



Integrated energy and ancillary services optimized management and risk analysis within a pay-as-bid market

Alberto Vannoni, Alessandro Sorce *

University of Genoa, Thermochemical Power Group, Via Montalegre 1, 16145 Genoa, Italy

HIGHLIGHTS

- Simultaneous dispatch optimization strategy considering energy and services market
- Integration of MILP and probabilistic Machine Learning service market model
- Operative tool for accurate technoeconomic assessments and daily management
- The novel risk propensity factor allows balancing expected profit and uncertainty
- Applied to an Italian Combined Cycle Gas Turbine in 2021, profits increase by 90%

ARTICLE INFO

Keywords:

Energy market
Ancillary service
Dispatch optimization
Risk assessment

ABSTRACT

In liberalized electricity markets, trading energy between generators and consumers occurs primarily on the Day-Ahead Market (DAM) one day in advance. However, the scheduled programs may not comply with grid requirements or real-time conditions. To ensure grid stability and sufficient reserves, system operators procure resources on the Ancillary Services Market (ASM). With the increasing share of renewable energy sources, many programmable generators are shifting their business model, from generating energy at base load to providing grid services. In this context, a DAM-based traditional approach to dispatch scheduling, widely adopted by existing techno-economics analysis, may result significantly suboptimal. This paper presents a novel model for dispatch optimization maximizing profits simultaneously on both the DAM and ASM, utilizing a mixed integer linear programming (MILP) formulation and a machine learning algorithm considering a pay-as-bid pricing system and predicting the probability of offer acceptance based on historical data. The proposed framework is modular and flexible, allowing for separate use of the MILP dispatch optimizer and the machine learning offer acceptance prediction model. A risk propensity factor is defined and the impact on the optimal bidding strategy, the expected profits, and their variability, is studied. A Montecarlo approach is used to evaluate the profits probability density function. The performance obtained (i.e. 20 min to optimize one week of operation of a Combined Cycle Gas Turbine) allows in applying the proposed methodologies for both long term energy system planning and daily production offer scheduling.

1. Introduction

In the current liberalized electricity markets, most energy is traded between generators and consumers (i.e., big consumers and energy retailers) on the Day-Ahead-Market (DAM) one day in advance of the real-time. Afterward, programs can be refined on the Intra-Day-Market (IDM) whose gate closure is one or some hours in advance of the real-time, typically volumes traded on IDM are significantly smaller than those on DAM and this market is neglected in this paper for the purpose

of simplicity. Nevertheless, programs defined by DAM (and IDM) may not comply with the grid requirements in terms of transmission capacity, moreover, real-time generation (especially from uncertain sources such as solar and wind) and demand may differ from the scheduled planning. As a consequence, the Transmission System Operator (TSO), to manage any grid congestion and guarantee the stability of the grid as well as adequate reserves, procures resources, in terms of services, on the Ancillary-Services-Market (ASM). Services providers are power plants, storage, or virtual power plant (VPP) aggregating generators, demand

* Corresponding author.

E-mail address: alessandro.sorce@unige.it (A. Sorce).

<https://doi.org/10.1016/j.apenergy.2024.123628>

Received 3 October 2023; Received in revised form 24 April 2024; Accepted 1 June 2024

Available online 18 June 2024

0306-2619/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

response loads, and storage [1–3]. Services may consist of regulation up, regulation down, spinning, and non-spinning reserves, local policies define which services are provided according to market-based mechanisms. Future regulatory frameworks may consider the possibility even for the Distribution System Operators (DSOs) to procure demand for services to flexible resources on the local scale by means of dedicated markets [4].

As the share of renewable energy sources, characterized by relevant uncertainty, increases, the demand for services is expected to follow, especially within the grids characterized by strict constraints in terms of transmission capacity. Besides the demand for services, revenue, profits, and cost (for the TSOs and DSOs) associated with this market increase as well. Consequently, understanding the dynamics driving the ASM is crucial to design effective policies aiming to reduce the costs for the TSOs and final users and to properly estimate the profit opportunities for the service providers.

Previous studies highlight how limiting techno-economic analyses to DAM and neglecting the economic opportunities from services provision, may lead to a significant underestimation of profitability indicators [5]. A widespread approach to the techno-economic assessment of market opportunities for power generators, or storage technologies, connected to the grid consists of performing a simulation of the real market behavior. For this purpose, an optimization algorithm seeks the best dispatch to maximize the profits net of all costs. Advanced strategies may include an integration of DAM and ASM and performing a robustness analysis considering the price uncertainties [6,7]. However, in those studies, the power plant is considered a price taker and the market is assumed to be based on Systems Marginal Price (SMP). This assumption does not cover all the cases, in fact, while DAM Marginal Price System is a standard in all the countries, many electricity markets (e.g., Italy and Germany [8]) implemented an ASM based on a Pay-as-bid (PAB) system. Additionally, the assumption of power plants as price takers is reasonable and generally accepted on the DAM but may be weak if the ASM is considered. In fact, the demand for services is strongly geospatial dependent and the competition among service providers within a specific grid knot to provide a specific product may be limited.

Generally, if the energy markets design is almost uniform across Europe, the procurement procedure for ancillary services is considerably heterogeneous. They may differ according to the energy or capacity base, the activation rule, product resolution, qualifiable providers, the settlement rule, or others. At the EU level, the commission establishes general guidelines [9], and the ENTSO-E annually surveys how the national grid codes implement them [10]. Both SMP and PAB settlement rule are common, different studies recommend SMP systems only if sufficient [11], thus, considering energy-based payments, 9 TSO implement PAB settlement for an automatic Frequency Restoration Reserve (aFRR), and 12 TSO for manual Frequency Restoration Reserve (mFRR) out of 36 surveyed [10].

Although the pay-as-bid scheme is widespread, it is far less investigated than SMP in literature and few studies have addressed the problem of optimizing the management strategy of a plant in this framework. Mazzi et al. proposed a two-stage linear approach within a pay-as-bid pricing scheme at the ASM stage for a price-tacker conventional producer, showing the capability of increasing profit expectations with respect to a sequential offering approach [12]. Previously Ren et al. published a study introducing the problem of offer strategy within pay-as-bid schemes [13] and studying the market behavior, generator profits, and consumer payments assuming optimized offering strategy of generators [14]. Other studies focus on optimally scheduling aggregators including storage [15,16]. Finally, a rigorous mathematical formulation of the integrated scheduling problem is provided by Swider resulting in the maximization of non-linear stochastic objective function [17].

This paper aims to develop a model of dispatch optimization by means of a linear formulation, to maximize the overall profits on DAM

and ASM, solved through a Mixed Integer Linear Programming (MILP) algorithm in order to determine the best bidding strategy both on DAM and ASM. The proposed approach focuses on maximizing the sum of DAM profits, on which uncertainties are neglected for the sake of simplicity, and expected profits on the ASM (i.e., the product of the probability of offer acceptance and the profits realized if the offer were accepted). The probability of offer/bid acceptance is predicted by a data-driven model trained on Italian real ASM data. The model approach is similar to that presented in [18], but the training dataset is here extended from 5 to 216 production units and not limited to offers but includes bids to downward regulation.

The novelty of this work consists of the simultaneous optimization of power generators dispatch in the day-ahead and ancillary services markets. A modular approach combining a linear formulation of the integrated energy-ancillary services dispatch problem within a pay-as-bid ASM and a model for the offers/bid classification as accepted or rejected and the assessment of probability associated with each class is here proposed. The proposed model is a machine learning algorithm, selected among different options, according to the defined goodness criteria. However, these two core elements are independent of each other and both of them can be used separately in the future. Whatever model for offer/bid acceptance probability can be integrated with the MILP approach to the optimization problem, on the other hand, the described machine learning model and the proposed goodness criteria represent a change of perspective with respect to existing literature, not only forecasting the market closure price but explicitly quantifying opportunities for the services providers. An additional factor describing the risk propensity of the operator is also included. Moreover a novel function is proposed to model propensity to risk. This is pivotal if the ASM is designed according to a pay-as-bid logic. The proposed model and approach, here applied to generators, can be also extended to planning the operation of those resources able to provide services to the grid.

Moreover, the integrated dispatch optimizer can be adopted for a twofold purpose, on one hand, it is an effective tool to optimally manage existing power plants, on the other hand, it is able to properly quantify the economic value of flexibility from programmable power plants demonstrating to be an essential tool for reliable techno-economic analysis of future investments.

In order to demonstrate the developed methodology described in the first four sections, in the Section 5 a combined cycle gas turbine (CCGT) F-class of 400 MW is assumed as an example of a power generator for the purpose of this paper, this is done considering that this kind of machine represents the majority of the installed Italian production units based on the natural gas combustion. Section 2 describes the MILP formulation of the problem including start-up costs and the impact of off-design performances considering the effect of partial load operation and GT inlet air temperature variation. Section 3 focuses on the machine learning model to predict the probability of offer acceptance; different algorithms have been compared to select the best at minimizing the misclassification when the error on the ASM revenue is zero. The following section introduces the concept of risk propensity and illustrates how the objective function can be adjusted according to the risk-averse or risk-seeking attitude of the operator in light of the uncertain acceptance of ASM offer especially considering a pay-as-bid settlement rule. Finally, the last section presents and discusses some yearly simulations exploring how considering the ASM opportunities affects also the DAM bidding strategy. Moreover, the impact of the risk propensity factor β is discussed through a Montecarlo approach to quantify how the probability density function of profits varies with the operator risk propensity on such an uncertain market as the ASM.

2. MILP problem formulation

The optimization approach presented in this paper represents a novel approach that implements a MILP for the simultaneous optimization of

both the market strategies (on the day ahead and the ancillary services market). It takes a relevant step further from the subsequential methodology presented in [5], where the DAM dispatch is optimized first and the following ASM optimization is limited by the residual flexibility margin imposed by the plant capacity and minimum load. A simultaneous optimization could lead to a non-optimal solution on the single market, but the expected optimum is reached by summing the two markets together. The following subsection describes the implemented approaches that is also schematized in Fig. 1 that includes references to the section of this paper which deepen specific functionalities.

2.1. Mixed integer linear programming formulation

The MILP approach addresses the dispatch optimization of a generic power generator whose power output has been discretized into $nLoad$ operational modes including the stand still condition represented by a null output, 0 MW (0%), and the maximum capacity. The discretization does not require to be evenly spaced so that if a specific technology requires to operate at the least at a minimum load it will be sufficient to discretize its power output including a null output, the minimum load, the maximum capacity, and a variable number of modes at intermediate loads. It should be considered that the power output resolution has a great impact on the time required to solve the problem as quantified in [19].

