Automatic Risk Adjustment for Profit Maximization in Renewable Dominated Short-Term Electricity Markets

Bottieau, J.[a,*], Bruninx, K.[b,c], Sanjab, A.[c,d], De Grève, Z.[a], Vallée, F.[a], Toubeau, J-F.[a]

[a] *Power Systems & Markets Research Group, Department of Electrical Power Engineering Unit, University of Mons, Boulevard Dolez 31, 7000 Mons, Belgium*
[b] *KU Leuven Energy Institute, Division Applied Mechanics and Energy Conversion, Department of Mechanical Engineering, KU Leuven, Celestijnenlaan 300 box 2421, B-3001 Leuven, Belgium*
[c] *EnergyVille, Thor Park 8310-8320, 3600 Genk, Belgium*
[d] *Flemish Institute for Technological Research (VITO), Unit Energy Technology, Boeretang 200, Mol, Belgium*

Abstract

State-of-the-art trading strategies in short-term electricity markets employ risk awareness for reducing, inter alia, their exposure to the volatility of electricity prices. To ensure an optimal balance between risk and profit, risk-aversion parameters are traditionally fine-tuned via an offline out-of-sample analysis. Such a computationally-intensive analysis is typically run once, which yields time-invariant risk policies. Instead, this paper proposes the use of machine learning to select, in an online fashion, optimal risk-aversion parameters. This novel automatic risk-tuning approach offers the benefit of continuously adjusting the risk policy based on the dynamically changing market operating conditions. Our approach is tested on two risk-aversion parameters, i.e., the confidence level of the conditional value-at-risk and the budget of uncertainty, respectively considering scenario-based and robust optimization frameworks. A set of performed case studies – focusing on the very short-term dispatch of a market actor participating in electricity markets – using real-world market data from the Belgian power system demonstrate the ability of the proposed methodology to outperform traditional offline risk policies.

Keywords: Electricity markets, machine learning, imbalance settlement, stochastic optimization, risk management

1. Introduction

While competing in liberalized electricity markets, actors adapt their short-term dispatch decisions based on their expectations of future market outcomes to maximize their profit [1]. These decisions are made while facing market uncertainties stemming from, e.g., the prevailing market prices, which expose the actor to financial risks [2], where financial risk is defined as the possibility that an actor's financial outcome deviates adversely from its expectation [3]. In this context, stochastic decision support tools, including financial risk management, allow improved scheduling decisions in short-term electricity markets, giving the possibility for actors to control the risks associated with their positions [4].

Two distinct methodologies can be considered for trading strategies: i) *performance satisfying methods*, such as information-gap theory decision [5, 6], which ensure a minimum acceptable profit, and ii) *performance maximization methods*, which, in contrast, maximize the expected profit of the market actor given its representation of the uncertainty space. Examples of performance maximization methods include stochastic programming [7–9], robust optimization [10, 11] and chance constrained programming [12, 13]. In performance maximization methods, the market actors take two successive decisions when employing risk awareness: i) the selection of their risk attitude, i.e., the setting of their financial risk management (e.g., their own risk averseness level), which determines their risk policy, and ii) their scheduling/dispatch decision, which maximizes their expected profit given their risk policy.

Risk awareness in performance maximization methods has attracted a high-level of interest within the power systems community, including several studies on conventional electricity generation [14–16], energy storage systems [17, 18], weather-dependent generation [7, 8, 19, 20], demand-response [12, 21, 22], and hybrid power plants [9–11, 23–26]. Overall, following a scenario-based stochastic optimization framework, most authors incorporate risk measures, e.g., the conditional value-at-risk (CVaR), in the objective function to assign higher weights to the scenarios with lowest profits [27]. On the other hand, robust optimisation-based approaches allow the ability to adjust the conservativeness of the decisions by varying the budget

*Corresponding author
Email addresses: jeremie.bottieau@umons.be (Bottieau, J.), kenneth.bruninx@kuleuven.be (Bruninx, K.), anibal.sanjab@vito.be (Sanjab, A.), zacharie.degreve@umons.ac.be (De Grève, Z.), francois.vallee@umons.ac.be (Vallée, F.), jean-francois.toubeau@umons.ac.be (Toubeau, J-F.)

of uncertainty, i.e., the size and shape of the uncertainty set [11].

Currently, existing research efforts in performance maximization methods focus on evaluating the optimal risk policy through an offline process. This process entails varying the risk-aversion parameters – e.g., the confidence level of the CVaR metric or the budget of uncertainty – on extensive out-of-sample evaluations to select the optimal risk attitude – see e.g., [12, 22, 26, 28]. Such a computationally-intensive analysis is typically run once, which does not allow capturing the dependency of the optimal risk policy on the dynamically changing market operating conditions. In this sense, in all the aforementioned approaches, there is, to the best of the authors' knowledge, no systematic way to autonomously and dynamically adjust the selection of the risk-aversion parameters. Indeed, the out-of-sample analysis typically yields only time-invariant risk policies, even though the real-time power system conditions, and, as a result, the financial risk and the optimal risk policy, may significantly vary over time.

This paper aims at addressing this fundamental limitation by leveraging Machine Learning (ML) techniques to adjust, at each decision step, the risk policy of an actor based on the current state of its expected market outcomes. Practically, an ML-based module is designed to estimate the time-specific out-of-sample economic performance of different risk attitudes *based on past trading sessions*, allowing to determine autonomously the most optimal online risk attitude. In addition, our approach offers two further benefits: i) it can be applied for any risk-aware performance maximization methods– e.g., we consider both CVaR-based stochastic and robust optimization frameworks, and ii) it can be incorporated in all types of regular decision procedures of market participants, e.g., bidding strategies in the real-time markets [29].

In this paper, we apply our innovative automatic risk adjustment tool, considering a specific, high-risk market environment, i.e., the very-short term dispatch of a strategic market participant that exploits opportunities in the single price imbalance settlement mechanism (see Section 2.1). The risk policy is crucial in this application since the system conditions are highly volatile and difficult to predict. In addition, actors are exposed to significant financial penalties in case of sub-optimal decisions. In this regard, we extend our bi-level framework considering offline risk policies [30] to an online risk-adjusted framework, which considers both CVaR-based stochastic and robust optimization frameworks.

The contribution of our proposed automatic risk adjustment tool applied within the single price imbalance framework is, hence, threefold:

1. We leverage the self-learning abilities of ML techniques to dynamically and preemptively adjust the risk policy of an actor based on the current state of its expected market outcomes. The risk policy is progressively updated and improved based on past trading sessions. We illustrate the effectiveness of this approach in a detailed case study using data from a real-life power system. We focus on short-term electricity markets, for CVaR-based stochastic and robust optimization problems, combined with offline and online risk-policy based strategies.

2. We establish an extensive benchmark of competitive ML-based techniques (neural networks, ensemble trees and K-nearest neighbours) for the online selection of an optimal risk policy based on a realistic case study.

3. The implementation of our theoretical models on actual electricity market settings corroborate the key goal of the single price imbalance settlement mechanism, by reducing the system imbalance, and consequently limiting corrective actions at the real-time balancing stage. Indeed, the obtained results show that the actor increases its operating profit, while the imbalance of the power system is reduced.

Our results show that the online risk policies supported by ML tend to exhibit better performance than the optimal offline choice, especially in the context of robust optimization problems. Moreover, the k-nearest neighbours method emerges from the ML benchmark as a suited candidate for selecting the risk policy, balancing simplicity and profitability.

The rest of the paper is organized as follows. Section 2 explains the participation of a market actor in the European single price imbalance settlement and describes the proposed automatic risk adjustment tool within its decision-making process. This automatic risk adjustment tool is further detailed in Section 3. Then, Section 4 provides the mathematical formulations of the subsequent risk-aware performance maximization methods. We present and compare the different offline and online risk policies in Section 5. Finally, conclusions are presented in Section 6.

2. Market and Decision-Making Frameworks

In Section 2.1, we present our electricity market application, which is the near real-time decision problem faced by a market player that has the ability to deviate from its position in day-ahead electricity markets to support the real-time system balancing within a single imbalance pricing scheme. These intentional deviations allow providing balancing services not defined in the standard balancing products. Then, Section 2.2 introduces the associated risk-aware decision-making process of the market player, with an emphasis on the role of the proposed automatic risk adjustment tool.

2.1. Market Application

Recent European directives have increasingly encouraged the harmonization towards a single pricing rule across

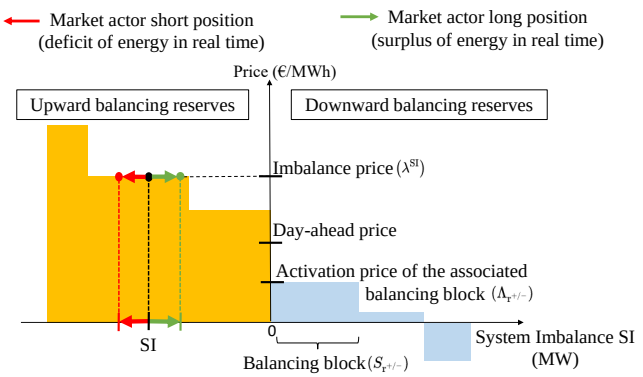


Figure 1: Impact of the imbalance position of a market actor on the single imbalance pricing mechanism in case of generation shortage.

