

A cluster-based optimization approach to support the participation of an aggregator of a larger number of prosumers in the day-ahead energy market

José Iria^{a,b,*}, Filipe Soares^a

^a Centre for Power and Energy Systems (CPES), INESC TEC, Porto, Portugal

^b Faculty of Engineering of University of Porto (FEUP), Porto, Portugal

ARTICLE INFO

Keywords:

Aggregators
Electricity markets
Prosumers
Flexibility
Clustering

ABSTRACT

Optimizing the participation of a large number of prosumers in the electricity markets is a challenging problem, especially for portfolios with thousands or millions of flexible resources. To address this problem, this paper proposes a cluster-based optimization approach to support an aggregator in the definition of demand and supply bids for the day-ahead energy market. This approach consists of two steps. In the first step, the aggregated flexibility of the entire portfolio is computed by a centroid-based clustering algorithm. In the second step, the supply and demand bids are defined by an optimization model that can assume the form of a deterministic or a two-stage stochastic problem. A case study of 10,000 prosumers from the Iberian market is used to evaluate and compare the performance of the bidding optimization models with and without pre-clustering. The numerical results show that the optimized bidding strategies outperform an inflexible strategy by more than 20% of cost savings. The centroid-based clustering algorithm reduces effectively the execution times of the bidding optimization problems, without affecting the quality of the energy bids.

1. Introduction

1.1. Motivation

The current electricity system does not consider the active participation of small consumers. The consumers are passive elements of the system. However, the recent developments in smart home technologies promise to transform passive consumers into active prosumers. The smart home technologies include smart appliances, sensors, photovoltaic (PV) systems, home energy management systems (HEMSs) equipped with communication, monitoring and control functionalities. These functionalities empower the prosumers and turn them into active and flexible elements of the system. The smart home technologies can render financial benefits to the prosumers since they create an opportunity for aggregators trading the prosumers' flexibility in the electricity markets. However, they also bring new challenges for the aggregators. One of the challenges is to transform the automation functionalities of the smart home technologies into products that can be traded in the electricity markets. Another challenge is to manage a large number of flexible resources in an effective and timely way since typical portfolios can reach thousands or millions of resources. This paper proposes a solution to these two new challenges.

1.2. Related work

The aggregator is a service provider that gathers small prosumers to trade their flexibility in the electricity markets [1,2]. Depending on the requirements of each electricity market, the aggregator may participate in the energy and/or reserve markets. The focus of this paper is the day-ahead (DA) energy market. The related work in this topic covers decision-aid methodologies to support aggregators in the definition of DA energy bids. The optimization of flexible resources, such as electric vehicles (EVs), thermostatically controlled loads (TCLs) and shiftable loads (SLs) is well documented in the literature. For instance, Alahäivälä et al. [3] formulated a deterministic model to optimize demand bids taking into account the flexibility of the TCLs. Chen et al. [4] included the uncertainty of the DA energy prices in the optimization of the demand bids. In the electric mobility context, Bessa and Matos [5,6] proposed two deterministic optimization models to define demand bids. One of the optimization models considers the aggregated information of the EVs, while the second comprises the individual information of the EVs. Both optimization models adopt point forecasts of DA energy prices and driving patterns to optimize the charging consumption of the EVs. Vagropoulos and Bakirtzis [7] proposed a two-stage stochastic model that incorporates the uncertainties of mobility patterns and electricity prices. These three works did not consider vehicle-to-grid (V2G), which was considered by Momber et al. [8].

* Corresponding author at: Centre for Power and Energy Systems (CPES), INESC TEC, Porto, Portugal.

E-mail address: jpiria@inesctec.pt (J. Iria).

<https://doi.org/10.1016/j.epsr.2018.11.022>

Received 16 July 2018; Received in revised form 16 October 2018; Accepted 30 November 2018

Available online 03 January 2019

0378-7796/ © 2018 Elsevier B.V. All rights reserved.