The problem consists of determining the best load (among those considered by the discretization) scheduled after the DAM closure and the best offer or bid to present on the ASM for each time step t . Basically, the generator will be allowed to operate according to $nMod$ operational modes considering the effect of both markets. With $nMod = nLoad^2$, in fact for each load scheduled after the DAM closure, the generator can present an offer to increase its power output to any higher load, or it can present a bid to lower the output according to the discretization. Eventually, it can maintain its schedule without any offer or bid on the ASM.

The best operational mode (i.e., the combination of DAM load and offer/bid on the ASM) is that which maximizes the overall expected profits considering both markets. Eq. (1) defines the objective function.

$$ExpProf = \sum_{t=1}^{n \text{ hours}} Profits_t^{DAM} + ExpProfits_t^{ASM} - ExpC_{SU_t} \quad (1)$$

The total expected profits must be maximized and are the sum of three main components:

- Profits on the DAM [€], as defined by eq. (2), are the difference between the electricity price and the cost of generating the electricity [€/MWh] times the quantity sold in this market [MWh]. The

electricity price on the DAM is assumed to be known in advance so that there is no uncertainty about this profit

$$Profits_t^{DAM} = (pr_{el}^{DAM} - COE_t^{DAM}) \cdot Q_t^{DAM} \quad (2)$$

- Expected profits on the ASM [€]. Eqs. (3) and (4) define the profits associated with each offer (3) and bid (4) on the ASM, then for each simulation interval, the optimizer will determine if it is more beneficial to offer or bid energy. If the load scheduled on DAM is equal to the one in ASM, no operations are done in the latter market and this term is zero. The expected profits, eq. (5), take also into account the probability of offer/bid acceptance.

$$Profits_t^{ASM} = (pr_{el}^{ASM} - COE_t^{ASM}) \cdot (Q_t^{ASM} - Q_t^{DAM}) + (COE_t^{DAM} - COE_t^{ASM}) \cdot Q_t^{DAM} \quad (3)$$

$$Profits_t^{ASM} = (COE_t^{DAM} - pr_{el}^{ASM}) \cdot (Q_t^{DAM} - Q_t^{ASM}) + (COE_t^{DAM} - COE_t^{ASM}) \cdot Q_t^{ASM} \quad (4)$$

$$ExpProfits_t^{ASM} = Profits_t^{ASM} \cdot ACC_{prob_t} \quad (5)$$

Q_t^{ASM} represent the generation quantity scheduled for the time t after the ASM closure, so that the term $(Q_t^{ASM} - Q_t^{DAM})$ is the quantity traded on the ASM for a generic service. Eqs. (3) and (4) include two terms, the first is directly associated with the energy traded, and the second relies on the typical generator efficiency dependency on the load (i.e., the off-design performance) so the COE varies as the load changes after the ASM closure. Referring to the energy quantity associated with the service provision the specific acceptance probability, the eqs. (3–5) are independent applicable to any service type.

- The expected start-up costs [€]. The start-up in many cases has a relevant cost, in terms of fuel and components life consumption due to thermal and mechanical stresses associated with this transient. However, even if the cost of start-up at the time t is known, there is no certainty about the offers/bids acceptance which can modify the generator status (on/off) at the time t and at the time $t-1$. Because of that, the expected SU costs must be defined as in eq. (6). Eqs. (7) and (8) define the SU probability, where the LoadDAM is the load after the DAM closure and LoadASM is the load that the power plant should adopt if the presented offer/bid is accepted.

$$ExpC_{SU_t} = C_{SU_t} \cdot SU_{prob_t} \quad (6)$$

$$SU_{prob_t} = On_{prob_t} \cdot Off_{prob_{t-1}} = On_{prob_t} \cdot (1 - On_{prob_{t-1}}) \quad (7)$$

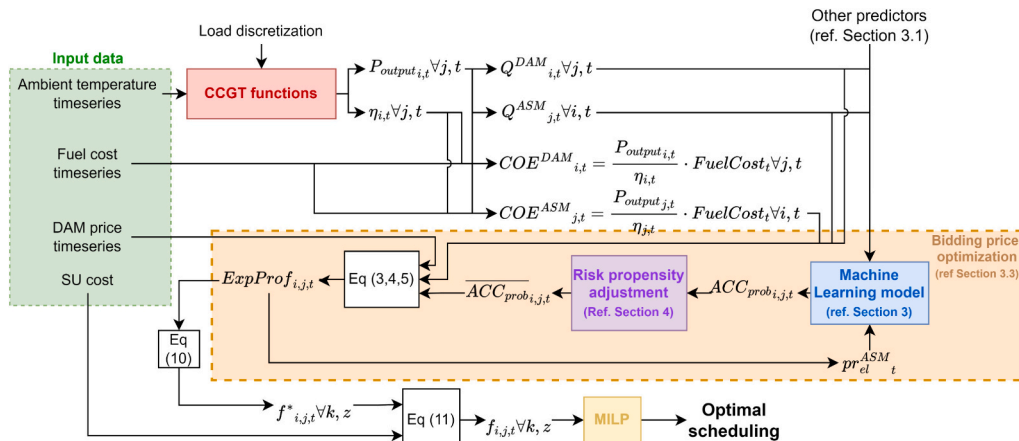


Fig. 1. Proposed methodology flow chart.

$$On_{prob_t} = \begin{cases} 1, & Load^{DAM} > 0 \text{ and } Load^{ASM} > 0 \\ ACC_{prob_t}, & Load^{DAM} = 0 \text{ and } Load^{ASM} > 0 \\ 1 - ACC_{prob_t}, & Load^{DAM} > 0 \text{ and } Load^{ASM} = 0 \\ 0, & Load^{DAM} = 0 \text{ and } Load^{ASM} = 0 \end{cases} \quad (8)$$

Eqs. (1),(6),and (7) state clearly that the profits at the time t depend on the probability of “on” status, thus on operational mode (i.e., the combination of DAM schedule and offer/bid presented on the ASM), at the time t and $t-1$. This problem cannot be solved directly through a linear programming approach. To manage it, the number of operational modes that can be selected by the scheduler is increased to $nMod^* = nMod^2 = nLoad^4$. Each mode* univocally defines an operational mode at the time t and at the time $t-1$ so that profits and expected start-up cost at time t are determined by the mode* at time t .

Eq. (9) defines the MILP problem, where the dimension of f is a currency [€], the first $nLoad^4 \cdot nInterval$ elements of x are boolean and the last $nInterval$ elements are dimensionless probability.

$$\max_x f^T \cdot x, \text{ subject to } \begin{cases} x(1 : nMod^* \cdot nInterval) \text{ are integers} \\ 0 \leq x \leq 1 \\ A \cdot x \leq b \end{cases} \quad (9)$$

Even if most linear programming solvers require f and x to be 1-D vectors, is possible to implement the problem by initializing f as a 5-D array, f^* , reshaping it as a column vector before calling the solver, f^{**} , and adding the C_{SU} . Then reshaping x back to 5 dimensions once the solution has been found. The 5 dimensions are relative to:

- $Load^{DAM}$ at time t , i index
- $Load^{ASM}$ at time t , j index
- $Load^{DAM}$ at time $t-1$, k index
- $Load^{ASM}$ at time $t-1$, z index
- Time interval, t index

So that:

$$f_{ij,k,z,t}^* = (pr_{el}^{DAM} t - COE_{it}^{DAM}) \cdot Q_{it}^{DAM} + Profits_{ij,t}^{ASM} \cdot ACC_{prob_{ij,t}} \forall k, z \quad (10)$$

After f^* has been implemented and reshaped to the column vector f^{**} further elements (as many as the considered time intervals) must be added to obtain the vector f used in eq. (9).

$$f = \begin{bmatrix} f^{**} \\ -C_{SU} \end{bmatrix} \quad (11)$$

According to the equations reported above, the solution x will be a vector of n elements, with $n = nLoad^4 \cdot nInterval + nInterval$. The first $nLoad^4 \cdot nInterval$ elements are constrained to be 0 or 1, $x^*(i,j,k,z,t) = 1$ if and only if the relative mode is selected.¹ The last $nInterval$ elements of x represent auxiliaries variables, to take into account the expected start-up cost as defined by eq. (6) their value must be constrained to the probability that a start-up occurs at that time according to the selected mode*.

MILP formulation, eq. (9), reports also the constraints to which the problem is subjected. In addition to integer constraints and upper and lower bounds, which need no further explanation, the problem is subject to linear inequality constraints defined by the matrix A and the column vector b . A is sized n -by- m and b has a length equal to m , where m is the number of constraints, $m = 3 \cdot nInterval + nInterval \cdot nLoad^2$.

- The first $nInterval$ constraints of A impose that the start-up auxiliary variable at the time t is equal to, or greater than, the start-up probability at the time t . Of course, the optimal value of the auxiliary variables is the minimum allowed, since it multiplies C_{SU} . However, imposing an equality constraint is much more expensive in terms of

computational effort. So the constraint is expressed by the inequality of eq. (10) and the solver will select the minimum allowed value

$$x_{nInterval \cdot nLoad^4 + t} \geq SU_{prob_{ij,k,z,t}} \quad (12)$$

- The second $2 \cdot nInterval$ constraints of A impose that at least one, and no more than one, operating mode* is selected for the time t

$$\sum_{z=1}^{nLoad} \sum_{k=1}^{nLoad} \sum_{j=1}^{nLoad} \sum_{i=1}^{nLoad} x_{ij,k,z,t}^* \leq 1, \forall 1 \leq t \leq nInterval \quad (13)$$

- The last $nInterval \cdot nLoad^2$ constraints impose the consistency between the time t and $t-1$. For the first time interval ($t = 1$) the consistency is imposed with respect to the status $t = 0$ which is imposed as a boundary condition.

$$x_{ij,a,b,t}^* = x_{a,b,k,z,t-1}^*, \forall 2 \leq t \leq nInterval \text{ and } \forall (1,1) \leq (a,b) \leq (nLoad, nLoad) \quad (14)$$

$$x_{ij,a,b,1}^* = x_{t=0}^*, \forall (1,1) \leq (a,b) \leq (nLoad, nLoad) \quad (15)$$

Defined the problem and all the constraints the problem can be solved, the last section of this paper shows some outputs of the problem solved in MATLAB by the `intlinprog` function [20].

2.2. Time discretization and simulation horizon

Time discretization in many markets is performed on an hourly basis, but the resolution adopted can be increased (some service markets work with 15-min time intervals) without any change in the given mathematical formulation.