Europe, in which all positive and negative imbalances of market actors are settled at a unique price [31]. This mechanism provides near real-time financial incentives for imbalances of market actors that restore the system balance, while penalizing those that enforce the imbalance of the system. In fact, as part of the imbalance settlement process, each agent with an unbalanced portfolio is charged with an imbalance fee, which depends on the actual grid balancing costs for restoring the system balance. It allows the Transmission System Operator (TSO) to recover the costs from the activation of balancing reserves, while transferring the financial risk to the market actors.

We define the system imbalance (SI) by the aggregated (net) imbalance position of all other market actors. A negative system imbalance defines a generation shortage in the system, while a positive system imbalance reflects a generation surplus. The TSO restores the system imbalance by activating pre-allocated operating reserve capacity and remunerates actors providing that capacity through the balancing market. The imbalance price (λ^{SI}) of each imbalance period, typically 15 minutes, is based on the marginal price associated with the most expensive activated upward reserve provider or least expensive activated downward reserve provider.

In case of generation shortage (SI<0), the TSO activates upward balancing reserves, which drives the imbalance price towards a high price regime (typically exceeding the day-ahead market prices). In contrast, if the system is in generation surplus (SI>0), downward balancing reserves are activated, thereby lowering the imbalance price regime (typically leading to an imbalance price below the day-ahead market prices). Depending on the system imbalance and the market participant's own imbalance position, a market actor faces a financial payoff or penalty. This is illustrated in Fig. 1 in case of generation shortage. A short position of the agent aggravates the shortage of the system. Hence, the market actor in this position would have to pay an expensive price to the TSO for the additional activation of upward balancing reserves. If the agent has a long position, its imbalance decreases the overall balancing needs. In that case, the actor would receive an attractive price for each 'excess' MWh. Similarly, when the system experiences a generation surplus (SI>0), this mechanism incentivizes market actors to adopt a short position.

The imbalance position of the market participant may also imply a regime-switching behaviour of the imbalance price. Indeed, under the system conditions in Fig. 1, a too long position of the market actor could change the sign of the SI. By adopting such an excessive long position, the market actor would be remunerated at a very low, even negative, imbalance price for its surplus of energy. These imbalance price regime switching effects, which negatively impact the participation of the market actor, can be hedged using a bi-level approach, which allows capturing the impact of the market actor's actions on the price formation of the imbalance settlement market.

The publication of transparent real time market data is crucial for reaching the full benefit of this single pricing scheme. For example, in this regard, the Belgian TSO provides day-ahead information with near real-time updates on the available balancing offers (capacity and activation costs) and the system imbalance position for each 15-min period [32]. Hence, different levels of balancing prices ($\Lambda_{r^{+,-}}$) corresponding to the activation of the balancing blocks ($S_{r^{+,-}}$) can be retrieved from these published data to reconstruct the imbalance pricing mechanism (as showcased in Fig. 1), and, thus, the associated imbalance price [33, 34].

2.2. Decision-Making Process

When optimizing its imbalance position, the market player can make use of a risk-aware stochastic decision support tool in its quarter-hourly routine; see Fig. 2. Indeed, its use is motivated by the significant variability and uncertainty exhibited by the system imbalance, arising from, e.g., renewable-based or load forecasting errors. Hence, based on a risk attitude ϵ, the decision support tool aims at providing an optimal imbalance position for the market player, while indicating the associated in-sample objective outcome $f_{\epsilon}^{\mathrm{IS}}$ for the next market period. As shown in Fig. 2, we further support the market player in its decision-making process by adding our automatic risk adjustment tool. This tool relies on a ML-based approach to adjust the risk attitude of the market player; see more details in Section 3. The selected risk attitude is then used in the subsequent risk-aware stochastic optimization tool to calculate its optimal position in the imbalance settlement market.

Once the imbalances of all market players are settled, an out-of-sample objective outcome $f_{\epsilon}^{\mathrm{OS}}$ can be generated in an ex-post analysis by confronting the optimised imbalance position of the market actor at a given risk attitude ϵ with the actual realization of the system imbalance. This computation can stem from real-life market outcomes or complex market simulators. The latter option is used in this paper (as shown in Fig. 2).

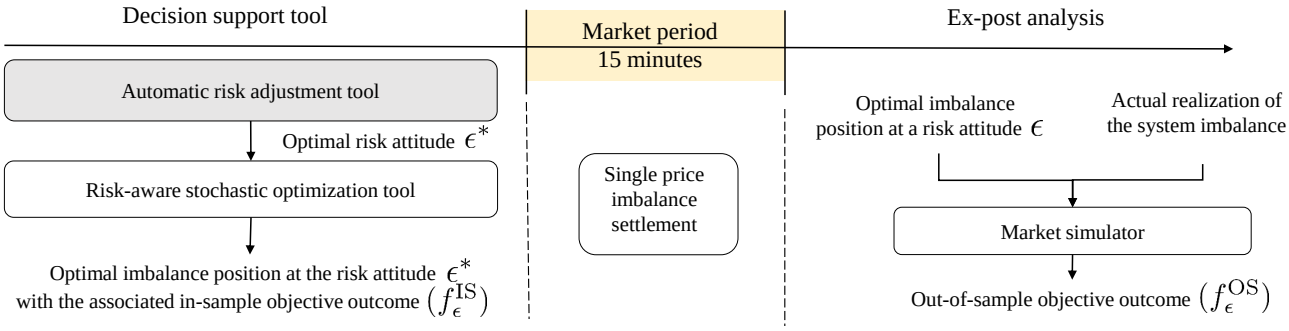


Figure 2: The integration of the proposed automatic risk adjustment tool within the decision-making process of an actor participating in the single price imbalance settlement. The decision support tool is run sequentially (96 times a day) at the start of each 15 minutes imbalance settlement period.

3. Automatic Risk Adjustment Tool

There exists inevitably a gap between an in-sample objective outcome (f_ϵ^{IS}) provided by a model, i.e., the optimal objective value of a risk-aware stochastic optimization tool at a risk-aversion parameter ϵ, and the corresponding out-of-sample objective outcome (f_ϵ^{OS}) obtained when the actual realization of the uncertainty is revealed. The gap between in-sample and out-of-sample objective outcomes can arise from either a misrepresentation of the uncertainty [35] or a simplified representation of the market environment. Indeed, a risk-aware stochastic optimization tool is generally composed of two stages (respectively described in Sections 4.1 and 4.2 for our application): (i) an uncertainty model, which allows identifying a probabilistic representation of the future possible realizations of random variables, and (ii) a risk-aware stochastic optimization model, which mathematically expresses the market environment in which the actor operates, and generates optimal decision outcomes at a pre-defined risk attitude, based on the prior representation of uncertainties.

Hence, the in-sample objective value, provided by the risk-aware stochastic optimization tool, channels the understanding of both the uncertainty and risk-aware optimization models at a given risk-aversion parameter and at a specific decision stage. The purpose of our proposed ML-based module is to add an additional learning stage on top of the in-sample objective value to provide an early estimate of the out-of-sample objective value. Practically, the module is designed as an ML-based regression model, which predicts online approximations of the out-of-sample objective outcomes, based on which the most suited risk-aversion parameter and, thus, the optimal risk-aware decision variables are computed.

3.1. Training Procedure

To capture the misspecification of the uncertainty and risk-aware optimization models, the ML-based module must be firstly trained on a database $\mathbf{D}$ that maps the in-sample objective outcomes with the actual ones. As shown in Fig. 3, this necessary learning stage is represented by Step ($\mathbf{A}$) which generates a database $\mathbf{D}$, i.e., the inputs $\mathcal{F}_n^{\mathrm{IS}} = \{f_{\epsilon,n}^{\mathrm{IS}},\ \forall \epsilon \in \mathcal{E}\}$ and outputs $\mathcal{F}_n^{\mathrm{OS}} = \{f_{\epsilon,n}^{\mathrm{OS}},\ \forall \epsilon \in \mathcal{E}\}$, on the $n = 1, ..., N^{\mathbf{D}}$ anterior time steps, whose relationship must be learnt for different risk-aversion parameters $\epsilon \in \mathcal{E}$. The out-of-sample objective outcomes $\mathcal{F}_n^{\mathrm{OS}}$ can be generated in an ex-post analysis for each risk-aversion level ϵ. Then, in Step ($\mathbf{B}$), the objective is to optimize the parameters θ of the ML model g_θ such that we accurately map the outputs $\mathcal{F}_n^{\mathrm{OS}}$ to the given inputs $\mathcal{F}_n^{\mathrm{IS}}$:

$$\theta^* = \arg\min_\theta \sum_{n=1}^{N^{\mathbf{D}}} L(g_\theta(\mathcal{F}_n^{\mathrm{IS}}), \mathcal{F}_n^{\mathrm{OS}}), \tag{1}$$

where $L(.,.)$ is a user-defined loss function that quantifies how well the model fits the data.

