Electric Power Systems Research 194 (2021) 107095

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

Reducing fossil fuel-based generation: Impact on wholesale electricity market prices in the North-Italy bidding zone

Marco G. Flammini [a], Giuseppe Prettico [a,*], Andrea Mazza [b], Gianfranco Chicco [b]

[a] European Commission, Joint Research Centre (JRC), Ispra, VA, Italy
[b] Politecnico di Torino, Dipartimento Energia, "Galileo Ferraris", Corso Duca degli Abruzzi, 24, 10129 Torino, TO, Italy

ARTICLE INFO

Keywords:
Electricity price
Electricity market
Energy scenarios
Generator bidding
Machine learning
Distribution sampling

ABSTRACT

Decarbonisation policies aim at reducing fossil fuel based generation in favour of cleaner renewable energy sources. Changes in the generation mix to supply future electricity demand will require tools capable to emulate the bidding behaviour of new generation plants. Price forecasting tools lacking this feature and only based on historical data time series might soon become not satisfactory for this scope. This paper presents a methodology that, by considering hourly electricity generation offers (price, volumes) datasets, allows simulating future electricity wholesale's prices. This is done by taking into account new generation units and the dismissing of old (coal-based) units according to the demand and generation forecasts in the European Ten Year Network Development Plan (TYNDP) 2030 scenarios. Machine learning, clustering and distribution sampling techniques are used in this work to finally estimate prices distribution in 2030 in the biggest bidding zone of the Italian market. The results suggest that the prices obtained in the different scenarios do converge to those estimated by the TYNDP. The approach used bypasses the need to have access to all the transactions of a given market. Probability distributions are in fact enough in the proposed methodology to achieve similar results to those based on full knowledge of transaction datasets.

1. Introduction

Increasing concerns over global warming have induced important technological changes aimed at making society living more sustainable and cities less polluted [42]. Scientific evidences [51] are calling for urgent cross-sector solutions, such as heating/cooling and transport electrification, among others, to cut emissions and mitigate climate change.

For the European Union, the 2030 climate and energy framework package [24] set the targets for the year 2030 to a 40% reduction in the global emissions from greenhouse gas compared to 1990 levels, 27% minimum Renewable Energy Sources (RES) share in gross final energy consumption, and 27% minimum energy savings compared to business-as-usual scenario. Further updates of the European target for 2030 increased to 32% the minimum RES share in gross final energy consumption [27].

The emerging situations that could arise in the future are extremely variable. For this reason, probabilistic models where the uncertain parameters are characterised by probability distributions and the results are presented as probability density functions of the observed variables, might result ineffective to tackle such problems. In practice, large-scale uncertainty needs to be addressed through *scenario analyses*, in which different scenarios are analysed in parallel and, if needed, the results are then merged by giving a user-defined weighting factor to each scenario.

In recent years, future energy scenarios have been built for whole continents. In Europe, the Ten Year Network Development Plan (TYNDP) defined by ENTSO-e – the European Network of Transmission System Operators for Electricity [20], sets specific scenarios for the years 2025, 2030 and 2040.

The International Renewable Energy Agency (IRENA), whose renewable energy roadmap (REmap) programme includes 70 countries worldwide, which account for about 90% of the global energy use, has also set further scenarios [13, 43].

Generally, the designed scenarios take into account decarbonisation objectives and the consequent deployment of RES, trying to limit the costs of the energy transition and enabling access to electricity in a secure and continuous manner [20]. Decarbonisation implies the reduction of the generation based on fossil fuels, in favour of larger levels of RES installation. In several areas, this reduction would lead to substantial changes in the outcomes of the electricity markets, due to the

* Corresponding author.
E-mail address: Giuseppe.PRETTICO@ec.europa.eu (G. Prettico).

https://doi.org/10.1016/j.epsr.2021.107095
Received 20 August 2020; Received in revised form 21 December 2020; Accepted 4 February 2021
Available online 24 February 2021

displacement of conventional power plants in favour of RES-based generation with null or low bid prices and thus higher despatch priority. However, the diffusion of converter-interfaced RES exacerbates other issues that require the provision of ancillary services to a larger extent, such as lower generation availability [39, 48, 58], possible higher economic volatility [54], as well as reduction of the system inertia needed for smoothing dynamic phenomena in the networks [1, 29].

In the transition towards a more decarbonised electricity supply, the RES diffusion will have a significant growth when RES will reach cost parity with fossil fuel production. However, in this process it is also important to consider that the fossil fuel prices could be adapted to the future conditions, thus slowing down the decarbonisation process with respect to what could be foreseen at present [32]. In the discussion on the policies that can support the decarbonisation process, the options considered [65] have been carbon tax, innovation subsidies, incentives for RES, and elimination of current subsidies to support fossil fuels. Two viable options have been identified, one with high carbon tax and innovation subsidies for RES, and the other one with lower carbon tax and elimination of existing subsidies to fossil fuels; the latter has been deemed politically more feasible.

The results shown in Cardoso Marques et al. [10] for a study carried out in ten European countries from 1990 to 2014 indicate that RES cannot satisfy electricity consumption without considering electricity generation from fossil fuels. A substantial increase of electricity in the end-use sectors is also expected, mainly with the diffusion of heat pumps and electric vehicles [43]. For the future, the diversification of the energy mix also requires a higher extent of reshaping the electrical demand, including the management of electric mobility and the promotion of self-consumption for prosumers, as well as more energy storage capability, and the establishment of cross-border markets.

The definition of the energy mix in a region depends on various aspects, ranging from fuel availability and cost to the specific policies in place to promote lower environmental impact and higher energy security. It strongly depends on social and political decisions as well, e.g., the acceptance of nuclear power plants in a given region [2]. All these aspects need to be taken into account in the defined scenarios. The energy mix in the considered area (bidding zone), together with the existing market rules, have an important effect on the prices that are cleared hourly in the power exchanges throughout the year.

Because of that, previous data, such as clearing quantities and clearing prices obtained from electricity markets, could only partially assess future estimates on electricity prices. The main reason behind this is the fact that changes in the energy mix cannot be modelled based on extrapolations of previous trends. It becomes thus important to consider not only historical time series, but also the foreseen bidding structure (new suppliers active in the future).

With these considerations in mind, in this paper, focusing on wholesale day-ahead electricity market (DAEM), real hourly data (prices/quantities) coming from hundreds of generators to determine future market clearing prices are analysed. The proposed methodology relies on several steps:

1) Identify unknown generators from their bidding behaviour[1] through an appropriate machine learning technique.
2) Determine[2] meaningful price profiles for clusters of representative power plants (CCGT, Coal, Hydro, Wind and Photovoltaic).
3) Use well-established 2030 scenarios (mainly forecasting hourly demand and generation mixes) as an input of the model.
4) Gather insights on the capabilities of the foreseen generation mix to satisfy future electricity demand and evaluate related costs (future electricity prices).

The methodology is applied to an actual case: the biggest of the six market zones of Italy[3] (i.e., North-Italy). Historical bidding data of each generator have been retrieved from the Gestore dei Mercati Energetici (GME) platform [33]. However, due to the lack of information on the generation technology (i.e., CCGT, Coal, Hydro, Wind and Photovoltaic) in this dataset, a machine learning approach based on XGBoost Classifier algorithm [44] has been used. This makes it possible to reconstruct meaningful results for the generator bidding in the DAEM. Looking forward, three main scenarios are presented to understand how the future generation capacity park will cope with the demand, and how its mix will influence the final day-ahead market price. The scenarios assumed are the ones adopted in the Ten Year Network Development Plan [20], used by ENTSO-e.

The paper is organized as follows. Section 2 describes the different steps of the proposed methodology used for the determination of future DAEM electricity prices. Section 3 shows the methodology applied to a concrete case as the North-Italy bidding zone. Section 4 discusses the main findings. Finally, the last section recalls some highlights and anticipates future lines of research on this matter.

2. Methodology

This section describes the proposed methodology to determine future electricity prices in the scenario of interest. The proposed approach can be considered as a hybrid one, which combines techniques from statistical analysis and computational intelligence. The computational intelligence is limited to the machine-learning algorithm used in the determination of the technology from the bidding behaviour of the power plants in the DAEM. The statistical model accuracy depends on the quality of data and for the specific case on the ability to include important factors, as historical demand, load estimation, weather conditions and fuel prices. These points are discussed in the next sections.

2.1. Overview of the methodology

Differently from models based on the hourly cleared prices of market sessions, the aim is to calculate the cleared market prices by matching all

[1] The estimation of the bidding strategies cannot be based on the knowledge of cost information, because this is private information of the electricity market players that submit their bids and offers for generation and demand. For this purpose, in this paper a procedure for determining the behaviour of the players on the basis of publicly available data on the bids and offers is formulated and applied.

