the national level have been proposed by the European Commission (EC) and by ENTSOs³. The European commission proposed a EU Reference Scenario 2020, which has been developed to be consistent with the EU energy objectives. While the 2020 targets for greenhouses gases have been met, additional efforts will be required to fulfill the 2030 objectives according to projections from the European Energy Agency [70]. ENTSOs scenarios are declined the Ten-Year Network Development Plans (TYNDP) published every two-years. The 2020 plans presented three scenarios: "National Trends" is based on the national energy and climate plans (NECPs) published by member states, the two other scenarios ("Global Ambition" and "Distributed Energy") are built to be compliant with the 1.5°C target of the Paris Agreement and the EU's climate targets for 2030. These last two scenarios differ in term of technology choices, "Global Ambition" is driven by centralized generation as opposed to "Distributed Energy", which relies on smaller generation units and gives a greater role to prosumers (i.e. consumers that actively manage their consumption based on power system constraints and market conditions). The IEA, EC and ENTSOs scenarios share two common traits: (i) a progressive increase in electricity demand due to electrification, (ii) an evolution in the composition of the electricity mix both in terms of installed capacity and generation (a decrease of fossil fuels and an increase of renewables).

France

In France, two comprehensive sets of scenarios until 2050 horizon have been published in fall 2021: one set published by the transmission system operator (RTE), the other by the agency ecological transition (ADEME). RTE scenarios enable to consider different shares of renewables and nuclear energy in the electricity mix. The best scenario from an economic perspective is to deploy renewable energy sources at a faster pace and to also build new nuclear power plants in order to keep a significant share of dispatchable

³ENTSOs encompasses the European associations for the cooperation of transmission system operators for electricity (ENTSO-E) and for gas (ENTSOG)

generation. On the demand side, multiple possible evolutions are explored (energy sufficiency, growing industries, different levels of electrification, low energy efficiency due to significant rebound effect, hydrogen as key energy vector). For most scenarios, electricity consumption is projected to increase, as electricity would replace fossil fuels in many applications.

The ADEME scenarios will be broader in scope as they will consider all energy sources and multidimensional analyses (technico-economical and environmental criterias). The final results are not public yet, especially the conclusions of this study regarding the power system have not been published as of December 2021. Still, it is interesting to note the four paths that are explored by ADEME to reach carbon neutrality by 2050. The first emphasises the importance of reconsidering our consumption habits. The second highlights the benefits of stronger cooperation between the public and private sectors, especially at the scale of regions. The third explore a green growth path relying on energy efficiency and decarbonized energy. The fourth focus on circular economy, which would allow to keep most of our consumption habits but with a lower environmental impact.

Remarks

All of these scenarios indicate that weather will have an increasingly important influence on power systems. Temperature, wind, solar irradiance and precipitation will impact both demand and supply (heating and cooling demand, renewable production, water resources for the operation of nuclear power plants, etc.). The problem is that weather will become less predictable because of climate change, so oversizing future power systems might be needed to guarantee resource adequacy.

In any case, there is a general consensus that power systems will have to evolve significantly and fast. The bottleneck, which limit the extent or speed of the transition is not clearly identified. But, we know for sure that massive investments are required to make the proposed transition possible. Financial resources from different streams will have to be mobilized and a significant share of investment portfolios will have to be allocated to green solutions to make the transition possible.

2.3 Electricity markets in Europe

2.3.1 Structure of the wholesale electricity market

The value chain of power markets

Electricity is sold and bought on markets at different time frames for a given location. Having multiple markets respond to different technical use cases (unit commitment, forecast correction, ...). From a financial perspective, it allows to hedge and to speculate in order to manage risks and to maximize profits. As a simplified overview, the value chain of electricity markets is summarized in Table 2.2.

Electricity markets are also commonly called power markets in the private sector. To differentiate between power and energy, the terms availability product (MW) and an activation product (MWh) are sometimes used. When not specified otherwise in the thesis, we refer to the trading of energy.

	Futures and forwards	Day-ahead auction	Intraday	Balancing
Time frame	- Up to several years ahead	- Day before delivery	- Last hours before delivery	- Last minutes before delivery
Main use case	- Buy and sell in advance to secure business against price fluctuations	- Optimise and balance portfolio - Nominate volumes to be fed into the grid	- Adjust for short-term changes	- Deliver committed volumes - Balance the grid to match supply and demand

Table 2.2: Simplified representation of the value chain of power markets [71]

For trading occurring more than a day before electricity delivery, derivatives markets enable market participants to hedge their position or speculate on price evolution. Base futures are one of the most common contracts of derivative markets. Such contract corresponds to a single price for all hours for a longer period (e.g. a month) for a single location. More complex financial instruments are also proposed on derivative markets (e.g. options, swaps).

For trading occurring less than 24 hours before electricity delivery, intraday markets give an opportunity for market participants to trade electricity through auctions or continuous trading. The MTU can be shorter (e.g. 15 minutes). These markets make it possible to account for the latest forecast corrections (demand, weather) and sudden outages. As such, intraday markets are complementary to day-ahead markets and are becoming increasingly popular, notably due to new uncertainties in power systems. Finally, balancing markets are activated a few minutes before physical delivery and they ultimately ensure that supply and demand are matched. For these markets, algorithmic trading is becoming prevalent.

The above-mentioned markets give an overview of pool-based electricity trading on exchanges. In addition, over-the-counter (OTC) transactions allows market participants to buy and sell electricity through standard⁴ or specific⁵ bilateral contracts. These contracts between two business entities are less transparent and, as such, less prone to scientific study. Corporate Renewable Power Purchase Agreement (CPPA) are growing

⁴<https://efet.org/standardisation/standard-contracts-gas-power/>

⁵<https://efet.org/standardisation/specific-contracts/>

fast, mainly due to large companies wanting to procure green energy⁶⁷ and the will of RES producers to hedge their production. As other OTC transactions, we can also mention RTE's block exchange service, which allows balancing responsible parties to trade in order to avoid paying penalties by ensuring that their portfolio is balanced.

The Agency for the Cooperation of Energy Regulators (ACER) publishes the list of standard energy contracts [72] offered by organized market place (i.e. exchanges of brokers)⁸.

Some wholesale energy transactions also occur out of the market. These transactions organized by public authorities, such as energy regulators, are specifically developed in response to energy policies. In France, these concern RES and nuclear energy. Historical RES projects can still benefit from feed-in tariff (i.e. regulated fixed prices regardless of the demand). Some new projects, such as offshore wind farms, are developed through tenders where the installed capacity is decided by public policies and the price depends on investor bids. Regarding nuclear, companies can buy nuclear electricity from EDF for a fixed price defined by public authorities thanks to the ARENH mechanism (*Accès Régulé à l'Electricité Nucléaire Historique*).

Price formation across markets

According to liberal vision advocated by the European Federation of Energy Traders (EFET) [74], the formation of electricity prices must be seen in light of these multiple markets corresponding to different time frames. Trading is described as a way to manage risks of imbalance costs that reflect the difference between sell and purchase of electricity compared to the power uses that are effectively observed. These risks and opportunities are progressively arbitrated more than years in advances and until real time. Trading modalities (markets, time frames, volumes, prices) should be determined by market participants as long as they are based on commercial, technical and/or economic assessment. Excessive restrictions or fear of potential sanctions would "harm the ability of the market to let price spikes materialise and to properly respond to this signal. As a result, necessary investments in flexible capacity and innovative energy services may not happen, as associated development and investment cost may never be recovered."

The price observed on day-ahead auction markets is especially interesting as it is typically unique for a given hour and location and, as such, can serve as a common, transparent price signal. It is worth noting that this price signal commonly serves as an indicator for pricing on other markets, such as forward markets (e.g. in [8]).

⁶<https://leveltenenergy.com/blog/energy-procurement/corporate-renewable-energy-power-purchase-agreements-europe/>

⁷<https://resource-platform.eu/>

⁸Based on [73], the main types of contracts are classified as follows: auction, continuous, forward style contract, future style contract, option style contract, option on a forward, option on a future, option on a swap, spread, swap (financial)

Pan-European spot markets

The creation of an internal energy market in the EU is an ongoing process that leads to more integration and harmonization. A detailed economic assessment of the European single market in electricity is proposed in [75], in which the impact of institutional changes on prices, security of supply, the environment, and innovation are notably examined.

The most notable evolution for spot markets (i.e. day-ahead and intraday markets) are the creation of pan-European cross zonal Single Day-Ahead Coupling (SDAC) and Single Intraday Coupling (SIDC) in compliance with the Capacity Allocation and Congestion Management Regulation (CACM Regulation). The goal is to create a single pan European cross zonal day-ahead electricity market. In order to allow a better use of power system resources, technical changes are currently implemented such as: (i) using flow-based capacity calculation to optimize the use of interconnexion, (ii) transitioning from a Market Time Unit (MTU) of one hour to fifteen minutes in order to promote a better use of flexibility solutions [76]. The SDAC will be further presented in the next chapter of this thesis.

Energy transition and market design

Both TSOs [77] and energy regulators [78] question whether the current market conditions and market design will provide the investment signals needed to ensure long-term security of supply. The reason is that a higher share of RES with low operational costs risks to depress wholesale electricity prices and push dispatchable assets out of the merit order, thus making them unprofitable even though some are needed for the security of supply. Capacity market is the main solution proposed to mitigate this risk.

In the context of the current and future EU power system, effective congestion management and the adequate provision of ancillary services are recurring topics for which adapted market solutions are considered [77]. Definitive solutions are not adopted yet, but local flexibility markets emerge as a complementary solution to these issues [79]. These local markets are meant to compensate for the low spatial granularity inherent to the zonal market design. Nevertheless, the coexistence of zonal and local markets set the condition for undesired increase-decrease gaming (inc-dec), where "market parties anticipate the redispatch market and bid strategically in the zonal market" [80].

2.3.2 Institutions and regulation

Energy trading is facilitated by organized market places. Although power exchanges operated as monopolies for given geographical areas, they are now in concurrence. For example, in addition to EPEX Spot, Nordpool and Nasdaq are now operating in the French market. The list of all market places is maintained by energy regulators [81].

Central counterparties (CCP), also known as clearing houses, are a common and sometimes even mandatory intermediary of transactions. By acting as both seller and buyer of all contracts they reduce risks for all market participants and are meant provide

robustness to the financial system [82, 83, 84]).

Transmission system operators (TSOs) are ultimately in charge of ensuring the security of supply and they make the physical settlement of transactions possible. Since trading also occurs across borders, interconnection capacities have to be allocated and made available for market participants. This effort is coordinated through the Joint Allocation Office (JAO), which is in charge of the Single Allocation Platform (SAP) for all European TSOs. [85]. JAO computations rely on data and insights provided by the regional security coordinators (RSC)⁹.

Energy markets are monitored and regulated by National Regulatory Agency (NRAs), such as the *Commission de Régulation de l'Énergie* (CRE) in France. At the European level NRAs work together in the Agency for the Cooperation of Energy Regulators (ACER).

The action of ACER and NRAs on electricity and gas markets is complementary to more global considerations. The actors of market surveillance are European Securities and Market Authorities (ESMA) at the EU level and the *Autorité des Marchés Financiers* (AMF) at the French level. Aspects related to competition are not under the responsibility of a single European Agency, but national agency such as the *Autorité de la concurrence*, cooperates through the European Competition Network (ECN) [86].

The functioning of markets as prescribed by the EU network codes and guidelines are explained in detail in [87].

Market participants have to comply with multiple regulations: REMIT (Regulation of the Wholesale Energy Market Integrity and Transparency EU 1227/2011), EMIR (European Market Infrastructure Regulation EU 648/2012), MAR (Market Abuse Regulation EU 596 /2014), MIFID II (Markets in Financial Instruments Directive II 2014/65/EU)... [88]. The details of the regulatory requirements are complex [89] and out of the scope of this thesis. Nevertheless, considering the implications of REMIT, which is dedicated to energy markets and under the authority of ACER gives interesting insights to understand the actual functioning of markets and some associated risks. ACER publishes and updates guidance on REMIT and specify the forbidden behaviors [90]. Sanctions are imposed in case of market manipulation, insider trading or when obligations are not met; the list of enforced decisions is also maintained and made public [91].

2.3.3 Market participants

The full list of market participants published by ACER [92] consists of around 15000 entries. As of 2019, 302 market participants are active on EPEX Spot [93]. These participants have different profiles, as shown in Figure 2.7.

Not all participants are interested in the physical procurement of electricity. In theory, pure speculation is not necessarily detrimental to the functioning of the market and can even play a useful role. Speculators also play an active role in the market by taking risky bets on short- and mid- term price evolutions without being fundamentally interested in

⁹The RSCs are: Coreso, Transmission System Operator Security Cooperation, Security Coordination Centre, Nordic RSC, Baltics RSC and Southeast Electricity Network Coordination Center

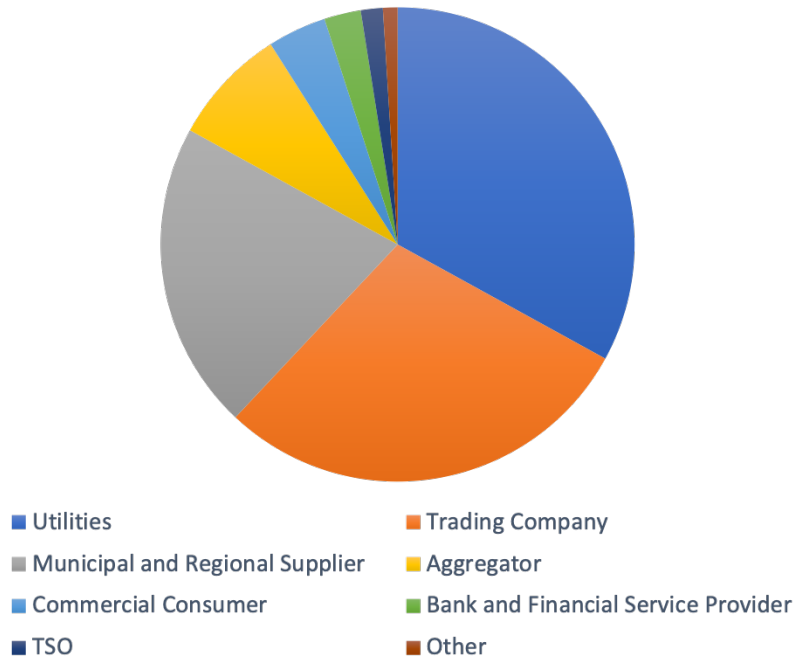


Figure 2.7: Market participants on EPEX Spot by category [94]

the underlying commodity. Doing so, they improve market liquidity, contribute to price discovery and assume risks that other market participants do not wish to take. [95]

With the development of new markets and the increase in temporal granularity, trading becomes more complex. It is now at least partially automated: 47% of traded volumes and 65% of the number of trades are executed through application programming interface (API) on EPEX Spot. The share of automated trading, including high-frequency trading, is rising. This also makes possible to apply machine learning techniques on electricity markets. EPEX Spot presents the future of power trading as follows: "For years software specialists have been working on the next big step in the field, which is the involvement of Artificial Intelligence. This opens the possibility to program software to achieve not only a certain goal, such as profit maximisation but to learn how to achieve this goal. The software acquires knowledge through training and trial-and-error explorations, it maps complex inputs and is eventually able to plan ahead, to predict and execute the optimal actions needed to achieve the goals that have been set" [93]. The fact that markets can formally be seen as a game (with an objective, a defined set of rules, participants, constraints) is an interesting feature for machine learning. These data-driven methods have proven to achieve better results than humans at certain tasks and they are improving fast [96, 97]. One issue is that such tools can come with unexpected strategies that are not transparent [98]. These strategies could have impacts on traded volumes and price formation, including an increased risk of tacit algorithmic collusion [99].

2.4 Conclusions

The energy transition is a necessity that must happen at a global scale and at an accelerating rate. The European Union shows determination in promoting policies to guide this transition. A desirable path towards a sustainable future should not be exclusively based on technologies, but also consider other aspects such as energy sufficiency and social equity. Nevertheless, technology will play a key role, especially low-carbon electricity. The power system will have to significantly evolve in the coming decades under the impulsion drivers such as renewables, batteries, hydrogen and demand-side management. The current use of fossil fuel, notably for heating, cooling, and mobility applications, will have to be replaced; sector coupling will facilitate the decarbonization. Flexibility solutions are also set to play an increasingly important role to reduce carbon intensity and the power grid will be essential for the integration of these solutions. Different paths towards net zero for future of power systems have been proposed at worldwide, continental and national scales; we emphasized how these prospective scenarios converge with regards to the strategic place given to electrification and renewables.

The pan-European energy market is developed with the intent of facilitating this transition by structuring energy and financial flows between market participants. The electricity covers multiple geographical areas and time frames, which enable market participants to manage their risks and generate revenues when opportunities arise. Market coordination is conducted by supra-national institutions at the political and regulatory levels as well as by the collaborative works of transmission system operators and nominated market operators that we presented in this chapter. The profile of market participants and the evolution of their trading preferences have been examined. We observed that not all traders are asset-backed and we discussed the growing share of algo trading in electricity markets. We noted that electricity markets must be carefully designed to ensure that they are efficient in the short- to mid- term for both consumers and producers, while also promoting adequate investments for a greener futures. The actual decoupling between economic growth and environmental damages will have to be closely monitored and the measures taken should be adapted accordingly.

Chapter 3

Price formation on day-ahead electricity markets

Contents

3.1	Operation of day-ahead markets	34
3.1.1	Setting prices that maximize social welfare	34
3.1.2	Impact of marginal costs on the merit order	36
3.1.3	Types of market orders	38
3.2	Interconnection of day-ahead markets	39
3.2.1	Bidding zones	39
3.2.2	Trading between market zones	41
3.2.3	Liquidity and price convergence in Europe	42
3.3	Stylized facts of electricity prices	45
3.3.1	Predictability of prices on day-ahead markets	45
3.3.2	Time series analysis of prices	45
3.3.3	Analyzing the influence of price drivers	48
3.3.4	Statistical modeling of prices	50
3.4	Conclusions	54

In the previous chapter, we presented the current state and the expected evolution of the European power sector and wholesale electricity markets. Building on this knowledge, we will now study more specifically the day-ahead electricity market: how is it structured and what are the implications for the price formation mechanism. After having considered the rules and theoretical functioning of the European day-ahead market, we will present the main features of prices that are empirically observed, which reflect the formation mechanism as well as the trading constraints and strategies of market participants. Through analyses of historical market prices and modeling of the influence of price-drivers with purely data-driven models, we will reveal that markets are predictable to some extent. This understanding of the day-ahead market will lead us to the next chapter of this thesis on the modeling of electricity prices thanks to a data-driven structural approach.

3.1 Operation of day-ahead markets

This section describes the mechanisms enabling the price formation on the day-ahead market across Europe and the market conditions needed for an efficient price formation.

3.1.1 Setting prices that maximize social welfare

Market clearing

The supply demand equilibrium is found through a closed auction, also called blind auction, where market participants submit orders without knowing the bids made by other exchange members. A basic order to sell or to buy consists of a volume and a price for a given time and location. In Europe, these locations correspond to bidding zones and this market design, called zonal, is presented in detail in subsection 3.2.1. In addition to basic orders, complex orders containing additional constraints can also be submitted by market participants as described in subsection 3.1.3. All orders are sent to a nominated electricity market operator (NEMO) one day before physical delivery and before a deadline called gate closure time (GCT). Then, EUPHEMIA[100], a single price coupling algorithm, determines which orders are executed and which ones are rejected. Based on the results of the computation, a market clearing price is defined for each market time unit, which is currently equal to one hour.

Uniform pricing

All accepted orders are traded at the market clearing price, regardless of the submitted order price¹. This market design is called uniform pricing (or marginal pricing²), as

¹Italy is divided in multiple bidding zones and has a slightly different approach. The supply side receives money according to the market clearing price in their zone, but all Italian consumers pay the same price (i.e. the PUN, “Prezzo Unico Nazionale”) based on the weighted average of supply prices in all Italian bidding zones.

²Commission regulation (EU) 2015/1222 (CACM), Article 38

opposed to pay-as-bid.

Social welfare definition

EUPHEMIA is designed to maximize social welfare, in other words, to maximize the economic surplus which is the sum of the consumer surplus, the producer surplus, and the congestion income. The producer and consumer surplus correspond to the difference between the money market participants were ready to pay or receive based on their orders, and the actual market clearing price. This means that all executed sell orders have an order price below the market clearing price. Reciprocally, all executed buy orders have an order price above the market clearing price. The congestion income is a function of energy transfers between bidding areas and accounts for the tariff rates through interconnectors. The congestion income calculation depends on the methodology adopted to define transfer between bidding areas and if tariffs are implemented. To give an intuition, a congestion income approximate to the difference between the money paid by supplier at one end of interconnectors and the money received by the consumers at the other end.

Merit order

The social welfare maximization objective means that bids are executed according to the merit order. In the simplified case, for each period and bidding zone, the orders proposing to sell at the lowest price and to buy at the highest price are accepted and the total volume of energy exchanged is maximized. In practice, the acceptance of orders also depends on technical constraints related to the interconnection between bidding zones and on the existence of order types that reflect additional constraints requested by market participants regarding the execution of their orders to ensure that they are feasible and profitable. The merit order is typically represented by the supply curve that is constructed by sorting all hourly orders of a bidding zone by ascending bidding price. Reciprocally, the aggregated demand curve is built using the demand orders sorted by descending price. Although the market participants are initially unaware of the bids made by others in order to reduce the risk of collusion, the anonymized supply and demand curves are published after market clearing so that these data can latter be used to conduct market analyses and to develop models of price forecasts. The market clearing resulting from the intersection of the supply and the demand curves is schematically represented in Figure 3.1.

Shorter market time unit

The electricity flows in power systems are varying at a faster pace, due to fluctuations on the supply side (weather-driven electricity production from renewable energy sources) and on the demand side (flexibility provided by industrial and residential consumers). A well-designed market has to account for and reflect these new dynamics. The Capacity Allocation and Congestion Management regulation (CACM) therefore imposes to reduce

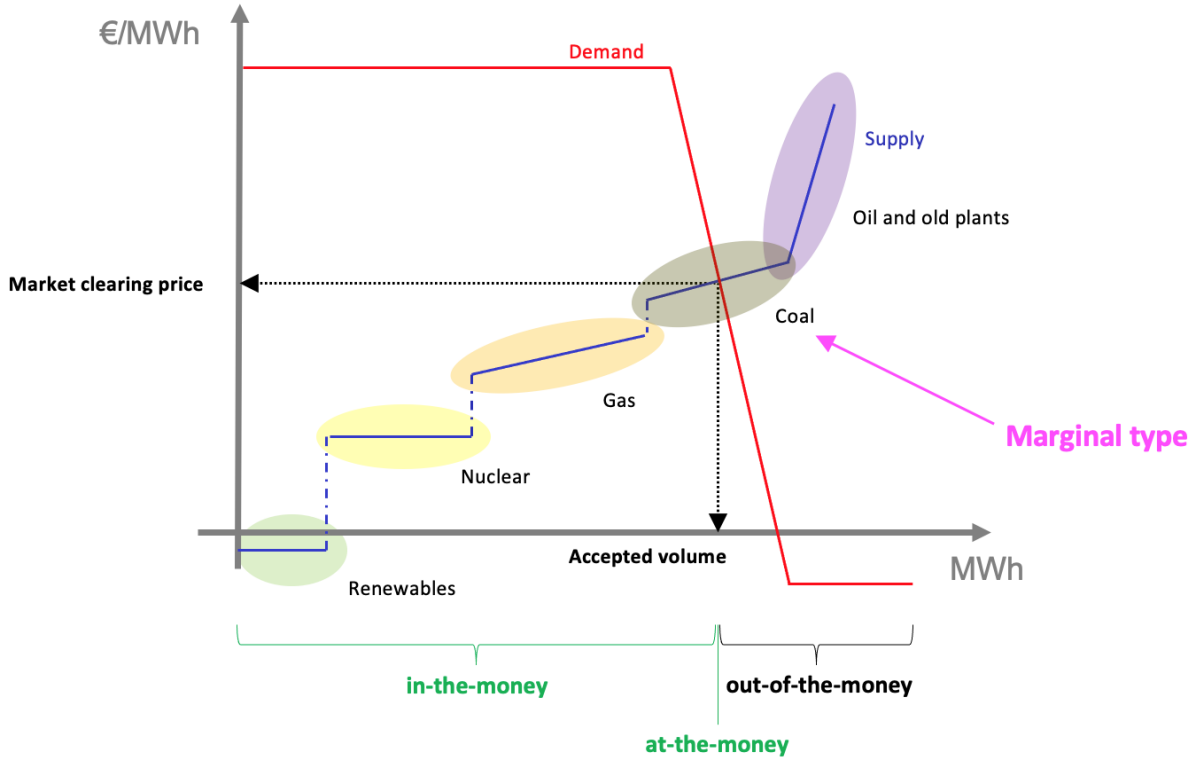


Figure 3.1: Simplified representation of the market clearing, based on [101]

the market time unit from one hour to fifteen minutes. Shorter time frames enable to derive benefit from flexible assets that are capable to quickly respond to unexpected events [102]. The all NEMOs COMMITTEE, with the support of ENTSO-E, is working on the development of a new scalable version of EUPHEMIA, which would allow for this change to be implemented.

3.1.2 Impact of marginal costs on the merit order

Bids based on marginal cost

As illustrated on Figure 3.1, supply orders are typically assumed to correspond to production from power system assets. The fact that financial transactions on day-ahead markets correspond to real electricity transfers in the grid increase the relevance of this assumption³. Nevertheless, a bid in zonal markets does not necessarily correspond to a specific power plant: for example, power producers can make use of the synergies within a portfolio (e.g. virtual power plants), and retailers can resell electricity previously bought in surplus on forward and future markets.

³such transactions for which the commodity is actually traded are said to be "physically settled" as opposed to financial settlements allowed on other markets

Evolution of the merit order

In a day-ahead market with perfect competition, the rational strategy for market participants is to bid according to their short-run marginal cost of production for suppliers, and to bid based on marginal utility for consumers. While all criteria for an idealized perfect market competition are never met in practice, the marginal cost remains the main factor to consider when thinking about electricity price formation.

First of all, it allows us to understand the influence of CO₂ price on the power system. As polluting power plants become more expensive to operate due to higher CO₂ prices, the merit order can change: even if the actual extraction and transport cost of coal is cheaper than the costs for natural gas, a sufficiently high CO₂ price will place natural gas before coal in the merit order, as coal-based production is more emissive.

Moreover, power plants running on fossil fuels can be pushed out of the merit order by renewable energy sources, regardless of CO₂ prices. Since renewables have low marginal costs, they are typically the first technologies in the merit order. This first position results in a shift of the supply curve to the right, which can make fossil fuels unnecessary to meet the demand during hours when the share of renewables is significant.

In addition, subsidies of renewables can also impact price formation on day-ahead markets, especially in the presence of feed-in tariff support schemes. When fixed subsidies are given for the supply electricity regardless of market conditions, a profitable bidding strategy is to auction orders at a price equal to the difference between the marginal cost and the fixed subsidy. In some cases, this difference can even be negative and lead to market prices below EUR 0 per MWh. Such subsidies paid regardless of electricity demand could be seen as a form of market distortion (curtailment would sometimes be more economically suitable under optimal operation). Feed-in-premium are therefore becoming the preferred solution over feed-in-tariff in Europe. The subsidy corresponding to feed-in-premium is variable and its amount depends on market conditions, thus encouraging renewable energy producers to better account for price signals.

Uniform pricing with a high share of renewables

As mentioned above, renewable energy sources are characterised by very low operational costs and, as their share significantly increases, question arise regarding the robustness of the current market design.

If power plants using fossil fuels are rarely in-the-money due to high operation costs, they might be unprofitable and become stranded assets. While the phasing-out of fossil fuel is a goal of the energy transition, the power system in its current state rely on thermal generation to ensure security of supply. The need for highly emissive dispatchable assets should be reduced in the future by the deployment of storage and demand-side flexibility, but this substitution is a long process. The financial issue faced by investors in peak generation assets when they fail to generate revenues that will cover both the capital expenditures (CAPEX) and the operational expenditures (OPEX) is called the "missing money problem". It stems from the fact that in energy-only markets, such as the day-ahead market, producers are paid only for the electricity they supply and not for the

power capacity they offer to the grid. With energy-only market design, peak generators need to generate significant revenues during price spikes, which occur occasionally during scarcity situations. Since the frequency of occurrence and the price amplitude of these spikes are hard to predict and since no guarantee is given that they will actually cover the investment costs, energy-only markets risk to not provide sufficient incentives for investors in peak generation, thus potentially jeopardizing the long-term security of supply. One proposed solution to ensure security of supply has been to develop capacity markets in order to generate additional revenues for dispatchable power plants.

In addition, wholesale market prices not set by thermal power plants might even become a problem for the renewable energy producers. When RES are dominant and in absence of collusion, the day-ahead market price based on the marginal cost of renewables might be enough to cover the low OPEX but insufficient to generate enough profits with regard to their high CAPEX. The fact that, in theory, the more renewables the lower the day-ahead price is called the "cannibalisation effect". In addition to subsidies, a market-based solution has been proposed to ensure additional revenues for renewable energy producers: the guarantees of origins (GOs). These green certificates prove that electricity is generated from low-carbon sources and can be sold to B2B customers, which want to prove to end-consumers or regulators that efforts have been made with regards to the energy transition.

Conversely, when gas or coal remain oftentimes marginal although most electricity is generated from low OPEX sources (renewables and/or nuclear), the producer surplus can be perceived as too big and socially unacceptable. Consumers do not easily understand nor accept that they have to pay high electricity prices even if most electricity is relatively cheap to produce, especially since taxes helped to finance the development of these baseload energy sources. The apparent paradox of the price formation has been criticized both by citizens and by politicians during the post-covid demand growth when gas prices surged, notably in Spain due to the high share of renewables in the energy mix and in France due to the dominance of nuclear power. Temporary remedies have been proposed to manage the effects of this energy crisis while maintaining the benefits of the European internal energy market (resource pooling, price signals, ...). For example, the French government decided to provide stimulus checks to low-income citizens.

3.1.3 Types of market orders

Orders to sell and buy electricity on the day-ahead market must be in a defined price range. The minimal price is called "price floor" and the maximum value is called "price cap". These values are set to facilitate the daily resolution of the EUPHEMIA algorithm in limited time. Nevertheless, the price floor and cap must not impose constraints on price formation. As of 2021, the floor is set to EUR -500/MWh and the cap is set to EUR 3000/MWh. Regulatory mechanisms are in place to reevaluate these thresholds should they be reached too often and thus imposing a limiting factor on price formation.

The simplest market order consist in indicating a sense (supply or demand), a volume, a price limit, a market time unit and a bidding zone. In addition, power exchanges offer the possibility to submit complex orders and block orders.

For a complex order, a market participant indicates additional conditions that have to be fulfilled and the hourly orders become interdependent. Two main requirements can be imposed: a minimum income condition (MIC) and a load gradient condition. The public description of the EUPHEMIA algorithm [100] describes these requirements with precision:

"Generally speaking, the Minimum Income economical constraint means that the amount of money collected by the order in all periods must cover its production costs, which is defined by a fix term (representing the startup cost of a power plant) and a variable term multiplied by the total assigned energy (representing the operation cost per MWh of a power plant). [...]"

Generally speaking, the Load Gradient constraint means that the amount of energy that is matched by the hourly sub-orders belonging to a Load Gradient order in one period is limited by the amount of energy that was matched by the hourly sub-orders in the previous period. There is a maximum increment / decrement allowed (the same value for all periods)."

Block order is another type that offers the possibility to add a coupling between hours and to set minimum acceptance ratios. For such orders, the principle is similar in spirit to the minimum income condition, but instead of imposing a condition on the total daily revenue, the condition is on volume, where a minimal amount has to be reached for the order to remain valid.

3.2 Interconnection of day-ahead markets

3.2.1 Bidding zones

Zonal market

Electricity markets can be organized based on one of the two main designs: zonal and nodal [103].

Zonal pricing, which has been adopted in Europe, sets a unique price for a large area called a bidding zone (typically a country or a region).

Nodal pricing, also known as locational marginal pricing (LMP), is common in North America. These markets feature a higher spatial granularity since a price is set for each bus. As a consequence, transmission constraints are better taken into account and play a greater role in price formation. Also, in some cases, the market power of a few electricity producers over a single node can be more significant than it would be over a larger zone with more competing resources; this market power could impact price formation.

The advantages and disadvantages of zonal and nodal models of energy markets are discussed more extensively in [104].

Delineation of bidding zones

When market participants buy or sell electricity in a bidding zone, they supply or consume electricity at the location of their choice within an area specified in the submitted

bid. With regards to the day-ahead market, technical constraints imposed by the electrical grid within a zone are not considered. Bidding zones mainly correspond to a whole country as shown in Figure 3.2. In some cases, bidding zones can span multiple countries (e.g. Germany and Luxembourg are in the same bidding zone in order to ensure sufficient liquidity) or, on the contrary, a country can be split into multiple zones (e.g. in Sweden to reflect the supply, demand and network heterogeneity within the country).

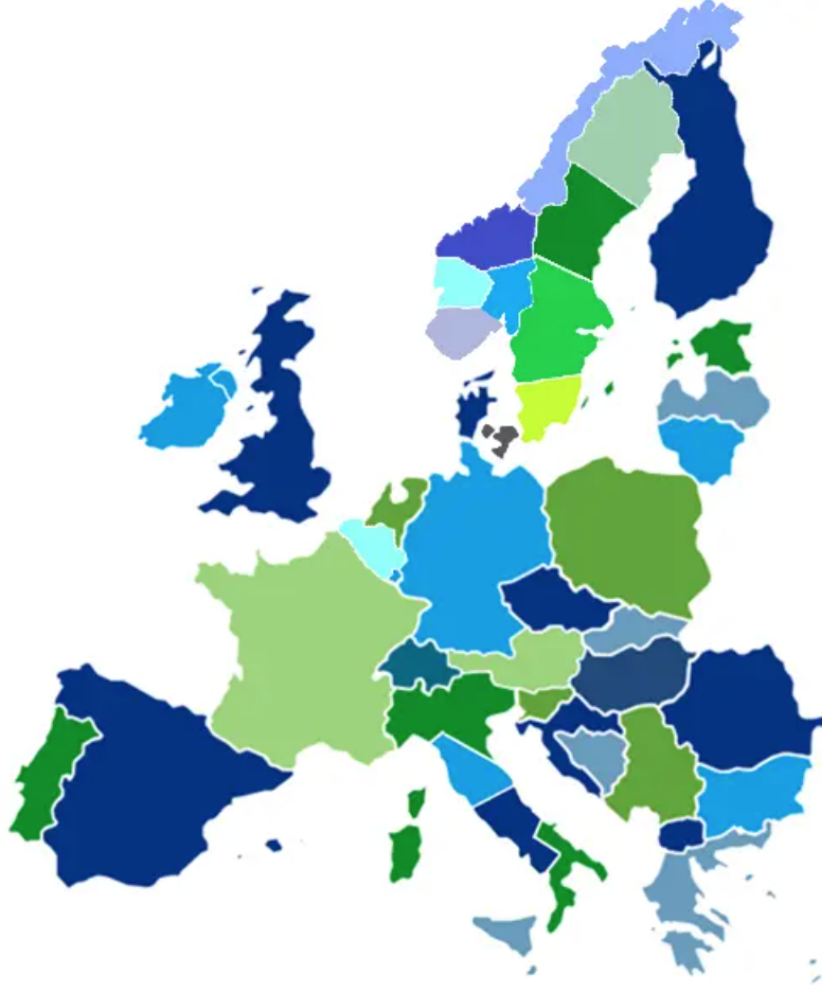


Figure 3.2: The bidding zone configuration in Europe as of September 2020 [105]

Interconnection of bidding zones

Bidding zones are interdependent as they are physically interconnected by lines and cables. The interconnection network is expanding across Europe to facilitate resource pooling, which benefits both the integration of low-carbon energy sources and the security of supply. The EUPHEMIA algorithm is solved in two-steps in order to account for inter-zone trading. First, the market clearing is solved for each bidding zone independently

(doing so also has the advantage of making the process more robust as it allows to apply fallback procedures when the coupling across Europe fail due to technical issues). Second, the use of interconnectors is optimized with regards to economic benefits while accounting for power grid constraints.

Congestion management in zonal markets

Internal network congestions within a bidding zone are problematic for the considered zone due to the incapacity to deliver electricity as traded. Such congestions are also problematic for the European system as a whole (e.g. wind production exported from Denmark to Poland might be hindered by unexpected congestions within the German power grid).

Even without considering additional investments in interconnection capacity, internal congestion can be minimized by changing the configuration of bidding zones to account for actual grid constraints. Nevertheless, defining new bidding zones that would be less aligned with political borders is a sensitive topic, for example regarding the proposal to split Germany into multiple zones.

Alternatively, Europe could adopt a nodal market design as implemented in other countries. Since the spatial granularity is much higher in nodal markets, congestion management is intrinsically better handled by the market itself. Nodal and zonal markets both have comparative advantages and drawbacks, but regardless of the benefits that such design change could provide for Europe, a switch is almost impossible before the end of this decade since it would imply considerable adaptations in term of regulation, governance, procedures and IT systems. Besides the current development of pilot projects for local markets that have been discussed in the previous chapter, the main cross-border relevant remedial actions to handle network congestion are redispatching and countertrading [106]. As per European regulation [107]: "countertrading" means a cross zonal exchange initiated by system operators between two bidding zones to relieve physical congestion [and] 'redispatching' means a measure activated by one or several system operators by altering the generation and/or load pattern in order to change physical flows in the transmission system and relieve a physical congestion."

3.2.2 Trading between market zones

Margin available for cross-zonal trade (MACZT)

The installed transfer capacity of an interconnector is a static value, but its available capacity is dynamic as it depends on physical phenomenon, which vary according to the state of the power system (Kirchhoff's laws). The precise assessment of this available capacity is a complex task for transmission system operators since this evaluation depends on dedicated network models, demand and consumption forecasts, as well as security margin. Such margin are necessary because of unavoidable forecast deviations, which are due to uncertainty and contingencies (e.g. sudden outages). Nevertheless, as explained earlier, the network capacity allocation should not be too conservative in order

to allow for an efficient use of resources within the internal energy market. This is why the European legislation imposes that at least 70% of the transmission capacity must be offered for cross-zonal trade by 2025.

Interconnection capacity has to be acquired through explicit auction, in which market participants have to actively procure this capacity in addition to trading energy. Alternatively, the allocation can be implicit when the multizone market clearing jointly allocate energy and transmission capacity to market participants. The allocation method vary depending on the considered border and time frame. The second option, which ensures a better overall usage of the available capacity, is a key component of market coupling. The generalization of market coupling is a central element for the development of the internal electricity market for the day-ahead time frame.

Cross-border capacity allocation methods

Two main approaches are used to calculate the cross-zonal capacity for the day-ahead market time frame: available transfer capacity (ATC) and flow-based market coupling (FBMC). A summary of the methods presented in detail in [108] is proposed thereafter.

The Available Transfer Capacity (ATC) method is performed before the trading of electricity by market participants in order to evaluate the maximum cross-zonal commercial exchange that can be allowed. A bidding zone is represented by a single node and the ATC is computed for each link between nodes. ATC computations are performed independently for each link and the network constraints within a zone are not considered. These two simplifications do not allow for a precise representation of the grid constraints and force to adopt overly conservative values in order to ensure secure grid operation.

The Flow-Based Market Coupling (FBMC) method allows to overcome some of the drawbacks of the ATC method. Bidding zones are modeled more precisely: the main internal grid constraints are accounted for. Moreover, the joint calculation of capacity for multiple links gives a more accurate representation of the capacity that is actually available. FBMC is performed at two stages. The first stage takes place before market clearing, as for ATC, to inform the flow domain of possible use of cross-zonal transmissions lines and cables. The second stage occurs during market clearing as the optimization algorithm determines simultaneously the execution of bids and the implicit allocation of network capacity for market participants (hence the name "market coupling").

3.2.3 Liquidity and price convergence in Europe

A well-functioning market needs to be liquid, which means that electricity can be purchased or sold quickly without significantly impacting the price. High trading volumes is a condition for liquidity as it indicates that buyers and sellers are interested in transactions through day-ahead markets. In 2019, 1489 TWh have been traded on day-ahead markets in total in Europe. The distribution of daily volumes for each month of 2019 is summarized in the boxplot presented in Figure 3.3. We observe variations at the seasonal and daily timescales, which are due to the evolution of electricity demand throughout the year.

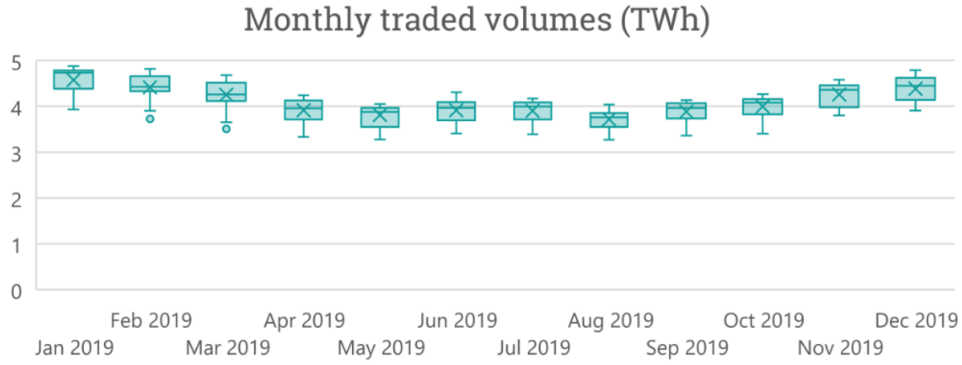


Figure 3.3: Monthly traded volumes on day-ahead markets in the EU, 2019 [109]

Market liquidity can be measured with the churn factor, which is a metric of that gives an indication of the relative ‘size’ of the market compared to its physical size. It is defined as the overall volume traded through exchanges and brokers expressed as a multiple of physical consumption. The churn-factors for the day-ahead markets by country from 2015 to 2019 is given in Figure 3.4. According to the European Union Agency of the Cooperation of Energy Regulators (ACER), the relatively low churn factors in some countries can be explained by the prevalence of alternatives to day-ahead markets to procure electricity: for example, the Regulated Access to Incumbent Nuclear Electricity (ARENH) in France or the use of Power Purchase Agreements (PPAs) to trade renewables over the counter in Nordic countries. On the opposite, countries such Ireland, Northern Ireland and Greece, for which the day-ahead market is the unique solution to trade electricity in a time frame of a day before delivery have a churn factor close to one. ACER also notes that the modest year-on-year changes in day-ahead market liquidity for most countries indicate that these markets are mature.

The actual development of the internal electricity market is not only be assessed through the lens of liquidity, but also by analyzing the evolution of price convergence between different market areas. Network expansion, the increase in commercial cross-zonal capacity and the development of market coupling lead to an improved use of resources in the internal energy market, which is mechanically reflected in wholesale prices [110]. The improvement of price convergence can be seen in Figure 3.5: for most regions, we observe an increase in the number of hours when the prices are similar in all bidding zones.

Nevertheless, price convergence depends on multiple factors, such as the generation mix and as well as the level of RES penetration, and might not be significant in all areas. The market integration should not be assessed from price convergence alone, since additional technical factors should be accounted for when analyzing these developments (e.g. internal balancing needs, as in [111]).

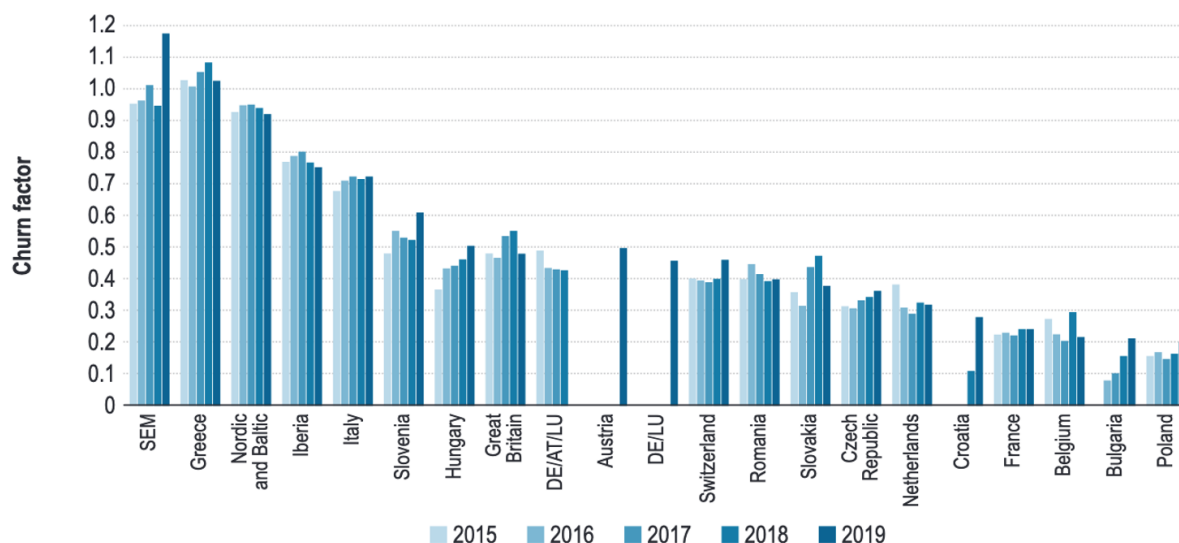


Figure 3.4: Churn factors in major European day-ahead markets, 2015-2019 (The missing data for DE/AT/LU in 2019 results from the split of this unique bidding zone in two: Germany-Luxembourg and Austria. "SEM" refers to the Irish Single Energy Market) [110]

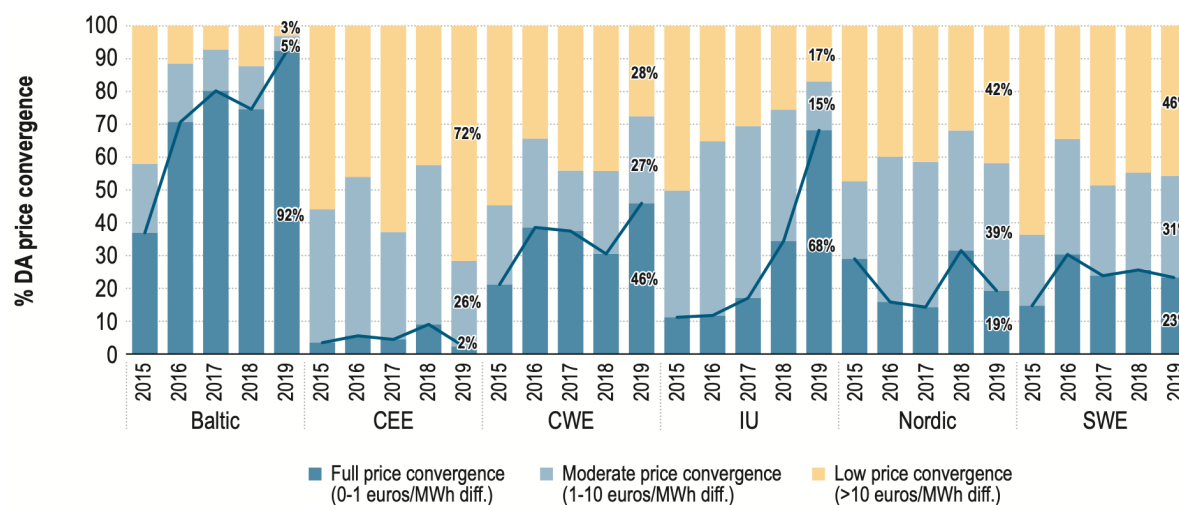


Figure 3.5: Price convergence on European day-ahead markets [110]. Central-East Europe (CEE): the Czech Republic, Hungary, Poland and Slovakia; Central-West Europe (CWE): Austria, Belgium, France, Germany/Luxembourg and the Netherlands; IU region: the Republic of Ireland and the United Kingdom; Nordic region: Denmark, Finland, Norway and Sweden; South-West Europe (SWE): France, Portugal and Spain

3.3 Stylized facts of electricity prices

3.3.1 Predictability of prices on day-ahead markets

Time series of hourly electricity prices on day-ahead markets display typical features that are well documented in the scientific literature [112, 113, 114].

Finding reliable patterns is essential as it empirically confirms that prices are not chaotic. The inherent constraints imposed by the physics of power system, the deterministic price formation mechanism of the market clearing, and the decisions of market participants based on sound economic reasoning lead to the emergence of relatively predictable wholesale prices. Nevertheless, the forecast precision is necessarily limited by the fact that wholesale prices ultimately depend on stochastic trading decisions; neither the information considered by market participants nor their decision processes can be perfectly anticipated.

Moreover, the specificity of electricity as a commodity (difficulty to store it at scale, low elasticity of the demand, grid-based delivery, etc.) and the specificity of the electricity market (oligopolistic supply-side, strategic importance of electricity for nations, etc.) imply that dedicated models should be developed in order to accurately capture these particularities.

Analyses of time series of prices reveal a dependency of a market outcome on the prices observed during other time steps. These analyses can then be completed by considering the influence of price-drivers on market outcome. These two kinds of analyses are presented at the beginning of this section. Thereafter, a last subsection introduces key concepts for modeling of electricity prices. The main steps used to model prices are presented through the examples of statistical and machine learning models, then the benefits and limits of such purely data-driven approaches are discussed.

The phenomena are illustrated thanks to data related to the French market for the four-year period spanning from 2015 to 2018. The corresponding time series of hourly prices on day-ahead market is shown in Figure 3.6.

3.3.2 Time series analysis of prices

Volatility

Electricity prices are very volatile compared to the price of other commodities, which means that they vary a lot over time. The variation of prices from one market time unit to the next is called "return". A return can be defined as an absolute or as a relative variation of prices. Volatility is statistical indicator of the dispersion of returns over a defined period. High volatility indicates that there is an interest in analysing the evolution of prices with a high time granularity and to go beyond aggregated measures (mean, median, etc.). These price variations generate risks and opportunities for market participants.