3.2. Online Risk Attitude Selection

In the online stage (after the model has been trained), the ML-based module is used to predict approximations of the out-of-sample objective outcomes $\hat{\mathcal{F}}_{\mathrm{new}}^{\mathrm{OS}} = \left\{\hat{f}_{\epsilon,\mathrm{new}}^{\mathrm{OS}},\ \forall \epsilon \in \mathcal{E}\right\}$ for different risk-aversion parameters $\epsilon \in$

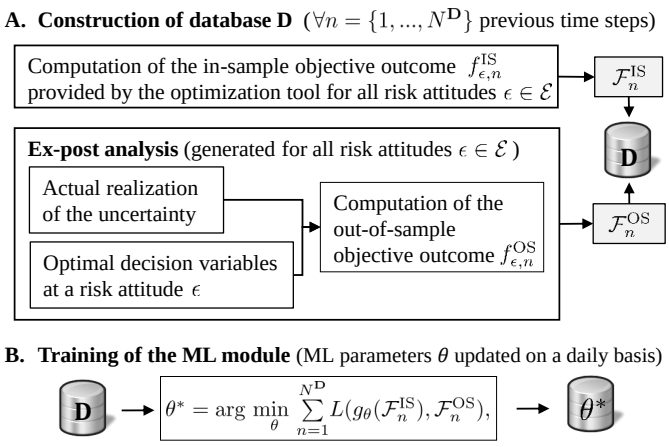


Figure 3: The training procedure of the ML-based module within a supervised framework.

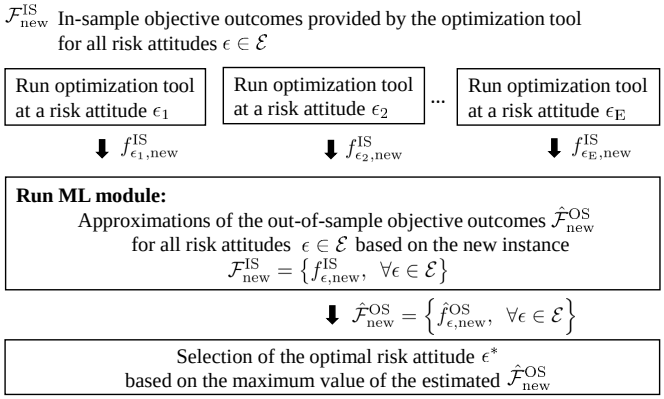


Figure 4: Illustration of the automatic risk adjustment approach. In the online phase, when a new decision stage is performed, the trained ML-based module estimates the out-of-sample objective outcomes $\hat{\mathcal{F}}^{\text{OS}}_{\text{new}}$ based on the new in-sample objective values $\mathcal{F}^{\text{IS}}_{\text{new}}$, provided by the risk-aware optimization tool for different risk attitudes $\epsilon \in \mathcal{E}$. Then, the optimal risk attitude ϵ^* is selected based on the maximum value of $\hat{\mathcal{F}}^{\text{OS}}_{\text{new}}$.

$\mathcal{E}$ on a new instance $\mathcal{F}^{\text{IS}}_{\text{new}} = \left\{f^{\text{IS}}_{\epsilon,\text{new}},\ \forall \epsilon \in \mathcal{E}\right\}$. Then, the most suited risk-aversion parameter, and consequently the most optimal risk-aware decision variables, can be selected based on the maximum value of the estimated out-of-sample objective outcomes. This process is showcased in Fig. 4.

3.3. Machine-Learning Models

In this paper, five ML models are assessed and compared, including a linear model (LR), a shallow multi-layer perceptron (1-MLP), random forest (RF), gradient boosted decision trees (GBDT) and the k-nearest neighbours (k-NN), for the $E = |\mathcal{E}|$ risk attitudes, where $|\mathcal{E}|$ stands for the cardinality of the set $\mathcal{E}$. The first four models, i.e., LR, 1-MLP, RF and GBDT, represent a snapshot of well-established ML techniques, as exemplified by their common use in contests such as the Global Energy Forecasting Competition [36]. The k-NN method provides another simple and yet competitive ML alternative, which differs from the four other methods by the fact that no model parameters θ need to be trained. In complement, the five ML models have already shown a high accuracy for approximating the objective outcomes of real-time operation processes [37, 38].

The LR model g_θ is simply expressed as $\mathbf{A}\mathcal{F}^{\text{IS}}_{\text{new}} + \mathbf{b}$, where the model parameters θ (to be optimized) are the slopes $\mathbf{A}_{(E\times E)}$ and intercepts $\mathbf{b}_{(E\times 1)}$. The loss function $L(.,.)$ is based on the least-squares criterion.

The 1-MLP model is the traditional architecture of neural networks, where the input information $\mathcal{F}^{\text{IS}}_{\text{new}}$ is propagated through a hidden layer containing H processing units (neurons). Each neuron consists of the application of a non-linear activation function $k_h(.)$, e.g., the rectified linear unit (ReLU) function, on the weighted sum of the inputs:

$$y_h = k_h\left(\sum_{\epsilon\in\mathcal{E}} w^{\text{in}}_{h\epsilon} f^{\text{IS}}_{\epsilon,\text{new}}\right) \qquad \forall h \in H. \tag{2}$$

Then, the output vector $\hat{\mathcal{F}}^{\text{OS}}_{\text{new}}$ is given by the application of the linear function $k_o(.)$ on the hidden units $y_.$:

$$\hat{f}^{\text{OS}}_{\epsilon,\text{new}} = k_o\left(\sum_{h=1}^{H} w^{\text{out}}_{\epsilon h} y_h\right) \qquad \forall \epsilon \in \mathcal{E}. \tag{3}$$

Using the mean square error as loss function $L(.,.)$, the backpropagation algorithm can be used to optimize the network weights $w^{\{\text{in,out}\}}_.$.

RF and GBDT are ensemble methods based on decision trees. A decision tree $\mathcal{T}$ is a hierarchical model combining a sequence of simple logical tests, e.g., the comparison between a numeric input and a threshold value, that recursively splits the paired dataset $\mathbf{D}$ into distinct (smaller) subsets. Hence, the parameter θ_m of a binary decision tree at a node m is the split s_m that yields locally the *sharpest* partition of the subset $\mathbf{D}_m$ at the node m into a left and right nodes, respectively noted m_l and m_r. The 'sharpness' of a split s_m at a node m can be measured by the decrease of an impurity function $i(.)$:

$$\Delta i(s_m) = i(m) - \frac{N^{\mathbf{D}_{m_l}}}{N^{\mathbf{D}_m}} i\left(m_l\right) - \frac{N^{\mathbf{D}_{m_r}}}{N^{\mathbf{D}_m}} i\left(m_r\right), \tag{4}$$

where $N^{\mathbf{D}_{\{m,m_l,m_r\}}}$ are the numbers of instances contained in $\mathbf{D}_{\{m,m_l,m_r\}}$, respectively. In a regression-based framework, $i(.)$ is usually derived from the variance:

$$i(m) = \frac{1}{N^{\mathbf{D}_m}} \sum_{f^{\text{OS}}_{\epsilon,n}\in\mathbf{D}_m} \left(f^{\text{OS}}_{\epsilon,n} - \overline{f^{\text{OS},\mathbf{D}_m}_{\epsilon}}\right)^2, \tag{5}$$

where $\overline{f^{\text{OS},\mathbf{D}_m}_{\epsilon}}$ is the mean value of the outputs at the node m for a risk-aversion parameter ϵ. Such a $i(.)$ results, thus, in assigning, at each terminal node z, the average value of the corresponding subset of outputs $f^{\text{OS}}_{\epsilon,n} \in \mathbf{D}_z$. Hence, the prediction of a new instance consists in (i) identifying the terminal node to which it belongs and (ii) retrieving the forecast value assigned to the corresponding node:

$$\hat{f}^{\text{OS}}_{\epsilon,\text{new}} = \mathcal{T}_\epsilon\left(\mathbf{S}_\epsilon, \mathcal{F}^{\text{IS}}_{\text{new}}\right) \qquad \forall \epsilon \in \mathcal{E}. \tag{6}$$

where $\mathbf{S}_\epsilon$ define the splits of the decision tree $\mathcal{T}_\epsilon$ at a risk-aversion parameter ϵ.

In RF, an ensemble of N^{RF} decision trees are independently grown using the aforementioned approach, with the particularity that each split of each tree is constructed based on a random subsample of the data set and a random subset of features, allowing the reduction of the variance of the entire model. A new prediction is obtained by averaging the outcomes of each decision tree:

$$\hat{f}^{\text{OS}}_{\epsilon,\text{new}} = \frac{1}{N^{\text{RF}}} \sum_{i=1}^{N^{\text{RF}}} \mathcal{T}_{\epsilon,i}\left(\mathbf{S}_{\epsilon,i}, \mathcal{F}^{\text{IS}}_{\text{new}}\right) \qquad \forall \epsilon \in \mathcal{E}. \tag{7}$$

As for the GBDT method, the trees are not built independently but rather in an additive fashion. For the p-th iteration, the GBDT model g_θ is written as:

$$g_\theta^p(\mathcal{F}_{\text{new}}^{\text{IS}}) = g_\theta^{p-1}(\mathcal{F}_{\text{new}}^{\text{IS}}) + \alpha \mathcal{T}^p\left(\mathcal{F}_{\text{new}}^{\text{IS}}\right), \tag{8}$$

where α is the learning rate. At each stage, the additional tree $\mathcal{T}^p$ is updated to minimize the residuals of the $p-1$ previously generated trees:

$$\mathcal{T}^p = \arg\min_{\mathcal{T}} \sum_{n=1}^{N^{\mathbf{D}}} L\left(f_{\epsilon,n}^{\text{OS}}, g_\theta^{p-1}\left(\mathcal{F}_n^{\text{IS}}\right) + \mathcal{T}\left(\mathcal{F}_n^{\text{IS}}\right)\right), \tag{9}$$

where $L(.,.)$ is the mean square error function, and $N^{\mathbf{D}}$ is the number of instances contained in the database $\mathbf{D}$. The overall approach is similar to the gradient descent algorithm in which the added tree $\mathcal{T}^p$ is optimized to leverage the prediction errors of its predecessors.

