Working paper

2022-02

Statistics and Econometrics
ISSN 2387-0303

Optimal day-ahead offering strategy for large producers based on market price response learning

Antonio Alcántara, Carlos Ruiz

uc3m | Universidad Carlos III de Madrid

Optimal day-ahead offering strategy for large producers based on market price response learning

Antonio Alcántara* Carlos Ruiz†

April 22, 2022

Abstract

In day-ahead electricity markets based on uniform marginal pricing, small variations in the offering and bidding curves may substantially modify the resulting market outcomes. In this work, we deal with the problem of finding the optimal offering curve for a risk-averse profit-maximizing generating company (GENCO) in a data-driven context. In particular, a large GENCO's market share may imply that her offering strategy can alter the marginal price formation, which can be used to increase profit. We tackle this problem from a novel perspective. First, we propose a optimization-based methodology to summarize each GENCO's step-wise supply curves into a subset of representative price-energy blocks. Then, the relationship between the market price and the resulting energy block offering prices is modeled through a Bayesian linear regression approach, which also allows us to generate stochastic scenarios for the sensibility of the market towards the GENCO strategy, represented by the regression coefficient probabilistic distributions. Finally, this predictive model is embedded in the stochastic optimization model by employing a constraint learning approach. Results show how allowing the GENCO to deviate from her true marginal costs renders significant changes in her profits and the market marginal price. Furthermore, these results have also been tested in an out-of-sample validation setting, showing how this optimal offering strategy is also effective in a real-world market contest.

Keywords: Stochastic programming, Constraint learning, Data-driven optimization, Electricity market, Optimal pricing strategy

*Department of Statistics, University Carlos III of Madrid, Avda. de la Universidad, 30, 28911 Leganés, Spain. Email: antalcan@est-econ.uc3m.es

†Department of Statistics and UC3M-BS Institute for Financial Big Data (IFiBiD), University Carlos III of Madrid, Avda. de la Universidad, 30, 28911 Leganés, Spain. Email: caruizm@est-econ.uc3m.es

1 Introduction

As digitization and automation processes advance in the so-called fourth technological revolution, the treatment and use of large amounts of data for decision-making is of great importance. From pandemic management to financial investment or electricity consumption planning, data play a key role to make informed and optimal decisions, and in most cases, under high levels of uncertainty. Indeed, this data-driven perspective is inspiring the development of state-of-art optimization modeling techniques and efficient solution algorithms.

The assumption to know the input parameters with complete certainty has been the fundamental hypothesis for deterministic optimization techniques tackling complex decision-making problems, based on linear, nonlinear, integer formulations, or a combination of these (Murty, 1994). However, this assumption is barely fulfilled in real contexts and hence, uncertainty must be considered within the optimization process. One of the most employed approaches in this regard is stochastic programming, where the model incorporates the estimated probability distribution of the uncertain parameters (Birge & Louveaux, 2011). In particular, stochastic programming considers the following problem:

$$\min_{x \in \mathcal{X}} \mathbb{E}\left[c(x; Y)\right] \tag{1}$$

where $x \in \mathcal{X} \subset \mathbb{R}^{d_x}$ represents the decision variables, $Y \in \mathcal{Y} \subset \mathbb{R}^{d_y}$ are the parameters that characterize the problem, $c(x;Y) : \mathbb{R}^{d_x} \times \mathbb{R}^{d_y} \to \mathbb{R}$ is the cost function, and $\mathbb{E}\left[\cdot\right]$ represents the expected value over the distribution of Y.

We want to further extend this setting and exploit the case where auxiliary information (covariates $\theta \in \Phi \subset \mathbb{R}^{d_\theta}$) is available to help modelling the complex response of Y. Thus, we try to set a decision problem in a data-driven context. In particular, lets assume again that x is our decision variable, while Y is a response from a complex system (e.g., market price) that conditions our objective function $c(x; Y|f, \theta)$. Notice how the function f is now used to relate the response Y with the contextual information θ. Furthermore, lets assume that we have access to historical observations of the type $\mathcal{D} = \{(x_1, y_1, \theta_1), \ldots, (x_N, y_N, \theta_N)\}$ where each of the N samples includes the decision variable x, the response y, and the covariates θ.

The problem to be treated can be generalized as:

$$x\left(f, \theta\right) \in \operatorname*{argmin}_{x \in \mathcal{X}} \mathbb{E}[c(x; Y|f, \theta)] \tag{2}$$

Different perspectives have tried to tackle the use of contextual information for decision-making in an optimal way. One of the most common approaches in the literature is to follow a Predict and Optimize strategy. That is, we learn the relationship between Y and θ through a predictive model f by employing a dataset $\mathcal{D}$ of past observations. Then, when a new value θ is given, $Y = f(\theta)$ is computed and used within the optimization problem to set optimal decisions x. However, this strategy has some key drawbacks: the use of the point prediction fails to capture the associated uncertainty level, and the function f is not aware of the optimization model's behavior.

For this second issue, an integrated approach to find functions f that also lead to good prescriptions is addressed in Elmachtoub and Grigas (2022), but only for linear objective and prediction functions. In Ban and Rudin (2019), the newsvendor problem is tackled by using machine learning models to predict optimal decisions as a direct function of the observed θ. One disadvantage of the proposed strategy is to reach potentially infeasible decisions in a test dataset.

Recently, Bertsimas and Kallus (2020) introduced the so-called Predictive to Prescriptive two-step approach, where the first step focus on training machine learning models to predict Y from a given θ. Nevertheless, in the second step, a Sample Average Approximation (SAA) is solved with the weights dictated by the prediction model for that particular observation. For instance, using kNN as the prediction function f, for any θ, k nearest neighbors are computed in the training set, and a SSA is solved only with these k neighbors to find the optimal decisions $x(\theta)$. An in-deep review of these approaches can be found in Mundru (2019).

Furthermore, in Muñoz et al. (2022) a bi-level framework is proposed to fit a parametric model to those data that are specifically tailored to maximize the decision value, while accounting for possible feasibility constraints. In Bertsimas et al. (2019) and Esteban-Pérez and Morales (2021), parametric approaches are left behind to focus on estimating a complete conditional distribution of the side information to make robust decisions over x. Whereas in Bertsimas et al. (2019), the approach is conceived as a two-step procedure, in Esteban-Pérez and Morales (2021), a single-step method is derived.

However, even if the above works deal with a data-driven approach for decision making, there is one specific setting that needs to be specially addressed: when decisions x have a direct influence on the response Y, which also conditions the cost function (and hence, can be treated as another decision variable), jointly with the rest of contextual information and associated uncertainty. Thus, we want to embed these types of interactions, related to an extended predictive model $y = f^{\mathcal{D}}(x, \theta)$, within our optimization problem to capture the relationship between a complex response, our optimal decisions, and the contextual information. This process is referred to as Constraint Learning (Fajemisin et al., 2021), a topic that has recently gained attention in the literature. Some works have studied how to embed linearizable machine learning models within the optimization problem. For instance, in Paulus et al. (2021) and Yang and Bequette (2021), neural networks are employed to learn the constraints, while in Maragno et al. (2021) and Mišić (2020) tree-based methods were studied.

Nevertheless, despite these "black-box" methods give a high-quality performance regarding prediction accuracy, they lack the sought of explainability that a simpler method, like a classical linear regression, can provide in real applications. Besides, a proper uncertainty characterization around the point prediction is not considered. For these reasons, we will extend this setting to explicitly account for uncertainty and risk aversion in the decision-making process.

In particular, the work by Pérez-Santalla et al. (2022) addresses the uncertainty and explainability under a constraint learning approach, but under a stylized and simulation-based application. However, we want to consider a fully data-driven context. That is, dealing with a real-world application with large amounts of data, employing a Bayesian linear regression approach to model uncertainty around coefficient distributions while trading-off explainability and prediction power, and performing an out-of-sample validation of the stochastic optimal solutions.

Particularly, considering a two-stage stochastic programming framework, we address the following model:

$$\min_{x,y\in\mathcal{S}(\theta;\xi)} \mathbb{E}\left[c(x,y|\theta;\xi)\right] \tag{3}$$

where, $c(\cdot|\theta;\xi):\mathbb{R}^{d_x}\times\mathbb{R}^{d_y}\times\mathbb{R}^{d_\theta}\times\mathbb{R}^{d_\xi}\to\mathbb{R}$, $x\in\mathbb{R}^{d_x}$ can be considered first stage and $y\in\mathbb{R}^{d_y}$ second stage decision variables, $\theta\in\mathbb{R}^{d_\theta}$ accounts for contextual information known when the first stage decisions take place, and $\xi\in\mathbb{R}^{d_\xi}$ gathers the exogenous uncertain data (random vector) characterizing our problem. $\mathbb{E}[\cdot]$ may stand for expectation, but we can also consider any other risk measure, e.g., we focus on Conditional Value-at-Risk (CVaR). Without loss of generality, we assume that the feasible region for x and y, i.e., $\mathcal{S}(\theta;\xi)$ depends on the contextual information and on the uncertain data as follows:

$$\mathcal{S}(\theta;\xi)=\begin{cases} g(x,y,\theta;\xi)\le 0\\ y=f^{\mathcal{P}}(x,\theta;\xi)\end{cases} \tag{4}$$

Where $g(\cdot):\mathbb{R}^{d_x}\times\mathbb{R}^{d_y}\times\mathbb{R}^{d_\theta}\times\mathbb{R}^{d_\xi}\to\mathbb{R}^{g}$ is a constraint mapping. In particular, $f^{\mathcal{P}}(x,\theta;\xi):\mathbb{R}^{d_x}\times\mathbb{R}^{d_\theta}\times\mathbb{R}^{d_\xi}\to\mathbb{R}^{d_y}$ represents the predictive model that links the first and the second stage decision variables. Note that it depends on the realization of the contextual information, together with the uncertainty associated with its training. Indeed, we are interested in predictive models $f^{\mathcal{P}}(\cdot)$ with accurate probabilistic characterization of their parameters.

To tackle (3) in practice, we can employ a SAA-based approach:

$$\min_{x,y_\omega} \sum_{\omega=1}^{\Omega} \pi_\omega c_\omega(x,y_\omega|\theta;\xi_\omega) \tag{5a}$$

s.t.

$$g_\omega(x,y_\omega,\theta;\xi_\omega)\le 0 \quad \forall\omega \tag{5b}$$

$$y_\omega=\hat{f}^{\mathcal{P}}_\omega(x,\theta;\xi_\omega) \quad \forall\omega \tag{5c}$$

where uncertainty is approximated by a discrete set of scenarios $\omega=1,\ldots,\Omega$ with an assigned probability π_ω (we may also account for the risk-averse case via risk-adjusted probabilities). Each scenario represents a possible realization of the uncertain vector ξ, i.e., ξ_ω. Similarly the second stage decision variables y can be considered scenario dependent, i.e., y_ω. Moreover, the estimation of the predictive function, is also scenario dependent, i.e., $\hat{f}^{\mathcal{P}}_\omega(\cdot)$, as it is conditioned by the particular realization of ξ_ω which affects its own parameters.

The assumption of y_ω being a second stage decision variable dependent on first stage decisions x, external factors and uncertain data is suitable for markets where participants may have some degree of market power. For instance, in electricity markets based on uniform marginal pricing, suppliers and consumers submit their offers and bids (first stage decisions) to the market operator. The intersection of the aggregated offering and bidding curves will generate an hourly marginal price for the market (which can be considered a second-stage decision for a player with sufficient market power). As we can see, decisions made at a first instance by the market agents have a direct impact on the marginal price formation and, therefore, on the profit they can obtain.