Concerning the simulation horizon, it should be considered that the assumption of a known price on the DAM is no longer reliable as $nInterval$ increases covering a period of many days. To simulate longer periods they are divided into subperiods, covering a time horizon on which is reasonable to assume the DAM price as known, typically 1 or 2 days. Defined this period then the simulation time window is progressively shifted. This approach is a well-known practice for MILP problems, successfully applied by [19,21]. For each optimization, only the first day is kept, while the scheduling related to the exceeding hours will be overwritten by the first day of the following optimization. The solution relative to the last time interval considered (e.g., the 24th) is imposed to the following optimization as the $x_{t=0}^*$ constraint. The sliding window width adopted in this paper is 36 h accordingly to the optimization made in [19].

3. Machine learning approach for the prediction of offers and bids acceptance

The previous subsection described the MILP approach including the objective function and the constraints. It is clarified how the uncertainty of success in the service market plays a crucial role in eqs. (5),(8), and (10). This section is focused on how to assess the $ACC_{prob_{ij,t}}$ (i.e., for each DAM schedule, each possible offer/bid, and each time interval t). While the described MILP methodology is general and can be applied to different contexts, to predict the ASM offer/bid probability is needed to select a specific case study since the local market designs and the available data impose to differentiate the approach. For this purpose, the Italian ASM (*Mercato dei Servizi di Dispacciamento*) is selected.

3.1. Italian case study: Raw data and pre-processing

A previous work [5] describes the Italian market design and how each unit

(e.g., power generator, storage facility, or virtual aggregator) can

¹ x^* is the reshaped 5-D solution excluding the auxiliary variables.

present offers and bids showing its availability to revise the schedule, defined after the DAM and the IDM closure, upward or downward. The Italian ASM is organized in two phases: the first, defined “ex-ante”, requires BSPs to submit offers/bids by 17:00 on day D-1, which are possibly accepted by the TSO, at least 105 min in advance of the real time, to relieve congestion within the bidding zones and ensure the availability of adequate FRR and RR margins. Time resolution of “ex-ante” phase is hourly. While during the second, real time phase, of the Italian Balancing Market (BM), offers and bids (with 15 min resolution) are selected with the aim of maintaining the balance between electricity injections and withdrawals, relieving real-time congestions, and ensuring or restoring FRR and, if needed, RR margins [22]. For the purpose of this paper only the first phase, ASM “ex-ante” is considered.

On the *Gestori dei Mercati Energetici* (GME) website [23], a public domain of the presented offers and bids is available. Data are available as a zipped file for every single day, containing a .xml file for each market. Each .xml file reports the fields described in Table 1.

The described raw data are processed and selected to create the datasets on which the machine learning algorithms are trained and tested. The period covers the last complete four years (from 2018 to 2021) the number of offers/bids is in the order of 10^7 . Only predefined offers/bids and only accepted or rejected, have been considered. The analysis is also limited to the 216 units that it was possible to identify (99 gas-fired generators, 19 coal-fired, 5 oil-fired, and 93 hydro). The complete list of identified units active on the ASM is available in [24]. To train proper models the predictors are selected or derived from the raw data and gathered as follows:

- Offer/bid specificity predictors: PURPOSE_CD, SCOPE
- Offer/bid time predictors: INTERVAL_NO, MONTH (month on which the offer is submitted), DAY_TYPE (H for holidays, B for weekdays)
- Offer/bid strategy predictors: ADJ_QUANTITY_NO and ADJ_ENERGY_PRICE, so that the fulfillment of the grid code is assumed
- Market indicator predictors: MGAS_MGP (the spot gas market price on the day the bid/offer is presented), PUN (Single National electricity price on the DAM on the hour on which the offer/bid is presented), Zonal_price (zonal electricity price on the DAM on the hour the offer/bid is presented), data published at [23]
- Units predictors: Voltage (the voltage level of the grid to which the unit is connected [kV]), Lat, and Long (latitude and longitude degrees of the capital of the administrative province in which the unit is located), data collected by [24]

Different models are trained including or not including coordinates as latitude and longitude, which will be referred to as GEO and NOGEO respectively. Different degree of model optimization is also tested differentiating between “coarse” and “fine” models:

- Coarse: 5 cross-validation folders, maximum 30 function evaluations during optimization, MATLAB automatic selection of optimized hyperparameters
- Fine: 15 cross-validation folders, maximum 50 function evaluations during optimization, all hyperparameters are optimized

Including geographical information can have a twofold effect. On one side the geospatial dependency of the ASM opportunities is widely acknowledged, so including this information in the Machine Learning algorithms would allow keeping it into consideration. On the other hand, it could lead to overfitted models, especially for Tree and Ensemble model the risk is to train specific sub-models for each latitude and longitude couple (i.e., each unit). The geographical information is provided to the algorithm as the coordinates of the province capital. So that the information is general enough to mitigate the overfitting risk, the province can describe the opportunities, as assumed by [25], since often the administrative boundaries coincide with orographic barriers which have constrained the grid development in the past.

Concerning the training and test split, in order to test the ability of the model to predict the probability of acceptance in different years, data from one year forms the test dataset, and data from the other years are used for training. To include in the training set offers/bids from both exceptionally high and low gas price market scenarios, 2019 is selected to be the test dataset, and 2018, 2020, and 2021 the training. In fact, market prices in 2020 were very low, because of the COVID-19 pandemic measures, while in 2021 (especially during Q4) prices were extremely high because of the beginning of the current energy crisis. To reduce computational time training is carried out on stratified partitions equal to 1%, 5%, and 10% of the whole dataset.

3.2. Models assessment and selection

Different models have been trained with the purpose of comparing each other.

In addition to the already mentioned features (GEO/NOGEO, fine/coarse), models are distinguished for the algorithm, and the following options have been tested, within brackets is reported the MATLAB function name:

- Classification tree (fitctree)
- Ensemble classifier (fitcensemble)
- Discriminant Analysis classifier
- Naive Bayes (NB) classifier (fitcnb)
- k-Nearest Neighbors (kNN) classifier (fitcknn)

The hyperparameter options are optimized in order to minimize cross-validation classification errors. Once the model training has been optimized the test dataset is used to assess the model goodness. For each observation in the test dataset the prediction of the target class and the posterior probability of being classified as ACC is assessed. If the posterior probability is higher than the threshold value² an ACC label is assigned, otherwise REJ label. Fig. 2 reports the posterior ACC probability of prediction on the test dataset for the Ensemble NOGEO coarse model, trained on the 5% partition of the whole training dataset, assumed as an example for the following paragraphs.

The red distribution is relative to the offers/bids which actually have been rejected, the blue distribution is relative to the offers/bids that have been accepted. The distribution is normalized on the overall number of observations reporting the same actual label. E.g., 77% percent of the offers/bids which have been actually rejected are predicted to have an acceptance probability between 0% and 2% (first red bar on the left). Generally, it can be appreciated how the actual rejected offers/bids are predicted to have a lower probability of acceptance (almost all the actually rejected offers/bids report an ACC posterior probability lower than 10%–20%). Conversely, actually accepted observations are foreseen to have a higher probability to be accepted, even if probabilities >50% are very rare and 10% of them are foreseen to have an almost null (<2%) probability.

Looking at Fig. 2, it is clear how a further step in optimizing the model can be taken by selecting the best threshold value T . If $T = 0$, all the offers will be classified as ACC, which means that the True ACC rate (TACCR) is equal to 1, but also the False ACC rate (FACCR) is equal to 1. It follows that the profits realized on the market are overestimated. Otherwise if $T = 1$ all the observations are classified as rejected, implying a null TrueACC rate and False ACC rate, and an underestimation of profits. The best classifier for the purpose described in this section is the classifier that minimizes FACCR and the error in estimating revenues, and profits maximizing the TACCR.

$$TACCR = \frac{TACC}{TACC + FREJ} \quad (16)$$

² During the training the threshold value is set to 0.5.

Table 1
list of fields included in the .xlm files and relative description.

Names	Values	Description
PURPOSE_CD	OFF or BID	Indicates if the purpose is to increase or decrease the injection to the grid
TYPE_CD	STND or REG	Indicates whether the offer/bid is predefined or current
STATUS_CD	ACC, REJ, INC, REP, REV, or SUB	Indicates the status of the offer/bid after the market execution: accepted, rejected, inadequate, replaced, revoked, or submitted
MARKET_CD	MSD	Identify the market
UNIT_REFERENCE_CD	string	Identification code of the market unit
INTERVAL_NO	Integer, from 1 to 24 (25 in the day of daylight-saving time to standard time switch)	Identify the hourly time interval
BID_OFFER_DATE_DT	YYYYMMDD	The date on which the bid/offer refers
TRANSACTION_REFERENCE_NO	Integer	Offer identifier
QUANTITY_NO	Double	Volume [MWh] offered/bid
AWARDED_QUANTITY_NO	Double	Volume awarded on the market
ENERGY_PRICE_NO	Double	Price [€/MWh] offered/bid
PARTIAL_QTY_ACCEPTED_IN	Y/N	Indicate if the offer/bid has been partially accepted, N if the AWARDED_QUANTITY_NO is equal to 0 or ADJ_QUANTITY_NO, Y otherwise.
ADJ_QUANTITY_NO	Double	Adjusted offer/bid volume to fulfill the constraints imposed by the grid code
ADJ_ENERGY_PRICE_NO	Double	Adjusted offer/bid price to fulfill the constraints imposed by the grid code
GRID_SUPPLY_POINT_NO	String	Grid supply point to which the unit is associated
ZONE_CD	String	Bidding zone to which the unit belongs
AWARDED_PRICE_NO	Double	Awarded price on the market, in Italy the ASM is a pay-as-bid market, so the awarded price is equal to the adjusted price
OPERATORE	String	Market operator, typically the Balancing Response Party
SUBMITTED_DT	YYYYMMDDhhmmssff	Date and time on which the offer/bid has been submitted
SCOPE		Type of service offered ^{1,2} : RS: "secondary reserve", i.e., aFRR AC: start-up or shutdown (depending on PURPOSE) CA: change configuration (for 2 GT + 1 ST CCGTs only) GRx: x th step for "other services" including mFRR, RR, congestion relieving

¹ The Italian TSO Terna does not provide an official correspondence between its own nomenclature and that one indicated by the ENTSO-E. However scientific community propose the correspondence reported in Table 1 [30,31]

² There is no product for FCR since "primary reserve" is mandatory for programmable large-scale power plant and not subjected to market mechanism in Italy. [32]

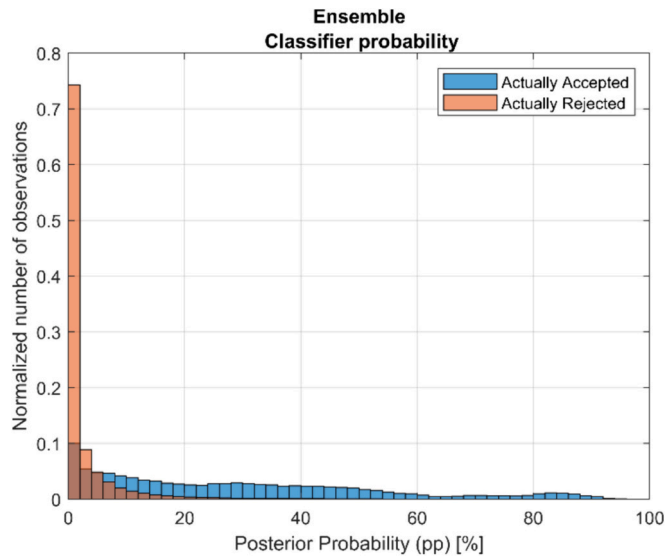


Fig. 2. Posterior ACC probability distributions for Ensemble NOGEO coarse classifier.

$$FREJR = \frac{FACC}{TREJ + FACC} \quad (17)$$

Where TACC, FACC, FREJ, and TREJ are respectively the observations correctly classified as accepted, misclassified as accepted, correctly classified as rejected, and misclassified as rejected. Fig. 3 reports the ROC curve for the same model as the previous fig. $T = 0$ corresponds to the point in the upright corner, increasing T to 0.02 the TACCR decrease to 77%, and the FACC to 10%. With reference to Fig. 2, this case corresponds to a threshold value between the first and the second bar.