Nomenclature**Abbreviations**

DA	Day-ahead
EV	Electric vehicle
HEMS	Home energy management system
IL	Inflexible load
MIBEL	Iberian electricity market
PV	Photovoltaic
RT	Real-time
SL	Shiftable load
TCL	Thermostatically controlled load
AR	Arrival of the electric vehicle
DE	Departure of the electric vehicle
HE	Heating
CO	Cooling
CU	Curtailement
$\hat{\cdot}$	Point forecast
\wedge, \vee	Discharging, charging
$+, -$	Positive, negative imbalances

Indices and sets

$j \in J$	Scenarios
$t, w \in T$	Time interval
v	Type of load and generator {EV, TCL, SL, IL, PV}
$i \in I^v$	Loads
$k \in K$	Clusters
$n \in N_k$	Objects $N = \{N_1, \dots, N_k\}$
$T^v \subset T$	Sub-set of time intervals
$O \subset$	Set of time intervals of active occupancy

$T^{SL} \subset \bar{T}^{SL} \subset T$ Availability to start, complete a program (sets)

Parameters

C	Thermal capacitance (kWh/°C)
COP	Coefficient of cooling performance
EER	Coefficient of heating performance
D	Number of time intervals of a program
\bar{P}	Maximum electric power (kW)
Pr	Power profile (kW)
$\overline{SOC}, \underline{SOC}$	Maximum, minimum state-of-charge (kWh)
R	Thermal resistance (°C/kW)
x	Object
λ	Price (€/kWh)
π	Probability of occurrence of the scenarios
η	Efficiency
ξ	Weight of the representative flexible loads
ϑ	Heat gains and losses not modeled explicitly (°C)
Δt	Duration of time interval t (1h)
$\bar{\theta}, \underline{\theta}$	Maximum, minimum temperatures (°C)
θ°	Outdoor temperature (°C)

Variables

E	Energy (kWh)
ΔE	Imbalance (kWh)
P	Electric Power (kW)
SOC	State-of-charge (kWh)
θ	Temperature (°C)
ψ	Sets the beginning of a program (binary)
μ	Centroid

Mohsenian-Rad [9] extended the portfolio of the EV aggregator to other SLs, such as dishwashers and washers. Iria et al. [2,10] added generation units to the aggregator's portfolio, in order to exploit synergies between demand and generation resources. The management of generation and demand resources enables the participation of the aggregator in the demand and supply sides of the DA energy market. In the scope of local energy markets, the authors of Refs. [11–13] proposed peer-to-peer approaches to promote energy trading between prosumers. The trading mechanisms rely on the generation surplus of the prosumers to enable the exchange of services between them.

The above mentioned bidding optimization models [2–9] consider the aggregation of a few hundreds of flexible resources. The optimization of a small group of flexible resources does not present a computational problem. On the other hand, optimizing a very large number of individual flexible resources is challenging since it may add millions of variables and constraints to the optimization problems, making them hard to solve or even computationally intractable [14,15]. Various approaches have been proposed to reduce the size of the optimization problems by representing the flexibility of the aggregated resources through generalized batteries [15–17] and polytopes [14,18,19]. Nonetheless, none of these aggregation approaches has been successfully integrated into bidding optimization problems.

1.3. Contributions

This paper proposes a novel cluster-based optimization approach to support an aggregator of prosumers in the definition of demand and supply bids for the DA energy market. This approach consists of two steps. In the first step, the aggregated flexibility of the TCLs, EVs, and SLs is computed by a centroid-based clustering algorithm. In the second step, the supply and demand bids are defined by an optimization model that can

assume the form of a deterministic or a two-stage stochastic problem. The differences between the optimization approaches are how the information of the prosumers is modeled and the number of market stages considered in the optimization process. The two-stage stochastic model optimizes the participation of the aggregator in the DA and real-time (RT) market stages and models the uncertainties of the prosumers through a set of scenarios. The deterministic model considers point forecasted information and optimizes the participation of the aggregator in the DA market. The two bidding optimization models are compared and evaluated.