In this work, we focus on this setting and study the optimal offering strategy from the perspective of a risk-averse large producer (or generating company, GENCO) participating in a day-ahead electricity market. We address this problem from a completely data-driven approach where large amounts of data are available from historical supply and purchase offers, and the resulting market outcomes. For that purpose, first, we propose a novel optimization-based technique to get an adequate subset of blocks summarizing the step-wise offer curves of each GENCO. Then, a statistical prediction model is employed to set the relationship between supply offering prices and the resulting marginal market price. This statistical model will allow generating meaningful scenarios for the marginal price response concerning the GENCO's offers. Finally, the resulting model will be embedded in the decision-making process, which will result in a computationally efficient stochastic optimization model. Optimal offering prices will be tested through an out-of-sample methodology.

In the case of power producers, the problem of setting an optimal offering strategy has been extensively discussed in the literature. Multilevel models have been used as a standard approach to set optimal strategies that maximize GENCOS' profits. For instance, in Ruiz and Conejo (2009), the solution for the problem is based on a bi-level equilibrium approach. Kardakos et al. (2015) deals with the problem of setting an offering strategy for a virtual problem plant with a stochastic bi-level approach, whereas in Pandžić et al. (2013), a two-stage stochastic approach was employed to set the optimal strategy of a virtual power plant selling and purchasing energy from the day-ahead and the balancing markets seeking to maximize its profit.

More recently, Xiao et al. (2021) tackles the problem with a single-level mixed-integer linear programming approach, where Conditional Value at Risk (CVaR) is used for risk management and ARIMA models to generate scenarios in the optimization problem. In Han et al. (2018), the optimal strategy for a photovoltaic power plant is set by a bi-level stochastic program, dealing with the uncertainty of the competitors and its photovoltaic output. Finally, in Chen et al. (2019), an Extreme Learning Machine is employed to find the relationship between the prosumer strategy and the obtained profits and costs in distribution grids. This is, indeed, an example of applying constraint learning in an energy market context. However, these approaches do not consider the data-driven uncertainty, especially that inherent within the parameters of the forecasting model. Moreover, to the authors knowledge, none of these approaches have been tested under out-of-sample validation schemes like the one presented in this work.

In summary, the main contributions of this work concerning the state-of-art on these topics are the following:

– to develop a complete data-driven optimization model for a risk averse GENCO offering strategy, making use of massive real world datasets.

– to propose an optimization-based reduction technique to summarize past realizations of the market participants hourly offering curves.

– to extend the standard constraint learning methodology by including a Bayesian linear regression approach to take the inherent predicted model uncertainty into account.

– to show the validity of the obtained optimal pricing strategy in a real world out-of-sample application based on the Spanish electricity market.

The structure of this article is as follows. Section 2 will describe in detail the proposed methodology, from the optimal discretization of supply curves to the marginal price formation and the Bayesian regression employed to characterize the relationship between the marginal price and the GENCO's strategy. Section 3 will introduce the stochastic optimization problem for the GENCO. Then, a case study will be shown in Section 4, including out-of-sample testing of the optimal strategy. Finally, Section 5 will draw the main conclusions of this work.

2 Marginal price and supply curves characterization

As it has been mentioned before, in many electricity markets the hourly marginal price is formed by the intersection of the demand and supply curves bidded and offered by the consumers and generators, respectively. In this work, we aim to characterize the optimal offering strategy of a profit-maximizing large generating company (GENCO). In particular, we will study, from a data-driven perspective, the potential ability of the GENCO to alter the formation of market-clearing hourly prices. For that reason, we need to deeply analyze and characterize her historical offering curves and the resulting market outcomes.

2.1 Optimal discretization of supply curves

In general, in day ahead electricity markets based on uniform marginal pricing, data from the hourly supply curve of each GENCO is composed of production units' power blocks with their corresponding offering prices, i.e., prices that each unit is willing to accept to produce that amount of power for one hour. In a fair, transparent, and audited market, this price must reflect the marginal generating cost of that production unit.

To build the GENCOS' aggregated supply curve for each hour, production blocks are ordered by their price, from the lowest to the largest. Regarding the energy quantity, a cumulative sum is performed so that each price faces the cumulative quantity of energy of those blocks with prices below. This renders an increasing step-wise curve. An illustrative example can be seen in Figure 1, where for a given hour and particular GENCO, red points represent her production units with their price and cumulative quantity within the supply curve.

For tractability of the subsequent predictive model, we propose to summarize first each aggregated supply curve into a smaller number of blocks by using an optimization approach (6). The aim is to obtain a minimum-error-based discretization of the GENCO's hourly supply curve.

$$\min_{C_b,\delta_b} \quad \sum_{b=1}^{B} |C_b - P_b^R| q_b^R \tag{6a}$$

s.t.:

$$0 \le C_b - C_{b-1} \le \delta_n M^P \quad b = 2, \ldots, B \tag{6b}$$

$$0 \le C_b - C_0 \le \delta_1 M^P \quad b = 1 \tag{6c}$$

$$\sum_{b=1}^{B} \delta_b = |I| - 1 \tag{6d}$$

$$\delta_b \in \{0, 1\} \quad b = 1, \ldots, B \tag{6e}$$

In this problem, we seek to obtain $|I|$ grouped power blocks from a total of B original ones. C_b is the energy price, P_b^R is the real price observed, and q_b^R is the amount of energy of each block b. The objective function of the model (6a) aims to minimize the absolute distance between the real prices P_b^R and the optimized ones C_b, weighted by energy quantity of the block. Constraints (6b) and (6c) allow to assign the same price C_b to all the values that belong to the same grouped block. The sum of δ_b in constraint (6d) ensures that $|I| - 1$ cuts are obtained, and therefore $|I|$ grouped blocks, since δ_b is defined as a binary variable (6e).

The solution to this problem will let us know the different prices C_b and positions δ_b where a step in the curve can be set to obtain an optimal cut and therefore, an optimal grouped block. As an example, let's assume we want to get $|I| = 7$ grouped blocks from the curve (red dots) presented in Figure 1. The optimization problem will render the solid black step-wise curve as a solution. In this way, dashed vertical lines represent the different cuts, and solid horizontal lines the energy price C_b associated with each grouped block. Notice that, for simplicity, an extra vertical line has been added in the last production unit to set the last block.

As it can be seen, this solution will appropriately summarize the information within the original supply curve, as it results in a finer discretization for those parts of the curve with higher price variations.

2.2 Marginal price formation from discretized curves

Once we obtain the optimal discretization of the GENCO supply curve, the same procedure can be applied to the rest of the offers in the market, that is, the supply curves from the competitors. Considering both the GENCO's and competitors' supply curves, and ordering all the blocks by their price, the cumulative energy quantity can be computed to build the aggregated discretized market supply curve.

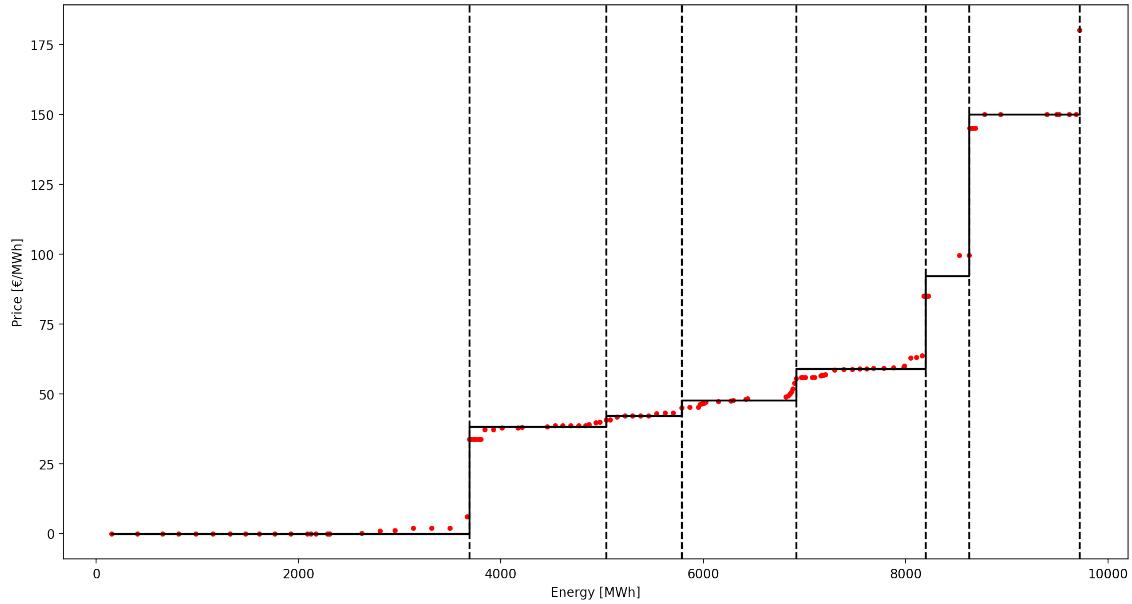


Figure 1: Supply curve discretization example for June 3rd 2017 at 15:00. Source: OMIE

Note that it is also possible to achieve a summarized curve by directly discretizing the original aggregated supply curve. However, splitting GENCO's units and the ones from the competitors will allow us to know the weight of the producer within the market and to get insights on how its pricing strategy can modify the market outcomes.

Nevertheless, to set a marginal price a demand curve is also needed. For the sake of simplicity, we will assume an inelastic demand curve, which will cut the supply curve and will establish the price producers will be paid for each MWh of energy offered below this price.

An example of this procedure is shown in Figure 2, where the increasing step-wise curve represents the aggregated market supply curve, the vertical line is the inelastic demand, and the green and red colors identify the blocks offered by the GENCO (main producer) and the rest of the competitors, respectively. We can see a total dispatched energy of 24000 MWh at a price of 42 €/MWh. Notice how the marginal price is set by a block from the main producer in this case. This motivates the study of how small variations in the offered price can directly change the marginal market price.

However, in real day-ahead electricity markets, the price resulting from the cut of the submitted demand and supply curves is not necessary the resulting market marginal price (e.g., Spain and Portugal). In fact, it is common to observe that the final supply curve from the market suffers changes (withdraw of some operation units) due to the system operator's ex-pot verification of technical constraints. We can take as an example the market curves presented in Figure 3. There, two different hourly marginal price formation procedures are shown within the same day in the Iberian Electricity Market (MIBEL, 2022). Differences between sale offers and matching (dispatched) sale offers can be easily

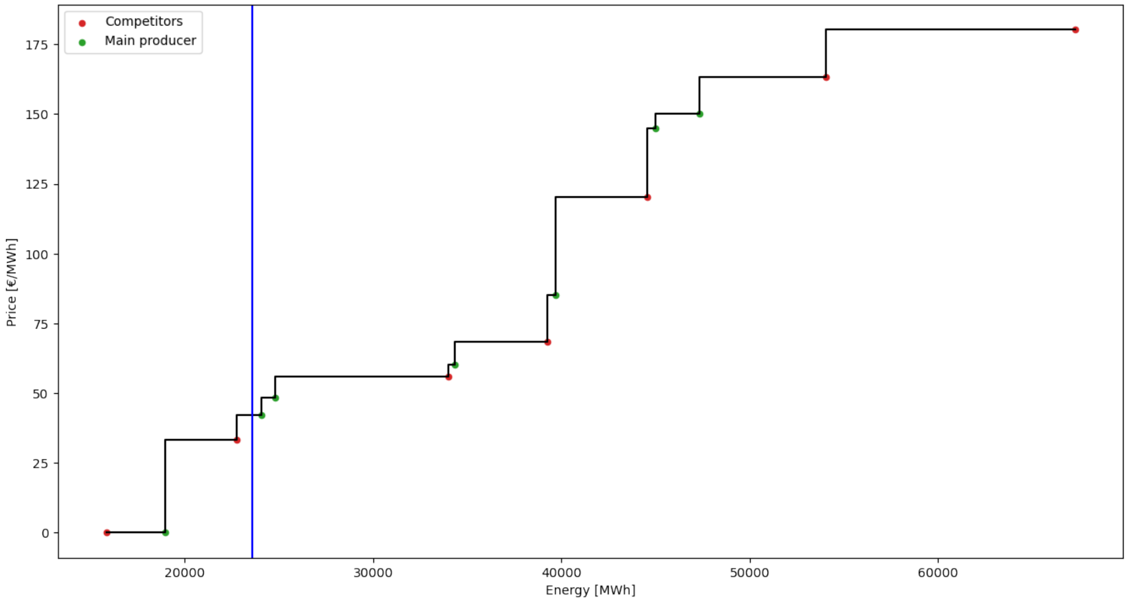


Figure 2: Example of marginal price formation on the discretized market

noticed at hour 15.