The reader can figure out T as a vertical threshold line moving from

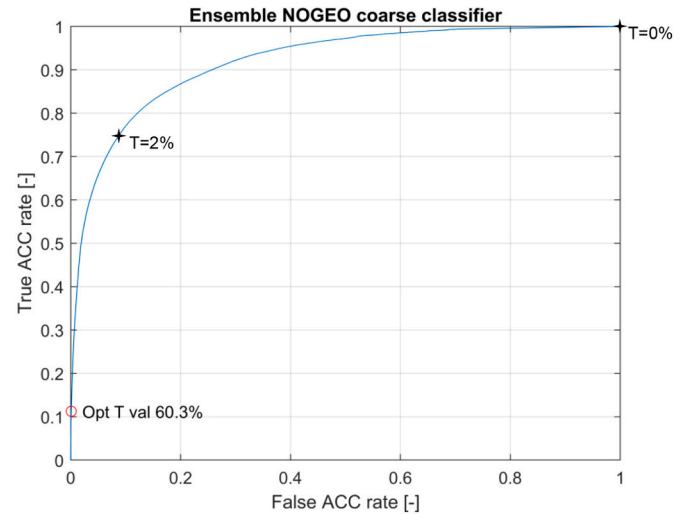


Fig. 3. Receiver Operating Characteristic curve for Ensemble NOGEO coarse classifier.

left to right in Fig. 2. The red bars on the left of the threshold value are correctly classified as rejected (TREJ), and the red bars on the right are misclassified as accepted (FACC). Analogously, the blue bars on the line are misclassified as rejected (FREJ) while the blue bars on the right are correctly classified as accepted (TACC). Figuring out to move the threshold line from the left to the right we move from the upright corner to the bottom left corner along the ROC curve in Fig. 3. It is common to refer as a perfect classifier to a model which can separate the acceptance probability distributions of Fig. 2 so that would exist at least one T to which corresponds $FACCR = 1$ and $TACCR = 1$. In this case, the area under the curve (AUC in Fig. 3) will be 1. AUC is assumed as a parameter to assess the real model.

It was pointed out how this model aims to correctly assess the unit's profits and that for $T = 0$ the profits will be overestimated and for $T = 1$ underestimated. For every trained model exist an optimal T for which the expected profits are equal to the actual profits on the test dataset. Therefore is relevant to assess how the model performs for $T = T_{opt}$.

Assessing the profits error is not trivial, since profits require taking into account the cost of electricity, which is dependent on the single generator, the load, the spot fuel price, and many other variables. Two formulas, normalized on the actual value, are here proposed, corresponding to the continuous and dashed lines of Fig. 4 respectively. The first, eq. (16), takes into account only the offers revenues. The second, eq. (17), includes a simplified calculation of bids profits in which the cost of electricity is assumed to be the DAM clearing price. Both values are the difference between the expected values of the model and the actual value from historical data. To compute the expected data it is needed to assess the ACC_{prob} for all the offers which differs from the posterior probability (pp) based on T . Eqs. (18) and (19) impose that $ACC_{prob} = 0$ for $pp = 0$, $ACC_{prob} = 1$ for $pp = 1$, and $ACC_{prob} = 0.5$ for $pp = T$, rescaling the probability.

$$Err_{RevOFF} = (pr_{aw} \cdot Q_{aw} \cdot ACC_{prob})_{OFF \text{ all}} - (pr_{aw} \cdot Q_{aw})_{OFF \text{ actually ACC}} \quad (9)$$

$$Err_{RevOFFProfBID} = Err_{RevOFF} + (Q_{aw} \cdot (pr_{DAM} - pr_{aw}) \cdot ACC_{prob})_{BID \text{ all}} - (Q_{aw} \cdot (pr_{DAM} - pr_{aw}))_{BID \text{ actually accepted}} \quad (10)$$

$$ACC_{prob} = pp^a \quad (11)$$

$$a = \frac{\log(0.5)}{\log(T)} \quad (12)$$

Since most of the offers/bids are commonly rejected, the dataset is highly unevenly distributed among the ACC and REJ classes, then the classification error is less significant as criterium (all the models report similar value around 1.5%). Moreover, TACCR can be very low turning out to be the most critical goodness parameter. As appreciable in Fig. 4, the TACCR trend is decreasing against T , a perfect classifier would show a step trend and TACCR = 1 for $T = T_{opt}$, which means that all the offers/bids are correctly classified leading to a zero error in revenue estimation.

Assuming the AUC, the classification error, and the TACCR, for $T = T_{opt}$ as the main model goodness indicators, the following figures outline the comparison between the trained models. Besides the goodness of prediction, the prediction time is extremely important since the probability of offer/bid acceptance is optimized and assessed several times to schedule the optimal dispatch. Then to keep the scheduling process in a reasonable amount of time, models with low prediction time must be prioritized.

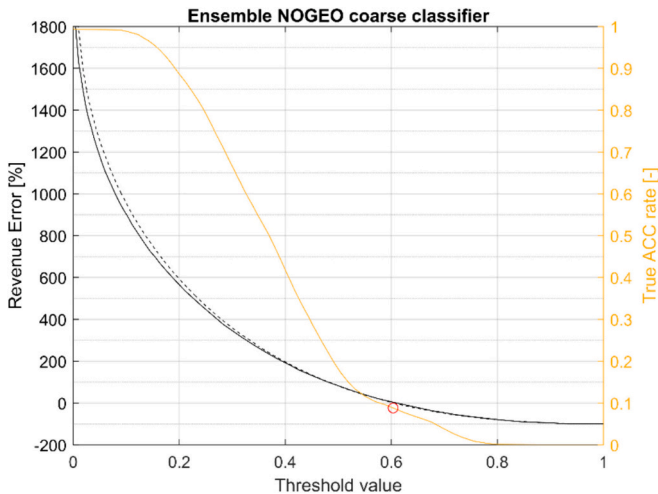


Fig. 4. Revenue error and True Accepted Rate vs Threshold value.

In Fig. 5, kNN and Naive Bayes models are not shown in the figures since the firsts were critical to train on such a large data set, and an *out of memory* error causes MATLAB to stop³ while the Naive Bayes ones take too long time to train. The training duration is limited to 24 h, afterwards, the process is forced to stop. Classification error is around 1.5% for almost all the models since the majority of offers/bids are rejected is not possible to appreciate a real difference in this parameter. AUC values in the order of 0.9 are reached by Tree classifiers and some Ensemble classifiers. Ensemble classifiers are expected to perform better than Trees since they combine many Tree classifiers, however, some of them report low values in AUC or classification errors higher than the average. This is because Ensembles' hyperparameters optimization requires more time than the Trees, so the final Tree classifier is close to the best that is possible to train on that data, while the Ensembles may be far away from the optimum hyperparameters array when the process reaches the maximum allowed training time. The TACCR is up to 0.45 for the Ensemble GEO fine model trained on the 10% partition of the whole training dataset, however, the same model reports the highest classification error.

Table 2 reports all the mentioned parameters for the best models of each training partition. The Tree Coarse GEO model trained on the 5% partition was selected and used in section 4, because of its good trade-off between goodness indicator and prediction time.

3.3. Model of acceptance prediction application within the dispatch optimizer

The acceptance probability prediction by algorithms is then used in the dispatch optimization described in Section 2 as a fundamental term of eqs. (6) and (8). To predict the probability of acceptance of a new observation ($ACC_{prob,i,j,t}$), the algorithms need all the predictors. With reference to the previously reported subdivision:

- Offer/bid specificity predictors: PURPOSE_CD is OFF if $Load_{i,t}^{DAM} < Load_{i,t}^{ASM}$. BID otherwise. SCOPE is selected according to the case study, in the following paragraphs is imposed GR1 (so generically "other services", including congestion resolution, mFRR, and RR)
- Offer/bid time predictors: INTERVAL_NO, MONTH (month on which the offer is submitted), DAY_TYPE (H for holidays, B for weekdays) are imposed by the time t
- Market indicator predictors: MGAS_MGP (the spot gas market price on the day the bid/offer is presented), PUN (Single National electricity price on the DAM on the hour on which the offer/bid is presented), Zonal_price (zonal electricity price on the DAM on the hour the offer/bid is presented), are input to the dispatch optimization, so they depend on the time t and the case study location
- Units predictors: Voltage (the voltage level of the grid to which the unit is connected [kV]), Lat, and Long (latitude and longitude degrees of the capital of the administrative province in which the unit is located), are depending on the case study location.
- Offer/bid strategy predictors: ADJ_QUANTITY_NO is imposed by the $Load_{i,t}^{DAM}$ and $Load_{i,t}^{ASM}$. ADJ_ENERGY_PRICE must be optimized to satisfy the expected profits in the ancillary services market

The only free predictor is the offer/bid price that, since the considered ASM is a pay-as-bid market, influences not only the probability of acceptance but also directly the ASM profits and so their product, the expected ASM profit eq. (5).

The following three figures describe the price optimization by means of an example relative to an offer up to a residual capacity of 200 MWh/

³ MATLAB 2019b was used on a computer reporting the following features. Processor Intel(R) Xeon(R) E-2176G CPU @ 3.70GHz, Memory 16 GB, OS Microsoft Windows 10 64bit.