[2] The electricity price determination procedure used in this paper does not belong to the electricity price forecasting models, as there is no historical time series of electricity prices to analyse. For more details on electricity price forecasting, the reader is directed to specific literature reviews [9, 63] and articles that consider solution approaches using statistical methods [15], probabilistic methods [49], approaches based on computational intelligence [46, 62], and hybrid methods [4, 14, 50].

[3] The Italian DAEM is currently divided into six market zones which have the same price if no congestions are present among them The European framework for reviewing the existing bidding zone configuration has been defined by the Regulation known as the Guideline on Capacity Allocation and Congestion Management [25].

the bids of the generators (adjusted quantities) with the hourly demand of electricity in the zone. This fact provides a more holistic view on the generation features typical of the market in exam. Moreover, all the bids become a fundamental input to model the behaviour of the different classes of active generators. The proposed methodology relies on the following steps:

1 *Classification of the generators*: based on a cross-checking exercise with a partial database on generator technologies [59], for 70% of the active power plants in the market in exam their generation technology is identified. For the remaining ones, a machine-learning algorithm is used to match them with one of the potential technologies with very high accuracy. The classification process is explained in detail in Section 2.2.
2 *Price and quantity distributions to estimate generators behaviour*: once every generator has been matched with one of the known technologies (e.g. CCGT, Coal, Hydro, Wind and Photovoltaic), a distribution of common prices/quantities bids is built for each category (e.g. CCGT, Coal, Hydro, Wind and Photovoltaic) based on several variables (capacity, day, hour).
3 *Consider future capacity and demand scenarios*: future scenarios are then used as inputs to model the generation mix and the demand in a future year of reference (e.g., 2030 in the case analysed).
4 *Future plants bidding sampling*: in order to have a more robust bidding behaviour,[4] new plants (to be built) are added to the generation park and price/quantity couples are generated by sampling from the distributions obtained by plants in the same cluster. Improvements in terms of efficiency in the generation are taken into account as well. These usually correspond to a shift of the price distributions, as it will be explained in later sections.
5 *Determination of the electricity price ranges*: based on the new generators bids, and on the future forecast demand, the future market is cleared and the relative prices are derived. This process is done for 52 Wednesdays (which represent a full year) and the final results are averaged over a single day.

Fig. 1 shows a schematic chart of the proposed methodology. Starting from the top left, the two databases on generator bids (in the DAEM) and the database that provides information about the technology of (a subset of) the same generators are matched. Moving to the right (on the same level), based on the machine-learning algorithm described in Section 3.3, the remaining generators are classified from a technological point of view. After that all the generators have been matched with a given technology, the distributions of prices and quantities are estimated for each technology cluster on the basis of the hour of the day for each Wednesday of the year.[5] These distributions become fundamental when future scenarios are considered (bottom part of Fig. 1, for the year 2030). In fact, in the market model, the future offers of the current generators, together with those coming from the new plants (to be built as specified in the considered scenario), are all based on the identified distributions and on some other aspect, e.g., fuel and CO2 price forecasts. After the future DAEM is cleared for each hour of reference, the electricity prices curves are shown and the result discussed (bottom right part of Fig. 1).

2.2. Classification of generators

The database containing all the hourly transactions occurred in the DAEM for a given period of time (e.g., one or more years), including fully accepted, adjusted and rejected offers, is the starting point of the study. Each transaction includes a series of parameters.[6] Since the type of technology used by the generators to produce electricity is not included in this database, other databases are needed to establish this match, in particular, the database that associates the Unit reference number to the technology. To this aim, the ENTSO-e transparency platform [21] can help by providing matching up to 73% of the bids with a specific unit reference number to a technology type. Unfortunately, this matching exercise does not fully cover all the generators active in the market in exam. To cope with it, the missing generators with no technology associated are estimated by using the scalable end-to-end tree boosting system XGBoost Classifier [11]. Detailed information on the XGBoost classifier and on its application on the specific problem is reported in Section 3.3.

2.3. Price and quantity distributions to estimate generators behaviours

After the matching power plant-generation technology is accomplished, a statistical analysis is carried-out for each technology class. The aim of this step is that of obtaining, for each hour of the day, the distribution of price and quantity for each category. The obtained distributions provide useful hints on the generators' behaviour and on their characteristics.

2.4. Consider future capacity and demand scenarios

Three scenarios are considered in this study for the year 2030: two of them are taken from the TYNDP 2018 developed by ENTSO-e, and one from the Directorate General of Energy of the European Commission:

- ST: The Sustainable Transition scenario developed in the TYNDP 2018, is built upon the collection of data from the transmission system operators (TSO).
- DG: The Distributed Generation scenario developed in the TYNDP 2018 is based on the projections from the IEA Energy Outlook 2016, and it foresees significant innovation of small-scale generation and residential storage technologies as a key driver in climate action.
- EUCO: The EUCO scenario, constructed on the European Commission figures is based on the European core policy, created using Price-Induced Market Equilibrium System (PRIMES) model [17] and the EU reference scenario of 2016 [12] with emphasis on large-scale renewables, mostly photovoltaic.

From a general perspective, in the ST scenario, a sustainable CO2 reduction is achieved by replacing coal-burning plants with gas-based plants. The electrification of heat and transport sector is smaller when compared with the other two mentioned scenarios. In the DG scenario, prosumers are at the centre of the energy transition. Smart grids

[4] Establishing an optimal bidding strategy in the electricity market is outside the focus of this paper. Various models, typically based on optimization methods [47], perfect competition [35], or game theory concepts [8, 30, 38, 41, 57], have been used in the literature for optimal bidding. Optimisation methods are based on maximising the profit of an individual player, with the challenging task of estimating the bidding strategies of the other players. Game theory-based methods model the interactions between a player and the other players, by using different models. In the Nash-Cournot model, the quantities produced are considered as strategic variables [7, 28, 34, 53, 61]. The Stackelberg model, similar to the Nash-Cournot model, uses a leader-follower approach in which the leader (dominant) player improves its strategic variables first; then, follower players change their strategic variables [40, 45, 55, 56]. In the supply function equilibrium model, the strategy used by the producers depends on a supply function that links bidding quantity and bidding price [3, 6, 52, 64].

[5] Wednesdays are traditionally assumed in Italy as conventional days to represent the reference system operation.

[6] ur - Unit reference number: identification code of the unit; m - Merit order number, based on market solution algorithm; qa - Offered electricity quantity in MWh; ps - Offer price in €/MWh; pa - Awarded market clearing price in €/MWh; z - Market zone: zone code to which the unit belongs to; x - Status code: accepted/rejected.

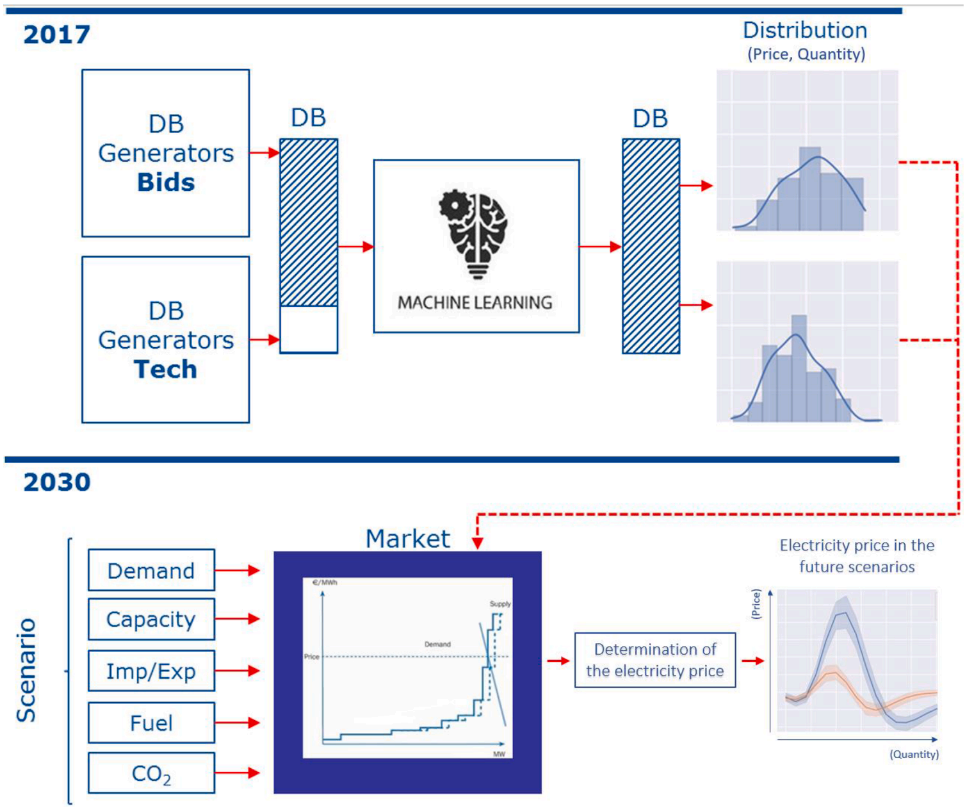

Fig. 1. Schematic view of the proposed methodology.

technology, decentralized development and electric vehicles penetration foster a rapid decarbonisation, which consequently increases the demand compared to EUCO and ST scenarios. In the EUCO scenario, the decarbonisation of the transportation sector is achieved by electric and gas vehicle growth [20].