The proposed cluster-based approach differs from works [2–10] by considering different sources of flexibility (PVs, TCLs, SLs, and EVs) in the joint optimization of energy bids. It also extends the bidding range of the aggregator to the supply and demand sides of the energy market, empowering the aggregator with two bidding options. The bidding strategies allow the prosumers to value their flexibility by transforming it into market products under the form of energy bids. These strategies may also prevent prosumers from investing in dedicated storage devices to store excess of renewable generation. Furthermore, the cluster-based optimization approach transforms the flexibility of thousands of resources into market products without relying on decentralized approaches. The centroid-based clustering algorithm reduces the size of the bidding optimization problems, both in terms of variables and constraints, making them tractable and fast for a large number of flexible resources. It preserves the original representation of the flexible resources and their main properties, contrarily to other works that adopt generalized battery models [15–17] and polytopes [14,18,19].

1.4. Paper organization

The remaining paper is organized as follows. Section 2 describes the electricity market and aggregator frameworks. Section 3 presents the

deterministic optimization and two-stage stochastic optimization models. Section 4 presents the centroid-based clustering algorithm. Sections 5 and 6 describe the case study and results. Finally, Section 7 presents the conclusions.

2. Problem description

2.1. Framework of the energy market

The participation of the aggregator in the electricity market follows the rules of the Iberian market (MIBEL). The MIBEL is a sequential market, where energy is traded first and reserves are contracted afterward. It covers the Portuguese and Spanish control areas.

The DA energy market of MIBEL is a single-price and double-sided auction, where market agents submit supply and demand hourly bids that cover the 24 h of the next day. The price (€/MWh) and volume of energy (MWh) of each hour are set by the EUPHEMIA algorithm (European DA market solver [20]). The gate closure is at the 12th hour and the clearing prices are published at the 13th hour. Before the 14th hour, the physical bilateral contracts are added to the cleared offers of the DA energy market. Afterward, the transmission system operator (TSO) performs congestion management and then proceeds to the generation of viable energy schedules. In case of being detected network problems, the TSO may use technical or market-based methods. Technical-based methods include adjusting network topology or other network elements, such as transformer taps. Market-based methods include countertrade and market splitting.

Days after delivery, the energy transitions are settled. The aggregator pays the energy bought and charges the energy sold in the day-ahead energy market. The deviations between DA market commitments and RT realizations are settled at imbalance prices. Portugal and Spain use the two-price system to value energy imbalances. The positive and negative imbalance prices result from a complex function of energy and reserve prices, as described in Ref. [22]. Fig. 1 shows the timeline of the DA energy market.

2.2. Participation of the aggregator in the day-ahead energy market

The aggregator assumes the role of retailer and supplier since it sells and buys energy. The aggregator acts as a price taker by submitting non-priced demand and supply bids to the DA energy market. This means that supply bids are presented at floor price (0 €/MWh), while demand bids are submitted at market cap price (180 €/MWh).

Fig. 1 shows the sequence of the aggregator's tasks. Before the 12th hour, the aggregator defines and submits the demand and supply bids. A cluster-based optimization approach generates the energy bids. Sections 3 and 4 describe the bidding optimization models and the centroid-based clustering algorithm.