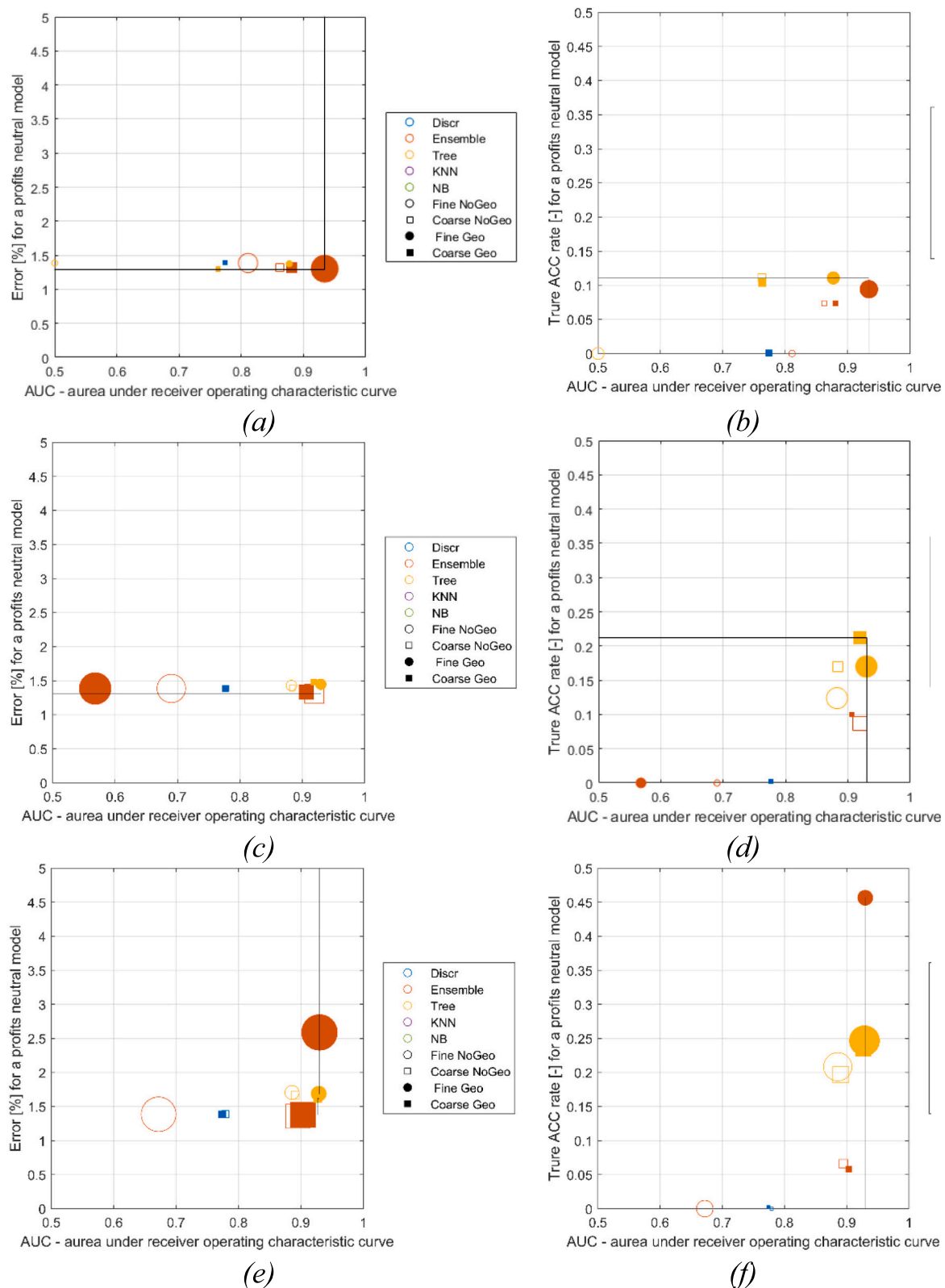


Fig. 5. Models goodness indicator parameters. Models trained on 1% ((a) and (b)), 5% ((c) and (d)), and 10% ((e) and (f)) stratified partition of the whole training dataset. The figures on the left report the classification error vs AUC, and the size of the markers is proportional to the training time. The figures on the right report the true accepted rate vs AUC, while the markers' size indicates the prediction time.

Table 2

Best models goodness parameters.

Partition	Model	AUC	Classification Error	TACCR	Prediction time
1%	Ensemble Fine GEO	0.934	1.30%	0.09	776 s
5%	Ensemble Coarse GEO	0.907	1.33%	0.10	15 s
5%	Tree Coarse GEO	0.919	1.48%	0.21	476 s
10%	Ensemble Fine GEO	0.929	2.59%	0.46	433 s
10%	Tree Geo Coarse	0.927	1.62%	0.24	1215 s
10%	Tree Fine GEO	0.929	1.69%	0.25	4918 s

h. The reported example optimizes revenues so that no assumptions on the COE are needed, while the process implemented within the dispatch optimizer is analogous but optimizes the profits. Fig. 6 shows the revenues contour if the offer was accepted with respect to the quantity and the price. Of course, the higher the quantity and the price the higher are revenues. Fig. 7 concerns the acceptance probability as output from the

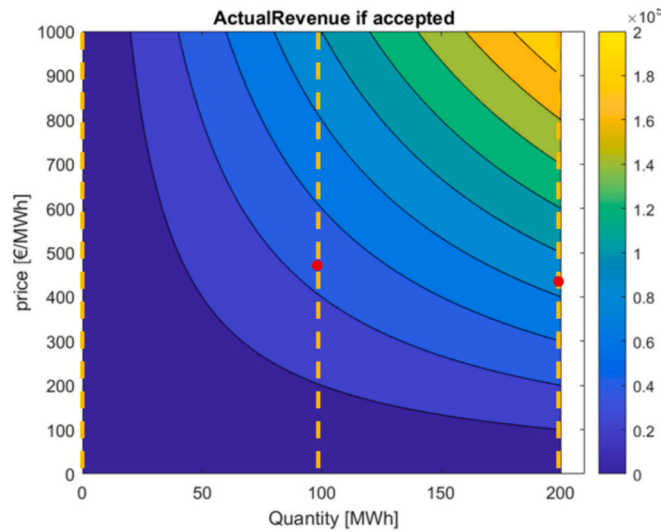


Fig. 6. ASM revenues [EUR] contour vs the offered quantity and price. Yellow dashed lines represent the constraint imposed by the i and j indices. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

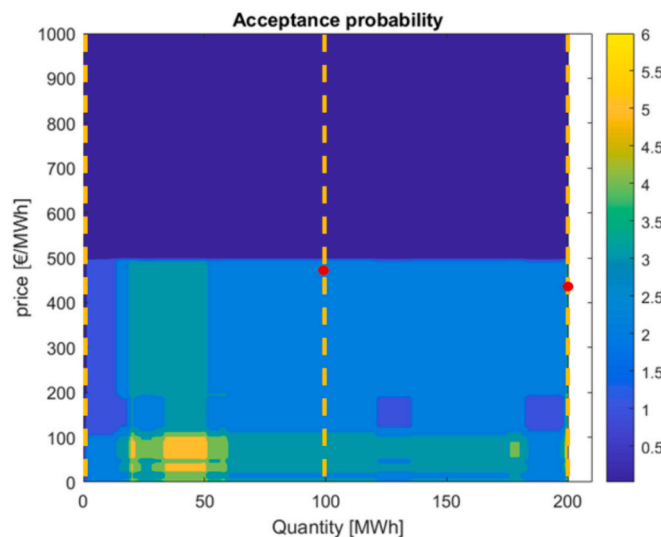


Fig. 7. ASM offer acceptance probability [%] contour vs the offered quantity and price. Yellow dashed lines represent the constraint imposed by the i and j indices. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

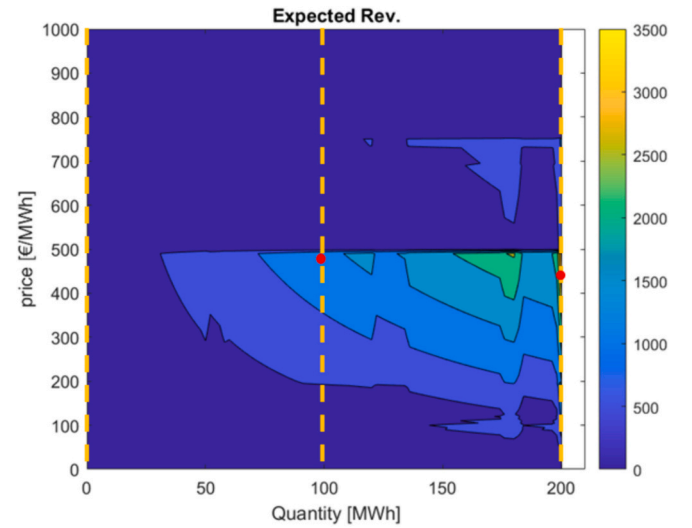


Fig. 8. Expected ASM revenues [EUR] contour vs the offered quantity and price. Yellow dashed lines represent the constraint imposed by the i and j indices. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

machine learning model. Finally, Fig. 8 is the product of the two previous, i.e., the expected revenues.

Within the dispatch MILP optimizer, the offer/bid optimization is carried out on the expected profits and constrained according to the generator load discretization. Thus to the quantity corresponding to $Load_{i,t}^{DAM}$ and $Load_{j,t}^{ASM}$. As a matter of example, in these figures, the optimization is constrained on the yellow dashed lines, since the discretization allows to offer 0, 100, or 200 MWh/h on the ASM. Red dot markers indicate the maximum expected revenues, and it is possible to appreciate that it corresponds neither to the revenues nor to the acceptance probability maxima. The $pr_{el}^{ASM,t}$, corresponding to that optimum, is selected and used directly in eqs. (3) and (4) and to determine $ACC_{prob,t}$ in eqs. (5), (8), and (10).

4. Modeling the propensity to risk

Because of the high potential profits of the ASM, the optimization results may indicate it is worth sacrificing some certain profits on the DAM, reserving some capacity to be offered on the ASM to seek uncertain but high profits. Thus, besides the approach described in Section 2 which maximizes the expected profits (i.e., the multiplication of profits and the related probability), when performing a simultaneous optimization of dispatch of a power generator both on DAM and ASM, it may be worth introducing the propensity to the risk of the operator facing the uncertainty of offers and bids acceptance in the ASM. So that it can decide to adopt a more conservative strategy, preferring DAM certainties and risking in the ASM only if foreseen profits are relevant, or for a more risky approach pursuing higher overall profits even less certain accepting the possibility of a worse scenario. Risk propensity was first introduced to assess the best bidding strategy only on a pay-as-bid market by Sadeh [26], while Kazempour applied the principle of *Markowitz Frontier* to the problem of a DAM-ASM integrated scheduler [27]. In both these approaches the authors modeled the risk introducing in the objective function a penalty taking into account the variability in profits associated with uncertain factors. These terms are multiplied by a factor, ranged $[0, \infty)$, describing the operator's reluctance to risky strategies. Xiao et al. [28] develop a scheduling strategy considering the impacts of windfall profits, proposing a holistic approach to both risk-averse and seeking behavior. This work adopts in the objective function two parameters to describe the risk propensity and adversity separately, these parameters are the weights assigned to the value-at-best and the

conditional value at risk respectively, while the complementary to 1 of these parameters sum is multiplied by the expected profits.