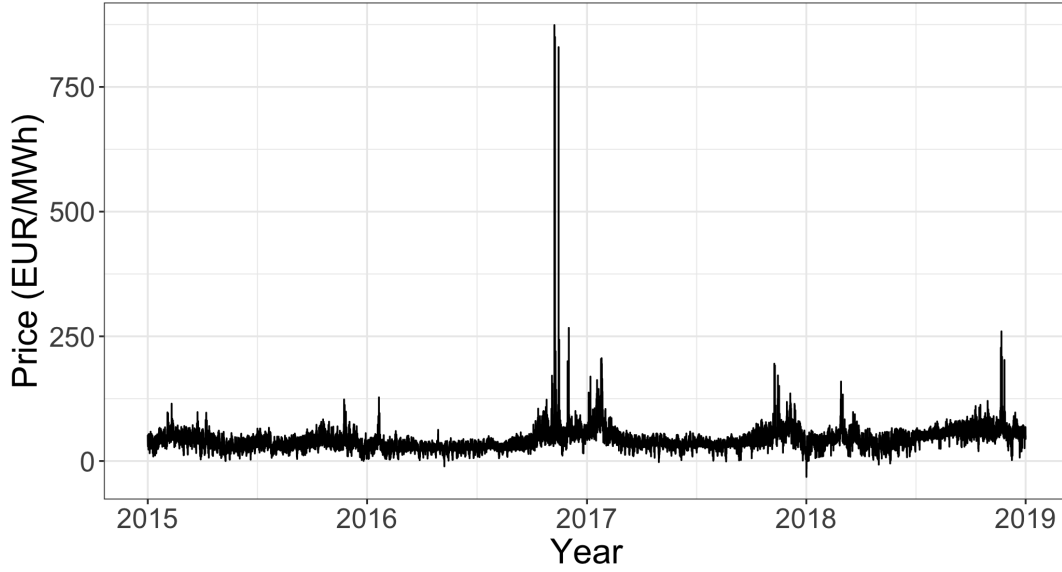


Figure 3.6: Hourly electricity prices on the French day-ahead market

Multiseasonal price patterns

Wholesale electricity prices display periodicity at three timescales: daily, weekly and yearly as shown in Figure 3.7. The patterns of prices reflect the underlying periodicity of power system conditions.

At the daily timescale, the demand variations are due to habits throughout the day (waking-up, commuting, work, cooking, sleep, etc.), the supply variations are due to weather patterns (peak solar production in midday, fast ramping up of flexible production assets around dinner, etc.). Another interesting aspect to consider at this timescale is the impact of gate closure time. Since all bids must be submitted before a single deadline for all hours of the next day, the latest demand and supply forecasts are more uncertain for the hours at the end of the day than at the beginning. Market participants might have to adjust their bidding according to the confidence they have in the forecasts and the way they manage imbalance risks.

At the weekly timescale, the variations are mainly due to the alternation between workdays and week-end. This not only impacts the consumption habits, but also some production processes which ramp-up at the beginning of the week and ramp-down before the week-end.

Finally, at the yearly time scale, the price variations can be mainly attributed to changes in weather (heating, cooling, solar and wind production, hydro stock, etc.). We also note that the availability of generation units is not constant throughout the year since maintenance of nuclear power plants are typically scheduled during summer when demand is lower.

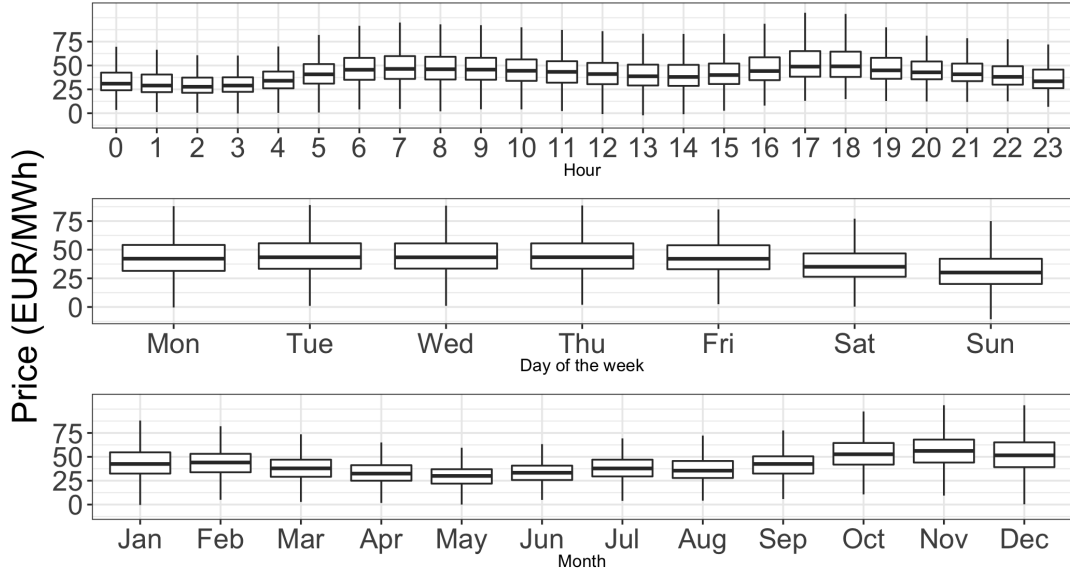


Figure 3.7: Boxplots of electricity prices on the French day-ahead market for the period 2015-2018 as a function of the hour, the day of the week, and the month

Spikes and mean-reversion

Power systems are built to ensure security of supply throughout the year. Consequently, they can appear oversized most of the time and electricity is provided by baseload units. Nevertheless, scarcity situations occur occasionally due to outages, extreme weather of events or any other kind of disruptive events. Although such situations generally do not last for more than a few hours, this new supply-demand equilibrium can still put the power system under stress. This stress is typically signaled by price spikes that significantly and abruptly raise the value of electricity. Spikes, as opposed to jumps, imply that prices quickly decrease after a short period and return to the usual range corresponding to the baseload situation. Electricity prices are therefore said to display mean-reversion.

Negative prices

We previously mentioned that negative wholesale prices can be observed as a consequence of subsidies for renewables. Another reason that can trigger negative prices is the lack of flexibility of certain power assets compared to residual demand variations. Power plants are constrained by upward and downward ramp limits. Fast and recurrent variations accelerate the aging of assets and can even cause safety issues. Starting and stopping a generation unit also comes with additional operating costs and technical constraints (duration of starting or stopping phases, mandatory refractory period between a stop and a new start). Finally, generation units are designed to operate around an optimal set point and deviation from this value range decreases conversion efficiency. For all these reasons, fast variations of power production are sometimes impossible or very

expensive. As a consequence, it is sometimes economically sensible for power producers to bid negative prices during some hours, if this ensure them to keep producing instead of being replaced by low-OPEX units. In addition, negative bids can be due to the necessity to close a trading position (i.e. too much energy has been bought on long-term markets and must be sold back).

3.3.3 Analyzing the influence of price drivers

We presented so far the main characteristics of time series of prices and proposed explanations for the causes of these features. Such affirmations can be confirmed by performing further analyses, which consist in studying the link between price-drivers and market outcomes. Here, we keep a single dependent variables Through illustrative examples, we will now highlight how simple studies can reveal the influence that key factors have on price formation. Three aspects will be considered thereafter: the aggregated demand, the nature of electricity supply, and the margin of supply capacity with regards to demand. The link between hourly prices and these variables is highlighted by regressions using the Loess procedure [115] (i.e. locally estimated scatterplot smoothing).

As demand for electricity is still largely price-insensitive and supply capacity does not vary on short time scales, there is a positive correlation between aggregated demand and market prices ($R^2 = 27.5\%$), as shown in Figure 3.8.

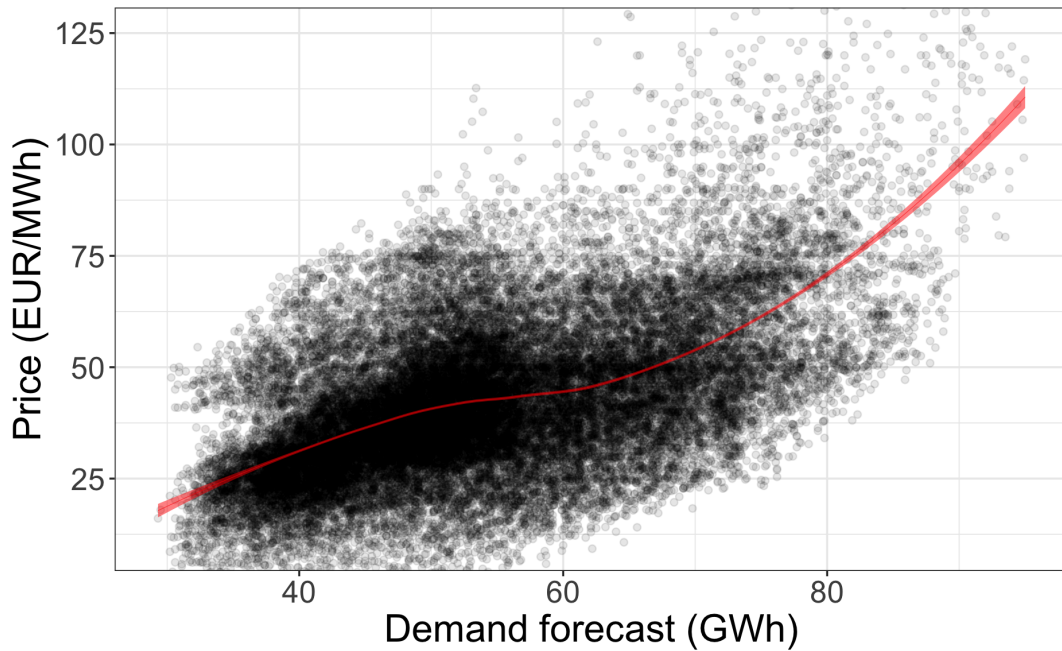


Figure 3.8: Hourly electricity prices on the French day-ahead market as a function of the forecasted demand for the period 2015-2018. Prices are represented by black dots; the Loess regression with a 95% confidence interval is represented by a red line.

As explained in the previous section, the type of resources mobilized to produce elec-

tricity influence market prices through the merit order effect. The negative correlation between the share of weather-driven renewable electricity (vRES) production and market prices is illustrated in Figure 3.9 ($R^2 = 18.7\%$). For this analysis, we could have chose to use the forecasted production of vRES instead of the actual production values, but we chose the latter due to data quality issues with the forecasts for the studied period.

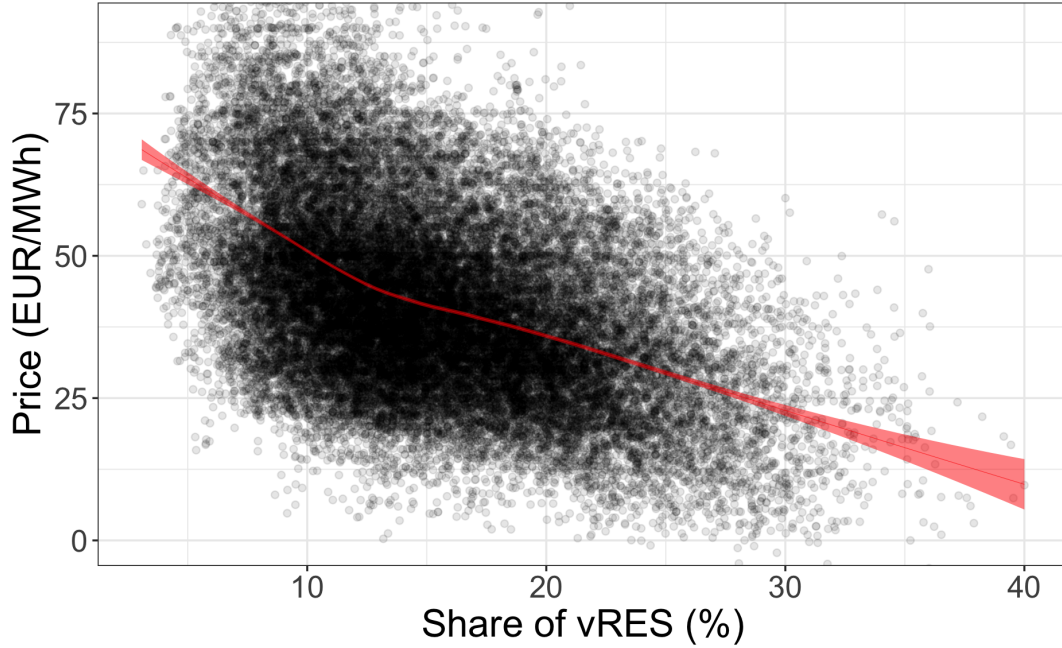


Figure 3.9: Hourly electricity prices on the French day-ahead market as a function of the share of demand covered by variable renewable energy (vRES) for the period 2015-2018. Prices are represented by black dots; the Loess regression with a 95% confidence interval is represented by a red line.

Another way to assess the impact of power system constraints on wholesale prices is to consider the scarcity of supply capacity with regards to demand. To do so, we can build a proxy variable called the "supply margin", which is equal to the difference between the available capacity of dispatchable generation units in a bidding zone and the aggregated residual demand⁴. The link between supply margin and electricity prices is displayed and assessed in Figure 3.10 ($R^2 = 45.1\%$).

The presented analyses can be expanded by enriching the datasets with additional factors, which relative influence can also be jointly analyzed. Such analyses are informative of the link between the state of power system and market prices, and as such they help to identity price drivers that can be used for modeling endeavors. Models are complementary to analyses as they allow us not only to study historical data, but also to generate new time series of prices that can serve various end goals (counterfactual analyses, forecasts, etc.).

⁴the residual demand is defined here as the difference between forecasted consumption and the contribution of weather-driven supply

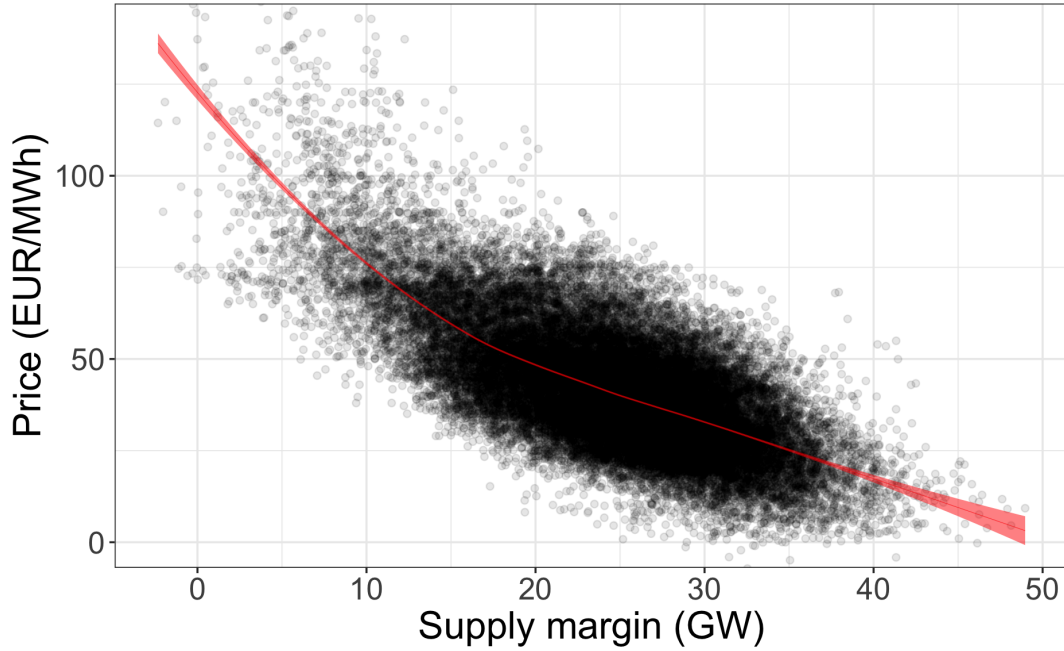


Figure 3.10: Hourly electricity prices on the French day-ahead market as a function of the supply margin for the period 2015-2018. Prices are represented by black dots; the Loess regression with a 95% confidence interval is represented by a red line.

3.3.4 Statistical modeling of prices

Modeling prices: key principles

The whole dataset (electricity prices and their drivers) is typically used as a single part to conduct analyses such as the ones that we presented. However, for modeling studies, such dataset is typically split in at least two distinct parts. The first part is used for the so-called "training phase" during which the model parameters are estimated through a regression between the outcome variable and explanatory variables (which respectively correspond to electricity prices and price drivers here). Using these estimated parameters and the price drivers from the second part of the dataset, the model can then be used to simulate prices during the "testing phase". The quality of the simulation can be assessed against the observed prices thanks to visualization (plotting both the observation and simulation) and the use of statistical metrics. This practice contributes to avoid overfitting and inform us about the predictive value of a model.

To conduct this modeling study, we must now define a train-test split, select explanatory variables, choose statistical models and propose metrics to evaluate the simulation.

We propose to assess the forecasting performance of the models on the last year of the dataset (i.e. 2018) and to use the three remaining years for training (i.e. from 2015 to 2017). For the explanatory variables, we keep the price drivers presented in the previous analyses: forecasted demand, share of vRES and supply margin.

Concerning the models, we propose to compare one of the simplest statistical ap-

proach used as a naive reference to a supervised machine learning method that is widely used for forecasting in various fields⁵. To do so, we will perform multiple linear regressions and random forest regressions, using the implementation in R programming language (respectively the function "lm" from the package "stats" v.3.6.1 and the function "rf" from the package "randomForest" v4.6). The goal is to introduce these models and to illustrate the performance of such approaches on simplified cases. These two models will then be applied again as a benchmark in order to evaluate the performance of our proposed model. The linear model is simply defined in 3.1 where $\hat{\Pi}_t$, α_i^{lm} , x_i , respectively represent the modeled electricity prices, regression coefficients and regressors. Random forests model consist in multiple decision trees which cannot be simply represented in a compact equation; the details are presented in [116], where the model is described as follows: "Random forests are a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. The generalization error for forests converges a.s. to a limit as the number of trees in the forest becomes large. The generalization error of a forest of tree classifiers depends on the strength of the individual trees in the forest and the correlation between them. [...]. Internal estimates monitor error, strength, and correlation and these are used to show the response to increasing the number of features used in the splitting. Internal estimates are also used to measure variable importance. These ideas are also applicable to regression."

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (\Pi_t - \hat{\Pi}_t)^2} \quad (3.1)$$

In terms of metrics used to verify the quality of simulations on out-of-sample tests, we primarily consider the root mean square error ($RMSE$), which assesses the global forecast accuracy. In addition, to estimate whether the model has accurately captured the price dynamics (i.e. the hourly variations), we evaluate the difference between the standard deviation of the observed time series of prices and the simulated one (Δsd). These metrics, which evaluate the performance of the forecast $(\hat{\Pi}_t)_{t \in T}$ against the observed prices $(\Pi_t)_{t \in T}$ are defined as follows:

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (\Pi_t - \hat{\Pi}_t)^2} \quad (3.2)$$

$$\Delta sd = sd((\Pi_t)_{t \in T}) - sd((\hat{\Pi}_t)_{t \in T}) \quad (3.3)$$

Simulation results

The results of the simulation for both models are presented in Figure 3.11 and Figure 3.12. We observe than in both cases, prices are underestimated for the second part of

⁵By selecting these two examples, we can evaluate the benefits and drawbacks of adding model complexity while using the same regressors.

the year. This is likely due to the impact of the increase in gas and CO₂ prices, which are not considered among the input variables. This highlights the importance of carefully selecting price drivers, so additional factors will be considered in the remaining of the thesis. Due to the limited number of variables chosen for this introductory example, the forecasting performances are quite similar for both models, as even a simple linear model succeeds in extracting information from the forecasted demand, the share of vRES and then supply margin. We obtain the same $RMSE$ in both cases equal to 19.8. Nevertheless, the machine learning better captures the dynamics than the linear model ($\Delta sd = 3.0$ for the linear model and $\Delta sd = 2.1$). We observe that the random forests algorithm is especially better at capturing price spikes.

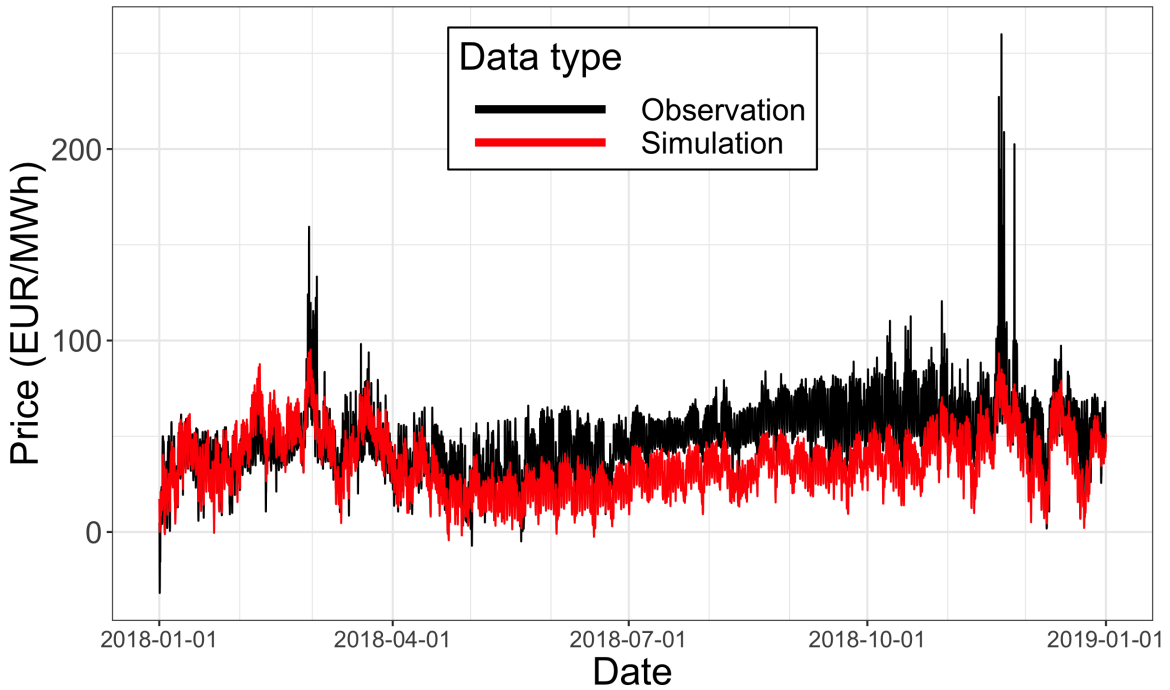


Figure 3.11: Simulation of prices on the French day-ahead market: linear model

We stress the importance to study the results on out-of-sample datasets, since in-sample predictions can result from an overfit. In this example, we obtained significantly better performances during training compared to the testing phase, especially with random forest (for the linear model: $RMSE = 13.6$ and $\Delta sd = 5.3$, for random forest: $RMSE = 6.3$ and $\Delta sd = 3.2$).

We also studied two variations of these models; the metrics obtained out-of-sample are discussed here. First, to better account for the specificities of prices for each hour of the day and day of the week, we split the global dataset into 168 independent parts (corresponding to 24×7 hours) in order to train 168 independent models. In practice, we start by only keeping the hours corresponding to Monday at 1 a.m. to train and test a model, then we repeat the process of the 167 remaining subsets. The results obtained

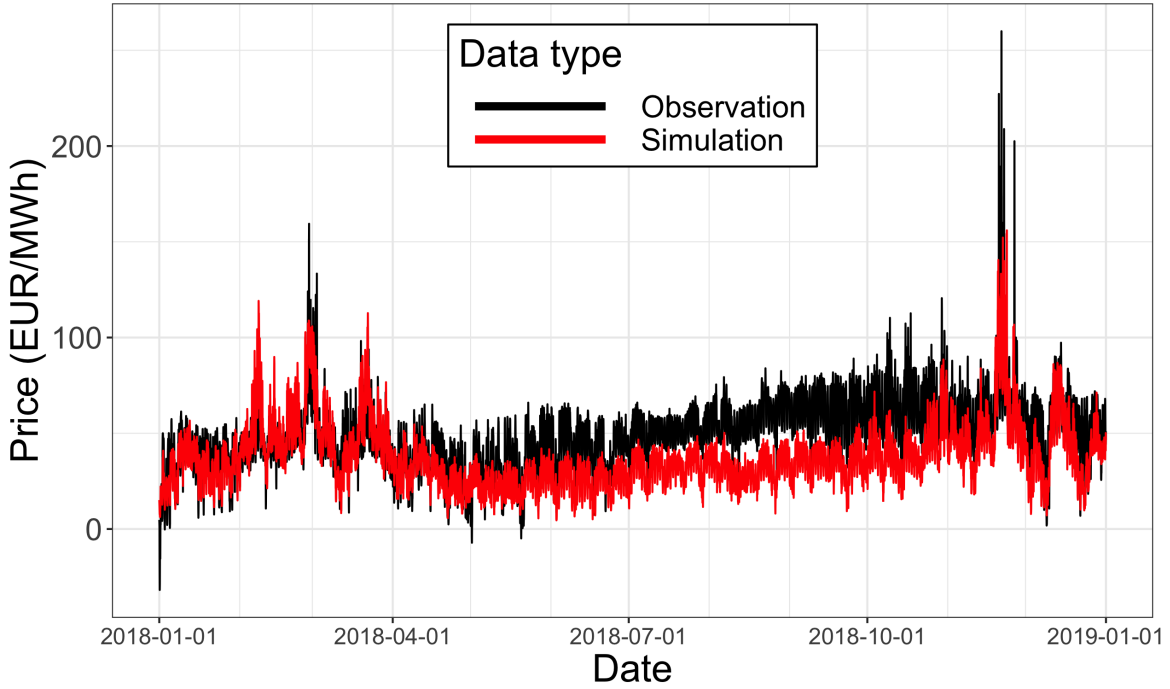


Figure 3.12: Simulation of prices on the French day-ahead market: random forest

with each model are concatenated in order to compute metrics with this reconstructed testing dataset. Doing this splitting method implies that data are more homogeneous for each model, but each training dataset is smaller. Overall, we observe a degradation of the performance compared to the default case, which seems to indicate that the size of the training dataset plays an important role (for the linear model: $RMSE = 18.8$ and $\Delta sd = 2.2$, for random forest: $RMSE = 18.8$ and $\Delta sd = 3.4$). Second, we evaluated the effect of accounting for additional data by including day-ahead prices of the neighboring countries to the set of explanatory variables. Since these prices are well correlated to the French prices, we observed significant performance improvements due to the ability of statistical models to capture these correlations (for the linear model: $RMSE = 5.6$ and $\Delta sd = 0.1$, for random forest: $RMSE = 4.3$ and $\Delta sd = 0.4$).

Remarks

As shown in this subsection, even the simplest statistical model and a general purpose machine learning algorithm are able to extract information from price drivers and can be useful to predict electricity prices. Alternative statistical and machine learning models, such as vector auto regression or neural networks, can also be applied and finely tuned, but such approaches would still not allow to account for the marginal pricing mechanism of day-ahead markets that reflects the supply-demand equilibrium. As this feature is important for prospective studies and econometric analyses, we propose in the next

chapter to explore the structural modeling methodology used to study price formation with a bottom-up approach. Accounting for the main features of the market mechanism through this modeling approach enables to consider the impact of the evolution of the electricity mix. Structural models also introduces transparent modeling hypotheses, which are a source of trust for decision-makers. Still, we recognized the benefits of statistical approaches and will propose a method to retrieve information from historical datasets in the framework of structural models.

3.4 Conclusions

The European day-ahead electricity market relies on a complex mechanism to ensure the maximization of social welfare. A double-blind auction organized by power exchanges enables to set a unique market price for each hour and each bidding zone. The EUPHEMIA algorithm is responsible for the market clearing across Europe, which determines the executed bids according to the merit order. The construction of the internal energy market continues in Europe thanks to more harmonization (e.g. Capacity Allocation and Congestion Management regulation), integration (e.g. Single Day-Ahead Coupling), and adaptations to accommodate for a higher share of renewables (e.g. shortening of the market time unit). An overview of the key steps of the day-ahead clearing process that have been presented is illustrated in Figure 3.13.

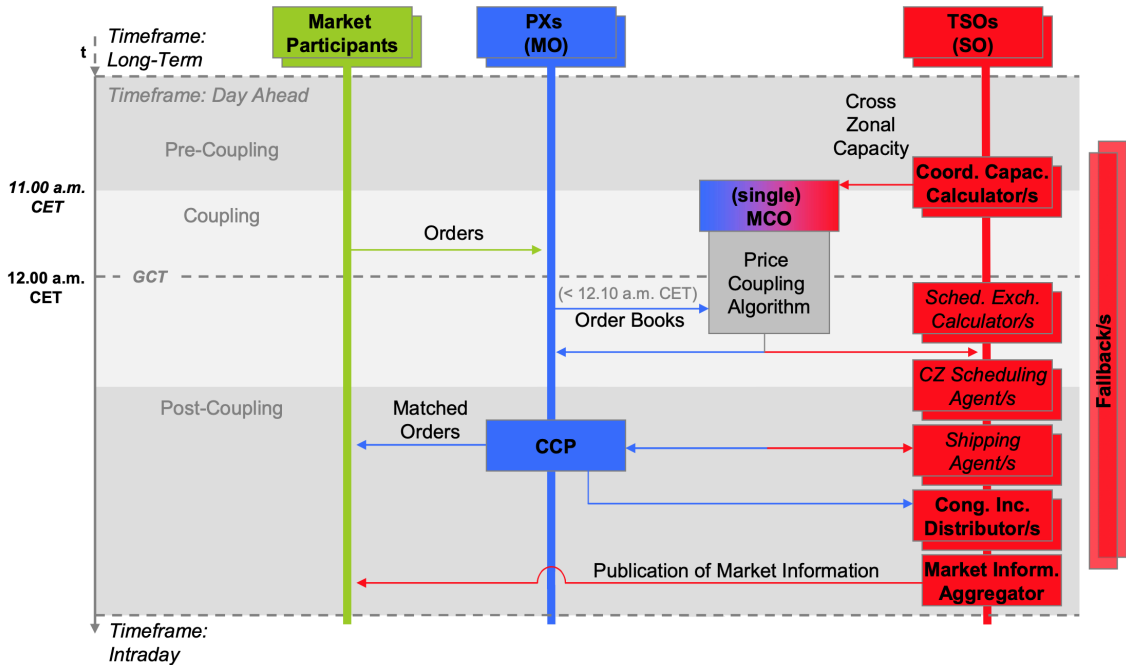


Figure 3.13: Key steps of the day ahead process (PX: power exchange, MO: market operator, SO: system operator, GCT: gate closure time, MCO: market coupling operator, CCP: central counterparty) [117]

The internal energy market enables to pool resources across Europe, thus assets can be used in an efficient fashion. The high level of liquidity and the converge of prices between various geographical areas are indicators of the good functioning of the market. Nevertheless, challenges remain: the zonal market design and uniform pricing have drawbacks that must be addressed by measures complementary to the day-ahead market. Congestion management, system adequacy, and social acceptability are key issues for the future of the power system and electricity markets. The evolution of the cost structure of the electricity mix, which includes renewables with low OPEX and fossil fuel technologies with higher short-run marginal costs, brings us to reconsider the benefits and drawbacks of marginal pricing.

The specificity of electricity as a commodity (grid-bound delivery, difficulty to store at scale, etc.) is the reason for the characteristics of prices on day-ahead markets. Analyzing the dynamics of electricity prices reveal the existence of a high volatility, of periodicities at daily, weekly and yearly timescales, of spikes followed by mean-reversion, and of negative prices. Further analyses confirm the link between the state of the power system and prices observed on wholesale markets. Statistical and machine learning models can be used for forecasting, especially for short-term horizons, but they do not allow us to fully leverage our understanding of electricity markets. Structural models, which reflect the price formation mechanism using a bottom-up approach, enables to conduct more detailed studies, therefore they will be the object of the next chapter.

Chapter 4

Data-driven structural modeling of prices

Contents

4.1	Literature review	58
4.2	Problem statement	60
4.3	Methodology	61
4.3.1	Market clearing with technical constraints	61
4.3.2	Parametrization of marginal costs	63
4.3.3	Bias correction	64
4.3.4	Estimation algorithm	64
4.4	Validating the approach on a case study	66
4.4.1	Method implementation	66
4.4.2	Data	69
4.4.3	Results	70
4.4.4	Discussion	76
4.5	Conclusions	80

In the previous chapters, we presented and analyzed key concepts for the modeling of prices on day-ahead electricity markets (how these markets work, which factors influence price formations, how historical data and numerical models can be exploited to simulate time series of prices, etc.). In this chapter, we explore a new category of models, called structural or fundamental models, that reflects the real price formation mechanism (uniform pricing and marginal pricing). To begin with, we discuss methods proposed in the scientific literature and approaches used by the private sector. Building on the review of literature, we then propose an innovative approach to leverage data-driven approaches within this class of optimization models. We analyse the benefits of the proposed method through a case study, which make use of real historical data. This chapter gives a detailed presentation of the core methodology, while model extensions and applications to prospective studies will be explored in the next chapter.

4.1 Literature review

Various approaches have been proposed in the literature to explore the relationship between energy prices and their drivers [6]. Prices can be simulated directly as a function of exogenous regressors (e.g. electricity demand, available capacity) employing time series models, artificial neural networks and regression trees, as presented in the previous chapter. These purely statistical models rely on observed correlations between variables and have good forecasting accuracy, especially for short-term horizons [118] when the power system composition remains the same. Nevertheless, to reflect a structural change in the system composition, both the supply-demand equilibrium and the associated price formation mechanism have to be explicitly taken into account in the regression models.

The equilibrium is usually modeled with an economic dispatch or unit commitment approach, which results in a cost-effective use of production assets to meet demand in power systems. The optimization solution provides a model for market prices for a given delivery period, either taking the marginal cost of the most expensive production mean with non-zero production, or equivalently, using the Lagrange multiplier associated with the demand constraint [4], also called the shadow price. These types of models are popular among transmission system operators (e.g. ENTSO-E uses ANTARES), political institutions (e.g. the European Commission uses METIS) and consulting companies (e.g. Cornwall Insight uses Plexos). A detailed review of modeling tools is proposed in [119] and a systematic mapping of power system models in [120]. For these models, generation units are typically grouped by technology and clustered in market areas, which are interconnected to partially model grid constraints. The models are used for long-term studies regarding system adequacy to estimate future system costs and to assess the potential benefits of new investments or market design adjustments. Monte Carlo simulations are used to evaluate the impact of different demand scenarios and meteorological conditions (temperature, rain, wind speed, solar radiation). However, because the models solve a global optimization problem they implicitly adopt the view of a unique central operator. Also, these models require the consideration of numerous techno-economic parameters that are rarely publicly available and that are hard to collect and estimate precisely (how-

ever, it is worth noting that the European Commission recently released some datasets used for METIS). Lastly, these models are generally not designed to simulate market prices. To our knowledge, their capacity to mimic real price dynamics is not evaluated. In the case of deviations between simulated and observed prices, the latter are not used for model tuning.

The equilibrium can also be studied from a multi-agent perspective, where the interactions between owners of different asset portfolios play a key role. Methods to reveal hidden variables related to bids associated with power supply assets can for example rely on a Bayesian approach [121] or on inverse optimization approaches [122], [123]. In these references, real historical data have not been used, but the studies are instead based on simulated test cases. More generally, this class of models is a popular choice for studying qualitative issues rather than for simulating long time series of real market prices [6] and, as such, the objective is usually not to exploit observed market data. Nevertheless, in nodal markets, more detailed technical and economic data about individual units and transmission lines can be leveraged through inverse optimization, as in [124], where the Midcontinent ISO electricity market (MISO) in North America is studied. In [124] learning from historical data primarily enables to infer information about grid usage, whereas our main focus is on supply bids.

Alternatively, the system equilibrium can be considered from a more market-oriented perspective by simulating supply curves [125]. In this case, for a given hour, the supply curve corresponds to the aggregation of market orders and not only depends on the merit order (i.e. the ranking of production units according to their short-run marginal cost), but also on the bidding decisions of market participants. The technical constraints that would induce temporal coupling, usually present in unit commitment models, are not included. Merit order models are also used to support business or policy development (e.g. EMMA developed by Neon Neue Energieökonomik and Power2Sim developed by Energy BrainPool). Such models can be classified as fundamental (or structural) models according to the electricity price forecast taxonomy proposed in [6]. Fundamental models are used for prospective studies as well as for ex-post analyses (e.g. [126] and [11]). A survey of structural models is proposed in [5].

For structural models, a key step is supply curve modeling. This is generally obtained as an aggregation of market orders, but it can also result from a statistical model as in [127], where the curve is built as a function of a normalized load and observed market prices. This purely statistical model is then adjusted to account for the evolution of fuel and emission costs for coal. The X-model presented in [7] and used in [128] proposes a piecewise model for the supply curve. Each local model is initially associated with a price level for which a volume of energy is determined through a statistical regression. Regressors at time t include solar and wind production or planned generation. An interpolation is used to build the final supply curve. The concept of the X-model has recently been extended in [129], where the supply and demand curves are considered jointly in order to reduce computation time and improve accuracy. While these piecewise models try to mimic the standard merit order approach, they do not include the effects of availability and production costs. On the contrary, a model inspired by [130] is

proposed in [8] for coupled electricity markets, where the supply curve is constructed by stacking sell orders for which the price offers are a function of the supply margin (i.e. the difference between the total available capacity and the demand) and the production cost by technology. An objective of the paper by these authors is the analysis of forward markets, which requires a specific form of price offer that reduces the realism of the day-ahead price dynamics in simulations. In [131], a clearing market price procedure, which also includes complex orders, is proposed. The bids are not parametrized and no learning algorithm is introduced since the case study is based on the Spanish market, where day-ahead market bids details are disclosed contrary to most European electricity markets. While the approach is relevant for ex post analyses, it would need to be extended to be used for prospective studies. In [9], specific attention is given to the variability realism of simulations. Their model starts with a description of the supply curve per technology through the short-run marginal cost. Price offers are then reduced in the case of low demand (reflecting for example must-run constraints) and increased for high demand (inducing a scarcity rent). The adjustment of bids is still done at the technology level without considering individual production units. References [10] and [11] also propose to model hourly prices using the merit order approach at the technology level to realistically model the price dynamics on the German day-ahead market. Both models make use of real data and linear regressions. The regressions are used in [10] to adjust the output of the merit order model to account for effects such as ramp constraints and strategic bidding, while [11] uses statistics to model the foreign trade balance. These models employ data to enhance realistic simulations, but they do not propose to directly use statistical calibration for the supply curve modeling.

4.2 Problem statement

As presented in the literature review above, structural models are beneficial for the conduct of prospective studies needed to support long-term decisions. By design, these models aim at capturing the price formation mechanism according to explicit technoeconomic hypotheses, and as such, these models enable to simulate electricity prices from various scenarios of evolution of the power systems, even if analogous conditions have not been observed yet. Although purely data-driven approaches (e.g. decision trees and artificial neural networks used in machine learning) are fundamentally not robust against major systemic changes and less suitable for such long-term studies, the benefits that data-driven approaches can bring should not be overlooked, since empirical verification relying on data from real observations is essential to ensure the validity of the formulated hypotheses. In other words, exploiting real data ensures that the theory is coherent with the market phenomena observed in practice. Although multiple solutions exist to model the full power system and unit constraints in order to perform studies for long-run simulations, few methodologies have been proposed to extract information from observed day-ahead prices in order to tune the internal parameters of such models. The novelty of this approach is that information extracted from in-sample datasets (e.g. past observations) becomes an integral part of the model, which can be used for out-of-sample

forecasts (e.g. prospective studies).

In the remaining of this chapter, we propose a structural model in which the supply curve is constructed with a bottom-up approach. Market orders are associated with production units and their prices are parametrized, which means that we can leverage available market and power system data. To develop this joint statistical and optimization approach, we introduce a method to simulate hourly electricity prices on day-ahead markets that meets two objectives: (i) reflect the techno-economic constraints imposed by the power system that induce the price formation according to the economic theory of marginal pricing under idealized conditions of perfect competition, and (ii) exploit market and power system data to calibrate and evaluate the model so that the simulations reflect the prices observed in practice. The method is presented and then validated using data related to the French day-ahead market and power system for the period from 2015 to 2018.

4.3 Methodology

4.3.1 Market clearing with technical constraints

In this section, we present the formulation of the optimization problem that emulates the market clearing mechanism and integrates power system constraints. The considered power system is composed of interconnected market areas. No assumptions are made on the nature of the energy mix.

The link between the technical constraints of the power system and electricity prices is modeled through a simple economic dispatch including a stock constraint for hydro power. This constraint introduces a temporal coupling that independently prevents a resolution for each time step. More technical constraints could be added here, though it is not the purpose of this chapter to add too much complexity here. For a more complete formulation and discussion on economic dispatch, see e.g. the review paper [132]. The formulation could be further extended to a unit commitment as presented in the review paper [133]. An elastic demand is introduced in the formulation making it closer to a market clearing with technical constraints than a pure economic dispatch. The cost function that is optimized depends on the marginal costs $\pi_{t,i}^P$ of production units, and the marginal utility $\pi_{t,j}^L$ for consumption units. The problem is formulated as a social welfare maximization, i.e. a maximization of the producer surplus and the consumer surplus:

$$\underset{P, L, E}{\text{maximize}} \quad \sum_{t \in T} \left(\sum_{j \in J} \pi_{t,j}^L L_{t,j} - \sum_{i \in I} \pi_{t,i}^P P_{t,i} \right) \quad (4.1a)$$

$$\text{subject to} \quad \sum_{i \in I_z} P_{t,i} + \sum_{z' \in Z} E_{t,z,z'} = \sum_{j \in J_z} L_{t,j} \quad \forall (t, z) \in T \times Z, \quad (4.1b)$$

$$E_{t,z,z'}^{min} \leq E_{t,z,z'} \leq E_{t,z,z'}^{max} \quad \forall (t, z, z') \in T \times Z \times Z, \quad (4.1c)$$

$$L_{t,j}^{min} \leq L_{t,j} \leq L_{t,j}^{max} \quad \forall (t, j) \in T \times J, \quad (4.1d)$$

$$P_{t,i}^{min} \leq P_{t,i} \leq P_{t,i}^{max} \quad \forall (t, i) \in T \times I, \quad (4.1e)$$

$$\sum_{(t,i) \in T \times I_z^{hydro}} P_{t,i} \leq S_{T,z}^{hydro} \quad \forall z \in Z \quad (4.1f)$$

where

- (1a) is the social welfare maximization objective
- (1b) is the power balance constraint
- (1c) is the interconnection capacity constraint
- (1d) is the consumption unit limits constraint
- (1e) is the generating unit limits constraint
- (1f) is the hydro stock constraint

We denote $\Pi_{t,z}^{MC}$ as the Lagrange multiplier associated with the constraint (4.1b). In the case without temporal coupling, the marginal costs $\pi_{t,i}^P$ are simply the sell prices of market participants and $\Pi_{t,z}^{MC}$ is the result of a market clearing. In a pure market vision, the set of production units I would be the set of orders rather than a set of production means. With a technical constraint, such as the hydro stock constraint, the sell price of hydro is the sum of the marginal cost $\pi_{t,i}^P$ and the Lagrange multiplier $\lambda^{(1f)}$ associated with the stock constraint (4.1f), also called the value of water. The interest of this kind of formulation compared to a pure market clearing is that a sell price taking into account technical constraints is computed internally by the model. In the general case, Lagrange multipliers associated with all internal technical constraints included in the model can be allocated to production means, so that we can write the sell price of production unit i at time t as $\pi_{t,i}^P + \lambda_{t,i}^{IC}$.

This optimization approach is widely used to simulate market prices from marginal costs and considering the technical description of the electric system (e.g. availability of production units, hydro stock constraint, transmission constraints). By construction, it reflects the day-ahead market clearing mechanism. However, it does not account for strategic bidding, information asymmetry and market power that can be observed in practice. In addition, financial aspects can also impact the formation of prices, such as the existence of multiple markets, which not only allows participants to sell and buy on the day-ahead market, but makes doing so more interesting for them regarding their profit expectation and risk management policies.

4.3.2 Parametrization of marginal costs

The concluding remark of the previous subsection highlights that, in practice, sell prices are not only influenced by the well-defined techno-economic constraints IC that are accounted for in Eq. (4.1), but also by strategic decisions of market participants that impact price formation. The contribution of these other factors is denoted $\epsilon_{t,i}$. We finally propose the following additive decomposition of the sell prices $\hat{\pi}_{t,i}^{sell}$, which accounts for the contribution of both $\lambda_{t,i}^{IC}$ and $\epsilon_{t,i}$:

$$\hat{\pi}_{t,i}^{sell} = \pi_{t,i}^P + \lambda_{t,i}^{IC} + \epsilon_{t,i} \quad (4.2)$$

The model's internal constraints IC are explicit, so the modeler can directly assess the effect that adding or removing constraints has on simulated prices (e.g. adding ramp constraints to the existing ones). The modeler can also easily evaluate the impact of the limits imposed (e.g. by changing the values of the hydro stock constraint or the maximal available production capacity constraint). But this approach is not applicable for $\epsilon_{t,i}$ since it is not practical to quantify a priori the impact on hourly day-ahead prices of economic factors such as strategic bidding, information asymmetry, market power and the co-existence of multiple wholesale electricity markets. Using a data-driven approach on historical time series helps us to empirically quantify the influence of these factors. So, we propose to build a statistical linear model $\hat{\pi}_{t,i}^P$ for $\pi_{t,i}^P + \epsilon_{t,i}$. $\hat{\pi}_{t,i}^P$ is a marginal cost proxy that accounts for both an objective short-run marginal cost of production unit i ($\pi_{t,i}^P$) and the influence of strategic factors ($\epsilon_{t,i}$). The main drivers of these marginal cost proxies include: fuel prices, emission prices, global supply margins (i.e. the difference between the available capacity and the demand), and the rank of each production unit within its production class, and thus we propose the parametrization described in Eq. (4.3). Fuel and emission prices allow us to account for the evolution of operating costs (OPEX). The contribution of the margin term reflects an important part of the strategic behavior that can occur during scarcity situations when market participants add a price premium to their bids. The rank of a production unit is introduced to reflect the heterogeneity within a production class since not all units bid the exact same price (e.g. due to various levels of efficiency and flexibility).

$$\begin{aligned} \hat{\pi}_{t,i}^P &= \alpha_{c(i),0} + \alpha_{c(i),\mathcal{F}} \cdot \mathcal{F}_{t,c(i)} + \alpha_{c(i),\mathcal{E}} \cdot \mathcal{E}_t + \alpha_{c(i),\mathcal{M}} \cdot \mathcal{M}_t + \alpha_{c(i),\mathcal{R}} \cdot \mathcal{R}_i \\ &= \boldsymbol{\alpha}_{c(i)}^\top \mathbf{X}_{t,i} \end{aligned} \quad (4.3)$$

where the parameters of the model are only a function of the production unit class $c(i)$:

- $\alpha_{c(i),x}$ - Calibrated parameters by class $c(i)$ for each price driver x
- $\mathcal{F}_{t,c(i)}$ - Fuel price at time t for class $c(i)$
- $\mathcal{E}_t$ - Emission price at time t
- $\mathcal{M}_t$ - Supply margin at time t
- $\mathcal{R}_i$ - Proxy variable for the rank of production unit i in class $c(i)$

Since the sell price is obtained by adding $\lambda_{t,i}^{IC}$ to $\hat{\pi}_{t,i}^P$, the price that we propose is also a statistical model of sell orders. Using the marginal cost proxy within the optimization problem defined in Eq. (4.1) results in simulating a day-ahead market price that we denote by $\hat{\Pi}_{t,z}^{MC}$.

The market clearing with technical constraints and the parametrization of marginal costs presented here is directly applicable for markets with zonal pricing. The main ideas could also be considered for nodal markets, but the approach would have to be adapted to account for the differences highlighted in the introduction. Here, we succinctly introduce some points to be considered for this adaptation. Since transmission constraints play a greater role, an additional price driver could be considered in the formulation of the marginal cost proxy so that the parameter associated with these restrictions reflects the impact of congestion observed in practice. Also, due to the lower number of generation units, the parameter values could be specific for each unit (as opposed to production classes) in order to better account for bidding strategies. Information retrieved at the unit level would then be more precise, but the optimization problem might become unstable due to the high number of variables.

4.3.3 Bias correction

We also introduce a bias correction term $\beta_{t,z}$ to finally obtain the simulated market prices $\hat{\Pi}_{t,z}$ studied in this chapter that we define as follows:

$$\begin{aligned}\hat{\Pi}_{t,z} &= \hat{\Pi}_{t,z}^{MC} + \beta_{t,z} \\ &= \hat{\Pi}_{t,z}^{MC} + \beta_z^\top \mathbf{Y}_t\end{aligned}\tag{4.4}$$

where for each region z , $\beta_z \in \mathbb{R}^{24 \times 7}$ is a parameter, and for each time step t $\mathbf{Y}_t$ is the 24×7 dimensional vector with $Y_t[(h, d)] = 1$ if $hour(t) = h$ and $weekday(t) = d$; otherwise $Y_t[(h, d)] = 0$. This bias correction term is similar in spirit to that proposed in [10]. However, we made this part less sophisticated because the correction is applied after the market clearing meaning that it would not be sensitive to technical changes in the electric system or structural changes in the market rules. This allows us to account for additional dependencies not represented in this simplified market clearing formulation (e.g. additional temporal dependencies) in order to simulate price dynamics more accurately. For further use in prospective applications, we believe this term should be kept as small as possible.