2.3. Flow of information in the day-ahead stage

Fig. 2 describes the flow of information between the aggregator, prosumers, TSO, and electricity market during the DA stage. Six DA actions are identified, as follows:

1. communication of the prosumers information. The prosumers information includes parameters and state-of-operation of the appliances, prosumer's preferences, end-user engagement information, generation, consumption, etc.;
2. definition of demand and supply bids for the DA energy market. The aggregator defines the bids through one of the bidding optimization models described in Section 3;
3. submission of the energy bids (MWh and €/MWh);
4. energy market clearing performed by the EUPHEMIA platform. The market operator participates in this process;
5. the TSO performs congestion management. The exchange of

information includes cleared bids (from the EUPHEMIA platform) and viable energy schedules (from the TSO);

6. communication of the viable energy schedules (MWh) and clearing prices (€/MWh);

It is important to stress that only the optimization of the energy bids is covered in this work. The RT operation is outside of the scope of this paper. Examples of algorithms to manage the operation of flexible resources in RT can be found in Refs. [2,5,6].

2.4. Interface between the aggregator and prosumer

The aggregator interfaces with the prosumer through a HEMS. In the DA stage, the HEMS collects and communicates prosumers information, as described in Fig. 3. In this framework, the HEMS has the following functionalities: metering the consumption and generation of appliances; monitoring and acquiring the state-of-operation of appliances; enabling the exchange of information between the prosumer and aggregator; engaging the prosumer into cost-efficiency behaviors and services; and controlling flexible resources according to set-points communicated by the aggregator during the RT phase. The prosumer can interface with the HEMS through a smartphone application.

This paper considers four sources of flexibility: TCLs, EVs, SLs, and PVs. The EVs are sources of demand and generation flexibilities (V2G), while the TCLs and SLs are only sources of demand flexibility and the PV units are only sources of generation flexibility. Being a source of demand flexibility means that the resource is capable of increasing, decreasing or shifting consumption. On the other hand, a source of generation flexibility is capable of increasing and decreasing generation. The aggregator transforms the load and generation flexibilities of the prosumers into energy bids with high economic value, as demonstrated in the results section.

2.5. Remuneration and rewarding mechanisms

The aggregator signs a contract with the prosumers to automate and exploit their load and generation flexibilities in the electricity market. In exchange, the aggregator may offer attractive financial rewards, such as cheaper retailing prices and discounts on the monthly bill. Not all the prosumers are motivated by financial rewards. Financial motivations depend mainly on the socioeconomic status of the prosumers. However, the prosumer can be motivated by social and environmental incentives [23]. The environmental incentives promote “green” values, such as positive feelings towards the use of renewable energy sources or the reduction of CO₂ emissions. The social incentives promote fun and personal satisfaction.

Therefore, the rewarding mechanism adopted by the aggregator

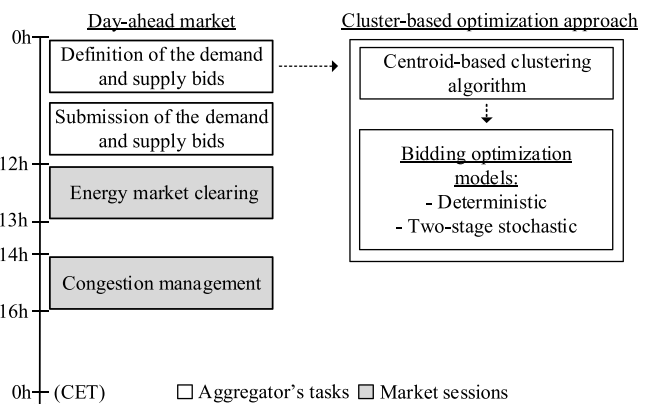


Fig. 1. The sequence of the aggregator's tasks and relevant market sessions of MIBEL.

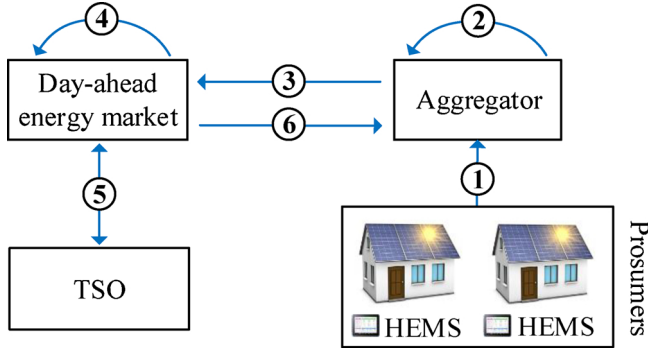


Fig. 2. Flow of information between the aggregator, prosumers, TSO, and electricity market during the DA stage.