In this paper to include the risk propensity in the MILP formulation reported in Section 2, a single β risk propensity factor is introduced. β ranges from 0, corresponding to a riskless strategy, to 1. It directly impacts the adjusted acceptance probability which is now used in eqs. (5), (8), and (10). The proposed formulation does not affect the weights of the different terms of the objective function, allowing the optimizer to work at its best. It models the irrational attitudes of an optimistic operator, who hopes that his offers will be accepted despite low probabilities, analogous to the buyer of a lottery ticket, and a pessimistic operator, who fears that his offers will be rejected, even if unlikely.

The adjusted acceptance probability $\overline{ACC}_{prob_{ij,t}}$ is then a function of the acceptance probability $ACC_{prob_{ij,t}}$, as returned by the associated model (e.g., the machine learning model described in Section 3 and β according to the following equation:

$$\overline{ACC}_{prob_{ij,t}} = ACC_{prob_{ij,t}}^{\frac{1}{\beta}} - 1 \quad (20)$$

Eq. (20) guarantees that if $\beta = 0$, the power exponent is infinite so that the adjusted probability is 0 and the MILP scheduler never sacrifices the DAM opportunities in favor of ASM. If $\beta = 1$, the power exponent is 0 and the adjusted probability is always 1, so the optimizer considers the ASM as all offers/bids would be accepted. It is worth noticing that bounds to the offer/bid prices are needed, otherwise the optimizer would select infinite prices if \overline{ACC}_{prob} is closer to 1. So, the offer prices have been ceiled at 3000 €/MWh and bids, complaint with the Italian regulatory framework, floored at 0 €/MWh.

Fig. 9 shows how the adjusted probability varies according to eq. (20), it is important to note how $\beta = 0.5$ corresponds to a straight line not affecting the probability returned from the probability prediction model. So $\beta = 0.5$ is a neutral propensity to the risk in this case the objective function is the same as Section 2.

5. Results

The model of optimal dispatch described in Section 2, using the algorithm for the probability prediction of ASM offers/bids selected in Section 3, is here applied to a specific case study to show how it works and its potentialities.

The case study assumes a 400 MW CCGT, and the GT load is discretized to 0%, 45%, 60%, 75%, and 100%. The CCGT off-design and nominal efficiency are modeled as in [29]. The CCGT is assumed to be in

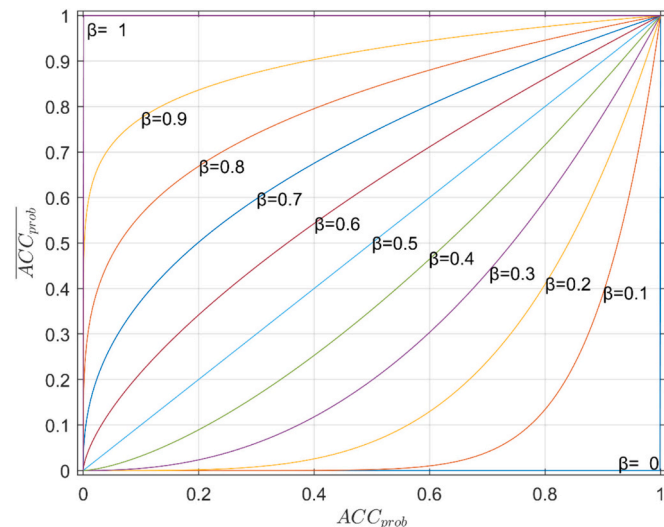


Fig. 9. adjusted acceptance probability trends varying β .

Turbigo (province of Milan, Italy) 43.52°N, 8.74°E, where a real power plant owned by IREN Energia is located.⁴ The location belongs to the NORD electricity market bidding zone and is connected to the high-voltage (380 kV) electricity grid. Input data (electricity price, natural gas price, CO₂ allowance cost, and ambient temperature are the historical real data time series). The scheduling process takes approximately 20 min for one week⁵.

5.1. DAM and ASM simultaneous optimization

First is proposed a comparison between a DAM-only scheduler and the proposed DAM-ASM integrated scheduler on week 15, 2018. This week is selected as it presents a relevant variety of situations to describe all the potential advantages of combined DAM and ASM optimization. At this point, the operator's propensity to the risk is neglected and the objective function is described by eq. (1).

The DAM model of optimal dispatch schedules the CCGT start-up when the sum of profits from the following operating hours can pay the cost of the start-up back. Then the power plant is run full load if the COE is lower than the price of electricity, in such a situation the more power is sold the higher the revenues and the profits, as it can be seen in Fig. 10. If the COE is slightly higher than the electricity price and the power plant is on, full load is still the best option to limit the losses, since lowering the load implies an absolute saving in fuel but a lower efficiency so higher COE and specific losses. The minimum load is adopted when the difference between electricity price and COE is negative, if the overall amount of losses in such an unprofitable period is lower than the start-up cost that should be paid later to exploit new favorable market hours. The power output is not constant neither at full nor at minimum load since it depends on the ambient temperature.

Fig. 11 reports the scheduling optimized to maximize the sum of DAM profits (black line in the figure) and the expected ASM profit, the red line in the figure shows the sum of the two contributions. The blue bars indicate the quantity sold in the DAM, while the white and red bars are the quantity offered or bid in the ASM. The ASM bar red filling is proportional to the probability of acceptance of the relative offer/bids: e.g., a full red bar indicates that such quantity has a 100% probability to be accepted in the ASM, otherwise an almost empty bar indicates that the acceptance probability is low. On April 6, the two scheduling differs because of the previous days operations, according to the optimal DAM-only scheduling, the CCGT is on at April 6 00:00, so it is managed to minimize the losses during low price hours and maximize the profits during profitable hours, according to the strategy described in the previous paragraph. Considering also the ASM, the best option is to present any offers in the DAM until the last high price hours, in facts is more profitable to offer energy on the ASM, if the offer will be accepted the profits will be higher if it will be rejected the plants will remain off and the start-up cost is avoided.

April 7 is characterized by unprofitable prices in the DAM, however, the model predicts a good probability (40–50%) of bid acceptance in the ASM, so the strategy is to sell in DAM and then bid to buy the energy at a lower price on the services market.

Another interesting scheduling is the one for the second half of April 9. These hours are characterized by good prices in the DAM, so the DAM-optimizer scheduled to run the power plant full load, nevertheless prices on the ASM can be much higher so the integrated optimizer imposes to sell an intermediate load in the DAM and offer the residual reserve in the ASM. Even if the probability of acceptance is not high, ASM prices are higher, so it is worth risking losing certain profits from the DAM and betting for offer acceptance on the ASM. If the acceptance probability reaches zero then is better to sell the whole available capacity in the

⁴ In Turbigo is actually located a 2 + 1 CCGT with an installed capacity of 800 MW, however, to reduce the complexity of the problem it is imposed to be a 1 + 1 configuration of 400 MW.

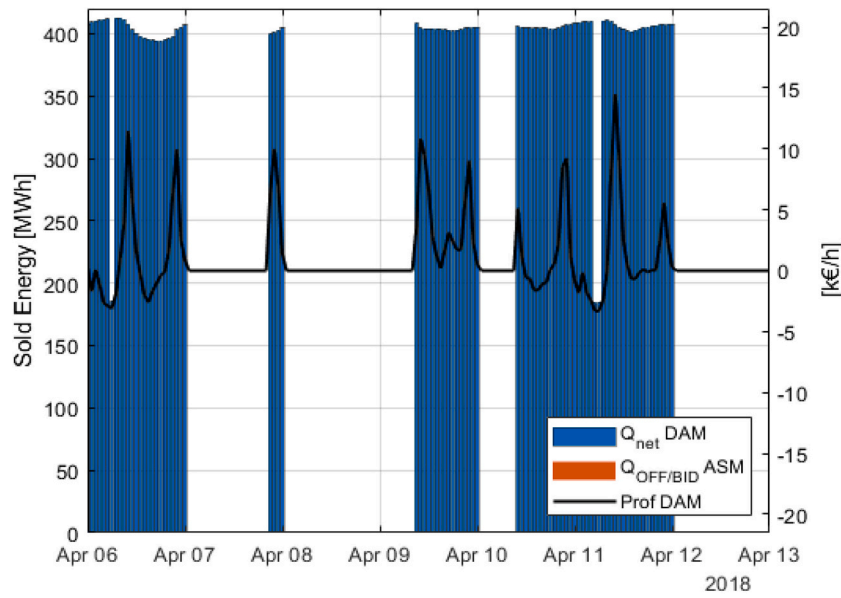


Fig. 10. DAM-only scheduler, between April 6 and April 13, 2018.

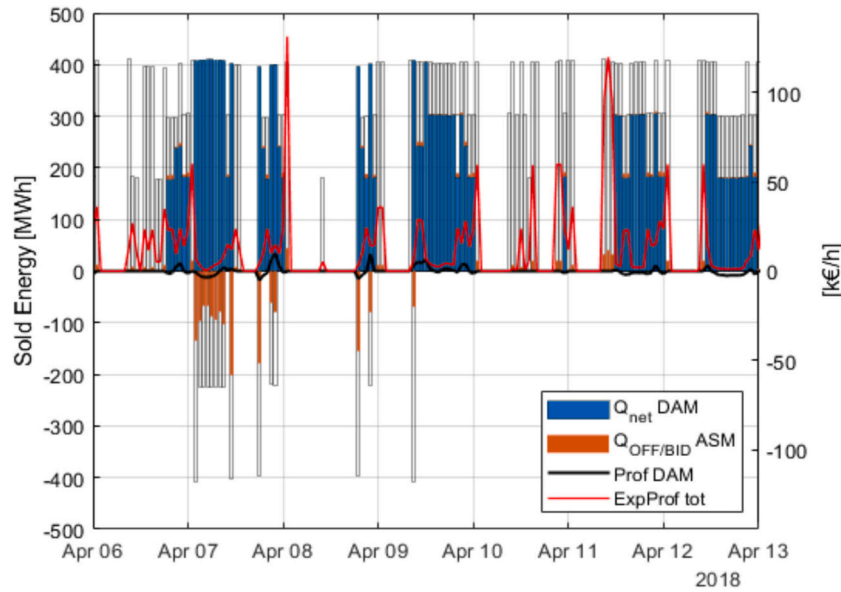


Fig. 11. DAM and ASM integrated scheduler, between April 6 and April 13, 2018.

