



HAL
open science

Data-driven Structural Modeling of Electricity Price Dynamics

Valentin Mahler, Robin Girard, Georges Kariniotakis

► **To cite this version:**

Valentin Mahler, Robin Girard, Georges Kariniotakis. Data-driven Structural Modeling of Electricity Price Dynamics. 2021. hal-03445396

HAL Id: hal-03445396

<https://hal.science/hal-03445396>

Preprint submitted on 23 Nov 2021

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.

Data-driven Structural Modeling of Electricity Price Dynamics

Valentin Mahler^{a,b,*}, Robin Girard^a, Georges Kariniotakis^a

^a*MINES ParisTech, PSL University, Center for processes, renewable energies and energy systems (PERSEE), 1 Rue Claude Daunesse, CS 10207, 06904 Sophia-Antipolis, France*

^b*ADEME - French agency for ecological transition, 20 avenue du Grésillé - BP 90406, 49004 Angers Cedex 01, France*

Abstract

In many countries, electricity prices on day-ahead auction markets result from a market clearing designed to maximize social welfare. For each hour of the day, the market price can be represented as the intersection of a supply and demand curve. Structural market models reflect this price formation mechanism and are widely used in prospective studies guiding long-term decisions (e.g. investments and market design). However, simulating the supply curve in these models proves challenging since estimating the sell orders it comprises (i.e. offer prices and corresponding quantities) typically requires formulating numerous techno-economic hypotheses about power system assets and the behaviors of market participants. In this paper we propose a method for the parameterization of sell orders associated with production units. The estimation algorithm for this parametrization makes it possible to mitigate the requirement for analytic formulation of all of the above-mentioned aspects and to take advantage of the ever-increasing volume of available data on power systems (e.g. technical and market data). Parametrized orders also offer the possibility to account for various factors in a modular fashion, such as the strategic behavior of market participants. The proposed approach is validated using data related to the French day-ahead market and power system, for the period from 2015 to 2018.

Keywords: Day-ahead markets, Electricity prices, Structural market model, Prospective studies, Power systems

*Corresponding author

Email address: valentin.mahler@mines-paristech.fr (Valentin Mahler)

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

Subscripts

i	index for a production unit in set I
j	index for a consumption unit in set J
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

1. Introduction

1.1. Context

Modern societies operate under the assumption that high security of electricity supply is guaranteed [1]. Over the last decades, electricity markets have been seen as a pertinent way to provide the supply necessary for the equilibrium of the system [2].

Studying and anticipating the evolution of wholesale market prices helps private and public stakeholders to better evaluate the opportunities and risks they face [3] [4] [5]. For example, renewable energy sources (RES) power

plants, when not under a subsidy regime, generate their revenue through participation in electricity markets. It is useful for the design of the RES investment to have estimates of the evolution of electricity prices to assess the payback and associated risks. Making estimates based on the evolution of electricity prices in the years to come is highly challenging due to uncertainties. This calls for the development of numerical models able to account for the effect of factors of different natures on prices. Such models can help to gain a better understanding of the mechanisms underlying the formation of prices, resulting in more rational decisions regarding investments, operation and maintenance, more appropriate market designs, and easier market surveillance. In practice, wholesale electricity prices vary according to the supply-demand equilibrium established by market participants, while considering the techno-economic constraints of the power system assets and in compliance with a strict regulatory framework that defines the underlying market mechanism. Thus, the influencing factors of price dynamics include: evolution of demand, composition of the energy mix, market design, and strategic behavior of market participants.

In deregulated energy markets, multiple possibilities are offered to procure or supply electricity. Participants trade on specialized exchange markets where buy and sell orders are pooled, and also trade over-the-counter through bilateral contracts. Each transaction is associated with a price that is not necessarily disclosed. 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 published by the power exchanges. This price results from the clearing of all buy and sell orders, submitted one day before physical delivery. The usual design of this market is presented in [6] and the technical implementation for the case of Europe is described in [7]. It is worth noting that this price signal commonly also serves as an indicator for pricing on other markets, such as forward markets (e.g. in [8]).

As mentioned above, in different applications like RES investments it is beneficial to have an idea about the evolution of electricity prices. To achieve this requires being able to model the market mechanism and the evolution of the factors influencing price dynamics. For this purpose, a growing volume of data is available and helps to reveal the effects observed in practice. This is especially true now that power system data (e.g. static data, such as the installed capacity by generation technology, and dynamic data, such as the hourly outages of production units) are more accessible than ever due to

supportive regulation (e.g. data available on the ENTSO-E Transparency Platform [9] as specified by the regulation [10]) and thanks to data sharing initiatives (e.g. Open Power System Data [11]). The availability of market data (e.g. time series of commodity prices, aggregated supply and demand curves) is also improving, but the costs associated with such financial data remain a concern in the EU [12] and in the US [13].

For different applications in the context of energy transition it is nevertheless necessary to look beyond global trends in electricity prices and consider their modeling with a high temporal resolution (i.e. hourly) by accounting for techno-economic constraints related to the operation of the power system. This reveals situations where the intraday variations of prices impact the performance of the considered solutions; this could for example be the case when studying incentives to develop flexibility both on the supply and demand sides (e.g. storage, demand side response). In the case of less flexible production technologies driven by weather conditions (e.g. solar and wind power) or limited by operational constraints (e.g. start and stop costs or ramp limits for thermal units), being able to simulate their exposure to low or even negative prices helps to assess their financial performance from the point of view of an investor.

1.2. Modeling prices on day-ahead markets

Various approaches have been proposed in the literature to explore the relation 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. These purely statistical models rely on observed correlations between variables and have good forecasting accuracy, especially for short-term horizons [14] 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 [15], also called the shadow price. These types of models are popu-

lar 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 [16] and a systematic mapping of power system models in [17]. 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 (however, it is worth noting that the European Commission recently released some data sets 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 [18] or on inverse optimization approaches [19], [20]. 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 not to exploit observed market data.

Alternatively, the system equilibrium can be considered from a more market-oriented perspective by simulating supply curves [21]. 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. [22] and [23]). A survey of structural models is proposed in [24].

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 [25], 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 [26] and used in [27] 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 [28], 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 [29] 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 [30], 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 [31], 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. Refer-

ences [32] and [23] 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 [32] to adjust the output of the merit order model to account for effects such as ramp constraints and strategic bidding, while [23] 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.

1.3. Key contributions

As presented in this introduction, 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 techno-economic 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. In light of the scientific literature and extending [33], 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 propose the following main contributions in this paper:

1. We propose a new 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 (ii) exploit market and power system data to calibrate and evaluate the model so that the simulations reflect the prices observed in practice.

2. We propose advanced approaches for the calibration of this structural model. The calibration process allows us to determine the value of parameters by production class, to modulate the importance given to market prices depending on their range, and to account for the observed supply curves' prices.
3. We validate the proposed method using data related to the French day-ahead market and power system, for the period from 2015 to 2018, and we show the benefits that statistical calibration can bring to structural models.