4.3.4 Estimation algorithm

The parameters of Eq. (4.3) and Eq. (4.4) must be estimated from real data. For a time step t for which we know that the production unit $i^*(t)$ is marginal, the at-the-money sell price establishes the simulated market price, which can be rewritten as:

$$\begin{aligned}\hat{\Pi}_{t,z} &= \hat{\pi}_{t,i^*(t)}^P + \lambda_{t,i^*(t)}^{IC} + \beta_{t,z} \\ &= \alpha_{c(i^*(t))}^\top \mathbf{X}_{t,i^*(t)} + \lambda_{t,i^*(t)}^{IC} + \beta_z^\top \mathbf{Y}_t\end{aligned}\tag{4.5}$$

We start the algorithm with $\beta_z = \mathbf{0}$, and $\alpha_{c(i)} = \alpha_{c(i)}^0$ as an initial value.

Market Clearing Step From a set of marginal cost proxies we can solve Eq. (4.1). This leads to an estimation $\hat{\Pi}_{t,z}$ of $\Pi_{t,z}$, and to the identification of marginal production units. This allows us to group time steps according to the value of the marginal class $c(i^*(t))$ denoting the associated partition as $(T_c)_{c \in C}$.

In more detail, this first step of the algorithm consists in simulating prices using the model with the current values of its parameters. Employing the simulation outcome, we identify which generation unit is marginal for each of the simulated hours (i.e. the unit with the highest sell price and a non-zero production value). We can then group time steps according to the marginal class. For example, if a gas unit is marginal at time $T_{gas} = \{t_1, t_{12}, t_{25}, t_{28}\}$ and a nuclear unit is marginal at time $T_{nuclear} = \{t_3, t_4, t_{11}, t_{22}, t_{23}\}$, we will create groups of hours accordingly. These groups are then used in the next step of the estimation algorithm.

Parameter update through statistical estimation Residuals $\Pi_{t,z} - \hat{\Pi}_{t,z}$ can be used to estimate a new value for β_z . Meanwhile, α_{c_0} is obtained by a linear regression of $(\mathbf{X}_{t,i})_{(t,i) \in T_{c_0} \times I_z}$ on the observed price $(\Pi_{t,z})_{t \in T_{c_0}}$. Coming back to the example given above, in order to determine the values of α_{gas} , we perform the regression only with the subset of input data corresponding to T_{gas} and not with the whole dataset (and similarly for the other production classes).

Constraints can be added at this regression stage, concerning for example the sign of coefficients, or values that are known (and not estimated) such as $\alpha_{c(i),\varepsilon}$, which is directly obtained from the average class emissions. We iterate until a stable RMSE is obtained.

Interpretation of the algorithm A natural justification of this algorithm comes from the following decomposition of the RMSE:

$$RMSE_z^2 = \frac{1}{T} \sum_{c_0 \in C} \sum_{t \in T_{c_0}} (\alpha_{c_0}^\top \mathbf{X}_{t,i^*(t)} + \lambda_{t,i^*(t)}^{IC} + \beta_z^\top \mathbf{Y}_t - \Pi_{t,z})^2 \quad (4.6)$$

It is important to note that the error here is not just a quadratic function of α and β_z because, when these parameters change, the merit order changes and this affects the partition $(T_c)_{c \in C}$. In the proposed algorithm, we alternate between two steps. The first is an estimation of $(T_c)_{c \in C}$ from the value of α and β_z . It can be interpreted as a computation of the local linearization of the cost function. In the second, we estimate a new value for the parameters α and β_z according to this linearization. This shows that our algorithm is similar to a gradient descent on the cost function.

4.4 Validating the approach on a case study

We implemented a method following the described approach, with R programming language, using data relating to the French market from 2015 to 2018. In our case study, all production units present in the power system are included in the optimization problem defined in Eq. (4.1), i.e. we assume that all units participate in the day-ahead market. There are several reasons for this choice. The first is that our aim is to estimate prices with a variability reflecting the composition of the whole production system. In addition, even if only part of the production in the electric system participates in the day-ahead market, the remaining units that do not participate still influences prices through other markets that we do not want to model explicitly here. Finally, no public data are available on the exact composition of the production units participating in the market, while production unit participation in the electric system is publicly available information. As we consider a single bidding zone case here, the subscript z will be omitted in this section. Application to a multizone case would not change the algorithm as it will be detailed in the next chapter.

4.4.1 Method implementation

Overview

For both training (i.e. the estimation of model parameters) and out-of-sample testing, considering a whole calendar year allows us to observe simultaneously the daily, weekly and seasonal variabilities. In addition, this long timescale allows us to observe the effect of imposing a yearly constraint on the usage of the global hydro stock.

We simulate prices in a single bidding zone and incorporate the effect of the neighboring interconnected zones as additional orders in both the supply and demand curves. For the numerical application, we use the observed net transfer capacity values as order quantities (both as $P_{t,i}^{max}$ to account for the possibility to import and as $L_{t,j}^{max}$ to account for the possibility to export) and the foreign day-ahead market prices as order prices ($\pi_{t,i}^P$ and $\pi_{t,j}^L$ respectively for supply and demand). No minimal volumes for import or export are imposed (i.e. $P_{t,i}^{min}$ and $L_{t,j}^{min}$ are equal to 0 MWh).

Figure 4.1 gives an overview of the case study.

Supply

For the orders that constitute the supply curve, we consider the production units of five production classes: nuclear, fossil gas, fossil hard coal, fossil oil and hydro water reservoir. Each order corresponds to a production unit with an installed capacity of over 100 MW. The quantities supplied each hour correspond to the available capacity $P_{t,i}^{max}$, which varies based on the reported planned and forced outages of the production units. For each production class, units with a capacity of less than 100 MW are aggregated into a single order whose quantity only depends on the installed capacity; the potential unavailabilities are not accounted for. $P_{t,i}^{min}$ is set to 0 MWh for all units.

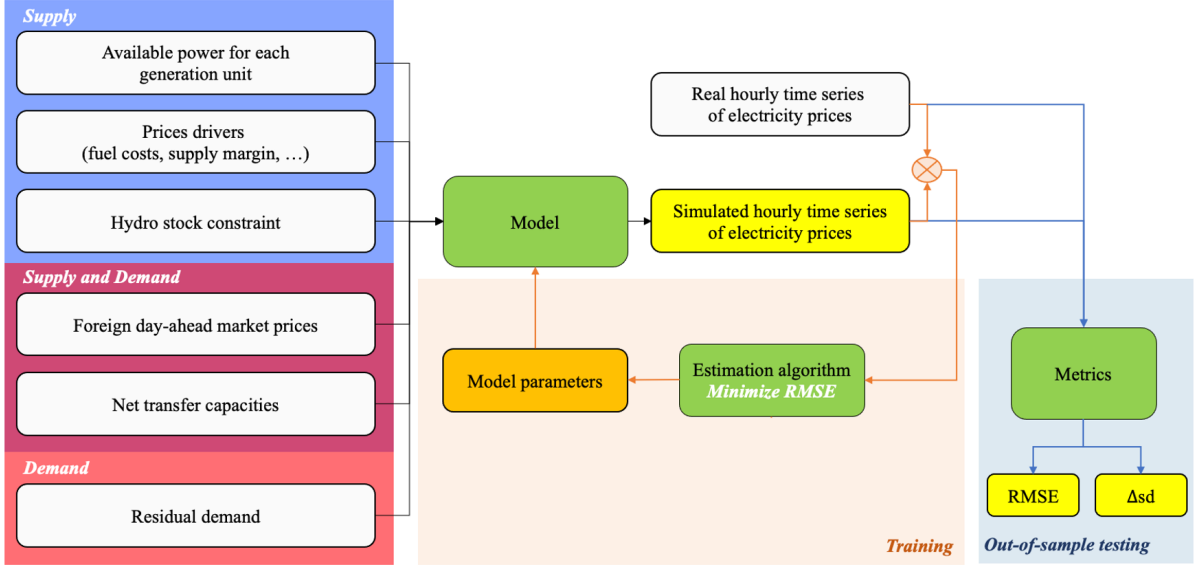


Figure 4.1: Case study overview

In the ENTSO-E taxonomy, hydro is classified according to three subcategories: (i) run-of-river and poundage, (ii) pumped storage, and (iii) water reservoirs. The hourly production of run-of-river and poundage, as well as the hourly production and consumption of pumped storage, are among the components which are subtracted from the hourly day-ahead consumption forecast in order to obtain the aggregated residual demand used as model input in the case study, as detailed in subsection 4.4.1. For the sell orders corresponding to supply, we only consider water reservoirs.

The price drivers considered by production class are summarized in Table 4.1. Note that the variation in uranium prices over the studied period is neglected compared to the fixed operating costs of nuclear power plants and that the ranks are based on the nominal power of production units. The rank in our model refers to a proxy variable indicative of the merit order place of a unit within its production class. This variable is used to account for differences in short-run marginal costs between units, which are due to different levels of efficiency and specific operational constraints (the lack of publicly available information at the unit level motivates the use of a proxy variable). To define the variable for the rank of production ($\mathcal{R}_i$) we consider each production class $c(i)$ separately and we sort its production units i by ascending order of nominal power P_i^{max} , then we compute $\mathcal{R}_i$ by recurrence: $\mathcal{R}_0 = P_0^{max}$ and $\mathcal{R}_i = \mathcal{R}_{i-1} + P_i^{max}$. This construction is motivated by the fact that we consider large power plants to be more efficient than smaller ones (especially for the French nuclear power plants for which the nominal power is correlated to the installation date).

Production class (c)	Unit ranking ($\mathcal{R}$)	Supply margin ($\mathcal{M}$)	Fuel cost ($\mathcal{F}$)	Emission cost ($\mathcal{E}$)
Hydro water reservoir	x	x		
Nuclear	x	x		
Fossil gas	x	x	x	x
Fossil hard coal	x	x	x	x
Fossil oil	x	x	x	x

Table 4.1: Price drivers of marginal cost proxies considered by production class

Parameter estimation

For each class, we initialize $\alpha_{c,0}$ introduced in Eq. (4.3) to be constant and equal to an average short-run marginal cost (i.e. the cost associated with the production of additional energy output). To avoid over-fitting, we constrain the regression to impose a sign on the coefficients $\alpha_{c,x}$ during calibration, so that an increase in fuel cost or a decrease in supply margin leads to an increase in energy price. We do not calibrate the parameters $\alpha_{c(i),\mathcal{E}}$ associated with emissions and choose instead to rely on data from RTE, the French transmission system operator. The data used for the estimation of the parameters associated with the unit ranking, the supply margin and the fuel costs are presented in subsection 4.4.2. Alternatively, we could have chosen to set the parameters $\alpha_{c(i),\mathcal{F}}$ associated with fuel prices using public data, since the parameters correspond to the average energy conversion efficiency by production class. We chose instead to estimate the parameters associated with fuel prices in order to mitigate against uncertainty regarding the available data. First, for each class, different values of efficiencies are published (these values depend on technology, operating points, etc.). Second, the actual cost of fuel procurement depends on undisclosed hedging strategies of electricity producers, so this cost can differ from the variable $\mathcal{F}_{t,c(i)}$, which corresponds to the fuel price on a specific market and at a specific time. As both the values of fuel costs and the price parameters (e.g. $\alpha_{gas,\mathcal{F}}$) are uncertain, we applied the estimation approach to set the values of $\alpha_{c(i),\mathcal{F}}$ instead of relying on public data. We note that both $\alpha_{c(i),\mathcal{E}}$ and $\alpha_{c(i),\mathcal{F}}$ can be predefined or estimated depending on modeling choices and the reliability of available data.

To estimate β introduced in Eq. (4.4), we compute the average of the residuals over the training set by hour and day of the week (i.e. 24x7 offset values) and we then apply this correction on out-of-sample tests.

Residual demand

We consider an aggregated residual demand for the bidding zone $L_{t,j}^{max}$ with a price set to the day-ahead market price cap (i.e. $\pi_{t,j}^L$ is equal to EUR 3,000 per MWh). The price cap value does not set a market price in this case study as no supply shortage was observed. Since this residual demand j is always fulfilled in the case study, $L_{t,j}$ is equal to $L_{t,j}^{max}$.

for all time steps. For the simulation, we set the value of $L_{t,j}^{min}$ to 0 MWh, but the same results would have been obtained for all values between 0 MWh and $L_{t,j}^{max}$. The demand $L_{t,j}^{max}$ results from the difference between the day-ahead consumption forecast by RTE and the contribution of supply assumed to be a price-taker for the studied period and location. Here, this supply is composed of fatal power production (solar, wind and run-of-river hydro), co-generation (observed production according to a baseload profile), storage (we consider only the contribution of pumped-storage hydroelectricity, not batteries, etc.) and production classes with low installed capacity (biomass). Including storage as a production unit in the initial optimization problem would not change the algorithm, but as already mentioned, it is not the objective of this chapter to include complex technical constraints (model extensions will be explored in the next chapter).

Market clearing

We solve Eq. (4.1) in two steps, which corresponds to dualizing the hydro stock constraint. In the first step, we maximize social welfare using the orders for each hour separately and we obtain a time series of market prices. In the second step, we adjust the availability of hydro power plants so that they produce only at higher prices within the limit of the annual hydro stock constraint, then we solve the market clearing anew as in the first step. This two-step procedure is reproduced at each iteration of the whole estimation algorithm.

4.4.2 Data

The raw datasets used for the simulation and results analysis are listed in Table 4.2.

The hydro stock $S_{T,z}^{hydro}$, computed using the dataset "Hourly production by production class", is equal to the sum of the hydro water reservoir production over one year.

Two remarks can be made about the datasets:

- The UK National Balancing Point (NBP) was used instead of the French *Point d'échange de gaz* (PEG) for gas prices due to data availability, but these prices are considered to be correlated enough for our modeling purposes.
- All datasets were collected from the web, except Meteo France's average atmospheric temperatures, which are not open data but available under special conditions for research purposes. However, this dataset is only used here to analyze simulation errors; therefore, the core approach proposed can be developed and evaluated using publicly available data only.

¹<https://transparency.entsoe.eu/>

²<https://www.rte-france.com/eco2mix/telecharger-les-indicateurs>

³<https://www.services-rte.com/>

⁴<http://developpement-durable.bsocom.fr>

⁵<https://www.erce.energy/graph/uk-natural-gas-nbp-spot-price>

⁶https://markets.businessinsider.com/commodities/historical-prices/co2-european-emission-allowances/euro/1.12.2014_1.2.2019

Dataset	Source
Installed capacity by production class	ENTSO-E transparency platform ¹
Installed capacity of production units ($P_{nom} > 100\text{MW}$)	
Unavailability of production units ($P_{nom} > 100\text{MW}$)	
Prices on day-ahead markets	Open Power System Data [134]
Day-ahead consumption forecast	RTE (Eco2Mix platform) ²
Hourly production by production class	
CO2 emissions by production class	
Net Transfer Capacity to and from neighboring bidding areas (NTC)	RTE (Data portal) ³
Operating costs by production class	RTE/ADEME study [135]
Import costs of coal and oil in France	French Ministry for the Ecological Transition (PEGASE) ⁴
Gas market prices	ERCE ⁵
CO2 European Emission Allowances	Business Insider ⁶
Atmospheric temperature	Meteo France

Table 4.2: Data sources for the case study

4.4.3 Results

The simulation results are obtained by out-of-sample testing. To do so, one year is used for training and then we test over each of the remaining years (we repeat this process for the four calendar years in the dataset). In this way, each year is used three times as a testing set. The final result for each year, presented throughout the chapter, is the average of the ensemble of the three alternative simulations in which this year is used as a testing set (e.g. using the models trained with the 2015, 2016 and 2017 datasets, we simulate three distinct time series of prices for 2018 and then we calculate the hourly average).

Presentation of the simulated time series

The results obtained using the method and data described above are shown in Figure 4.2. In addition to the model's output (i.e. time series with hourly resolution), we analyzed its variability by computing the average prices by day of the week and hour of the day. The temporal variations appear to be globally captured by the model.

A focus on a period of one month presented in Figure 4.3 gives a more detailed representation of the dynamics of day-ahead prices.

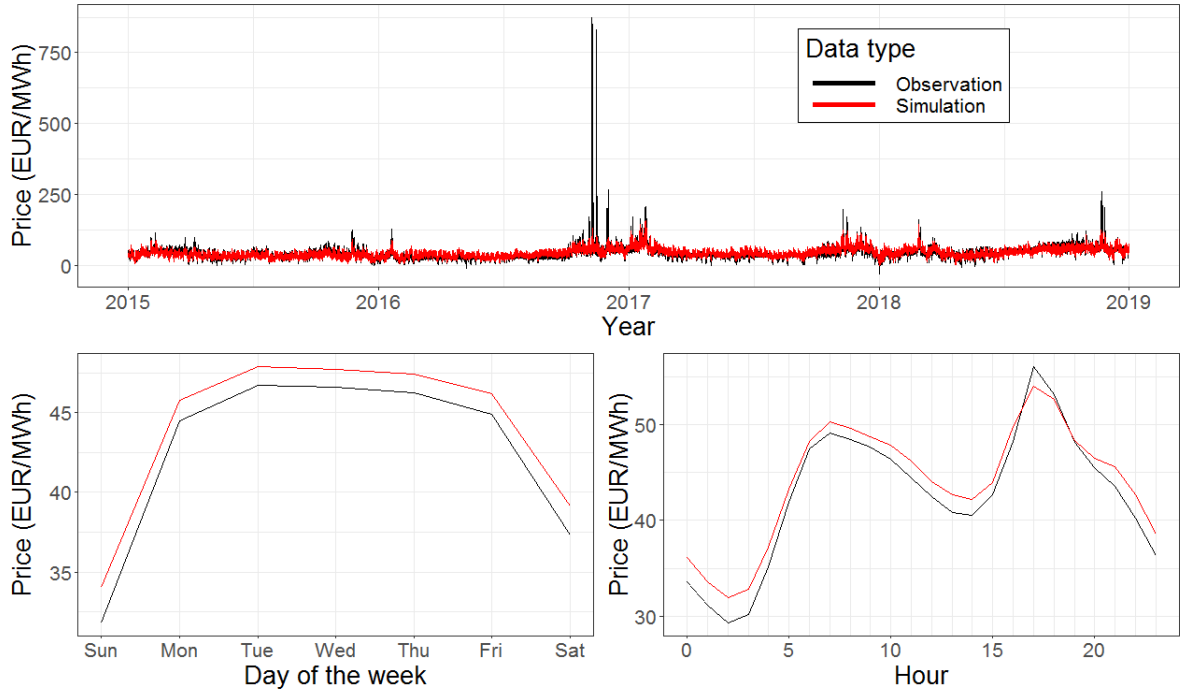


Figure 4.2: Time series of prices on the French day-ahead market over four years (top), average prices by day of the week (bottom-left) and average prices by hour of the day (bottom-right)

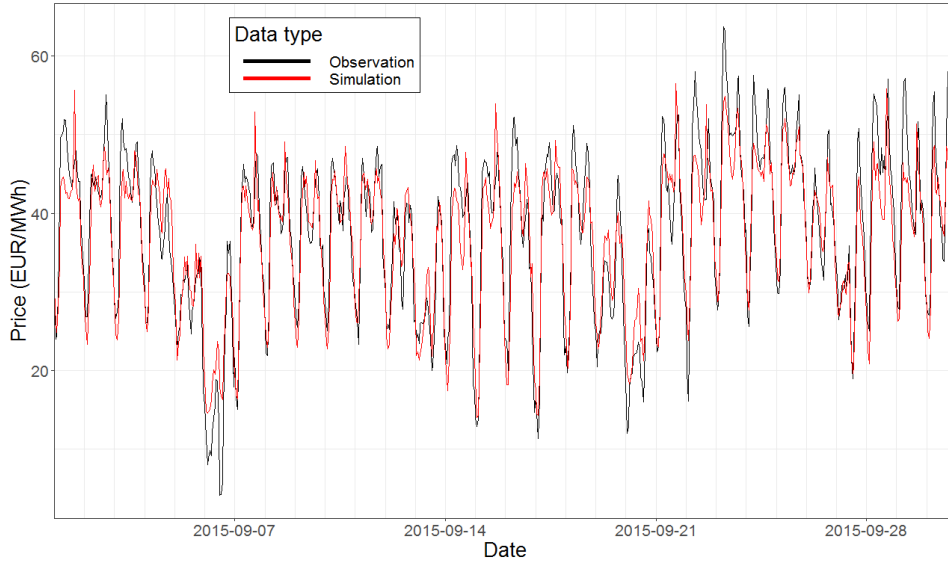


Figure 4.3: Time series of prices on the French day-ahead market over one month

Evaluation of the simulations

To analyze the results, we first compute the metrics and we obtain $RMSE = 10.8$ and $\Delta sd = 4.9$ over the whole test four-year test period. If we limit our analysis to

the year 2018, we obtain a $RMSE = 7.5$, which is a significant improvement over the simple tests performed with linear and random forests models in the previous chapter. Nevertheless, we observe a deterioration of the performance regarding the estimation of the price dynamics, especially compared to random forests, as we obtain $\Delta sd = 3.6$ with our model.

To go beyond the analysis of metrics, we plot the distribution of our simulated prices against the distribution of observed prices. The results are displayed with smoothed density estimates in Figure 4.4. To make the visualization clearer, the outliers above EUR 150/MWh are not displayed. While the prices are correctly modeled on average (the mean of observed prices is equal to EUR 42.6/MWh and it is equal to EUR 44.0/MWh for the simulated prices), we observed that the dispersion of prices is underestimated by our model. So the error measured by Δsd is not only stemming from price spikes but from the whole distribution. In the next chapter, we will explore advanced parameter estimation techniques to mitigate this effect. The left tail appears to be overestimated; this could be partially due to the use of the actual production values of weather-driven renewables instead of using the forecasted values as the market participants would do.

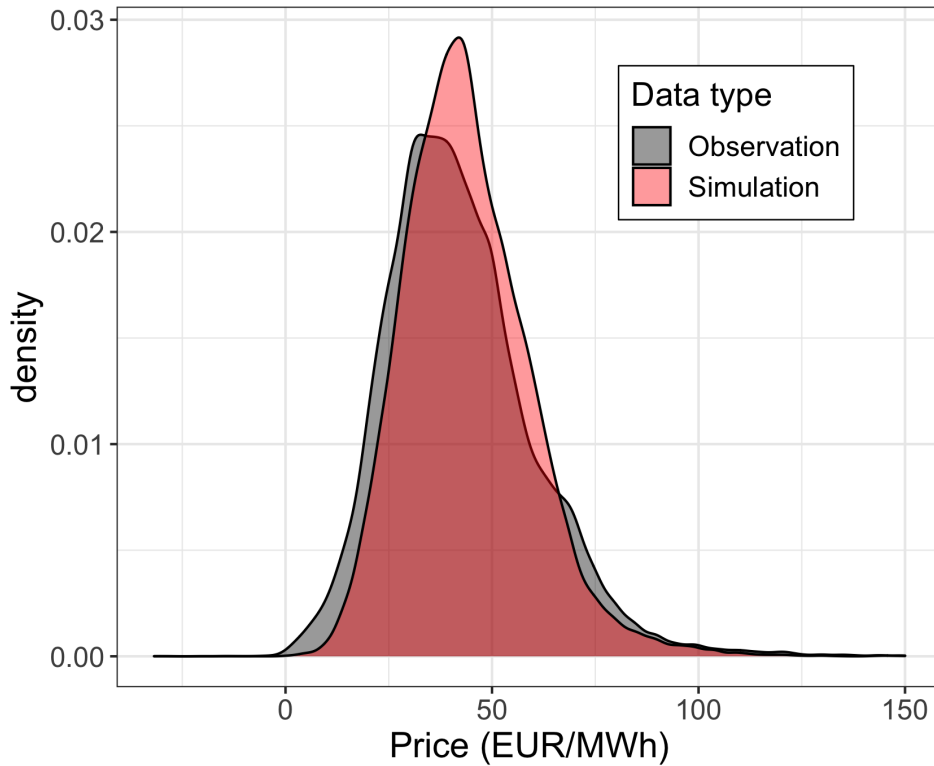


Figure 4.4: Density plot of prices on the French day-ahead market from 2015 to 2018

Studying residuals, i.e. the difference between observed and simulated prices, is also indicative of areas of improvement. If we consider the average value of residuals by month, as shown in Figure 4.5, we observe that simulations tend to overestimate prices during the first three trimesters of the year, while prices are underestimated from

October to December. Expanding the model to better capture seasonal variations and having more differentiated values throughout the year would help to correct the price distribution effects discussed above. One of the reason for this limitations could be that residual demand in our model jointly consider consumption and vRES production, while market participants adapt their bidding strategy depending on the composition of this residual demand.

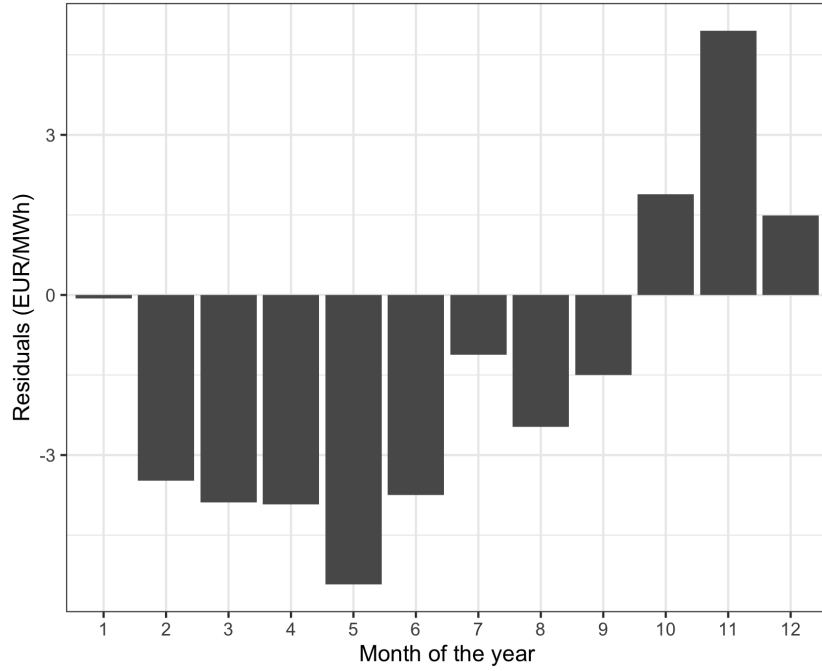


Figure 4.5: Mean of residuals by month from 2015 to 2018

Price offers decomposition

In this subsection, we propose to have a closer look at the simulated price formation. We consider the modeled hourly orders that the market clearing prices (i.e. the orders "at-the-money", which correspond to marginal units).

First of all, we analyze how often each production class set market prices. For this analysis, we consider interconnections as a production class. Hydro water reservoir are not represented as their contribution is considered in the case study to always be price-taker, where the hydro stock is strategically used during hours corresponding to the highest market prices. The relative frequencies of marginality by production class are summarized in Table 4.3. We observe some expected results: nuclear serves as a baseload most of the time, interconnections set prices quite often as the true observed foreign prices are used in the case study are more reliable for than the prices simulated for the other productions classes, and prices are set by fossil fuels for the remaining 15% of hours.

Production class	Marginality frequency
Nuclear	61.2%
Interconnections	23.8%
Fossil gas	8.1%
Fossil hard coal	3.9%
Fossil oil	3.1%

Table 4.3: Frequency of marginality by production class

We can then look into the average decomposition of prices by production class. The values are summarized in Table 4.4. In order to compare the relative influence of each factor, we provide the product between the parameter associated with a price driver and the value of the price driver, so that all values are in EUR/MWh. We observe that the price of orders from nuclear units are impacted by the variation of supply margin and also by which unit is last called with the production class. The decomposition of electricity prices set by gas unit reveal that the operating costs play an important role in price formation and, as expected, the average order prices are higher than for nuclear. More surprisingly, the prices for coal and oil seem underestimated, although the estimation of parameters still reveal important features such as the high dependency of oil price to the supply margin. The performance of the estimation algorithm might be degraded for these two productions classes due to the relatively low number of production units in France and the insufficient amount of hours for which these fossil fuels is marginal in order to better estimate these values.

Production class	Intercept	Unit ranking	Supply margin	Fuel cost	Emission cost	Order price
Interconnections	n.a.	n.a.	n.a.	n.a.	n.a.	46.8
Nuclear	89.5	-21.7	-24.6	n.a.	n.a.	43.2
Fossil gas	55.5	-0.10	-23.0	21.2	3.95	57.7
Fossil hard coal	32.9	-1.18	-15.3	11.5	8.89	36.8
Fossil oil	89.6	-7.86	-162.0	120.	7.23	46.9

Table 4.4: Average contribution of each price driver to in the additive decomposition of order prices by production class (values in EUR/MWh)

The bias correction term β partially compensate for the underestimation of prices set by coal and oil. As we can see in Figure 4.6, this bias correction plays a more important role during the weekday hours around dinner time, which typically correspond to higher electricity consumption in France and are associated with a significant ramp of demand.

Model parameters could be better estimated by increasing the size of the training dataset, but having a longer time horizon would create new challenges for the management of hydro stock. We suppose here that a total hydro stock is set initially and can be used for the duration of the training and testing period. For a longer horizon, additional considerations would have to be accounted for to handle a dynamic stock that can be

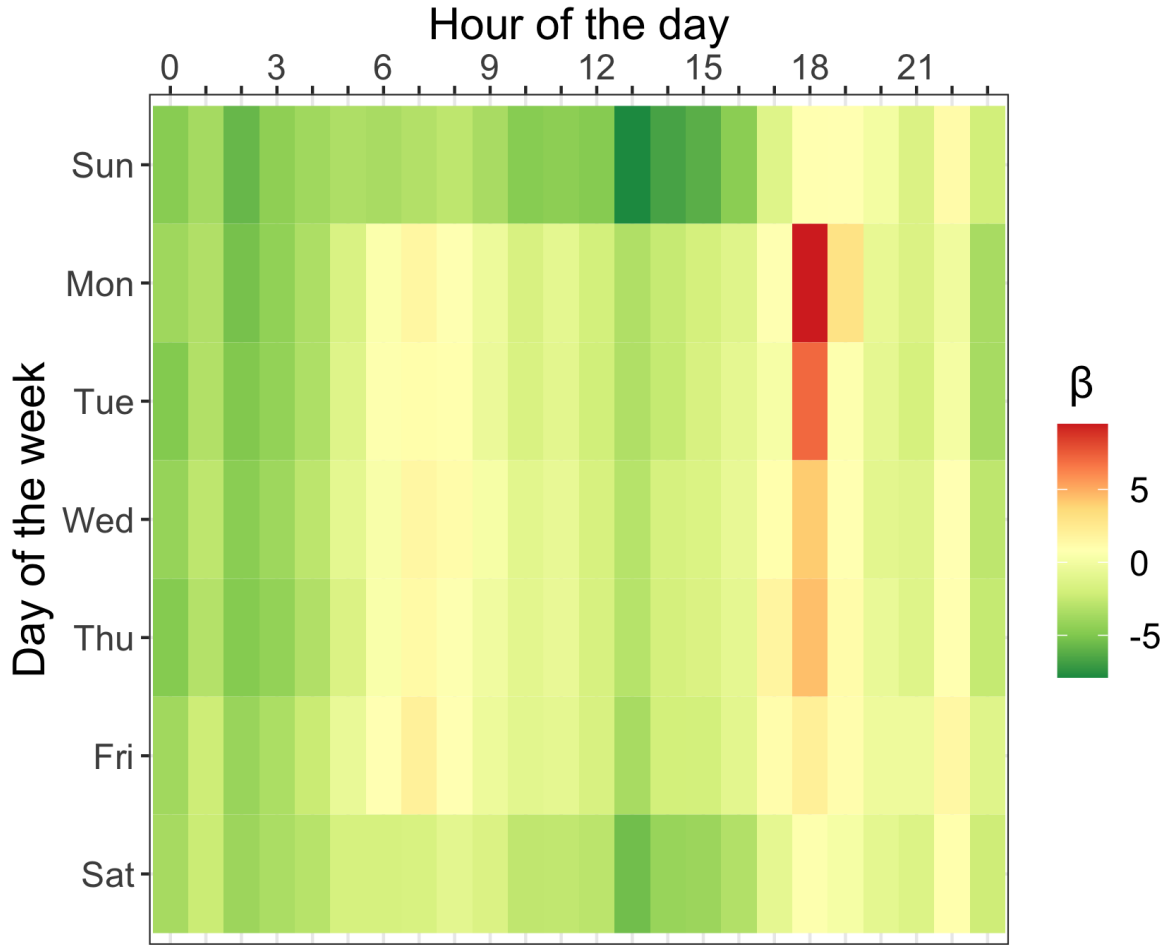


Figure 4.6: Bias correction by hour of the day and day of the week

filled thanks to precipitations. Another approach for a better estimation of model parameters would be to add additional constraints on each model parameters, for example by imposing minimal and maximal values. The risk of doing so would be to have a form of overfitting, where the modeler imposes too many hypotheses in order to match the datasets too closely, thus losing the benefits of statistical learning. Nevertheless, the formulation of the parametrization and methods for a finer estimation of parameters could be the subject of future work.

Comparison to benchmark

We aim to (i) assess the benefits of the model calibration to estimate the contribution of statistics to our structural model, and (ii) compare our performance to purely data-driven methods (i.e. linear regression as a reference for statistical modeling and random forests as a reference for machine learning modeling). The results are summarized in Figure 4.7.

We observe that model calibration, which corresponds to our reference case, improves the performance for our two metrics. Also, when we consider only a basic dataset consisting of power system data as an input for the statistical and machine learning models (i.e. excluding information about prices and costs), our calibrated model tends to perform better. Nevertheless, when considering an extended dataset for calibration, which includes power system and market data (fuel costs and day-ahead prices in the neighboring bidding zones), the purely data-driven approaches tend to be better at exploiting correlations between prices on different markets.

Although purely data-driven approaches perform well in this case study, they are given only as a reference since they do not fulfill the objectives set in our problem statement, which is to propose a suitable method for prospective studies. As these methods exploit the outcome of past observed situations and do not intrinsically consider the price formation mechanism, they are less relevant to study the impact in terms of prices of long-term scenarios where major system changes are considered. As requirements for prospective studies, the model must offer the possibility to estimate the potential consequences of changes in electricity mix or market design. By construction, structural models are well-suited to integrate the technical constraints imposed by the electricity mix, and the optimization problem can be extended beyond the current formulation if needed. Nevertheless, our data-driven approach would not be suitable for scenarios which radically differ from the training dataset (e.g. in the case of 100% renewable energy). First of all, in case of radical changes, the behavior of market participants will differ from the behavior inferred from historical data. In addition, the bias correction term $\beta_{t,z}$ in Eq. (4.4) reflects a periodicity of spot prices that would also change in case of a major evolution of the energy mix. Regarding the market design, our model represents the price formation induced by a zonal double-sided auction market with uniform pricing. The impact of the design of other electricity markets on the day-ahead market is not the object of this case study, but it could be the focus of additional analyses under the current framework. For example, in the case of France, the ARENH mechanism forces EDF to sell nuclear electricity to competitors in order to mitigate the effect of EDF's historical monopoly on the French energy supply (up to 100 TWh of electricity at EUR 42 per MWh from 2015 to 2018). Market participants might want to resell a share of this energy at a higher price. To do so, they would only have to bid the ARENH price since the day-ahead market is pay-as-clear (this would correspond to an additional offer in our model and the availability of nuclear plants would then be reduced accordingly). Sensitivity analyses can be performed in order to assess how the ARENH price, the total ARENH volume, or the percentage of ARENH resold could impact prices on the French day-ahead market.

4.4.4 Discussion

Impact of unexpected events

The occurrence of an unexpected event can influence selling and procurement requirements. Here we assess the impact on simulation errors due to temperature anomalies

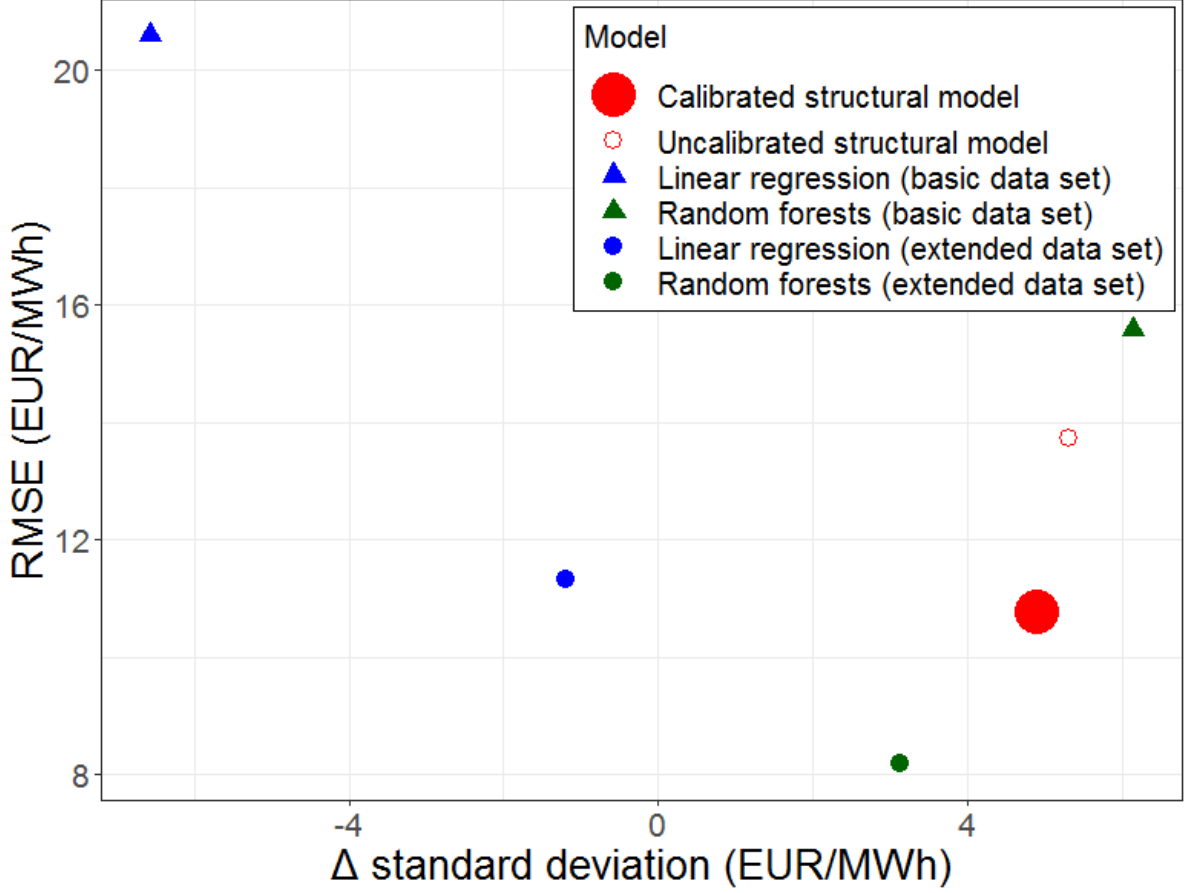


Figure 4.7: Performance comparison of benchmarked models

(i.e. the difference between the observed atmospheric temperature and the long-term average) and the impact of unexpected unavailability of production units. The hourly temperatures are spatially averaged for the French bidding zone in order to consider a single representative temperature for each hour. Regarding the production unavailability published on the ENTSO-E transparency platform, we distinguish between planned outages (e.g. maintenance) and forced outages (e.g. plant failure), where only the second category is considered as being unexpected. The evolution of unavailability is shown in Figure 4.8.

We perform regressions between the prediction errors of simulated prices on the out-of-sample dataset and our variables quantifying unexpected events using the Loess procedure [115] (i.e. locally estimated scatterplot smoothing). For the univariate cases, a coefficient of determination R^2 of 0.10% is obtained for the temperature anomalies and 1.23% for the impact of forced outages. In the bivariate case, the coefficient of determination is not improved compared to considering only forced outages ($R^2 = 1.14\%$). The results for forced outages are presented in Figure 4.9.

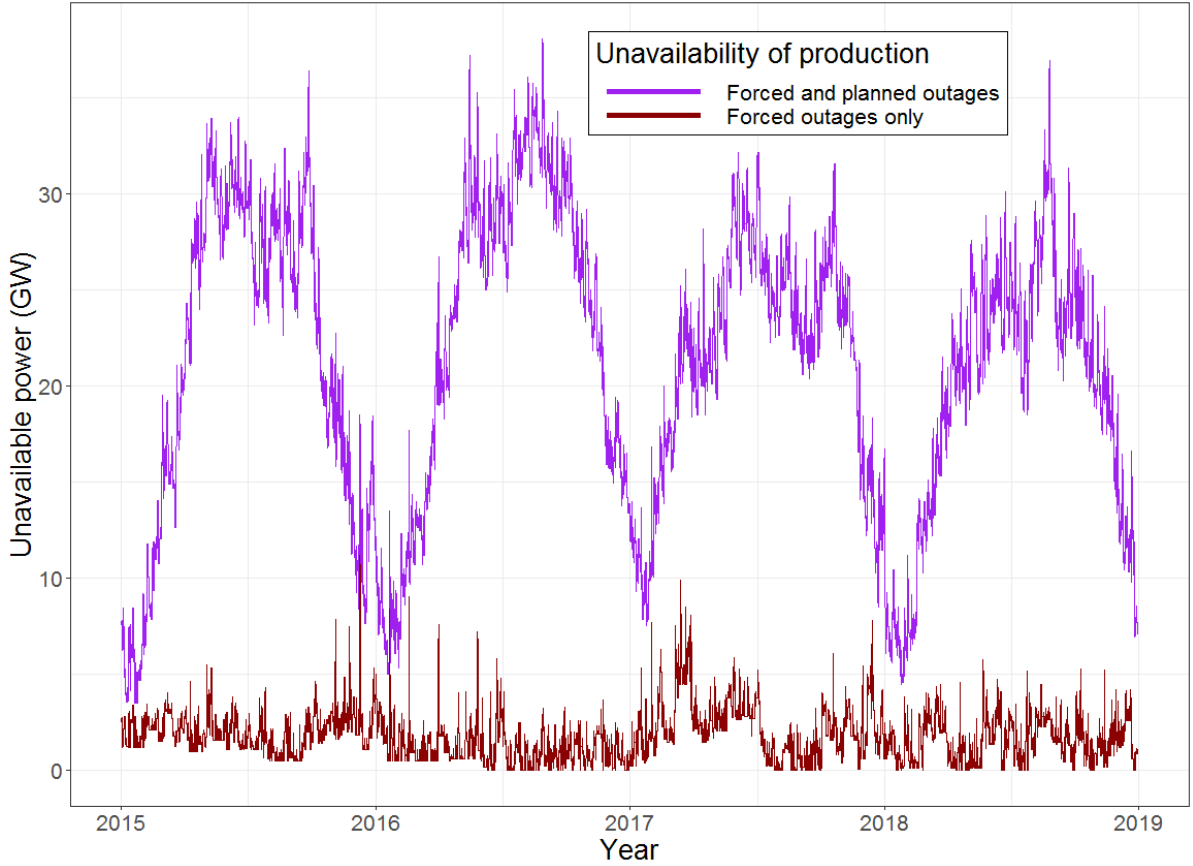


Figure 4.8: Unavailability of production units

General remarks

The case study illustrates how the method can be implemented to simulate prices on the French day-ahead market. The model makes it possible to simulate price dynamics and calibration improved performance. Exploiting fundamental data proves interesting to simulate these prices over the long run. We note that the French market is characterized by a dominance of nuclear power plants operated by a single electricity producer (EDF). This makes the French case particular and further analyses of other markets could be the purpose of further work.

Moreover, the model captures the main characteristics of price dynamics but underestimates the prices in case of spikes. The formulation of sell order prices could be extended to include non-linear price increases in case of limited supply margins. More generally, the procedure to estimate model parameters can be extended, for example to consider additional datasets (real supply curves) and to finely tune the importance given to the dynamics of historical prices during training phase, as we will see in the next chapter.

Finally, for prospective studies, scenarios could be generated to assess the potential

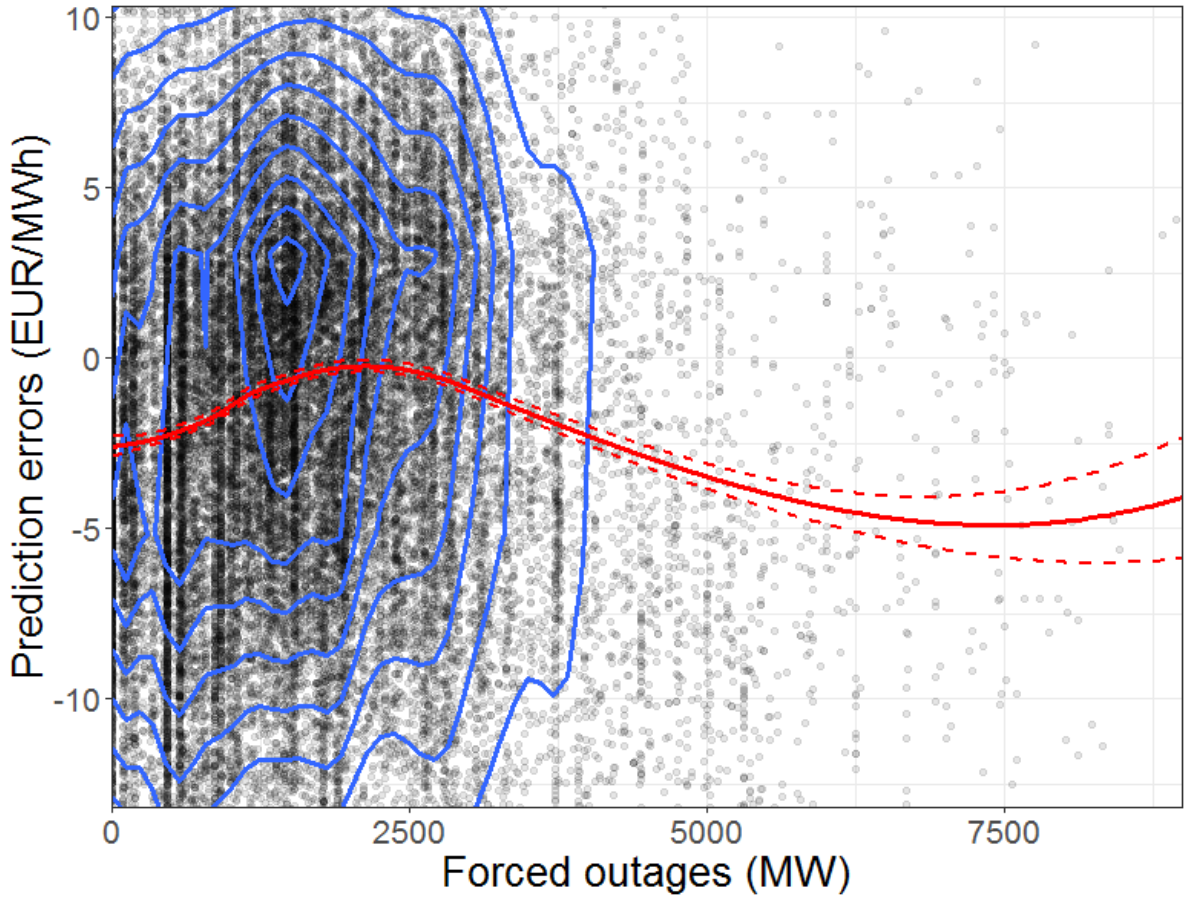


Figure 4.9: Impact of forced outages on prediction errors. Hourly values (black dots), their corresponding densities (blue lines) and locally estimated scatterplot smoothing, i.e. Loess (the estimation is shown by the solid red line and the 95% confidence interval by the dashed red lines)

impact of the evolution of fundamentals, such as the demand, energy mix or fuel and emission costs. For the French market, an evolution of the energy mix could be considered in light of the studies made by RTE to reach carbon neutrality in 2050 rather by mainly developing nuclear power, or by significantly increasing the share of renewable energy, or a balanced combination of both. Results for this distant future have to be studied critically, but analyzing simulated prices during the upcoming transition years leading to these significant changes can be informative for decision-makers. For detailed studies at the European level, scenarios could for example be based on the Ten-Year Network Development Plans (TYNDP) published every two years by ENTSO-E. When using scenarios, some price-sensitive factors, such as flows in interconnections and storage usage, should be made part of the optimization problem. The aim here is to ensure that these variables remain internal to the model and thus become dependent on spot prices. To do so, multiple bidding zones can be simulated at once in order to compute

interconnection flows with the model. Moreover, the optimization problem presented by Eq. (4.1) could be enhanced with additional constraints in order to actively manage storage assets, so that storage management also becomes internal to the model (such constraints typically account for capacity, state of charge, efficiency and ramp limits of storage assets). Such extensions will be explored in the next chapter.

4.5 Conclusions

Long-term simulations of wholesale electricity prices help to support investment and policy decisions. Structural models of day-ahead markets, which enable to account for both the price formation mechanism and techno-economic constraints of the power system, are commonly used in the industry for prospective studies. These models are built upon theoretical foundations and make it possible to find an optimal price corresponding to the maximization of social welfare. Nevertheless, empirical testing of models contributes to verify that the theory reliably accounts for the observations, and data-driven approaches can help to reduce the gap between theory and practice. While many models developed for electricity price forecasting are presented in the academic literature, methods specifically designed for long-term considerations that propose to combine optimization and statistical approaches are less common. In this chapter, we proposed a model which meets the aforementioned criteria. Having a bottom-up approach with a high granularity starting from production units for the supply side allows us to introduce a detailed and modular formulation of the market orders, which is suitable to properly leverage the power system and market data. We introduced a calibration method of structural model parameters capable of exploiting real data in a differentiated fashion by considering the production classes separately. The method has been validated using data relating to the French market from 2015 to 2018. The global price dynamics, especially the hourly and weekly variations, has been captured by the calibrated model on out-of-sample tests. However the amplitude of price spikes is underestimated and could be the focus of future work for the more specific formulation of the orders' prices. The case study shows that the calibration of parameters using real data improves the accuracy of the simulations. Moreover, the calibrated model reaches performances on historical data that are close to the ones obtained using purely data-driven methods, while also fulfilling our requirement regarding the possibility to explicitly model changes in the electricity mix or market design. In the next chapter, some advanced parameter estimation methods will be introduced and some limitations of the model with regards to its use for prospective studies will be mitigated.