2.5. Future plants bidding sampling

Every scenario considered foresees changes in the generation mix and in the demand curve. From the generation side, several facts need to be taken into account: new plants (new capacity foreseen) will be built, and those considered too old or too polluting might shut down their activity. Note that in the considered scenarios the additional capacity (e.g., in 2030) is given in aggregated terms. This poses the problem of how to divide this capacity between the new plants to be built (Section 3). Additionally, information on electricity import/export from neighbouring market zones and fuel and CO2 prices changes with respect to the reference year (2017) must be considered. To model the bids of the new plants a statistical sampling method based on the distributions found per each technology class are used.

2.6. Determination of the electricity price ranges

Once the previous steps have been tackled, future supply is matched with the demand hourly profile provided by the scenario under consideration. The matching between supply and demand determines future market clearing prices. It is worth mentioning that the following assumptions have been considered: the revenue mechanism in the DAEM remains unchanged, no generators constraints (ramp up, minimum up time, etc.), no transmission constraints and the weather conditions along the year are comparable with those of the reference year. Both assumptions will be discussed more in detail in the next sections.

3. Case study: North-Italy

In this section, after providing some key figures about the Italian electricity market, the previously introduced methodology is applied to one of the six Italian market zones: the North-Italy zone. The reason behind this choice is that this zone is the most interesting one from an energy point of view. This zone in fact contains the most industrialised Italian regions and has several cross-border connections with other countries.

3.1. Electricity market: the Italian context

The Italian electricity market is divided into four main markets: day-ahead (MGP), intraday (MI), ancillary service (MSD ex-ante) and balancing (MB) [19]. The day-ahead market amounted in 2017 to 17.9 billion euros. In terms of energy volumes, this market is definitely the predominant one, because 85% of the electricity supplied is in fact traded in this market. For this reason, the study presented is focused solely on this market. In the day-ahead market, the generators offers (or bids) are ranked for each hour, day and market zone, on a merit order structure, in ascending order of price, by reporting the amount of energy offered taking into account zonal constraints [37]. Currently, six bidding market zones exist in Italy: i) North, ii) Central North, iii) Central South, iv) South, v) Sicily, and vi) Sardinia. Moreover, there are seven foreign virtual zones interconnected with Italy: i) Austria, ii) Corsica, iii) France, iv) Greece, v) Slovenia, vi) Switzerland, and vii) Malta. These foreign virtual zones are interchanging energy through transmission lines with the different Italian zones. Based on the cleared zonal prices (which is the same for all the market zones if no inter-zonal grid congestion is experienced) the national Italian single price (Prezzo Unico Nazionale, PUN) is obtained through EUPHEMIA (Pan-European Hybrid Electricity Market Integration Algorithm). It calculates the day-ahead electricity price across Europe by allocating cross border transmission capacity [23]. The PUN corresponds to the weighted average of the prices obtained in the six market zones. The zonal price is determined by the marginal technology fixing the price over the zone, and it is the clearing price at which all the accepted supply offers are evaluated. The day-ahead market, which is an auction market, is managed by the Market Operator called Gestore Mercati Energetici (GME) that co-operates with the Italian Transmission System Operator (TSO). Regarding the year 2017 in North-Italy, the average zonal price has been 54.4 €/MWh, with a standard deviation of 18.4 €/MWh and a maximum hourly price of 206.1 €/MWh. The total volume exchanged in the same zone has been of 159.2 TWh for the whole year, with an hourly average value of 18.1 GWh, a standard deviation of 4.4 GWh, and a maximum hourly quantity of 27.9 GWh [36].

By giving a look at the generation mix in the market zone under study, the total installed power of the thermal power plants, divided in Coal and Combined-Cycle Gas Turbines (CCGT), covers respectively 1.7 GW and 24.5 GW [60]. Differently from other market zones, hydro power plants (including water reservoir, run-of-river and pumped storage) have a remarkable role in this zone. Collectively, they reach a total installed power of 16.6 GW. The on-shore wind power plant capacity is limited to only 117 MW. The photovoltaic generation has an installed nominal capacity of 8.7 GW, but only 6.4% of it (560.9 MW), namely, those plants with at least 10 MVA, are connected to the high voltage network and are bidding in the day-ahead market [37]. The remaining 94% PV power production is connected to the distribution grid and thus is seldom observable and non-tradable.

3.2. Electricity market: future scenarios

Table 1 reports, for each scenario introduced in Section 2.4, the total amount of power foreseen for the year 2030 per technology type. The first column (2017) reports the current nominal power installed in the North-Italy zone for each technology (rows). As it can be observed from Table 1, in all the paved scenarios, the on-shore wind installation is limited to around 300 MW. Regarding the hydro power plant production, it seems to reach a saturation level at 19,455 MW for all the considered scenarios. An important fact emerges as well: given the expected decrease of capacity coming from coal power plants, photovoltaic plants (but not only that) will need to be massively installed in the North-Italy zone. Among the three scenarios, the ST scenario is the only exception with respect to this. Despite the paved changes, in terms of energy mix, the thermal sector will still account in this scenario for around 30% of the energy production [22].

3.3. Application of the methodology

This section goes through the different steps of the proposed methodology applied to an actual case. All the relevant details are provided here.

3.3.1. Identification of the generators technology

From the GME website is possible to gather hourly information about generators bidding and about hourly demand. For the year 2017, the dataset contains 13 million transactions, 32% of which are rejected offers. Meaning that, these offers were above the cleared price in one or more time intervals. As mentioned, due to the lack of information about the technology of generation of each supplier in this dataset, the GAUDI[7] dataset from Terna – the Italian TSO [59], and the ENTSO-e transparency platform [21] were used to match up to around 73% of the suppliers with a unique generation technology. This partial dataset is indicated in the following as the "known dataset". The remaining suppliers' technologies have been inferred through the implementation of an XGBoost Classifier. Originally developed in C++ language, XGBoost was chosen over other classifiers due to its faster and more reliable performance when properly tuned [5]. This machine learning method is a decision tree model that allows a classification of the sample (with respect to a target variable) based on a set of input features. Tree models present a high flexibility which make them capable of capturing complex non-linear relationships. On the other side, they are prone to memorising the noise present in a dataset. To reduce this tendency, ensemble methods are implemented in the used algorithms. With reference to this, XGBoost relies upon pruning techniques to avoid the over-fitting of the dataset. Over-fitting would in fact imply that the model obtained from the training data is too close to them and not reliable for new data to be classified (or predicted) based on them. More in detail, the model itself aims at predicting the Y target variable based on the multiple features inputs X_i of the dataset. The goal of the training process is that of finding the best parameters (weights, βi) that best fit Y with X_i. Starting from the root node, containing the whole sample, the splitting into sub-nodes is based on the chosen features. Decision trees use multiple algorithms to decide to split a node in two or more sub-nodes. A benchmarking on the best feature-based splitting is usually done through the calculation of certain functions as Gini, Entropy, Chi Square, etc. Stopping criteria, as minimum samples for a node split, minimum samples for a terminal node (leaf) are used to stop the tree growing in order to avoid the over-fitting issue. While these kinds of halting mechanisms are checked at each level of the tree, pruning techniques rely on a posteriori check. This means that the tree is let to grow and then various levels are compared and pruned based on the function used to do the comparison. Coming back to the specific case, the following features (X_i) have been considered and listed in order of relevance:

- g generation capacity: identification of the code of the unit
- m merit order number
- n_q awarded energy quantity in MWh
- n_p awarded energy price in euro/MWh
- h time interval (hour to which the offer refers)
- u_f utilisation factor of each generator along the year

In a preliminary phase of this study all the features available from the original database of the GME, plus the utilization factor that was derived from the raw data, were taken. After running the classification algorithm, these features were ordered by importance (by assigning to each of them a weight in percentage, the sum of the shares being 100%). Using a threshold of 1% (that is, weight < 1%) several features which had a low impact on the target variable were discarded. As a consequence of that, a reduction of the computational effort of the machine learning algorithm was observed. Among the features removed there were (i) the day of the bids, and (ii) the bilateral contract that indicates if the bid was based on a long-term agreement before the opening of the market.

The output of the XGBoost classifier is a categorical class, which in the case under study corresponds to the possible technology type (CCGT, Coal, etc.). As for other classifiers, XGBoost uses some training data to infer how given input variables are related to the target class (i.e., the technology type). To this aim, part of the known dataset is used to train the classifier. The accuracy of the model has afterwards been estimated by means of a k-Fold Cross-Validation procedure. With respect to other methods, this procedure offers less biased and less optimistic estimates of the model skills. This cross validation works as follows: it shuffles the

Table 1
Amount of total capacity divided per technology in three presented scenarios.