1.4. Paper structure

The remainder of this article is structured as follows: in section 2, we present our methodology, then we illustrate the concept on a use case in section 3. The advantages and limitations of the method are discussed in section 4. Finally, conclusions and some perspectives for further work are proposed in section 5.

2. Problem statement and methodology

2.1. 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 paper to add too much complexity here. For a more complete formulation and discussion on economic dispatch, see e.g. the review paper [34]. The formulation could be further extended to a unit commitment as presented in the review paper [35]. 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:

$$\text{maximize}_{P, L, E} \quad \sum_{t \in T} \left(\sum_{j \in J} \pi_{t,j}^L L_{t,j} - \sum_{i \in I} \pi_{t,i}^P P_{t,i} \right) \quad (1a)$$

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

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

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

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

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

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 (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 (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.

2.2. Parametrization of marginal costs

To account for the above-mentioned effects, an explicit modeling of e.g. futures and forwards markets, internal strategic bidding would dramatically increase the complexity of the approach and require data that are generally unavailable. This would result in model constraints not included in our initial model, which we can call external constraints. These constraints would lead to adding Lagrange multipliers to the marginal cost to generate sell prices. Using the preceding formulation, the sell prices would be $\pi_{t,i}^P + \lambda_{t,i}^{IC} + \lambda_{t,i}^{EC}$, with $\lambda_{t,i}^{EC}$ unknown Lagrange multipliers associated with external constraints.

We have observed that it is not possible to formally compute $\lambda_{t,i}^{EC}$; since the main drivers of this term include: global supply margin (i.e. the difference between the available capacity and the demand), fuel price, emission price and the rank of each production unit within its production class, we propose to build a statistical linear model $\hat{\pi}_{t,i}^P$ for $\pi_{t,i}^P + \lambda_{t,i}^{EC}$:

$$\begin{aligned}\hat{\pi}_{t,i}^P &= \alpha_{c(i),0} + \alpha_{c(i),\mathcal{F}} \cdot \mathcal{F}_{t,c(i)} + \alpha_{c(i),\mathcal{E}} \cdot \mathcal{E}_t + \alpha_{c(i),\mathcal{M}} \cdot \mathcal{M}_t + \alpha_{c(i),\mathcal{R}} \cdot \mathcal{R}_i \\ &= \boldsymbol{\alpha}_{c(i)}^\top \mathbf{X}_{t,i}\end{aligned}\quad (2)$$

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 $\hat{\pi}_{t,i}^P$ is no longer a simple marginal cost, we propose to call it an extended marginal cost. Since the sell price is obtained by adding $\lambda_{t,i}^{IC}$, the price that we propose is also a statistical model of sell orders. Using the extended marginal cost within the optimization problem defined in Eq. (1) results in simulating a day-ahead market price that we denote by $\hat{\Pi}_{t,z}^{MC}$.

2.3. 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 paper that we define as follows:

$$\begin{aligned}\hat{\Pi}_{t,z} &= \hat{\Pi}_{t,z}^{MC} + \beta_{t,z} \\ &= \hat{\Pi}_{t,z}^{MC} + \boldsymbol{\beta}_z^\top \mathbf{Y}_t\end{aligned}\tag{3}$$

where for each region z , $\boldsymbol{\beta}_z \in \mathbb{R}^{24 \times 7}$ 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 [32]. 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.

2.4. Metrics

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 order to also 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). Our algorithm can simultaneously optimize $RMSE$ and Δsd and generate the associated Pareto front. The metrics

used to evaluate the model performance of the forecast $(\hat{\Pi}_{t,z})_{t \in T}$ against the observed prices $(\Pi_{t,z})_{t \in T}$ are defined as follows:

$$RMSE_z = \sqrt{\frac{1}{T} \sum_{t=1}^T (\Pi_{t,z} - \hat{\Pi}_{t,z})^2} \quad (4)$$

$$\Delta sd_z = sd((\Pi_{t,z})_{t \in T}) - sd((\hat{\Pi}_{t,z})_{t \in T}) \quad (5)$$

2.5. Estimation algorithm

The parameters of Eq. (2) and Eq. (3) 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:

$$\begin{aligned} \hat{\Pi}_{t,z} &= \hat{\pi}_{t,i^*(t)}^P + \lambda_{t,i^*(t)}^{IC} + \beta_{t,z} \\ &= \boldsymbol{\alpha}_{c(i^*(t))}^\top \mathbf{X}_{t,i^*(t)} + \lambda_{t,i^*(t)}^{IC} + \beta_z^\top \mathbf{Y}_t \end{aligned} \quad (6)$$

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

Market Clearing Step. From a set of extended marginal costs we can solve Eq. (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 $(T_c)_{c \in C}$.

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, $\boldsymbol{\alpha}_{c_0}$ is obtained by a linear regression of $(\mathbf{X}_{t,i})_{(t,i) \in T_{c_0} \times I_z}$ on the observed price $(\Pi_{t,z})_{t \in T_{c_0}}$. 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 $\alpha_{c(i),\mathcal{E}}$, 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:

$$RMSE_z^2 = \frac{1}{T} \sum_{c_0 \in C} \sum_{t \in T_{c_0}} (\boldsymbol{\alpha}_{c_0}^\top \mathbf{X}_{t,i^*(t)} + \lambda_{t,i^*(t)}^{IC} + \boldsymbol{\beta}_z^\top \mathbf{Y}_t - \Pi_{t,z})^2 \quad (7)$$

It is important to note that the error here is not just a quadratic function of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}_z$ 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 $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}_z$. 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 $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}_z$ according to this linearization. This shows that our algorithm is similar to a gradient descent on the cost function.

2.6. Introducing variability into the cost function

The root mean square error is known to give results that are smoother than observations. This can be an issue for the kind of application we foresee (i.e. analyzing the variability evolution of market prices). 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:

$$RMSE_z = \left(\frac{1}{T} \sum_{c_0 \in C} \sum_{t \in T_{c_0}} \omega_{t,z} (\boldsymbol{\alpha}_{c_0}^\top \mathbf{X}_{t,i^*(t)} + \lambda_{t,i^*(t)}^{IC} + \boldsymbol{\beta}_z^\top \mathbf{Y}_t - \Pi_{t,z})^2 \right)^{\frac{1}{2}} \quad (8)$$

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

$$\frac{1}{|T_{c_0}|} \sum_{t \in T_{c_0}} \Pi_{t,z} = \frac{1}{\sum_{t \in T_{c_0}} \omega_{t,z}} \sum_{t \in T_{c_0}} \omega_{t,z} \Pi_{t,z} \quad (9)$$

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

2.7. 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:

$$\Pi_{t,z}^\delta = \mathcal{S}_{t,z}(V_{t,z}^\delta) \quad (10)$$

The process is illustrated in Fig. 1.