Chapter 5

Extensions and applications of the model

Contents

5.1	Further studies of the estimation method	82
5.1.1	Introducing variability into the cost function	82
5.1.2	Leveraging information from observed supply curves	84
5.1.3	Sensitivity analysis of initial conditions	86
5.1.4	Discussion	87
5.2	Model validation for prospective studies	88
5.2.1	Non-regression testing	88
5.2.2	Ramp constraints	88
5.2.3	Active storage management	90
5.2.4	Multi-zone simulation	91
5.2.5	Discussion	94
5.3	Model applications	96
5.3.1	Scenarios for prospective studies	96
5.3.2	Prospective studies	98
5.3.3	Discussion	101
5.4	Conclusions	102

In the previous chapter, we proposed a method to simulate electricity prices on day-ahead markets. We observed that the method could be improved in two notable ways. First, further considerations can be given to the estimation algorithm in order to simulate prices more accurately. Second, the formulation of the market optimization problem itself can be extended in order to propose a global methodology that mitigates some remaining limitations with regards to the use of the model for prospective studies. In this chapter, we will explore and discuss these two extensions of the model.

This validation of the extended model on historical data finally allows us to propose some applications. More precisely, the model is used to conduct scenario-based prospective studies regarding possible evolutions of electricity prices; these prospective studies are the object of the last section of this chapter.

5.1 Further studies of the estimation method

The outcome of the parameter estimation method depends on the way the regression is performed and the choice of data. We propose three complementary studies to explore how alternative estimations ultimately impact the final out-of-sample simulation results. In the first study, we extend the regression by introducing weights, which influence of the variability of the simulated time series of prices. In the second study, we introduce a method to create additional synthetic datasets that account for information embedded in real supply curves. In the third study, we analyze to which extent the initial assumptions regarding the merit order impact the simulation outcome.

5.1.1 Introducing variability into the cost function

The root mean square error is known to give results that are smoother than observations. We therefore propose a modification of the cost function that implicitly approaches a minimization of $RMSE_z + \gamma \Delta sd_z$. We propose to introduce weights $(\omega_{t,z})_{t \in T}$ in the cost function as a variability hyperparameter of the model:

$$RMSE_z = \left(\frac{1}{T} \sum_{c_0 \in C} \sum_{t \in T_{c_0}} \omega_{t,z} (\alpha_{c_0}^T \mathbf{X}_{t,i^*(t)} + \lambda_{t,i^*(t)}^{IC} + \beta_z^T \mathbf{Y}_t - \Pi_{t,z})^2 \right)^{\frac{1}{2}} \quad (5.1)$$

such that the price barycenter is not modified, i.e. so that :

$$\frac{1}{|T_{c_0}|} \sum_{t \in T_{c_0}} \Pi_{t,z} = \frac{1}{\sum_{t \in T_{c_0}} \omega_{t,z}} \sum_{t \in T_{c_0}} \omega_{t,z} \Pi_{t,z} \quad (5.2)$$

By giving more weight to hours with extreme low or high prices, we thus decrease the value of Δsd_z .

We apply this method to our case study and show in Figure 5.1 how tuning the variability hyperparameter of the model during calibration can effect the out-of-sample

performance as measured by the $RMSE$ and Δsd metrics. This hyperparameter is equal to the sum of the weights ω_i introduced in Eq. (5.1), which are added to the minimal and maximal price values used for estimating the parameters α_{c_0} for each production class. We ensure that the weights are added in a balanced way so that the mean price of the set remains unchanged.

In other words, weights are used during the training phase when we estimate the α parameters of the sell prices. Let us recall that datasets are split according to the class of the production unit considered to be marginal. For each production class, when considering only the relevant time steps according to the marginality criterion, we have a subset of observed prices and the corresponding explanatory variables. We identify the time corresponding to the lowest and highest market prices and we assign extra weights only to these two time steps. We call the sum of added weights the variability hyperparameter. The weights are shared between these two points so that the average of observed prices remains unchanged after assigning these new weights.

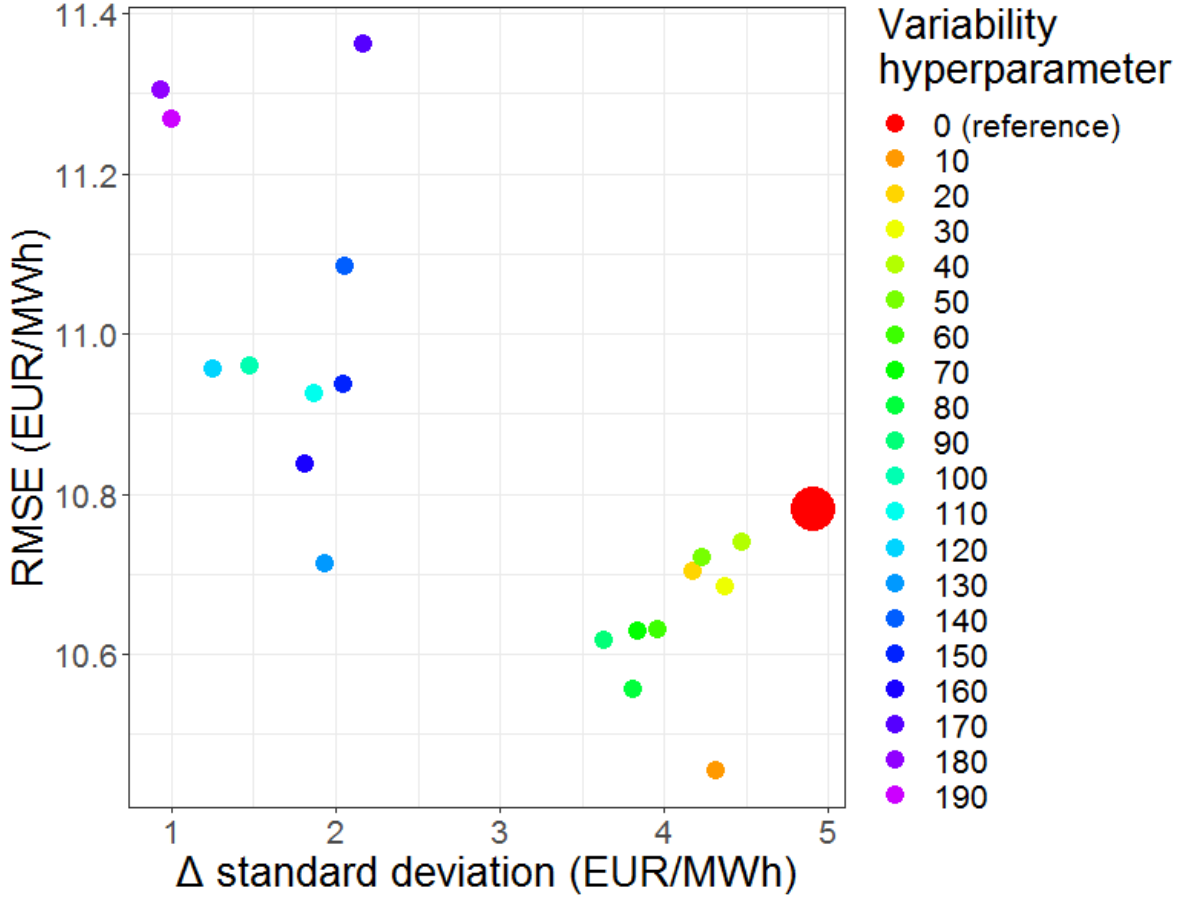


Figure 5.1: Out-of-sample performance for different values of the variability hyperparameter

We observe in Figure 5.1 that hyperparametrization can help to improve the dynamics

of the simulation (i.e. the variation of the simulated prices is more similar to the variation of the observed prices) at the expense of lower global accuracy as measured by the RMSE.

5.1.2 Leveraging information from observed supply curves

The estimation algorithm makes use of observed market prices, implying that only the price information given by the at-the-money order is exploited for each time step. However, the supply curves bring valuable information that has not been taken into account in our method so far. As day-ahead auctions are generally blind (i.e. the identity of market participants is not made public, and it is not even possible to determine which of the production units is actually participating in a market), information cannot simply be retrieved by matching orders to production units. We propose a methodology to exploit the information contained in observed supply curves $\mathcal{S}_{t,z}$, which translates into an extension of the set of observed prices, called synthetic prices. A small deviation, δ , is applied to the real market clearing volume $V_{t,z}$ in order to generate synthetic prices $\Pi_{t,z}^\delta$ so that:

$$\Pi_{t,z}^\delta = \mathcal{S}_{t,z}(V_{t,z}^\delta) \quad (5.3)$$

The process is illustrated in Figure 5.2.

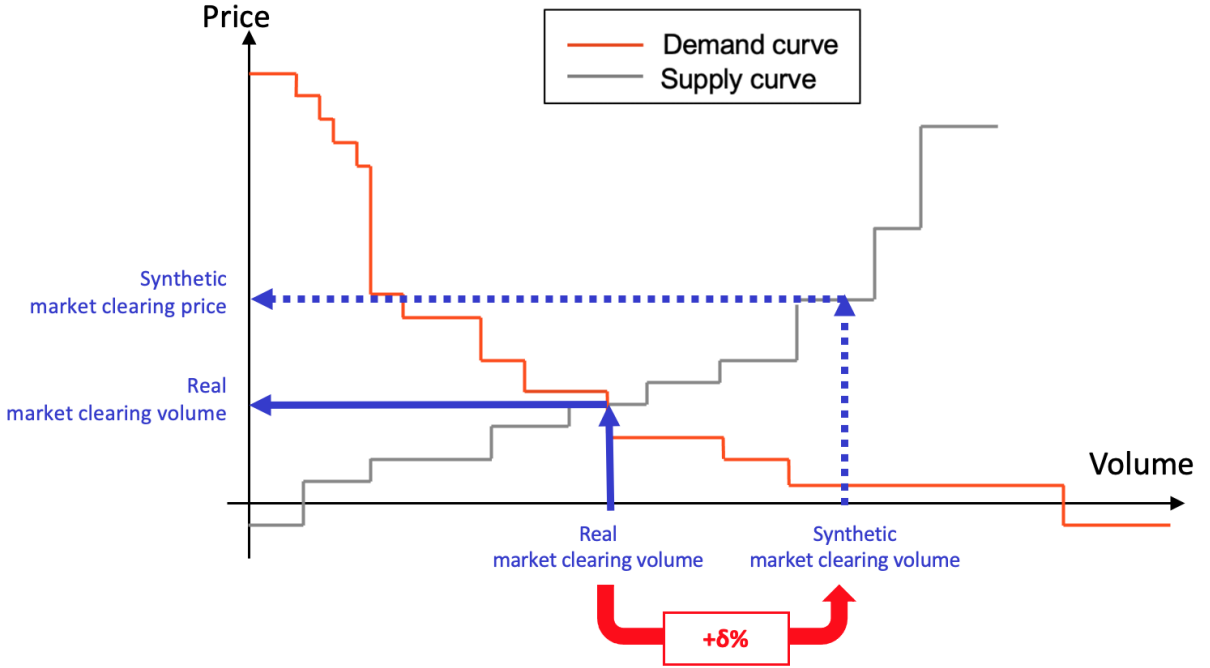


Figure 5.2: Data augmentation principle to generate synthetic prices from supply curves

Each relative variation δ creates a new synthetic dataset used to estimate model parameters (i.e. the datasets are only used for in-sample training). As we want our synthetic prices to reflect realistic counterfactual cases, we consider that the relative

variation of the hourly volumes is caused by the same variation of residual demand. Consequently, the supply margin $\mathcal{M}_{t,z}$ is also modified in the synthetic dataset since it depends on the residual demand. Finally, the synthetic dataset is composed of the new prices $(\Pi_{t,z}^\delta)_{t \in T}$, the new residual demand, and the price drivers $(\mathbf{X}_{t,i}^\delta)_{t \in T}$, where all variables of $\mathbf{X}_{t,i}$ remain unchanged except $\mathcal{M}_{t,z}^\delta$.

We now apply this estimation method to our case study. Due to data availability issues, we applied the presented technique on the supply curves only for the year 2015. We tested the performance on the year 2016 where more price spikes are observed. In the reference case, the spikes are not fully captured, which lead to degraded performances compared to the other simulated years as measured by our two metrics. As training sets, we consider 12 additional synthetic years with a relative variation of hourly traded energy volumes from -3% to +3% by steps of 0.5%. In Figure 5.3, we represent the performance obtained when averaging the simulated prices for 2016 obtained using the models trained with different synthetic datasets (e.g. the case "demand variation 1.0%" is obtained by averaging the outputs obtained with the models trained using the datasets -1.0%, -0.5%, 0.0%, 0.5% and 1.0%).

Contrary to the results obtained with different variability hyperparameter values, the data augmentation approach tends to increase the global accuracy at the expense of less realistic price variations.

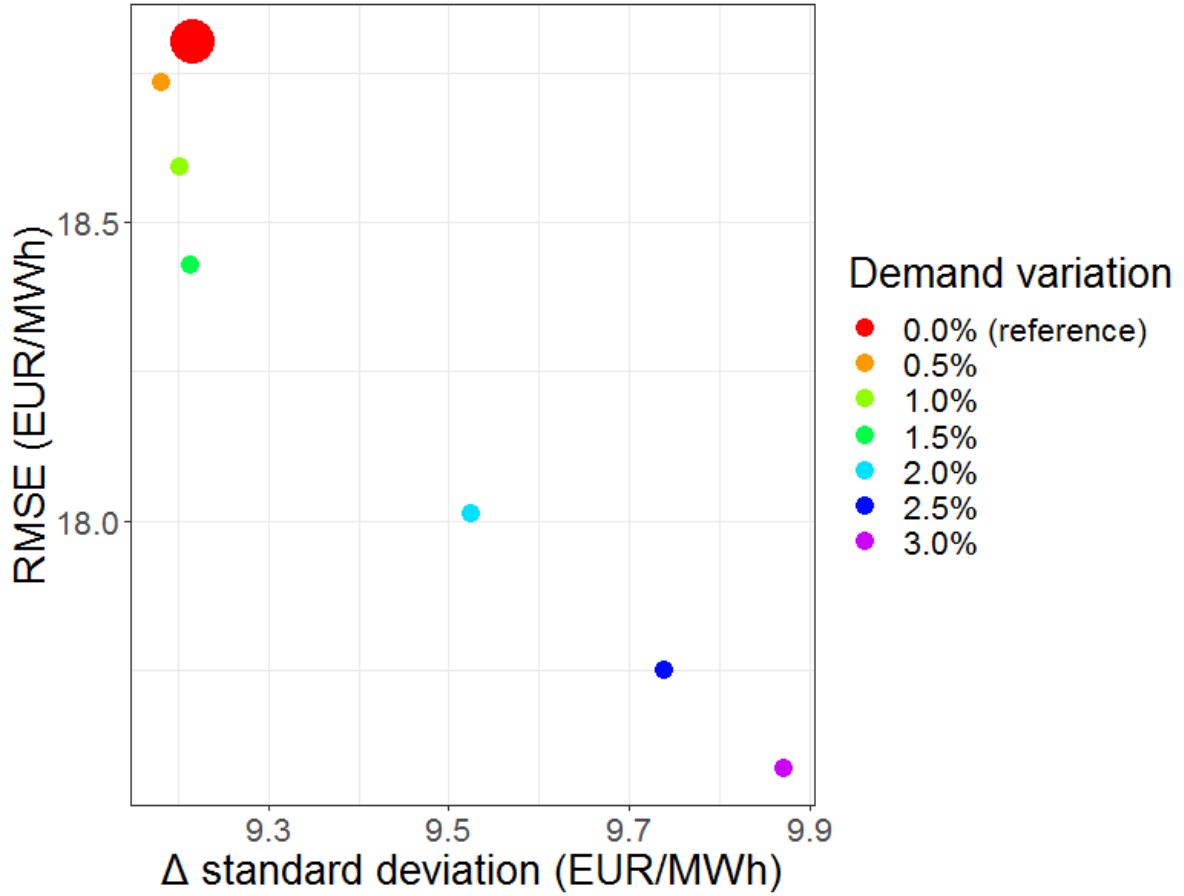


Figure 5.3: Out-of-sample performance for different augmented datasets (training set: 2015, testing set: 2016)

5.1.3 Sensitivity analysis of initial conditions

By initializing the $\alpha_{c,0}$ parameters to a different value for each production class, we make assumptions regarding the merit order (i.e. the energy is supplied in priority by the production class with the lowest cost, then the second lowest, and so on). This initial choice has an impact on the parameter estimation procedure. The reason is that the identification of the marginal unit for each hour, which results from the solving of the market clearing problem, directly depends on the estimated merit order. Thus, we perform a sensitivity analysis to investigate how alternative hypotheses regarding the initial merit order affect the final result of the calibration. To do so, we initially assign five distinct constant prices to our five production classes (i.e. nuclear, hydro, gas, coal and oil) and consider all possible permutations, then we analyze the initial merit order for the best decile in terms of $RMSE$ (see Figure 5.4).

The results illustrate that the model initialization should be consistent with common economic assumptions (e.g. nuclear to constitute a less expensive baseload and oil for

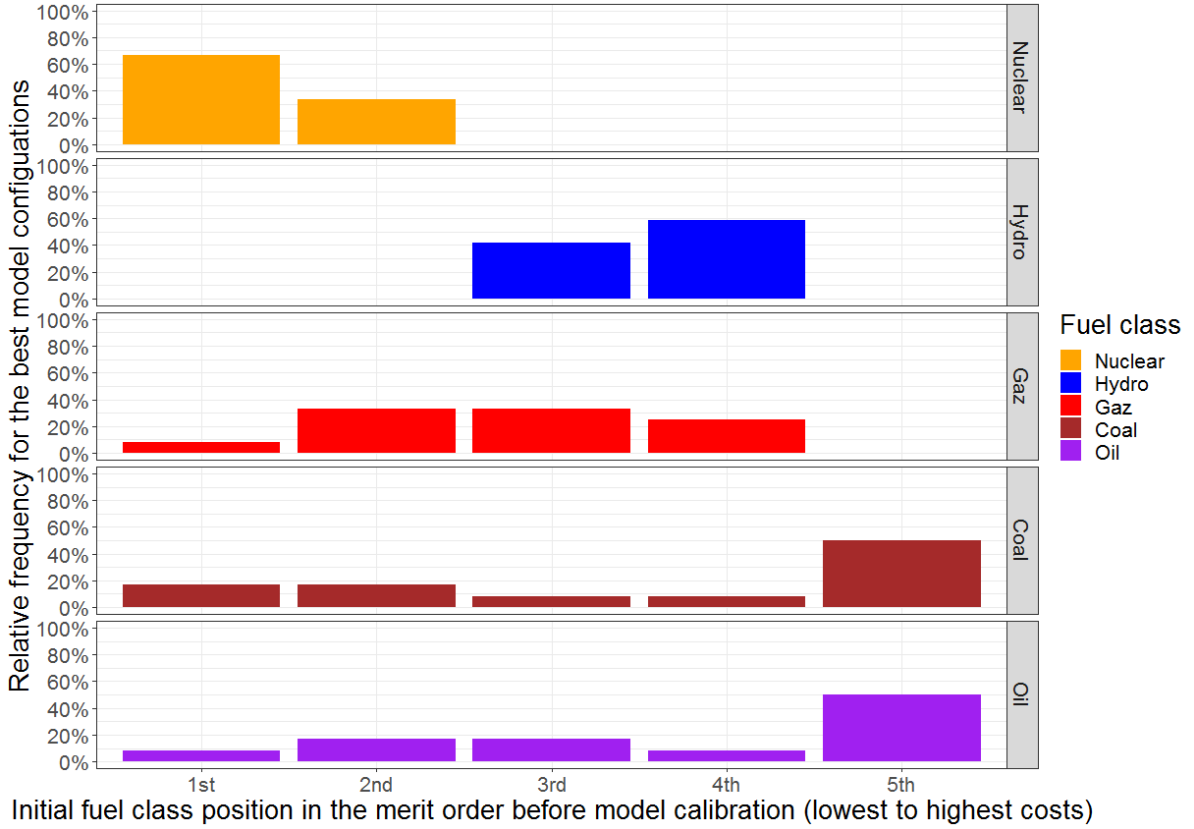


Figure 5.4: Best initial merit orders (first decile)

more expensive peaks). The place of hydro in the merit order is likely due to the strategic use of stock during high-price hours. Therefore, the initialization of the merit order during the training phase must be considered carefully by the modeler.

5.1.4 Discussion

The advanced calibration techniques proposed in this section impact model performance over the whole test period, but a more detailed analysis of the simulation outcomes reveal no significant impact during hours corresponding to price spikes. This is likely due to the linear formulation of the marginal cost proxy: the presented calibration techniques allow us to vary the importance of the points in the training dataset (hyperparametrization) or to consider alternative training datasets (data augmentation), without modifying the marginal cost proxy formulation. Thus, the training phase enables us to estimate parameters that provide a relevant compromise that is relevant for most hours, but still underperform during price spikes. Additional model extensions regarding the parametrization of orders, for example by introducing non-linear relationships between price offers and market drivers (piecewise-linear, quadratic, etc.), could be the subject of future work.

5.2 Model validation for prospective studies

In this section, our model will be incrementally complexified in order to fulfil additional requirements for prospective studies.

5.2.1 Non-regression testing

To conduct the remaining studies presented in this chapter, we chose to use Pyomo, a Python-based optimization modeling language. This framework ensures that the model can be extended in a robust fashion.

Before considering extensions for the model and its usage, we first want to verify the results obtained with Pyomo are similar to those previously obtained with R when running under similar conditions (data and modeling assumptions). As we ultimately want to extend the model and introduce additional factors, this verification is not performed through an exact transposition of the R code into Python, but in a way that allows us to subsequently built upon the verified models.

Two main adjustments are made to ensure that the modeling can be generalized. First, we chose to preferably make use of ENTSO-E data retrieved from the transparency platform. Using a common source for all European countries ensures that all considered countries can be modeled with similar levels of accuracy and that discrepancies between data sources are kept to a minimum. Second, we do not consider generation units individually but rather the aggregated availabilities by production types. The main advantage of using aggregated production types instead of individual production units is the reduction of the numbers of variables. A drawback of this approach is that differences between units have to be considered through a alternative mechanism in order to ensure that the dynamics of time series of day-ahead prices is properly captured. This is why we now consider a new additive term for the price of a supply order, which is directly proportional to the energy delivered by a production class.

As illustrated in Figure 5.5, we are able to replicate similar simulation for the price dynamics. In this illustration, the year 2015 has been chosen for the training dataset and the test is performed on 2018. Previously, we obtained $RMSE = 11.9$ ($RMSE = 12.2$ with R) and $\Delta sd = 4.9$ ($\Delta sd = 2.5$ with R). Although the results are not strictly identical due to differences with regards to model implementation, we consider that the newly obtained simulations validate the this non-regression test.

When then simulate our reference case to which model extensions will be compared against. To do so, using the Pyomo model, we train the model using the dataset corresponding to the year 2015 and we test for the years 2016, 2017 and 2018. The results of the simulation are displayed in Figure 5.6. We obtain $RMSE = 13.5$ and $\Delta sd = 6.6$.

5.2.2 Ramp constraints

The dynamics of electricity prices depends on the flexibility of the power system and its ability to respond to temporal variations. Generation units cannot instantaneously ramp up production from zero to nominal capacity, nor can they ramp down unconstrained.

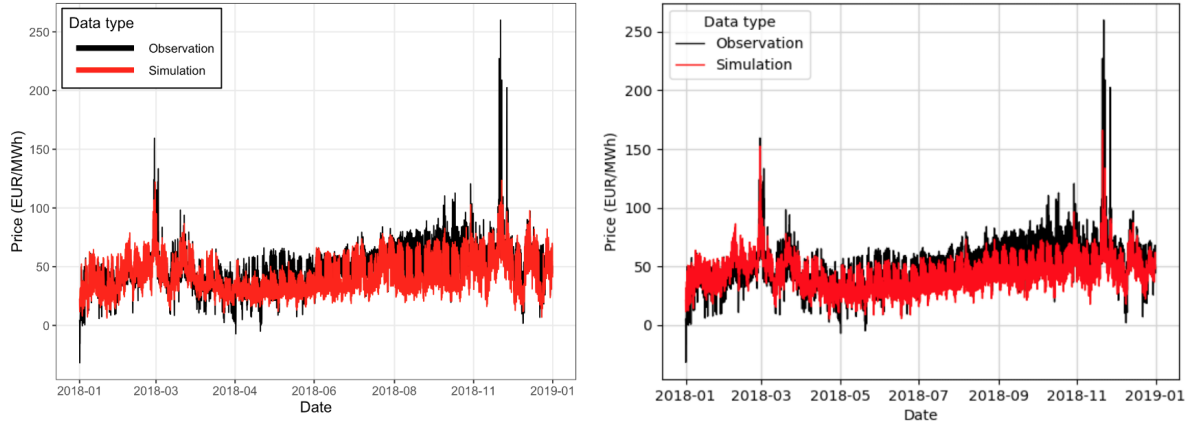


Figure 5.5: Comparison between results obtained in the previous case study with R (left) and the results obtained under similar conditions with Python (right)

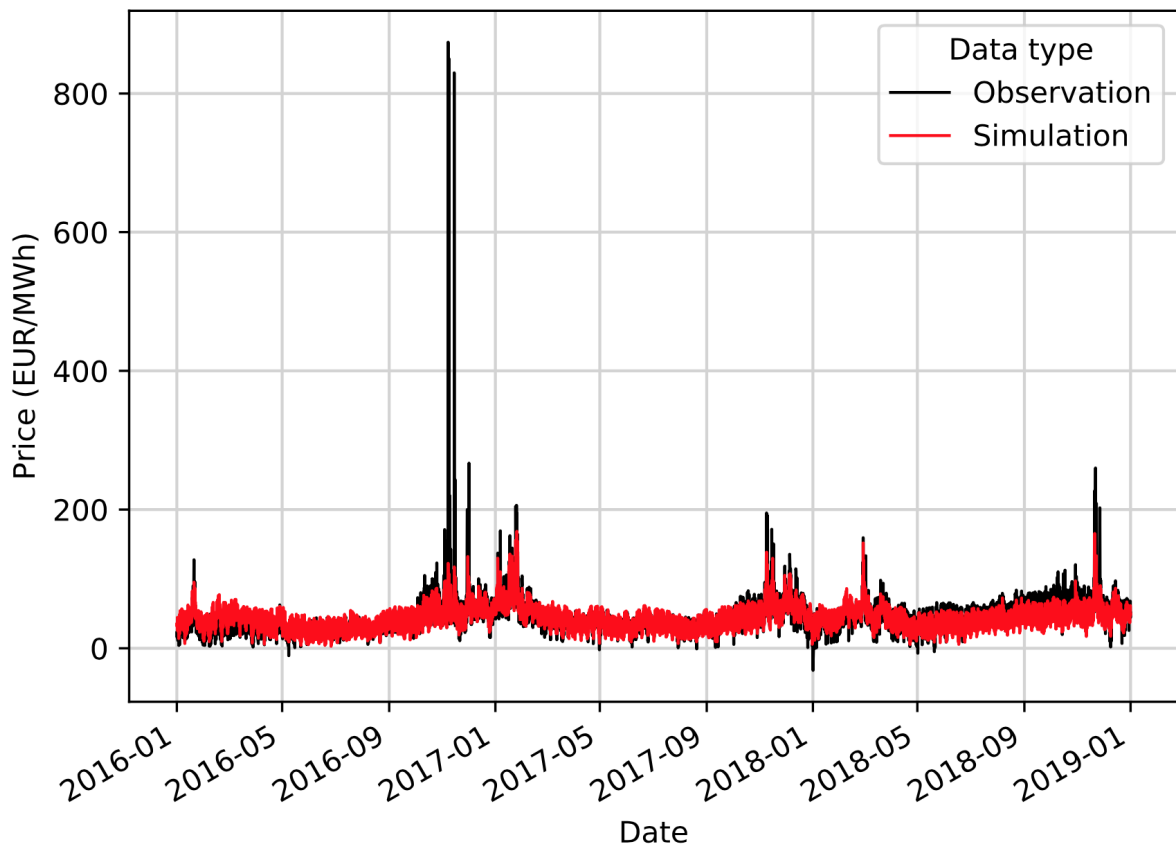


Figure 5.6: Simulation of French day-ahead prices (reference case)

These limitations are mainly due to thermal and mechanical limits. Thus, the value

of these limits are specific to each power generation technology and, by extension, the limits depend on fuel type.

Market participants modulate production as needed and chose to activate the available flexibilities within their portfolio. While the decisions concern individual power plants, global effects can be observed when considering the aggregated production by fuel type as they reflect both economic and technical constraints. By considering historical data, we can then estimate the maximal ramping up and ramping down capacity. This estimation has been done using data related to the French power system for the period 2015-2018. We chose to keep a single representative ramp limit by technology based on this numerical estimation. To ensure that the model is relevant for prospective studies in which the installed capacity does not remain constant, we represent these constraints as hourly power variation relative to the installed capacity. The obtained values are summarized in Table 5.1.

Production class	Ramp limits (% of installed capacity per hour)
Hydro	80
Nuclear	20
Gas	85
Coal	65
Oil	50

Table 5.1: Ramp limits by production class

The additional ramp-down and ramp-up constraints are respectively added to the optimization problem as follows:

$$-r_{c(i)}^{down} \cdot P_i^{max} \leq P_{t,i} - P_{t,i-1} \quad \forall (t, i) \in T \times I \quad (5.4)$$

$$P_{t,i} - P_{t,i-1} \leq r_{c(i)}^{up} \cdot P_i^{max} \quad \forall (t, i) \in T \times I \quad (5.5)$$

After running the simulations, we observed that including ramp constraints does not significantly impact on the simulation outcome: for the out-of-sample test period, we now obtain a $RMSE = 13.4$ ($RMSE = 13.5$ without the ramp constraints) and $\Delta sd = 6.8$ ($\Delta sd = 6.6$ without the ramp constraints). This modification alone is thus insufficient to properly capture the price spikes, as the maximum simulated price change is below 1% (EUR 168.7/MWh without ramp and EUR 169.7/MWh with ramps).

5.2.3 Active storage management

In the case study presented in the previous chapter, the contribution of pumped storage hydro power plants was assumed to be known. The real aggregated production and consumption values were subtracted from the demand in order to compute a residual demand. As the use of such plants depends on market prices, storage management should

instead be an internal model variable for prospective studies. To do so, the contribution of pumped hydro storage is not included in the residual demand anymore and additional constraints are added to the optimization model as follows:

$$S_{t,k}^{level} = S_{t-1,k}^{level} \cdot (1 - d_k) + S_{t,k}^{in} \cdot \eta_k^{in} - S_{t,k}^{out} \cdot \eta_k^{out} \quad \forall (t, k) \in T \times K \quad (5.6)$$

$$S_{t,k}^{level} \leq C_k^{max} \quad \forall (t, k) \in T \times K \quad (5.7)$$

$$S_{t,k}^{in} \leq P_k^{max} \quad \forall (t, k) \in T \times K \quad (5.8)$$

$$S_{t,k}^{out} \leq P_k^{max} \quad \forall (t, k) \in T \times K \quad (5.9)$$

The first constraint correspond to the recursive definition of the stock level $S_{t,k}^{level}$ for the technology k at time t . The stock level depends on the its value at time $t - 1$ (including the spontaneous dissipation at each time step by a factor $1 - d_k$), and of the input or output flows ($S_{t,k}^{in}$ and $S_{t,k}^{out}$) corrected by the conversion efficiency (η_k^{in} and η_k^{out}). The second and third constraints reflect the power limits of the technology. The fourth constraint reflects its capacity limit in term of energy.

In this study, we only consider the contribution of pumped storage hydro power plants to grid-scale energy storage. All storage units are aggregated under a single variable similar to what has been done for production units. Thanks to this problem formulation, additional technologies (e.g. hydrogen) can easily be added to the model if needed. The technical parameters used to model the pumped hydro storage for the year 2015 are summarized in Table 5.2.

Variable	Value
P^{max}	4965 MW
C^{max}	34755 MWh
d	0
η^{in}	87%
η^{out}	87%

Table 5.2: Technical constraints for the aggregation of the French pumped hydro power plants in 2015

Even though this model variant is less reliant on historical data, the out-of-sample performances are not degraded and the model still performs well ($RMSE = 13.5$ and $\Delta sd = 6.9$).

5.2.4 Multi-zone simulation

The final validation test consists in ensuring that the model is also suitable to simulate prices even when an extended spatial scope is considered. In the case study, the market prices in neighboring bidding zones were supposed to be known and only the flows in interconnectors were simulated. Here, we want to simultaneously compute prices in multiple zones instead. We keep the same train-test split (i.e. 2015 for training and 2016

to 2018 for the test), but we consider three countries: France, Belgium and Germany. The time series of prices for these three countries are presented in Figure 5.7. We still proceed by incrementally adding complexity to the model, which means that ramp constraints are accounted for and that storage is still actively managed.

This time, we observe significant differences in terms of simulation results. Since we no longer rely on historical prices for neighboring countries, which are correlated to the French ones, the performance is globally degraded ($RMSE = 17.0$) and we observe that the amplitude of price spikes is even less accurate than before. Nevertheless, in this case, the model still tend capture the dynamics of prices and the metrics even improves ($\Delta sd = 0.7$). The study of the Belgian ($RMSE = 23.0$, $\Delta sd = -0.0$) and German ($RMSE = 20.0$, $\Delta sd = -2.2$) cases reveal that the model converges but that additional work would be needed to improve the performance for extreme values (both for price spikes and negative prices).

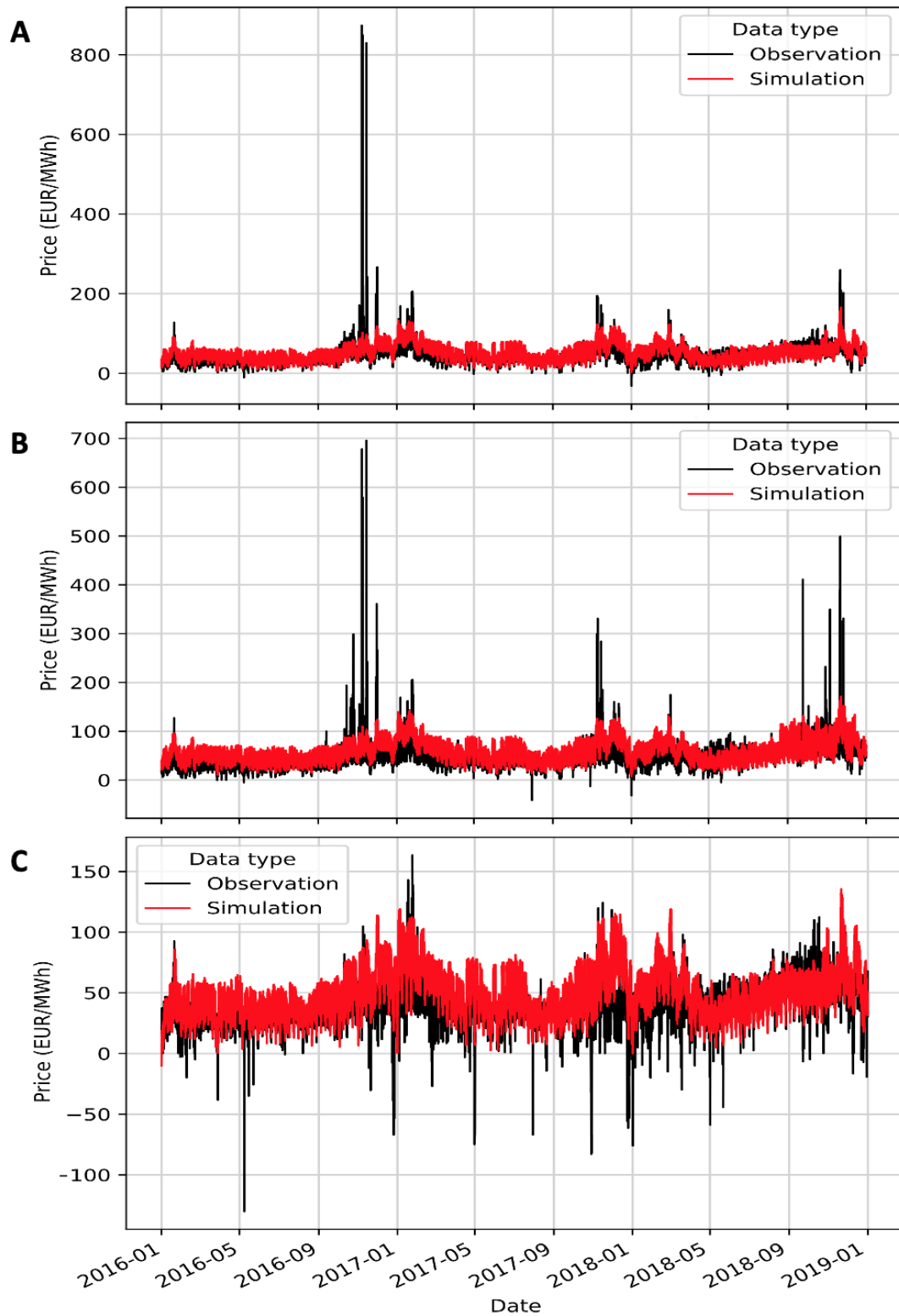


Figure 5.7: Multizone case: (A) French prices, (B) Belgian prices, (C) German prices

5.2.5 Discussion

In this section, we explored how the model can be enhanced to be more representative of the power system constraints and the functioning of the internal European energy market. While relying on less historical data tend to decrease forecast accuracy, the overall dynamics of prices is still captured by the model. The added complexity comes at the cost of more computation time and additional data requirement. The time required to generate the model with Pyomo and to solve it is reported in Table 5.3. The simulation were performed on a laptop with the following hardware characteristics: MacBook Pro (Retina, 13-inch, Early 2015), processor 3,1 GHz Dual-Core Intel Core i7 and Memory 16 GB 1867 MHz DDR3. In terms of software, the laptop run with macOS Big Sur (v11.4), Python v3.8.8 was used to implement the model, and the optimization problem was solved using the solver Mosek v9.2.35. Table 5.3 reveals that adding ramp constraints to the single zone case increases the computation time by +7%. Then, actively managing storage increase the time by an additional +47%. Finally, simulating two more bidding zones as for consequence a +263% increase in computation time compared to the previously simulated case. Thus, this added complexity has an important impact in terms of computation time and, most importantly, the increase is not proportional to the number of bidding zones. Solving this problem at scale is not a problem in theory (the modeling principles are identical for all 3+ bidding zones cases), but properly simulating all European bidding zones would require to consider alternative computing frameworks, for example by making use of distributed computing.

Test case	Computation time
Single zone (reference case)	39.4 seconds
Ref + Ramp constraints	42.0 seconds
Ref + Ramps + Active storage management	61.6 seconds
Multizone with all constraints	223.4 seconds

Table 5.3: Computing time to generate and solve the optimization problem for different test cases

The relative performance of the models can be compared by visualizing the distribution of prices (Figure 5.8) and the evaluation metrics (Figure 5.9). These figures highlights the similarities between the reference case, the case with ramps and the case with ramps and active storage management. For the multizone case, the greater price dispersion, which lead to an improvement of the Δsd metrics, can interpreted in light of the fact that units from foreign countries can set the market price so that more variable prices are simulated.

Some operational constraints of the power system are still not represented in this version of the model. Scaling down to the generation unit level would allow to include the start-stop constraints that have an impact on flexibility and thus on price formation. Nevertheless, achieving to model the correct dependency between time steps at such a fine granularity might not be achieved in practice and the aggregated ramp constraints can be considered as a proxy for all temporal constraints. Similarly, further extensions can be

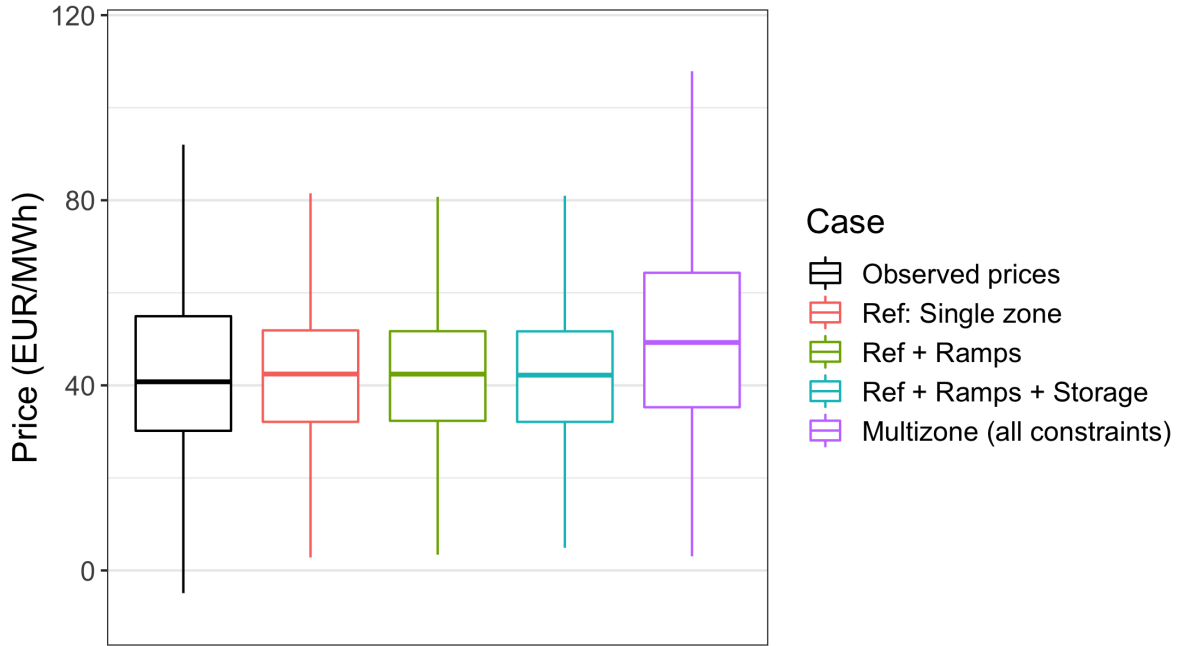


Figure 5.8: Boxplot: Comparison of the simulated cases

brought to the model with regards to grid operation (reserve, N-x security constraints, line congestion, etc.). Such extensions would help to bring the model closer to the representation of flow-based market clearing. Without further modeling tasks, reserve requirements can be approximated by modulating the values of residual demands and/or availability of power plants. Conversely, modeling the power grid more accurately would be more challenging; finding the right level of precision is not straightforward. Not all lines and cables can be represented due to computing constraints, thus some criteria to select the critical ones have to be defined. Also, the available capacity of grid assets depends on decisions taken by TSOs and collecting reliable data about such decisions is another issue.

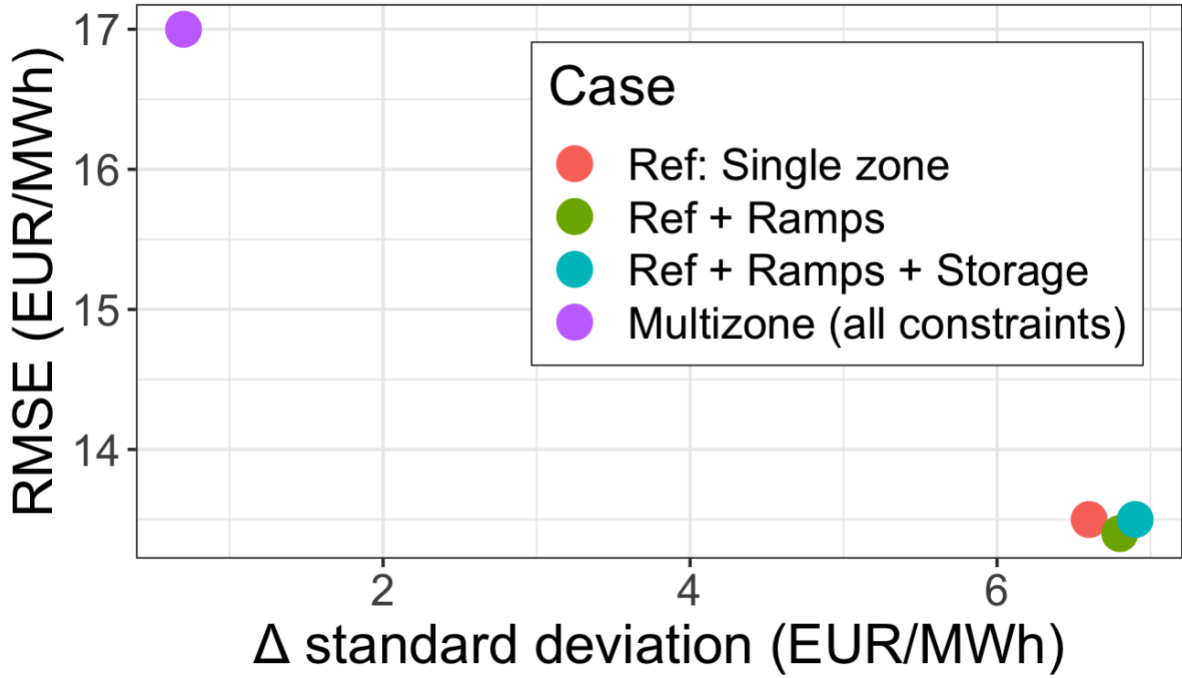


Figure 5.9: Metrics plots: Comparison of the simulated cases

5.3 Model applications

The proposed model can finally be applied to conduct prospective studies. In this section, we first present the scenario approach, then we run the simulations and discuss the implications of the price outcomes.

5.3.1 Scenarios for prospective studies

Benefits of the scenario approach

One common approach to forecast the future of power systems relies on generation expansion planning, which enables to find an optimal path for the development of new assets under specific constraints and a given time horizon. Although such optimization models are informative, in practice, the development of power systems does not follow a globally optimal trajectory as it is driven by multiple forces (decisions of market participants, policies, etc.). "What if" scenarios constitute an interesting alternative since they allow us to explore multiple possible futures corresponding to selected hypotheses. By varying input data in scenarios, uncertainty can also be accounted for thanks to Monte Carlo simulations (i.e. simulations are repeatedly performed using random sampling from input datasets). Here, we propose a few illustrative examples in order to highlight to which extent our modeling approach makes it possible to estimate the evolution of prices depending on future evolutions of the power system and commodity markets.

Factors to include in scenarios

Before building scenarios, let us recap the main factors that impact price formation and are accounted for in our model. We group them in the following categories: supply, demand, storage, grid and commodity prices. In terms of supply, we account for the installed capacity and availability for assets using different fuel types. The flexibility constraints and hydro stock constraints are modeled as well. Regarding demand, or more precisely, the residual demand, we not only account for the evolution of consumption but also for the contribution of weather-driven renewables. Concerning storage, we have shown that various technologies can be modeled, such as pumped-hydro and hydrogen, each having its own technical characteristics in terms of capacity, power and efficiency. With regards to the grid, our model contains less details as internal constraints do not constitute a main price-driver in zonal markets; nevertheless, the interconnection capacity between bidding zone enables to partially account for grid congestion. Last but not least, market prices for commodities such as oil, coal, gas and greenhouse gases emissions are among the factors in our model that influence electricity price formation.

Building a reference scenario

The prospective studies presented in the next section are built according to common principles shared across all scenarios. We chose a time horizon of one year to account for seasonal effects. For each scenario, multiple yearly time series are simulated using the Monte Carlo approach. Two interpretations can be given to these time series: whether as multiannual simulations by considering a simple concatenation of model outcomes, or as ensemble forecasts that reflect the uncertainty stemming from the stochastic variables.

Static data Static data, i.e. variables for which the value remains constant over the one-year period, are derived from the historical datasets (e.g. installed capacity). For the static model inputs, we take the values for the year 2018 as reference. For the model parameters, we chose to leverage the whole historical dataset now that the model has been validated. Thus, we estimated the parameters for each of the four years of our historical dataset, then we took the average of the four years to run the prospective simulations.

Time-dependent data To build hourly time series that we can use as input data for our model, we consider two complementary methods: (i) generating yearly profiles and (ii) generating stochastic inputs thanks to resampling techniques. In both cases, our goal is to account for the seasonal variability of the time series. The aggregated national consumption and the commodity prices are generated using the profile approach. That is to say, the values are derived from the whole historical dataset by computing the mean hour by hour, leading us to 8760 different values as we exclude the 29th of February. For the availability factor of production units and for the load factor of weather-driven renewables, we chose to rely on the resampling approach. Using this method, we define a Monte Carlo case by randomly drawing from historical data. Similar to the profile

approach, we only draw values from the same hour of the year of the historical data. This sampling operation is repeated multiple times in order to define different cases within a scenario.

Alternative scenarios For the prospective studies, we propose to first consider the reference scenario developed with the methodology described above (a scenario comprise different variants, each corresponding to a Monte Carlo case). Then, we analyse alternative scenarios derived from the reference case. These alternative scenarios are built by applying multiplicative or additive factors on chosen variables according to our modeling hypotheses.

5.3.2 Prospective studies

Reference scenario

For the reference scenario, we consider a increase in electrical demand and installed renewable capacities for wind and solar, as it is the case for most net zero scenarios. Compared to 2018, the scenario represents a +20% increase in demand and an increase of +50% for wind and solar installed capacity. We simulate this scenario with 30 Monte Carlo cases. As it is not practical to display all simulation outcomes over a period of one year, we propose to visualize the results for the first week of June and the first week of December, as show in Figure 5.10 and Figure 5.11. We observe that we still have all seasonal variabilities (hourly, weekly and yearly) and that the Monte Carlo approach allow us to represent uncertainty since the different values of prices are obtained for each hour.