	2017	2030_{ST}	2030_{DG}	2030_{EUCO}
Coal (MW)	1689	560	0	935
CCGT (MW)	16,938	18,067	18,607	22,838
Hydro (MW)	16,630	19,455	19,455	19,455
Wind (MW)	117	312	312	300
PV (MW)	8703	10,382	22,377	15,281
Demand (TWh)	159	206	214	179

[7] This dataset contains the data of the registered production units in Italy.

dataset randomly; splits it into k groups; each group $k = 1...K$ is divided in training and test; finally, the model is fit and evaluated [44].

Fig. 2 shows a scheme of the machine learning process. The process is subdivided into four phases: pre-processing, learning, evaluation and prediction. In the pre-processing the known dataset (73%) is split into training (80%) and test set (20%). In the learning phase, a weight β_i is assigned to each feature X_i. The internal function evaluates, through an iterative process, the possible target technology type based on the weights given to each feature. After the model has been trained, the test set is used to calculate the model accuracy.

The validation of the results is carried out by constructing the confusion matrix, which helps visualise the performance of the algorithm. The confusion matrix is the metric used to measure the performance of the classifier. It is used on the XGBoost to understand its accuracy and performance based on the training model. In the case under study, the confusion matrix has size 8 × 8, since 8 are the meaningful variables/categories. As columns indicates the predicted values and rows indicate the real ones, on the main diagonal are represented those predictions (in percentage) which correctly matched the corresponding real values. For instance, obtaining a diagonal matrix would correspond to the case of perfect prediction. Outside the diagonal are represented instead those predictions which were incorrectly classified as a different category. For instance, looking at category 4 (Wind generation), the diagonal value (element [4,4] in the confusion matrix) shows that 90% of the times the bidding behaviour of a UNIT_REFERENCE_NO variable was recognized and assigned correctly to the corresponding category (Wind). On the same row the confusion matrix indicates that 6% of the times Wind bidding behaviour was confused with a Geothermal plant, 1% of the times respectively with CCGT, Hydro Pumped and Hydro Pumped Storage.

Table 2 highlights the most relevant features needed to be able to detect correctly the type of technology.

The initial GME dataset was thus split into two portions, the known dataset (73%) and the unknown dataset (27%). Additionally, the known dataset was divided into a training set (80%) and a test set (20%). This procedure is quite common in machine learning in order to train the algorithm in use. Based on it, the XGBoost classifier model was fit on this 80% of the dataset and its accuracy was tested on the remaining known 20%. The model accuracy η reaches 91% (with a standard deviation of 3% on the known dataset). In a nutshell, this means that the proposed algorithm is able to detect correctly 9 out of 10 power plants only based on their bids in the market. For the sake of clarity, Fig. 3 shows the confusion matrix for each identified technology. The matching between the technologies and the numbers reported in the figure is the following: 0 for Coal, 1 for CCGT, 2 for total Hydro, 3 for Geo-thermal, 4 for Wind, 5 for PV, 6 for Hydro Pumped, 7 for Hydro run-of-river and 8 for Hydro Pumped Storage.

After all the suppliers in the dataset have been assigned a generation technology from the classifier, another validation exercise has been done by comparing the total power installed in the zone (which is known to the TSO) with that coming from the obtained dataset. Table 3 highlights this benchmark. The comparison shows that the dataset values coming from the classification are slightly lower than those of the TSO. This is due in our opinion to the fact that the total capacity for each technology is calculated by summing up the maximum energy offered during the year by each unit (or supplier). This difference is higher for thermal generators, and it might be attributed to current Grid Codes. According to the latter, thermal generators are generally not allowed to bid all their rated capacity which might be needed for balancing services which ensure the security of supply, e.g., frequency containment reserve [26].

With respect to the PV technology, it is worth mentioning that since 94% of the 8703 MW installed have a capacity below 10 MVA, they do not bid in the DAEM [37]. For this reason, the value shown in Table 3 corresponds to 6.4% of the one listed in the first column of Table 1.

3.3.2. Suppliers' bid behaviour analysis

In order to reconstruct the day-ahead market for the year 2030, it is fundamental to give a close look at the current suppliers' features. Currently, there are 617 generators in the North-Italy zone bidding in the day-ahead market, 17% of which from neighbouring countries (Austria, France, Slovenia and Switzerland). In all the presented scenarios for 2030, Italy is an importing country, with a slight average (all scenarios) increase of the import to 20.6%, compared to 18.7% in 2017. For this reason, in the following, the same profiles of imported and exported energy have been considered multiplied only by a coefficient for each foreign market zone as specified in the TYNDP 2018.

Fig. 4 shows the average normalised load demand for all the Wednesdays in the year 2017 supplied from national generators, that is, no cross-border imports have been considered. An important feature of the machine learning algorithm is that it is able to capture the hourly share for each technology. In Fig. 4, the reddish colours are used for technologies that can be grouped under the thermal power plants category. Blue shades are used to group hydro, pumped, pumped storage and wind technologies, and finally orange is used for PV generation. Differently, from the other Italian bidding zones, the North-Italy zone is characterised by a low share of renewable technologies connected at the high voltage level. In fact, an almost constant level of energy supply from thermal power plants can be seen in the same figure (between 20% and 40% of the average demand). The residual load, defined as the difference between total demand and the generation coming renewable technologies, has a relatively flat profile. This seems to indicate that the thermal power plant operation is not currently stressed (steep increases or decreases of the generated electricity) by renewables in this market zone. Looking forward to the year 2030 (Table 1), expected higher shares of renewables could definitely have an impact on traditional generators' operation. However, it can still be assumed that this fact will not affect much the thermal power plant bidding.

Given these premises, below it is explained how the distributions of price/quantity pairs have been derived for the different classes of generator active in the North-Italy bidding zone.

It is worth mentioning that differently from other European Power Exchange markets, e.g., NordPool, [23], in Italy there is no block order bid option. A block order is a conditional offer (all or nothing) that a market participant does for a set of consecutive periods (more than 1 hour). This is done from thermal producers in order to recover start-up costs and other costs, which could not be recovered otherwise. Due to the lack of this option, suppliers in Italy bid at 0 €/MWh for some hours to ensure their consecutive production. In the dataset used, bids of this kind especially for CCGT and Coal, amount to 47% and 32% respectively of the total. To avoid biased in the derived distributions, bids at 0 €/MWh have been neglected. Figs. 5 and 6 show the resulting price/quantity distributions for these two classes of generators (CCGT and Coal). In particular, Fig. 4 reports the price/quantity probability distribution profile in 2017 based on all the offers from CCGT. The figure shows on the horizontal axis the quantity supplied in MWh, and the corresponding price offered in €/MWh on the vertical axis. Prices in the range of 30 - 50 €/MWh occur 55% of the times. Furthermore, 38% of the bids do not exceed 25 MWh.

From Fig. 5, three main peaks can be identified: the first, the biggest one at 10 MWh, and other two smaller ones at 50 MWh and 130 MWh. These three values suggest which are the three most common sizes of CCGT plants active in the zone considered. Then the same magnitudes are considered to simulate the representative power plants to be built in 2030. Considering only three sizes of plants per technology helps us speed up the computation to match future supply and demand. Note that this factor can be changed if one wants to consider more than three sizes.

Moving to coal power plants, Fig. 6 illustrates the price/quantity probability distribution for the set of all active generators. Differently from CCGT, price and quantity ranges are much lower in this case. Almost 90% of the bids lay in the range 30 - 50 €/MWh and the average quantity is close to 11 MWh. Concerning PV power plants, they always

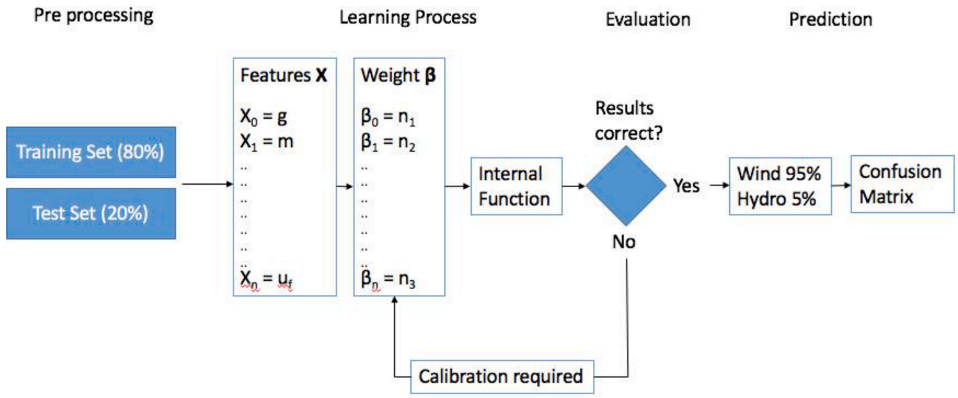

Fig. 2. Schematic view of machine learning process.