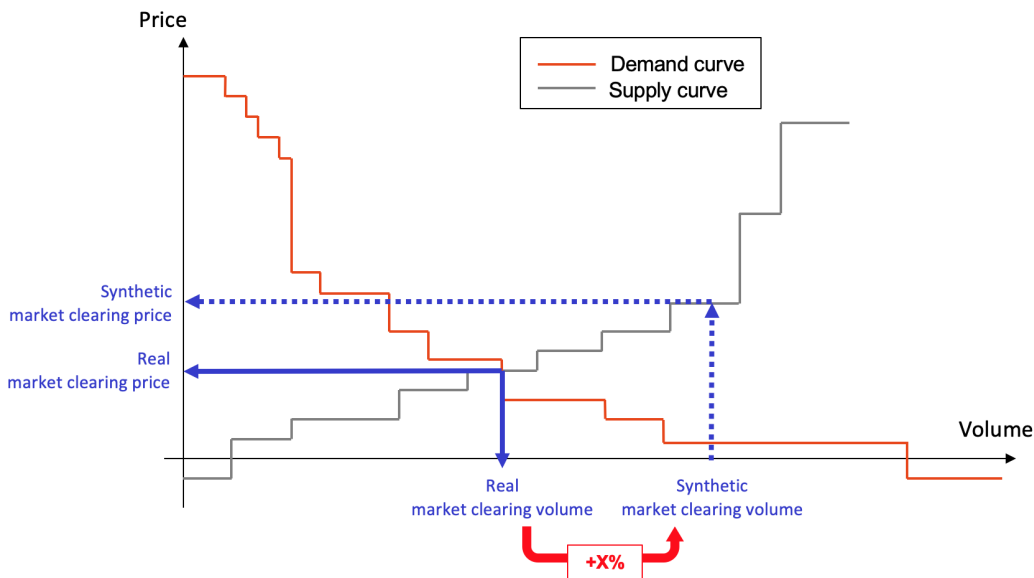


Figure 1: Data augmentation principle to generate synthetic prices from supply curves

For consistency, the same percentage change is also applied to demand and we account for the effect of this variation on the supply margin $\mathcal{M}_{t,z}$.

Therefore we associate $\mathbf{X}_{t,i}^\delta$ with $\Pi_{t,z}^\delta$, where only $\mathcal{M}_{t,z}^\delta$ has been modified and all other factors remain unchanged.

3. 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. (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, but will be the purpose of further research.

For the simulations, we split the global data set into four parts corresponding to the four calendar years, where each year is used for model training and out-of-sample testing. As final results, we consider the average of the simulated values (e.g. using the models trained with the 2015, 2016 and 2017 data sets, we simulate three time series of prices for 2018 and then calculate the hourly average). In this section, we first present the method implemented, then the data used, and finally we show the results obtained.

3.1. Method implementation

3.1.1. 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 and the foreign day-ahead market prices as order prices.

Fig. 2 gives an overview of the case study.

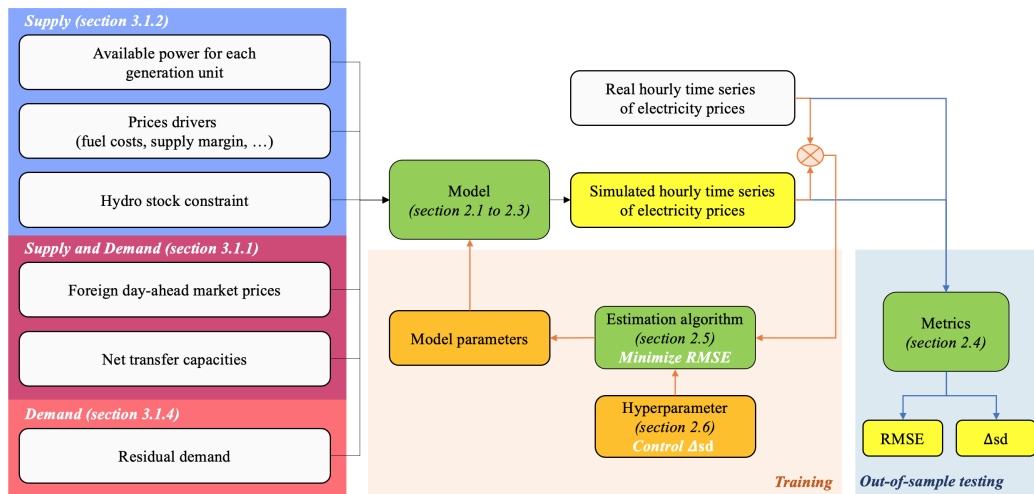


Figure 2: Case study overview

3.1.2. 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.

The price drivers considered by production class are summarized in Table 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.

Production class (c)	Unit ranking (\mathcal{R})	Supply margin (\mathcal{M})	Fuel cost (\mathcal{F})	Emission cost (\mathcal{E})
Hydro	x	x		
Nuclear	x	x		
Gas	x	x	x	x
Coal	x	x	x	x
Oil	x	x	x	x

Table 1: Price drivers of extended marginal costs considered by production class

3.1.3. Parameter estimation

For each class, we initialize $\alpha_{c,0}$ introduced in Eq. (2) to be constant and equal to an average short-run marginal cost (i.e. the cost associated to 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 \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 3.2.

For the reference case presented here, we estimate the model parameters with all weights ω_i equal to 1 and we do not exploit the real supply curves. In a second step, we also assess the impact of weight variations as presented in subsection 2.6, and of exploiting the supply curves as presented in subsection 2.7.

To estimate β introduced in Eq. (3), 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.

3.1.4. 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. This demand 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 ac-

ording to a baseload profile), storage (pumped-storage hydroelectricity) 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 paper to include complex technical constraints.

3.1.5. Market clearing

We solve Eq. (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.

3.2. Data

The raw data sets used for the simulation and results analysis are listed in Table 2.

Data set	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 [36]
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 [37]
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
Hourly supply and demand curves	EPEX SPOT
Atmospheric temperature	Meteo France

Table 2: Data sources for the case study

Two remarks can be made about the data sets:

- 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 data sets were collected from the web, except for EPEX Spot's supply and demand curves, and Meteo France's average atmospheric

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

temperatures, which are not open data but commercially available or under special conditions for research purposes. However, these two data sets are used here to further validate extensions of the core methodology (respectively for data augmentation and error analysis), therefore the core approach proposed can be developed and evaluated using publicly available data only.

3.3. Results

3.3.1. Reference case

The results obtained using the method and data described above are shown in Fig. 3. 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 more detailed analysis relying on the presented metrics features in the next subsection.