We also compare the distribution of simulated prices to our reference year of 2018. The comparison is represented in Figure 5.12 in which the mean hourly prices for all simulation is computed for the distribution. We observe, that our scenario lead to an overall increase of prices. In the absence of additional storage, the 50% increase in installed capacity of wind and solar do not put downward pressure on prices due to the 20% increase in demand. For the 2018 reference year, the share of variable renewable energy sources (vRES) in the total French electricity consumption was 16.8% and this value remained similar for the prospective studies with an average share of vRES equal to 15.7%, which suggests that additional investments efforts in vRES might be needed to counterbalance the projected increase in electricity demand. In 2018, the median price was EUR 49.9/MWh, the mean price was EUR 50.2/MWh and a standard deviation of EUR 18.5/MWh. For our prospective simulation in this reference case, we obtain: median price equal to EUR 66.5/MWh, mean price equal to EUR 69.8/MWh and standard deviation equal to EUR 23.7/MWh. We note that the increase variability stemming from more weather-driven renewables lead to more dispersed prices.

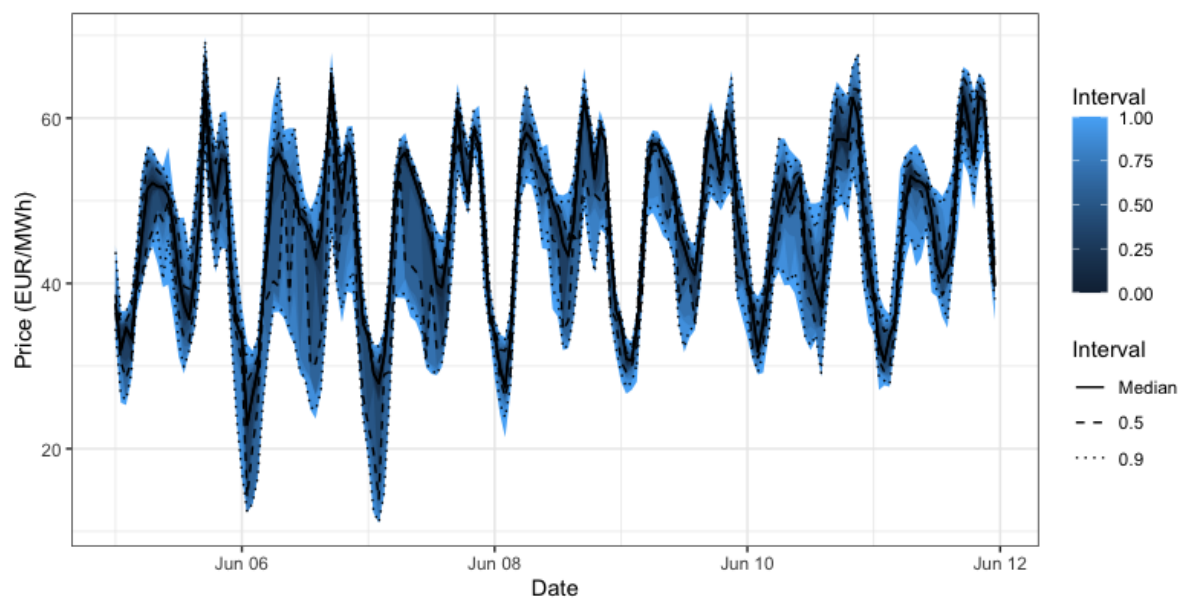


Figure 5.10: Monte Carlo simulations of prices on the French day-ahead market (reference prospective scenario, first week of June)

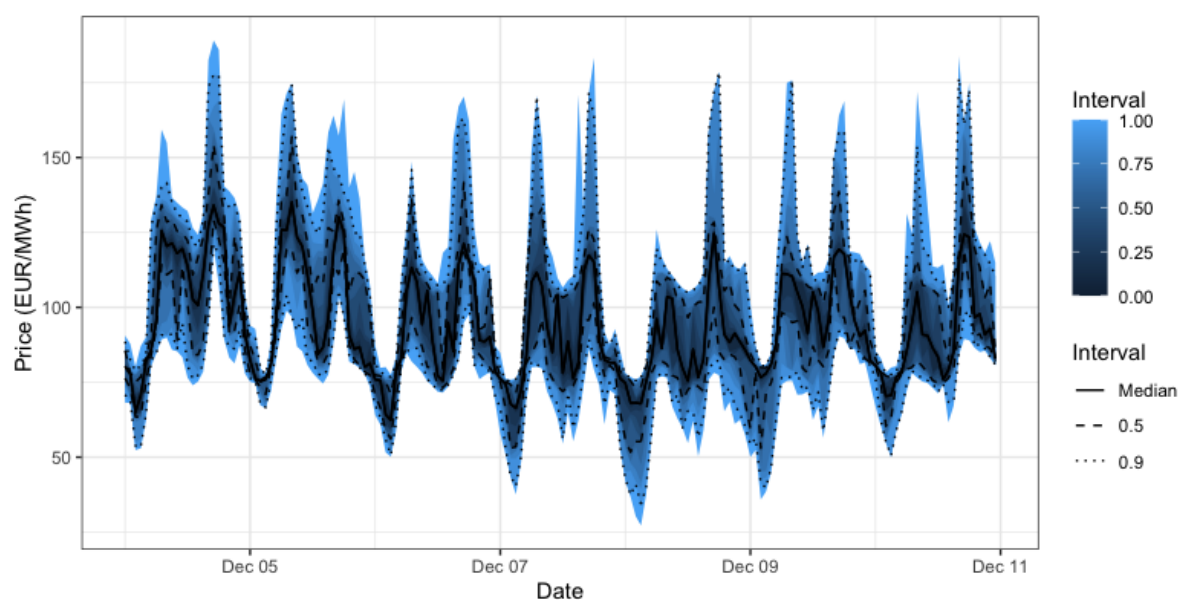


Figure 5.11: Monte Carlo simulations of prices on the French day-ahead market (reference prospective scenario, first week of December)

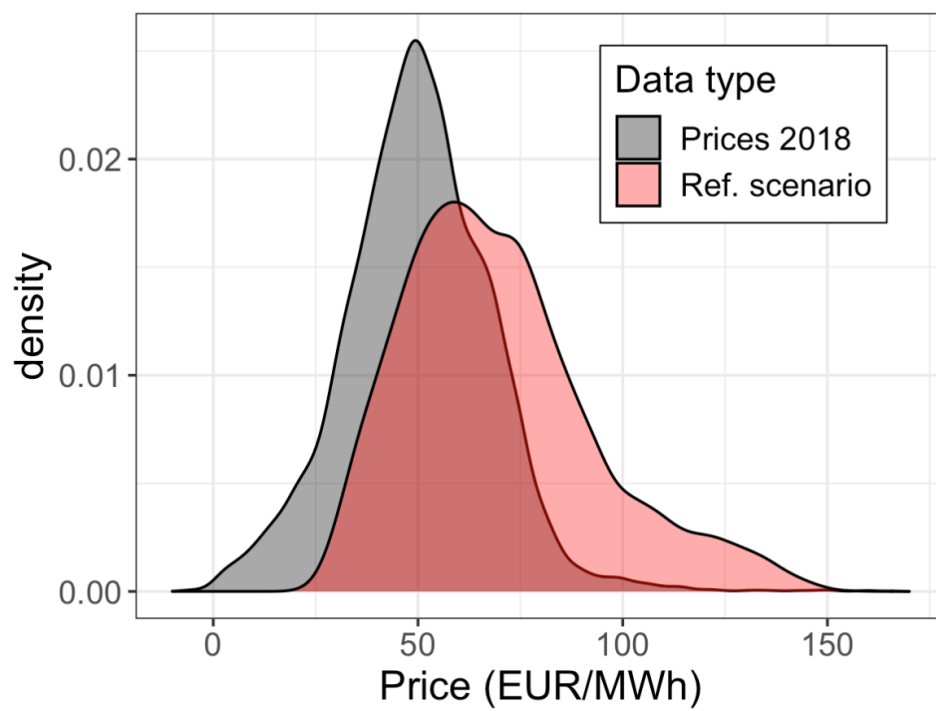


Figure 5.12: Distribution of electricity prices on the French day-ahead market

Alternative scenarios

We study three alternative scenarios that vary from our reference case in order to evaluate how electricity prices would be impacted by these changes according to our projections.

Case A: Nuclear policies For the first case, we consider the impact of the expected nuclear policies for Belgium, Germany and France. Thus, in this scenario the installed nuclear capacity in Belgium and Germany are equal to zero, while the French installed capacity is increased by 10 GW, which represent six new EPR nuclear reactors.

Case B: Hydrogen deployment For the second case, we consider the deployment of Hydrogen in Belgium, Germany and France to complement the increase of renewable capacity (500 MW of Hydrogen per country, with a capacity of 5000 MWh per country).

Case C: High operationg costs For this third case, we consider a tripling of commodity costs (oil, coal, gas and CO₂) in order to evaluate how this would impact electricity prices.

We represent the simulation results obtained for all these cases alongside the reference scenario in Figure 5.13 in the form of boxplots in order to compare the outcomes. We observe that price range of the "Nuclear policies" and "Hydrogen deployment" scenarios are quite similar to our reference scenario. For the "Nuclear policies" scenario, the median price increase compared to our reference indicates that overall the additional capacity appears to be insufficient to curbe prices due to the simultaneous Belgian and German phase-outs. Regarding the "Hydrogen deployment", we observe no significant impact on prices. This can be attributed to the techno-economic assumptions taken regarding storage. We assume operating costs of EUR 20/MWh for pumped-hydro and EUR 50/MWh for hydrogen, while we also consider a round-trip efficiency of 76% for pumped-hydro as opposed to 40% for hydrogen. Thus, the hydrogen does not play a key role in price formation according to our simulations. Finally, for the "High operating costs" scenario, we observe both a increase of the median electricity price and of its dispersion. But thanks to renewables and nuclear the tripling of fossil fuel and CO₂ cost is not fully transposed to electricity prices.

5.3.3 Discussion

In this section, we propose a method to build scenario that account for uncertainty as well as visualizations to assess the variability of time series of prices and to compare multiple scenarios.

The simulated prices appear to be credible with regards to our scenarios. Nevertheless, as shown during the validation phase on historical data, it is likely that we are underestimating the frequency and amplitude of price spikes. Thus, the scenarios outcome might be too conservative.

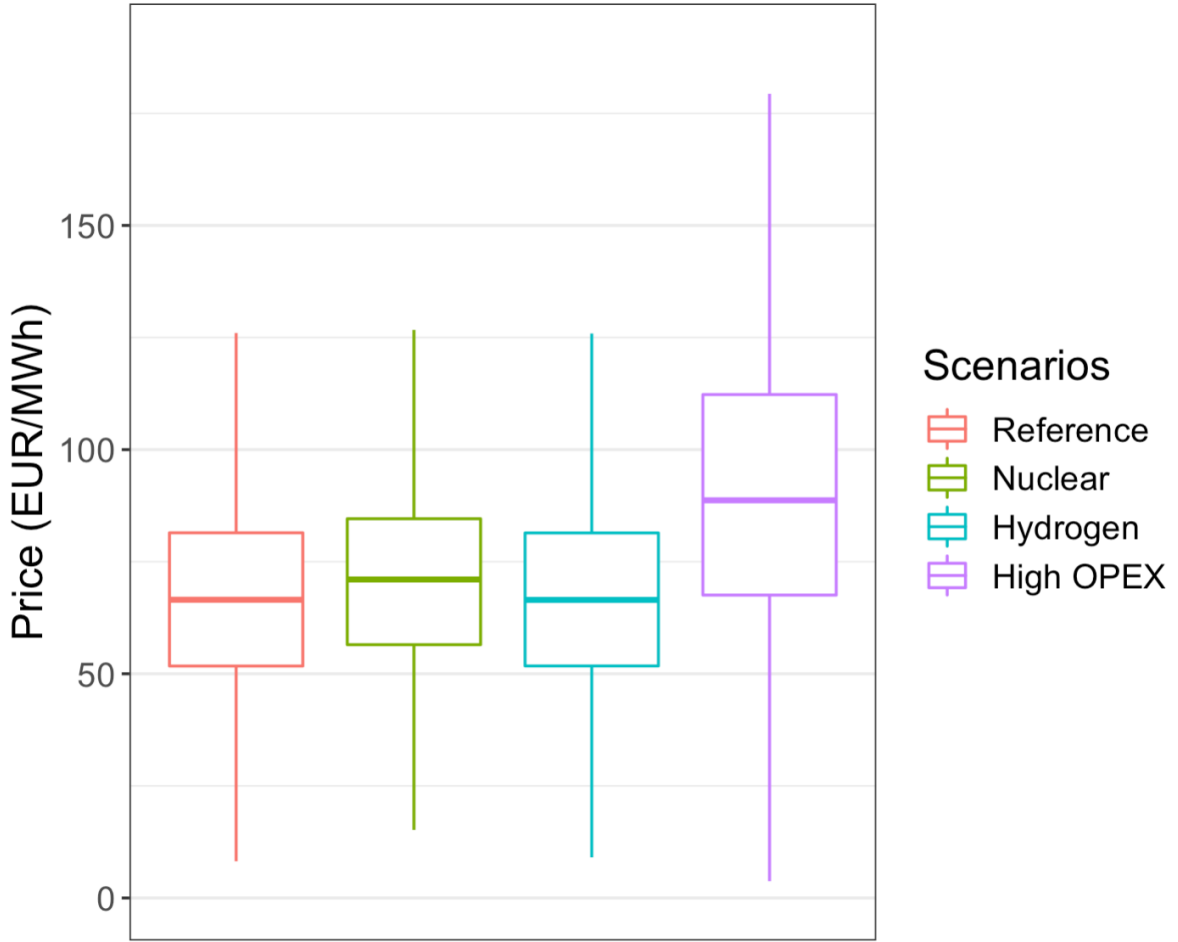


Figure 5.13: Boxplot of French prices on day-ahead markets for the prospective scenarios

Now that a global approach has been proposed to validate our data-driven structural model and to use with prospective scenarios, future work could focus on expanding the spatial scope of the simulations or to transpose of these modeling principles to more established tools. Doing so, and using a simulation hardware with adequate memory and computing power, studies at the European level could be conducted.

One remaining limitation of the modeling is that renewables are assumed to be price-taker. Some additional data and insights should be collected to more adequately reflect bidding strategies of renewable energy producers in case of high penetration rate.

5.4 Conclusions

In this chapter, we proposed extensions and final validation steps that prepared our model in view of being applied to prospective studies.

We analyzed how the estimation method for the model parameters can be enhanced to be finely tuned thanks to the introduction of a new hyperparameter and how we can account for additional market data. We highlighted that the estimation of parameters is sensitive to the values used for initialization. We observed that the parametrization can be further enhanced by introducing non-linear models of order prices.

To make the model better reflects the functioning of the European power system and market, we made sure to account for the flexibility potential of key technologies, we made the management of storage assets an internal part of the optimization problem, and we verified that the model is still operational for multizone simulations. Having mitigated the main limitations of the model allowed us to propose some applications in the form of prospective studies.

For the prospective studies, we proposed a method to build scenarios that draw from historical data in order to keep some realistic operating constraints while also accounting for uncertainty thanks to Monte Carlo simulations. This modeling strategies enabled us to consider the impact of policy decisions (nuclear development of phase-out), industrial development (installation of hydrogen capacities) and market circumstances (high fuel and CO₂ prices). Throughout the modeling proposed in this thesis, we analyzed key trends and factors that influence the development of prices on the day-ahead markets. This led us to formulate a framework that account for the market clearing mechanism and to reflect on ways to leverage historical data. These reflections and simulation outcomes gives us indications on possible future development of markets. Nevertheless, it remains that the electricity market, being a complex adaptative system, is not easily predictable over long time frames. As a matter of fact, key factors impacting predictability include: (i) "how well we understand the factors that contribute to it, (ii) how much data is available, (iii) how similar the future is to the past, and (iv) whether the forecasts can affect the thing we are trying to forecast." [136] Through the modeling of price formation, the identification of reliable data sources, and the analysis of past and future market developments, this chapter concludes our endeavor to gain a better understanding of electricity prices and to make them more predictable.

Chapter 6

Concluding remarks

Contents

6.1	General conclusions	106
6.2	Perspectives and future work	107

6.1 General conclusions

In this thesis, we studied the European Electricity market, with a special emphasis on day-ahead markets.

Electricity plays an essential role in the global decarbonization effort. As a versatile energy carrier, it can replace fossil fuels in a variety of applications, notably for mobility, heating, and cooling. The deployment and adequate use of low-carbon solutions is not only dependent on technological choices and public policies, it also depends on the economics of the energy transition. Electricity transactions are facilitated by wholesale electricity markets, which provide price signals indicating the evolution of the value of electricity at specific locations and times. Understanding the price formation mechanism contributes to a better anticipation of possible future developments, which supports rational decision making. Through the exploitation of open data, the link between the state of the power system and market prices can be modeled in order to empirically estimate the contribution of complementary price factors, thus making it possible to simulate realistic time series of wholesale electricity prices.

Since the market is a complex adaptive system, we want to make sure that the key insights we expect to remain valid in the long run are effectively accounted for in the model. Relying exclusively on data would not allow us to differentiate between structural and conjectural market features; conversely, models that replicate the market clearing mechanism allow us to explicitly reflect the price formation mechanism. This modeling practice also has the benefit of actively questioning our understanding of price formation and thus fueling fruitful discussions regarding the effect of price drivers.

Structural models have been proposed in the scientific literature and they are commonly used in industry, notably for adequacy assessments. They are based on the assumption of an optimal use of power system assets from a techno-economic perspective. This hypothesis is legitimate but does not encompass a multitude of other effects that influence prices, including the bidding strategies of market participants in case of imperfect competition. Since an explicit account of all factors influencing prices is not feasible, alternative modeling strategies must be considered. We chose to use historical power system and market data to reduce the gap between economic theory and the reality of price formation. A few approaches with a similar rationale have been proposed by researchers; being inspired by their approaches, we suggested a new method that built upon their endeavors. Here, we proposed to integrate the influence of price drivers into the optimization problem in a modular and parametrized fashion. The influence of selected factors is quantitatively estimated thanks to a dedicated algorithm. The overall methodology allows us to keep the control over the techno-economic hypotheses we want to include in the model without neglecting insights that can be gained from the use of historical datasets. One of the main benefits is that the information is then embedded in the estimated parameters so that the tuned model can be used on other datasets, for example for counterfactual analyses or prospective studies.

Case studies show that the dynamics of electricity prices is globally captured by the model, especially the seasonal variations at the daily and weekly timescales. The results are analyzed on out-of-sample data that differ from the training set, which prevents the

analysis from being distorted by a strong interpretation of overfitting phenomena. We observed that the model could still be improved to better simulate price spikes; potential improvements are discussed in the next section. We also emphasized the influence of stochastic events on the quality of our simulations and we therefore made sure to account for uncertainty in subsequent prospective studies. Moreover, the estimation of model parameters thanks to the proposed algorithm improves the quality of simulations over a simple initialization based on techno-economic hypotheses. We have shown how the estimation method can be enhanced to finely tune the weighting of data points or by retrieving information from real supply curves in order to improve the overall accuracy and the dynamics of simulations. Still, the estimation method could be further developed to be more accurate for production classes that do not often set market prices during the training period. Through successive improvements, we also examined how the modularity of the model formulation enabled us to propose extensions to consider additional technical factors (e.g. ramp constraints and strategic management of storage). Doing so, we account for the interdependency of prices across multiple hours, which is not the case for structural models in which each hourly pair of supply and demand curves is considered separately. Likewise, we also demonstrated that the method can be applied for multizone cases, which correspond to the actual configuration of the interconnected internal market in Europe.

Our model not only allows us to gain a better understanding of electricity price formation through the simulation of historical situations, it also lets us study a range of possible future electricity prices thanks to dedicated scenarios. These scenarios make it possible to transcribe potential policy decisions and technological upgrades into custom datasets. The scenarios, which reflect the assumptions or wishes of the modeler, can also include some intrinsic uncertainty corresponding to stochastic factors. For example, the load factor of solar and wind or the unavailability of conventional power plants can be accounted for through Monte Carlo simulations. Building scenarios and simulating prices using our model enables us to estimate and compare the impact of investment pathways through prospective studies.

6.2 Perspectives and future work

Further studies of the electricity market and its evolution can be performed. We suggest here a few directions to build upon this work.

The model itself can be extended further to account for additional technical constraints, such as start and stop constraints for power plants. Performing more precise studies at the power plant unit level would be an opportunity to assess to what extent proprietary technical information about individual units can be inferred from public data. Likewise, the correspondence between individual bids and units could be further studied instead of adopting a more aggregated approach through the supply curve. Still, scaling down precisely is not always possible, and capturing aggregated effects should remain the focus for the simulation of market prices. Regarding the modeling choices for the supply side, the parametrization formulation could be replaced by non-linear terms

in order to better capture price spikes, or regime-switching could be proposed to change the value of parameters during scarcity situations. A next step could be to keep the linear formulation but to explore the potential of performing local regressions during the training phase instead of a single regression per production class. Moreover, a more complex and accurate implementation of the demand side modeling could be examined. While we mostly considered a price inelastic residual demand, additional cases could be studied in which the value of demand side response would be explicitly tested. Similarly, renewables have been considered to be mostly price-takers in our case studies, but with a higher penetration rate of weather-driven production assets, this assumption might have to be reevaluated. To do so, specific bidding models could be developed to estimate relevant strategies for renewable energy producers with market orders above their marginal production costs.

The method could also be further applied for studies spanning larger spatial-temporal scopes. Open data initiatives offer datasets that cover an increasingly longer period and the quality of datasets keeps improving, which would prove interesting for a more precise estimation of model parameters. Using these data and running the model at the scale of Europe would also be interesting to study price convergence phenomenon. In addition, the procurement of proprietary data, such as power prices on futures markets, could also improve the relevance of simulations. For the construction of scenarios at the European scale, detailed analyses of political decisions and industrial trends would bring additional value. Specific analyses of the prospective studies could also be performed in order to assess the expected revenues generated from new investments and to define public policies, notably fiscal policies regarding the need for new subsidies, taxes or contracts for difference (CfD).

A last avenue to extend these studies would be to consider more closely the evolution of the functioning of the day-ahead market and its interaction with other markets. The evolution of the day-ahead market and the intraday market could be analyzed jointly in order to assess how the respective trading volumes impact prices. Also, more speculative discussions could be conducted to estimate the extent to which the rise of trading bots and the 15-minute time frame could influence price formation, and then how such insights could be incorporated into our model.

Appendix A

Résumé du chapitre 1 : Introduction

A.1 Contexte

L'électricité : le besoin d'un équilibre constant L'électricité est l'un des piliers essentiel au monde moderne. Les sociétés industrielles sont construites sur l'hypothèse que la demande d'électricité sera satisfaite et que les pannes seront rares. Des structures technico-économiques fiables, guidées par des politiques énergétiques, sont mises en place pour garantir que la production correspond à la consommation à tout moment.

D'un point de vue technique, les unités de consommation se fournissent en électricité auprès de moyens de production par l'intermédiaire du réseau électrique. Ces transferts d'énergie doivent toujours être équilibrés à des échelles de temps courtes, ce qui nécessite une coordination complexe. 2021 a permis de rappeler que cet équilibre peut être fragile, même dans les pays développés (par exemple, les effets de vague de froid au Texas [2]) et que les prix de l'énergie peuvent varier brutalement sous l'effet d'un choc sur l'équilibre offre-demande (par exemple, la reprise économique après les premières vagues de Covid-19 [3]). Par ailleurs, l'équilibre à long terme dépend d'investissements adéquats pris par anticipation.

Équilibre par les marchés de gros de l'électricité Sur les marchés libéralisés de l'énergie, l'équilibre entre l'offre et la demande est respecté à la fois d'un point de vue physique (c'est-à-dire que la consommation est égale à la production) et d'un point de vue financier (c'est-à-dire que l'énergie est achetée et vendue).

Les marchés de gros de l'électricité permettent aux producteurs et aux consommateurs de s'entendre sur les quantités d'électricité à échanger et de fixer les prix correspondants. Les acteurs du marché échangent l'électricité sur des marchés organisés ou de gré à gré en fonction de leurs préférences, qui dépendent notamment des horizons temporels et des points de livraison. La dynamique de l'offre et de la demande est complétée par des politiques fiscales (subventions et taxes) ainsi que par un large éventail de réglementations, qui sont définies par les pouvoirs publics pour guider l'évolution du

secteur de l'électricité.

Un équilibre sous contrainte Trouver un équilibre à court et à long terme est une tâche complexe qui est directement liée aux trois dimensions du trilemme énergétique¹ : (i) la sécurité énergétique, (ii) l'équité énergétique et (iii) la durabilité environnementale. Dans le cadre du paradigme sociétal actuel, le système électrique doit être suffisamment robuste pour garantir la sécurité énergétique, et non surdimensionné pour éviter le gaspillage des ressources. Cette fiabilité a un coût, et une compensation financière peut donc inciter les acteurs du marché à fournir et à déployer des technologies adéquates. Dans le même temps, l'électricité est considérée comme une nécessité, et un accès équitable à tous doit être garanti ; une énergie inabordable peut même déclencher des troubles sociaux. Néanmoins, la consommation d'énergie liée à la technologie est également responsable des dommages que nous causons à l'environnement, qui se sont multipliés depuis la révolution industrielle. L'utilisation inconsidérée des matériaux et de l'énergie entraîne un effondrement de la biodiversité qui ne peut être ignoré. Ainsi, l'offre et la demande doivent non seulement être équilibrées, mais elles doivent également rester dans des limites physiques acceptables.

L'importance du prix de l'électricité Le prix associé à l'électricité est probablement l'indicateur le plus significatif de la valeur que nous lui attribuons. Considérée comme un signal, l'évolution des prix indique les risques et les opportunités. Comprendre les mécanismes de formation des prix pour anticiper leur évolution future nous aide à définir des stratégies d'investissement et des politiques publiques pertinentes. Pour cela, il faut identifier les principaux facteurs de formation des prix et leur influence. Dans l'Union européenne, le secteur de l'électricité est marqué par deux tendances majeures : la transition énergétique, et le développement d'un marché intérieur de l'énergie totalement intégré. Ces deux tendances ont un impact direct sur les prix de gros de l'électricité et seront présentées dans cette thèse.

A.2 Problématique

A.2.1 Question de recherche

Les modèles numériques nous permettent d'étudier les prix de manière quantitative. En s'appuyant sur les théories économiques et les hypothèses de formation des prix, on peut construire des modèles empiriques. Pour cela, il faut utiliser des données réelles sur le système électrique et les marchés de l'énergie, rendues possibles par les données ouvertes, qui sont de plus en plus répandues grâce à une réglementation favorable et à des initiatives de partage des données. Année après année, l'ensemble des données disponibles s'accroît et leur qualité s'améliore (plus grande granularité, standardisation, etc.). L'exploitation de ces données est essentielle pour développer des modèles pertinents des prix de l'électricité.

¹Terme défini par le Conseil mondial de l'énergie

Le contexte d'incertitude quant à l'avenir des prix de l'électricité et les possibilités offertes par leur modélisation pour mieux comprendre la formation des prix nous amènent à la question centrale suivante qui motive cette thèse : "Comment utiliser au mieux les données publiques disponibles afin de modéliser la relation entre les facteurs de prix et la dynamique des prix de gros de l'électricité tout en tenant compte du mécanisme réel de formation des prix ?".

A.2.2 Pertinence

La transition énergétique implique l'émergence de nouveaux usages et de nouvelles technologies. La rentabilité des solutions déployées dépend du prix de l'électricité sur les marchés de gros, mais aussi d'autres sources de revenus, comme les services auxiliaires, les marchés de capacité et les garanties d'origine. Les prix de l'électricité varient en fonction de l'heure (par exemple, les prix sont généralement plus élevés le jour que la nuit, notamment en raison des variations de la demande) et du lieu (par exemple, les zones d'enclenchement). Les prix peuvent être volatils et varier fortement d'une heure à l'autre, de sorte qu'une granularité temporelle élevée est nécessaire dans la modélisation pour effectuer des évaluations précises.

L'étude des prix de l'électricité sur les marchés est notamment pertinent lorsque l'on s'intéresse au développement des énergies renouvelables et du stockage, ainsi qu'à l'utilisation de sources de flexibilité pour la gestion du réseau électrique. L'étude des prix de gros est également riche en enseignements sur l'évolution des prix de détail.

A.2.3 Délimitation du sujet

La modélisation des marchés de gros de l'électricité est un sujet vaste. Cette thèse s'intéresse tout particulièrement aux marchés journaliers européens. L'attention porte sur les séries temporelles de prix au pas de temps horaire.

A.3 État de l'art et analyse des écarts

L'équilibre offre-demande est généralement représenté par une approche ascendante, où le prix résulte de l'utilisation optimale des actifs du système électrique tout en tenant compte des contraintes technico-économiques. De tels modèles, où l'ordre de mérite joue un rôle clé dans la formation des prix, sont appelés modèles structurels dans la littérature scientifique ; une étude des modèles structurels est proposée dans [5]. Cette approche est particulièrement adaptée aux études à long terme, car la méthode peut tenir compte de l'évolution du système électrique et de la conception du marché. Ces modèles sont largement utilisés, mais ils nécessitent de formuler de nombreuses hypothèses afin d'identifier les coûts marginaux à court terme. De plus, les simulations ne conduisent pas nécessairement à des prévisions de prix qui reflètent la dynamique observée du marché, puisque cette classe de modèles n'est pas principalement conçue pour capturer le comportement

stratégique des participants au marché. Une analyse détaillée de la littérature scientifique est présentée au début du chapitre 4. Nous examinons ici les travaux saillants de la littérature et identifions comment nous pourrions nous appuyer sur ces travaux antérieurs.

Si la prévision des prix de l'électricité est globalement un domaine de recherche bien développé [6], peu de modèles ont été spécifiquement conçus pour des horizons temporels d'un an ou plus, ce qui correspond aux études à long terme [7]. Parmi ce type de modèle, une approche pour les marchés couplés de l'électricité est proposée dans [8], où la courbe d'offre est construite par empilement d'ordres de vente pour lesquels les offres de prix sont fonction de la marge d'offre (c'est-à-dire la différence entre la capacité totale disponible et la demande) et du coût de production par technologie. Si l'approche proposée reflète de nombreux aspects du fonctionnement du marché et du système électrique, elle ne vise pas à simuler précisément la dynamique horaire des prix. À l'inverse, dans [9], une attention particulière est accordée au réalisme de la variabilité des simulations. Leur modèle se base sur une description de la courbe d'offre par technologie à travers le coût marginal à court terme. Les offres de prix sont ensuite réduites dans le cas d'une demande faible (reflétant par exemple les contraintes d'arrêt et de démarrage) et augmentées pour une demande élevée (induisant une rente de rareté). Les auteurs utilisent des données réelles pour simuler des prix réalistes, mais leur méthode ne tient compte ni des contraintes techniques qui sont typiquement incluses dans les problèmes d'attribution des unités dans la production d'électricité, ni de l'influence des autres marchés voisins. Par ailleurs, des méthodologies complémentaires sont proposées dans [10] et [11] pour modéliser les prix horaires en utilisant l'approche de l'ordre de mérite au niveau technologique afin de modéliser de manière réaliste la dynamique des prix sur le marché journalier allemand. Les deux modèles utilisent des données réelles et des régressions linéaires. Les régressions sont utilisées dans le modèle [10] pour ajuster la sortie du modèle d'ordre de mérite afin de tenir compte d'effets tels que les contraintes de rampe et les enchères stratégiques, tandis que le modèle [11] utilise des statistiques pour modéliser la balance du commerce extérieur. Ces modèles utilisent des données pour renforcer le réalisme des simulations, mais ils ne proposent pas d'utiliser directement la calibration statistique au sein même du problème d'optimisation de manière paramétrée, ce qui limite potentiellement la modularité de la formulation du problème.

L'approche basée sur les données peut contribuer à atténuer les limites typiques des modèles structurels. Les données du système électrique et du marché contiennent des informations sur les facteurs de prix, qui ne sont généralement pas incluses dans les modèles d'optimisation, comme la prime de prix ajoutée au coût marginal en cas de faible marge d'approvisionnement. Ces données peuvent être utilisées pour renforcer la fiabilité des simulations et aller au-delà des théories économiques.

A.4 Objectifs et contributions

L'objectif principal de cette thèse est de proposer des méthodes permettant d'exploiter la richesse des données disponibles, en particulier les données du marché de l'électricité,

afin d'améliorer l'approche du modèle structurel utilisé pour simuler les séries temporelles des prix de l'électricité sur les marchés journaliers. La nouveauté de notre modèle réside dans son objectif de combler le fossé entre (i) les méthodes d'optimisation qui estiment le coût total minimal de la fourniture d'électricité, et (ii) les méthodes basées sur les données qui visent à capturer les tendances observées dans la façon dont le marché résout efficacement ce problème d'allocation des ressources. Le rapprochement des deux approches nous permet de nous appuyer sur des bases théoriques solides sans négliger les effets observés dans la pratique. Le modèle est développé dans l'intention d'être adapté à l'estimation des prix futurs à long terme sur la base de scénarios d'évolution des marchés de l'électricité et des systèmes électriques. Les objectifs peuvent être résumés comme suit :

- Mettre en évidence les tendances clés du système électrique et des marchés de l'électricité de l'UE, qui façonneront l'évolution future des prix de gros de l'électricité.
- Exploiter les données ouvertes afin d'étudier la formation des prix de l'électricité
- Proposer des modèles innovants qui combinent l'optimisation et les approches statistiques afin de simuler des séries chronologiques réalistes de prix.
- Formaliser le lien quantitatif entre les facteurs de prix et les prix du marché par une formulation d'optimisation convexe
- Proposer une paramétrisation des offres de prix afin de rendre compte des contraintes technico-économiques et des comportements stratégiques
- Définir un algorithme pour estimer les paramètres de prix tout en tenant compte des données réelles et des effets de l'ordre de mérite
- Valider les méthodes proposées sur plusieurs années sur une étude de cas réelle
- Utiliser les modèles pour des études prospectives basées sur des scénarios impliquant des évolutions possibles du contexte énergétique.

A.5 Plan du document

Dans le chapitre 2, nous commençons par un aperçu global du système énergétique mondial et, plus précisément, nous présentons le rôle actuel des combustibles fossiles.

Dans le chapitre 3, nous nous intéressons aux marchés journaliers de l'électricité, qui représentent un élément clé du marché intérieur européen de l'énergie. Nous détaillons le fonctionnement de ces marchés, en particulier le mécanisme qui sous-tend la formation des prix.

Dans le chapitre 4, nous proposons d'aller au-delà d'un modèle purement basé sur les données en tenant compte explicitement de nos connaissances sur le fonctionnement des systèmes et des marchés électriques dans notre approche de modélisation.

Dans le chapitre 5, nous nous appuyons sur le modèle proposé au chapitre précédent en permettant à l'algorithme d'estimation de prendre en compte d'autres données économiques pertinentes (c'est-à-dire les courbes d'offre) et de mieux refléter la variabilité des séries temporelles de prix. En utilisant le modèle validé, des applications sont finalement proposées. Une méthodologie utilisée pour construire des scénarios pour les études prospectives est présentée. Les résultats des études basées sur des scénarios sont discutés à la lumière des forces et des limites de la méthode.

Le chapitre 6 conclut cette thèse. Les principales observations et contributions sont résumées. À partir des résultats obtenus jusqu'à présent, des perspectives sont offertes pour les recherches futures.

A.6 Liste des communications scientifiques

Cette recherche a été présentée à un public d'experts du marché de l'électricité (revue à comité de lecture et conférence internationale). La thèse a également été présentée à un public plus large de scientifiques intéressés par divers sujets liés à la transition énergétique lors de sessions de posters.

Par ailleurs, le fait de remporter le concours de prévision EEM20 avec une équipe composée de membres de notre laboratoire a donné lieu à des communications non directement liées au sujet de cette thèse. L'objectif de cette compétition était de proposer des prévisions probabilistes horaires de la production éolienne agrégées à des niveaux régionaux.

La liste des communications scientifiques est présentée ci-après.

Revue à comité de lecture:

- *Data-driven Structural Modeling of Electricity Price Dynamics* - Valentin MAHLER, Robin GIRARD, Georges KARINIOTAKIS - Energy Economics, Elsevier, 2022, 107, pp.105811, doi:10.1016/j.eneco.2022.105811 (<https://hal.archives-ouvertes.fr/hal-03542564>)

Conférences internationales:

- *Simulation of day-ahead electricity market prices using a statistically calibrated structural model* - Valentin MAHLER, Robin GIRARD, Sébastien BILLEAU, Georges KARINIOTAKIS - 16th International Conference on the European Energy Market (<https://hal.archives-ouvertes.fr/hal-02351932/document>)
- *Probabilistic Forecasting of Regional Wind Power Generation for the EEM20 Competition: a Physics-oriented Machine Learning Approach* - Kévin BELLINGUER, Valentin MAHLER, Simon CAMAL, Georges KARINIOTAKIS - 17th International Conference on the European Energy Market (<https://hal.archives-ouvertes.fr/hal-02952589/document>)

- *Forecasting regional wind production based on weather similarity and site clustering for the EEM20 Competition*, Kévin BELLINGUER, Valentin MAHLER, Simon CAMAL - 40th International Symposium on Forecasting – ISF20 (<https://hal.archives-ouvertes.fr/hal-03157849>)

Wébinaire:

- *EEM20 Forecasting Competition: Probabilistic forecast of aggregated wind production with hourly resolution* - Kévin BELLINGUER, Valentin MAHLER, Simon CAMAL - International Energy Agency Wind Task 36 (<https://youtu.be/n0m3S18Zwtk>)

Posters :

- *Modélisation de l'évolution des prix sur le marché de l'électricité dans un contexte de forte pénétration des énergies renouvelables* - PhD student seminar - French Agency for Ecological Transition (ADEME) - March 2019 - Angers, France
- *Simulation of day-ahead electricity market prices using a statistically calibrated structural model* - Summer School on Data-Driven Analytics and Optimization for Energy Systems - Technical University of Denmark, Center for Electric power and Energy (DTU CEE) - June 2019 - Copenhagen, Denmark (<https://hal.archives-ouvertes.fr/hal-03445344/document>)
- *Modélisation de l'évolution des prix de l'électricité dans un contexte de forte pénétration des énergies renouvelables* - PhD students seminar - Doctoral school ISMME (Ingénierie des Systèmes, Matériaux, Mécanique, Énergétique) - September 2020 - Online

Appendix B

Résumé du chapitre 2 : Transition énergétique et marchés européens de l'électricité

Pour bien étudier les marchés de l'électricité, il faut d'abord bien comprendre l'offre, la demande et la manière dont les échanges sont rendus possibles. C'est ce que nous nous proposons de faire dans ce chapitre. L'objectif est non seulement de présenter le système tel qu'il est aujourd'hui, mais aussi la manière dont il est susceptible d'évoluer. Pour ce faire, une approche multidisciplinaire est adoptée. Des chiffres et des études clés sont présentés afin de fonder les analyses sur des données quantitatives.

Nous commençons par présenter la consommation énergétique mondiale et son lien avec l'Anthropocène. La place spécifique de l'électricité parmi toutes les sources d'énergie est mise en évidence et son rôle spécifique dans le cadre de la transition énergétique est expliqué. Après cet aperçu global, nous nous concentrons sur le système électrique européen et sur la manière dont l'électricité est produite, consommée et fournie aux utilisateurs finaux. Enfin, nous analysons comment les marchés de gros facilitent le commerce de l'électricité en Europe.

B.1 Le lien entre climat et énergie

B.1.1 Contribution des combustibles fossiles au changement climatique

Afin de réduire significativement les risques et les impacts du changement climatique, l'accord de Paris, adopté en décembre 2015, a fixé l'objectif de maintenir "l'augmentation de la température moyenne mondiale bien en dessous de 2°C par rapport aux niveaux préindustriels". L'augmentation des températures moyennes mondiales est notamment due aux émissions fossiles de CO₂, qui sont les principales sources de gaz à effet de serre (GES) émis [12]. Un budget carbone, c'est-à-dire la quantité maximale de dioxyde de carbone (CO₂) qui peut encore être libérée dans l'atmosphère avant de compromettre un

objectif climatique, peut être estimé grâce à des modèles climatiques. Le budget carbone de 2°C est incompatible avec les tendances de consommation de charbon, de pétrole et de gaz naturel observées depuis le début de la révolution industrielle [13, 14]. Ce budget impose non seulement de réduire la consommation de combustibles fossiles, mais aussi de considérer une part élevée des réserves de carbone comme imbrûlable [15]. Néanmoins, nous ne sommes absolument pas sur la bonne voie pour réduire le surplus d'émissions [12]. L'ampleur du changement climatique dépendra de la rapidité avec laquelle nous réduirons ces émissions et de la mesure dans laquelle nous parviendrons à capter les gaz à effet de serre émis. Selon les projections, la mise en œuvre d'aucune politique climatique devrait entraîner une augmentation de la température moyenne mondiale par rapport aux niveaux préindustriels d'environ 4,5°C d'ici 2100. Les politiques actuelles, si elles sont correctement mises en œuvre, devraient conduire à une augmentation d'environ 3 degrés Celsius. Les changements dans le secteur de l'énergie sont essentiels, car la consommation d'énergie est responsable des trois quarts des émissions mondiales de gaz à effet de serre [16].

B.1.2 Les leviers pour réduire les émissions de combustibles fossiles

La réduction des émissions implique de diminuer le produit entre la consommation finale d'énergie et l'intensité des émissions. Idéalement, l'intensité des émissions devrait être estimée grâce à une évaluation du cycle de vie complet afin de tenir compte des émissions en amont, telles que les pertes de conversion de l'énergie primaire en énergie finale.

La consommation finale d'énergie dépend de la population mondiale et de la consommation par habitant (en gardant à l'esprit que la variabilité entre les individus est considérable : "Les émissions des 1% les plus riches de la population mondiale représentent plus de deux fois la part combinée des 50% les plus pauvres" [12]). Le juste niveau de consommation par habitant renvoie au concept de sobriété énergétique [18], qui est difficile à quantifier et à transposer dans des politiques publiques acceptables. Il est également peu probable que la sobriété énergétique soit spontanément adoptée par la plupart des individus [19].

L'intensité des émissions peut être réduite en utilisant de l'électricité produite à partir de sources à faible teneur en carbone (principalement l'hydroélectricité, le nucléaire, l'éolien et le solaire). L'électrification à grande échelle du secteur du chauffage et de la réfrigération ainsi que du secteur des transports est donc un défi industriel permanent et une composante essentielle de la transition énergétique [20].

B.1.3 La transition énergétique en Europe

Pour atteindre la neutralité climatique d'ici à 2050, la Commission européenne s'est fixé pour objectif de réduire les émissions de gaz à effet de serre d'au moins 40% en 2030 par rapport à 1990. Cet objectif de 2030 devrait être atteint grâce à une part d'énergie renouvelable d'au moins 32% et à une efficacité énergétique de 32,5% par rapport à 1990

[29]. Cette ambition est susceptible d'être revue à la hausse avec le prochain paquet Fit for 55, qui fait l'objet de discussions entre les États membres. Pour atteindre les objectifs de 2030, chaque pays développe sa propre stratégie nationale. L'approche choisie est détaillée dans les plans nationaux pour l'énergie et le climat (PNEC) [30], qui doivent être conformes aux stratégies nationales à long terme destinées à atteindre l'accord de Paris à l'horizon 2050 [31]. L'Agence internationale de l'énergie a procédé à un examen complet de la politique énergétique de l'UE, qui a abouti à la publication d'un rapport détaillé sur la transformation du système énergétique et la sécurité énergétique [32].

L'échange de droits d'émission de carbone est l'un des principaux instruments des politiques de transition. Une quantité limitée d'émissions de gaz à effet de serre par les secteurs de l'électricité, de l'industrie et de l'aviation est autorisée dans l'UE. Le prix du marché des émissions évolue en fonction de l'offre et de la demande au sein du système communautaire d'échange de quotas d'émission. Le plafond du nombre de quotas d'émission diminue chaque année sur une base réglementaire. Pour les installations fixes telles que les centrales électriques, le plafond à l'échelle de l'UE a été réduit de 1,74% par an pour la période 2013-2020 et est réduit d'un facteur annuel de 2,2 pour la période 2021-2030. Pour l'année 2021, des émissions correspondant au potentiel de réchauffement planétaire d'environ 1,6 milliard de tonnes de CO₂ sont autorisées [33].

B.2 Le système électrique européen

B.2.1 Demande

Comme dans le reste du monde, l'électricité représente une part mineure de la consommation finale d'énergie en Europe (environ 10% de la consommation totale), tandis que les énergies fossiles sont dominantes. Cela signifie que le potentiel d'électrification est encore conséquent en Europe. Globalement, la consommation d'électricité est assez stable sur la période 2000-2018. L'électricité est principalement consommée par le secteur industriel (environ 38%), les services (environ 29%) et les ménages (environ 28%). Les principaux contributeurs industriels sont (i) la chimie et la pétrochimie, (ii) les machines, (iii) le fer et l'acier, (iv) l'alimentation, les boissons et le tabac, et (v) le papier, la pâte à papier et l'imprimerie.

La composition de la demande d'électricité va évoluer en raison de l'électrification des utilisations finales et du développement de nouveaux usages. Les principaux moteurs de l'électrification sont le chauffage, la ventilation et la climatisation (pompes à chaleur), la chaleur industrielle (chaudières, fours), l'e-mobilité (véhicules électriques) et la production d'hydrogène. La gestion de la demande est susceptible d'influencer les futurs modèles de consommation temporelle d'électricité. Des réductions de coûts et d'émissions de CO₂ peuvent être obtenues en tenant compte des contraintes d'approvisionnement lors de l'utilisation de l'électricité.

B.2.2 Approvisionnement

Le mix de production évolue en fonction de la demande, de la capacité installée, des conditions météorologiques et des contraintes technico-économiques. La part des combustibles fossiles dans la production d'électricité diminue progressivement en Europe et est remplacée par des sources d'énergie renouvelables[39].

Au cours des deux dernières décennies, une augmentation significative de la capacité installée mondiale a été observée, notamment pour les centrales éoliennes, solaires photovoltaïques (PV) et au gaz, comme le montre la figure B.1. Cette évolution peut être comprise à la lumière de l'évolution du coût énergétique nivelé (LCOE) non subventionné pour la production à l'échelle industrielle [40].

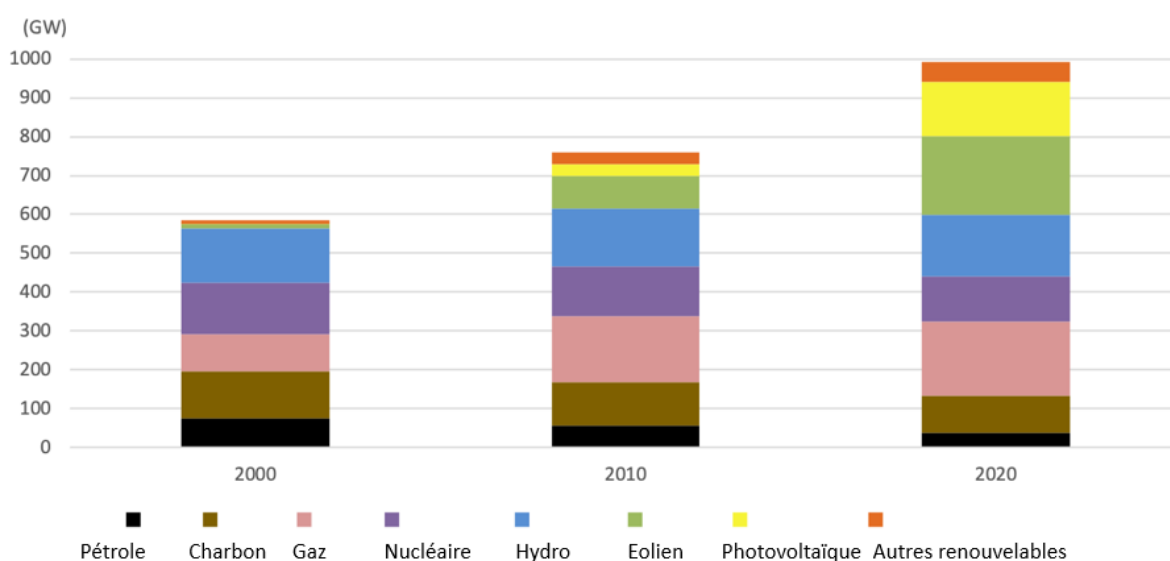


Figure B.1: Capacité électrique installée dans l'Union européenne [44]

B.2.3 Stockage

Sous l'impulsion de la conjonction de nouveaux besoins et de l'innovation technologique, le stockage prend une place de plus en plus importante dans les systèmes électriques. Le stockage hydroélectrique par pompage est le principal stockage d'énergie dans l'UE, à tel point qu'il pourrait être considéré comme le seul significatif (c'est la seule technologie de stockage mentionnée dans la plateforme de transparence pour le marché paneuropéen [52]). Le Power-to-gas, notamment l'hydrogène vert (c'est-à-dire l'hydrogène produit grâce à de l'électricité à faible teneur en carbone), peut être utilisé pour le stockage de l'énergie à long terme ; mais il devra concurrencer d'autres usages, qui pourraient être plus intéressants en termes de coûts et de réduction du carbone (procédés chimiques et industriels, transports, ...)[53]. Les batteries deviennent compétitives pour le stockage à

court et moyen terme et progressent rapidement à la fois comme stockage de masse et comme solution distribuée "derrière le compteur" [54].