Table 2
Machine learning input parameter importance to predict the technology type.

Feature relevance	Parameter (%)
u_f utilization factor	35.8
g generation capacity	23.5
n_p energy price	15.6
m merit order	14.8
n_q awarded quantity	7.5
h interval	2.8

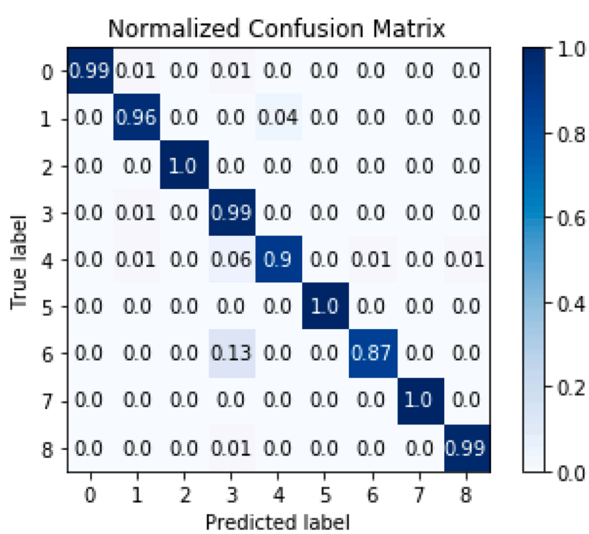

Fig. 3. Confusion matrix of the XGBoost classifier.

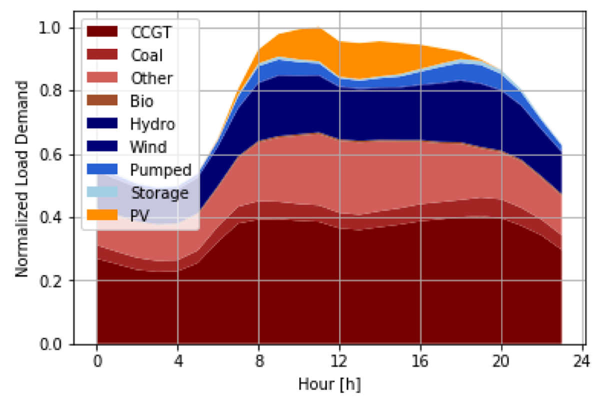

Fig. 4. Normalized load demand, not considering cross border imports.

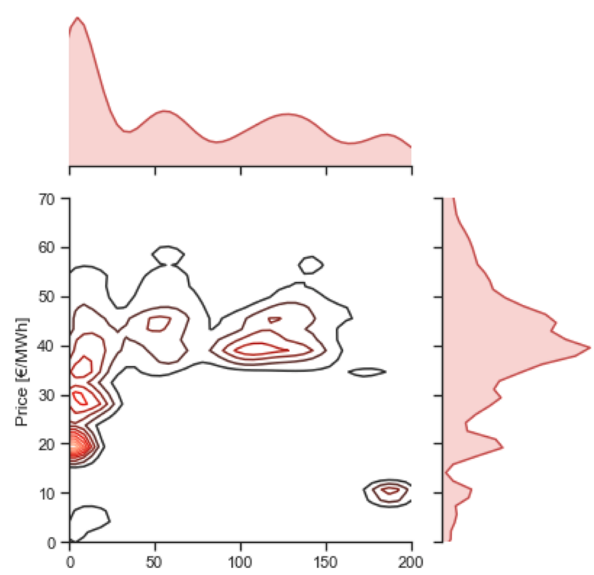

Fig. 5. Price/quantity probability distribution of CCGT in the year 2017 in the North-Italy market zone.

Table 3
Installed capacity benchmark between the dataset used after XGBoost and Terna's data (year 2017).

Technology	Dataset	Terna
Coal (MW)	1550	1689
CCGT (MW)	16,130	16,938
Hydro(MW)	16,580	16,630
Wind (MW)	117	118
PV (MW	550	556

bid at 0 €/MWh, thus having the highest acceptance priority in the market. By looking at Figs. 5 and 6, it can be checked from the dataset that the most dominant price peaks, for both CCGT and Coal, coincide with the average clearing price of the North-Italy zone: 54.4 €/MWh.

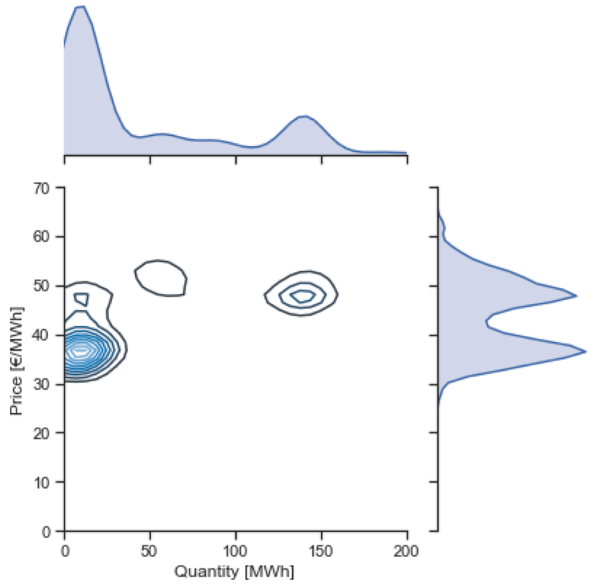

Fig. 6. Price/quantity probability distribution of Coal in the year 2017 in the North-Italy market zone.

This information seems to suggest that these plants are those setting the day-ahead market clearing price in this zone.[8]

With respect to hydro power plants, the most frequent peaks coincide with 5, 20 and 35 €/MWh, while wind power plants, bid on average at around 15 €/MWh. These power plants are subject to premium for difference incentives, as established in two Italian Decrees (Decreto Ministeriale - DM 6/2/2012, and DM 23/6/2016). These incentives have two different mechanisms, for a new power plant (in €/MWh) or for a refurbishment (in €/kW). The incentives are categorised based on the power plant size. The average incentive, based on the capacity of the power plant in the North-Italy zone, is 55 €/MWh. This incentive ends up in two possible situations:

- first, if the bid price is below the zonal market price, an extra revenue is passed to the electricity bills to cover the gap with the zonal market price [36];
- second, if the bid price is above the zonal market price, no extra remuneration is provided (max value 205 €/MWh).

In the latter case, this may indicate that wind power plants are well aware about their position in the merit order market situation, and thus are exploiting the maximum profit out of it [18].

Based on the categorisation of the different set of suppliers active in the market, one of the first issues tackled is that of reducing the number of generators to a smaller number per each generation technology. After this mapping is done, one can sample from a smaller number of the built distributions (price, quantity, hour, …). For instance, for the CCGT, the representative power plants have been grouped according to three capacity sizes (0 - 25 MW, 25 - 75 MW, > 75 MW) identified in the previous section. As stressed, this partitioning has been based on the three peaks observed in CCGT power plants (Fig. 5). This means that each CCGT generator is mapped to one of these three types in such a way that only a tractable number of discrete frequency distributions is used. To validate the proposed sampling approach, a Monte-Carlo simulation has been run (with 1000 runs), for 52 Wednesdays in 2017. The results are compared with the historical data of 2017. In a nutshell, the real offers are substituted by a random sampling over the probability distributions obtained for each class of generator. The results are shown in Fig. 7.

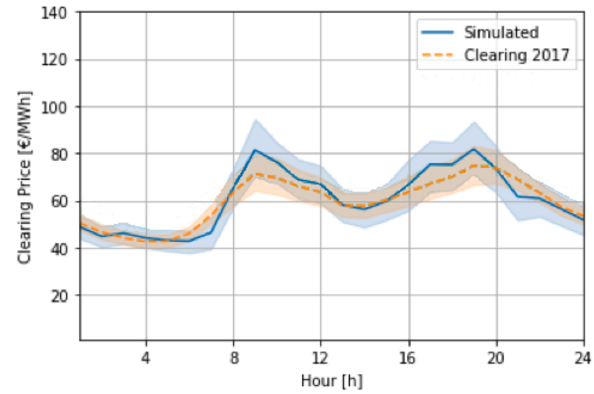

Fig. 7. Price results comparison between simulated and real data from 2017 to 2018.

The coloured bands, in light orange, around the average values contain the historical data of the years 2017. Moreover, to further validate the results of the proposed sampling approach, the total cost of the simulated Wednesdays have been compared with the historical data of the year 2017. Regarding the year 2017, the total cost in the day-ahead market, calculated by multiplying the average hourly zonal clearing price with the total yearly hour demand (for the North-Italy zone), is 1.54 B€. The relative difference with the simulated data is +0.92% for 2017. Based on these considerably small errors, the proposed model is assumed to be accurate enough to provide realistic electricity prices in the North-Italy market zone in future scenarios. It is worth stressing that the proposed method has moved from a high number of operators and bids to a simpler system, based on a random sampling over the probability distributions obtained for each class of generator. From a modelling perspective, this approach can reveal quite helpful when dealing with complex systems where for instance the power flow equations needs to be considered.