The k-NN algorithm relies on the concept of learning by analogy: based on the creation of a paired dataset, the prediction of a new instance is carried out on the average value of the k nearest neighbours in the dataset where the closeness condition is determined based on a distance metric (e.g., the euclidean distance). For a new instance, the k-NN estimate is given by:

$$\hat{\mathcal{F}}_{\text{new}}^{\text{OS}} = g(\mathcal{F}_{\text{new}}^{\text{IS}}) = \frac{1}{k}\sum_{n=1}^{N^{\mathbf{D}}} x_n \mathcal{F}_n^{\text{OS}}, \tag{10}$$

where $x_n \in \{0,1\}$ depending on whether or not $\mathcal{F}_n^{\text{IS}}$ is among the k-nearest neighbours of $\mathcal{F}_{\text{new}}^{\text{IS}}$ and $N^{\mathbf{D}}$ is the number of instances within the database $\mathbf{D}$.

In practice, new information is continuously revealed and the training procedure must be updated over time. In this context, an appealing feature of k-NN is that it seamlessly supports online updates [39]. Its prediction performance is simply improved by adding new instances to the dataset, while the other ML models need the additional Step (**B**) to calculate their optimal ML parameters θ^*. In our case study focusing on single price imbalance settlement markets, the parameters θ^* are updated on a daily basis resulting in a balance between computational burden and model precision.

In general, each of the ML models are characterized by hyper-parameters, which are (task-dependent) parameters reflecting the complexity of the model, e.g., the number of hidden units of neural networks, the number of basic trees in RF or the number of K neighbours. These values are typically estimated using a grid search approach, which is embedded within a cross validation scheme [40].

4. Risk-aware Stochastic Optimization Tool

In this section, we present the risk-aware stochastic optimization tool of our electricity market application. First, we propose two methods for modeling the uncertainty of the system imbalance: (i) a deterministic and set-based method, which is used for the robust optimization framework, and (ii) a scenario-based method, which is applied in the CVAR-based stochastic optimization framework. Both methods rely on a probabilistic forecasting tool for predicting the conditional quantiles of the system imbalance. Second, we introduce the mathematical formulation of both risk-aware stochastic optimization models, i.e., the robust optimization and CVaR-based stochastic frameworks, solved by the market actor. The optimization model relies on a bi-level approach to endogenously anticipate the imbalance price distribution based on the system imbalance uncertainty and the optimized imbalance position of the market player. This (single time-step) optimization framework is run sequentially (96 times a day) at the start of each 15 minutes imbalance settlement period, by exploiting the last measured values of the system imbalance communicated by the system operator.

4.1. System Imbalance Uncertainty Modeling

We leverage our probabilistic forecasting tool from [30], which relies on a sequence-to-sequence recurrent neural network architecture, to forecast the future system imbalance. The uncertainty is endogenously captured by training the forecaster using the quantile regression method. The inputs of the forecaster are, among others, lagged measurements of the system imbalance, electrical load and power production, day-ahead forecasts of renewable generation and electrical load, and schedules of conventional generation (which are publicly available on the Belgian TSO's website [32]). On the other hand, the outputs of the forecaster are the q-quantiles $\hat{\text{SI}}^{(q)}$ for $q \in \mathcal{Q} = \{0.01, 0.05, 0.25, 0.5, 0.75, 0.95, 0.99\}$, i.e., $\mathbb{P}(\text{SI} \leq \hat{\text{SI}}^{(q)}) = q$. Hence, following [41], for each quarter hour, an empirical cumulative distribution function (ECDF) is estimated based on the resulting discrete set of q-quantiles through cubic spline interpolation.

For the **robust optimization framework**, we define an uncertainty set $\mathcal{U}$ as a box bounded by a symmetric pair of q-quantiles from the empirical ECDF. Hence, symmetric lower and upper bounds $\left\{\hat{\text{SI}}^{(q)}, \hat{\text{SI}}^{(1-q)}\right\}$ can be selected, each combination ensuring a certain probabilistic guarantee that the future system imbalance is realized within the uncertainty set:

$$\mathcal{U} = \left\{\text{SI} \in \mathbb{R}^n \ : \ \hat{\text{SI}}^{(q)} \leq \text{SI} \leq \hat{\text{SI}}^{(1-q)}\right\}, \tag{11}$$

In this approach, the risk-aversion parameter $\epsilon = 1-2q$ determines the confidence interval of the uncertainty set. Hence, a larger ϵ yields a larger (and more conservative) uncertainty set, while a smaller value leads to a smaller uncertainty set.

For the **CVaR-based stochastic optimization framework**, following [41], a Monte Carlo approach is applied to generate scenarios from the ECDF. This ECDF is derived from the outputs of our state-of-the-art probabilistic forecasting tool. For each optimization period, a

set Ω of $N = 100$ scenarios is constructed, each of them associated with a probability of occurrence $p = 1/N$. In our experiments, increasing further the number of scenarios do not improve the results of the CVaR-based stochastic optimization framework. The financial risk associated with each scenario is then controlled through the CVaR metric (see Section 4.2).

4.2. Risk-Aware Stochastic Optimization Model

The participation of the actor in the single price imbalance settlement is modeled in a bi-level optimization framework in which the market actor anticipates the price of the imbalance settlement mechanism (λ^{SI}) by having at its disposal (i) a list, typically made available by the TSO [32], of the submitted balancing offers and the associated activation prices, and (ii) its own probabilistic description of the future system imbalance volume. The upper level (12a)-(12b) aims at defining the optimal imbalance position ($e^{\text{imb},+/-}$) of the market actor and the lower level (12c)-(12f) performs the single imbalance pricing mechanism. In a **deterministic setting**, the problem reads as:

$$\max_{\substack{e^{\text{imb},+/-},\lambda^{\text{SI}} \\ \geq 0}} \; f^{\text{IS}}\left(e^{\text{imb},+/-}, \lambda^{\text{SI}}\right) \tag{12a}$$

$$\text{s.t.} \;\; e^{\text{imb},+/-} \in \Pi_{\text{UL}} \tag{12b}$$

$$\min_{\Theta_{\text{LL}}=\{s_{r^+/-}\}} \mathcal{C}(\Theta_{\text{LL}}) = \sum_{r^+\in R^+} \Lambda_{r^+} s_{r^+} - \sum_{r^-\in R^-} \Lambda_{r^-} s_{r^-} \tag{12c}$$

$$\text{s.t.} \sum_{r^+\in R^+} s_{r^+} - \sum_{r^-\in R^-} s_{r^-} = - (e^{imb,+} - e^{imb,-}) - \hat{\text{SI}} : \lambda^{SI} \tag{12d}$$

$$s_{r^+} \leq S_{r^+} : \mu_{r^+} \qquad \forall r^+ \in R^+ \tag{12e}$$

$$s_{r^-} \leq S_{r^-} : \mu_{r^-} \qquad \forall r^- \in R^- \tag{12f}$$

$$-s_{r^+} \leq 0 : \gamma_{r^+} \qquad \forall r^+ \in R^+ \tag{12g}$$

$$-s_{r^-} \leq 0 : \gamma_{r^-} \qquad \forall r^- \in R^- \tag{12h}$$

where $\Theta_{\text{LL}} = \{s_{r^+}, s_{r^-}\}$ are the primal variables of the lower-level problem, while the set $\{\mu_{r^+}, \mu_{r^-}, \lambda^{\text{SI}}, \gamma_{r^+}, \gamma_{r^-}\}$ are the dual variables. The objective function (12a), maximizing the profit of the market actor, is computed as:

$$f^{\text{IS}}\left(e^{\text{imb},+/-}, \lambda^{\text{SI}}\right) = (\lambda^{SI} - C^+)e^{\text{imb},+} - (\lambda^{SI} - C^-)e^{\text{imb},-} \tag{13}$$

where C^+ and C^- define the cost structure of the asset.

Practically, the market actor is incentivized to adopt a long position $e^{\text{imb},+}$ when $\lambda^{SI} > C^+$, and to favor a short position $e^{\text{imb},-}$ when $\lambda^{SI} < C^-$. Constraint (12b) ensures that these decisions comply with the technical margins Π_{UL} of the agent.

The lower-level reflects the costs minimization problem (12c) in which the TSO carries out the merit-order-based activation of operating reserves (where more economic offers are activated first). The offers $r^+ \in R^+$ and $r^- \in R^-$ are respectively activated at a price Λ_{r^+} and Λ_{r^-} to compensate the negative and positive imbalances. The imbalance price λ^{SI} is the price associated with the last (marginally activated) offer, which is endogenously obtained from the dual variable of the constraint (12d). The latter guarantees that the activated amount of reserves exactly offsets the anticipated imbalances caused by all other actors ($\hat{\text{SI}}$), while accounting for the strategic participation of the market participant ($e^{\text{imb},+/-}$). Finally, the set of constraints (12f) ensures that the activated balancing volumes s_{r^+} and s_{r^-} do not violate the energy limits (i.e., capacity S_{r^+} and S_{r^-} offered at an earlier stage).