Hence, when analyzing historical market data, it is important to consider this fact by statistically characterizing the differences between the matching supply curve and the offered one. This will allow us to validate the optimal strategy in real life (out-of-sample). That is, we can set the optimal blocks with real data and use competitors' blocks, jointly with the demand and the approximated displacement of the supply curve to study the differences in profit according to the producer risk aversion and other relevant features.

2.3 Marginal price characterization model

To strategically derive the GENCO's optimal offered prices within her supply curve (one for each block) is necessary to find a function accurate enough to predict the response of the marginal electricity price. That is, assuming that the strategic GENCO can alter the market marginal price, we will characterize it as a function of its price block offers and other external covariates, such as different marginal price lags, the expected demand, or levels of renewable energy production.

Lets assume we have a dataset of N observations of the type $\mathcal{D} = \{(y_1, \mathbf{X}_1, \boldsymbol{\theta}_1), \ldots, (y_N, \mathbf{X}_N, \boldsymbol{\theta}_N)\}$, where $y \in \mathbb{R}$ represents the marginal price, $\mathbf{X} = \{X^1, \ldots, X^{d_x}\}$ are the block price offers done by the GENCO, and $\boldsymbol{\theta} = \{\theta^1, \ldots, \theta^{d_\theta}\}$ are the values for the rest of auxiliary variables (covariates) the marginal price depends on, known at the time of the decision making.

In this way, it is possible to characterize the marginal price y as a linear function of $\mathbf{X}$ and $\boldsymbol{\theta}$, as shown in (7).

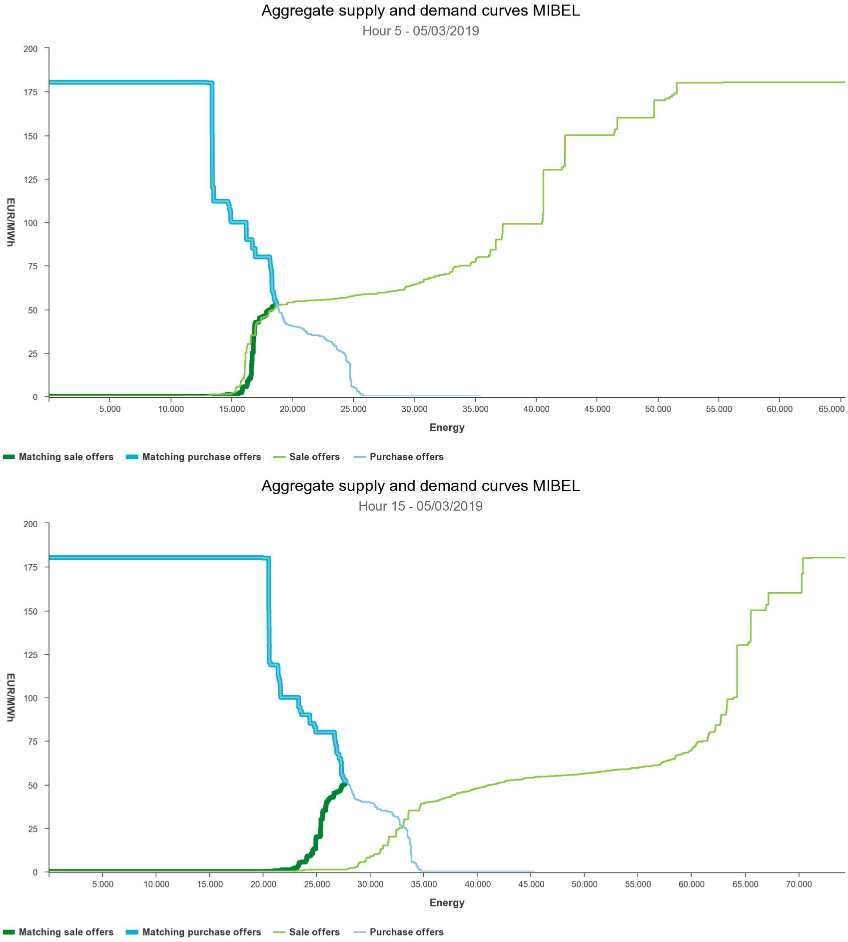


Figure 3: Marginal price formation in two different hours in the MIBEL

$$y = \beta_0 + \sum_{i=1}^{I} \beta_i X^i + \sum_{j=I+1}^{I+J} \beta_j \theta^j \tag{7}$$

where the different linear regression coefficients within $\boldsymbol{\beta} = \{\beta_1, \ldots, \beta_I\}$ can be estimated from $\mathcal{D}$ through the classical Ordinary Least Square approach. This approach is simple and it would allow us to characterize the uncertainty over the coefficients by assuming their distribution from the Hajek-Sidak CLT. However, in a data-driven context it seems more convenient to directly fit the distribution of the coefficients using existing empirical methods in the literature. For that reason, a Bayesian approach can be used to model the marginal price through Bayesian linear regression. In this way:

$$y \sim N(\boldsymbol{\beta}(\mathbf{X}, \boldsymbol{\theta}), \boldsymbol{\sigma} I)$$

Estimation of regression model coefficient distribution with the Bayesian approach can be done by iterating at a marginal posterior. A posterior distribution is calculated by multiplying the prior distribution and the likelihood function. In particular, the process of obtaining these parameters is done by using Markov Chain Monte Carlo (MCMC) algorithm. One of the commonly used algorithms in MCMC is Gibbs Sampling, and several packages have been developed to automatize this process, such as the *PyMC3* library in Python (Salvatier et al., 2016) or *rstanarm* package in R (Goodrich et al., 2020).

2.4 Scenario generation through coefficient distribution

Regarding the stochastic optimization problem, we seek to model the uncertainty associated with the coefficients of the regression model. In this case, we acknowledge that the marginal price sensitivity β_i to the block prices offered by the GENCO is not deterministic, but characterized by an associated probability distribution.

This fact means that, for example, there will be some scenarios where the price of one block will make the marginal price increase more than in others. It is even possible that those β_is with a high uncertainty level can be either positive or negative within the scenarios.

In this context, the data-driven Bayesian linear regression approach exposed in Section 2.3 will allow us to estimate the parameters of a normal distribution for each of the regression coefficients in the linear model. Focusing on the coefficients related to block prices, each one of the I of them will follow a normal distribution:

$$\hat{\beta}_i \sim N(\mathbb{E}[\hat{\beta}_i], \hat{\sigma}_{\beta_i}), \quad \forall i \in I$$

Random realizations from these distributions can be sampled to generate scenarios for the stochastic optimization model. For the rest of the covariates, a point estimate of the coefficients can be assumed, and their expected value will be directly embedded in the optimization problem.

3 Stochastic Optimization Problem

In this section, we describe the optimization problem that is employed by the GENCO to derive her optimal offering strategy. The problem is formulated using a risk-averse two-stage stochastic approach in which the marginal price of the market is explicitly characterized through constraint learning.

3.1 Notation

The notation employed to formulate the stochastic problem is described in this subsection for quick reference.

Indices and sets:

- I: Set of energy blocks, indexed by i.

– T: Set of hourly periods within a day, indexed by t.

– Ω: Set of stochastic scenarios, indexed by ω.

Variables:

– P_t^i: Price offered for block i at time t.

– $Q_{t,\omega}^i$: Quantity of energy produced from block i at time t and scenario ω.

– $u_{t,\omega}^i$: Binary decision variable indicating whether price from block i at time t and scenario ω is below the marginal electricity price or not.

– s_ω, η: Auxiliary variables for CVaR formulation.

– $\lambda_{t,\omega}$: Marginal market price at time t and scenario ω.

– Q_t^{ren} : Offered quantity at price zero from renewable resources at time t.

Parameters:

– $\beta_{t,\omega}^i$: Regression coefficient of the i-th block price predictor at time t and scenario ω, sampled from their corresponding distribution.

– $\hat{D}_t$: Sum of regression coefficients multiplied by the rest of exogenous regressors within the linear model (demand, wind and solar energy forecasting with their respective lags, marginal price lags, and calendar variables), at time t.

– C_t^i : Marginal generating cost for block i at time t.

– σ_t^i: Allowed variability for the i-th block price offer with respect to its generating cost at time t.

– π_ω : Probability assigned to each scenario ω.

– α: Fraction of the profit distribution to be used in the CVaR calculation.

– χ: Weight assigned to the CVaR against the expected profit.

3.2 Formulation

The GENCO knows in advance of sending his offers the maximum energy quantity $Q_t^{\text{Max } i}$ she can produce for each time and energy block, and its associated cost C_t^i. In an efficient market, the producer would directly send these offers (true marginal cost of production) which can be imposed in the current model by fixing $\sigma_t^i = 0$. However, positive values of parameter σ_t^i will let us adjust the degree in which producers with market power can modify their offers and deviate from perfect competition.

As described in Section 2.3, the behavior of the marginal price, concerning her offers, is learned through constraint learning (Maragno et al., 2021) from the historical relationship between these variables and the rest of the covariates. In particular, a Bayesian linear regression model will let us generate the needed scenarios for the risk-averse two-stage stochastic programming formulation (8). Uncertainty will be related to the market price sensitivity towards the main producer offers, as exposed in Section 2.4. Sampling from the coefficient distributions of the production blocks $\beta^i_{t,\omega}$ will drive uncertainty into the marginal price $\lambda_{t,\omega}$ and the quantity of energy finally produced $Q^i_{t,\omega}$. Hence, the GENCO's aim is to make an optimal decision over the price she offers for each block (P^i_t) at the first stage of the problem, with a direct impact on its expected profits and risks.

The stochastic model is formulated as follows:

$$\max_{\Theta} \quad (1-\chi)\sum_{\omega\in\Omega}\pi_\omega\sum_{t\in T}\left[\lambda_{t,\omega}Q^{ren}_t+\sum_{i\in I}(\lambda_{t,\omega}Q^i_{t,\omega}-C^i_tQ^i_{t,\omega})\right]+\chi\left(\eta-\frac{1}{\alpha}\sum_{\omega\in\Omega}\pi_\omega s_\omega\right) \tag{8a}$$

s.t.

$$Q^{ren}_t = Q^{\text{Max ren}}_t \quad \forall t \tag{8b}$$

$$u^i_{t,\omega}\in\{0,1\} \quad \forall i,t,\omega \tag{8c}$$

$$\lambda_{t,\omega}-P^i_t \le u^i_{t,\omega}M \quad \forall i,t,\omega \tag{8d}$$

$$P^i_t-\lambda_{t,\omega}\le(1-u^i_{t,\omega})M \quad \forall i,t,\omega \tag{8e}$$

$$Q^i_{t,\omega}=u^i_{t,\omega}Q^{\text{Max } i}_t \quad \forall i,t,\omega \tag{8f}$$

$$\lambda_{t,\omega}=\beta_0+\beta^{ren}Q^{ren}_t+\sum_i\beta^i_{t,\omega}P^i_t+D_t \quad \forall t,\omega \tag{8g}$$

$$C^i_t-\sigma^i_t\le P^i_t\le C^i_t+\sigma^i_t \quad \forall i,t \tag{8h}$$

$$0<P^i_t\le P^{i+1}_t \quad \forall i,t \tag{8i}$$

$$\eta-\sum_t\left[\lambda_{t,\omega}Q^{ren}_t+\sum_i(\lambda_{t,\omega}Q^i_{t,\omega}-C^i_tQ^i_{t,\omega})\right]\le s_\omega \quad \forall\omega \tag{8j}$$

$$0\le s_\omega \quad \forall\omega \tag{8k}$$

where $\Theta=\{P^i_t, Q^i_{t,\omega}, u^i_{t,\omega}, \eta, s_\omega\}$ is the set of optimization variables.