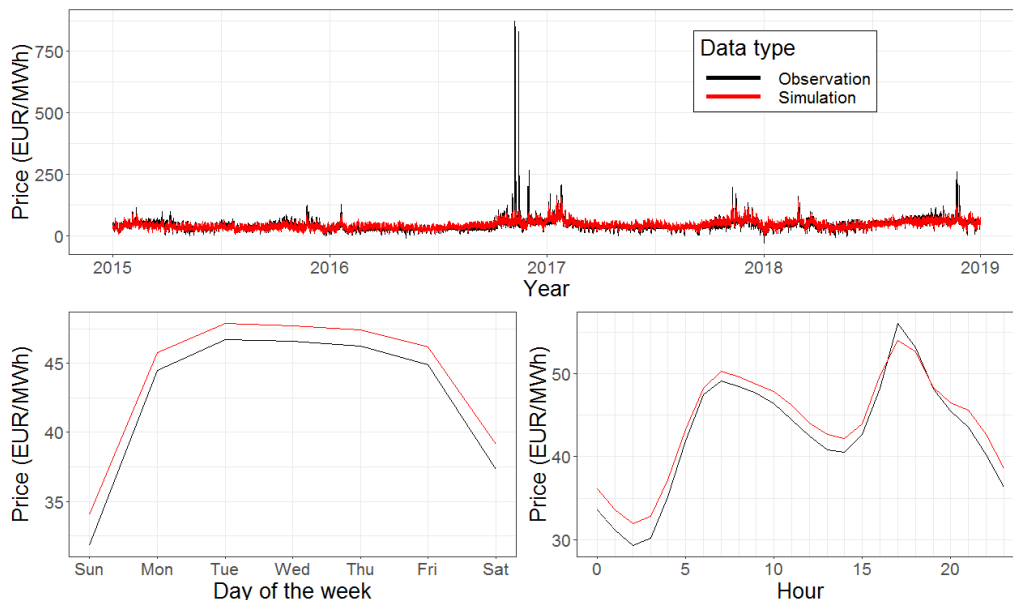


Figure 3: Time series of prices on the French day-ahead market over four years (top), average prices by day of the week (bottom-left) and average prices by hour of the day (bottom-right)

A focus on a period of one month presented in Fig. 4 gives a more detailed representation of the dynamics of day-ahead prices.

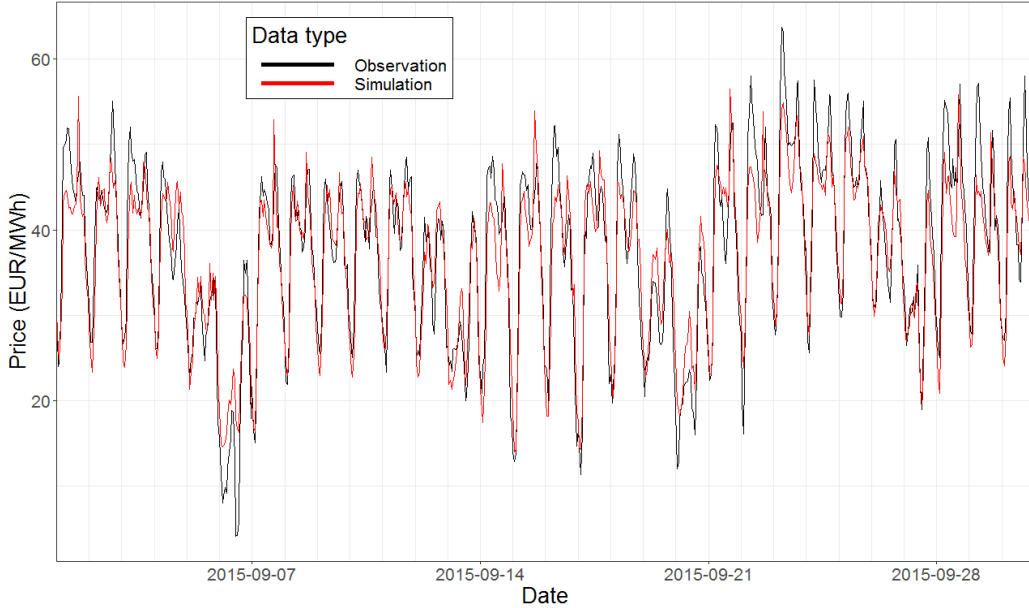


Figure 4: Time series of prices on the French day-ahead market over one month

3.3.2. Advanced calibration

Hyperparametrization. In Fig. 5, we show how tuning the variability hyperparameter of the model during calibration can improve 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. (8), 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.

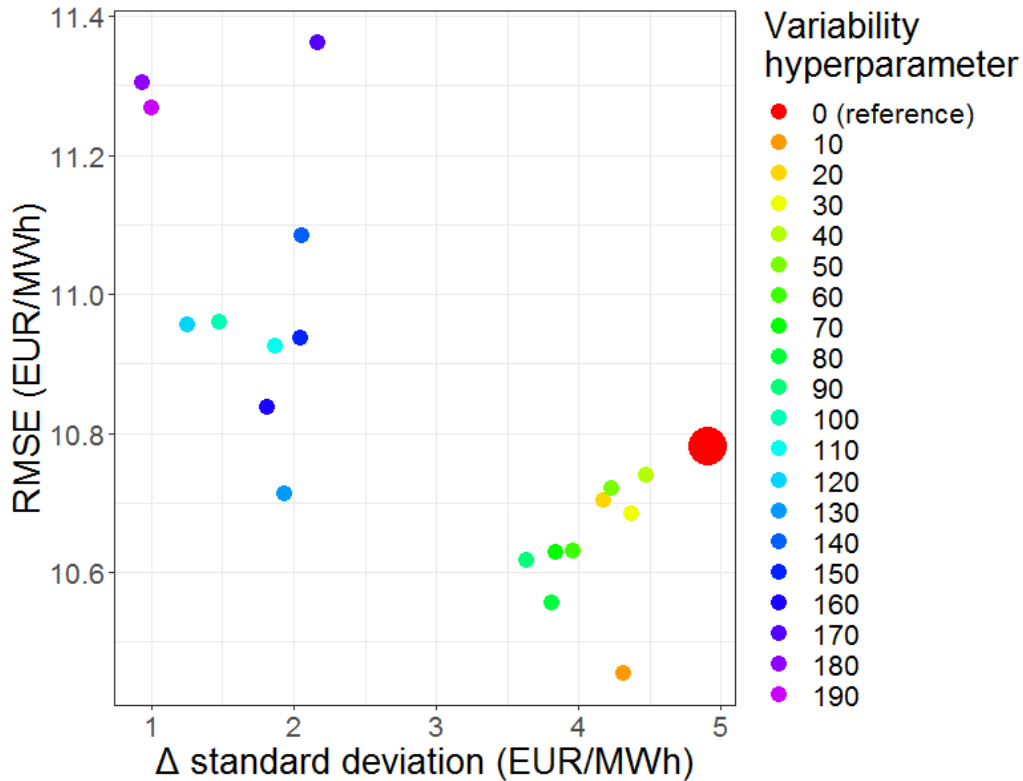


Figure 5: Out-of-sample performance for different values of the variability hyperparameter

We observe in Fig. 5 that hyperparameterization 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.