DAM, as it happens for two hours in the morning.

On April 11 and 12 the strategy is similar but in many hours is better to offer only a fraction of the residual available capacity, since offering the whole amount would imply a lower probability in this case. On April 11 is interesting to look at how the start-up time is delayed, the DAM-only optimizer switches it on earlier to exploit a high-price hour at the end of the morning and this justifies the slight losses in the central day hours. But considering also the ASM the best option is to address the ASM, where the prices are higher and probability acceptance is moderately high, previously the profitability of the DAM is stable to sell in this market the baseload.

Finally, on April 10 the power plant is off, but some energy is offered in the ASM, and the expected profits are considerable, up to 50 k€/h while the cost associated to the start-up event (that depends also on the cost of gas and the electricity price) are about 25 k€. Thus, the viability of this strategy depends on the ASM potential price, sometimes, e.g., on April 8 except for one hour, the best option is to not offer any quantity,

since even in case of acceptance the profits would be lower than the start-up cost.

5.2. Risk propensity analysis

The last results focus on the impact of the propensity to risk that is investigated considering the objective function as described in Section 4. The first week of each month of 2021 is simulated to make a projection of techno-economic indicators on a yearly basis, on each time interval is performed a Montecarlo simulation with 200,000 samples to simulate the different scenarios (i.e., if each offer/bid is actually accepted or rejected). 200,000 uniformly distributed numbers between 0 and 1 are generated for each time interval, and the relative offer/bid is considered accepted in all those scenarios in which the acceptance probability (as returned by the probability prediction model and not adjusted by eq. (20)) is higher than the generated number. The described approach allows generating as an output, not only the generally expected profits as

in eq. (1) but the probability distribution of profits, whose median, quartiles, 5th and 95th percentiles are assumed to be the key indicators.

For each week, 12 optimizations have been run, since β ranges from 0 to 1 with step 0.1, and a benchmark DAM-only optimization is run to quantify the contribution of ASM to the profits. In the last case, the output is not a probability distribution but a single number. Fig. 12 shows a graphical representation of the distribution of the profit at different β .

Table 3 reports the dependency of the main indicators on the risk propensity factor. It must be remarked that if the adjusted probability is used in the objective function, then the actual estimated value ACC_{prob} must be used to compute the expected KPIs of the scheduled strategy.

If the dispatch is optimized only considering the DAM the earning 85.14 M€. If the opportunities in the ASM are considered but no risk is taken ($\beta = 0$) the expected profits increase. This case is equivalent to a subsequential approach to the optimization of dispatch which optimize first the DAM, it is worth noticing that the DAM profits are 85.14 M€ as in the previous case, and then randomly seizing the residual opportunities offering the remaining flexibility margins on the ASM increasing the expected overall profits. The expected value, as in eq. (1), is 92.04 M€, slightly lower than the median of the probability distribution of 92.82 M€.

As β increases, some certain DAM profits decrease in favor of greater, but not certain, profits in the ASM. In the projection of Fig. 12, DAM profits dropped significantly only beyond $\beta = 0.4$, while the increase in ASM profits is appreciable even earlier. In fact, slightly increasing the propensity to the risk from 0 to 0.4, the DAM scheduling is basically not affected but the residual reserve is offered/bidden at more favorable prices and, at this stage, the increase in profits overcomes the decrement in acceptance probability. Moreover, in Fig. 9 is possible to appreciate that transfer curves to the adjusted probability are not evenly spaced and for low β and low ACC_{prob} are very close. The expected profits peak for $\beta = 0.5$, the value in which the objective function is not distorted by the propensity to risk. For this value DAM profits are 24.1% lower than the optimal DAM baseline but the ASM itself (considering the expected

value) is worth 114.4% of the realized profits with respect to the benchmark DAM only optimization, with an overall increase in profits of 90.2%. Increasing further β the drop of profits from DAM is more relevant and is not compensated by the expected increment from the ASM. It is interesting to observe that different percentiles of the profits distribution peak at different β values, with the expected profit and the median profits confirm that $\beta = 0.5$ is the best option when looking for the maximum expected profit. Worst scenarios (i.e., lower percentiles) show better performance in the case of more conservative strategies (lower β). Finally, because of the imposed bound to the offer/bid prices the profits do not change significantly beyond $\beta = 0.7$; It can be noticed that propensity to risk parameter, β , could be used to assess and then balance the expected profits and their variability in order to find a robust solution.

6. Conclusion

In the current electricity market, flexibility is an essential feature of new business models for programmable power plants, storage, and demand response loads. Consequently, is important to develop a methodology able to quantify the real economic value of flexibility. The approach proposed by this paper can be adopted to assess the viability of many solutions oriented to a flexibility enhancement, which is often not certain because, beside the high investment costs that they imply, there are difficulties in making reliable prediction of the different revenues opportunities. Traditional approaches based only on the Day Ahead Market, DAM, may neglect a relevant part of profits coming from the Ancillary Service Market, ASM. Thus, the proposed optimizer may be adopted for two different purposes. The first is the preliminary viability assessment on a long-term to investigate the opportunity to invest in such a resource, installed in such a location. The second is the daily optimization of the existing power plant or storage equipment to operate it at its best.

The MILP model is described in detail, alongside a machine learning approach to estimate the probability of offers/bids acceptance on the ancillary services markets. Then, the developed scheduler is compared

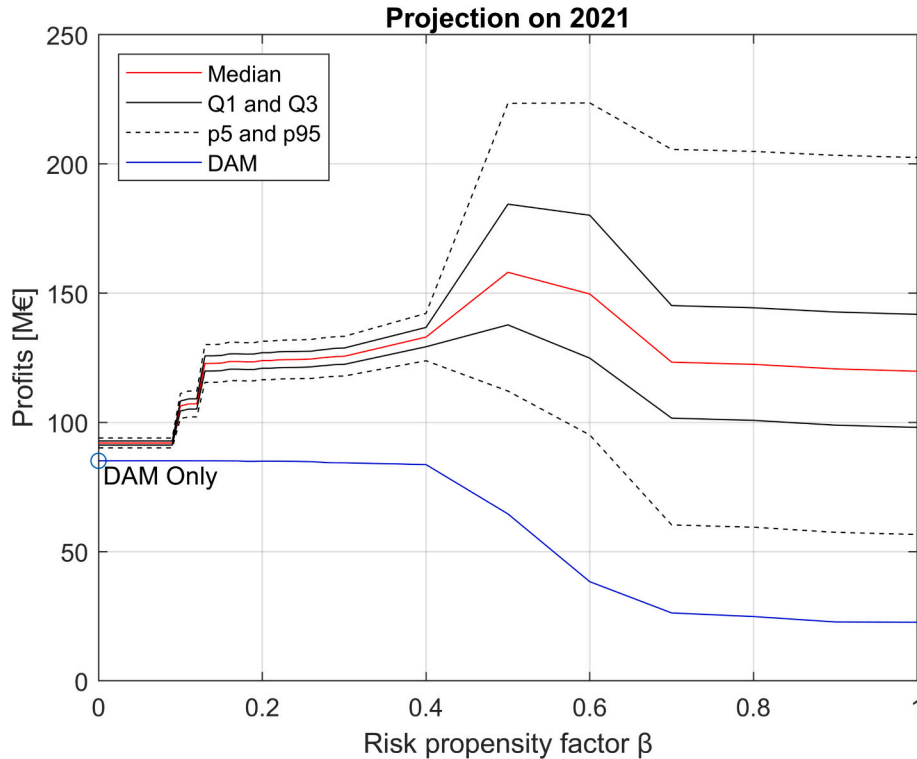


Fig. 12. Impact of risk propensity factor on the profits probability distribution, yearly projection on 2021.

Table 3

Risk propensity factor impact on the main operational indicators, year projection in 2021.

	DAM only	$\beta = 0$	$\beta = 0.1$	$\beta = 0.2$	$\beta = 0.3$	$\beta = 0.4$	$\beta = 0.5$	$\beta = 0.6$	$\beta = 0.7$	$\beta = 0.8$	$\beta = 0.9$	$\beta = 1$
Profits DAM [M€]	85.14	85.14	85.14	84.98	84.36	83.68	64.60	38.37	26.28	24.89	22.80	22.68
Expected Profits ASM [M€]	0.00	6.90	21.20	38.99	41.24	49.33	97.40	115.46	98.69	99.24	99.59	98.82
Expected Profits [M€]	85.14	92.04	106.34	123.97	125.60	133.01	162.00	153.83	124.97	124.12	122.38	121.50
p95 Profits [M€]		92.03	106.37	123.92	125.57	133.01	158.08	149.68	123.30	122.46	120.71	119.80
Q3 Profits [M€]		91.24	104.42	120.95	122.45	129.27	137.73	124.85	101.64	100.79	98.94	98.06
Median Profits [M€]		92.82	108.27	126.98	128.77	136.76	184.43	180.13	145.18	144.35	142.71	141.80
Q1 Profits [M€]		90.14	101.50	116.51	117.94	123.86	112.13	95.12	60.33	59.45	57.47	56.63
p5 Profits [M€]		93.97	111.15	131.43	133.27	142.18	223.44	223.59	205.64	204.84	203.33	202.48

to a traditional DAM MILP scheduler, highlighting the increased opportunities for profits in the ancillary services market. The results show that if these markets are considered to schedule the best power plant dispatch, the flexibility value is properly awarded. While, from the grid operator's point of view is clear how the availability of a programmable power plant to fluctuating load gives a relevant contribution to managing grid congestions, regulating the frequency, and properly meeting the demand. The optimization logic is explained allowing a full understanding of the problem complexity and the developed tool potentialities.

The core element of the proposed scheduling model is the data-driven algorithm for the prediction of a specific offer/bid probability of acceptance. The algorithm must perform good predictions in order to provide a reliable schedule since it is used to optimize the offer/bid price ($nLoad-1$)- $nLoad-t$ times, thus each optimization calls the predictor several times and a fast-to-predict model is essential to keep the optimization time within an acceptable value.

The development of data-driven algorithms for offer/bid acceptance probability is the issue on which future works can focus more. The predictors list can be varied looking for the best possible model, the objective function of the training optimization can be improved by setting the goodness indicator defined by this paper to search for the best hyperparameters of each classifier. Then training optimization should prioritize fast-to-predict models, for this purpose other machine learning algorithms can be tested further than those listed here.