The objective function (8a) represents the weighted sum of the GENCO expected value and the CVaR of her profit. CVaR will be employed to measure the risk taken by the producer, which equals the expected value of $100\alpha\%$ scenarios with the lowest profit. In particular, we use the linear formulation of the CVaR proposed by Rockafellar and Uryasev (2000). The parameter $\chi\in[0,1]$ is used to model the risk aversion level of the producer. Thus, when χ is equal to zero the producer acts as a risk-neutral decision-maker. On the other hand, when χ is equal to one, the producer can be considered risk-averse, and his decisions will be focused on improving the left tail of the profit distribution. The expected profit is computed as the sum of the profits over the set of scenarios multiplied by their respective probabilities π_ω. The profit is composed of the revenues of the dispatched energy blocks, paid at a marginal price $\lambda_{t,\omega}$, minus their production costs.

Constraint (8b) establishes that the produced renewable energy Q_t^{ren} at time t equals the estimated one in order to send block offers to the pool $Q_t^{\text{Max ren}}$. Constraints (8c)-(8e) model whether the price requested for one block P_t^i is lower than the marginal price $\lambda_{t,\omega}$, or not. In particular, (8c) assigns variable $u_{t,\omega}^i$ a binary domain while (8d) and (8e) are employed through the big-M method in order to assign $u_{t,\omega}^i$ a value of one if the price of the block i is under the marginal price, and zero otherwise.

With this assigned value of $u_{t,\omega}^i$, equation (8f) will establish the dispatched energy. That is, $Q_{t,\omega}^i$ will be equal to the estimated energy that can be produced $Q_t^{\text{Max } i}$ for the block i if the price requested for that block is under the marginal price. On the other hand, the value of $Q_{t,\omega}^i$ will be zero if the price of that block is over the marginal price. That is, it is not a dispatched energy block.

Equation (8g) represents the embedded linear model: the learned constraint. In this case, as in a classical linear model, the marginal price will be estimated as the sum of the intercept β_0, plus the quantity of renewable energy multiplied by its coefficient $\beta^{ren} Q_t^{ren}$, plus each block price (decision variable) multiplied by their stochastic coefficient variables (generated scenarios), i.e., $\beta_{t,\omega}^i P_t^i$. Finally, it is included the value composed by the sum of the rest of the covariates multiplied by their coefficients, i.e., D_t.

Constraint (8h) determines the producer flexibility of setting a block price P_t^i above or below its true marginal cost C_t^i, as a function of parameter σ_t^i. In particular, the impact of σ_t^i will be used to study how the GENCO can increase her profit by modifying the offering block prices. Equation (8i) ensures an increasing offering curve, a condition that is required in most electricity markets (see Figure 1). Finally, constraints (8j) and (8k) will be employed in order to characterize the CVaR. In particular, the value of η would be equal to the Value at Risk at the optimal solution to the problem (8).

One of the advantages of this formulation is that can be easily transformed into a mixed-integer linear problem, as the only non-linear term is the product between two variables: $\lambda_{t,\omega} Q_{t,\omega}^i$, which according to (8f) is equivalent to $\lambda_{t,\omega} u_{t,\omega}^i Q_t^{\text{Max } i}$. The product $\lambda_{t,\omega} u_{t,\omega}^i$ involves a continuous variable and a binary one, which can be linearized without approximation (big-M approach).

4 Case study

The main goal of this case study is to analyze the optimal bidding strategy of a large GENCO in a data-driven context. For this purpose, real-world data will be employed during the study, jointly with an out-of-sample validation, to show how price variations in GENCO's offered energy blocks can modify the marginal market price and make her profit increase.

4.1 The dataset

In this case study, we will focus on a large GENCO within the Spanish electricity market. This GENCO produces around 25% of the energy in the Spanish system. As exposed in Sections 2 and 3, we will employ different types of data to characterize the problem.

Firstly, in order to obtain discretized energy blocks and compute a marginal price with the inelastic demand, the hourly supply curves for each GENCO in the day-ahead market are necessary. This information is open-access and provided by the designated electricity market operator for the Iberian Peninsula, OMIE (OMIE, 2022). Two years of data have been collected on an hourly frequency, from June 2017 to June 2019. This data includes day-ahead hourly supply curves (similar to Figure 1) for the production units from the large GENCO as well as from the competitors. That is, more than 30 million observations were processed. Hourly data from June 1st, 2017 to May 31st, 2019 will be mainly employed to train our linear model to predict the day-ahead marginal price. June 2019 will be used as a test, applying an out-of-sample validation approach.

The selection of this time range is due to two main reasons: there were no atypical external factors that made electricity price increase or decrease from the average price of the decade, and there were no significant changes in the renewable power installed capacity that could bias the analysis.

Regarding the covariates employed in the predictive model (price lags, demand, solar, and wind energy forecasts with their respective lags), they are obtained through the information system of Red Eléctrica Española (the Spanish system operator), ESIOS (ESIOS, 2022). All this information is assumed to be known in advance at the time of the decision-making.

4.2 Applied methodology

4.2.1 Market dicretization

The first step in the methodology process is to optimally discretize the offering supply curves. As exposed in Section 2, we will consider a specific supply curve discretization for the GENCO and another one for the rest of the competitors. This allows better characterizing the GENCO's curve and gaining technological insights, which will be useful in designing the marginal price prediction model.

All the hourly supply curves will be discretized into 7 different blocks, both for the main producer and for the competitors. For each block, we assume its energy quantity and price represent how much quantity can be produced and at what cost. We consider 7 blocks as a reasonable number of blocks to capture the main functional properties of the supply curves, while not increasing in excess the dimensionality of the mixed-integer linear problem (8). The first block will always represent the amount of estimated renewable production (at zero cost), and the last two blocks, high-cost generating technologies that, as the historical data reflects, are never marginal. An example of this characterization was shown in Figure 1.

Once the discretized market supply curve is obtained, we approximate the hourly aggregated demand curve by an inelastic one. We get this value from the estimated hourly demand provided by ESIOS. This inelastic demand, jointly with the aggregated supply curve will render an intersection point, that sets a first marginal price with its corresponding total dispatched energy (Figure 2).

However, we know that this first intersection can be far from the resulting market price value, as technical restrictions come into play. These change the shape of the supply curve, creating a gap between the intersection of the curves and the resulting hourly marginal price (Figure 3).

This is a phenomenon that we can not precisely quantify without a detailed physical description of the electricity system, but that must be considered somehow by the GENCO, as it may modify the resulting market outcomes. Nevertheless, by using historical data, we have statistically characterized this difference, and used it to replicate the resulting supply curve including these technical corrections. The difference in energy quantities will allow us to displace our discretized supply curve and give a more accurate estimation of the market marginal price. After this displacement is applied, we compute our new intersection, obtaining the final market marginal price and the dispatched energy.

For a fair out-of-sample model validation, the real hourly displacement due to the technical restrictions cannot be known in advance by the GENCO. Thus, as a proxy, we will employ the mean displacement of the last two months, grouped by hourly periods. We will see that this is an effective strategy to characterize this phenomenon.

4.2.2 Marginal price predictive model

Now that hourly marginal prices from the discretized supply curves have been computed, we seek an adequate predictive model (8g), i.e., to estimate the marginal price as a function of the offering prices and several covariates, to be embedded within the stochastic optimization problem (8). There are two main approaches regarding model selection. Firstly, we could have chosen linearizable machine learning models (Tree-based methods, KNNs, etc) to get the minimum possible error on the prediction. However, with these models, we would lose explainability and appropriate probabilistic characterization of the uncertain parameters. Hence, we have chosen to use a linear regression model through a Bayesian approach. In this way, the model will be interpretable, and random scenarios will be obtained through the price block coefficients' probabilistic distributions.

For this model, the following predictors have been employed:

- Quantity of renewable energy offered by the large GENCO power producer.
- Block prices (6 decision variables) offered by the GENCO. These decision variables include all the block prices except the one related to the renewable energy, that is offered at zero price.
- Demand estimation for the respective hour, and 7 lags (24, 48, ..., 168 hours), 24 hour rolling mean, maximum and minimum.
- Wind power estimation for the respective hour, and 7 lags (24, 48, ..., 168 hours), 24 hour rolling mean, maximum and minimum.
- Solar power estimation for the respective hour, and 7 lags (24, 48, ..., 168 hours), 24 hour rolling mean, maximum and minimum.
- 7 marginal price lags (24, 48, ..., 168 hours).

– Calendar covariates: dummies for day of the week and month, and binary variable for holidays.

Therefore, the dataset will be formed by a total of 70 different predictors for the hourly market marginal price model, being trained in the two-year hourly period stated above.

Table 1 summarizes the model performance on the training set. We show the Mean Absolute Error (MAE) and the Root Mean Squared Error (RMSE) as the performance metrics for both the full model and the model without decision variables (offering block prices) as predictors. MAE is computed as the average absolute deviation of the predictions from the real price values, and RMSE as the squared root of the mean squared prediction error. Furthermore, a cross-validation approach has been included using 6 folds (4 months per fold). Results have been computed using the mean of the coefficient probability distributions (as in classical linear regression models). They show an acceptable error value, and a slight improvement of the performance when GENCO's offering block prices are added as features ("Full model"). Furthermore, the model is not excessively overfitted, as the cross-validation error is similar to the one we get employing the complete training set.

Table 1: Full linear model and model without decision variables performance metrics.

	Training MAE	Training RMSE	Cross-val MAE	Cross-val RMSE
Full model	8.22	10.69	8.65	11.04
Model w/o decision vars.	8.32	10.81	8.88	11.23

As a final note, we have also analyzed the impact of the standardized coefficients for the decision variables predictors concerning the rest of the standardized predictors. That is, we have computed $\left(|\beta^{ren}| + \sum_i |\beta^i|\right) / \sum_j |\beta^j| \times 100$, where β^i represents the coefficients related to the block prices (decision variables), and the set of $\beta^j\ \forall j$ depicts the complete set of regression coefficients for all the covariates. Notice that, for this analysis, the point estimation of the coefficients was employed, that is, the expected value of the distribution fitted utilizing Bayesian linear regression. This computation throws a value of 4.84%.

With this result, jointly with the model performance, we assert that the linear model, beyond being simple and interpretable, is accurate enough to predict the marginal price of the market making use of the considered decision variables (price offered per block) and external covariates. For these reasons, we will assume the selected producer is a large GENCO, whose decisions do affect the marginal price, and hence, her profits.

4.2.3 Scenario generation

The last step is to generate scenarios $\omega \in \Omega$ for the stochastic optimization problem. As it has been stated in Section 2.4, the scenario uncertainty will come from the coefficient distributions of the Bayesian linear regression model. That is, the effect of one block price over the marginal price is not fixed: it is conceived as stochastic.

A total of 200 scenarios have been generated by sampling from the block coefficients estimated normal distributions (Table 2).

Table 2: Mean and standard deviation estimated parameters for normally distributed price block coefficients.