A **robust optimization-based** equivalent of problem (12) reads as:

$$\max_{\substack{\sigma, e^{\text{imb},+/-}, \lambda_v^{\text{SI}} \\ \geq 0}} \; \sigma \tag{14a}$$

$$\text{s.t.} \quad \sigma \leq f^{\text{IS}}\left(e^{imb,+/-}, \lambda_v^{\text{SI}}\right), \quad \forall v = \{V_1, V_2\} \in \mathcal{U}, \tag{14b}$$

$$(12b) \;\; - \;\; (12f), \quad \forall v = \{V_1, V_2\} \in \mathcal{U}, \tag{14c}$$

where $\mathcal{U}$ is the uncertainty set, which can be tailored to better reflect the risk behavior of the agent (Section 4.1). In this robust approach, the objective function (14a) immunizes the operation strategy against the worst-case realization of the system imbalance contained in the uncertainty set $\mathcal{U}$. Practically, an auxiliary variable σ is introduced for representing the worst-case profit $f^{\text{IS}}(.)$ through the additional constraint (14b), and, the continuity of the uncertainty set $\mathcal{U}$ is managed by considering only the finite number of vertices $v = \{V_1, V_2\}$ contained in $\mathcal{U}$ [30].

An alternative approach is to rely on a scenario-based framework, in which risk can be incorporated through the CVaR measure. This allows quantifying the level of trading risk, and to reduce the volatility of the profit among the set of scenarios Ω. At the given confidence interval $\epsilon \in [0, 1]$ (representing a risk-aversion parameter), the CVaR_ϵ is defined as the expected profit of the $(1-\epsilon)\times 100\%$ worst scenarios. The **CVaR-based stochastic program** can then be defined as:

$$\max_{\substack{\zeta, \eta_w, e^{imb,+/-}, \lambda_w^{\text{SI}} \\ \geq 0 \quad \geq 0}} \; \zeta - \frac{1}{1-\beta} \sum_{w\in\Omega} \pi_w \eta_w \tag{15a}$$

$$\zeta - f^{\text{IS}}\left(e^{imb,+/-}, \lambda_w^{\text{SI}}\right) \leq \eta_w, \quad \forall w \in \Omega, \tag{15b}$$

$$(12b) \;\; - \;\; (12f), \quad \forall w \in \Omega, \tag{15c}$$

where ζ is the Value-at-Risk (VaR), i.e., the (1-ϵ)-quantile of the profit distribution of scenarios. The non-negative auxiliary variable η_w corresponds to the difference between the VaR ζ and the market actor profit $f^{\text{IS}}(.)$ if it is positive. A larger risk-aversion parameter ϵ leads to more conservative decisions (as only few worst scenarios are considered), while a smaller value entails more risk-neutral decisions (as the set of scenarios is widened). Note that a trade-off between the expected profit and the CVaR metric, i.e., a weighted average of both in the objective function, is also doable.

In order to solve the resulting bi-level problems (14) and (15) by off-the-shelf optimization solvers, they can be converted into an equivalent mixed-integer linear programming (MILP) formulation using the following steps [30]: (i) the convex (linear and continuous) lower-level problem is replaced by its Karush-Khun-Tucker (KKT) optimality conditions, (ii) the non-linearities of the complementarity conditions within the KKT conditions are equivalently expressed as a set of mixed-integer linear constraints using a Big-M approach, and (iii) the bilinear term in the objective function is replaced by the related equivalent linear expression from the strong duality equation. In step (ii), the selection of Big-Ms values can be a challenging task, since improper values may lead to highly suboptimal solutions [42]. Here, Big-Ms values are determined based on the economic or physical upper bounds of their associated variables (Appendix A). The resulting MILP formulation is solved by the traditional branch-and-cut algorithm. If one encounters computational issues, advanced decomposition techniques could be used to speed up the solution procedure [43].

5. Case Study

The probabilistic prediction of the system imbalance is performed at the start of each quarter hour of the test set, i.e., the year 2018 including 35040 data points, using the Belgian data published in [44], in accordance with the methodology in [30]. The forecaster is trained using historical data from 2014 until end of December 2016, and is stabilized with regularization techniques (e.g., early stopping) using the year 2017 as a validation set to avoid overfitting.

All the market data required for the risk-aware optimization tool are also available in [44]. At each time step, the feasible region of the market actor portfolio Π_{UL} is constrained by (upward and downward) 30 MWh energy limits. The cost parameters C^+ and C^- are respectively set to 50 and 30 €/MWh ensuring a consistent imbalance position regarding the imbalance price [30]. For both risk-aware optimization methods, we consider 11 risk attitudes for the actor corresponding to the following $\epsilon \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.98\}$ risk-aversion parameters. This allows covering the entire range of possible risk attitudes, from the most risk-seeking ($\epsilon = 0$) to the most risk-averse ($\epsilon = 0.98$) ones.

To emphasize the learning ability of the ML tools supporting the selection of a risk policy, these models are trained only on the data available during the year 2018, which are also published in [44]. The first half of January is used to tune the hyper-parameters of the models using a grid search embedded within a cross validation scheme, and the latter are then recalibrated on a daily basis, using the new information revealed over time. The final configurations of the five ML regression models (along with their search spaces) are:

- An LR model.
- A 1-MLP model with $H = 20$ neurons using early stopping, and rectified linear units (ReLU) as activation functions. The search range of H was $\{10, 20, 50, 100\}$.
- An RF model with $N^{RF} = 100$ and a maximum depth of 5. We have varied the maximum depth between $\{3, 5, 8\}$ for the grid search.
- A GBDT model, in which $\alpha = 0.05$, the maximum depth is 3, and the number of iterations is determined by using early stopping. The search range of α was $\{0.3, 0.2, 0.1, 0.05\}$, and we also varied the maximum depth between $\{3, 5, 8\}$
- A k-NN model with $k = 1000$, where the search range of k was $\{1, 100, 500, 1000, 1250\}$.

Interestingly, it should be noted that the hyperparameter k can be naturally interpreted. If $k = 1$, the selection of the risk policy is only based on the ex-post economic performance of the closest instance within the database (where the closeness condition is computed based on the in-sample objective values using the euclidean distance). On the other hand, if k is set to the number of instances within the database, the risk policy's selection of the k-nn method would coincide with the one obtained using the traditional offline approach.

The results are computed over the period spanning from 15 January until the end of December 2018. The notation RO-Q stands for the robust optimization, while SP-CVaR denotes the CVaR-based stochastic optimization. The time step Δt is equal to an imbalance settlement period, i.e., 15 minutes. The performance of the proposed methodology is principally evaluated through one indicator, which is averaged over all the quarter hours of the test set: the actual profit $f^{\text{OS}}(.) = (\lambda^{SI} - C^+)e^{\text{imb},+} - (\lambda^{SI} - C^-)e^{\text{imb},-}$, where the imbalance price is computed via the ex-post single imbalance pricing mechanism. The ex-post single imbalance pricing mechanism is modelled using the cost minimization problem (12c) accounting for the actual system imbalance and the imbalance position of the market participant ($e^{\text{imb},+/-}$).

All the experiments have been conducted on an Intel Core i7-8850H CPU running at 2.60 GHz and with 16.0 GB of RAM, and coded in Python 3.6. The probabilistic forecasting model and the 5 ML models are implemented using the Scikit-learn and Keras packages. The MILP formulations for the risk-aware optimization models are written using the Pulp package and solved using the Gurobi 8.1.1 solver.

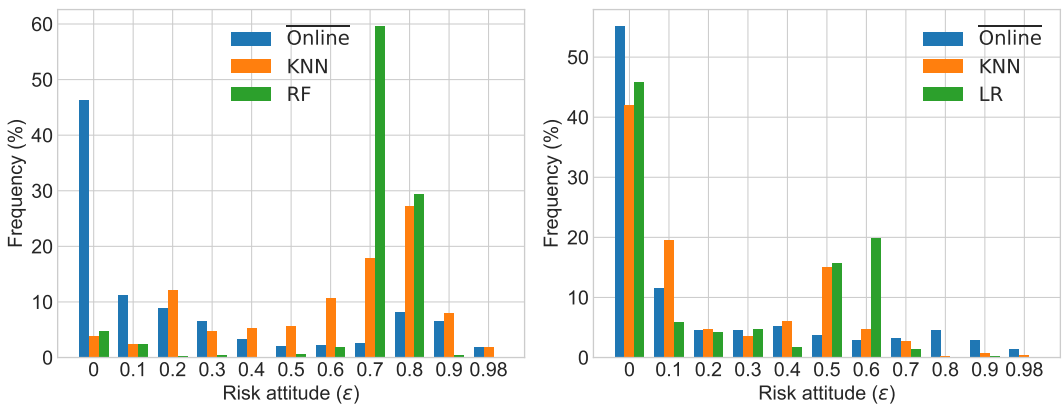
5.1. Automatic Risk Policies

Table 1 shows the evolution of the ex-post profits achieved by the different ML techniques over several time periods of the test set for RO-Q and SP-CVaR. These results are put into perspective with (i) Online: the ideal

Table 1: Rolling average of the out-of-sample profits $f^{OS}(.)$ over different time periods. The percentage values indicate the variation of the out-of-sample profit with respect to the $\overline{\text{Offline}}$ strategy.