Data augmentation. Due to data availability issues, we applied the technique presented in subsection 2.7 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 Fig. 6, we represent the performance obtained when averaging the simulated prices for 2016 obtained using the models trained with different

synthetic data sets (e.g. the case "demand variation 1.0%" is obtained by averaging the outputs obtained with the models trained using the data sets -1.0%, -0.5%, 0.0%, 0.5% and 1.0%).

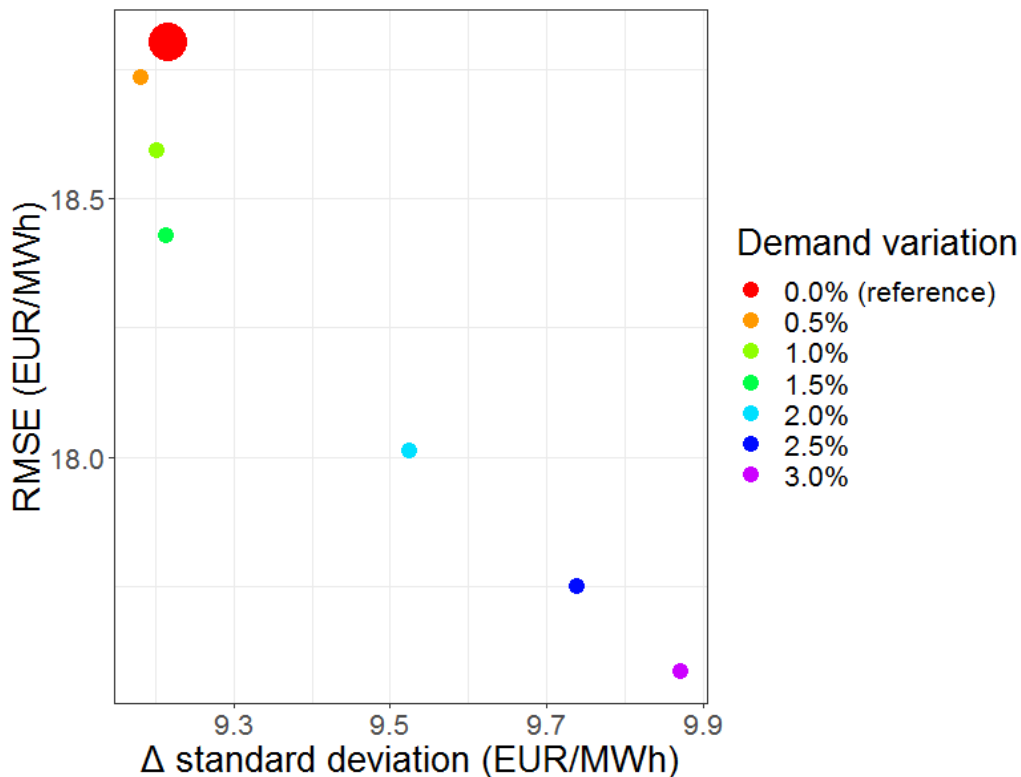


Figure 6: Out-of-sample performance for different augmented data sets (training set: 2015, testing set: 2016)

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.

3.3.3. 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 Fig. 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 data set 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 data set 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 the introduction of this paper, 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. 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. This additional electricity is currently sold at a fixed price and given total volume. The energy acquired thanks to ARENH can then be sold back on day-ahead markets to generate profits from this arbitrage or to minimize the financial losses due to oversupply. Sensitivity analyses based on the volume and price of ARENH and its indirect consequence on the day-ahead market can be conducted using structural models.

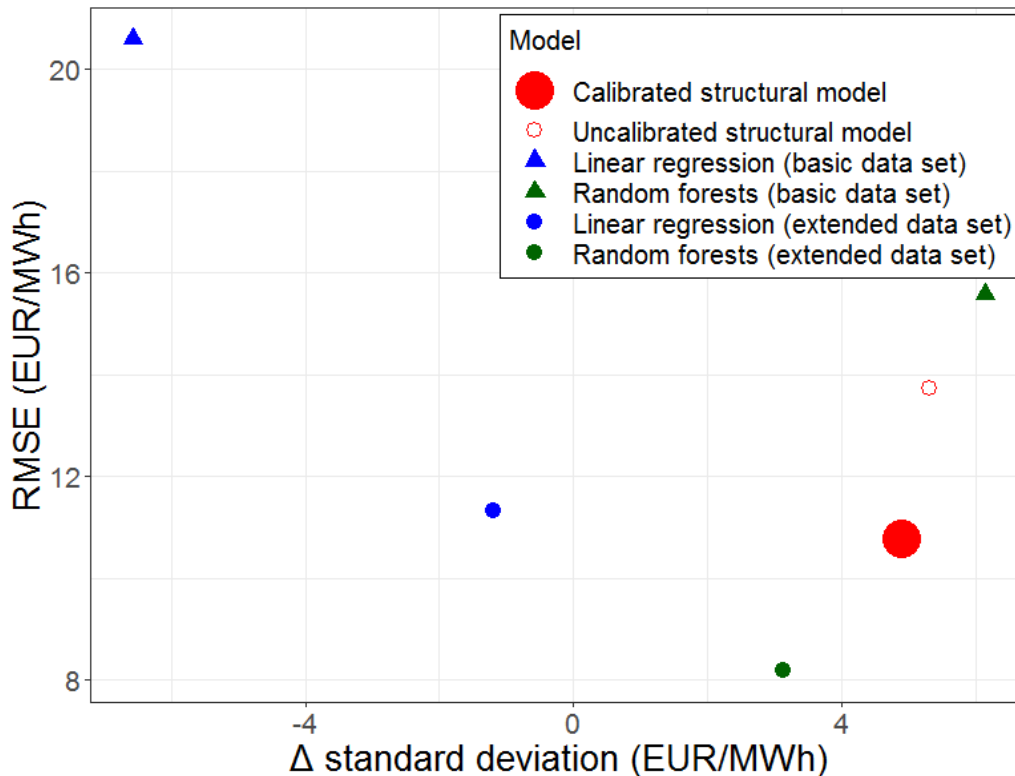


Figure 7: Performance comparison of benchmarked models

4. Sensitivity analysis and discussion

4.1. Sensitivity analysis of initial conditions

When we initialize the parameters $\alpha_{c,0}$ to a different value for each production class, we make assumptions on 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, since the marginal unit for each hour determined while solving the market clearing problem directly depends on this merit order. In this section, we perform a sensitivity analysis to investigate how alternative cost scenarios affect the initial merit order and as consequence 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 Fig. 8).