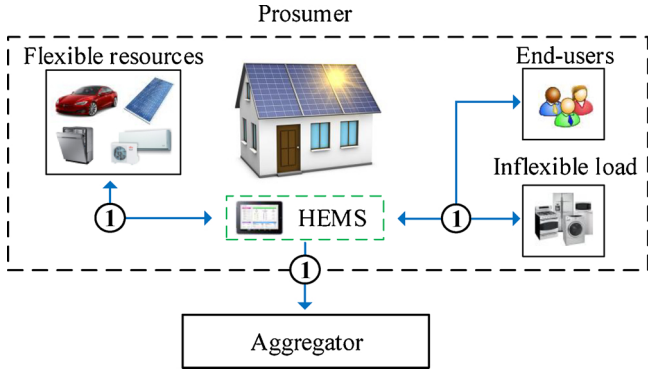


Fig. 3. Interface between the aggregator and prosumer.

should be “tailor-made”, i.e. adapted to the characteristics and preferences of the prosumers. The definition of the rewarding mechanism is outside the topic of this work. This paper is focused on the analysis of the possible economic advantages for the prosumers from actively participating in the wholesale market through an aggregator.

3. Day-ahead bidding optimization models

3.1. Two-stage stochastic model

The two-stage stochastic optimization model (1)–(16) defines demand and supply bids for the DA energy market. The main characteristics of the optimization model are described below:

1. **two-stage**: the optimization model defines energy bids for the DA energy market, taking into account multiple possible RT outcomes of generation and consumption. The DA market is the first-stage and the RT phase is the second-stage. The optimization models the net cost of trading energy in both DA and RT market stages;
2. **stochastic**: the uncertainties of the prosumers are incorporated in the optimization model through a set of scenarios. The scenarios model possible RT outcomes of PV generation, consumption, outdoor temperature and behaviors and preferences of the prosumers (e.g., house occupancy and plug-in availability of the EV).

Table 1

Variable parameters and sets of the thermostatically controlled load.

Symbol	Unit	Name	Description
θ^o	°C	Outdoor temperature	The aggregator signs a contract with a weather service provider to acquire forecasts/scenarios of outdoor temperatures.
O	–	House occupancy	The HEMS uses sensors to collect house occupancy information. Based on this information, the aggregator computes scenarios through a seasonal naïve forecasting algorithm.
ϑ	°C	Heat gains and losses	The heat gains and losses result from human activity, solar radiation and other loads. They can be estimated by the aggregator using least squares optimization [24].

3.1.1. Objective function

The two-stage stochastic optimization model is formulated as a minimization problem. The aim is to minimize the net cost of the aggregator buying and selling energy in both DA and RT market stages. The objective function (1) has two terms:

1. the first term is the net cost of selling and buying energy E_t^{DA} at forecasted price $\hat{\lambda}_t^E$ in the DA energy market. Negative values of E_t^{DA} are selling bids and positive values are buying bids;
2. the second term defines the expected imbalance cost, due to deviations between RT realizations and DA market commitments. A negative imbalance $\Delta E_{j,t}^-$ (kWh) is a surplus of demand or a shortage of generation. A positive imbalance $\Delta E_{j,t}^+$ is a shortage of demand or a surplus of generation. Energy imbalances are valued at negative $\hat{\lambda}_t^-$ or positive $\hat{\lambda}_t^+$ forecasted imbalance prices;

$$\text{Min} \sum_{t \in T} \left(\hat{\lambda}_t^E E_t^{DA} + \sum_{j \in J} \pi_j (\hat{\lambda}_t^- \Delta E_{j,t}^- - \hat{\lambda}_t^+ \Delta E_{j,t}^+) \right) \quad (1)$$

The uncertainties are modeled through a set of scenarios $j \in J$ with a probability of occurrence π_j . The optimization horizon is defined by a set of time intervals $t \in T$ of duration 1 h Δt (h).