[€/Δt]	$\overline{\text{Offline}}$	LR	1-MLP	RF	GBDT	k-NN	$\overline{\text{Online}}$
RO-Q	$\epsilon = 0.7$						
1 day	127.6	107.8 (▼15.5%)	51.1 (▼60%)	99 (▼22.4%)	102.4 (▼19.7%)	124.6 (▼2.3%)	317.8 (▲149%)
1 week	255	247.9 (▼2.8%)	271.6 (▲6.5%)	282.4 (▲10.7%)	255.9 (▲0.3%)	234.6 (▼8%)	633.6 (▲148.5%)
1 month	198.2	205.5 (▲3.7%)	202.1 (▲2%)	223.1 (▲12.6%)	201.8 (▲5%)	208 (▲5.1%)	528.4 (▲166.6%)
3 months	340.5	362.4 (▲6.4%)	349.2 (▲2.6%)	377 (▲10.7%)	367.5 (▲7.9%)	360.9 (▲6%)	696.4 (▲104.5%)
6 months	343.1	343.1 (▲0%)	341.5 (▼0.5%)	366.5 (▲6.8%)	360.4 (▲5%)	361.6 (▲5.4%)	693.3 (▲102.1%)
1 year	281.2	288.3 (▲2.5%)	294.1 (▲4.6%)	312.9 (▲11.3%)	312 (▲11%)	315.7 (▲12.3%)	622.4 (▲121.3%)
SP-CVaR	$\epsilon = 0.3$						
1 day	135.2	112.1 (▼17%)	131 (▼3.1%)	105.3 (▼22.1%)	127.6 (▼5.6%)	123.12 (▼8.9%)	289 (▲113.8%)
1 week	323.5	319.2 (▼1.3%)	312.4 (▼3.4%)	297.1 (▼8.1%)	286.6 (▼11.4%)	311.7 (▼3.7%)	594.1 (▲83.6%)
1 month	238.8	242 (▲1.3%)	240 (▲0.5%)	234.2 (▼1.9%)	229 (▼4.1%)	245.7 (▲2.9%)	496.9 (▲108.1%)
3 months	396.5	410.9 (▲3.6%)	398.5 (▲0.5%)	408.3 (▲3%)	396.9 (▲0.1%)	416.9 (▲5.1%)	672.3 (▲69.6%)
6 months	375.1	393.2 (▲4.8%)	380 (▲1.3%)	392.6 (▲4.7%)	382.7 (▲2%)	398.8 (▲6.3%)	662.1 (▲76.5%)
1 year	327	338.1 (▲3.4%)	327.9 (▲0.3%)	337.3 (▲3.1%)	333.3 (▲1.9%)	342.7 (▲4.8%)	515.7 (▲57.7%)

choice of the risk attitude at each quarter hour which is determined ex-post when the actual conditions are revealed and (ii) $\overline{\text{Offline}}$: the single choice of a risk-aversion parameter ϵ based on the ex-post economic performance of the different risk-aversion parameters $\mathcal{E}$ (from risk-seeking $\epsilon = 0$ towards risk-averse $\epsilon = 0.98$) during the first half of January. Over the 1-year period, the (omniscient) risk policies $\overline{\text{Online}}$ for each optimisation methodologies allow to improve the real revenues by over 121% and 57 % compared to their counterparts $\overline{\text{Offline}}$, stressing the relevance of adopting a dynamic risk attitude. As expected, the optimal risk policy $\overline{\text{Online}}$ for robust approach leads to greater ex-post profits than the SP-CVaR. This gap can be explained by the difference in the risk attitude at $\epsilon = 0$.



(a) RO-Q (b) SP-CVaR

Figure 5: Histograms of the risk-aversion parameters selected over the 1-year period.

At this risk-aversion parameter, the robust approach is purely deterministic and is extremely risk-seeking which leads to huge rewards in case of perfect information, while SP-CVaR is only risk-neutral and still takes into account extreme scenarios in its decision-making, preventing to fully leverage the added value of a perfect forecast.

Overall, the proposed online selection of the risk attitudes (guided by the best ML models) respectively improves the ex-post economic performance by 12.3% and 4.8% for RO-Q and SP-CVaR in comparison with their counterparts $\overline{\text{Offline}}$ over the 1-year period, which highlights the added value of better informing stochastic optimization tools with tailored risk-aversion parameters. In addition, the ability of ML to dynamically select the optimal risk parameter is analyzed in Fig. 5, by showing the frequency at which each risk-aversion parameter was chosen over the test set. For clarity, only the two best ML models are represented, i.e., k-NN and RF for RO-Q, and k-NN and LR for SP-CVaR.

Regarding RO approaches, the 1 year-based results in Table 1 show that an online support in the construction of the uncertainty set is a key element to fully leverage their potential. Indeed, we see that each ML model allows outperforming the economic gains of the strategy $\overline{\text{Offline}}$, emphasizing the importance of the size of the uncertainty set and its effect on the economic performance for robust-based optimization formulations. Logically, the simplest model LR gives less insights about the selection of the risk-aversion parameter than more advanced ML techniques. In particular, the single parametric models LR and 1-MLP show lower economic performances than the ensemble methods (RF and GBDT). Interestingly, the k-NN technique, which is simple and intuitive, shows a high suitability for our application as it outperforms the other ML techniques over the 1-year period. Additionally, Fig. 5a demonstrates that both ML techniques (i) select predominantly (45% and 89% of the time for respectively k-NN and RF) the optimal risk policies $\overline{\text{Offline}}$ at $\epsilon = \{0.7, 0.8\}$ and (ii) timely deviate 16% (k-NN) and 7.5% (RF) of the time towards riskier strategies at $\epsilon = \{0, 0.1, 0.2, 0.3\}$. These distributions explain the gap in ex-post profit with respect to the optimal (omniscient) strategy $\overline{\text{Online}}$, which requires a greedy approach that adopts more than 45% of the time the most risky strategy.

Concerning SP-CVaR, none of the ML-based strategies provide worse ex-post profit than the offline one over the period of one year, but their added value in terms of ex-post profits is lower in comparison with the robust case. Surprisingly, the simple LR model provides a better ex-post performance than the other (more complex) parametric ML models (1-MLP, RF and GBDT), while k-NN devises the most optimal online risk strategy. Fig. 5b) gives a first rationale behind such results by showing the vision of the risk strategies supported by ML: they adopt regularly the risk-neutral strategy (more than 40% of the time), while sparsely selecting risk-averse strategies at $\epsilon = \{0.5, 0.6\}$. This kind of risk management, recommended by the optimal one $\overline{\text{Online}}$, appears to be more challenging to implement for the more advanced ML models. Indeed, the latter are inclined to adopt a more conservative behaviour, which penalize them more severely in the SP-CVaR case.

Regarding the evolution over time of the RO-Q and SP-CVaR ex-post profits, it can be seen that the ML-based risk strategies have the ability to learn and capture the adequate risk attitude rapidly, which is reflected by the positive gains after only a day on the field. In the same vein, for both optimization tools, the machine learning approaches exhibit better economic performance than the offline one after only one month. However, the gap between ML-based strategies and the approach $\overline{\text{Offline}}$ does not widen over time after the first month of use. It tends to show that specific calibration methods for ML models have to be developed to take full advantage of the new data that are constantly revealed over time [45]. More particularly, the k-NN method is less affected by the calibration method. The results of Table 1 indicate that the k-NN method is emerging over time as the most suited method for supporting the automatic risk adjustment strategy of a market player.

The ML-based approach is more computational intensive than $\overline{\text{Offline}}$ strategies as it requires the prior computations of the in-sample profits $\mathcal{F}^{\text{IS}} = \{f_\epsilon^{\text{IS}}, \forall \epsilon\}$ as inputs. In our case study, the averaged computation time of an in-sample profit f_ϵ^{IS} is 0.02s for the robust optimization framework and 1.1s for the CVaR-based stochastic optimization framework with a duality gap of 1% imposed. Concerning the inference time of the LR, 1-MLP, RF, GBDT and k-NN methods, their averaged time for outputting the most optimal risk attitude are around 0.05ms, 27ms, 112ms, 5ms and 1ms, respectively. Overall, the prior computations of the in-sample profits $\mathcal{F}^{\text{IS}}$ are the most time consuming, however, this issue can be alleviated through parallel computing. There still exists a gap in computation times between RO-Q and SP-CVaR, e.g., RO-Q is 50 times faster than SP-CVaR in our application. If the computation time is a hard constraint, the RO-Q approach guided by the k-NN model provides a viable alternative to the $\overline{\text{Offline}}$ SP-CVaR approach. This allows reducing the ex-post profits differential between RO-Q and $\overline{\text{Offline}}$ SP-CVaR from 16.3% ($\overline{\text{Offline}}$ RO-Q) to 3.6% (RO-Q with k-NN model) over the entire year.

Finally, the impact of the imbalances of the agent on the performance of the TSO balancing dispatch procedure is studied in Table 2. To that end, the cost $\mathcal{C}^{os} = \sum_{r^+ \in R^+} \Lambda_{r^+} s_{r^+} + \sum_{r^- \in R^-} \Lambda_{r^-} s_{r^-}$ and the activated energy $V^{\text{act}} = \sum_{r^+ \in R^+} s_{r^+} + \sum_{r^- \in R^-} s_{r^-}$ of each quarter hour are retrieved from the ex-post analysis and are averaged over the test set. Results show an average drop of 20% and 40% for, respectively, the averaged activation of balancing energy and the costs, thereby reducing the TSO's corrective actions at the real-time balancing stage.