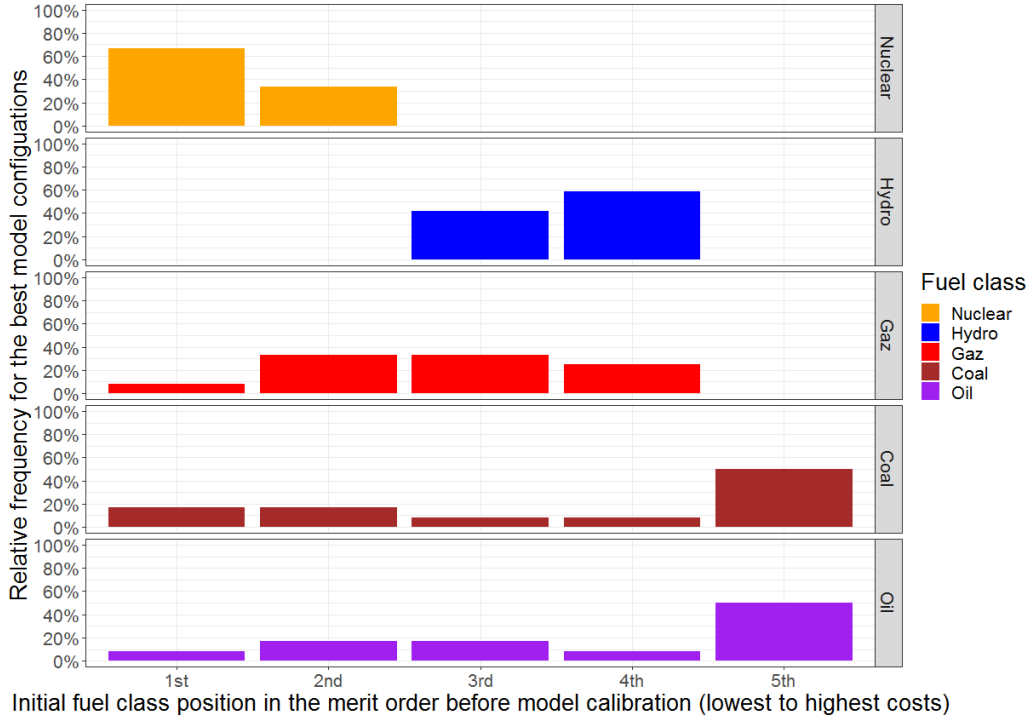


Figure 8: Best initial merit orders (first decile)

The results illustrate that the model initialization should be consistent (e.g. nuclear to constitute a less expensive baseload and oil for 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, since the initial merit order impacts the assessment of hourly marginal production classes, the initialization must be considered carefully.

4.2. 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 (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 Fig. 9.

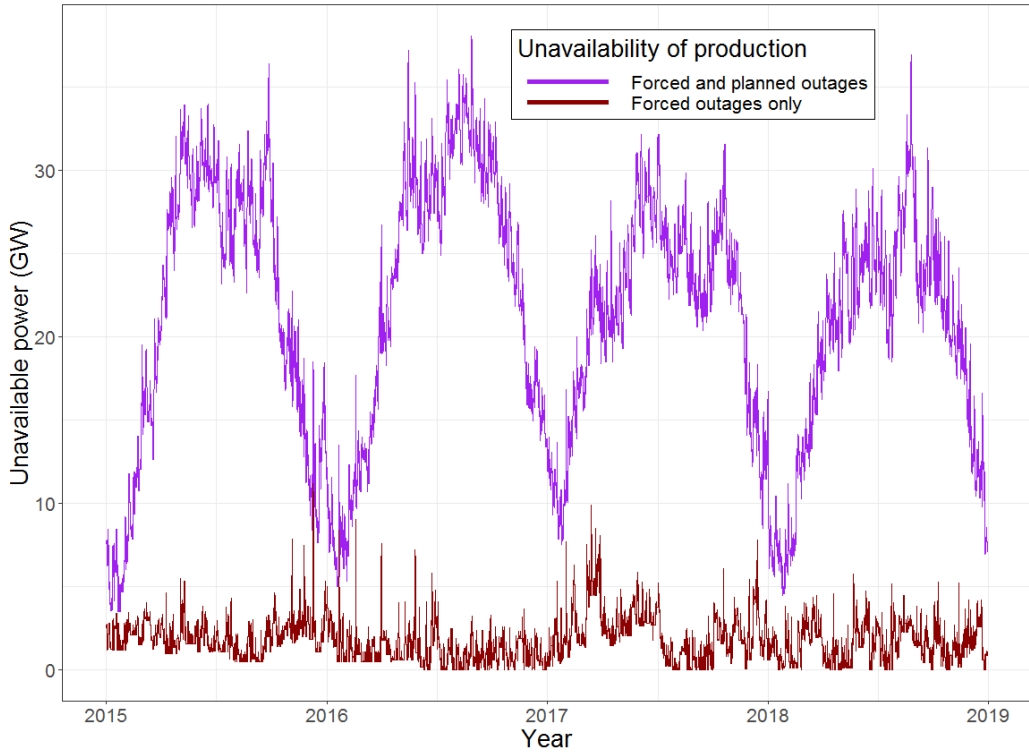


Figure 9: Unavailability of production units

We perform regressions between the prediction errors of simulated prices on the out-of-sample data set and our variables quantifying unexpected events using the Loess procedure [38] (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 Fig. 10.