Finally, the last section introduces a risk propensity factor, this parameter is used to account for the risk attitude of the operator, as β increases the expected profits show a maximum while the variability increases since more uncertainties are introduced in the optimal offer strategy. This analysis confirmed that the expected profit found in the neutral optimization, $\beta = 0.5$, represents the maximum obtainable. However, selecting a risk propensity factor lower than 0.5 a lower profit value could be traded for a lower uncertainties if a more prudential strategy is adopted.

Nomenclature

Acronyms and abbreviations.

ACC	Accepted
aFRR	Automatic Frequency Restoration Reserve
ASM	Ancillary Services Market
CCGT	Combined Cycle Gas Turbine
DAM	Day Ahead Market
DSO	Distribution System Operator
EU	European Union
FACC	False ACC
FCR	Frequency Containment Reserve
FREJ	False REJ
GT	Gas Turbine
IDM	Intraday Market
kNN	k-Nearest Neighbors
mFRR	Manual Frequency Restoration Reserve
MILP	Mixed Integer Linear Programming

NB	Naive Bayes
PAB	Pay as Bid
REJ	Rejected
ROC	Receiver Operating Characteristic
RR	Replacement Reserve
SMP	System Marginal Price
ST	Steam Turbine
TACC	True ACC
TREJ	True REJ
TSO	Transmission System Operator

Variables

β	Risk propensity Factor
A	MILP inequalities constraints matrix
ACC _{prob}	Offer Acceptance Probability
AUC	Area under the curve
b	MILP inequalities constraints array
COE	Cost of Electricity
C _{su}	Start-up cost
Err	Error
ExpC _{su}	Expected Start-up cost
ExpProf	Expected Profits
f	MILP optimization vector
FREJR	False REJ rate
m	number of MILP inequality constraints
n	size of MILP solution array
nLoad	number of load discretization steps
nLoad	number of time intervals
nMod	Operational modes
Off _{prob}	Probability of off status
On _{prob}	Probability of on status
p5, p95	5th and 95th percentiles
pp.	Posterior probability
pr	Price
Q	Quantity
Q1, Q3	First and Third quartiles
SU _{prob}	Probability of start-up
T	Threshold value
TACCR	True ACC rate
x	MILP solution array

Subscripts and superscripts

t	time index
i	$Load^{DAM}$ at time t index
j	$Load^{ASM}$ at time t index
k	$Load^{DAM}$ at time $t-1$ index
opt	Optimal
z	$Load^{ASM}$ at time $t-1$ index

CRedit authorship contribution statement

Alberto Vannoni: Writing – original draft, Visualization, Software,

Methodology, Data curation, Conceptualization. **Alessandro Sorce:** Writing – review & editing, Supervision, Methodology.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Alessandro Sorce reports financial support was provided by University of Genoa.

Data availability

Data will be made available on request.

Acknowledgments

This project has received funding from the European Union's Horizon 2022 green research and innovation programme under Grant Agreement

No 101096992, SINNOGENES. (<https://sinnogenes.eu>).



References

- [1] Nosratabadi SM, Hooshmand R-A, Gholipour E. A comprehensive review on microgrid and virtual power plant concepts employed for distributed energy resources scheduling in power systems. *Renew Sustain Energy Rev* 2017;67: 341–63. <https://doi.org/10.1016/j.rser.2016.09.025>.
- [2] Panda S, Mohanty S, Rout PK, Sahu BK. A conceptual review on transformation of micro-grid to virtual power plant: issues, modeling, solutions, and future prospects. *Int J Energy Res* 2022;46:7021–54.
- [3] Ali J, Massucco S, Silvestro F. Architec Virtual Power Plant Ancillary Services CIREC Workshop - Ljubljana, 7–8 June 2018 2018:7–8.
- [4] European Parliament, European Council. Directive (EU) 2019/944 of the European Parliament and of the council of 5 June 2019 on common rules for the internal market for electricity and amending directive 2012/27/EU. 2019.
- [5] Vannoni A, Garcia JA, Guedez R, Sorce A. Ancillary Service Potential for Flexible Combined Cycles. In: ASME Turbo Expo 2021: Turbomachinery Technical Conference and Exposition. Volume 4: Controls, Diagnostics, and Instrumentation; Cycle Innovations; Cycle Innovations: Energy Storage; Education; Electric Power. Virtual, Online. June 7–11, 2021.; 2021. <https://doi.org/10.1115/GT2021-59587>.
- [6] Zhou Y, Wei Z, Sun G, Cheung KW, Zang H, Chen S. A robust optimization approach for integrated community energy system in energy and ancillary service markets. *Energy* 2018;148:1–15. <https://doi.org/10.1016/j.energy.2018.01.078>.
- [7] Kazemi M, Zareipour H, Amjadi N, Rosehart WD, Ehsan M. Operation scheduling of battery storage Systems in Joint Energy and Ancillary Services Markets. *IEEE Trans Sustain Energy* 2017;8:1726–35. <https://doi.org/10.1109/TSTE.2017.2706563>.
- [8] Wang Q, Zhang C, Ding Y, Xydys G, Wang J, Østergaard J. Review of real-time electricity markets for integrating distributed energy resources and demand response. *Appl Energy* 2015;138:695–706. <https://doi.org/10.1016/j.apenergy.2014.10.048>.
- [9] European Commission. COMMISSION REGULATION (EU) 2017/ 2195 - of 23 November 2017 - establishing a guideline on electricity balancing. n.d.
- [10] ENTSO-E. Survey on Ancillary services procurement Balancing market design. 2022. p. 2022.
- [11] Elia Group. Remuneration of mFRR & mFRR Capacity: pay-as-bid vs. pay-as-cleared. 2020.
- [12] Mazzi N, Kazempour J, Pinson P. Price-taker offering strategy in electricity pay-as-bid markets. *IEEE Trans Power Syst* 2018;33:2175–83. <https://doi.org/10.1109/TPWRS.2017.2737322>.
- [13] Ren Y, Galiana FD. Pay-as-bid versus marginal pricing-part I: Strategic generator offers. *IEEE Trans Power Syst* 2004;19:1771–6. <https://doi.org/10.1109/TPWRS.2004.835638>.
- [14] Ren Y, Galiana FD. Pay-as-bid versus marginal pricing-part II: market behavior under strategic generator offers. *IEEE Trans Power Syst* 2004;19:1777–83. <https://doi.org/10.1109/TPWRS.2004.835650>.
- [15] Nosratabadi SM, Hooshmand RA, Gholipour E. Stochastic profit-based scheduling of industrial virtual power plant using the best demand response strategy. *Appl Energy* 2016;164:590–606. <https://doi.org/10.1016/j.apenergy.2015.12.024>.
- [16] Kuttner L. Integrated scheduling and bidding of power and reserve of energy resource aggregators with storage plants. *Appl Energy* 2022;321:119285. <https://doi.org/10.1016/j.apenergy.2022.119285>.
- [17] Swider DJ. Simultaneous bidding in day-ahead auctions for spot energy and power systems reserve. *Int J Elec Power Energy Syst* 2007;29:470–9. <https://doi.org/10.1016/j.ijepes.2006.11.005>.
- [18] Bovera F, Blaco A, Rancilio G, Delfanti M. Assessing the accuracy of different machine learning classification algorithms in forecasting results of Italian ancillary services market. International conference on the European energy market, EEM, vol. vols. 2019-Septe, IEEE; 2019, p. 0–4. doi: <https://doi.org/10.1109/EEM.2019.98916497>.
- [19] Vasylyev A, Vannoni A, Sorce A. Best Practices For Electricity Generators And Energy Storage Optimal Dispatch Problems. ASME Turbo Expo 2023: Turbomachinery Technical Conference and Exposition. Volume 4: Controls, Diagnostics, and Instrumentation; Cycle Innovations; Cycle Innovations: Energy Storage; Education; Electric Power. June 26–30, 2023., Boston: 2023.
- [20] Mathworks. intlinprog. Mixed-integer linear programming (MILP) n.d. <https://it.mathworks.com/help/optim/ug/intlinprog.html#bts3f9f-f>. accessed April 12, 2022.
- [21] Fang T, Lahdelma R. Optimization of combined heat and power production with heat storage based on sliding time window method. *Appl Energy* 2016;162:723–32. <https://doi.org/10.1016/j.apenergy.2015.10.135>.
- [22] Terna. Report on Balancing in accordance with establishing a guideline on electricity balancing. 2022.
- [23] GME Download page n.d. https://www.mercatoelettrico.org/En/Download/DownloadDati.aspx?val=MGP_Prezzi (accessed November 25, 2020).
- [24] Vannoni A. Flexible heat and power generation: market opportunities for combined cycle gas turbines and heat pumps. *Univ Genoa* 2022. doi: <http://hdl.handle.net/11567/1083022>.
- [25] Colbacchini R. Development of a visualization and analysis tool for public domain data of the Italian electricity market. Politecnico di Milano 2014. <https://www.politesi.polimi.it/handle/10589/91024>.
- [26] Sadeh J, Mashhadi HR, Latifi MA. A risk-based approach for bidding strategy in an electricity pay-as-bid auction. *Eur T Electr Power* 2009;19:39–55. <https://doi.org/10.1002/etep.212>.
- [27] Kazempour SJ, Moghaddam MP, Haghifam MR, Yousefi GR. Risk-constrained dynamic self-scheduling of a pumped-storage plant in the energy and ancillary service markets. *Energy Convers Manage* 2009;50:1368–75. <https://doi.org/10.1016/j.enconman.2009.01.006>.
- [28] Xiao D, Lin Z, Chen H, Hua W, Yan J. Windfall profit-aware stochastic scheduling strategy for industrial virtual power plant with integrated risk-seeking/averse preferences. *Appl Energy* 2024;357. <https://doi.org/10.1016/j.apenergy.2023.122460>.
- [29] Sorce A, Giugno A, Marino D, Piola S, Guedez R. Analysis of a Combined Cycle Exploiting Inlet Conditioning Technologies for Power Modulation. Proceedings of ASME Turbo Expo 2019: Turbomachinery technical conference and exposition, June 17–21 2019, Phoenix, Arizona, USA: 2019. doi: <https://doi.org/10.1115/GT2019-91541>.
- [30] Rancilio G. Abate the barriers: opening the electricity markets to storage and RES. Politecnico di Milano; 2022. <https://www.politesi.polimi.it/handle/10589/188700>.
- [31] Schwidtal JM, Agostini M, Bignucolo F, Coppo M, Garengo P, Lorenzoni A. Integration of flexibility from distributed energy resources: mapping the innovative italian pilot project UVAM. *Energies (Basel)* 2021;14. <https://doi.org/10.3390/en14071910>.
- [32] Terna. Codice di Rete Allegato A15 -PARTECIPAZIONE ALLA REGOLAZIONE DI FREQUENZA E FREQUENZA-POTENZA. 2021.