Table 2: TSO Balancing actions and cost for different risk strategies (the percentage values indicate the variation of the balancing actions and cost with respect to the non-participation of a market player).

	$\mathbb{E}\left(V^{\text{act}}\right)$ [MW/Δt]	$\mathbb{E}(\mathcal{C}^{os})$ [€/Δt]
No Participation	24.36	1350.11
RO-Q with k-NN	19.22 (▼21.1%)	782.34 (▼42%)
SP-CVaR with k-NN	19.39 (▼20.4%)	717.83 (▼46.8%)

5.2. *Offline Risk Policies*

Table 3 presents the out-of-sample performance of the three optimization tools for all possibilities ($E = 11$) of offline risk strategies over the entire test set.

Interestingly, the risk approach Offline for the SP-CVaR case showed in Table 1 has induced a miss-selection of the risk-aversion parameter in regard to the actual optimal offline strategy in Table 3, i.e., $\epsilon = 0.3$ instead of $\epsilon = 0$. It results in a decrease in the ex-post profits by 3.25%. This highlights the difficulty of arbitrary fixing an offline risk strategy having only a short-term vision on the out-of-sample economic performance of the risk-aware optimization tools.

It is also worth mentioning that the SP-CVaR approach yields the best performance for the risk-neutral strategy, which stems from two reasons: (i) the forecaster [30] has been tailored for such an application and it, thus, provides 'high-quality' scenarios, and (ii) based on these scenarios, this framework optimizes the profit in expectation, which is efficient for decision-making procedure that occur very regularly in time. In particular, SP-CVaR outperforms the robust optimization method, with a relative increase of the actual profits of around 20% on average at their optimal risk-aversion parameter. It should be stressed that the SP-CVaR guided by the k-nn model still outperforms the risk-neutral SP-CVaR.

6. Conclusion

In this paper, an automatic approach, aiming at continuously adjusting the conservativeness of the decisions arising from risk-aware optimization techniques, is proposed. This novel approach is applied on the very short-term dispatch of a market actor exploiting opportunities in the European single price imbalance settlement mechanism.

Extensive numerical analyses, using real-world market data from the Belgian power system over one year, demonstrates the ability of the proposed approach to achieve efficient online risk-adjusted strategies for robust-based and CVaR-based stochastic optimizations. More particularly, the k-NN technique has been identified as a suited ML candidate to support these risk-aware optimization methods for preemptively devising the risk attitude. More specifically, both RO-Q and SP-CVaR approaches guided by the k-nn model have presented promising results as this had led to a respective increase of 12.3% and 4.8% in the ex-post profits compared with their offline risk policy-based counterparts.

The integration of the proposed approach in another market framework deserves further research efforts. As an example, the application of this approach in the joint day-ahead bidding strategy in energy and reserve markets can bring valuable insights regarding how a market player adapts its risk strategy over multiple time steps.

7. Acknowledgements

This work is supported by the energy transition funds project 'EPOC 2030-2050' organised by the Belgian FPS economy, S.M.E.s, Self-employed and Energy.

8. Data Availability Statement

The data that support the findings of this study are openly available in [44].

Appendix A. Supplementary Mathematical Expression

After replacing the lower-level problem (12c) by its Karush-Kuhn-Tucker conditions and linearising the complementary slackness conditions using the big-M reformulation, an equivalent expression can be obtained with the following mixed-integer linear constraints:

$$0 \leq -s_{r^+} + S_{r^+} \leq z^1_{r^+} M^1_{r^+}, \qquad \forall r^+ \in R^+, \tag{A.1a}$$
$$0 \leq \mu_{r^+} \leq (1 - z^1_{r^+}) M^2_{r^+}, \qquad \forall r^+ \in R^+, \tag{A.1b}$$
$$0 \leq -s_{r^-} + S_{r^-} \leq z^1_{r^-} M^1_{r^-}, \qquad \forall r^- \in R^-, \tag{A.1c}$$
$$0 \leq \mu_{r^-} \leq (1 - z^1_{r^-}) M^2_{r^-}, \qquad \forall r^- \in R^-, \tag{A.1d}$$
$$0 \leq s_{r^+} \leq z^2_{r^+} M^3_{r^+}, \qquad \forall r^+ \in R^+, \tag{A.1e}$$
$$0 \leq \gamma_{r^+} \leq (1 - z^2_{r^+}) M^4_{r^+}, \qquad \forall r^+ \in R^+, \tag{A.1f}$$
$$0 \leq s_{r^-} \leq z^2_{r^-} M^3_{r^-}, \qquad \forall r^- \in R^-, \tag{A.1g}$$
$$0 \leq \gamma_{r^-} \leq (1 - z^2_{r^-}) M^4_{r^-}, \qquad \forall r^- \in R^-, \tag{A.1h}$$
$$\sum_{r^+ \in R^+} s_{r^+} - \sum_{r^- \in R^-} s_{r^-} = -(e^{imb,+} - e^{imb,-}) - \hat{\text{SI}}, \tag{A.1i}$$
$$\Lambda_{r^+} - \lambda^{\text{SI}} + \mu_{r^+} = \gamma_{r^+}, \qquad \forall r^+ \in R^+, \tag{A.1j}$$

Table 3: Average of the out-of-sample profits $f^{OS}_{\epsilon}(.)$ for the different offline risk strategies over the 1-year period.

Model Type	Unit [€/Δt]	Risk-aversion parameter (ϵ)										
		ϵ=0	ϵ=0.1	ϵ=0.2	ϵ=0.3	ϵ=0.4	ϵ=0.5	ϵ=0.6	ϵ=0.7	ϵ=0.8	ϵ=0.9	ϵ=0.98
RO-Q	$\mathbb{E}(f^{OS}_{\epsilon})$	124.1	185.1	227.4	251.8	263.4	270.6	275.4	281.2	280.5	221.2	100.5
SP-CVaR	$\mathbb{E}(f^{OS}_{\epsilon})$	338	335	331,9	327	315	300,1	292,1	283,2	244,2	175,7	129

M_j^i	$i=\{1,3\}$	$i=2$	$i=4$
$j=r_+$	S_{r^+}	$\max\limits_{r^+\in R^+}\ (\Lambda_{r^+})-\Lambda_{r^+}$	$-\min\limits_{r^-\in R^-}\ (\Lambda_{r^-})+\Lambda_{r^+}$
$j=r_-$	S_{r^-}	$-\min\limits_{r^-\in R^-}\ (\Lambda_{r^-})+\Lambda_{r^-}$	$\max\limits_{r^+\in R^+}\ (\Lambda_{r^+})-\Lambda_{r^-}$

Table A.1: big-Ms values

$$-\Lambda_{r^-}+\lambda^{\mathrm{SI}}+\mu_{r^-}=\gamma_{r^-}, \qquad \forall r^-\in R^-, \tag{A.1k}$$

where $M^1_{r^+}$, $M^2_{r^+}$,$M^3_{r^+}$, $M^4_{r^+}$, $M^1_{r^-}$, $M^2_{r^-}$, $M^3_{r^-}$, $M^4_{r^-}$ are large positive constants and $z^1_{r^+}$, $z^2_{r^+}$, $z^3_{r^+}$, $z^4_{r^+}$, $z^1_{r^-}$, $z^2_{r^-}$, $z^3_{r^-}$, $z^4_{r^-}$ are binary variables.

The big-Ms values used in our case studies are shown in Table A.1. They are determined based on the economic or physical upper bounds of their associated variables. For instance, the constraints related to the dispatch of operating reserves s_{r^+} (or s_{r^-}), i.e., eq. (A.1a), (A.1c), (A.1e), (A.1g), are naturally bounded by the energy limits of blocks S_{r^+} (or S_{r^-}). Consequently, the associated big-Ms values $M^1_{r^+}$ and $M^3_{r^+}$ (or $M^1_{r^-}$ and $M^3_{r^-}$) are fixed by the energy limits of block S_{r^+} (or S_{r^-}). The same logic can be applied for determining $M^2_{r^+}$, $M^4_{r^+}$, $M^2_{r^-}$ and $M^4_{r^-}$.

References

[1] D. S. Kirschen, G. Strbac, Fundamentals of Power System Economics, 2nd Ed., Wiley, West Sussex, U.K., 2018.

[2] R. Dahlgren, Chen-Ching Liu, J. Lawarree, Risk assessment in energy trading, IEEE Trans. Power Syst. 18 (2) (2003) 503–511.

[3] M. K. AlAshery, D. Xiao, W. Qiao, Second-order stochastic dominance constraints for risk management of a wind power producer's optimal bidding strategy, IEEE Trans. on Sust. Energy 11 (3) (2020) 1404–1413.

[4] A. Conejo, M. Carrion, J. Morales, Decision Making Under Uncertainty in Electricity Markets, Operations Research and Management Science, Springer, New York, 2010.

[5] K. Zare, A. J. Conejo, M. Carrión, M. P. Moghaddam, Multi-market energy procurement for a large consumer using a risk-aversion procedure, Electr. Pow. Syst. Res. 80 (1) (2010) 63–70.