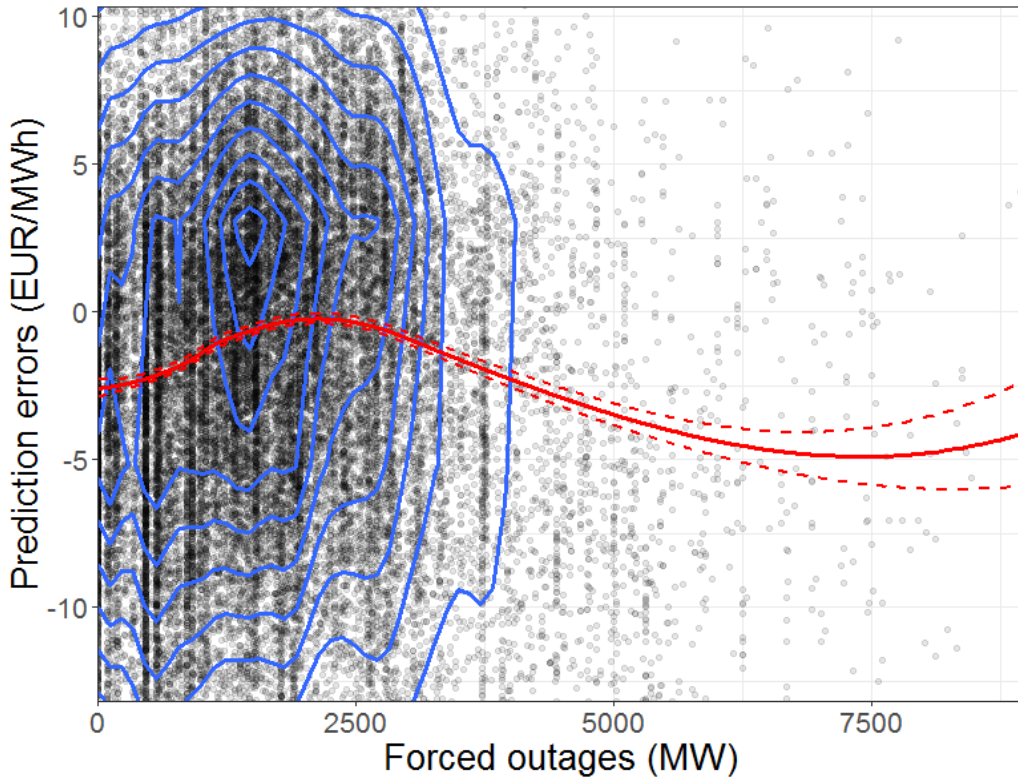


Figure 10: Impact of forced outages on prediction errors. Hourly values (black dots), their corresponding densities (blue lines) and locally estimated scatterplot smoothing, i.e. Loess (the estimation is shown by the solid red line and the 95% confidence interval by the dashed red lines)

4.3. 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 calibrate improved performance. Exploiting fundamental data proves interesting to simulate these prices over the long run. For prospective studies, scenarios could be generated to assess the potential impact of the evolution of fundamentals, such as the demand, energy

mix or fuel and emission costs. In addition, to account for cross-zonal effects, the price in neighboring countries could be included in scenarios, or multi-zone simulations could be performed. Finally, 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.

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 such 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 paper, we propose 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, emphasizing key market conditions and integrating additional price information from real supply curves data. 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. A sensitivity analysis revealed the importance of initially configuring the model with a realistic merit order to ensure that the model parameters are properly calibrated. For future work, the model can be extended by incorporating additional technical constraints, such as ramp rate limits of generation units, and by modelling additional market orders corresponding to the demand side flexibility. Finally, scenarios corresponding to the likely or desired evolution of the power system can be formulated and be used with our calibrated model thanks to Monte-Carlo simulations in order to estimate the possible consequences of the considered changes on wholesale electricity prices.

Acknowledgements

This work was partially supported by a PhD grant from the French agency for ecological transition (ADEME) and the Association pour la Recherche et le Développement des Méthodes et Processus Industriels (ARMINES). The authors would 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.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] S. Küfeoğlu, Economic Impacts of Electric Power Outages and Evaluation of Customer Interruption Costs, Ph.D. thesis, Aalto University, iSSN: 1799-4942 (electronic) (2015).
URL <https://aaltodoc.aalto.fi:443/handle/123456789/17867>
- [2] K. Mayer, S. Trück, Electricity markets around the world, *Journal of Commodity Markets* 9 (2018) 77–100. doi:10.1016/j.jcomm.2018.02.001.
URL <http://www.sciencedirect.com/science/article/pii/S2405851318300059>

- [3] T. W. Parkinson, Market Price Risks of Merchant Generation, *The Electricity Journal* 17 (4) (2004) 33–48. doi:10.1016/j.tej.2004.01.010.
URL <http://www.sciencedirect.com/science/article/pii/S1040619004000399>
- [4] S. J. Deng, S. S. Oren, Electricity derivatives and risk management, *Energy* 31 (6) (2006) 940–953. doi:10.1016/j.energy.2005.02.015.
URL <http://www.sciencedirect.com/science/article/pii/S0360544205000496>
- [5] R. Gross, W. Blyth, P. Heptonstall, Risks, revenues and investment in electricity generation: Why policy needs to look beyond costs, *Energy Economics* 32 (4) (2010) 796–804. doi:10.1016/j.eneco.2009.09.017.
URL <http://www.sciencedirect.com/science/article/pii/S0140988309001832>
- [6] R. Weron, Electricity price forecasting: A review of the state-of-the-art with a look into the future, *International Journal of Forecasting* 30 (4) (2014) 1030–1081. doi:10.1016/j.ijforecast.2014.08.008.
URL <http://www.sciencedirect.com/science/article/pii/S0169207014001083>
- [7] NEMO COMMITTEE, EUPHEMIA Public Description - Single Price Coupling Algorithm (Oct. 2020).
URL <http://www.nemo-committee.eu/assets/files/euphemia-public-description.pdf>
- [8] C. Alasseur, O. Féron, Structural price model for coupled electricity markets, *Energy Economics* 75 (2018) 104–119. doi:10.1016/j.eneco.2018.07.018.
URL <http://www.sciencedirect.com/science/article/pii/S0140988318302676>
- [9] L. Hirth, J. Mühlenpfordt, M. Bulkeley, The ENTSO-E Transparency Platform – A review of Europe’s most ambitious electricity data platform, *Applied Energy* 225 (2018) 1054–1067. doi:10.1016/j.apenergy.2018.04.048.
URL <http://www.sciencedirect.com/science/article/pii/S0306261918306068>

- [10] 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 (Jun. 2013). URL <http://data.europa.eu/eli/reg/2013/543/oj/eng>
- [11] F. Wiese, I. Schlecht, W.-D. Bunke, C. Gerbaulet, L. Hirth, M. Jahn, F. Kunz, C. Lorenz, J. Mühlenpfordt, J. Reimann, W.-P. Schill, Open Power System Data – Frictionless data for electricity system modelling, *Applied Energy* 236 (2019) 401–409. doi:10.1016/j.apenergy.2018.11.097. URL <http://www.sciencedirect.com/science/article/pii/S0306261918318130>
- [12] ESMA launches consultation on cost of market data and consolidated tape. URL <https://www.esma.europa.eu/press-news/esma-news/esma-launches-consultation-cost-market-data-and-consolidated-tape>
- [13] R. H. Baker, J. Król, Petition for Rulemaking Regarding Market Data Fees and Request for Guidance on Market Data Licensing Practices; Investor Access to Market Data (2018). URL <https://www.sec.gov/rules/petitions/2018/petn4-728.pdf>
- [14] J. Lago, F. De Ridder, B. De Schutter, Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms, *Applied Energy* 221 (2018) 386–405. doi:10.1016/j.apenergy.2018.02.069. URL <http://www.sciencedirect.com/science/article/pii/S030626191830196X>
- [15] P. 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 ?, Ph.D. thesis, Université Paris Dauphine, PhD Thesis (Jul. 2007).
- [16] H.-K. Ringkjøb, P. M. Haugan, I. M. Solbrekke, A review of modelling tools for energy and electricity systems with large shares of variable renewables, *Renewable and Sustainable Energy Reviews* 96 (2018) 440–459. doi:10.1016/j.rser.2018.08.002.