B.2.4 Réseau électrique

Le réseau électrique joue un rôle clé dans la transition énergétique car il permet d'exploiter tout le potentiel des solutions de flexibilité. Une augmentation de la flexibilité à différentes échelles de temps faciliterait l'intégration des sources d'énergie renouvelables liées aux conditions météorologiques tout en garantissant la sécurité d'approvisionnement. Le réseau électrique est notamment un élément central du couplage des secteurs d'utilisation finale (transport, industrie, chauffage et refroidissement) et de l'intégration transversale (réseaux d'électricité, de gaz et de chaleur), qui sont bénéfiques en termes d'efficacité, de flexibilité, de fiabilité et d'adéquation [64].

B.2.5 Scénarios prospectifs

La voie vers des émissions nettes nulles en 2050 est incertaine et des scénarios à différentes échelles géographiques ont été proposés.

L'Agence internationale de l'énergie (AIE) propose une trajectoire à l'échelle mondiale. Des scénarios jusqu'en 2050 à l'échelle européenne avec des données détaillées sur les réseaux électriques au niveau national ont été proposés par la Commission européenne (CE) et par les ENTSOs¹. La Commission européenne a proposé un scénario de référence de l'UE en 2020, qui a été élaboré pour être cohérent avec les objectifs énergétiques de l'UE. Des efforts supplémentaires seront nécessaires pour atteindre les objectifs de 2030 selon les projections de l'Agence européenne de l'énergie [70]. En France, des ensembles de scénario à l'horizon 2050 ont été publiés par le gestionnaire du réseau de transport d'électricité (RTE) et par l'agence de la transition écologique (ADEME).

Tous ces scénarios indiquent que la météo aura une influence de plus en plus importante sur les systèmes électriques. La température, le vent, l'irradiation solaire et les précipitations auront un impact tant sur la demande que sur l'offre (demande de chauffage et de refroidissement, production renouvelable, ressources en eau pour le fonctionnement des centrales nucléaires, etc.) Le problème est que les conditions météorologiques deviendront moins prévisibles en raison du changement climatique, de sorte que le surdimensionnement des futurs systèmes électriques pourrait être nécessaire pour garantir l'adéquation des ressources.

Dans tous les cas, on s'accorde à dire que les systèmes électriques devront évoluer considérablement et rapidement. Les goulets d'étranglement qui limitent l'ampleur ou la vitesse de la transition ne sont pas clairement identifiés. Mais nous savons avec certitude que des investissements massifs sont nécessaires pour rendre la transition proposée possible. Pour rendre la transition possible, il faudra mobiliser des ressources

¹ENTSOs englobe les associations européennes pour la coopération des gestionnaires de réseaux de transport d'électricité (ENTSO-E) et de gaz (ENTSOG)

financières provenant de différents flux et allouer une part importante des portefeuilles d'investissement à des solutions vertes.

B.3 Marchés de l'électricité en Europe

B.3.1 Marchés spot paneuropéens

La création d'un marché intérieur de l'énergie dans l'UE est un processus continu qui conduit à davantage d'intégration et d'harmonisation. Une évaluation économique détaillée du marché unique européen de l'électricité est proposée dans [75], dans laquelle l'impact des changements institutionnels sur les prix, la sécurité d'approvisionnement, l'environnement et l'innovation est notamment examiné.

L'évolution la plus notable pour les marchés spot (c'est-à-dire les marchés journaliers et intrajournalier) est la création d'un couplage unique day-ahead (SDAC) et d'un couplage unique intraday (SIDC) paneuropéens et transzonaux, conformément au règlement sur l'attribution des capacités et la gestion de la congestion (règlement CACM). L'objectif est de créer un marché paneuropéen unique et transzonal de l'électricité à un jour. Afin de permettre une meilleure utilisation des ressources du système électrique, des changements techniques sont actuellement mis en œuvre tels que : (i) l'utilisation d'un calcul de capacité basé sur les flux pour optimiser l'utilisation de l'interconnexion, (ii) la transition d'une unité de temps de marché (MTU) d'une heure à quinze minutes afin de promouvoir une meilleure utilisation des solutions de flexibilité [76]. Le SDAC sera présenté plus en détail dans le prochain chapitre de cette thèse.

B.3.2 Transition énergétique et conception du marché

Tant les gestionnaires de réseaux de transport [77] que les régulateurs de l'énergie [78] se demandent si les conditions et la conception actuelles du marché fourniront les signaux d'investissement nécessaires pour garantir la sécurité d'approvisionnement à long terme. La raison en est qu'une part plus importante de renouvelables à faibles coûts d'exploitation risque de faire baisser les prix de gros de l'électricité et de pousser les actifs dispatchables hors de l'ordre de mérite, les rendant ainsi non rentables même si certains sont nécessaires à la sécurité de l'approvisionnement. Le marché de capacité est la principale solution proposée pour atténuer ce risque.

Dans le contexte du système électrique actuel et futur de l'UE, la gestion efficace de la congestion et la fourniture adéquate de services auxiliaires sont des sujets récurrents pour lesquels des solutions de marché adaptées sont envisagées [77]. Des solutions définitives ne sont pas encore adoptées, mais les marchés locaux de flexibilité apparaissent comme une solution complémentaire à ces questions [79]. Ces marchés locaux sont censés compenser la faible granularité spatiale inhérente à la conception du marché zonal. Néanmoins, la coexistence de marchés zonaux et locaux crée les conditions d'un jeu indésirable d'augmentation-diminution (inc-dec), où "les acteurs du marché anticipent le marché de réexpédition et font des offres stratégiques sur le marché zonal" [80].

B.3.3 Participants au marché

La liste complète des acteurs du marché publiée par ACER [92] comporte environ 15 000 entrées. En 2019, 302 participants de marché sont actifs sur EPEX Spot [93].

Tous les participants ne sont pas intéressés par l'achat physique d'électricité. En théorie, la spéculation pure n'est pas nécessairement préjudiciable au fonctionnement du marché et peut même jouer un rôle utile. Les spéculateurs jouent également un rôle actif sur le marché en prenant des paris risqués sur l'évolution des prix à court et moyen terme sans être fondamentalement intéressés par le produit de base sous-jacent. Ce faisant, ils améliorent la liquidité du marché, contribuent à la découverte des prix et assument des risques que les autres participants au marché ne souhaitent pas prendre [95]. Avec le développement de nouveaux marchés et l'augmentation de la granularité temporelle, le trading devient plus complexe. Elle est désormais au moins partiellement automatisée : 47% des volumes négociés et 65% du nombre de transactions sont exécutés via une interface de programmation d'applications (API) sur EPEX Spot. La part du trading automatisé, y compris le trading à haute fréquence, est en augmentation. Cela permet également d'appliquer des techniques d'apprentissage automatique sur les marchés de l'électricité. L'un des problèmes est que ces outils peuvent s'accompagner de stratégies inattendues qui ne sont pas transparentes [98]. Ces stratégies pourraient avoir des répercussions sur les volumes échangés et la formation des prix, notamment un risque accru de collusion algorithmique tacite [99].

B.4 Conclusions

La transition énergétique est une nécessité qui doit se faire à l'échelle mondiale et à un rythme accéléré. L'Union européenne fait preuve de détermination en promouvant des politiques pour guider cette transition. La voie souhaitable vers un avenir durable ne doit pas se fonder exclusivement sur les technologies, mais prendre en compte d'autres aspects tels que la sobriété énergétique et l'équité sociale. Néanmoins, la technologie jouera un rôle clé, en particulier l'électricité à faible teneur en carbone. Le système électrique devra évoluer de manière significative au cours des prochaines décennies sous l'impulsion de facteurs tels que les énergies renouvelables, les batteries, l'hydrogène et la gestion de la demande. L'utilisation actuelle de combustibles fossiles, notamment pour le chauffage, le refroidissement et les applications de mobilité, devra être remplacée ; le couplage sectoriel facilitera la décarbonisation. Les solutions de flexibilité sont également appelées à jouer un rôle de plus en plus important pour réduire l'intensité carbone et le réseau électrique sera essentiel pour l'intégration de ces solutions. Différentes voies vers le zéro net pour l'avenir des systèmes électriques ont été proposées à l'échelle mondiale, continentale et nationale ; nous avons souligné comment ces scénarios prospectifs convergent en ce qui concerne la place stratégique accordée à l'électrification et aux énergies renouvelables.

Le marché paneuropéen de l'énergie est développé dans l'intention de faciliter cette transition en structurant les flux énergétiques et financiers entre les acteurs du marché. Il couvre de multiples zones géographiques et échéances, ce qui permet aux acteurs du

marché de gérer leurs risques et de générer des revenus lorsque des opportunités se présentent. La coordination du marché est assurée par des institutions supranationales aux niveaux politique et réglementaire, ainsi que par les travaux collaboratifs des gestionnaires de réseau de transport et des opérateurs de marché désignés. Nous avons noté que les marchés de l'électricité doivent être conçus avec soin afin de garantir leur efficacité à court et moyen terme, tant pour les consommateurs que pour les producteurs, tout en favorisant des investissements adéquats pour un avenir plus vert. Le découplage effectif entre la croissance économique et les dommages environnementaux devra être suivi de près et les mesures prises devront être adaptées en conséquence.

Appendix C

Résumé du chapitre 3 : Formation des prix de l'électricité sur les marchés journaliers

Dans le chapitre précédent, nous avons présenté l'état actuel et l'évolution attendue du secteur électrique européen et des marchés de gros de l'électricité. Forts de ces connaissances, nous allons maintenant étudier plus spécifiquement le marché journalier de l'électricité : comment est-il structuré et quelles sont les implications pour le mécanisme de formation des prix. Après avoir présenté les règles et le fonctionnement théorique du marché journalier européen, nous étudierons les principales caractéristiques des prix observés empiriquement, qui reflètent le mécanisme de formation ainsi que les contraintes et les stratégies de négociation des acteurs du marché. Grâce à des analyses des prix historiques du marché et à la modélisation de l'influence des facteurs de prix à l'aide de modèles purement basés sur des données, nous révélerons que les marchés sont prévisibles dans une certaine mesure. Cette compréhension du marché journalier nous conduira au chapitre suivant de cette thèse à la modélisation des prix de l'électricité grâce à une approche structurelle guidée par les données.

C.1 Fonctionnement des marchés journaliers

C.1.1 Prix maximisant le bien-être social

Equilibre du marché

L'équilibre entre l'offre et la demande est trouvé par le biais d'une enchère fermée, également appelée enchère aveugle, où les participants au marché soumettent des ordres sans connaître les offres faites par les autres membres de la bourse. Un ordre de base de vente ou d'achat consiste en un volume et un prix pour un moment et un lieu donnés. En Europe, ces lieux correspondent à des zones d'enchères et cette conception du marché, appelée zonale. Outre les ordres de base, des ordres complexes contenant des contraintes supplémentaires peuvent également être soumis par les participants au marché. Tous les

ordres sont envoyés à un opérateur du marché de l'électricité désigné (NEMO) un jour avant la livraison physique et avant une heure limite. Ensuite, EUPHEMIA[100], un algorithme de couplage à prix unique, détermine quels ordres sont exécutés et lesquels sont rejetés. Sur la base des résultats du calcul, un prix de compensation du marché est défini pour chaque unité de temps du marché, qui est actuellement égale à une heure.

Tous les ordres acceptés sont négociés au prix de compensation du marché, quel que soit le prix de l'ordre soumis. On parle alors de prix uniforme.

Définition du bien-être social

EUPHEMIA est conçu pour maximiser le bien-être social, autrement dit, pour maximiser le surplus économique qui est la somme du surplus du consommateur, du surplus du producteur et du revenu de congestion. Le surplus du producteur et du consommateur correspond à la différence entre l'argent que les participants au marché étaient prêts à payer ou à recevoir en fonction de leurs ordres, et le prix réel de compensation du marché.

Ordre de mérite

L'objectif de maximisation du bien-être social signifie que les offres sont exécutées selon l'ordre de mérite. Dans le cas simplifié, pour chaque période et zone d'enchères, les ordres proposant de vendre au prix le plus bas et d'acheter au prix le plus élevé sont acceptés et le volume total d'énergie échangée est maximisé. La compensation du marché résultant de l'intersection des courbes d'offre et de demande est représentée schématiquement dans la figure C.1.

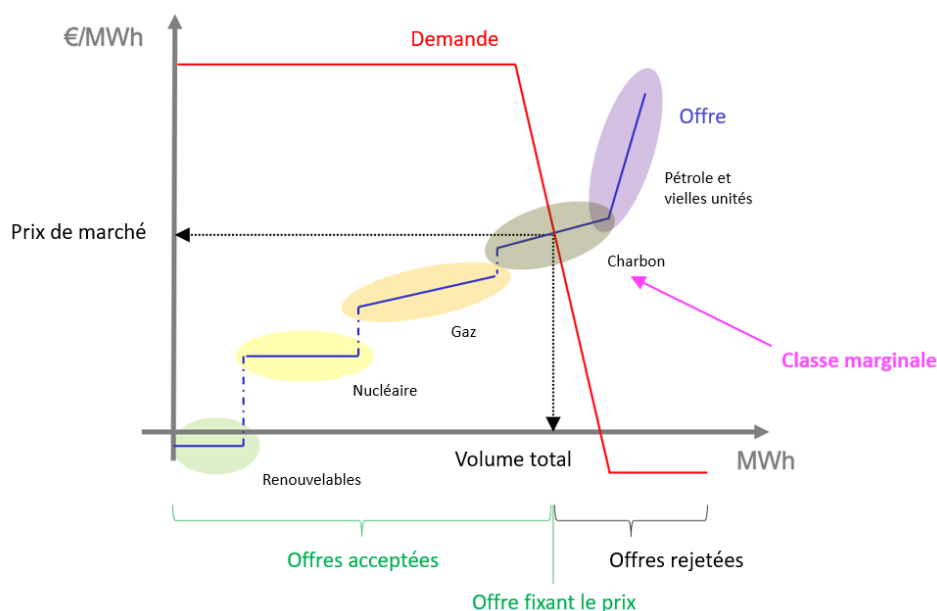


Figure C.1: Représentation simplifiée de la compensation du marché, basée sur [101]

C.1.2 Impact des coûts marginaux sur le merit order

Offres basées sur le coût marginal

Comme l'illustre la figure C.1, les ordres de fourniture sont généralement supposés correspondre à la production des actifs du système électrique. Le fait que les transactions financières sur les marchés journaliers correspondent à des transferts réels d'électricité dans le réseau accroît la pertinence de cette hypothèse.

Évolution de l'ordre de mérite

Dans un marché journalier où la concurrence est parfaite, la stratégie rationnelle des participants au marché consiste à faire des offres en fonction de leur coût marginal de production à court terme pour les fournisseurs, et à faire des offres en fonction de l'utilité marginale pour les consommateurs. Bien que tous les critères d'un marché idéal de concurrence parfaite ne soient jamais réunis dans la pratique, le coût marginal reste le principal facteur à prendre en compte dans la réflexion sur la formation des prix de l'électricité.

Une augmentation des prix du CO₂ ainsi qu'une part accrue d'énergie renouvelable à faible coût opérationnel auront donc tendance à faire baisser le nombre d'heure d'utilisation des centrales à charbon.

Types d'ordres de marché

L'ordre de marché le plus simple consiste à indiquer un sens (offre ou demande), un volume, une limite de prix, une unité de temps de marché et une zone d'enchères. Les bourses de l'électricité offrent également la possibilité de soumettre des ordres complexes et des ordres de bloc. Ces offres permettent notamment aux participants de s'assurer un volume de transaction ou un revenu minimum, ou bien de prendre en compte des contraintes relatives aux limites de rampe des moyens de production.

C.2 Interconnexion des marchés journaliers

C.2.1 Zones d'enchère

Marché zonal

Les marchés de l'électricité peuvent être organisés selon l'une des deux conceptions principales : zonale et nodale [103]. La tarification zonale, qui a été adoptée en Europe, fixe un prix unique pour une grande zone appelée zone d'enchère (généralement un pays ou une région).

Lorsque les acteurs du marché achètent ou vendent de l'électricité dans une zone d'enchère, ils fournissent ou consomment de l'électricité à l'heure de leur choix dans une zone spécifiée lors de la soumission de l'offre. En ce qui concerne le marché journalier, les contraintes techniques imposées par le réseau électrique à l'intérieur d'une zone ne

sont pas prises en compte. Les zones d'enchères correspondent généralement à un pays entier.

Interconnexion des zones d'enchère

Les zones d'enchère sont interdépendantes car elles sont physiquement interconnectées par des lignes et des câbles. Le réseau d'interconnexion se développe à travers l'Europe pour faciliter la mise en commun des ressources, ce qui profite à la fois à l'intégration des sources d'énergie à faible teneur en carbone et à la sécurité d'approvisionnement. L'algorithme EUPHEMIA est résolu en deux étapes afin de prendre en compte les échanges inter-zones.

Marge disponible pour le commerce interzonal

La capacité de transfert installée d'une interconnexion est une valeur statique, mais sa capacité disponible est dynamique car elle dépend de phénomènes physiques, qui varient en fonction de l'état du système électrique (lois de Kirchhoff). L'évaluation précise de cette capacité disponible est une tâche complexe pour les gestionnaires de réseaux de transport car cette évaluation dépend de modèles de réseaux dédiés, de prévisions de demande et de consommation, ainsi que de marges de sécurité. Ces marges sont nécessaires en raison des déviations inévitables des prévisions, qui sont dues à l'incertitude et aux imprévus (par exemple, des pannes soudaines).

La capacité d'interconnexion doit être acquise par le biais d'une vente aux enchères explicite, dans laquelle les participants au marché doivent se procurer activement cette capacité en plus du commerce de l'énergie. L'attribution peut également être implicite lorsque la compensation du marché multizone attribue conjointement l'énergie et la capacité de transmission aux participants au marché. La méthode d'allocation varie en fonction de la frontière et de la période considérées. La deuxième option, qui garantit une meilleure utilisation globale de la capacité disponible, est un élément clé du couplage des marchés. La généralisation du couplage des marchés est un élément central pour le développement du marché intérieur de l'électricité pour la période correspondant au marché journalier.

C.3 Caractéristiques des prix de l'électricité

C.3.1 Prévisibilité des prix sur les marchés journaliers

Les séries chronologiques des prix horaires de l'électricité sur les marchés journaliers présentent des caractéristiques typiques qui sont bien documentées dans la littérature scientifique [112, 113, 114].

La spécificité de l'électricité en tant que produit de base (difficulté à la stocker à grande échelle, faible élasticité de la demande, livraison sur réseau, etc.) et la spécificité du marché de l'électricité (offre oligopolistique, importance stratégique de l'électricité

pour les nations, etc.) impliquent que des modèles spécifiques soient développés afin de capturer précisément ces particularités.

Les phénomènes sont illustrés grâce à des données relatives au marché français pour la période de quatre ans allant de 2015 à 2018.

C.3.2 Analyse univariée des séries temporelles de prix

Les caractéristiques les plus notables des prix de l'électricité sont : (i) sa forte volatilité par rapport au prix d'autres commodités, (ii) la présence de périodicités aux échelles journalière, hebdomadaire et saisonnière, (iii) l'existence de pics de prix suivis de retour à la moyenne, (iv) la présence de prix négatifs.

C.3.3 Analyse multivariée : Influence des facteurs de prix

Étant donné que la demande d'électricité est encore largement insensible aux prix et que la capacité d'approvisionnement ne varie pas sur de courtes échelles de temps, il existe une corrélation positive entre la demande agrégée et les prix du marché ($R^2 = 27,5\%$).

Le type de ressources mobilisées pour produire de l'électricité influence les prix du marché par l'effet de l'ordre de mérite. La corrélation négative entre la part de la production d'électricité renouvelable et les prix du marché ($R^2 = 18,7\%$).

Une autre façon d'évaluer l'impact des contraintes du système électrique sur les prix de gros est de considérer la rareté de la capacité d'offre par rapport à la demande. Pour ce faire, nous pouvons construire une variable de substitution appelée "marge d'approvisionnement", qui est égale à la différence entre la capacité disponible des unités de production dispatchables dans une zone de soumission et la demande résiduelle agrégée ($R^2 = 45,1\%$).

C.3.4 Modélisation statistique multivariée des prix

Modélisation des prix : principes clés

L'ensemble des données (prix de l'électricité et leurs déterminants) est généralement utilisé d'un seul bloc pour mener des analyses telles que celles que nous avons présentées. Cependant, pour les études de modélisation, cet ensemble de données est généralement divisé en au moins deux parties distinctes. La première partie est utilisée pour ce que l'on appelle la "phase d'entraînement", au cours de laquelle les paramètres du modèle sont estimés par une régression entre la variable de résultat et les variables explicatives (qui correspondent ici respectivement aux prix de l'électricité et aux déterminants des prix). En utilisant ces paramètres estimés et les facteurs de prix de la deuxième partie de l'ensemble de données, le modèle peut ensuite être utilisé pour simuler les prix pendant la "phase de test". La qualité de la simulation peut être évaluée par rapport aux prix observés grâce à la visualisation (en traçant à la fois l'observation et la simulation) et à l'utilisation de mesures statistiques. Cette pratique contribue à éviter le surapprentissage et nous informe sur la valeur prédictive d'un modèle.

Pour cette étude, nous proposons d'évaluer la performance prévisionnelle des modèles sur la dernière année de l'ensemble de données (c'est-à-dire 2018) et d'utiliser les trois années restantes pour l'entraînement (c'est-à-dire de 2015 à 2017). Pour les variables explicatives, nous conservons les déterminants de prix présentés dans les analyses multivariées : demande prévue, part de renouvelable dépendant des conditions météorologiques et marge d'approvisionnement. Concernant les modèles, nous proposons de comparer une des approches statistiques les plus simples utilisées comme référence naïve à une méthode d'apprentissage automatique supervisée largement utilisée pour la prévision dans divers domaines. Pour ce faire, nous effectuerons des régressions linéaires multiples et des régressions par forêt d'arbres de décision (*random forest*), en utilisant l'implémentation dans le langage de programmation R (respectivement la fonction "lm" du paquet "stats" v.3.6.1 et la fonction "rf" du paquet "randomForest" v4.6). En ce qui concerne les métriques utilisées pour vérifier la qualité des simulations sur les tests hors échantillon, nous considérons principalement l'erreur quadratique moyenne (*RMSE*), qui évalue la précision de la prévision globale. De plus, pour estimer si le modèle a correctement capturé la dynamique des prix (c'est-à-dire les variations horaires), nous évaluons la différence entre l'écart-type de la série temporelle de prix observée et celle simulée (Δsd). Ces métriques, qui évaluent la performance de la prévision $(\hat{\Pi}_t)_{t \in T}$ par rapport aux prix observés $(\Pi_t)_{t \in T}$ sont définies comme suit :

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (\Pi_t - \hat{\Pi}_t)^2} \quad (C.1)$$

$$\Delta sd = sd((\Pi_t)_{t \in T}) - sd((\hat{\Pi}_t)_{t \in T}) \quad (C.2)$$

Résultats de la simulation

Les résultats de la simulation pour les deux modèles sont présentés dans la figure C.2 et la figure C.3. Nous observons que dans les deux cas, les prix sont sous-estimés pour la deuxième partie de l'année. Cela est probablement dû à l'impact de l'augmentation des prix du gaz et du CO₂, qui ne sont pas pris en compte parmi les variables d'entrée. Cela souligne l'importance de sélectionner soigneusement les déterminants de prix, c'est pourquoi des facteurs supplémentaires seront pris en compte dans la suite de la thèse. En raison du nombre limité de variables choisies pour cet exemple d'introduction, les performances de prévision sont assez similaires pour les deux modèles, car même un modèle linéaire simple réussit à extraire des informations de la demande prévue, de la part de vRES et ensuite de la marge d'approvisionnement. Nous obtenons le même *RMSE* dans les deux cas, égal à 19.8. Néanmoins, l'apprentissage automatique capture mieux la dynamique que le modèle linéaire ($\Delta sd = 3.0$ pour le modèle linéaire et $\Delta sd = 2.1$). Nous observons que l'algorithme de forêt d'arbres de décision est particulièrement performant pour capturer les pics de prix.

Nous soulignons l'importance d'étudier les résultats sur des ensembles de données hors échantillon, car les prédictions en échantillon peuvent résulter d'un surapprentissage.

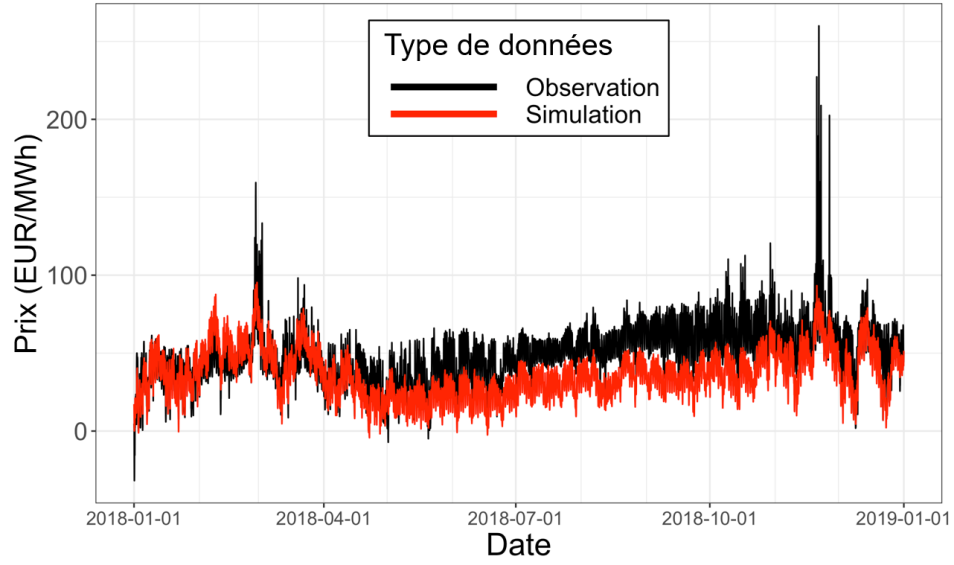


Figure C.2: Prix sur le marché journalier français : modèle linéaire multivarié

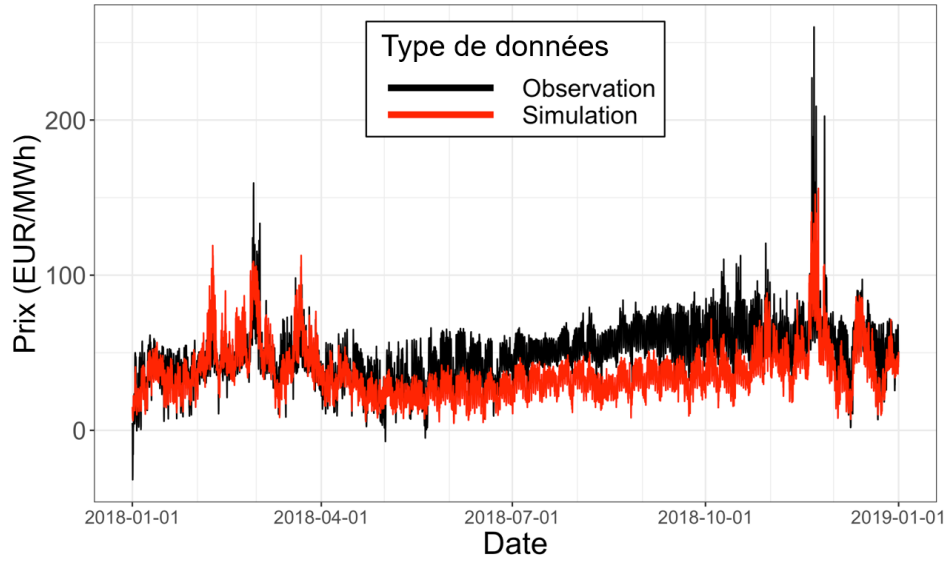


Figure C.3: Prix sur le marché journalier français : forêt d'arbres de décision

Dans cet exemple, nous avons obtenu des performances significativement meilleures lors de la phase d'apprentissage par rapport à la phase de test, notamment avec la forêt d'arbres de décision (pour le modèle linéaire : $RMSE = 13,6$ et $\Delta sd = 5,3$, pour la forêt d'arbres de décision : $RMSE = 6,3$ et $\Delta sd = 3,2$).

C.4 Conclusions

Le marché européen journalier de l'électricité s'appuie sur un mécanisme complexe pour assurer la maximisation du bien-être social. Une enchère en double aveugle organisée par les bourses d'électricité permet de fixer un prix de marché unique pour chaque heure et chaque zone d'enchère. L'algorithme EUPHEMIA est responsable de la compensation du marché à travers l'Europe, qui détermine les offres exécutées selon l'ordre de mérite. La construction du marché intérieur de l'énergie se poursuit en Europe grâce à davantage d'harmonisation (par exemple, la réglementation sur l'allocation des capacités et la gestion de la congestion), d'intégration (par exemple, le couplage unique journalier) et d'adaptations pour tenir compte d'une part plus importante d'énergies renouvelables (par exemple, le raccourcissement de l'unité de temps du marché).

Le marché intérieur de l'énergie permet de mettre en commun les ressources dans toute l'Europe, ce qui permet d'utiliser les actifs de manière efficace. Le niveau élevé de liquidité et la convergence des prix entre les différentes zones géographiques sont des indicateurs du bon fonctionnement du marché. Néanmoins, des défis subsistent : la conception du marché zonal et la tarification uniforme présentent des inconvénients qui doivent être traités par des mesures complémentaires au marché journalier. La gestion de la congestion, l'adéquation du système et l'acceptabilité sociale sont des questions clés pour l'avenir du système électrique et des marchés de l'électricité. L'évolution de la structure des coûts du mix électrique, qui comprend des énergies renouvelables à faible OPEX et des technologies à base de combustibles fossiles dont les coûts marginaux à court terme sont plus élevés, nous amène à reconsidérer les avantages et les inconvénients de la tarification marginale.

La spécificité de l'électricité en tant que produit (livraison liée au réseau, difficulté de stockage à l'échelle, etc.) est à l'origine des caractéristiques des prix sur les marchés journaliers. L'analyse de la dynamique des prix de l'électricité révèle l'existence d'une forte volatilité, de périodicités à l'échelle quotidienne, hebdomadaire et annuelle, de pics suivis d'un retour à la moyenne et de prix négatifs. D'autres analyses multivariées confirment le lien entre l'état du système électrique et les prix observés sur les marchés de gros. Les modèles statistiques et d'apprentissage automatique peuvent être utilisés pour la prévision, en particulier pour les horizons à court terme, mais ils ne nous permettent pas de tirer pleinement parti de notre compréhension des marchés de l'électricité. Les modèles structurels, qui reflètent le mécanisme de formation des prix selon une approche ascendante, permettent de mener des études plus détaillées, ils feront donc l'objet du prochain chapitre.

Appendix D

Résumé du chapitre 4 : Modélisation structurelle des prix basée sur les données

Dans ce chapitre, nous explorons une nouvelle catégorie de modèles, appelés modèles structurels ou fondamentaux, qui reflètent le mécanisme réel de formation des prix (prix uniforme et prix marginal).

D.1 Revue de la littérature

Diverses approches ont été proposées dans la littérature pour explorer la relation entre les prix de l'énergie et les facteurs influençant l'évolution des prix [6].

Pour les modèles structurels, une étape clé est la modélisation de la courbe d'offre. Celle-ci est généralement obtenue par agrégation des ordres de marché, mais elle peut aussi résulter d'un modèle statistique comme dans [127], où la courbe est construite en fonction d'une charge normalisée et des prix de marché observés. Ce modèle purement statistique est ensuite ajusté pour tenir compte de l'évolution des coûts des combustibles et des émissions pour le charbon. Le modèle X présenté dans [7] et utilisé dans [128] propose un modèle par morceaux pour la courbe d'approvisionnement. Chaque modèle local est initialement associé à un niveau de prix pour lequel un volume d'énergie est déterminé par une régression statistique. Les régresseurs au temps t comprennent la production solaire, la production éolienne et la production planifiée. Une interpolation est utilisée pour construire la courbe d'offre finale. Le concept du modèle X a récemment été étendu dans [129], où les courbes d'offre et de demande sont considérées conjointement afin de réduire le temps de calcul et d'améliorer la précision. Bien que ces modèles par morceaux tentent d'imiter l'approche standard de l'ordre de mérite, ils n'incluent pas les effets de la disponibilité et des coûts de production. Au contraire, un modèle inspiré de [130] est proposé dans [8] pour les marchés couplés de l'électricité, où la courbe d'offre est construite en empilant des ordres de vente pour lesquels les offres de prix sont une fonction de la marge d'offre (c'est-à-dire la différence entre la capacité totale disponible

et la demande) et du coût de production par technologie. Un objectif de l'article de ces auteurs est l'analyse des marchés à terme, qui nécessite une forme spécifique d'offre de prix qui réduit le réalisme de la dynamique des prix journaliers dans les simulations. Dans l'article [131], une procédure de prix de marché de compensation, qui inclut également des ordres complexes, est proposée. Les offres ne sont pas paramétrées et aucun algorithme d'apprentissage n'est introduit puisque l'étude de cas est basée sur le marché espagnol, où les détails des offres du marché journalier sont divulgués contrairement à la plupart des marchés européens de l'électricité. Si l'approche est pertinente pour les analyses ex post, il faudrait l'étendre pour l'utiliser dans des études prospectives. Dans [9], une attention particulière est accordée au réalisme de la variabilité des simulations. Leur modèle commence par une description de la courbe d'offre par technologie à travers le coût marginal à court terme. Les offres de prix sont ensuite réduites en cas de faible demande (reflétant par exemple les contraintes de production ininterrompue) et augmentées pour une demande élevée (induisant une rente de rareté). L'ajustement des offres se fait toujours au niveau de la technologie sans tenir compte des unités de production individuelles. Les références [10] et [11] proposent également de modéliser les prix horaires en utilisant l'approche de l'ordre de mérite au niveau technologique afin de modéliser de manière réaliste la dynamique des prix sur le marché journalier allemand. Les deux modèles utilisent des données réelles et des régressions linéaires. Les régressions sont utilisées dans le modèle [10] pour ajuster la sortie du modèle d'ordre de mérite afin de tenir compte d'effets tels que les contraintes de rampe et les enchères stratégiques, tandis que le modèle [11] utilise des statistiques pour modéliser la balance du commerce extérieur. Ces modèles utilisent des données pour améliorer le réalisme de simulation, mais ils ne proposent pas d'utiliser directement la calibration statistique pour la modélisation de la courbe d'offre.

D.2 Énoncé du problème

Bien qu'il existe de nombreuses solutions pour modéliser le système électrique complet et les contraintes des unités afin d'effectuer des études pour des simulations à long terme, peu de méthodologies ont été proposées pour extraire des informations des prix journaliers observés afin d'ajuster les paramètres internes de ces modèles. La nouveauté de cette approche est que les informations extraites des ensembles de données (par exemple, les observations passées) deviennent une partie intégrante du modèle, qui peut être utilisée pour les prévisions hors échantillon (par exemple, pour des études prospectives).

Dans le reste de ce chapitre, nous proposons un modèle structurel dans lequel la courbe d'offre est construite selon une approche ascendante. Les ordres de marché sont associés aux unités de production et leurs prix sont paramétrés, ce qui signifie que nous pouvons exploiter les données disponibles sur le marché et le système électrique. Pour développer cette approche conjointe de statistique et d'optimisation, nous introduisons une méthode de simulation des prix horaires de l'électricité sur les marchés journaliers qui répond à deux objectifs : (i) refléter les contraintes technico-économiques imposées par le système électrique qui induisent la formation des prix selon la théorie économique

du prix marginal dans des conditions idéalisées de concurrence parfaite, et (ii) exploiter les données de marché et du système électrique pour calibrer et évaluer le modèle afin que les simulations reflètent les prix observés en pratique. La méthode est présentée puis validée à l'aide de données relatives au marché journalier et au système électrique français pour la période de 2015 à 2018.

D.3 Méthodologie

D.3.1 Equilibre du marché avec contraintes techniques

Le lien entre les contraintes techniques du système électrique et les prix de l'électricité est modélisé par un simple dispatch économique incluant une contrainte de stock pour l'énergie hydraulique. Cette contrainte introduit un couplage temporel qui empêche une résolution pour chaque pas de temps individuellement. Des contraintes plus techniques pourraient être ajoutées ici, bien que ce ne soit pas l'objectif de ce chapitre d'ajouter trop de complexité ici. Pour une formulation et une discussion plus complètes sur la répartition économique, voir par exemple la revue de littérature [132]. La formulation pourrait être étendue à un unit commitment comme présenté dans la revue de littérature [133]. Une demande élastique est introduite dans la formulation, ce qui la rend plus proche d'une compensation de marché avec des contraintes techniques que d'un dispatching économique pur. La fonction de coût qui est optimisée dépend des coûts marginaux $\pi_{t,i}^P$ des unités de production, et de l'utilité marginale $\pi_{t,j}^L$ des unités de consommation. Le problème est formulé comme une maximisation du bien-être social, c'est-à-dire une maximisation du surplus du producteur et du surplus du consommateur :

$$\begin{aligned} \underset{P, L, E}{\text{maximize}} \quad & \sum_{t \in T} \left(\sum_{j \in J} \pi_{t,j}^L L_{t,j} - \sum_{i \in I} \pi_{t,i}^P P_{t,i} \right) \end{aligned} \quad (\text{D.1a})$$

$$\text{subject to} \quad \sum_{i \in I_z} P_{t,i} + \sum_{z' \in Z} E_{t,z,z'} = \sum_{j \in J_z} L_{t,j} \quad \forall (t, z) \in T \times Z, \quad (\text{D.1b})$$

$$E_{t,z,z'}^{min} \leq E_{t,z,z'} \leq E_{t,z,z'}^{max} \quad \forall (t, z, z') \in T \times Z \times Z, \quad (\text{D.1c})$$

$$L_{t,j}^{min} \leq L_{t,j} \leq L_{t,j}^{max} \quad \forall (t, j) \in T \times J, \quad (\text{D.1d})$$

$$P_{t,i}^{min} \leq P_{t,i} \leq P_{t,i}^{max} \quad \forall (t, i) \in T \times I, \quad (\text{D.1e})$$

$$\sum_{(t,i) \in T \times I_z^{hydro}} P_{t,i} \leq S_{T,z}^{hydro} \quad \forall z \in Z \quad (\text{D.1f})$$

où

- (1a) est l'objectif de maximisation du bien-être social
- (1b) est la contrainte d'équilibre des puissances
- (1c) est la contrainte de capacité d'interconnexion

- (1d) est la contrainte de limite des unités de consommation
- (1e) est la contrainte de limite des unités de production
- (1f) est la contrainte du stock d'hydroélectricité

Nous désignons par $\Pi_{t,z}^{MC}$ le multiplicateur de Lagrange associé à la contrainte (D.1b). Dans le cas sans couplage temporel, les coûts marginaux $\pi_{t,i}^P$ sont simplement les prix de vente des participants au marché et $\Pi_{t,z}^{MC}$ est le résultat d'un équilibre du marché. Dans une vision de marché pur, l'ensemble des unités de production I serait l'ensemble des commandes plutôt qu'un ensemble de moyens de production. Avec une contrainte technique, telle que la contrainte de stock d'hydro, le prix de vente de l'hydro est la somme du coût marginal $\pi_{t,i}^P$ et du multiplicateur de Lagrange $\lambda^{(1f)}$ associé à la contrainte de stock (D.1f), également appelé valeur de l'eau. L'intérêt de ce type de formulation par rapport à une pure compensation de marché est qu'un prix de vente tenant compte des contraintes techniques est calculé en interne par le modèle. Dans le cas général, les multiplicateurs de Lagrange associés à toutes les contraintes techniques internes incluses dans le modèle peuvent être affectés aux moyens de production, de sorte que l'on peut écrire le prix de vente de l'unité de production i au temps t comme $\pi_{t,i}^P + \lambda_{t,i}^{IC}$.

D.3.2 Paramétrage des coûts marginaux

Dans la pratique, les prix de vente ne sont pas seulement influencés par les contraintes technico-économiques bien définies IC qui sont prises en compte dans l'équation (D.1), mais aussi par les décisions stratégiques des participants au marché qui ont un impact sur la formation des prix. La contribution de ces autres facteurs est notée $\epsilon_{t,i}$. Nous proposons enfin la décomposition additive suivante des prix de vente $\hat{\pi}_{t,i}^{sell}$, qui tient compte de la contribution à la fois de $\lambda_{t,i}^{IC}$ et de $\epsilon_{t,i}$:

$$\hat{\pi}_{t,i}^{sell} = \pi_{t,i}^P + \lambda_{t,i}^{IC} + \epsilon_{t,i} \quad (D.2)$$

Les contraintes internes du modèle IC sont explicites, de sorte que le modélisateur peut évaluer directement l'effet de l'ajout ou de la suppression de contraintes sur les prix simulés (par exemple, en ajoutant des contraintes de rampe aux contraintes existantes). Le modélisateur peut aussi facilement évaluer l'impact des limites imposées (par exemple en changeant les valeurs de la contrainte du stock d'hydro ou de la contrainte de la capacité de production maximale disponible). Mais cette approche n'est pas applicable pour le $\epsilon_{i,t}$ car il n'est pas pratique de quantifier a priori l'impact sur les prix horaires journaliers de facteurs économiques tels que les enchères stratégiques, l'asymétrie d'information, le pouvoir de marché et la coexistence de plusieurs marchés de gros de l'électricité. L'utilisation d'une approche basée sur les données des séries temporelles historiques nous aide à quantifier empiriquement l'influence de ces facteurs. Nous proposons donc de construire un modèle statistique linéaire $\hat{\pi}_{t,i}^P$ pour $\pi_{t,i}^P + \epsilon_{t,i}$. L'approximation du coût marginal tient compte à la fois du coût marginal objectif à court terme de l'unité de production i ($\pi_{t,i}^P$) et de l'influence des facteurs stratégiques

($\epsilon_{t,i}$). Les principaux facteurs de ces approximations du coût marginal sont les suivants : prix du carburant, prix des émissions, marges d’approvisionnement globales (c’est-à-dire la différence entre la capacité disponible et la demande), et le rang de chaque unité de production dans sa classe de production, et nous proposons donc la paramétrisation décrite dans l’équation (D.3).

$$\begin{aligned}\hat{\pi}_{t,i}^P &= \alpha_{c(i),0} + \alpha_{c(i),\mathcal{F}} \cdot \mathcal{F}_{t,c(i)} + \alpha_{c(i),\mathcal{E}} \cdot \mathcal{E}_t + \alpha_{c(i),\mathcal{M}} \cdot \mathcal{M}_t + \alpha_{c(i),\mathcal{R}} \cdot \mathcal{R}_i \\ &= \boldsymbol{\alpha}_{c(i)}^\top \mathbf{X}_{t,i}\end{aligned}\tag{D.3}$$

où les paramètres du modèle sont uniquement une fonction de la classe d’unité de production $c(i)$:

- $\alpha_{c(i),x}$ - Paramètres calibrés par la classe $c(i)$ pour chaque déterminant de prix x .
- $\mathcal{F}_{t,c(i)}$ - Prix du carburant au moment t pour la classe $c(i)$.
- $\mathcal{E}_t$ - Prix des émissions au moment t .
- $\mathcal{M}_t$ - Marge d’approvisionnement au moment t .
- $\mathcal{R}_i$ - Variable de substitution pour le rang de l’unité de production i dans la classe $c(i)$.

Comme le prix de vente est obtenu en ajoutant $\lambda_{t,i}^{IC}$ à $\hat{\pi}_{t,i}^P$, le prix que nous proposons est également un modèle statistique des ordres de vente. L’utilisation du proxy de coût marginal dans le problème d’optimisation défini dans l’équation (D.1) aboutit à la simulation d’un prix de marché journalier que nous désignons par $\hat{\Pi}_{t,z}^{MC}$.

D.3.3 Correction des biais

Nous introduisons également un terme de correction de biais $\beta_{t,z}$ pour obtenir finalement les prix de marché simulés $\hat{\Pi}_{t,z}$ étudiés dans ce chapitre que nous définissons comme suit :

$$\begin{aligned}\hat{\Pi}_{t,z} &= \hat{\Pi}_{t,z}^{MC} + \beta_{t,z} \\ &= \hat{\Pi}_{t,z}^{MC} + \boldsymbol{\beta}_z^\top \mathbf{Y}_t\end{aligned}\tag{D.4}$$

où pour chaque région z , $\boldsymbol{\beta}_z \in \mathbb{R}^{24 \times 7}$ est un paramètre, et pour chaque pas de temps t $\mathbf{Y}_t$ est le vecteur de dimension 24×7 avec $Y_t[(h, d)] = 1$ si $hour(t) = h$ et $weekday(t) = d$; sinon $Y_t[(h, d)] = 0$.

D.3.4 Algorithme d’estimation

Les paramètres de l’équation (D.3) et de l’équation (D.4) doivent être estimés à partir de données réelles. Pour un pas de temps t pour lequel on sait que l’unité de production

$i^*(t)$ est marginale, le prix de vente at-the-money établit le prix de marché simulé, qui peut être réécrit comme suit :

$$\begin{aligned}\hat{\Pi}_{t,z} &= \hat{\pi}_{t,i^*(t)}^P + \lambda_{t,i^*(t)}^{IC} + \beta_{t,z} \\ &= \boldsymbol{\alpha}_{c(i^*(t))}^\top \mathbf{X}_{t,i^*(t)} + \lambda_{t,i^*(t)}^{IC} + \boldsymbol{\beta}_z^\top \mathbf{Y}_t\end{aligned}\tag{D.5}$$

Nous démarrons l'algorithme avec $\boldsymbol{\beta}_z = \mathbf{0}$, et $\boldsymbol{\alpha}_{c(i)} = \boldsymbol{\alpha}_{c(i)}^0$ comme valeur initiale.

Etape d'équilibre du marché À partir d'un ensemble d'approximations du coût marginal, nous pouvons résoudre l'équation (D.1). Cela conduit à une estimation $\hat{\Pi}_{t,z}$ de $\Pi_{t,z}$, et à l'identification des unités de production marginales. Ceci nous permet de regrouper les pas de temps en fonction de la valeur de la classe marginale $c(i^*(t))$ en désignant la partition associée par $(T_c)_{c \in C}$.

Mise à jour des paramètres par l'estimation statistique Les résidus $\Pi_{t,z} - \hat{\Pi}_{t,z}$ peuvent être utilisés pour estimer une nouvelle valeur pour $\boldsymbol{\beta}_z$. Parallèlement, $\boldsymbol{\alpha}_{c_0}$ est obtenu par une régression linéaire de $(\mathbf{X}_{t,i})_{(t,i) \in T_{c_0} \times I_z}$ sur le prix observé $(\Pi_{t,z})_{t \in T_{c_0}}$.

D.4 Validation de l'approche sur une étude de cas

Nous avons implémenté une méthode suivant l'approche décrite, avec le langage de programmation R, en utilisant des données relatives au marché français de 2015 à 2018. Dans notre étude de cas, toutes les unités de production présentes dans le système électrique sont incluses dans le problème d'optimisation défini dans l'équation (D.1), c'est-à-dire que nous supposons que toutes les unités participent au marché journalier. Comme nous considérons ici le cas d'une seule zone d'enchère, l'indice z sera omis dans cette section.

D.4.1 Implémentation de la méthode

D.4.2 Vue d'ensemble

Nous simulons les prix dans une seule zone de soumission et incorporons l'effet des zones voisines interconnectées comme ordres supplémentaires dans les courbes d'offre et de demande. Pour l'application numérique, nous utilisons les valeurs de capacité de transfert nettes observées comme quantités de commande (à la fois comme $P_{t,i}^{max}$ pour tenir compte de la possibilité d'importer et comme $L_{t,j}^{max}$ pour tenir compte de la possibilité d'exporter) et les prix du marché journalier étranger comme prix de commande ($\pi_{t,i}^P$ et $\pi_{t,j}^L$ respectivement pour l'offre et la demande). Aucun volume minimal d'importation ou d'exportation n'est imposé (c'est-à-dire que $P_{t,i}^{min}$ et $L_{t,j}^{min}$ sont égaux à 0 MWh).

La figure D.1 donne un aperçu de l'étude de cas.

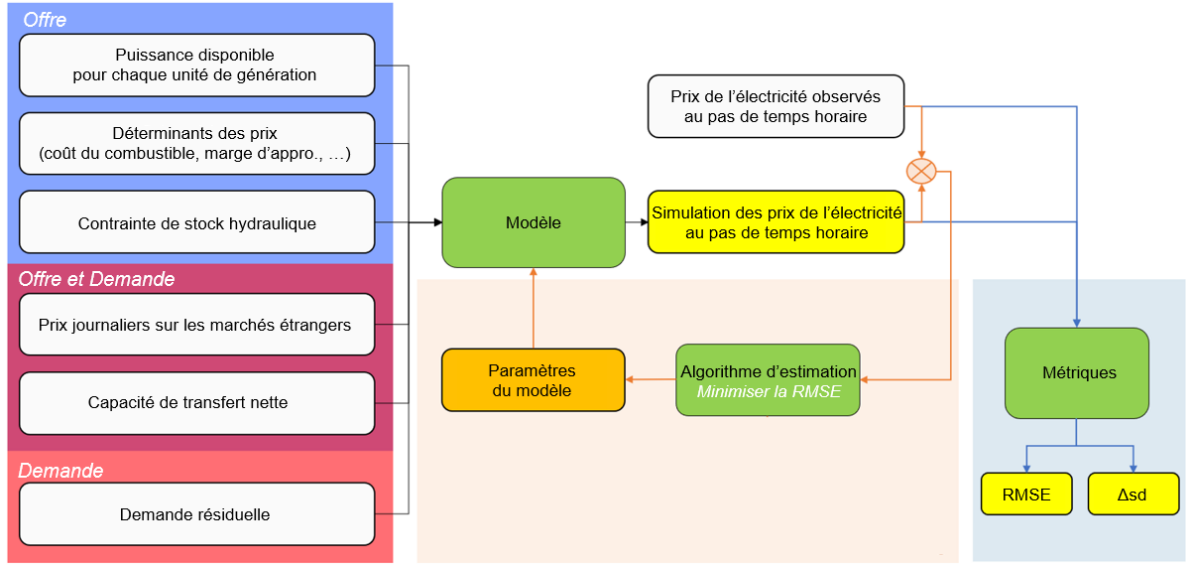


Figure D.1: Vue d'ensemble de l'étude de cas

D.4.3 Approvisionnement

Pour les ordres qui constituent la courbe d'approvisionnement, nous considérons les unités de production de cinq classes : nucléaire, gaz fossile, charbon fossile, pétrole fossile et réservoir d'eau hydraulique. Chaque commande correspond à une unité de production d'une capacité installée de plus de 100 MW. Les quantités fournies chaque heure correspondent à la capacité disponible $P_{t,i}^{max}$. La valeur de $P_{t,i}^{min}$ est fixée à 0 MWh pour toutes les unités.