[6] M. Kazemi, B. Mohammadi-Ivatloo, M. Ehsan, Risk-constrained strategic bidding of gencos considering demand response, IEEE Trans. Power Syst. 30 (1) (2015) 376–384.

[7] T. Dai, W. Qiao, Optimal bidding strategy of a strategic wind power producer in the short-term market, IEEE Trans. Sust. Energy 6 (3) (2015) 707–719.

[8] L. Baringo, A. J. Conejo, Offering strategy of wind-power producer: A multi-stage risk-constrained approach, IEEE Trans. Power Syst. 31 (2) (2016) 1420–1429.

[9] E. G. Kardakos, C. K. Simoglou, A. G. Bakirtzis, Optimal offering strategy of a virtual power plant: A stochastic bi-level approach, IEEE Trans. Smart Grid 7 (2) (2016) 794–806.

[10] M. Shabanzadeh, M.-K. Sheikh-El-Eslami, M.-R. Haghifam, The design of a risk-hedging tool for virtual power plants via robust optimization approach, Applied Energy 155 (2015) 766 – 777.

[11] M. Kazemi, H. Zareipour, M. Ehsan, W. D. Rosehart, A robust linear approach for offering strategy of a hybrid electric energy company, IEEE Trans. Power Syst. 32 (3) (2017) 1949–1959.

[12] K. Bruninx, H. Pandžić, H. Le Cadre, E. Delarue, On the interaction between aggregators, electricity markets and residential demand response providers, IEEE Trans. Power Syst. 35 (2) (2020) 840–853.

[13] J.-F. Toubeau, J. Bottieau, Z. De Greve, F. Vallee, K. Bruninx, Data-driven scheduling of energy storage in day-ahead energy and reserve markets with probabilistic guarantees on real-time delivery, IEEE Trans. Power Syst. (2020) 1–1.

[14] T. Li, M. Shahidehpour, Z. Li, Risk-constrained bidding strategy with stochastic unit commitment, IEEE Trans. Power Syst. 22 (1) (2007) 449–458.

[15] P. Sheikhahmadi, S. Bahramara, J. Moshtagh, M. Yazdani Damavandi, A risk-based approach for modeling the strategic behavior of a distribution company in wholesale energy market, Appl. Energy 214 (2018) 24 – 38.

[16] N. Shang, C. Ye, Y. Ding, T. Tu, B. Huo, Risk-based optimal power portfolio methodology for generation companies considering cross-region generation right trade, Appl. Energy 254 (2019) 113511.

[17] Y. Wang, Y. Dvorkin, R. Fernández-Blanco, B. Xu, T. Qiu, D. S. Kirschen, Look-ahead bidding strategy for energy storage, IEEE Trans. Sust. Energy 8 (3) (2017) 1106–1117.

[18] U. Yıldıran, İsmail Kayahan, Risk-averse stochastic model predictive control-based real-time operation method for a wind energy generation system supported by a pumped hydro storage unit, Appl. Energy 226 (2018) 631 – 643.

[19] V. Guerrero-Mestre, A. A. Sánchez de la Nieta, J. Contreras, J. P. S. Catalão, Optimal bidding of a group of wind farms in day-ahead markets through an external agent, IEEE Trans. Power Syst. 31 (4) (2016) 2688–2700.

[20] A. A. Sánchez de la Nieta, N. G. Paterakis, M. Gibescu, Participation of photovoltaic power producers in short-term electricity markets based on rescheduling and risk-hedging mapping, Appl. Energy 266 (2020) 114741.

[21] M. Song, M. Amelin, Price-maker bidding in day-ahead electricity market for a retailer with flexible demands, IEEE Trans. Power Syst. 33 (2) (2018) 1948–1958.

[22] H. Rashidizadeh-Kermani, M. Vahedipour-Dahraie, M. Shafiekhah, P. Siano, A regret-based stochastic bi-level framework for scheduling of DR aggregator under uncertainties, IEEE Trans. Smart Grid (to be published, 2020).

[23] M. S. Al-Swaiti, A. T. Al-Awami, M. W. Khalid, Co-optimized trading of wind-thermal-pumped storage system in energy and regulation markets, Energy 138 (2017) 991 – 1005.

[24] M. H. Abbasi, M. Taki, A. Rajabi, L. Li, J. Zhang, Coordinated operation of electric vehicle charging and wind power generation as a virtual power plant: A multi-stage risk constrained approach, Appl. Energy 239 (2019) 1294 – 1307.

[25] Y. Liu, Z. Shen, X. Tang, H. Lian, J. Li, J. Gong, Worst-case conditional value-at-risk based bidding strategy for wind-hydro hybrid systems under probability distribution uncertainties, Appl. Energy 256 (2019) 113918.

[26] S. Ghavidel, M. J. Ghadi, A. Azizivahed, J. Aghaei, L. Li, J. Zhang, Risk-constrained bidding strategy for a joint operation of wind power and caes aggregators, IEEE Trans. Sust. Energy 11 (1) (2020) 457–466.

[27] H. Höschle, H. Le Cadre, Y. Smeers, A. Papavasiliou, R. Belmans, An admm-based method for computing risk-averse equilibrium in capacity markets, IEEE Trans. Power Syst. 33 (5) (2018) 4819–4830.

[28] J.-F. Toubeau, Z. De Grève, F. Vallée, Medium-term multimarket optimization for virtual power plants: A stochastic-based decision environment, IEEE Trans. Power Syst. 33 (2) (2018) 1399–1410.

[29] M. Rahimiyan, L. Baringo, Strategic bidding for a virtual power plant in the day-ahead and real-time markets: A price-taker robust optimization approach, IEEE Trans. Power Syst. 31 (4) (2016) 2676–2687.

[30] J. Bottieau, L. Hubert, Z. De Grève, F. Vallée, J. Toubeau, Very-short-term probabilistic forecasting for a risk-aware participation in the single price imbalance settlement, IEEE Trans. Power Syst. 35 (2) (2020) 1218–1230.

[31] Commission Regulation (EU) 2017/2195 of 23 November 2017, Establishing a Guideline on Electricity Balancing, OJ L 312 (2017) 6–53.

[32] Elia Group, Data Download, `https://www.elia.be/en/grid-data/data-download-page` (2019).

[33] J. Dumas, I. Boukas, M. M. de Villena, S. Mathieu, B. Cornélusse, Probabilistic forecasting of imbalance prices in

the belgian context, 2019 16th International Conference on the EEM (2019) 1–7.

[34] J. Baetens, J. De Kooning, G. Van Eetvelde, L. Vandevelde, Imbalance price prediction for the implicit demand response potential evaluation of an electrode boiler, 4th Annual APEEN Conference on Energy Demand-Side Management and Electricity Markets, Proceedings (2019).

[35] J.-F. Toubeau, J. Bottieau, F. Vallée, Z. De Grève, Deep learning-based multivariate probabilistic forecasting for short-term scheduling in power markets, IEEE Trans. Power Syst. 34 (2) (2019) 1203–1215.

[36] T. Hong, J. Xie, J. Black, Global energy forecasting competition 2017: Hierarchical probabilistic load forecasting, Int. J. Forecast. 35 (2019) 1389–1399.

[37] G. Dalal, E. Gilboa, S. Mannor, L. Wehenkel, Unit commitment using nearest neighbor as a short-term proxy, 2018 PSCC (2018) 1–7.

[38] L. Duchesne, E. Karangelos, L. Wehenkel, Machine learning of real-time power systems reliability management response, 2017 IEEE Manchester PowerTech (2017) 1–6.

[39] G. Dalal, E. Gilboa, S. Mannor, L. Wehenkel, Chance-constrained outage scheduling using a machine learning proxy, IEEE Trans. Power Syst. 34 (4) (2019) 2528–2540.

[40] J.-F. Toubeau, T. Morstyn, J. Bottieau, K. Zheng, D. Apostolopoulou, Z. De Grève, Y. Wang, F. Vallée, Capturing spatio-temporal dependencies in the probabilistic forecasting of distribution locational marginal prices, IEEE Transact. Smart Grid 12 (3) (2021) 2663–2674.

[41] H. Quan, D. Srinivasan, A. Khosravi, Incorporating wind power forecast uncertainties into stochastic unit commitment using neural network-based prediction intervals, IEEE Trans. Neural Netw. Learn. Syst. 26 (9) (2015) 2123–2135.

[42] S. Pineda, J. Morales, Solving linear bilevel problems using big-ms: Not all that glitters is gold, IEEE Trans. Power Syst. 34 (3) (2019) 2469–2471.

[43] E. Moiseeva, M. R. Hesamzadeh, Strategic bidding of a hydropower producer under uncertainty: Modified benders approach, IEEE Trans. Power Syst. 33 (1) (2018) 861–873.

[44] J. Bottieau, K. Bruninx, A. Sanjab, Z. De Grève, F. Vallée, J. Toubeau, Database for automatic risk adjustment in short-term electricity markets (2020).
URL `https://doi.org/10.5281/zenodo.4265457`

[45] J.-F. Toubeau, P.-D. Dapoz, J. Bottieau, A. Wautier, Z. De Grève, F. Vallée, Recalibration of recurrent neural networks for short-term wind power forecasting, Electr. Power Syst. Res. 190 (2021) 106639.