URL <http://www.sciencedirect.com/science/article/pii/S1364032118305690>

- [17] Joint Research Centre (European Commission), Systematic mapping of power system models, Tech. rep., Publications Office (2017).
URL <https://data.europa.eu/doi/10.2760/422399>
- [18] L. Mitridati, P. Pinson, A Bayesian Inference Approach to Unveil Supply Curves in Electricity Markets, *IEEE Transactions on Power Systems* 33 (3) (2018) 2610–2620. doi:10.1109/TPWRS.2017.2757980.
URL <http://ieeexplore.ieee.org/document/8053811/>
- [19] C. Ruiz, A. J. Conejo, D. J. Bertsimas, Revealing Rival Marginal Offer Prices Via Inverse Optimization, *IEEE Transactions on Power Systems* 28 (3) (2013) 3056–3064. doi:10.1109/TPWRS.2012.2234144.
URL <https://ieeexplore.ieee.org/abstract/document/6423235>
- [20] R. Chen, I. C. Paschalidis, M. C. Caramanis, P. Andrianesis, Learning From Past Bids to Participate Strategically in Day-Ahead Electricity Markets, *IEEE Transactions on Smart Grid* 10 (5) (2019) 5794–5806, conference Name: *IEEE Transactions on Smart Grid*. doi:10.1109/TSG.2019.2891747.
- [21] R. Carmona, M. Coulon, D. Schwarz, Electricity price modeling and asset valuation: a multi-fuel structural approach, *Mathematics and Financial Economics* 7 (2) (2013) 167–202. doi:10.1007/s11579-012-0091-4.
URL <http://link.springer.com/10.1007/s11579-012-0091-4>
- [22] L. Hirth, What caused the drop in European electricity prices? A factor decomposition analysis, *The Energy Journal* 39 (1). doi:10.5547/01956574.39.1.lhir.
URL <http://www.iaee.org/en/publications/ejarticle.aspx?id=3031>
- [23] P. Beran, C. Pape, C. Weber, Modelling German electricity wholesale spot prices with a parsimonious fundamental model – Validation & application, *Utilities Policy* 58 (2019) 27–39. doi:10.1016/j.jup.2019.01.008.

URL <http://www.sciencedirect.com/science/article/pii/S0957178719300359>

- [24] R. Carmona, M. Coulon, A Survey of Commodity Markets and Structural Models for Electricity Prices, in: F. E. Benth, V. A. Kholodnyi, P. Laurence (Eds.), *Quantitative Energy Finance*, Springer New York, New York, NY, 2014, pp. 41–83. doi:10.1007/978-1-4614-7248-3_2. URL http://link.springer.com/10.1007/978-1-4614-7248-3_2
- [25] Y. He, M. Hildmann, F. Herzog, G. Andersson, Modeling the Merit Order Curve of the European Energy Exchange Power Market in Germany, *IEEE Transactions on Power Systems* 28 (3) (2013) 3155–3164. doi:10.1109/TPWRS.2013.2242497.
- [26] F. Ziel, R. Steinert, Electricity price forecasting using sale and purchase curves: The X-Model, *Energy Economics* 59 (2016) 435–454. doi:10.1016/j.eneco.2016.08.008. URL <http://www.sciencedirect.com/science/article/pii/S0140988316302080>
- [27] F. Ziel, R. Steinert, Probabilistic mid- and long-term electricity price forecasting, *Renewable and Sustainable Energy Reviews* 94 (2018) 251–266. doi:10.1016/j.rser.2018.05.038. URL <http://www.sciencedirect.com/science/article/pii/S1364032118303885>
- [28] S. Kulakov, X-Model: Further Development and Possible Modifications, *Forecasting* 2 (1) (2020) 20–35. doi:10.3390/forecast2010002. URL <https://www.mdpi.com/2571-9394/2/1/2>
- [29] R. Aïd, L. Campi, N. Langrené, A Structural Risk-Neutral Model for Pricing and Hedging Power Derivatives, *Mathematical Finance* 23 (3) (2013) 387–438. doi:10.1111/j.1467-9965.2011.00507.x. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9965.2011.00507.x>
- [30] A. A. Sánchez de la Nieta, J. Contreras, Quantifying the effect of renewable generation on day-ahead electricity market prices: The Spanish case, *Energy Economics* 90 (2020) 104841. doi:10.1016/j.eneco.2020.104841.

- URL <https://www.sciencedirect.com/science/article/pii/S014098832030181X>
- [31] K. R. Ward, R. Green, I. Staffell, Getting prices right in structural electricity market models, *Energy Policy* 129 (2019) 1190–1206. doi:10.1016/j.enpol.2019.01.077.
URL <http://www.sciencedirect.com/science/article/pii/S0301421519300564>
- [32] C. Pape, S. Hagemann, C. Weber, Are fundamentals enough? Explaining price variations in the German day-ahead and intraday power market, *Energy Economics* 54 (2016) 376–387. doi:10.1016/j.eneco.2015.12.013.
URL <http://www.sciencedirect.com/science/article/pii/S0140988316000037>
- [33] V. Mahler, R. Girard, S. Billeau, G. Kariniotakis, Simulation of day-ahead electricity market prices using a statistically calibrated structural model, in: 2019 16th International Conference on the European Energy Market (EEM), 2019, pp. 1–5, iSSN: 2165-4093. doi:10.1109/EEM.2019.8916273.
- [34] X. Xia, A. M. Elaiw, Optimal dynamic economic dispatch of generation: A review, *Electric Power Systems Research* 80 (8) (2010) 975–986. doi:10.1016/j.epsr.2009.12.012.
URL <http://www.sciencedirect.com/science/article/pii/S0378779610000027>
- [35] S. Y. Abujarad, M. W. Mustafa, J. J. Jamian, Recent approaches of unit commitment in the presence of intermittent renewable energy resources: A review, *Renewable and Sustainable Energy Reviews* 70 (2017) 215–223. doi:10.1016/j.rser.2016.11.246.
URL <http://www.sciencedirect.com/science/article/pii/S1364032116310140>
- [36] Open Power System Data, Data Package Time series, . Version 2019-06-05. doi:10.25832/time_series/2019-06-05.
- [37] RTE, ADEME, Signal prix du CO₂, Analyse de son impact sur le système électrique européen (Mar. 2016).

URL https://presse.ademe.fr/wp-content/uploads/2016/04/Signal-prix-du-CO2_finale_29-03-16.pdf

- [38] W. S. Cleveland, E. Grosse, W. Shyu, Chapter 8 - Local regression models, in: Statistical Models in S, Chapman and Hall/CRC, 1992, p. 624.