D.4.4 Estimation des paramètres

Pour chaque classe, nous initialisons $\alpha_{c,0}$ introduit dans l'équation (D.3) comme étant constant et égal à un coût marginal moyen à court terme (c'est-à-dire le coût associé à la production d'énergie supplémentaire). Pour éviter un ajustement excessif, nous contraignons la régression à imposer un signe aux coefficients $\alpha_{c,x}$ pendant la calibration, de sorte qu'une augmentation du coût du combustible ou une diminution de la marge d'approvisionnement entraîne une augmentation du prix de l'énergie. Nous ne calibrons pas les paramètres $\alpha_{c(i),\varepsilon}$ associés aux émissions et choisissons plutôt de nous appuyer sur les données de RTE, le gestionnaire du réseau de transport français.

Pour estimer les β introduits dans l'équation (D.4), nous calculons la moyenne des résidus sur l'ensemble d'apprentissage par heure et jour de la semaine (c'est-à-dire les valeurs de décalage 24x7) et nous appliquons ensuite cette correction sur les tests hors échantillon.

D.4.5 Demande résiduelle

Nous considérons une demande résiduelle agrégée pour la zone d'enchère $L_{t,j}^{max}$ avec un prix fixé au plafond de prix du marché journalier (c'est-à-dire que $\pi_{t,j}^L$ est égal à 3 000 EUR par MWh). Pour la simulation, nous avons fixé la valeur de $L_{t,j}^{min}$ à 0 MWh. La demande $L_{t,j}^{max}$ résulte de la différence entre la consommation prévue par RTE à J-1 et la contribution de l'offre supposée être preneuse de prix pour la période et le lieu étudiés. Cette offre est ici composée de la production d'électricité fatale (solaire, éolienne et hydraulique au fil de l'eau), de la cogénération (production observée selon un profil de base), du stockage (on ne considère que la contribution de l'hydroélectricité de pompage, pas les batteries, etc.) et des classes de production à faible puissance installée (biomasse).

D.4.6 Données

Les ensembles de données brutes utilisés pour la simulation et l'analyse des résultats sont répertoriés dans le tableau D.1.

Jeu de données	Source
Capacité installée par classe de production	Plateforme ENTSO-E ¹
Capacité installée des unités de production ($P_{nom} > 100\text{MW}$)	
Indisponibilité des unités de production ($P_{nom} > 100\text{MW}$)	
Prix sur les marchés journaliers	Open Power System Data [134]
Prévision J-1 de la consommation	RTE (Plateforme Eco2Mix) ²
Production horaire par classe de production	
Émissions de CO2 par classe de production	
Capacité nette de transfert vers et depuis les zones d'enchère (NTC)	RTE (portail de données) ³
Coûts d'exploitation par classe de production	Etude RTE/ADEME [135]
Coûts d'importation du charbon et du pétrole en France	Ministère français de la transition écologique (PEGASE) ⁴
Prix du marché du gaz	ERCE ⁵
Quotas d'émission européens de CO2	Business Insider ⁶
Température atmosphérique	Meteo France

Table D.1: Sources de données pour l'étude de cas

¹<https://transparency.entsoe.eu/>

²<https://www.rte-france.com/eco2mix/telecharger-les-indicateurs>

³<https://www.services-rte.com/>

⁴<http://developpement-durable.bsocom.fr>

⁵<https://www.erce.energy/graph/uk-natural-gas-nbp-spot-price>

⁶https://markets.businessinsider.com/commodities/historical-prices/co2-european-emission-allowances/euro/1.12.2014_1.2.2019

D.4.7 Résultats

Les résultats de la simulation sont obtenus par des tests hors échantillon. Pour ce faire, une année est utilisée pour la formation, puis nous testons sur chacune des années restantes (nous répétons ce processus pour les quatre années civiles de l'ensemble de données). De cette façon, chaque année est utilisée trois fois comme ensemble de test. Le résultat final pour chaque année, présenté tout au long du chapitre, est la moyenne de l'ensemble des trois simulations alternatives dans lesquelles cette année est utilisée comme ensemble de test (par exemple, en utilisant les modèles formés avec les ensembles de données 2015, 2016 et 2017, nous simulons trois séries temporelles distinctes de prix pour 2018, puis nous calculons la moyenne horaire).

D.4.8 Présentation des séries temporelles simulées

Les résultats obtenus à l'aide de la méthode et des données décrites ci-dessus sont présentés dans la figure D.2. En plus de la sortie du modèle (c'est-à-dire les séries temporelles avec une résolution horaire), nous avons analysé sa variabilité en calculant les prix moyens par jour de la semaine et par heure de la journée. Les variations temporelles semblent être globalement bien représentées par le modèle.

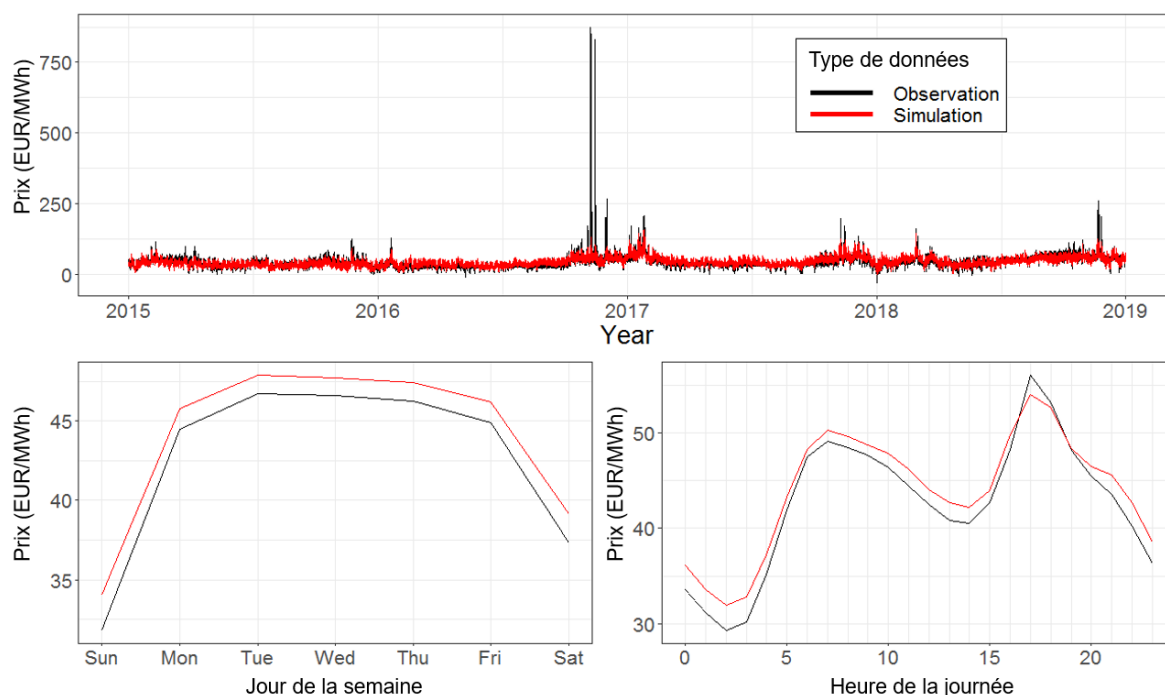


Figure D.2: Série temporelle des prix sur le marché journalier français sur quatre ans (en haut), prix moyens par jour de la semaine (en bas à gauche) et prix moyens par heure de la journée (en bas à droite)

Une mise en avant d'une période d'un mois présentée dans la figure D.3 donne une

représentation plus détaillée de la dynamique des prix journalier.

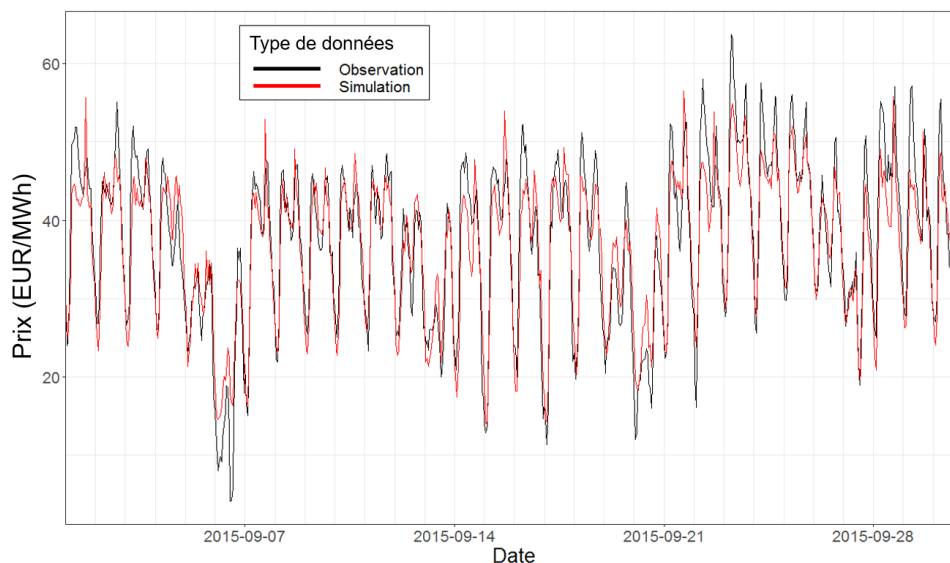


Figure D.3: Série temporelle des prix sur le marché journalier français sur un mois

D.4.9 Evaluation des simulations

Pour analyser les résultats, nous calculons d'abord les métriques et nous obtenons $RMSE = 10,8$ et $\Delta sd = 4,9$ sur l'ensemble de la période d'essai de quatre ans. Si nous limitons notre analyse à l'année 2018, nous obtenons un $RMSE = 7,5$, ce qui constitue une amélioration significative par rapport aux tests simples réalisés avec les modèles linéaire et de forêts aléatoires dans le chapitre précédent. Néanmoins, nous observons une détérioration des performances concernant l'estimation de la dynamique des prix, notamment par rapport aux forêts aléatoires, puisque nous obtenons $\Delta sd = 3,6$ avec notre modèle.

Pour aller au-delà de l'analyse des métriques, nous traçons la distribution de nos prix simulés par rapport à la distribution des prix observés. Les résultats sont affichés avec des estimations de densité lissées dans la figure D.4. Pour rendre la visualisation plus claire, les valeurs supérieures à 150 EUR/MWh ne sont pas affichées. Si les prix sont correctement modélisés en moyenne (la moyenne des prix observés est égale à 42,6 EUR/MWh et elle est égale à 44,0 EUR/MWh pour les prix simulés), nous avons observé que la dispersion des prix est sous-estimée par notre modèle. Ainsi, l'erreur mesurée par Δsd ne provient pas seulement des pics de prix mais de l'ensemble de la distribution. Dans le prochain chapitre, nous explorerons des techniques avancées d'estimation des paramètres pour atténuer cet effet.

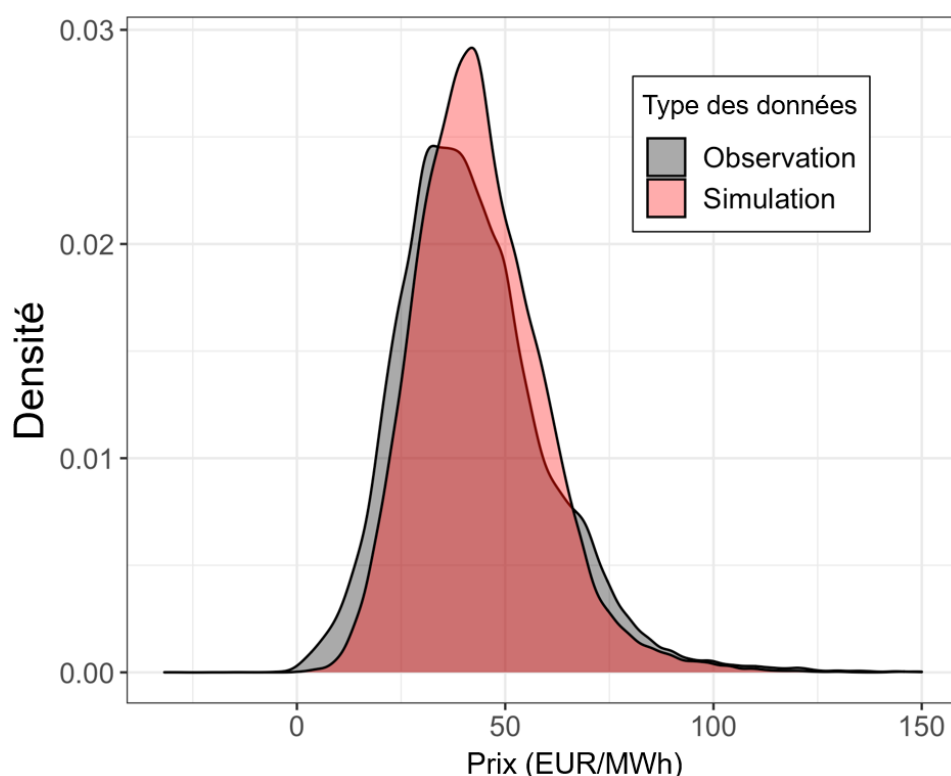


Figure D.4: Distribution des prix sur le marché journalier français de 2015 à 2018

D.5 Conclusions

Les simulations à long terme des prix de gros de l'électricité aident à soutenir les décisions d'investissement et de politique. Les modèles structurels des marchés journaliers, qui permettent de rendre compte à la fois du mécanisme de formation des prix et des contraintes technico-économiques du système électrique, sont couramment utilisés dans l'industrie pour les études prospectives. Ces modèles reposent sur des bases théoriques et permettent de trouver un prix optimal correspondant à la maximisation du bien-être social. Néanmoins, les tests empiriques des modèles contribuent à vérifier que la théorie rend compte de manière fiable des observations, et les approches basées sur les données peuvent aider à réduire l'écart entre la théorie et la pratique. Alors que de nombreux modèles développés pour la prévision des prix de l'électricité sont présentés dans la littérature académique, les méthodes spécifiquement conçues pour des considérations à long terme qui proposent de combiner des approches d'optimisation et statistiques sont moins courantes. Dans ce chapitre, nous avons proposé un modèle qui répond aux critères susmentionnés. Une approche ascendante avec une granularité élevée à partir des unités de production en ce qui concerne la courbe d'offre nous permet d'introduire une formulation détaillée et modulaire des ordres du marché, qui est appropriée pour exploiter correctement les données du système électrique et du marché. Nous avons introduit une méthode de calibration des paramètres du modèle structurel capable d'exploiter

les données réelles de manière différenciée en considérant les classes de production séparément. La méthode a été validée à l'aide de données relatives au marché français de 2015 à 2018. La dynamique globale des prix, notamment les variations horaires et hebdomadaires, a été correctement représentée par le modèle calibré sur des tests hors échantillon. Cependant l'amplitude des pics de prix est sous-estimée et pourrait faire l'objet de travaux futurs pour la formulation plus spécifique des prix des ordres. L'étude de cas montre que la calibration des paramètres en utilisant des données réelles améliore la précision des simulations. De plus, le modèle calibré atteint des performances sur les données historiques proches de celles obtenues par des méthodes purement basées sur les données, tout en répondant à notre exigence concernant la possibilité de modéliser explicitement les changements dans le mix électrique ou la conception du marché. Dans le chapitre suivant, certaines méthodes avancées d'estimation des paramètres seront introduites et certaines limitations du modèle en ce qui concerne son utilisation pour des études prospectives seront atténuées.

Appendix E

Résumé du chapitre 5 : Extensions et applications du modèle

Dans le chapitre précédent, nous avons proposé une méthode pour simuler les prix de l'électricité sur les marchés journaliers. Nous avons observé que la méthode pouvait être améliorée de deux façons notables. Premièrement, l'algorithme d'estimation peut être amélioré afin de simuler les prix avec plus de précision. Deuxièmement, la formulation du problème d'optimisation du marché lui-même peut être étendue afin de proposer une méthodologie globale qui atténue certaines limitations restantes en ce qui concerne l'utilisation du modèle pour des études prospectives. Dans ce chapitre, nous allons explorer et discuter ces deux extensions du modèle.

Cette validation du modèle étendu sur des données historiques nous permet enfin de proposer quelques applications. Plus précisément, le modèle est utilisé pour mener des études prospectives basées sur des scénarios concernant les évolutions possibles des prix de l'électricité ; ces études prospectives font l'objet de la dernière section de ce chapitre.

E.1 Etudes complémentaires de la méthode d'estimation

E.1.1 Introduire la variabilité dans la fonction de coût

On sait que l'erreur quadratique moyenne donne des résultats plus lisses que les observations. Nous proposons donc une modification de la fonction de coût qui se rapproche implicitement d'une minimisation de $RMSE_z + \gamma \Delta sd_z$. Nous proposons d'introduire des poids $(\omega_{t,z})_{t \in T}$ dans la fonction de coût comme hyperparamètre de variabilité du modèle :

$$RMSE_z = \left(\frac{1}{T} \sum_{c_0 \in C} \sum_{t \in T_{c_0}} \omega_{t,z} \left(\alpha_{c_0}^\top \mathbf{X}_{t,i^*(t)} + \lambda_{t,i^*(t)}^{IC} + \beta_z^\top \mathbf{Y}_t - \Pi_{t,z} \right)^2 \right)^{\frac{1}{2}} \quad (\text{E.1})$$

de telle sorte que le barycentre des prix ne soit pas modifié, c'est-à-dire de telle sorte que :

$$\frac{1}{|T_{c_0}|} \sum_{t \in T_{c_0}} \Pi_{t,z} = \frac{1}{\sum_{t \in T_{c_0}} \omega_{t,z}} \sum_{t \in T_{c_0}} \omega_{t,z} \Pi_{t,z} \quad (\text{E.2})$$

En donnant plus de poids aux heures où les prix sont extrêmement bas ou élevés, nous diminuons ainsi la valeur de Δsd_z .

Le réglage de l'hyperparamètre de variabilité du modèle pendant le calibrage peut affecter les performances hors échantillon mesurées par les métriques $RMSE$ et Δsd . Cet hyperparamètre est égal à la somme des poids ω_i introduits dans l'équation (E.1), qui sont ajoutés aux valeurs de prix minimales et maximales utilisées pour estimer les paramètres α_{c_0} pour chaque classe de production. Nous veillons à ce que les poids soient ajoutés de manière équilibrée afin que le prix moyen de l'ensemble reste inchangé.

Nous avons observé que l'hyperparamétrage peut contribuer à améliorer la dynamique de la simulation (c'est-à-dire que la variation des prix simulés est plus similaire à la variation des prix observés) au prix d'une précision globale plus faible mesurée par la RMSE.

E.1.2 Tirer parti des informations des courbes d'offre observées

L'algorithme d'estimation utilise les prix du marché observés, ce qui implique que seule l'information sur le prix donnée par l'ordre fixant le prix est exploitée pour chaque pas de temps. Cependant, les courbes d'offre apportent des informations précieuses qui n'ont pas été prises en compte dans notre méthode jusqu'à présent. Comme les enchères sur les marchés journaliers sont généralement aveugles (c'est-à-dire que l'identité des participants au marché n'est pas rendue publique, et qu'il n'est même pas possible de déterminer laquelle des unités de production participe réellement à un marché), l'information ne peut pas simplement être récupérée en faisant correspondre les ordres aux unités de production. Nous proposons une méthodologie pour exploiter l'information contenue dans les courbes d'offre observées $\mathcal{S}_{t,z}$, qui se traduit par une extension de l'ensemble des prix observés, appelés prix synthétiques. Une petite déviation, δ , est appliquée au volume réel de compensation du marché $V_{t,z}$ afin de générer des prix synthétiques $\Pi_{t,z}^\delta$ de sorte que :

$$\Pi_{t,z}^\delta = \mathcal{S}_{t,z}(V_{t,z}^\delta) \quad (\text{E.3})$$

Le processus est illustré sur la figure E.1.

Chaque variation relative δ crée un nouvel ensemble de données synthétiques utilisé pour estimer les paramètres du modèle (c'est-à-dire que les ensembles de données ne sont utilisés que pour l'entraînement en échantillon). Comme nous voulons que nos prix synthétiques reflètent des cas contrefactuels réalistes, nous considérons que la variation relative des volumes horaires est causée par la même variation de la demande résiduelle. Par conséquent, la marge d'offre $\mathcal{M}_{t,z}$ est également modifiée dans l'ensemble de données synthétiques puisqu'elle dépend de la demande résiduelle. Enfin, l'ensemble de données synthétiques est composé des nouveaux prix $(\Pi_{t,z}^\delta)_{t \in T}$, de la nouvelle demande résiduelle

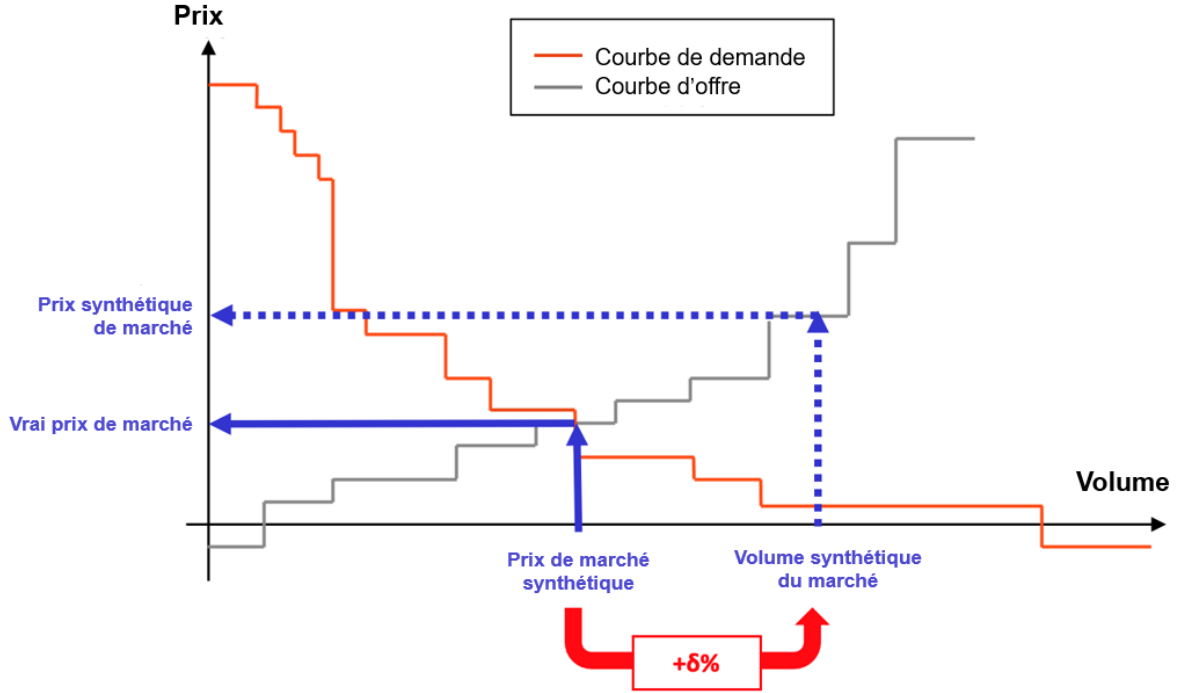


Figure E.1: Principe d'augmentation des données pour générer des prix synthétiques à partir des courbes d'offre

et des déterminants des prix $(\mathbf{X}_{t,i}^\delta)_{t \in T}$, où toutes les variables de $\mathbf{X}_{t,i}$ restent inchangées à l'exception de $\mathcal{M}_{t,z}^\delta$.

Contrairement aux résultats obtenus avec différentes valeurs d'hyperparamètres de variabilité, l'approche d'augmentation des données tend à augmenter la précision globale au détriment de variations de prix moins réalistes.

E.1.3 Analyse de sensibilité des conditions initiales

En initialisant les paramètres $\alpha_{c,0}$ à une valeur différente pour chaque classe de production, nous faisons des hypothèses concernant l'ordre de mérite (c'est-à-dire que l'énergie est fournie en priorité par la classe de production dont le coût est le plus bas, puis par la deuxième plus basse, et ainsi de suite). Ce choix initial a un impact sur la procédure d'estimation des paramètres. La raison est que l'identification de l'unité marginale pour chaque heure, qui résulte de la résolution du problème de compensation du marché, dépend directement de l'ordre de mérite estimé. Ainsi, nous effectuons une analyse de sensibilité pour étudier comment des hypothèses alternatives concernant l'ordre de mérite initial affectent le résultat final de la calibration. Pour ce faire, nous attribuons initialement cinq prix constants distincts à nos cinq classes de production (c'est-à-dire nucléaire, hydroélectrique, gaz, charbon et pétrole) et considérons toutes les permutations possibles, puis nous analysons l'ordre de mérite initial pour le meilleur décile en

termes de $RMSE$.

Les résultats montrent que l'initialisation du modèle doit être conforme aux hypothèses économiques courantes (par exemple, le nucléaire pour constituer une base peu coûteuse et le pétrole pour les pointes plus coûteuses). La place de l'hydroélectricité dans l'ordre de mérite est probablement due à l'utilisation stratégique du stock pendant les heures à prix élevé. Par conséquent, l'initialisation de l'ordre de mérite pendant la phase de formation doit être considérée avec soin par le modélisateur.

E.2 Validation du modèle pour les études prospectives

Dans cette section, notre modèle sera complexifié de manière incrémentale afin de répondre aux exigences supplémentaires des études prospectives. Pour réaliser les autres études présentées dans ce chapitre, nous avons choisi d'utiliser Pyomo, un langage de modélisation par optimisation basé sur Python. A présent, nous ne considérons pas les unités de production individuellement mais plutôt les disponibilités agrégées par types de production.

E.2.1 Contraintes de rampe

Les acteurs du marché modulent la production en fonction des besoins et choisissent d'activer les flexibilités disponibles au sein de leur portefeuille. Bien que les décisions concernent des centrales électriques individuelles, des effets globaux peuvent être observés en considérant la production agrégée par type de combustible car ils reflètent les contraintes économiques et techniques. En considérant les données historiques, nous pouvons alors estimer la capacité maximale de montée en puissance et de descente en puissance. Cette estimation a été réalisée en utilisant les données relatives au système électrique français pour la période 2015-2018. Nous avons choisi de conserver une seule limite de rampe représentative par technologie sur la base de cette estimation numérique. Afin que le modèle soit pertinent pour des études prospectives dans lesquelles la puissance installée ne reste pas constante, nous représentons ces contraintes comme une variation horaire de la puissance par rapport à la puissance installée.

Les contraintes supplémentaires de descente et de montée en régime sont respectivement ajoutées au problème d'optimisation comme suit :

$$-r_{c(i)}^{down} \cdot P_i^{max} \leq P_{t,i} - P_{t,i-1} \quad \forall (t, i) \in T \times I \quad (E.4)$$

$$P_{t,i} - P_{t,i-1} \leq r_{c(i)}^{up} \cdot P_i^{max} \quad \forall (t, i) \in T \times I \quad (E.5)$$

E.2.2 Gestion active du stockage

Dans l'étude de cas présentée au chapitre précédent, on a supposé que la contribution des centrales hydroélectriques à accumulation par pompage était connue. Les valeurs

réelles de production et de consommation agrégées ont été soustraites de la demande afin de calculer une demande résiduelle. Comme l'utilisation de ces centrales dépend des prix du marché, la gestion du stockage devrait plutôt être une variable interne du modèle pour les études prospectives. Pour ce faire, la contribution du stockage hydraulique par pompage n'est plus incluse dans la demande résiduelle et des contraintes supplémentaires sont ajoutées au modèle d'optimisation comme suit :

$$S_{t,k}^{niveau} = S_{t-1,k}^{niveau} \cdot (1 - d_k) + S_{t,k}^{in} \cdot \eta_k^{in} - S_{t,k}^{out} \cdot \eta_k^{out} \quad \forall (t, k) \in T \times K \quad (E.6)$$

$$S_{t,k}^{level} \leq C_k^{max} \quad \forall (t, k) \in T \times K \quad (E.7)$$

$$S_{t,k}^{in} \leq P_k^{max} \quad \forall (t, k) \in T \times K \quad (E.8)$$

$$S_{t,k}^{out} \leq P_k^{max} \quad \forall (t, k) \in T \times K \quad (E.9)$$

La première contrainte correspond à la définition récursive du niveau de stock $S_{t,k}^{niveau}$ pour la technologie k au temps t . Le niveau de stock dépend de sa valeur au temps $t - 1$ (incluant la dissipation spontanée à chaque pas de temps d'un facteur $1 - d_k$), et des flux d'entrée ou de sortie ($S_{t,k}^{in}$ et $S_{t,k}^{out}$) corrigés par l'efficacité de conversion (η_k^{in} et η_k^{out}). Les deuxième et troisième contraintes reflètent les limites de puissance de la technologie. La quatrième contrainte reflète sa limite de capacité en terme d'énergie.

E.2.3 Simulation multizone

Le dernier test de validation consiste à s'assurer que le modèle est également apte à simuler les prix même lorsqu'une étendue spatiale plus vaste est considérée. Dans l'étude de cas, les prix du marché dans les zones de soumission voisines étaient supposés être connus et seuls les flux dans les interconnexions étaient simulés. Ici, nous voulons plutôt calculer simultanément les prix dans plusieurs zones. Nous conservons la même répartition entraînement-test (c'est-à-dire 2015 pour l'entraînement et 2016 à 2018 pour le test), mais nous considérons trois pays : France, Belgique et Allemagne. Les séries temporelles de prix pour ces trois pays sont présentées dans la figure E.2. Nous procédons toujours par ajout progressif de complexité au modèle, ce qui signifie que les contraintes de rampe sont prises en compte et que le stockage est toujours géré activement.

Nous observons des différences significatives en termes de résultats de simulation. Comme nous ne nous appuyons plus sur les prix historiques des pays voisins, qui sont corrélés à ceux de la France, les performances sont globalement dégradées ($RMSE = 17,0$) et nous observons que l'amplitude des pics de prix est encore moins précise qu'auparavant. Néanmoins, dans ce cas, le modèle tend toujours à rendre compte de la dynamique des prix et la métrique s'améliore même ($\Delta sd = 0.7$). L'étude des cas belge ($RMSE = 23,0$, $\Delta sd = -0,0$) et allemand ($RMSE = 20,0$, $\Delta sd = -2,2$) révèle que le modèle converge mais qu'un travail supplémentaire serait nécessaire pour améliorer les performances pour les valeurs extrêmes (à la fois pour les pics de prix et les prix négatifs).

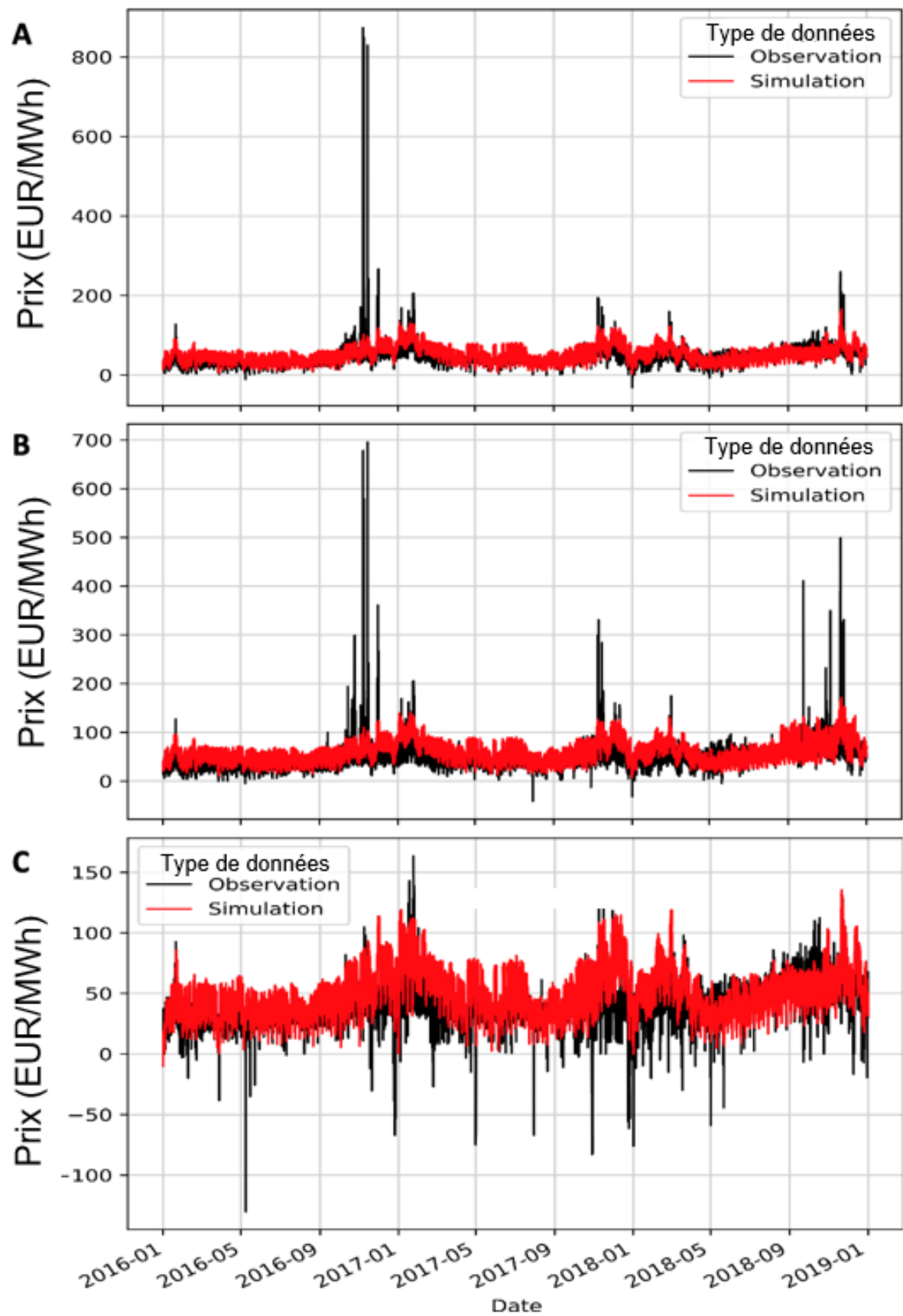


Figure E.2: Cas multizone : (A) prix français, (B) prix belges, (C) prix allemands

E.3 Applications du modèle

Le modèle proposé peut enfin être appliqué pour mener des études prospectives. Dans cette section, nous présentons d’abord l’approche par scénario, puis nous effectuons les simulations et discutons des implications des résultats des prix.

E.3.1 Scénarios pour les études prospectives

E.3.2 Construction d’un scénario de référence

Les études prospectives présentées dans la section suivante sont construites selon des principes communs à tous les scénarios. Nous avons choisi un horizon temporel d’un an pour tenir compte des effets saisonniers. Pour chaque scénario, plusieurs séries temporelles annuelles sont simulées en utilisant l’approche de Monte Carlo. Deux interprétations peuvent être données à ces séries temporelles : soit comme des simulations pluri-annuelles en considérant une simple concaténation des résultats du modèle, soit comme des prévisions d’ensemble qui reflètent l’incertitude issue des variables stochastiques.

Données statiques Les données statiques, c’est-à-dire les variables dont la valeur reste constante sur la période d’un an, sont dérivées des ensembles de données historiques (par exemple, la capacité installée). Pour les données statiques du modèle, nous prenons les valeurs de l’année 2018 comme référence.

Données dépendantes du temps Pour construire des séries temporelles horaires que nous pouvons utiliser comme données d’entrée pour notre modèle, nous considérons deux méthodes complémentaires : (i) la génération de profils annuels et (ii) la génération d’entrées stochastiques grâce à des techniques de rééchantillonnage. La consommation nationale agrégée et les prix des produits de base sont générés en utilisant l’approche par profil ; c’est-à-dire que les valeurs sont dérivées de l’ensemble des données historiques en calculant la moyenne heure par heure. Pour le facteur de disponibilité des unités de production et pour le facteur de charge des énergies renouvelables liées aux conditions météorologiques, nous avons choisi de nous appuyer sur l’approche de rééchantillonnage. Avec cette méthode, nous définissons un cas de Monte Carlo en tirant au hasard des données historiques. Comme pour l’approche par profil, nous tirons uniquement des valeurs de la même heure de l’année des données historiques. Cette opération d’échantillonnage est répétée plusieurs fois afin de définir différents cas dans un scénario.

E.3.3 Etudes prospectives

Pour le scénario de référence, nous considérons une augmentation de la demande électrique et des capacités renouvelables installées pour l’éolien et le solaire, comme c’est le cas pour la plupart des scénarios net zéro. Par rapport à 2018, le scénario représente une augmentation de +20% de la demande et une augmentation de +50% pour les capacités installées éoliennes et solaires. Nous simulons ce scénario avec 30 cas de Monte

Carlo. Comme il n'est pas pratique d'afficher tous les résultats de la simulation sur une période d'un an, nous proposons de visualiser les résultats pour la première semaine de juin, comme le montre la figure E.3. Nous observons que nous avons toujours toutes les variabilités saisonnières (horaire, hebdomadaire) et que l'approche de Monte Carlo nous permet de représenter l'incertitude puisque les différentes valeurs des prix sont obtenues pour chaque heure. La variabilité annuelle a également été observée dans nos résultats de simulation.

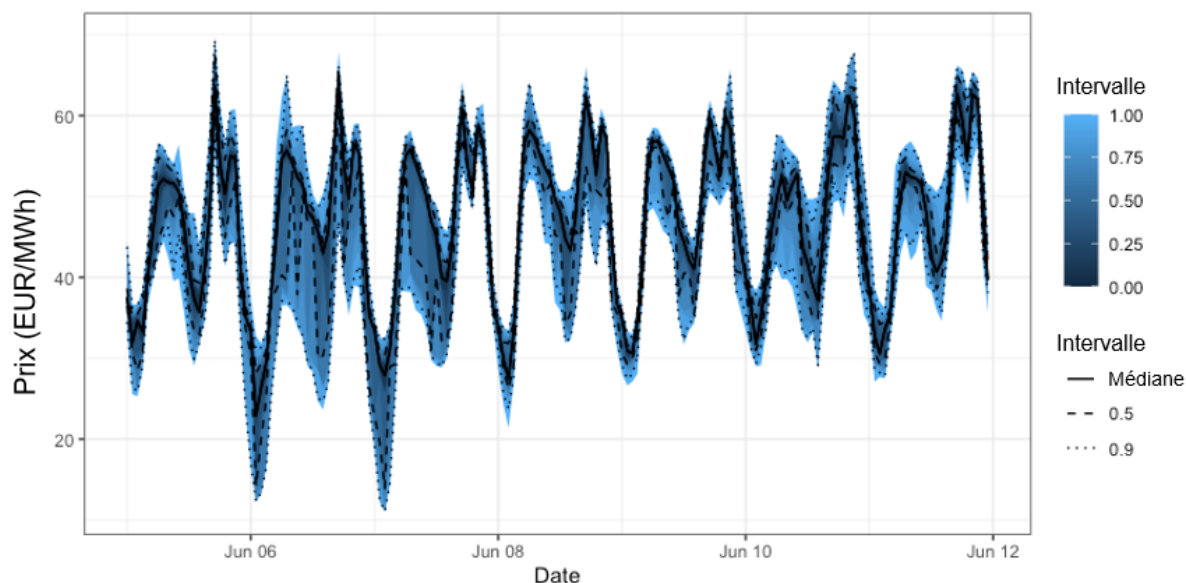


Figure E.3: Simulations Monte Carlo des prix sur le marché journalier français (scénario prospectif de référence, première semaine de juin)

Nous comparons également la distribution des prix simulés à notre année de référence, 2018. La comparaison est représentée dans la figure E.4 dans laquelle les prix horaires moyens pour toutes les simulations sont calculés pour la distribution. Nous observons que notre scénario conduit à une augmentation globale des prix. En l'absence de stockage supplémentaire, l'augmentation de 50% de la capacité installée de l'éolien et du solaire n'exerce pas de pression à la baisse sur les prix en raison de l'augmentation de 20% de la demande. Pour l'année de référence 2018, la part des sources d'énergie renouvelables variables (ERV) dans la consommation totale d'électricité française était de 16,8% et cette valeur est restée similaire pour les études prospectives avec une part moyenne de ERV égale à 15,7%, ce qui suggère que des efforts d'investissements supplémentaires dans les ERV pourraient être nécessaires pour contrebalancer l'augmentation prévue de la demande d'électricité. En 2018, le prix médian était de 49,9 EUR/MWh, le prix moyen de 50,2 EUR/MWh et un écart type de 18,5 EUR/MWh. Pour notre simulation prospective dans ce scénario de référence, nous obtenons : un prix médian égal à 66,5 EUR/MWh, un prix moyen égal à 69,8 EUR/MWh et un écart-type égal à 23,7 EUR/MWh. Nous

constatons que l'augmentation de la variabilité due à des énergies renouvelables plus dépendantes des conditions météorologiques conduit à des prix plus dispersés.

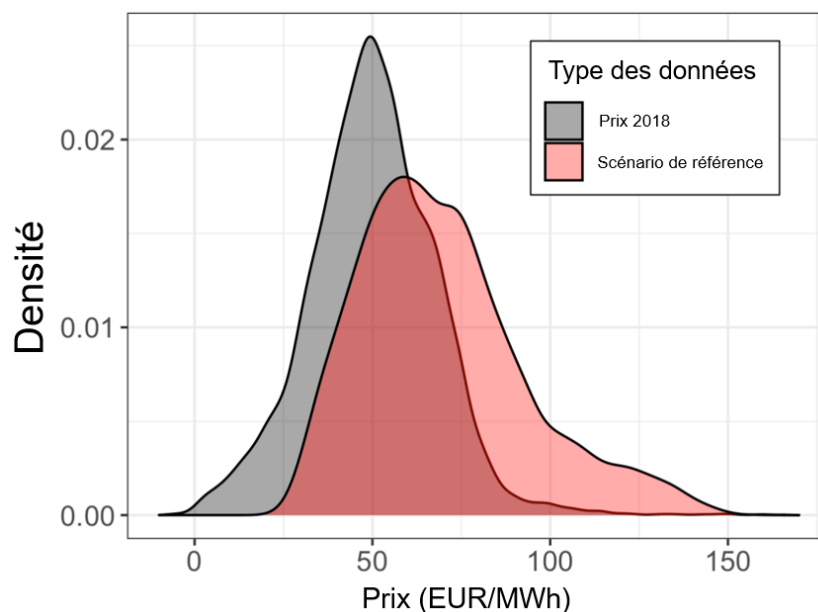


Figure E.4: Distribution des prix de l'électricité sur le marché journalier français

E.4 Conclusions

Dans ce chapitre, nous avons proposé des extensions et des étapes de validation finale qui ont préparé notre modèle pour être appliqué à des études prospectives.

Nous avons analysé comment la méthode d'estimation des paramètres du modèle peut être améliorée pour être finement ajustée grâce à l'introduction d'un nouvel hyperparamètre et comment nous pouvons prendre en compte des données de marché supplémentaires.

Pour que le modèle reflète mieux le fonctionnement du système électrique et du marché européen, nous nous sommes assurés de prendre en compte le potentiel de flexibilité des technologies clés, nous avons fait de la gestion des actifs de stockage une partie interne du problème d'optimisation, et nous avons vérifié que le modèle est toujours opérationnel pour les simulations multizone.

Pour les études prospectives, nous avons proposé une méthode pour construire des scénarios qui s'appuient sur des données historiques afin de conserver des contraintes d'exploitation réalistes tout en tenant compte de l'incertitude grâce aux simulations de Monte Carlo.

Appendix F

Résumé du chapitre 6 : Conclusion générale

F.1 Conclusion générale

Des modèles structurels ont été proposés dans la littérature scientifique et ils sont couramment utilisés dans l'industrie, notamment pour les évaluations d'adéquation. Ils sont basés sur l'hypothèse d'une utilisation optimale des actifs du système électrique d'un point de vue technico-économique. Cette hypothèse est légitime mais n'englobe pas une multitude d'autres effets qui influencent les prix, notamment les stratégies d'enchères des participants au marché en cas de concurrence imparfaite. Comme il n'est pas possible de rendre compte explicitement de tous les facteurs influençant les prix, il faut envisager des stratégies de modélisation alternatives. Nous avons choisi d'utiliser les données historiques du système électrique et du marché pour réduire l'écart entre la théorie économique et la réalité de la formation des prix. Quelques approches avec un raisonnement similaire ont été proposées par des chercheurs ; en nous inspirant de leurs approches, nous avons proposé une nouvelle méthode qui s'appuie sur leurs efforts. Ici, nous avons proposé d'intégrer l'influence des facteurs de prix dans le problème d'optimisation d'une manière modulaire et paramétrée. L'influence des facteurs sélectionnés est estimée quantitativement grâce à un algorithme dédié. La méthodologie globale nous permet de garder le contrôle sur les hypothèses technico-économiques que nous voulons inclure dans le modèle sans négliger les enseignements qui peuvent être tirés de l'utilisation d'ensembles de données historiques. L'un des principaux avantages est que ces informations sont ensuite intégrées dans les paramètres estimés, de sorte que le modèle ajusté peut être utilisé sur d'autres ensembles de données, par exemple pour des analyses contrefactuelles ou des études prospectives.

Les études de cas montrent que la dynamique des prix de l'électricité est globalement bien représentée par le modèle, en particulier les variations saisonnières aux échelles quotidiennes et hebdomadaires. Les résultats sont analysés sur des données hors échantillon qui diffèrent de l'ensemble d'entraînement, ce qui évite que l'analyse soit faussée par une forte interprétation des phénomènes de surapprentissage. Nous avons observé que le

modèle pouvait encore être amélioré pour mieux simuler les pics de prix ; les améliorations potentielles sont discutées dans la section suivante. Nous avons également souligné l'influence des événements stochastiques sur la qualité de nos simulations et nous avons donc veillé à prendre en compte l'incertitude dans les études prospectives ultérieures. Par ailleurs, l'estimation des paramètres du modèle grâce à l'algorithme proposé améliore la qualité des simulations par rapport à une simple initialisation basée sur des hypothèses technico-économiques. Nous avons montré comment la méthode d'estimation peut être améliorée en ajustant finement la pondération des points de données ou en récupérant des informations à partir de courbes d'approvisionnement réelles afin d'améliorer la précision globale et la dynamique des simulations. Néanmoins, la méthode d'estimation pourrait être encore développée pour être plus précise pour les classes de production qui ne fixent pas souvent les prix du marché pendant la période d'apprentissage. Par le biais d'améliorations successives, nous avons également examiné comment la modularité de la formulation du modèle nous a permis de proposer des extensions pour considérer des facteurs techniques supplémentaires (par exemple, les contraintes de rampe et la gestion stratégique du stockage). Ce faisant, nous tenons compte de l'interdépendance des prix sur plusieurs heures, ce qui n'est pas le cas des modèles structurels dans lesquels chaque paire horaire de courbes d'offre et de demande est considérée séparément. De même, nous avons également démontré que la méthode peut être appliquée à des cas multizone, qui correspondent à la configuration réelle du marché intérieur interconnecté en Europe.

Notre modèle permet non seulement de mieux comprendre la formation des prix de l'électricité par la simulation de situations historiques, mais il permet également d'étudier une gamme de prix futurs possibles de l'électricité grâce à des scénarios dédiés. Ces scénarios permettent de transcrire les décisions politiques et les évolutions technologiques potentielles dans des jeux de données personnalisés. Les scénarios, qui reflètent les hypothèses ou les souhaits du modélisateur, peuvent également comporter une part d'incertitude intrinsèque correspondant à des facteurs stochastiques. Par exemple, le facteur de charge du solaire et de l'éolien ou l'indisponibilité des centrales électriques conventionnelles peuvent être pris en compte par des simulations de Monte Carlo. L'élaboration de scénarios et la simulation des prix à l'aide de notre modèle nous permettent d'estimer et de comparer l'impact des voies d'investissement dans le cadre d'études prospectives.

Des études de cas montrent que la dynamique des prix de l'électricité est globalement bien représentée par le modèle, en particulier les variations saisonnières aux échelles quotidiennes et hebdomadaires. Les résultats sont analysés sur des données hors échantillon qui diffèrent de l'ensemble d'entraînement, ce qui évite que l'analyse soit faussée par une forte interprétation des phénomènes de surapprentissage. Nous avons observé que le modèle pouvait encore être amélioré pour mieux simuler les pics de prix ; les améliorations potentielles sont discutées dans la section suivante. Nous avons également souligné l'influence des événements stochastiques sur la qualité de nos simulations et nous avons donc veillé à prendre en compte l'incertitude dans les études prospectives ultérieures. Par ailleurs, l'estimation des paramètres du modèle grâce à l'algorithme proposé améliore la qualité des simulations par rapport à une simple initialisation basée sur des hypothèses

technico-économiques. Nous avons montré comment la méthode d'estimation peut être améliorée en ajustant finement la pondération des points de données ou en récupérant des informations à partir de courbes d'approvisionnement réelles afin d'améliorer la précision globale et la dynamique des simulations. Néanmoins, la méthode d'estimation pourrait être encore développée pour être plus précise pour les classes de production qui ne fixent pas souvent les prix du marché pendant la période d'apprentissage. Par le biais d'améliorations successives, nous avons également examiné comment la modularité de la formulation du modèle nous a permis de proposer des extensions pour considérer des facteurs techniques supplémentaires (par exemple, les contraintes de rampe et la gestion stratégique du stockage). Ce faisant, nous tenons compte de l'interdépendance des prix sur plusieurs heures, ce qui n'est pas le cas des modèles structurels dans lesquels chaque paire horaire de courbes d'offre et de demande est considérée séparément. De même, nous avons également démontré que la méthode peut être appliquée à des cas multizone, qui correspondent à la configuration réelle du marché intérieur interconnecté en Europe.