3.3.3. Supply and demand in 2030

As done for the year 2017, the supply curve for 2030 has been built by sampling from the distributions of the representative power plants. To take into account the different generation mix for the future in the considered scenarios, the number of suppliers' bids has been proportionally increased or reduced. For instance, in line with Table 1, in the case of CCGT for the DG scenario the number of offers is increased by a factor of 1.07 (18,067/16,938) compared to 2017. The opposite case occurs instead for coal power plants, with capacity in the year 2030 considerably lower than in 2017. For this specific case, the bids are sorted in descending order, thus from the most expensive bid to the least one. Then, bids are removed until the energy obtained from the new capacity is reached. By doing so, it is considered that the least efficient power plant has been substituted with a new one, or it has been refur-

Table 4
Fuel and CO2 prices for the different 2030 scenarios.

Scenario	Coal (€/GJ)	CCGT (€/GJ)	CO_2 (€/ton)
2017	2.3	6.1	18
2030_{ST}	2.7	8.8	84.3
2030_{DG}	2.7	6.9	50
2030_{EUCO}	4.3	8.8	27

[8] Unfortunately from the dataset it is not always possible to have the information on the technology setting the clearing price.

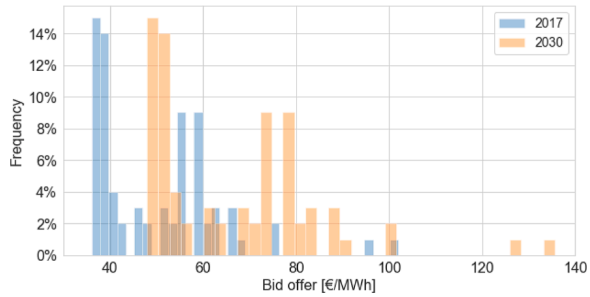

Fig. 8. Frequencies distribution for the CCGT technology for the year 2017 and 2030.

bished.[9] Additionally, apart from capacities changes, fuel and CO_2 costs have been taken into account as well (Table 4). For completeness, the equations showing how future prices are modified based on these factors are reported hereby:

$$p^{2030}_{gen_{s,gt}} = p^{2030}_{est,gt} + \Delta CO_{2_{s,gt}} + \Delta Fuel_{s,gt} \tag{1}$$

$$\Delta CO_{2_{s,gt}} = \left(p^{2030}_{CO_{2\,s,gt}} - p^{2017}_{CO_{2\,s,gt}}\right) \tag{2}$$

$$\Delta Fuel_{s,gt} = \left(p^{2030}_{Fuel_{s,gt}} - p^{2017}_{Fuel_{s,gt}}\right) \tag{3}$$

The emission factors for fuel combustion adopted are 0.341 tonnes of CO_2/MWh and 0.202 tonnes of CO_2/MWh, for Coal and CCGT power plants, respectively [16].

Fuel and CO_2 costs have been considered as follows. After identifying the coal and CCGT power plants that set the marginal price in the year 2017, as provided by GME for each hour of the year [33], fuel and CO_2 costs from the year 2017 were subtracted from the offered prices. The remaining values were considered as the average cost of these power plants. After the subtraction, the average cost accounts for 71% and 62% of the offered prices for coal and CCGT, respectively. From Table 4 it is evident, particularly in the ST scenario, that the CO_2 price would have a major impact on the costs of Coal and CCGT suppliers. Fig. 8 shows how a frequency price distribution for CCGT is shifted to the right once new elements as fuel price and CO_2 costs are considered.

With respect to the demand side, in the TYNDP three demand profiles were built based on the weather condition of three reference years (Fig. 9). They are respectively 1982 (dry condition), 1984 (normal condition), and 2007 (wet condition). The demand profiles take into account average domestic electrical heat pumps installation of 8%, which may have challenging effects on the distribution grids (ENTSO-e 2020). The EV growth is defined as very high for DG, high for EUCO and moderate for ST [20]. The DG scenario has the highest demand and peak ramp up in the early morning. The average demand amounts to 26.6 GWh, which compared to the year 2017 corresponds to an increase of 25%. Further devices increasing the expected demand are related to the boost in the implementation of heat pumps installation.

Furthermore, the new electricity demand accommodates a very high growth in flexibility demand, which flattens the demand profile during the day [20]. The growth in the morning peak demand can be attributed to the EV recharging occurring at the office location [31], when no specific incentives are given to consumers in order to shift the recharging time. In the future this price or time constraints elements will be key to allow a smarter charging mechanism.

Known the supply and demand, the proposed model to determine the electricity prices matches supply and demand orders in order to estimate the electricity price in the North-Italy market zone. Summarising, the

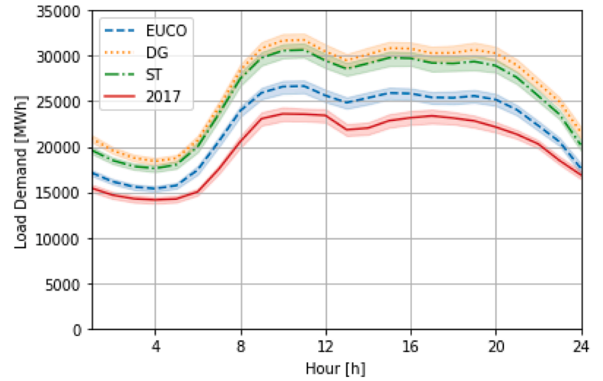

Fig. 9. Time series of the daily demand in the year 2017 and for the three scenarios in 2030.

final electricity price is a function of the demand, suppliers bid prices, import and export, fuel and CO_2 prices.

In particular, the electricity prices in the future have been calculated as follows: for each generator bids (price-quantity pairs) in 2017 new pairs of bids were built taking into account future fuel cost and carbon prices in 2030 as specified in the previous formulas. Additionally, new bids were built based on the new generators (to be built by 2030 as specified by the scenario under study) and on the price-quantity distributions obtained per each technology. The supply offers were then matched with the given demand (which is deterministic for each considered scenario). As a result of the matching the clearing price was obtained. As discussed throughout the manuscript, the bids coming from the new generators (in 2030) were based on the price-quantity distribution of the corresponding technology with the capacity levels opportunely scaled.

4. Results

This section presents the results and shows a qualitative benchmark analysis by comparing the results with those coming from the TYNDP. Firstly, the electricity marginal cost considered by TYNDP in the year 2017 at the European level is 38 €/MWh, which corresponds to 0.7 times the North-Italy market zone price of 54 €/MWh. To compare the prices from the TYNDP 2018 electricity marginal cost forecast at the European level, the same factor of 0.7 should perhaps be considered

The marginal clearing prices, the whole set of price and quantity pairs, of each generator, were sorted in ascending order. The marginal clearing price is by the generator bids matching demand, and this calculation is performed for each hour. A similar rejected ratio (30%) of bids was also obtained in future scenarios. When calculating the hourly

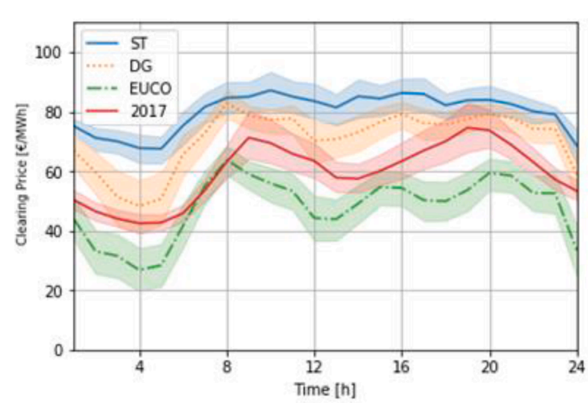

Fig. 10. EPF model price results for ST, DF and EUCO. Year 2017 is reported as reference.

[9] This can be seen as adding a given price for the CO2 emissions produced by the coal power plant.

marginal clearing prices, no physical constraints were taken into account. For the three scenarios, the average yearly prices correspond in the TYNDP 2018 to 86 €/MWh in ST, to 72 €/MWh in DG, and to 67 €/MWh in EUCO. These forecasts show that, even though higher shares of zero-marginal cost generators are included in the generation mix, a potential increase in the range of 19%−37% should be observable when compared to the 2017 electricity prices. This increase considering the 0.7 factor would correspond for the North-Italy zone to 122 €/MWh for ST, to 102 €/MWh for DG and to 95 €/MWh for EUCO scenarios, respectively. However, the prices resulting from the proposed model (Fig. 10) point out that the average marginal electricity price follows the same trend of the ENTSO-e forecasts. In fact, the ST scenario has an average price of 80 €/MWh (with a standard deviation of 17.3 €/MWh); the DG scenario reaches 71 €/MWh (with a standard deviation of 22 €/MWh); while, the EUCO scenario achieves a mean value of 48 €/MWh (with a standard deviation of 25.7 €/MWh), which is even lower than the prices of the year 2017. The high standard deviation for the EUCO scenario can be definitely related to the high share of PV generation in combination with the relatively low growth of demand considered for this scenario. The convergence of the resulting prices between the proposed model and the one used by ENTSO-e seems to suggest as well, that in the latter in ten years time the cross-border congestions will be reduced considerably having a unique price for all Europe.