	P_t^1	P_t^2	P_t^3	P_t^4	P_t^5	P_t^6
$\mathbb{E}[\hat{\beta}_i]$	-0.00537	0.11155	0.05026	0.00455	-0.01415	0.03217
$\hat{\sigma}_{\beta_i}$	0.00694	0.01620	0.01413	0.01181	0.00620	0.01614

As can be seen from the distributions, the prices offered for the second block have the biggest influence on increasing the marginal price of the market. On the other hand, the influence of the other blocks is more uncertain. One potential reason is that the first non-zero cost block is matched in most of the cases, so its price is not determinant in the market marginal price. Similarly, the most expensive energy blocks are rarely marginal in the day-ahead market.

4.3 Stochastic results

We solve the stochastic optimization problem (8) through the scenario generation described in the previous section. This optimization problem will be solved for each day within the month of June 2019. In the Spanish market, producers send their offers at 12:00 for every hour of the following day. For that reason, the quantity offered per block ($Q_t^{\text{Max } i}$) and its true estimated cost (C_t^i) will be determined by solving the model (6) to obtain the respective blocks for the GENCO.

The proposed optimal offering problem has been solved through a Python 3.9.12 implementation, using Pyomo 6.3 (Hart et al., 2017). The selected mathematical solver for all the computations was Gurobi (Gurobi Optimization, LLC, 2022) in its version 9.5. Besides, the computer employed included a CPU Intel Core i7 10700, RAM of 64 GB, and NVIDIA GeForce GTX 2060 graphic card. The computation time is dependent on the particular scenario set and the assigned parameters, especially σ_t^i, where a bigger value increases the decision space, and therefore, the computation time. In general, none of the following experiments took more than 30 minutes to reach global optimality.

As indicated, we use σ_t^i to quantify possible deviations with respect to C_t^i. For example, if the production cost of the first block at time t is 25€, and we set a variability of 10%, σ_t^i will have a value of 2.5€. From now on, we will directly refer to σ_t^i as this percentage. This variability has been applied to the first four non-zero cost blocks. For the remaining two blocks, σ_t^i was set to zero, as their cost is so high that their price never intervenes in the marginal price formation. Moreover, the α value affecting the CVaR formulation is set to 10%, that is, in the risk-averse case, we focus on improving the expected value of the scenarios below the 10th percentile of the profit distribution.

Regarding the rest of the covariates included in D_t that affect the constrained linear regression model (8g), they are also known at the time of the decision making, as the platform ESIOS offers open-access estimations of hourly demand, renewable production, etc. Thus, we consider this problem as a realistic data-driven approach.

Concerning the risk aversion level, we solve the optimization problem for the cases where the producer is risk-neutral ($\chi = 1$) and risk-averse ($\chi = 0$). In the computation process, we used χ values 0.001 and 0.999 instead of 0 and 1 for stability in the results. A code example for a σ_t^i level of 10% and $\chi = 0$ in openly available in a Github repository (Alcántara, 2022).

Firstly, we present a summary of the stochastic results for χ values (risk aversion level) of zero and one, and σ_t^i values of 0%, 5%, 10% and 15%, in Table 3. The idea of limiting σ_t^i to relatively small values comes from the employment of a Bayesian linear regression as the model to predict the day-ahead market price. We believe that, as we approximate a complex system (the electricity market) with a simple linear model, limiting GENCO price flexibility to a small value will make the linear approximation more effective. This hypothesis is later corroborated by the numerical results in the out-of-sample validation.

The first and second columns of Table 3 represent the price flexibility over the cost of production and the risk aversion level of the GENCO, respectively. The expected profit is computed as the mean daily profit over the 30 days of testing; the same testing period is used to derive the expected CVaR. The fifth column represents the expected learned marginal price during June 2019, whereas the sixth one reports the expected dispatched energy, that is, the energy to be produced as its offered price is under the market marginal price. Finally, the last four columns present the mean price offered for each of the first four blocks (decision variables).

Table 3: Result summary of the stochastic optimization problem.

Price flexibility	Risk aversion	$\mathbb{E}$[Profit] (€)	$\mathbb{E}$[CVaR] (€)	$\mathbb{E}[\lambda_{t,\omega}]$ (€/MWh)	$\mathbb{E}[Q_t^{ren}+\sum Q_{t,\omega}^i]$ (MWh)	$\mathbb{E}[P_t^1]$ (€/MWh)	$\mathbb{E}[P_t^2]$ (€/MWh)	$\mathbb{E}[P_t^3]$ (€/MWh)	$\mathbb{E}[P_t^4]$ (€/MWh)
$\sigma_t^i = 0\%$	$\chi = 0$	3478823.44	3039206.15	43.19	4045.88	37.23	49.69	60.84	77.51
	$\chi = 1$	3478823.44	3039206.15	43.19	4045.88	37.23	49.69	60.84	77.51
$\sigma_t^i = 5\%$	$\chi = 0$	3518082.54	3070042.51	43.63	4001.96	35.75	52.18	63.86	80.96
	$\chi = 1$	3515066.28	3075152.73	43.61	4023.92	35.51	52.16	63.87	73.78
$\sigma_t^i = 10\%$	$\chi = 0$	3553546.47	3098084.94	44.03	3985.29	34.66	54.44	66.69	83.78
	$\chi = 1$	3546894.81	3109359.34	44.01	4016.29	33.95	54.49	66.85	71.11
$\sigma_t^i = 15\%$	$\chi = 0$	3588804.76	3127173.84	44.42	3972.68	33.31	56.84	69.16	85.68
	$\chi = 1$	3578339.65	3141192.84	44.39	4012.44	32.48	56.89	69.43	71.38

By differentiating between levels of price flexibility, we can see how the expected profit increases as this flexibility does. This is due to the offering price adjustments that the GENCO employs to modify the marginal price of the market. Starting from a marginal price of 43.19€/MWh when the GENCO makes offers at production cost, we can see an increment of the marginal price up to 44.42€/MWh where the flexibility is 15% of the production costs.

There are also differences regarding the risk aversion level. When the producer is risk-averse ($\chi = 1$), expected profits are slightly lower but the CVaR improves, as the marginal price does not increase as much as in the case where the producer is risk-neutral ($\chi = 0$). The opposite occurs with the dispatched energy: the risk-averse GENCO tries to ensure the production even if it is at a lower marginal price. Therefore, the expected dispatched energy for the risk-neutral GENCO is lower than the risk-averse one.

In relation to the block prices offered, significant differences appear when varying the price flexibility and risk aversion levels. For example, as the flexibility increases, the mean price offered for the first block decreases, while the price for the second block increases. This can be caused by the necessity of the producer of ensuring one dispatched block (the first one) and, on the other hand, increasing the price of the second block to leave the competence behind and increase the marginal price. The same behavior can be seen for the third block price. Regarding the differences between risk aversion levels, most of the time, the risk-averse and risk-neutral GENCOs follow the same strategy, but in the risk averse case, she tries not to increase (or decrease) the block price that much compared to the risk neutral one. However, the opposite behavior is observed for the price of the fourth block.

What follows will graphically show the stochastic results for the complete testing period, studying the profit distribution and optimal block prices for different levels of risk aversion and price flexibility.

Firstly, in Figure 4, we show the expected daily profit distribution (left) and hourly block prices (right) when no strategic offering is allowed. That is, the price flexibility is 0% and the block prices are the production costs. With this level of price flexibility, the block prices show the real production cost during June 2019. These costs are obtained following the methodology exposed in Section 2.1. That is, we discretized the original supply curve of the large GENCO and assumed the obtained values for each block to be the true costs. Regarding the expected profit distribution, we can see how it fluctuates between 2 and 6 million euros per day. Most of the atypical expected profits are in the right tail of the distribution, with fewer of them in the left tail.

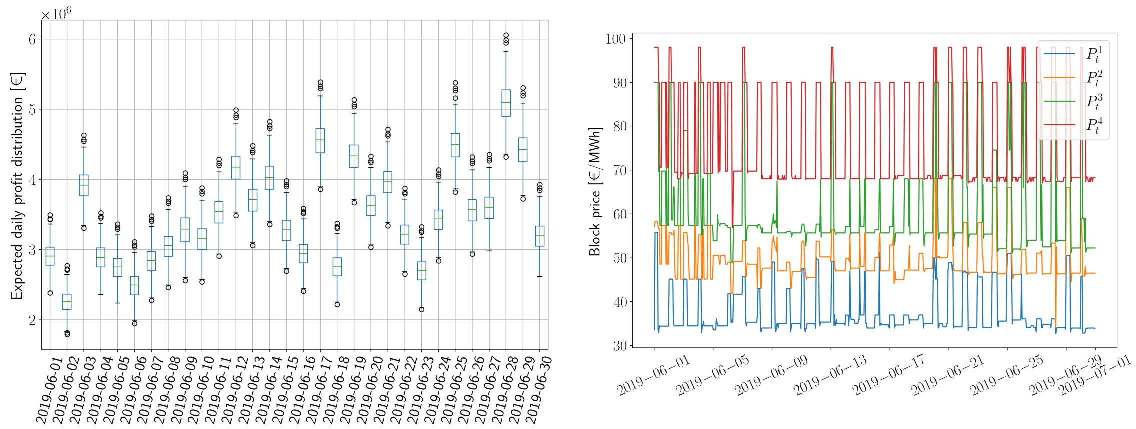
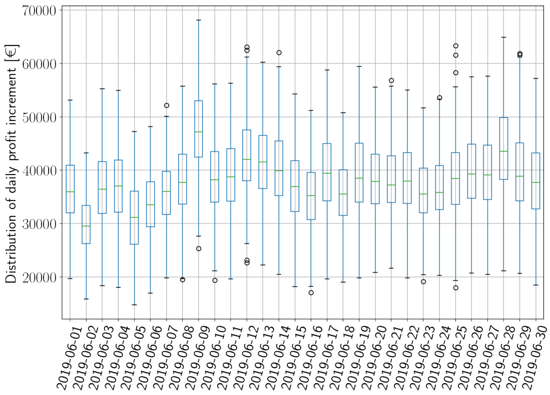
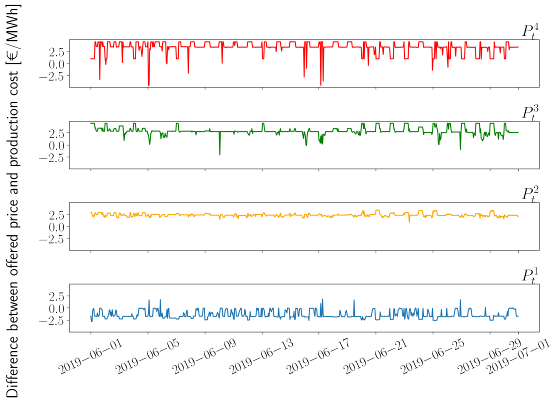


Figure 4: Expected daily profit distribution (left) and hourly block prices (production costs, right) for $\chi = 0, 1$ and $\sigma_t^i = 0\%$

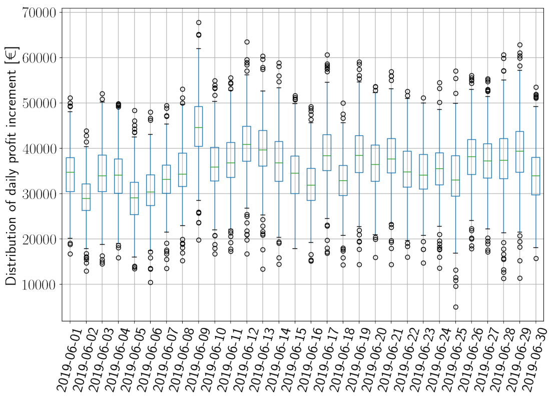
Next, we increase the price flexibility level up to 5% and recast the results by risk aversion level in Figure 5. We represent the risk-neutral GENCO in Figures 5(a) and (b), and the risk-averse in Figures 5(c) and (d). Figures 5(a) and (c) show the distribution of daily profit increments compared to the base case when the GENCO offers her energy at production costs. This means that, for the same scenario, we compute the difference between the expected profit at a price flexibility level of 5%, and a level of 0%. On the other hand, Figures 5(b) and (d) represent the difference between the optimal prices offered for each block for a price flexibility level of 5% and the base case (the production cost). The same descriptive setting is applied for the rest of analyzed cases with different flexibility levels in Figures 6 and 7.