Notre modèle permet non seulement de mieux comprendre la formation des prix de l'électricité par la simulation de situations historiques, mais il permet également d'étudier un ensemble de prix futurs possibles de l'électricité grâce à des scénarios dédiés. Ces scénarios permettent de transcrire les décisions politiques et les évolutions technologiques potentielles dans des jeux de données personnalisés. Les scénarios, qui reflètent les hypothèses ou les souhaits du modélisateur, peuvent également comporter une part d'incertitude intrinsèque correspondant à des facteurs stochastiques. Par exemple, le facteur de charge du solaire et de l'éolien ou l'indisponibilité des centrales électriques conventionnelles peuvent être pris en compte par des simulations de Monte Carlo. L'élaboration de scénarios et la simulation des prix à l'aide de notre modèle nous permettent d'estimer et de comparer l'impact des voies d'investissement dans le cadre d'études prospectives.

F.2 Perspectives et travaux futurs

D'autres études sur le marché de l'électricité et son évolution peuvent être réalisées. Nous suggérons ici quelques directions pour développer ce travail.

Le modèle lui-même peut être étendu pour tenir compte de contraintes techniques supplémentaires, telles que les contraintes de démarrage et d'arrêt des centrales électriques. La réalisation d'études plus précises au niveau de l'unité de la centrale électrique serait l'occasion d'évaluer dans quelle mesure des informations techniques exclusives sur les unités individuelles peuvent être déduites des données publiques. De même, la correspondance entre les offres individuelles et les unités pourrait être étudiée de manière plus approfondie au lieu d'adopter une approche plus agrégée à travers la courbe d'offre. Néanmoins, il n'est pas toujours possible de réduire précisément à cette échelle, et la représentation des effets agrégés doit rester le point central de la simulation des prix du marché. En ce qui concerne les choix de modélisation du côté de l'offre, la formulation de la paramétrisation pourrait être remplacée par des termes non linéaires afin de mieux saisir les pics de prix, ou un changement de régime pourrait être proposé pour

changer la valeur des paramètres pendant les situations de pénurie. Une prochaine étape pourrait être de conserver la formulation linéaire mais d'explorer le potentiel d'effectuer des régressions locales pendant la phase de formation au lieu d'une seule régression par classe de production. En outre, une mise en œuvre plus complexe et plus précise de la modélisation de la demande pourrait être examinée. Alors que nous avons principalement considéré une demande résiduelle inélastique au prix, d'autres cas pourraient être étudiés dans lesquels la valeur de la réponse à la demande serait explicitement testée. De même, les énergies renouvelables ont été considérées comme des preneurs de prix dans nos études de cas, mais avec un taux de pénétration plus élevé des actifs de production liés aux conditions météorologiques, cette hypothèse pourrait être réévaluée. Pour ce faire, des modèles d'enchères spécifiques pourraient être développés afin d'estimer les stratégies pertinentes pour les producteurs d'énergies renouvelables ayant des ordres de marché supérieurs à leurs coûts de production marginaux.

La méthode pourrait également être appliquée à des études portant sur des périodes spatiales et temporelles plus importantes. Les initiatives de données ouvertes offrent des ensembles de données qui couvrent une période de plus en plus longue et la qualité des ensembles de données ne cesse de s'améliorer, ce qui s'avérerait intéressant pour une estimation plus précise des paramètres du modèle. L'utilisation de ces données et l'exécution du modèle à l'échelle de l'Europe seraient également intéressantes pour étudier le phénomène de convergence des prix. En outre, l'obtention de données propriétaires, telles que les prix de l'électricité sur les marchés à terme, pourrait également améliorer la pertinence des simulations. Pour la construction de scénarios à l'échelle européenne, des analyses détaillées des décisions politiques et des tendances industrielles apporteraient une valeur ajoutée supplémentaire. Des analyses spécifiques des études prospectives pourraient également être réalisées afin d'évaluer les revenus attendus des nouveaux investissements et de définir les politiques publiques, notamment fiscales, concernant la nécessité de nouvelles subventions, taxes ou contrats de différence (CfD).

Une dernière piste pour prolonger ces études serait d'examiner de plus près l'évolution du fonctionnement du marché journalier et son interaction avec les autres marchés. L'évolution du marché journalier et du marché infrajournalier pourrait être analysée conjointement afin d'évaluer l'impact des volumes d'échanges respectifs sur les prix. De même, des discussions plus spéculatives pourraient être menées afin d'estimer dans quelle mesure l'essor des robots de trading et le délai de 15 minutes pourraient influencer la formation des prix, et ensuite comment de telles idées pourraient être intégrées dans notre modèle.

Bibliography

- [1] Sinan Küfeoğlu. “Economic Impacts of Electric Power Outages and Evaluation of Customer Interruption Costs”. en. ISSN: 1799-4942 (electronic). PhD thesis. Aalto University, 2015. URL: <https://aaltodoc.aalto.fi:443/handle/123456789/17867> (visited on 03/19/2020).
- [2] John Kemp. “Column: Worldwide energy shortage shows up in surging coal, gas and oil prices: Kemp”. In: *Reuters* (Sept. 27, 2021). URL: <https://www.reuters.com/business/energy/worldwide-energy-shortage-shows-up-surging-coal-gas-oil-prices-kemp-2021-09-24/> (visited on 09/28/2021).
- [3] “Why a predictable cold snap crippled the Texas power grid”. In: *Reuters* (Feb. 21, 2021). URL: <https://www.reuters.com/article/us-usa-weather-texas-power-insight-idUSKBN2AL00N> (visited on 09/28/2021).
- [4] Philippe Vassilopoulos. “Les prix des marchés de gros de l’électricité donnent-ils les bons signaux et les bonnes incitations pour l’investissement en capacité de production électrique ?” fr. PhD thesis. PhD Thesis: Université Paris Dauphine, July 2007. URL: https://www.cgemp.dauphine.fr/fileadmin/mediatheque/centres/cgemp/Theses_soutenues/Vassilopoulos%5C%20Resume.pdf.
- [5] René Carmona and Michael Coulon. “A Survey of Commodity Markets and Structural Models for Electricity Prices”. en. In: *Quantitative Energy Finance*. Ed. by Fred Espen Benth, Valery A. Kholodnyi, and Peter Laurence. New York, NY: Springer New York, 2014, pp. 41–83. ISBN: 978-1-4614-7247-6 978-1-4614-7248-3. DOI: 10.1007/978-1-4614-7248-3_2. URL: http://link.springer.com/10.1007/978-1-4614-7248-3_2 (visited on 12/14/2018).
- [6] Rafał Weron. “Electricity price forecasting: A review of the state-of-the-art with a look into the future”. In: *International Journal of Forecasting* 30.4 (Oct. 2014), pp. 1030–1081. ISSN: 0169-2070. DOI: 10.1016/j.ijforecast.2014.08.008.
- [7] Florian Ziel and Rick Steinert. “Electricity price forecasting using sale and purchase curves: The X-Model”. en. In: *Energy Economics* 59 (Sept. 2016), pp. 435–454. ISSN: 0140-9883. DOI: 10.1016/j.eneco.2016.08.008. URL: <http://www.sciencedirect.com/science/article/pii/S0140988316302080> (visited on 03/19/2020).

- [8] C. Alasseur and O. Féron. “Structural price model for coupled electricity markets”. In: *Energy Economics* 75 (Sept. 2018), pp. 104–119. ISSN: 0140-9883. DOI: 10.1016/j.eneco.2018.07.018. URL: <http://www.sciencedirect.com/science/article/pii/S0140988318302676> (visited on 12/12/2018).
- [9] K. R. Ward, R. Green, and I. Staffell. “Getting prices right in structural electricity market models”. en. In: *Energy Policy* 129 (June 2019), pp. 1190–1206. ISSN: 0301-4215. DOI: 10.1016/j.enpol.2019.01.077. URL: <http://www.sciencedirect.com/science/article/pii/S0301421519300564> (visited on 03/10/2020).
- [10] Christian Pape, Simon Hagemann, and Christoph Weber. “Are fundamentals enough? Explaining price variations in the German day-ahead and intraday power market”. en. In: *Energy Economics* 54 (Feb. 2016), pp. 376–387. ISSN: 0140-9883. DOI: 10.1016/j.eneco.2015.12.013. URL: <http://www.sciencedirect.com/science/article/pii/S0140988316000037> (visited on 10/14/2020).
- [11] Philip Beran, Christian Pape, and Christoph Weber. “Modelling German electricity wholesale spot prices with a parsimonious fundamental model – Validation & application”. en. In: *Utilities Policy* 58 (June 2019), pp. 27–39. ISSN: 0957-1787. DOI: 10.1016/j.jup.2019.01.008. URL: <http://www.sciencedirect.com/science/article/pii/S0957178719300359> (visited on 10/14/2020).
- [12] United Nations Environment Programme. *The emissions gap report 2020*. OCLC: 1226523442. 2020. ISBN: 978-92-807-3812-4. URL: <https://www.unep.org/emissions-gap-report-2020>.
- [13] V. Masson-Delmotte et al. *IPCC, 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by N. Caud. URL: <https://www.ipcc.ch/report/ar6/wg1/#FullReport>.
- [14] Pierre Friedlingstein et al. “Global Carbon Budget 2020”. In: *Earth System Science Data* 12.4 (Dec. 11, 2020). Publisher: Copernicus GmbH, pp. 3269–3340. ISSN: 1866-3508. DOI: 10.5194/essd-12-3269-2020. URL: <https://essd.copernicus.org/articles/12/3269/2020/> (visited on 10/04/2021).
- [15] Dan Welsby et al. “Unextractable fossil fuels in a 1.5C world”. In: *Nature* 597.7875 (Sept. 2021), pp. 230–234. ISSN: 1476-4687. DOI: 10.1038/s41586-021-03821-8. URL: <https://www.nature.com/articles/s41586-021-03821-8> (visited on 09/09/2021).
- [16] Hannah Ritchie and Max Roser. “CO₂ and Greenhouse Gas Emissions”. In: *Our World in Data* (May 11, 2020). URL: <https://ourworldindata.org/emissions-by-sector> (visited on 10/04/2021).
- [17] *IEA Sankey Diagram*. URL: <https://www.iea.org/sankey/#?c=World&s=Balance> (visited on 01/11/2022).

-
- [18] Carina Zell-Ziegler et al. “Enough? The role of sufficiency in European energy and climate plans”. In: *Energy Policy* 157 (Oct. 1, 2021), p. 112483. ISSN: 0301-4215. DOI: 10.1016/j.enpol.2021.112483. URL: <https://www.sciencedirect.com/science/article/pii/S0301421521003530> (visited on 09/29/2021).
 - [19] Oliver Richters and Andreas Siemoneit. “Growth imperatives: Substantiating a contested concept”. In: *Structural Change and Economic Dynamics* 51 (Dec. 1, 2019), pp. 126–137. ISSN: 0954-349X. DOI: 10.1016/j.strueco.2019.07.012. URL: <https://www.sciencedirect.com/science/article/pii/S0954349X19301742> (visited on 10/04/2021).
 - [20] Michael Child et al. “Sustainability guardrails for energy scenarios of the global energy transition”. In: *Renewable and Sustainable Energy Reviews* 91 (Aug. 1, 2018), pp. 321–334. ISSN: 1364-0321. DOI: 10.1016/j.rser.2018.03.079. URL: <https://www.sciencedirect.com/science/article/pii/S136403211830176X> (visited on 10/04/2021).
 - [21] Paul E. Brockway et al. “Energy efficiency and economy-wide rebound effects: A review of the evidence and its implications”. In: *Renewable and Sustainable Energy Reviews* 141 (May 1, 2021), p. 110781. ISSN: 1364-0321. DOI: 10.1016/j.rser.2021.110781. URL: <https://www.sciencedirect.com/science/article/pii/S1364032121000769> (visited on 03/17/2021).
 - [22] IEA. *Energy Technology Perspectives 2020 - Special Report on Carbon Capture Utilisation and Storage: CCUS in clean energy transitions*. Paris: OECD Publishing, 2020, p. 174. URL: <https://doi.org/10.1787/208b66f4-en>.
 - [23] S. C. Page, A. G. Williamson, and I. G. Mason. “Carbon capture and storage: Fundamental thermodynamics and current technology”. In: *Energy Policy*. New Zealand Energy Strategy 37.9 (Sept. 1, 2009), pp. 3314–3324. ISSN: 0301-4215. DOI: 10.1016/j.enpol.2008.10.028. URL: <https://www.sciencedirect.com/science/article/pii/S0301421508005776> (visited on 10/07/2021).
 - [24] Dominik Wiedenhofer et al. “A systematic review of the evidence on decoupling of GDP, resource use and GHG emissions, part I: bibliometric and conceptual mapping”. In: *Environmental Research Letters* 15.6 (June 2020). Publisher: IOP Publishing, p. 063002. ISSN: 1748-9326. DOI: 10.1088/1748-9326/ab8429. URL: <https://doi.org/10.1088/1748-9326/ab8429> (visited on 10/04/2021).
 - [25] Helmut Haberl et al. “A systematic review of the evidence on decoupling of GDP, resource use and GHG emissions, part II: synthesizing the insights”. In: *Environmental Research Letters* 15.6 (June 2020). Publisher: IOP Publishing, p. 065003. ISSN: 1748-9326. DOI: 10.1088/1748-9326/ab842a. URL: <https://doi.org/10.1088/1748-9326/ab842a> (visited on 10/04/2021).
 - [26] Thomas Wiedmann et al. “Scientists’ warning on affluence”. In: *Nature Communications* 11.1 (Dec. 2020), p. 3107. ISSN: 2041-1723. DOI: 10.1038/s41467-020-16941-y. URL: <http://www.nature.com/articles/s41467-020-16941-y> (visited on 06/22/2020).

- [27] Carl Folke et al. “Our future in the Anthropocene biosphere”. In: *Ambio* 50.4 (Apr. 1, 2021), pp. 834–869. ISSN: 1654-7209. DOI: 10.1007/s13280-021-01544-8. URL: <https://doi.org/10.1007/s13280-021-01544-8> (visited on 09/29/2021).
- [28] *Greenhouse gas emission intensity of electricity generation in Europe* — European Environment Agency. URL: <https://www.eea.europa.eu/data-and-maps/indicators/overview-of-the-electricity-production-3/assessment> (visited on 05/21/2021).
- [29] European Commission. *2030 climate & energy framework*. Climate Action - European Commission. Nov. 23, 2016. URL: https://ec.europa.eu/clima/policies/strategies/2030_en (visited on 05/21/2021).
- [30] *National energy and climate plans*. European Commission - European Commission. URL: https://ec.europa.eu/info/energy-climate-change-environment/implementation-eu-countries/energy-and-climate-governance-and-reporting/national-energy-and-climate-plans_en (visited on 05/21/2021).
- [31] *National long-term strategies*. European Commission - European Commission. URL: https://ec.europa.eu/info/energy-climate-change-environment/implementation-eu-countries/energy-and-climate-governance-and-reporting/national-long-term-strategies_en (visited on 05/21/2021).
- [32] International Energy Agency. *European Union 2020 - Energy Policy Review*. 2020. URL: https://iea.blob.core.windows.net/assets/ec7cc7e5-f638-431b-ab6e-86f62aa5752b/European_Union_2020_Energy_Policy_Review.pdf (visited on 05/21/2021).
- [33] *Emissions cap and allowances*. Climate Action - European Commission. Nov. 23, 2016. URL: https://ec.europa.eu/clima/policies/ets/cap_en (visited on 05/21/2021).
- [34] *Data - Energy - Eurostat*. URL: <https://ec.europa.eu/eurostat/web/energy/data> (visited on 05/21/2021).
- [35] European Commission. Statistical Office of the European Union. *Energy balance sheets: 2020 edition*. LU: Publications Office, 2020. URL: <https://data.europa.eu/doi/10.2785/68334> (visited on 05/21/2021).
- [36] *Energy statistics - an overview*. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Energy_statistics_-_an_overview (visited on 05/21/2021).
- [37] *electricityMap | Live CO₂ emissions of electricity consumption*. URL: <https://www.electricitymap.org/map> (visited on 05/21/2021).
- [38] *Electricity and heat statistics*. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Electricity_and_heat_statistics (visited on 05/21/2021).

-
- [39] DG Energy. *Quarterly report on European Electricity Markets*. Market Observatory for Energy Volume 13. European Commission, 2020. URL: https://ec.europa.eu/energy/sites/default/files/quarterly_report_on_european_electricity_markets_q4_2020.pdf (visited on 05/21/2021).
 - [40] Douglas Lazard. *Lazard's Levelized Cost of Energy Analysis - Version 14.0*. 2020. URL: <https://www.lazard.com/media/451419/lazards-levelized-cost-of-energy-version-140.pdf>.
 - [41] Fabian Gotzens et al. "Performing energy modelling exercises in a transparent way - The issue of data quality in power plant databases". In: *Energy Strategy Reviews* 23 (Jan. 1, 2019), pp. 1–12. ISSN: 2211-467X. DOI: 10.1016/j.esr.2018.11.004. URL: <https://www.sciencedirect.com/science/article/pii/S2211467X18301056> (visited on 05/21/2021).
 - [42] World Resources Institute. *Global Power Plant Database*. URL: <https://resourcewatch.org/data/explore/Powerwatch> (visited on 05/21/2021).
 - [43] Nuclear Energy Agency. *Small Modular Reactors: Challenges and Opportunities*. 2021. URL: https://www.oecd-neo.org/upload/docs/application/pdf/2021-03/7560_smr_report.pdf.
 - [44] *Installed capacity in the European Union, 2000-2010, and projections up to 2040 in the Stated Policies Scenario – Charts – Data & Statistics*. IEA. URL: <https://www.iea.org/data-and-statistics/charts/installed-capacity-in-the-european-union-2000-2010-and-projections-up-to-2040-in-the-stated-policies-scenario> (visited on 05/21/2021).
 - [45] *Marine renewable energy*. URL: https://ec.europa.eu/oceans-and-fisheries/ocean/blue-economy/marine-renewable-energy_en (visited on 05/21/2021).
 - [46] European Commission. *An EU Strategy to harness the potential of offshore renewable energy for a climate neutral future*. Nov. 19, 2021. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0741&from=EN> (visited on 05/21/2021).
 - [47] Emiliano Bellini. "Submarine cable to connect 10.5 GW wind-solar complex in Morocco to the UK grid – pv magazine International". In: *PV Magazine* (Apr. 22, 2021). URL: <https://www.pv-magazine.com/2021/04/22/submarine-cable-to-connect-10-5-gw-wind-solar-complex-in-morocco-to-the-uk-grid/> (visited on 05/21/2021).
 - [48] Pilar Sánchez Molina. "Chile wants to export green hydrogen to Port of Rotterdam – pv magazine International". In: *PV Magazine* (Mar. 19, 2021). URL: <https://www.pv-magazine.com/2021/03/19/chile-wants-to-export-green-hydrogen-to-port-of-rotterdam/> (visited on 05/21/2021).
 - [49] *Geothermal energy*. European Commission - European Commission. URL: https://ec.europa.eu/info/research-and-innovation/research-area/energy-research-and-innovation/geothermal-energy_en (visited on 05/21/2021).

- [50] European Association for Storage of Energy (EASE). *Energy Storage Applications Summary*. June 2020. URL: <https://ease-storage.eu/wp-content/uploads/2020/06/ES-Applications-Summary.pdf> (visited on 05/25/2021).
- [51] *Technologies / EASE: Why Energy Storage? / EASE*. URL: <https://ease-storage.eu/energy-storage/technologies/> (visited on 05/21/2021).
- [52] Lion Hirth, Jonathan Mühlenpfordt, and Marisa Bulkeley. “The ENTSO-E Transparency Platform – A review of Europe’s most ambitious electricity data platform”. en. In: *Applied Energy* 225 (Sept. 2018), pp. 1054–1067. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2018.04.048. URL: <http://www.sciencedirect.com/science/article/pii/S0306261918306068> (visited on 07/06/2020).
- [53] Olfa Tlili et al. “Hydrogen market penetration feasibility assessment: Mobility and natural gas markets in the US, Europe, China and Japan”. In: *International Journal of Hydrogen Energy* 44.31 (June 21, 2019), pp. 16048–16068. ISSN: 0360-3199. DOI: 10.1016/j.ijhydene.2019.04.226. URL: <https://www.sciencedirect.com/science/article/pii/S0360319919316805> (visited on 01/07/2022).
- [54] European Commission. *Study on energy storage – Contribution to the security of the electricity supply in Europe*. Jan. 7, 2020. URL: https://ec.europa.eu/energy/topics/technology-and-innovation/energy-storage_en (visited on 05/21/2021).
- [55] ENTSO-E. *ENTSO-E at a glance*. 2015. URL: https://eepublicdownloads.entsoe.eu/clean-documents/Publications/ENTSO-E%20general%20publications/entsoe_at_a_glance_2015_web.pdf (visited on 05/21/2021).
- [56] European Commission. *Regulation (EC) No 714/2009 of the European Parliament and of the Council of 13 July 2009 on conditions for access to the network for cross-border exchanges in electricity and repealing Regulation (EC) No 1228/2003*. URL: <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:211:0015:0035:EN:PDF>.
- [57] *ENTSO-E Official Mandates*. URL: <https://www.entsoe.eu/about/inside-entsoe/official-mandates/> (visited on 05/21/2021).
- [58] *Grid Map*. URL: <https://www.entsoe.eu/data/map/> (visited on 05/21/2021).
- [59] Jonas Hörsch et al. *PyPSA-Eur: An Open Optimisation Model of the European Transmission System (Dataset) (Version v0.1.0) [Data set]*. 2020. URL: <https://zenodo.org/record/3601882#.YKd-UeuxXUL> (visited on 05/21/2021).
- [60] *OpenMod - Transmission network datasets*. URL: https://wiki.openmod-initiative.org/wiki/Transmission_network_datasets (visited on 05/21/2021).
- [61] *Regulation (EU) 2019/943 of the European Parliament and of the Council of 5 June 2019 on the internal market for electricity (Text with EEA relevance.)* Legislative Body: EP, CONSIL. June 14, 2019. URL: <http://data.europa.eu/eli/reg/2019/943/oj/eng> (visited on 05/21/2021).

-
- [62] Jean Proriol. *Rapport d'information sur la sécurité et le financement des réseaux de distribution d'électricité*. 3307. Assemblée nationale. URL: <https://www.assemblee-nationale.fr/13/rap-info/i3307.asp> (visited on 05/21/2021).
- [63] Enedis OpenStreetMap France : signature d'une convention de partenariat. OpenStreetMap France. Mar. 1, 2021. URL: <https://www.openstreetmap.fr/convention-dechange-de-donnees-entre-enedis-et-openstreetmap-france/> (visited on 05/21/2021).
- [64] European Parliament. *Sector coupling: how can it be enhanced in the EU to foster grid stability and decarbonise?* URL: [https://www.europarl.europa.eu/RegData/etudes/STUD/2018/626091/IPOL_STU\(2018\)626091_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2018/626091/IPOL_STU(2018)626091_EN.pdf).
- [65] Thomas Heggarty et al. "Quantifying power system flexibility provision". In: *Applied Energy* 279 (Dec. 1, 2020), p. 115852. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2020.115852. URL: <https://www.sciencedirect.com/science/article/pii/S030626192031326X> (visited on 05/21/2021).
- [66] Hannele Holttinen et al. "System impact studies for near 100% renewable energy systems dominated by inverter based variable generation". In: *IEEE Transactions on Power Systems* (2020). Conference Name: IEEE Transactions on Power Systems, pp. 1–1. ISSN: 1558-0679. DOI: 10.1109/TPWRS.2020.3034924.
- [67] Chi K. Tse et al. "Circuits and Systems Issues in Power Electronics Penetrated Power Grid". In: *IEEE Open Journal of Circuits and Systems* 1 (2020). Conference Name: IEEE Open Journal of Circuits and Systems, pp. 140–156. ISSN: 2644-1225. DOI: 10.1109/OJCAS.2020.3020633.
- [68] TENNET. *Report on systemic issues*. Horizon 2020 – MIGRATE. URL: https://www.h2020-migrate.eu/_Resources/Persistent/9bf78fc978e534f6393afb1f8510db86e56a1177/MIGRATE_D1.1_final_TenneT.pdf (visited on 05/21/2021).
- [69] U.S Department of Energy and U.S. Department of Homeland Security. "Energy Sector-Specific Plan". In: (2015), p. 39. URL: <https://www.cisa.gov/sites/default/files/publications/nipp-ssp-energy-2015-508.pdf>.
- [70] *Trends and projections in Europe reports — European Environment Agency*. URL: <https://www.eea.europa.eu/themes/climate/trends-and-projections-in-europe> (visited on 10/06/2021).
- [71] Norton Rose Fulbright LLP in association with Europex. *The Impact of Financial Services Regulation on European Wholesale Energy Markets - A Post-MiFID II Analysis.pdf*. 2020. URL: <https://www.europex.org/wp-content/uploads/2020/09/The-Impact-of-Financial-Services-Regulation-on-European-Wholesale-Energy-Markets-A-Post-MiFID-II-Analysis.pdf> (visited on 05/21/2021).
- [72] *List of Standard Contracts – documents*. URL: <https://documents.acer-remit.eu/remittance-reporting-user-package/list-of-standard-contracts/> (visited on 05/21/2021).

- [73] ACER. *REMIT Guidance on the List of Standard Contracts*. Oct. 8, 2015. URL: https://documents.acer-remit.eu/wp-content/uploads/ACER_REMIT_SC_List_Guidance_v2.pdf (visited on 05/21/2021).
- [74] European Federation of Energy Traders (EFET). *Price formation and capacity withholding in light of Regulation (EU) 2019/943 and Regulation (EU) 1227/2011 - EFET position paper*. Jan. 31, 2020. URL: https://efet.org/Files/Documents/Electricity%20Market/General%20market%20design%20and%20governance/EFET_paper_free_formation_of_prices_31012020_final.pdf (visited on 05/21/2021).
- [75] Michael G. Pollitt. “The European Single Market in Electricity: An Economic Assessment”. In: *Review of Industrial Organization* 55.1 (Aug. 1, 2019), pp. 63–87. ISSN: 1573-7160. DOI: 10.1007/s11151-019-09682-w. URL: <https://doi.org/10.1007/s11151-019-09682-w> (visited on 05/01/2020).
- [76] *ALL NEMO COMMITTEE - SDAC*. URL: <http://www.nemo-committee.eu/sdac> (visited on 05/21/2021).
- [77] ENTSO-E. *Options for the design of European Electricity Markets in 2030 - Discussion Paper for Stakeholder Consultation*. Mar. 31, 2021. URL: https://eepublicdownloads.entsoe.eu/clean-documents/Publications/Market%20Committee%20publications/210331_Market_design%202030.pdf (visited on 05/21/2021).
- [78] Peter Cramton. “Electricity market design”. In: *Oxford Review of Economic Policy* 33.4 (Nov. 2, 2017), pp. 589–612. ISSN: 0266-903X, 1460-2121. DOI: 10.1093/oxrep/grx041. URL: <http://academic.oup.com/oxrep/article/33/4/589/4587939> (visited on 05/21/2021).
- [79] Tim Schittekatte and Leonardo Meeus. “Flexibility markets: Q&A with project pioneers”. In: *Utilities Policy* 63 (Apr. 1, 2020), p. 101017. ISSN: 0957-1787. DOI: 10.1016/j.jup.2020.101017. URL: <http://www.sciencedirect.com/science/article/pii/S0957178720300126> (visited on 11/04/2020).
- [80] Lion Hirth and Ingmar Schlecht. “Market-Based Redispatch in Zonal Electricity Markets”. In: *SSRN Electronic Journal* (2018). ISSN: 1556-5068. DOI: 10.2139/ssrn.3286798. URL: <https://www.ssrn.com/abstract=3286798> (visited on 06/10/2021).
- [81] *REMIT PORTAL*. URL: <https://www.acer-remit.eu/portal/organised-marketplaces> (visited on 05/21/2021).
- [82] *Central counterparty (CCP) - Emissions-EUETS.com*. URL: <https://www.emissions-euets.com/internal-electricity-market-glossary/548-central-counterparty-ccp> (visited on 05/21/2021).
- [83] Oliver Wyman and World Federation of Exchanges. *The Future of Clearing*, p. 46. URL: <https://www.oliverwyman.com/content/dam/oliver-wyman/v2/publications/2019/apr/The-Future-Of-Clearing.pdf>.

-
- [84] *ECC risk management Overview*. URL: <https://www.ecc.de/en/risk-management/overview> (visited on 05/21/2021).
- [85] *JAO about us*. URL: <https://www.jao.eu/about-us> (visited on 05/21/2021).
- [86] *European Competition Network (ECN) - EUR-Lex*. URL: <https://eur-lex.europa.eu/summary/EN/legissum:126109> (visited on 05/21/2021).
- [87] Tim Schittekatte, Valerie Reif, and Leonardo Meeus. *The EU electricity network codes (2020 ed.)* Technical Report. Accepted: 2020-07-03T14:46:16Z ISBN: 9789290849056. European University Institute, 2020. DOI: 10.2870/101042. URL: <https://cadmus.eui.eu/handle/1814/67610> (visited on 05/21/2021).
- [88] *Financial Style Regulation in Wholesale Energy: Continuing Complexity - Florence School of Regulation*. URL: <https://fsr.eui.eu/financial-style-regulation-in-wholesale-energy-continuing-complexity/> (visited on 05/21/2021).
- [89] Liebrich M Hiemstra. “Energy trading and the exchange of information between supervisors: effectiveness of fragmented supervision and information sharing”. In: *Journal of Energy & Natural Resources Law* 39 (2021). DOI: <https://doi.org/10.1080/02646811.2020.1841400>. URL: <https://www.tandfonline.com/doi/full/10.1080/02646811.2020.1841400> (visited on 05/21/2021).
- [90] ACER. *Guidance on the application of Regulation (EU) No 1227/2011 of the European Parliament and of the Council of 25 October 2011 on wholesale energy market integrity and transparency*. 5th edition. May 11, 2021. URL: <https://documents.acer-remit.eu/category/guidance-on-remit/> (visited on 05/21/2021).
- [91] *Enforcement decisions*. URL: <https://acer.europa.eu:443/en/remit/REMITATACER/Pages/Enforcement-decisions.aspx> (visited on 05/21/2021).
- [92] *REMIT PORTAL*. URL: <https://www.acer-remit.eu/portal/european-register> (visited on 05/21/2021).
- [93] EPEX Spot. *EPEX Spot 2019 annual report*. URL: https://www.epexspot.com/sites/default/files/download_center_files/Epex-spot-2019_200703_Planche.pdf (visited on 05/21/2021).
- [94] *Exchange Members / EPEX SPOT*. URL: <https://www.epexspot.com/en/exchangemembers#list-of-ecc-clearing-members> (visited on 05/21/2021).
- [95] Andrew Beattie. *Commodities Speculators: More Help Than Harm?* Investopedia. URL: <https://www.investopedia.com/articles/basics/09/the-function-of-speculators.asp> (visited on 05/21/2021).

- [96] Kai Arulkumaran, Antoine Cully, and Julian Togelius. “AlphaStar: an evolutionary computation perspective”. In: *Proceedings of the Genetic and Evolutionary Computation Conference Companion*. GECCO '19. New York, NY, USA: Association for Computing Machinery, July 13, 2019, pp. 314–315. ISBN: 978-1-4503-6748-6. DOI: 10.1145/3319619.3321894. URL: <https://doi.org/10.1145/3319619.3321894> (visited on 05/21/2021).
- [97] Julian Schrittwieser et al. “Mastering Atari, Go, chess and shogi by planning with a learned model”. In: *Nature* 588.7839 (Dec. 2020). Number: 7839 Publisher: Nature Publishing Group, pp. 604–609. ISSN: 1476-4687. DOI: 10.1038/s41586-020-03051-4. URL: <https://www.nature.com/articles/s41586-020-03051-4> (visited on 05/21/2021).
- [98] Joel Lehman et al. “The Surprising Creativity of Digital Evolution: A Collection of Anecdotes from the Evolutionary Computation and Artificial Life Research Communities”. In: *Artificial Life* 26.2 (2020), pp. 274–306. ISSN: 1530-9185. DOI: 10.1162/artl_a_00319.
- [99] OECD. *Algorithms and Collusion: Competition Policy in the Digital Age*. 2017. URL: <http://www.oecd.org/competition/algorithms-collusion-competition-policy-in-the-digital-age.htm>.
- [100] NEMO COMMITTEE. *EUPHEMIA Public Description - Single Price Coupling Algorithm*. English. Oct. 2020. URL: <http://www.nemo-committee.eu/assets/files/euphemia-public-description.pdf> (visited on 12/14/2020).
- [101] Arnaud de Latour. *Electricity markets*. Feb. 2016. URL: <https://cermics.enpc.fr/~delmas/Enseig/levy-delatour-market.pdf> (visited on 06/26/2019).
- [102] Stéphane Goutte and Philippe Vassilopoulos. “The value of flexibility in power markets”. In: *Energy Policy* 125 (Feb. 1, 2019), pp. 347–357. ISSN: 0301-4215. DOI: 10.1016/j.enpol.2018.10.024. URL: <https://www.sciencedirect.com/science/article/pii/S0301421518306785> (visited on 01/07/2022).
- [103] Klaus Mayer and Stefan Trück. “Electricity markets around the world”. In: *Journal of Commodity Markets* 9 (Mar. 2018), pp. 77–100. ISSN: 2405-8513. DOI: 10.1016/j.jcomm.2018.02.001. URL: <http://www.sciencedirect.com/science/article/pii/S2405851318300059> (visited on 10/22/2018).
- [104] Piotr F. Borowski. “Zonal and Nodal Models of Energy Market in European Union”. en. In: *Energies* 13.16 (Aug. 2020), p. 4182. ISSN: 1996-1073. DOI: 10.3390/en13164182. URL: <https://www.mdpi.com/1996-1073/13/16/4182> (visited on 09/09/2021).
- [105] *Electricity markets in the EU - Florence School of Regulation*. URL: <https://fsr.eui.eu/electricity-markets-in-the-eu/> (visited on 05/26/2021).
- [106] *ACER Decision 35-2020 on Core RDCT 35 - Annex I.pdf*. URL: <https://www.eles.si/Portals/0/Novice/DOKUMENTI/ACER%20Decision%2035-2020%20on%20Core%20RDCT%2035%20-%20Annex%20I.pdf> (visited on 11/04/2021).

-
- [107] European Commission. *Commission Regulation (EU) No 543/2013 of 14 June 2013 on submission and publication of data in electricity markets and amending Annex I to Regulation (EC) No 714/2009 of the European Parliament and of the Council Text with EEA relevance*. URL: <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:163:0001:0012:EN:PDF>.
 - [108] Kenneth Van den Bergh, Jonas Boury, and Erik Delarue. “The Flow-Based Market Coupling in Central Western Europe: Concepts and definitions”. In: *The Electricity Journal* 29.1 (Jan. 1, 2016), pp. 24–29. ISSN: 1040-6190. DOI: 10.1016/j.tej.2015.12.004. URL: <https://www.sciencedirect.com/science/article/pii/S104061901530004X> (visited on 10/26/2021).
 - [109] All NEMO Committee. *CACM Annual Report 2019*. July 1, 2020. URL: [http://www.nemo-committee.eu/assets/files/NEMO_CACM_Annual_Report_2019_final30062020\(3\)-942c4cb6c19d32485bfe6ab57f13581c.pdf](http://www.nemo-committee.eu/assets/files/NEMO_CACM_Annual_Report_2019_final30062020(3)-942c4cb6c19d32485bfe6ab57f13581c.pdf) (visited on 05/21/2021).
 - [110] ACER. *ACER Market Monitoring Report 2019 – Electricity Wholesale Markets Volume*. Feb. 12, 2021. URL: https://www.acer.europa.eu/Official_documents/Acts_of_the_Agency/Publication/ACER%20Market%20Monitoring%20Report%202019%20-%20Electricity%20Wholesale%20Markets%20Volume.pdf (visited on 05/21/2021).
 - [111] Angelica Gianfreda, Lucia Parisio, and Matteo Pelagatti. “Revisiting long-run relations in power markets with high RES penetration”. In: *Energy Policy* 94 (July 1, 2016), pp. 432–445. ISSN: 0301-4215. DOI: 10.1016/j.enpol.2016.02.006. URL: <https://www.sciencedirect.com/science/article/pii/S0301421516300477> (visited on 06/13/2022).
 - [112] Rafał Weron. “Stylized Facts of Electricity Loads and Prices”. In: *Modeling and Forecasting Electricity Loads and Prices*. Wiley Online Library, July 29, 2013. URL: <https://onlinelibrary.wiley.com/doi/10.1002/9781118673362.ch2> (visited on 11/03/2021).
 - [113] Angelica Gianfreda and Luigi Grossi. “Forecasting Italian electricity zonal prices with exogenous variables”. In: *Energy Economics* 34.6 (Nov. 1, 2012), pp. 2228–2239. ISSN: 0140-9883. DOI: 10.1016/j.eneco.2012.06.024. URL: <https://www.sciencedirect.com/science/article/pii/S0140988312001338> (visited on 10/19/2021).
 - [114] Cristina Ballester and Dolores Furió. “Effects of renewables on the stylized facts of electricity prices”. In: *Renewable and Sustainable Energy Reviews* 52 (Dec. 1, 2015), pp. 1596–1609. ISSN: 1364-0321. DOI: 10.1016/j.rser.2015.07.168. URL: <https://www.sciencedirect.com/science/article/pii/S1364032115008151> (visited on 05/28/2021).
 - [115] W. S. Cleveland, E. Grosse, and W.M. Shyu. “Chapter 8 - Local regression models”. In: *Statistical Models in S*. Chapman and Hall/CRC, 1992, p. 624. ISBN: 978-0-203-73853-5.

- [116] Leo Breiman. “Random Forests”. English. In: *Machine Learning* 45.1 (2001), pp. 5–32. ISSN: 0885-6125. DOI: 10.1023/A:1010933404324.
- [117] ENTSO-E. *Network code on Capacity Allocation & Congestion Management (CACM) - Supporting document*. Mar. 23, 2012. URL: https://eepublicdownloads.entsoe.eu/clean-documents/pre2015/consultations/Network_Code_CACM/120323_NetworkCodeConsultation_FINAL.pdf (visited on 05/27/2021).
- [118] Jesus Lago, Fjo De Ridder, and Bart De Schutter. “Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms”. In: *Applied Energy* 221 (July 2018), pp. 386–405. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2018.02.069. URL: <http://www.sciencedirect.com/science/article/pii/S030626191830196X> (visited on 10/22/2018).
- [119] Hans-Kristian Ringkjøb, Peter M. Haugan, and Ida Marie Solbrekke. “A review of modelling tools for energy and electricity systems with large shares of variable renewables”. In: *Renewable and Sustainable Energy Reviews* 96 (Nov. 2018), pp. 440–459. ISSN: 1364-0321. DOI: 10.1016/j.rser.2018.08.002. URL: <http://www.sciencedirect.com/science/article/pii/S1364032118305690> (visited on 05/16/2019).
- [120] Joint Research Centre (European Commission). *Systematic mapping of power system models*. en. Tech. rep. Publications Office, 2017. URL: <https://data.europa.eu/doi/10.2760/422399> (visited on 10/28/2020).
- [121] Lesia Mitridati and Pierre Pinson. “A Bayesian Inference Approach to Unveil Supply Curves in Electricity Markets”. en. In: *IEEE Transactions on Power Systems* 33.3 (May 2018), pp. 2610–2620. ISSN: 0885-8950, 1558-0679. DOI: 10.1109/TPWRS.2017.2757980. URL: <http://ieeexplore.ieee.org/document/8053811/> (visited on 01/29/2019).
- [122] C. Ruiz, A. J. Conejo, and D. J. Bertsimas. “Revealing Rival Marginal Offer Prices Via Inverse Optimization”. In: *IEEE Transactions on Power Systems* 28.3 (Aug. 2013), pp. 3056–3064. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2012.2234144. URL: <https://ieeexplore.ieee.org/abstract/document/6423235>.
- [123] Ruidi Chen et al. “Learning From Past Bids to Participate Strategically in Day-Ahead Electricity Markets”. In: *IEEE Transactions on Smart Grid* 10.5 (Sept. 2019). Conference Name: IEEE Transactions on Smart Grid, pp. 5794–5806. ISSN: 1949-3061. DOI: 10.1109/TSG.2019.2891747.
- [124] John R. Birge, Ali Hortaçsu, and J. Michael Pavlin. “Inverse Optimization for the Recovery of Market Structure from Market Outcomes: An Application to the MISO Electricity Market”. en. In: *Operations Research* 65.4 (Aug. 2017), pp. 837–855. ISSN: 0030-364X, 1526-5463. DOI: 10.1287/opre.2017.1606. URL: <http://pubsonline.informs.org/doi/10.1287/opre.2017.1606> (visited on 08/30/2021).

-
- [125] René Carmona, Michael Coulon, and Daniel Schwarz. “Electricity price modeling and asset valuation: a multi-fuel structural approach”. en. In: *Mathematics and Financial Economics* 7.2 (Mar. 2013), pp. 167–202. ISSN: 1862-9679, 1862-9660. DOI: 10.1007/s11579-012-0091-4. URL: <http://link.springer.com/10.1007/s11579-012-0091-4> (visited on 02/19/2019).
 - [126] Lion Hirth. “What caused the drop in European electricity prices? A factor decomposition analysis”. en. In: *The Energy Journal* 39.1 (Jan. 2018). ISSN: 01956574. DOI: 10.5547/01956574.39.1.lhir. URL: <http://www.iaee.org/en/publications/ejarticle.aspx?id=3031> (visited on 03/04/2019).
 - [127] Y. He et al. “Modeling the Merit Order Curve of the European Energy Exchange Power Market in Germany”. In: *IEEE Transactions on Power Systems* 28.3 (Aug. 2013), pp. 3155–3164. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2013.2242497.
 - [128] Florian Ziel and Rick Steinert. “Probabilistic mid- and long-term electricity price forecasting”. In: *Renewable and Sustainable Energy Reviews* 94 (Oct. 2018), pp. 251–266. ISSN: 1364-0321. DOI: 10.1016/j.rser.2018.05.038. URL: <http://www.sciencedirect.com/science/article/pii/S1364032118303885> (visited on 02/07/2019).
 - [129] Sergei Kulakov. “X-Model: Further Development and Possible Modifications”. en. In: *Forecasting* 2.1 (Mar. 2020), pp. 20–35. DOI: 10.3390/forecast2010002. URL: <https://www.mdpi.com/2571-9394/2/1/2> (visited on 11/03/2020).
 - [130] René Aïd, Luciano Campi, and Nicolas Langrené. “A Structural Risk-Neutral Model for Pricing and Hedging Power Derivatives”. en. In: *Mathematical Finance* 23.3 (2013), pp. 387–438. ISSN: 1467-9965. DOI: 10.1111/j.1467-9965.2011.00507.x. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9965.2011.00507.x> (visited on 02/19/2019).
 - [131] A. A. Sánchez de la Nieta and J. Contreras. “Quantifying the effect of renewable generation on day-ahead electricity market prices: The Spanish case”. en. In: *Energy Economics* 90 (Aug. 2020), p. 104841. ISSN: 0140-9883. DOI: 10.1016/j.eneco.2020.104841. URL: <https://www.sciencedirect.com/science/article/pii/S014098832030181X> (visited on 03/26/2021).
 - [132] X. Xia and A. M. Elaiw. “Optimal dynamic economic dispatch of generation: A review”. en. In: *Electric Power Systems Research* 80.8 (Aug. 2010), pp. 975–986. ISSN: 0378-7796. DOI: 10.1016/j.epsr.2009.12.012. URL: <http://www.sciencedirect.com/science/article/pii/S0378779610000027> (visited on 01/27/2021).
 - [133] Saleh Y. Abujarad, M. W. Mustafa, and J. J. Jamian. “Recent approaches of unit commitment in the presence of intermittent renewable energy resources: A review”. en. In: *Renewable and Sustainable Energy Reviews* 70 (Apr. 2017), pp. 215–223. ISSN: 1364-0321. DOI: 10.1016/j.rser.2016.11.246. URL: <http://www.sciencedirect.com/science/article/pii/S1364032116310140> (visited on 11/10/2020).

- [134] Open Power System Data. “Data Package Time series”. In: . Version 2019-06-05 (2019). DOI: 10.25832/time_series/2019-06-05.
- [135] RTE and ADEME. *Signal prix du CO₂, Analyse de son impact sur le système électrique européen*. French. Mar. 2016. URL: https://presse.ademe.fr/wp-content/uploads/2016/04/Signal-prix-du-CO2_finale_29-03-16.pdf (visited on 09/21/2020).
- [136] R.J. Hyndman and G. Athanasopoulos. “Forecasting: principles and practice”. In: *Forecasting: Principles and Practice*. 3rd Edition. Melbourne, Australia: OTexts. URL: <https://otexts.com/fpp3/> (visited on 12/28/2021).

RÉSUMÉ

Les marchés de l'électricité se sont développés au cours des dernières décennies à travers le monde afin faciliter les échanges entre producteurs et consommateurs. Du fait de l'évolution des prix de gros de l'électricité en fonction de l'équilibre offre-demande, de nouveaux risques et opportunités émergent constamment, non seulement pour les acteurs du marché mais aussi pour la société dans son ensemble. L'évolution des prix est particulièrement incertaine dans le contexte de la transition énergétique en raison de changements considérables au sein du système électrique, notamment l'importance croissante des renouvelables dans le mix électrique. Néanmoins, anticiper les possibles évolutions des prix est nécessaire pour prendre des décisions à long terme adéquates en ce qui concerne les investissements, la conception des règles de marché et la définition de politiques publiques.

L'objectif principal de cette thèse est de proposer une méthode pour simuler des séries temporelles pluriannuelles des prix de l'électricité sur les marchés journaliers. La nouveauté de notre modèle réside dans son objectif de rapprocher (i) les méthodes d'optimisation qui permettent d'estimer le coût total minimal pour fournir l'électricité, et (ii) les méthodes basées sur les données qui visent à inférer à partir des tendances observées la façon dont le marché résout efficacement ce problème d'allocation des ressources. Le rapprochement de ces deux approches permet de s'appuyer sur des fondements théoriques solides tout en considérant les effets observés en pratique.

Le modèle proposé tient compte des contraintes physiques qui s'imposent à la fourniture d'électricité, tout en reflétant le mécanisme réel de formation des prix sur les marchés journaliers (ordre de préséance économique et principe de la tarification marginale). Un avantage clé de notre méthode est qu'elle permet d'exploiter pleinement les données de marché (par exemple, les historiques de prix et les courbes d'offre), ainsi que les données du système électrique (par exemple, la disponibilité des unités de production et l'évolution de la demande). Pour cela, nous proposons une paramétrisation des ordres de vente complétée d'une méthode d'estimation statistique pour fixer les valeurs de ces paramètres. Cette technique de modélisation offre une grande modularité et nous permet de prendre en compte l'influence de plusieurs facteurs influant sur les prix de marché, tels que les enchères stratégiques qui diffèrent du coût marginal de production. De ce fait, nous sommes capables de simuler des séries temporelles de prix qui présentent des dynamiques représentatives des variations horaires observées pour les prix sur les marchés journaliers. Ce modèle validé est ensuite exploité à travers des études prospectives évaluant les potentielles évolutions des prix de gros de l'électricité dans le cadre de futurs scénarios technico-économiques.

MOTS CLÉS

Marché de l'électricité, Planification énergétique, Système électrique, Énergies renouvelables, Transition énergétique, Évolution des prix d'électricité, Analyse prospective

ABSTRACT

Electricity markets have been developed worldwide during the past decades to facilitate energy transaction. Since wholesale electricity prices evolve depending on the supply-demand equilibrium, risks and opportunities constantly emerge, not only for market participants but also for society as a whole. The evolution of prices is especially uncertain in the context of the energy transition due to considerable changes in the power system, notably the growing importance of renewables in the electricity mix. Nevertheless, anticipating possible price developments is needed to make adequate long-term decisions in terms of investments, market design, and public policies.

The main objective of this thesis is to propose a method to simulate multiannual time series of electricity prices on day-ahead markets. The novelty of our model lies in its goal to bridge the gap between (i) optimization methods that estimate the minimal total cost to supply electricity, and (ii) data-driven methods that aim at capturing the observed trends in how the market effectively solves this resource allocation problem. Bringing the two approaches together enables us to build upon robust theoretical foundations without neglecting the effects observed in practice.

The proposed model accounts for the physical constraints imposed on electricity delivery, while also reflecting the actual price formation mechanism on day-ahead markets (merit order and marginal pricing). A key advantage of our method is that it makes possible to fully exploit market data (e.g. historical prices, supply curves), as well as power system data (e.g. availability of generation units, evolution of demand). To this end, we propose a parametrization of sell orders, and a statistical estimation method to set the values of these parameters. This modelling technique offers great modularity and allows us to consider the influence of multiple price drivers, such as strategic bidding. As a result, we are able to simulate time series of prices, which display dynamics similar to the hourly price variations observed on day-ahead markets. This validated model is then exploited to conduct prospective studies in order to assess the potential evolution of wholesale electricity prices under future techno-economic scenarios.

KEYWORDS

Electricity markets, Power systems planning, Electric power system, Renewable energy, Energy transition, Electricity price dynamics, Prospective studies