3.1.2. Energy balance and market trading constraints

The market trading Constraint (2) sets the energy imbalances. Constraint (3) sets the expected net load in RT equal to the sum of the IL, PV generation, and flexible net load. The parameter ξ_i^v weighs the flexible loads $i \in I^v$. The weight ξ_i^v and set I^v are outputs of the centroid-based clustering algorithm. The electric power $P_{j,i,t}^v/P_{j,i,t}^v$ (kW) is positive for demand and negative for generation.

$$\Delta E_{j,t}^- - \Delta E_{j,t}^+ = E_{j,t}^{RT} - E_t^{DA}, \quad \forall j \in J, t \in T \quad (2)$$

$$E_{j,t}^{RT} = \sum_{v \in \{EV, TCL, SL\}} \left(\sum_{i \in I^v} \xi_i^v P_{j,i,t}^v \Delta t \right) + (P_{j,t}^{IL} + P_{j,t}^{PV}) \Delta t, \quad \forall j \in J, t \in T \quad (3)$$

3.1.3. Thermostatically controlled load constraints

The TCL consists of a split heat pump with heating and cooling capacities. The optimization requires four constraints. Constraints (4) and (5) set the ranges of the electric power in heating $P_{j,i,t}^{HE}$ (kW) and cooling $P_{j,i,t}^{CO}$ modes. The electric power is defined by $P_{j,i,t}^{TCL} = P_{j,i,t}^{HE} + P_{j,i,t}^{CO}$. Constraint (6) sets the temperature inside the room $\theta_{j,i,t+1}$ (°C). Constraint (7) ensures that the thermal comfort of the prosumer is satisfied in the hours of house occupancy $O_{j,i}$. The thermal comfort is defined by the temperature range $[\theta_i, \hat{\theta}_i]$.

$$0 \leq P_{j,i,t}^{HE} \leq \bar{P}_i^{HE}, \quad \forall j \in J, i \in I^{TCL}, t \in T \quad (4)$$

$$0 \leq P_{j,i,t}^{CO} \leq \bar{P}_i^{CO}, \quad \forall j \in J, i \in I^{TCL}, t \in T \quad (5)$$

$$\theta_{j,i,t+1} = \beta_i \theta_{j,i,t} + (1 - \beta_i) [\theta_{j,t}^o + R_i (COP_i P_{j,i,t}^{HE} - EER_i P_{j,i,t}^{CO})] + \vartheta_{j,i,t}, \quad \forall j \in J, i \in I^{TCL}, t \in T \quad (6)$$

$$\underline{\theta}_i \leq \theta_{j,i,t+1} \leq \bar{\theta}_i, \quad \forall j \in J, i \in I^{TCL}, t \in O_{j,i} \quad (7)$$

The TCL parameters can be variable or fixed. The variable parameters assume the form of scenarios and are computed on a daily basis. The variable and fixed parameters are described in Tables 1 and 2.

3.1.4. Electric vehicle constraints

The optimization of the EV requires six constraints. Constraints (8) and (9) set the ranges of the charging power $P_{j,i,t}^V$ (kW) and discharging power $P_{j,i,t}^A$. This formulation does not require the adoption of binary variables to avoid simultaneous charging and discharging due to the price structure of the objective function. Nonetheless, the binary variables can be added to Constraints (8) and (9), if necessary [2,25,26]. The electric power is defined by $P_{j,i,t}^{EV} = P_{j,i,t}^V - P_{j,i,t}^A$. Constraints (10) and (11) set the SOC (kWh) within its limits $[SOC_i, \bar{SOC}_i]$. Constraint (12) defines the target SOC at departure time $t_{j,i}^{DE}$. The aim is to ensure the preferences of the prosumers $t_{j,i}^{DE}$ and $SOC_{j,i}^{DE}$.