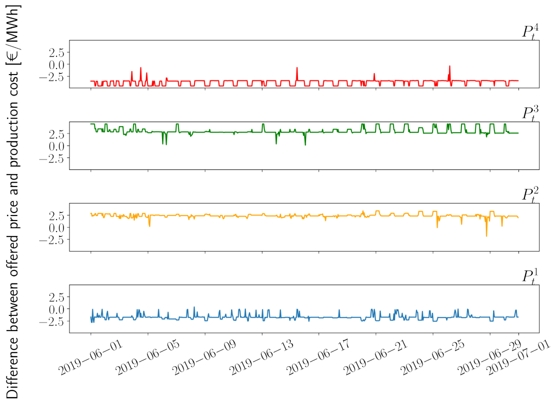
(a) Distribution of daily profit increment, $\chi = 0$, $\sigma_t^i = 5\%$



(b) Block prices vs. production costs, $\chi = 0$, $\sigma_t^i = 5\%$



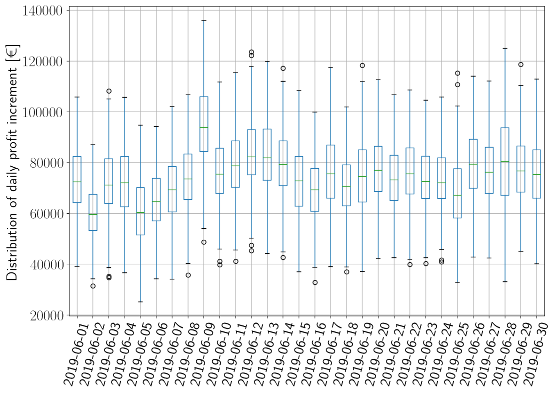
(c) Distribution of daily profit increment, $\chi = 1$, $\sigma_t^i = 5\%$



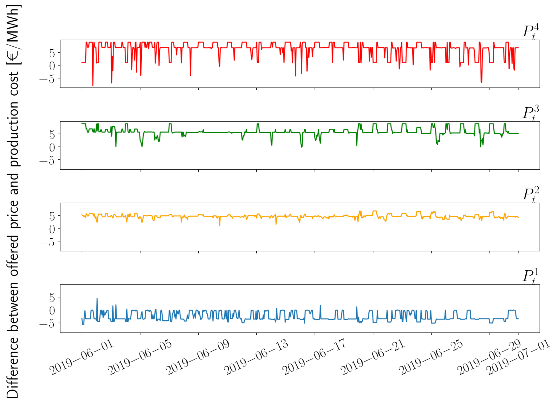
(d) Block prices vs. production costs, $\chi = 1$, $\sigma_t^i = 5\%$

Figure 5: Expected profit increment distribution and block prices for cases $\chi = 0, 1$ and $\sigma_t^i = 5\%$

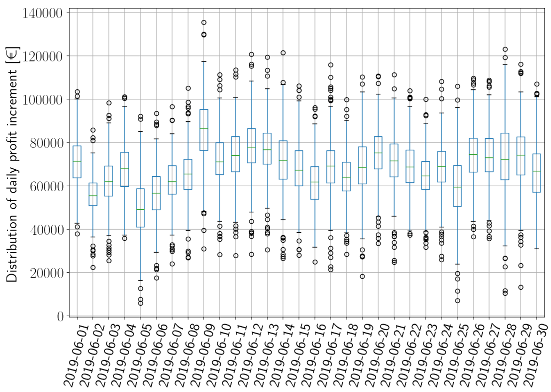
In Figure 5 we can see big differences regarding the profit increment distribution by risk aversion level. Although there are not so many differences in the mean of the distribution (this mean is slightly bigger for the risk-neutral GENCO), we can appreciate the change in the distribution shape. For instance, the distribution for the risk-neutral GENCO has a higher variance than the risk-averse case. Besides, we can notice how the risk-averse GENCO tries to improve the lower tail of the distribution, transforming worst-case profit increments into atypical values. Concerning the offering strategy, we can see the main differences in the price offered for the fourth block, where the risk-neutral GENCO tries



(a) Distribution of daily profit increment, $\chi = 0$, $\sigma_t^i = 10\%$



(b) Block prices vs. production costs, $\chi = 0$, $\sigma_t^i = 10\%$



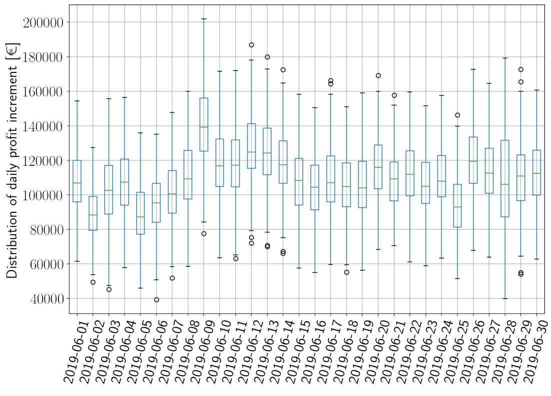
(c) Distribution of daily profit increment, $\chi = 1$, $\sigma_t^i = 10\%$

(d) Block prices vs. production costs, $\chi = 1$, $\sigma_t^i = 10\%$

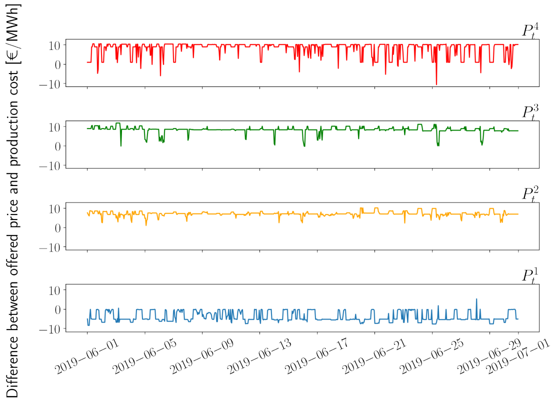
Figure 6: Expected profit increment distribution and block prices for cases $\chi = 0, 1$ and $\sigma_t^i = 10\%$

to set the maximum possible price while the risk-averse GENCO tries to minimize it. Furthermore, the risk-neutral GENCO has a higher variance for the prices of the first block compared to the risk-averse case.

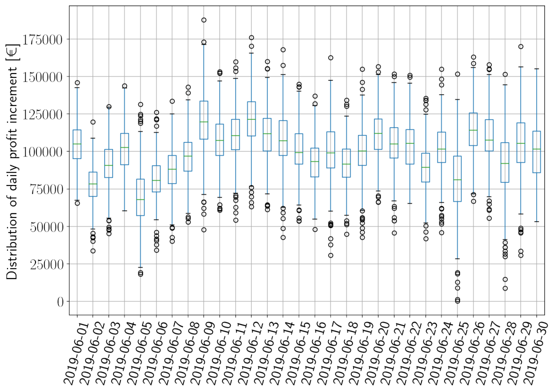
Following the increase of flexibility, we repeat our analysis for the case of $\sigma_t^i = 10\%$ in Figure 6. In this case, we can observe the same behavior regarding the distributions of daily profit increment. In general, the mean expected increment of profit goes from 40000€ in the former case to more than 60000€. Besides, the variance in the distribution increases for both the risk-neutral and the risk-averse GENCO. The latter continues to improve the lower tail of the profit increment distribution. In relation to the offered prices, both GENCOs take advantage of the price flexibility level by modifying their offers in a larger range concerning the production costs. The behavior regarding the fourth block continues to be the opposite between different levels of risk aversion, and the risk-neutral



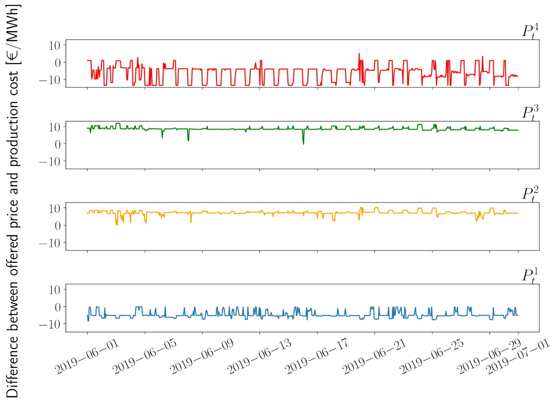
(a) Distribution of daily profit increment, $\chi = 0$, $\sigma_t^i = 15\%$



(b) Block prices vs. production costs, $\chi = 0$, $\sigma_t^i = 15\%$



(c) Distribution of daily profit increment, $\chi = 1$, $\sigma_t^i = 15\%$



(d) Block prices vs. production costs, $\chi = 1$, $\sigma_t^i = 15\%$

Figure 7: Expected profit increment distribution and block prices for cases $\chi = 0, 1$ and $\sigma_t^i = 15\%$

GENCO still adds more variability to the prices of the first block. This might be due to her risk aversion level, trying to get maximum profits even in the most secure case of a dispatched energy block.

We will finish this analysis with a 15% level of price flexibility case. Results are shown in Figure 7. We can appreciate the largest differences in this case. Regarding the expected profit increment distributions, mean values increase to more than 100000€ per day. The interquartile range continues to increase: in the risk-neutral case also up to 100000€ per day. The risk-averse GENCO is able to increase the first quartile to a value over 60000€, with a smaller variance on the profit increment distribution than the risk neutral GENCO. In relation to the block prices, the risk-averse GENCO tries to minimize the price of the fourth block and maximize the price of the third one, which suggests the target of joining the prices of both blocks. On the other hand, the risk-neutral GENCO keeps her block prices apart at high values. This behavior might be due to the risk the GENCO faces if a block of the competitor falls below his blocks and excludes her from dispatch. However, the risk-averse GENCO tries to ensure the first block while keeping the rest of the blocks at medium prices.

To finish this section of results, we will show the differences regarding the mean expected daily profit and the mean daily CVaR between different levels of GENCO's risk aversion. For this example, we set a price flexibility level of 10% and solve our stochastic optimization problems for $\chi = 0, 0.1, 0.2, \ldots, 1$. A graphical illustration of the efficient frontier can be seen in Figure 8.