Table 5 shows the total cost in the day-ahead market for the year 2030 scenarios. The total cost is calculated by multiplying the average hourly price with the total yearly demand supplied in that hour.

The results show an increase compared to 2017, except for the EUCO case where the total price is below the current one. In 2017 the total cost for Wednesday in the day-ahead North-Italy market zone amounted to 1.54 B€. This opposite trend may be caused by the lower demand forecast (compared to DG and ST scenario) as well as to a very high penetration of renewable energy generation. It is important to stress as well that the existing differences with the results of the TYNDP are obvious, due to the fact that is not really possible to compare the results for a single bidding zone (as North-Italy) with those obtained at European level.

Table 5
Fuel and CO_2 prices for the different 2030 scenarios.

Scenarios	Day-Ahead (B€)	St. Dev. (%)
2017	1.54	–
2030_{ST}	2.6	2.5
2030_{DG}	2.4	3.1
2030_{EUCO}	1.4	4.2

5. Conclusions

The energy system will meet the decarbonisation and sustainability goals if reliable tools can help policy makers to take evidence-based decisions. For this reason, methodologies able to provide future electricity prices by considering also the evolution of both the generation park and the demand are needed. This paper has presented a methodology that aims to overcome the limitation linked to the projection of past trends on the prices, by introducing a model aimed at emulating the market mechanism. The information regarding the bids in the biggest zone of the Italian day-ahead market have been matched with the relative technologies. The match has been found partially in a deterministic way, and partially by using a machine learning method called XGBoost Classifier that, after proper training, is able to indicate the technology associated to a particular generation unit. The approach has been validated by comparing the calculated capacities with the capacities existing in 2017, demonstrating its effectiveness. After the identification of the technologies associated to the bids, the most common energy/price pairs offers (organised into probability distributions) have been identified, by also providing information on the fossil fuel plants' size. This information has been essential to model those new generators that will be added (e.g., CCGT) to the generation park according to the three future TYNDP scenarios, or reduced accordingly (e.g., coal plants in all the scenarios). The results of the methodology applied to the North-Italy market zone, show that the average marginal electricity price follows the same trend of the ENTSO-e forecast, namely, the ST scenario has an average price of 80 €/MWh (against 86 €/MWh according to ENTSO-e); the DG scenario reaches 71 €/MWh (against 72 €/MWh); while, the EUCO scenario achieves a mean value of 48 €/MWh (against 67 €/MWh). By considering the EUCO scenario, it is worth nothing that the average price is lower than the one reported by ENTSO-e; however, this particular case shows a high standard deviation (around 25 €/MWh) due to the high share of PV generation in combination with the relatively low growth of demand. The way in which the PV production is considered in the two models can partially explain this difference. Another aspect that demonstrates the possible benefits of using the proposed model for policy making is the evaluation of the electricity costs in the future, comparing the current scenarios by changing the energy mix, considering similar weather conditions and similar bidding strategies with respect to the ones from which the data are taken. These peculiarities of the model will allow energy analysts to identify viable bidding strategies and run future scenarios avoiding the need to get private information on the generation assets. The proposed approach in fact allows taking into account the bidding behaviour of a myriad of generators without knowing exactly each financial transactions but only relying on probability distributions which are used to random sampling from them. Additionally, the proposed method helps simplify the modelling of the generation layer, which becomes fundamental when other layers (for instance, power flow equations, retailers modelling, energy communities preferences, etc.) are considered in the model, which definitely increase its complexity and make the finding of a solution harder.

CRediT Author Statement

Marco G. Flammini: Resources, Conceptualization, Methodology, Software Development, Writing: Editing & Review
Giuseppe Prettico: Conceptualization, Methodology, Formal analysis, Project Management, Writing: Editing & Review
Andrea Mazza: Writing- Review and Editing
Gianfranco Chicco: Writing- Review and Editing

Declaration of Competing Interest

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

References

[1] A. Adrees, J.V. Milanović, P. Mancarella, Effect of inertia heterogeneity on frequency dynamics of low-inertia power systems, IET Gener. Transm. Distrib. 13 (14) (2019) 2951–2958.
[2] T. Atalla, J. Blazquez, L.C. Hunt, B. Manzano, Prices versus policy: an analysis of the drivers of the primary fossil fuel mix, Energy Policy 106 (2017) 536–546.
[3] R. Baldick, R. Grant, E. Kahn, Theory and Application of linear supply function equilibrium in electricity markets, J. Regul. Econ. 25 (2004) 143–167.
[4] A. Bello, D.W. Bunn, J. Reneses, A. Munoz, Medium-term probabilistic forecasting of electricity prices: a hybrid approach, IEEE Trans. Power Syst. 32 (1) (2017) 334–343.
[5] C. Bentéjac, A. Csörgő, G. Martínez-Muñoz, A comparative analysis of gradient boosting algorithms, Artif. Intell. Rev. (2020), https://doi.org/10.1007/s10462-020-09896-5.
[6] C.A. Berry, B.F. Hobbs, W.A. Meroney, Understanding how market power can arise in network competition: a game theoretic approach, Util. Policy 8 (1999) 139–158.
[7] S. Borenstein, J. Bushnell, C. Knittel, A Cournot-Nash Equilibrium Analysis of the New Jersey Electricity Market, University of California Energy Institute, 1997.
[8] J. Bower, D. Bunn, Experimental analysis of the efficiency of uniform-price versus discriminatory auctions in the England and Wales electricity market, J. Econ. Dyn. Control 25 (3–4) (2001) 561–592.