$$0 \leq P_{j,i,t}^A \leq \bar{P}_i^{EV}, \quad \forall j \in J, i \in I^{EV}, t \in T_{j,i}^{EV} \quad (8)$$

$$0 \leq P_{j,i,t}^V \leq \bar{P}_i^{EV}, \quad \forall j \in J, i \in I^{EV}, t \in T_{j,i}^{EV} \quad (9)$$

$$SOC_{j,i,t+1} = SOC_{j,i,t} + (\eta_i^V P_{j,i,t}^V - P_{j,i,t}^A / \eta_i^A) \Delta t, \quad \forall j \in J, i \in I^{EV}, t \in T_{j,i}^{EV} \quad (10)$$

$$SOC_i \leq SOC_{j,i,t+1} \leq \bar{SOC}_i, \quad \forall j \in J, i \in I^{EV}, t \in T_{j,i}^{EV} \quad (11)$$

$$SOC_{j,i,t_{j,i}^{DE}} \geq SOC_{j,i}^{DE}, \quad \forall j \in J, i \in I^{EV} \quad (12)$$

The EV parameters can be variable or fixed. The variable parameters assume the form of scenarios and are computed on a daily basis by the aggregator. The aggregator uses a seasonal naïve forecasting algorithm to generate the scenarios. The fixed and variable parameters are described in Tables 3 and 4.

3.1.5. Shiftable load constraints

The SL can be a dishwasher or a washing machine. The scheduling of the SL requires two constraints. Constraint (13) sets the electric power $P_{j,i,t}^{SL}$ (kW). The binary variable $\psi_{j,i,t}$ schedules the beginning of the program. Constraint (14) ensures that only one program is started at each availability period $T_{j,i}^{SL}$.

$$P_{j,i,t}^{SL} = \sum_{w=0}^{D_{j,i}-1} Pr_{j,i,w}^{SL} \psi_{j,i,t-w}, \quad \forall j \in J, i \in I^{SL}, t \in T_{j,i}^{SL} \quad (13)$$

$$\sum_{t \in T_{j,i}^{SL}} \psi_{j,i,t} = 1, \quad \forall j \in J, i \in I^{SL} \quad (14)$$

The SL parameters are variable and assume the form of scenarios. The aggregator computes the scenarios using a seasonal naïve forecasting algorithm. The variable parameters are described in Table 5.

3.1.6. Photovoltaic generation constraints

The optimization of the PV generation requires two constraints. Constraint (15) sets the PV generation $P_{j,i,t}^{PV}$ (kW). Constraint (16) defines

the range of the curtailment of PV generation $P_{j,i,t}^{CU}$. Note that curtailing generation can make sense in energy markets with negative prices. The uncertainty of the PV generation is modeled through a set of scenarios $\{Pr_{j,i,t}^{PV}, \forall j, i\}$.

$$P_{j,i,t}^{PV} = P_{j,i,t}^{CU} - \sum_{i \in I^{PV}} Pr_{j,i,t}^{PV}, \quad \forall j \in J, t \in T \quad (15)$$

$$0 \leq P_{j,i,t}^{CU} \leq \sum_{i \in I^{PV}} Pr_{j,i,t}^{PV}, \quad \forall j \in J, t \in T \quad (16)$$

3.1.7. Inflexible load

The uncertainty of the IL is modeled through a set of scenarios of power profiles $\{Pr_{j,i,t}^{IL}, \forall j, i\}$. The inflexible power is given by $P_{j,i,t}^{IL} = \sum_{i \in I^{IL}} (Pr_{j,i,t}^{IL})$.