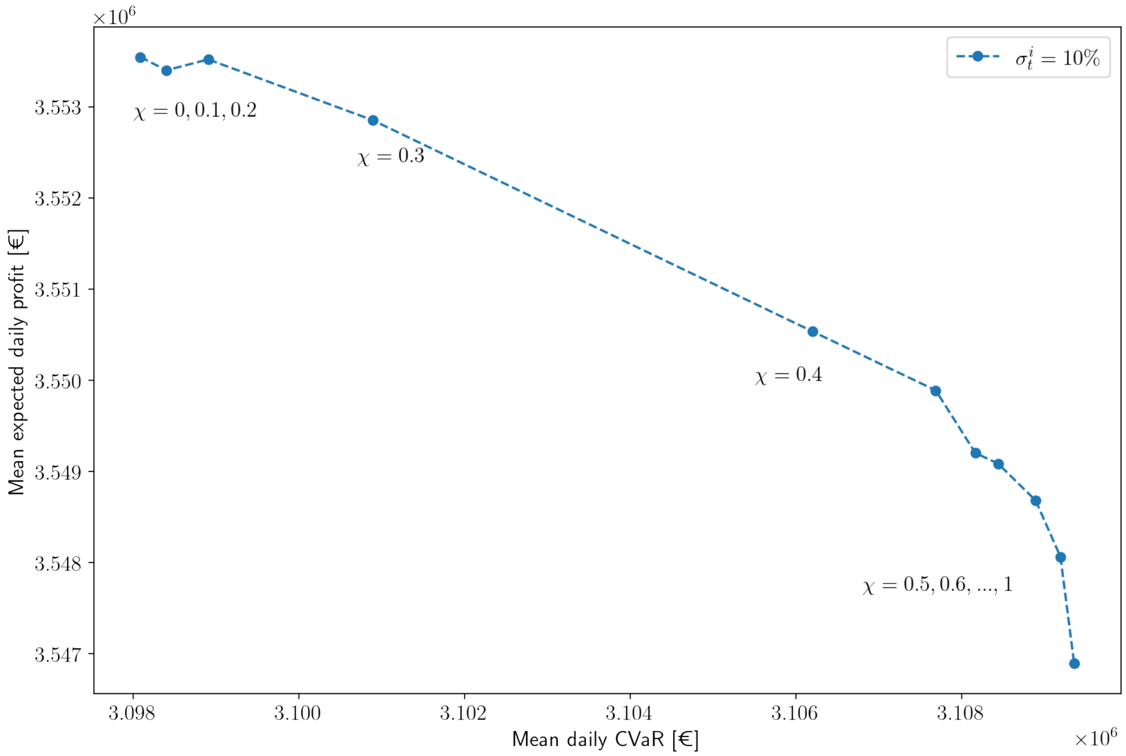


Figure 8: Expected profit versus CVaR for different values of χ. Price flexibility is set to 10%

As can be noticed, we first find a group of low levels of risk aversion, for those GENCOs with $\chi = 0, 0.1, 0.2$, where they all obtain similar expected values of profits and CVaR. With $\chi = 0.3$ and $\chi = 0.4$, a linear decrease of mean expected profit (and a linear increment of CVaR) begins. Finally, when χ is bigger than 0.5, the expected profit drops substantially and small improvements in the CVaR are made.

4.4 Out-of-sample model validation

As we have seen, the stochastic optimization model illustrates how allowing strategic price offering for a GENCO can increase the day-ahead market marginal price. Furthermore, we see differences in offering prices, marginal prices, and expected profits in relation to the risk aversion level of the GENCO.

This analysis was based on the assumption that the linear model (8g) was an accurate representation of the market price response to the GENCO' strategic offers. However, we aim to test if this stochastic optimal strategy would actually work in a real market. That is, we want to know what may occur when the GENCO sends her optimized supply curves to the Spanish day-ahead market. For this reason, we will follow an out-of-sample validation of the derived offering strategy, using data from June 2019.

For that purpose, the methodology will stand as follows:

1. Once the optimal strategy by the GENCO is derived from (8), send the resulting supply curve to the market. Block quantities will be $Q_t^{\text{Max } i}$ with prices P_t^i, jointly with Q_t^{ren} at zero cost.

2. Aggregate the GENCO supply curve with the one collected by the market operator from the rest of the market competitors (also discretizated in 7 blocks, see Section 2.2), and use the inelastic demand forecast from ESIOS to obtain the initial marginal price.

3. Displace the market supply curve employing the mean displacement of the last two months at its corresponding hour, to reproduce technical adjustments performed by the market operator.

After all these steps are done, we will finally get an estimated marginal price. Profits for the GENCO can be computed as the product of the dispatched quantities (those blocks below the resulting marginal price) times the marginal price minus the production costs of this energy.

These out-of-sample results have been computed for different levels of risk aversion (χ) and price flexibility (σ_t^i). Table 4 summarizes the main obtained insights. The third column of the table shows the average daily profit for the GENCO in June 2019. The fourth and fifth columns indicate the first and third quantiles over the profit distribution. We find the variance of the profit rightwards. Finally, the mean hourly marginal price is computed.

The first and one of the most important results that we can confirm out-of-sample is that allowing the large GENCO offers to deviate from marginal costs makes the market marginal price to increase. Besides, the risk-neutral GENCO makes this marginal price increase more than the risk averse GENCO.

For a risk-neutral GENCO, the maximum mean profit is obtained for price flexibility of 15%, whereas for risk averse GENCO the maximum profit takes place at a price flexibility level of 5%. Furthermore, we can see how the average profit for the risk-neutral GENCO

Table 4: Result summary of the out-of-sample model validation.

Price flexibility	Risk aversion	$\overline{\text{Profit}}$ (€)	Q_1 (€)	Q_3 (€)	Var[Profit] (€^{2})	$\overline{\lambda_t}$ (€/MWh)
$\sigma_t^i = 0\%$	$\chi = 0$	3565205.49	3303414.68	4060772.16	645124.68	43.82
	$\chi = 1$	3565205.49	3303414.68	4060772.16	645124.68	43.82
$\sigma_t^i = 5\%$	$\chi = 0$	3572792.23	3305175.26	4056156.26	659907.01	43.91
	$\chi = 1$	3570222.41	3305175.26	4047973.11	662875.81	43.88
$\sigma_t^i = 10\%$	$\chi = 0$	3574425.62	3302773.32	4047537.69	661838.42	43.94
	$\chi = 1$	3567050.32	3300685.63	4032516.38	672635.64	43.88
$\sigma_t^i = 15\%$	$\chi = 0$	3574699.76	3298282.72	4051551.90	666520.89	43.97
	$\chi = 1$	3559860.77	3294270.89	4028966.40	677172.50	43.88

is always above the case where no price flexibility is allowed. On the other hand, for a risk-averse GENCO, the mean profit is not improved at a flexibility level of 15%. In general, we can assume that there is some level of flexibility upon which profits stop increasing, this is where the offering curves from the competitors may start determining the marginal price.

Regarding profit quantiles and variance, no significant differences are appreciated. Only quantiles seem to decrease as the price flexibility increases, suggesting that more atypical values appear in the right tail of the profit distribution, making the mean profit to generally increase.

To finish this section, we will show graphically the daily GENCO profit increment in the test period for different levels of price flexibility and risk aversion concerning the true generating cost offering strategy. Figure 9 shows the daily out-of-sample profit increment for risk-neutral (up) and risk-averse (down) GENCO for different levels of price flexibility. Dashed lines represent the mean for each pricing strategy.

As can be seen, in most of the cases, the profit increment is slightly higher for cases where the price flexibility is different from zero, for both levels of risk aversion. Besides, on some specific days, this profit difference is increased on a higher scale, for instance, June 13th, 17th, 27th, or 28th. It is interesting to notice that, for the risk-neutral case, the mean profit increment is almost the same for price flexibility levels of 10% and 15%. However, for the risk-averse case, a large value of price flexibility strategy can even worsen the case where the offers are made at true generating cost. In this case, a low value of price flexibility, i.e., 5%, achieves the best results for the risk-averse large GENCO.

Furthermore, we can appreciate more differences when we show the average profit increment in an hourly basis (Figure 10). As in the case above, we represent the profit increment for several levels of price flexibility and risk aversion: risk-neutral (up) and risk-averse (down). We can see how the biggest amount of profit increment can be made from 8:00 to 10:00, and in some specific hours like 3:00, 13:00 to 15:00, 17:00, 18:00 20:00, 21:00, and 23:00 in a non-homogeneous way.

For the risk-neutral GENCO, we can notice how, on average, and in most of the hours of the day, a profit increment can be obtained by allowing price flexibility in her offering strategy. Besides, when the profit increases, it does it on a bigger scale when the maximum price flexibility is set. However, when the profit decreases with respect to offering at production costs, it decreases more for the case when the price flexibility is allowed up to 15%.

On the other side, the behavior of the hourly profit increments is similar in the case of a risk-averse GENCO. However, more hours where the profit does not increase appear. In the case where the profit suffers an increment, it does it in a similar way for different levels of price flexibility. Nevertheless, the case of a 5% price flexibility allows the risk-averse GENCO to soften the profit decrease.

These results show how the stochastic optimization model is optimistic regarding the profit increment. Nevertheless, the profit increments do occur, but at a lower scale than the stochastic model reported. This might be due to the local limitations of a linear predictive model, as large variations in the competitor block offers, that could change the market outcomes, cannot be taken into account. However, the hourly behavior of the profit increment provides valuable insights and suggests that an hourly price flexibility strategy could be taken into account to further increase GENCO's profits.

To sum up, this out-of-sample validation approach has allowed us to confirm in a real-life context how allowing a GENCO to deviate from its true generating costs makes her expected profits increase. Nevertheless, we have seen how an excessive value of the price flexibility may diminish this effect. Finally, we have learned that the highest increment in the profit can be achieved in some specific hours of the day.

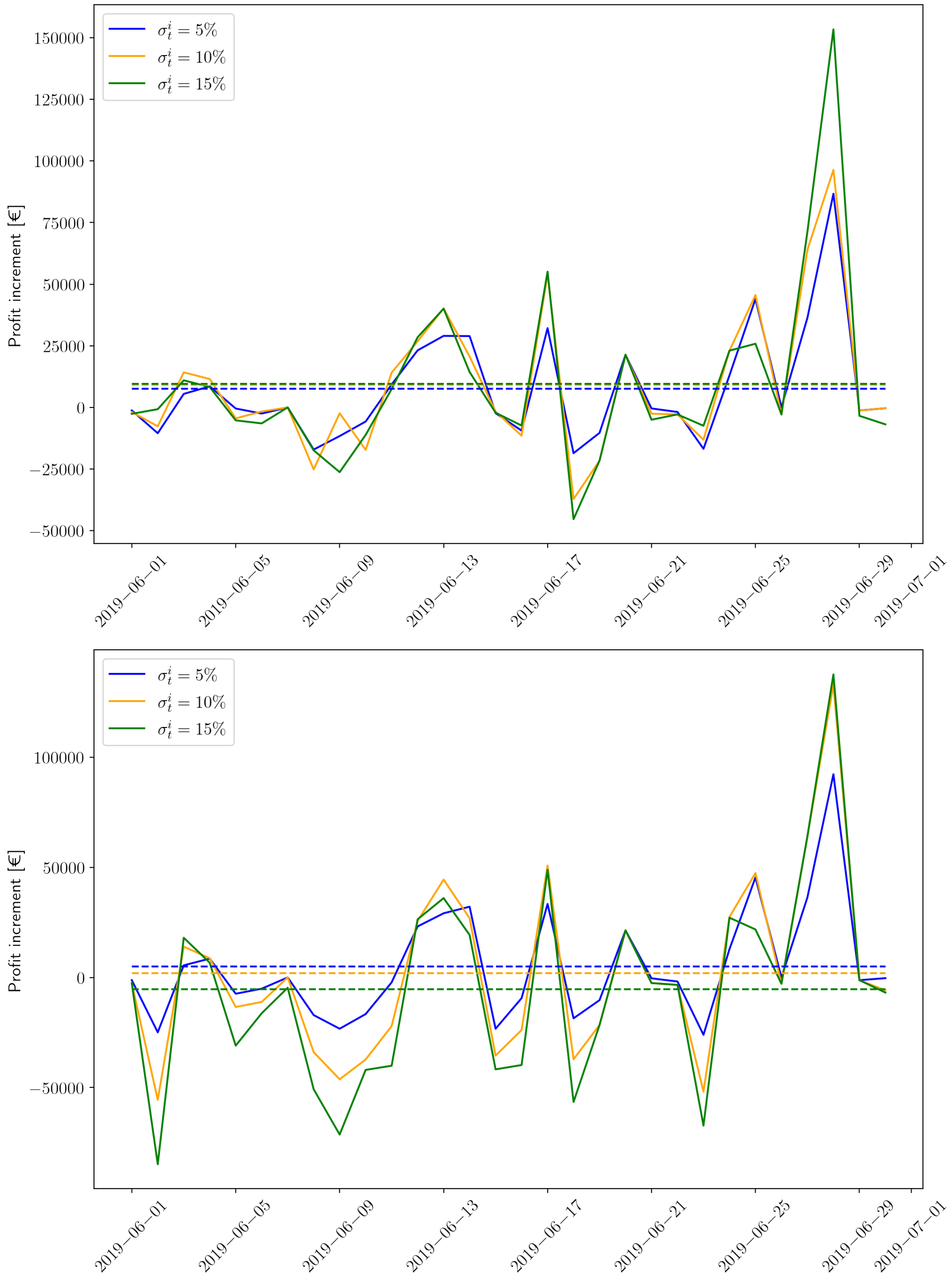


Figure 9: Daily out-of-sample profit increment for risk neutral (up) and risk averse (down) GENCO at different levels of price flexibility