[9] D.W. Bunn, Forecasting loads and prices in competitive power markets, Proc. IEEE 88 (2) (2000) 163–169.
[10] A. Cardoso Marques, J.A. Fuinhas, D.A. Pereira, Have fossil fuels been substituted by renewables? An empirical assessment for 10 European countries, Energy Policy 116 (2018) 257–265.
[11] T. Chen, C. Guestrin, Xgboost: a scalable tree boosting system, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16), 2016, pp. 785–794.
[12] P. Chi, S. Li, C. Xu, T. Zhang, J. Zhao, Y. Liu, Y. Wang, Y. Xie, Prime: a novel processing-in-memory architecture for neural network computation in reram-based main memory, in: Proceedings of the 43rd International Symposium on Computer Architecture, ISCA '16, IEEE Press, Piscataway, NJ, USA, 2016, pp. 27–39.
[13] S. Collins, D. Saygin, J.P. Deane, A. Miketa, L. Gutierrez, B. Ó Gallachóir, D. Gielen, Planning the European power sector transformation: the REmap modelling framework and its insights, Energy Strateg. Rev. 22 (2018) 147–165.
[14] A.J. Conejo, M.A. Plazas, R. Espinola, A.B. Molina, Day-ahead electricity price forecasting using the wavelet transform and ARIMA models, IEEE Trans. Power Syst. 20 (2) (2005) 1035–1042.
[15] J. Contreras, R. Espinola, F.J. Nogales, A.J. Conejo, ARIMA models to predict next-day electricity prices, IEEE Trans. Power Syst. 18 (3) (2003) 1014–1020.
[16] Covenant of Mayors, 2019. Technical Annex to the SEAP template: The emission Factors. Online: www.eumayors.eu/IMG/pdf/technical_annex_en.pdf (Accessed 16.07.2020).
[17] E3M Lab, 2020. The PRIMES model. Online: http://www.e3mlab.eu/e3mlab/index.php?option=com_content&view=category&id=35%3Aprimes&Itemid=80&layout=default&lang=en (Accessed 16.07.2020).
[18] Elemens, 2017. Wind Power in Italy: Inertia or Full Support? (Eolico in Italia: inerzia o full support?, in Italian). Online: https://www.i-com.it/wp-content/uploads/2017/05/eLeMeNS-RINNOVARE-PER-COMPETERE.pdf (Accessed 16.07.2020).
[19] Energy & Strategy Group, 2019. Electricity market report. Politecnico di Milano, Italy, Online: http://www.energystrategy.it/report.html (Accessed 16.07.2020).
[20] ENTSO-e, 2018. Ten-Year Network Development Plan (TYNDP) 2018 – Network Development Plan to 2025, 2030 and 2040. Online: http://tyndp.entsoe.eu/maps-data/ (accessed 16.07.2020).
[21] ENTSO-e, 2020a. Entso-e Platform. Online: https://transparency.entsoe.eu/ (Accessed 16.07.2020).
[22] ENTSO-e, 2020b. Entso-e Platform. https://tyndp.entsoe.eu/maps-data/. (Accessed 16.07.2020).
[23] EPEX Spot, APX, Belpex, Nord Pool Spot, OMIE, GME, OTE, 2013. Euphemia Public description, October 2013. http://www.apxgroup.com/wp-content/uploads/Euphemia-public-description-Nov-20131.pdf (Accessed 16.07.2020).
[24] European Commission, A Policy Framework For Climate and Energy in the Period from 2020 to 2030 Communication from the Commission to the European Parliament, the Council, The European Economic and Social Committee and the Committee of the Regions, Brussels, Belgium, 2014.
[25] European Commission, 2015. Commission Regulation (EU) 2015/1222 of 24 July 2015 Establishing a Guideline on Capacity Allocation and Congestion Management. Online: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32015R1222&from=EN (Accessed 16.07.2020).
[26] European Commission, Commission regulation (EU) 2017/2195 of 23 November 2017 establishing a guideline on electricity balancing, Off. J. Eur. Union (2017).
[27] European Commission, Directive (EU) 2018/2001 of the European parliament and of the council of 11 December 2018 on the promotion of the use of energy from renewable sources, Off. J. Eur. Union (2018) 82–209.
[28] B. Fanzeres, S. Ahmed, A. Street, Robust strategic bidding in auction-based markets, Eur. J. Oper. Res. 272 (31) (2019) 1158–1172.
[29] A. Fernández-Guillamón, E. Gómez-Lázaro, E. Muljadi, Á. Molina-García, Power systems with high renewable energy sources: a review of inertia and frequency control strategies over time, Renew. Sustain. Energy Rev. 115 (2019), 109369.
[30] R.W. Ferrero, S.M. Shahidehpour, V.C. Ramesh, Transaction analysis in deregulated power systems using game theory, IEEE Trans. Power Syst. 12 (3) (1997) 1340–1347.
[31] M.G. Flammini, G. Prettico, A. Julea, G. Fulli, A. Mazza, G. Chicco, Statistical Characterisation of the Real Transaction Data Gathered From Electric Vehicle Charging Stations, 166, Electric Power Systems Research, 2019, pp. 136–150.
[32] E. Foster, M. Contestabile, J. Blazquez, B. Manzano, M. Workman, N. Shah, The unstudied barriers to widespread renewable energy deployment: fossil fuel price responses, Energy Policy 103 (2017) 258–264.
[33] GME, 2020. Gestore Mercati Energetici. Online: www.mercatoelettrico.org (Accessed 16.07.2020).
[34] V.P. Gountis, A.G. Bakirtzis, Efficient determination of Cournot equilibria in electricity markets, IEEE Trans. Power Syst. 19 (4) (2004) 1837–1844.
[35] G. Gross, D.J. Finlay, G. Deltas, Strategic bidding in electricity generation supply markets, in: Proceedings of the IEEE Power Engineering Society Winter Meeting, New York, NY, 1999.
[36] GSE, 2019. Statistical report – Solar photovoltaic (Rapporto statistico – Solare fotovoltaico, in Italian), June 2019. Online: https://www.gse.it/documenti_site/Documenti%20GSE/Rapporti%20statistici/Solare%20Fotovoltaico%20-%20Rapporto%20Statistico%202018.pdf (Accessed 16.07.2020).
[37] GSE, 2020. Gestore Servizi Energetici. Online: www.gse.it/ (Accessed 16.07.2020).
[38] S. Hao, A study of basic bidding strategy in clearing pricing auctions, IEEE Trans. Power Syst. 15 (3) (2000) 975–980.
[39] M. Hermans, K. Bruninx, S. Vitiello, A. Spisto, E. Delarue, Analysis on the interaction between short-term operating reserves and adequacy, Energy Policy 121 (2018) 112–123.
[40] B.F. Hobbs, C.B. Metzler, J. Pang, Strategic gaming analysis for electric power systems: an MPEC approach, IEEE Trans. Power Syst. 15 (2) (2000) 638–645.
[41] B.F. Hobbs, K.A. Kelly, Using game theory to analyze electric transmission pricing policies in the U.S, Euro. J. Oper. Res. 56 (2) (1992) 154–171.
[42] International Energy Agency, 2018. World Energy Outlook 2018, https://www.iea.org/reports/world-energy-outlook-2018 (Accessed 16.07.2020).
[43] IRENA, 2018. Renewable Energy Prospects For the European Union. February 2018, https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2018/Feb/IRENA_REmap_EU_2018.pdf (Accessed 16.07.2020).
[44] G. James, D. Witten, T. Hastie, R. Tibshirani, An Introduction to Statistical Learning: With Applications in R. Volume 103 of Springer Texts in Statistics, Springer, New York, 2014, 2014.
[45] 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.
[46] J. Lago, F. De Ridder, B. De Schutter, Forecasting spot electricity prices: deep learning approaches and empirical comparison of traditional algorithms, Appl. Energy 221 (2018) 386–405.
[47] G. Li, J. Shi, X. Qu, Modeling methods for GenCo bidding strategy optimization in the liberalized electricity spot market – a state-of-the-art review, Energy 36 (8) (2011) 4686–4700.
[48] H. Nosair, F. Bouffard, Flexibility envelopes for power system operational planning, IEEE Trans. Sustain. Energy 6 (3) (2015) 800–809.
[49] J. Nowotarski, R. Weron, Recent advances in electricity price forecasting: a review of probabilistic forecasting, Renew. Sustain. Energy Rev. 81 (2018) 1548–1568.
[50] G. Osorio, M. Lotfi, M. Shafie-khah, V. Campos, J.P.S. Catalão, Hybrid forecasting model for short-term electricity market prices with renewable integration, Sustainability 11 (1) (2018) 1–15.
[51] M.L. Parry, O.F. Canziani, J.P. Palutikof, P. van der Linden, C.E. Hanson, Climate Change 2007: Impacts, Adaptation and Vulnerability, Working Group II Contribution to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, New York, NY, 2007. www.cambridge.org/9780521880107 (Accessed 16.07.2020).
[52] A.G. Petoussis, X.P. Zhang, S.G. Petoussis, K.R. Godfrey, Parameterization of linear supply functions in nonlinear AC electricity market equilibrium models – part I: literature review and equilibrium algorithm, IEEE Trans. Power Syst. 28 (2) (2013) 650–658.
[53] C. Ruiz, A.J. Conejo, J.D. Fuller, S.A. Gabriel, B.F. Hobbs, A tutorial review of complementarity models for decision-making in energy markets, EURO J. Decis. Process. 2 (2014) 91–120.
[54] A. Sapio, Greener, more integrated, and less volatile? A quantile regression analysis of Italian wholesale electricity prices, Energy Policy 126 (2019) 452–469.
[55] M. Shafiekhani, A. Badri, M. Shafie-khah, J.P.S. Catalão, Strategic bidding of virtual power plant in energy markets: a bi-level multi-objective approach, Int. J. Electr. Power Energy Syst. 113 (2019) 208–219.
[56] C.K. Simoglou, P.N. Biskas, A.G. Bakirtzis, Optimal self-scheduling of a dominant power company in electricity markets, Electr. Power Energy Syst. 43 (2012) 640–649.
[57] H. Singh, Introduction to game theory and its application in electric power markets, IEEE Comput. Appl. Power 12 (2) (1999) 18–20.
[58] L. Söder, E. Tómasson, A. Estanqueiro, et al., Review of wind generation within adequacy calculations and capacity markets for different power systems, Renew. Sustain. Energy Rev. 119 (2020), 109540.
[59] Terna, 2020a. Gaudi – Single registry Management of the Plants (Gestione Anagrafica Unica Degli impianti, in Italian), Version 11, February 2020. Online: https://www.terna.it/it/sistema-elettrico/gaudi (Accessed 16.07.2020).
[60] Terna, 2020b. Power Generation Plants (Impianti Di generazione, in Italian). https://www.terna.it/en (Accessed 16.07.2020).
[61] J. Valenzuela, M. Mazumdar, Cournot prices considering generator availability and demand uncertainty, IEEE Trans. Power Syst. 22 (1) (2007) 116–125.
[62] K. Wang, C. Xu, Y. Zhang, S. Guo, A. Zomaya, Robust big data analytics for electricity price forecasting in the smart grid, IEEE Trans. Big Data 5 (1) (2019) 34–45.
[63] R. Weron, Electricity price forecasting: a review of the state-of-the-art with a look into the future, Int. J. Forecast. 30 (4) (2014) 1030–1081.
[64] F. Zaman, S.M. Elsayed, T. Ray, R.A. Sarker, Co-evolutionary approach for strategic bidding in competitive electricity markets, Appl. Soft Comput. 51 (2017) 1–22.
[65] P. Zeppini, J.C.J.M. van den Bergh, Global competition dynamics of fossil fuels and renewable energy under climate policies and peak oil: a behavioural model, Energy Policy 136 (2020), 110907.