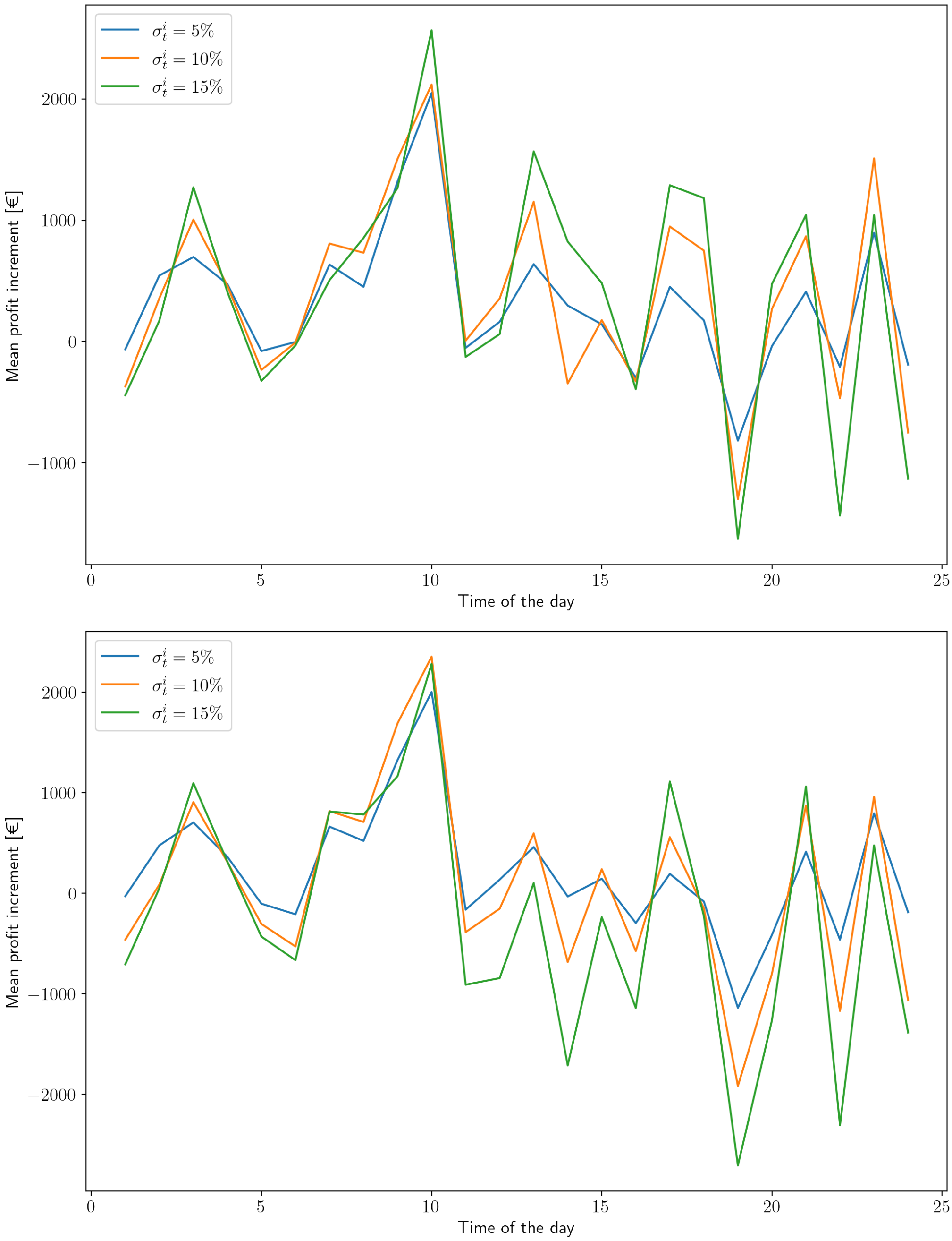


Figure 10: Hourly out-of-sample mean profit increment for risk neutral (up) and risk averse (down) GENCO at different levels of price flexibility

5 Conclusions

In this work, we have dealt with the problem of finding an optimal offering strategy for a large GENCO. That is, knowing the amount of energy that the GENCO can produce and at what cost, setting the prices of the offered energy blocks for a profit-maximizing strategy.

We present a data-driven methodology where the GENCO's supply curves and the ones from the rest of the competitors are optimally discretized. This discretization will allow us to get important insights from their offering strategy, reproduce the market clearing to compute the resulting market price, and ease an out-of-sample validation of the proposed stochastic optimization model.

The relationship between the hourly market marginal price and the block prices offered by the GENCO is modeled through a Bayesian linear regression approach. The advantage of this approach is to obtain an interpretable, fully linear model, from which different scenarios for the price coefficients (sensitivity of the marginal price to the GENCO supply curve) can be sampled from their posterior distribution. This linear model is embedded into a two-stage stochastic optimization model which accounts for risk aversion to derive the optimal supply curve to submit to the day-ahead market.

After an in-depth analysis, stochastic results have shown how allowing the GENCO to deviate from her marginal costs offers, results in a marginal market price increase, which also increases her profits. Besides, there are differences in the pricing behavior of the GENCO depending on her risk aversion level. In general, a risk-neutral GENCO achieves to increase the marginal price to a higher extent than the risk averse GENCO, but worsening the worst case profit scenarios.

One of the main novelties of this work is that the optimal offering strategy is tested out-of-sample. That is, we simulate the actual functioning of the market combining the demand and the offers from the GENCO and her competitors. We show how the proposed optimization model achieves a marginal price increment and a maximum profit for the risk-neutral GENCO for price flexibility levels of around 10% of the production costs, and at 5% for the risk-averse one. Besides, an hourly price flexibility strategy could be even more profitable for the GENCO. Furthermore, these results warn us about the importance of executing effective audits in markets with uniform pricing based on marginal technologies. In these type of markets, offers should be done at marginal generating costs. However, as we have shown, minor increments of the block prices may significantly increase the marginal price which will be transferred to the consumers.

Future work is focused on studying not only the price strategy for the GENCO but also allowing flexibility on the offered energy quantity. This would add more complexity to the optimization problem and the prediction model. Thus, state-of-art linearizable machine learning models should be tested, with the aim of not losing an adequate uncertainty characterization.

Acknowledgements

The authors gratefully acknowledge the financial support from MCIN/AEI/10.13039/501100011033, project PID2020-116694GB-I00 and from the FPU grant (FPU20/00916).

References

Alcántara, A. (2022). *opti-genco-offers* (Version 1.0.0). https://github.com/antonioalcantaramata/opti-genco-offers

Ban, G.-Y., & Rudin, C. (2019). The big data newsvendor: Practical insights from machine learning. *Operations Research*, *67*(1), 90–108.

Bertsimas, D., & Kallus, N. (2020). From predictive to prescriptive analytics. *Management Science*, *66*(3), 1025–1044.

Bertsimas, D., McCord, C., & Sturt, B. (2019). Dynamic optimization with side information. *arXiv preprint arXiv:1907.07307*.

Birge, J. R., & Louveaux, F. (2011). *Introduction to stochastic programming*. Springer Science & Business Media.

Chen, K., Lin, J., & Song, Y. (2019). Trading strategy optimization for a prosumer in continuous double auction-based peer-to-peer market: A prediction-integration model. *Applied energy*, *242*, 1121–1133.

Elmachtoub, A. N., & Grigas, P. (2022). Smart "predict, then optimize". *Management Science*, *68*(1), 9–26.

ESIOS. (2022). Red Eléctrica España. https://www.esios.ree.es

Esteban-Pérez, A., & Morales, J. M. (2021). Distributionally robust stochastic programs with side information based on trimmings. *Mathematical Programming*, 1–37.

Fajemisin, A., Maragno, D., & Hertog, D. d. (2021). Optimization with constraint learning: A framework and survey. *arXiv preprint arXiv:2110.02121*.

Goodrich, B., Gabry, J., Ali, I., & Brilleman, S. (2020). Rstanarm: Bayesian applied regression modeling via Stan. [R package version 2.21.1]. https://mc-stan.org/rstanarm

Gurobi Optimization, LLC. (2022). Gurobi Optimizer Reference Manual. https://www.gurobi.com

Han, X., Kardakos, E. G., & Hug, G. (2018). Offering strategy of a price-maker pv power plant: Multi-stage stochastic programming with probabilistic constraints. *2018 Power Systems Computation Conference (PSCC)*, 1–7.

Hart, W. E., Laird, C. D., Watson, J.-P., Woodruff, D. L., Hackebeil, G. A., Nicholson, B. L., Siirola, J. D., et al. (2017). *Pyomo-optimization modeling in python* (Vol. 67). Springer.

Kardakos, E. G., Simoglou, C. K., & Bakirtzis, A. G. (2015). Optimal offering strategy of a virtual power plant: A stochastic bi-level approach. *IEEE Transactions on Smart Grid*, *7*(2), 794–806.

Maragno, D., Wiberg, H., Bertsimas, D., Birbil, S. I., Hertog, D. d., & Fajemisin, A. (2021). Mixed-integer optimization with constraint learning. *arXiv preprint arXiv:2111.04469*.

MIBEL. (2022). Mercado Ibérico de la Electricidad. https://www.mibel.com/es/home_es/

Mišić, V. V. (2020). Optimization of tree ensembles. *Operations Research*, *68*(5), 1605–1624.

Mundru, N. (2019). *Predictive and prescriptive methods in operations research and machine learning: An optimization approach* (Doctoral dissertation). Massachusetts Institute of Technology.

Muñoz, M. A., Pineda, S., & Morales, J. M. (2022). A bilevel framework for decision-making under uncertainty with contextual information. *Omega*, *108*, 102575.

Murty, K. G. (1994). *Operations research: Deterministic optimization models*. Prentice-Hall, Inc.

OMIE. (2022). OMIE. https://www.omie.es

Pandžić, H., Morales, J. M., Conejo, A. J., & Kuzle, I. (2013). Offering model for a virtual power plant based on stochastic programming. *Applied Energy*, *105*, 282–292.

Paulus, A., Rolínek, M., Musil, V., Amos, B., & Martius, G. (2021). Comboptnet: Fit the right np-hard problem by learning integer programming constraints. *International Conference on Machine Learning*, 8443–8453.

Pérez-Santalla, R., Carrión, M., & Ruiz, C. (2022). Optimal pricing for electricity retailers based on data-driven consumers' price-response. *TOP*, 1–35.

Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of risk*, *2*, 21–42.

Ruiz, C., & Conejo, A. J. (2009). Pool strategy of a producer with endogenous formation of locational marginal prices. *IEEE Transactions on Power Systems*, *24*(4), 1855–1866.

Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. (2016). Probabilistic programming in python using pymc3. *PeerJ Computer Science*, *2*, e55.

Xiao, D., do Prado, J. C., & Qiao, W. (2021). Optimal joint demand and virtual bidding for a strategic retailer in the short-term electricity market. *Electric Power Systems Research*, *190*, 106855.

Yang, S., & Bequette, B. W. (2021). Optimization-based control using input convex neural networks. *Computers & Chemical Engineering*, *144*, 107143.