3.2. Deterministic model

The deterministic optimization model (4)–(18) defines demand and supply bids. The main characteristics of the optimization model are described below:

1. **single-stage**: the optimization models the net cost of trading energy in the DA market stage;
2. **deterministic**: the prosumers information is modeled through point forecasts instead of probabilistic information under the form of scenarios.

3.2.1. Objective function

The deterministic model is formulated as a minimization problem. The aim is to minimize the net cost of buying and selling energy in the DA energy market. The objective function is given by:

$$\text{Min} \sum_{t \in T} \hat{\lambda}_t^E E_t^{DA} \quad (17)$$

3.2.2. Constraints

Constraint (18) sets the energy bids equal to the sum of the IL, PV generation, and flexible net load.

$$E_t^{DA} = \sum_{v \in \{EV, TCL, SL\}} \left(\sum_{i \in I^v} \xi_i^v P_{i,t}^v \Delta t \right) + (P_t^{IL} + P_t^{PV}) \Delta t, \quad \forall t \in T \quad (18)$$

The constraints of the flexible resources are given by Eq. (4)–(16).

4. Centroid-based clustering algorithm

The k-means [27] is the centroid-based clustering algorithm used to compute the aggregated flexibility of the TCLs, EVs, and SLs. The k-means finds the centroid μ_k and assigns the objects x to the nearest cluster, such that the squared distances from the cluster are minimized. The objective function of k-means assumes the following form:

Table 2

Fixed parameters of the thermostatically controlled load.

Symbol	Unit	Name	Description
C	kWh/°C	Thermal capacitance	C and R can be computed by the aggregator using estimation techniques based on least squares optimization [23].
R	°C/kW	Thermal resistance	
β	–	Thermal constant	
$\underline{\theta}/\bar{\theta}$	°C	Min./max. temperatures of comfort	$\beta = e^{-\frac{\Delta t}{CR}}$ $\underline{\theta}/\bar{\theta}$ are set by the prosumer.
COP/EER	–	Coefficients of cooling/heating performance	COP/EER are provided by the manufacturing companies.
$\bar{P}^{CO}/\bar{P}^{HE}$	kW	Maximum electric power for cooling/heating	$\bar{P}^{CO}/\bar{P}^{HE}$ are provided by the manufacturing companies.

Table 3
Fixed parameters of the electric vehicle.

Symbol	Unit	Name	Description
η	–	Charging and discharging efficiency	All the fixed parameters are provided by the manufacturing companies.
\bar{p}^{EV}	kW	Charging and discharging power rate	
\underline{SOC}	kWh	Minimum state-of-charge	
\overline{SOC}	kWh	Maximum state-of-charge	

Table 4
Variable parameters and sets of the electric vehicle.

Symbol	Unit	Name	Description
t^{DE}	–	Departure time	t^{DE} is set by the prosumer.
t^{AR}	–	Arrival time	–
T^{EV}	–	Availability period	T^{EV} is defined by the period between the arrival and departure times $[t^{AR}, \dots, t^{DE}]$.
SOC^{AR}	kWh	SOC at arrival time	–
SOC^{DE}	kWh	SOC at departure time	SOC^{DE} is set by the prosumer.

$$\underset{N}{\operatorname{argmin}} \sum_{k=1}^K \sum_{x \in N_k} x - \mu_k^2 \quad (19)$$

In this case, the k-means selects the TCLs, EVs, and SLs capable of emulating the flexibility of the entire portfolio. The k-means algorithm is applied to each type of flexible load. The inputs of the k-means algorithm are the number of clusters K and parameters. The parameters are divided into technical and behavioral, as shown in Fig. 4. The technical parameters correspond to the physical properties of the appliances. The behavioral parameters represent the end-users' preferences. The outputs of the k-means algorithm are a set of representative flexible loads $i \in I^v$ and the respective weight ξ_i^v .

Fig. 5 shows the application of the k-means algorithm to a portfolio of 7 EVs. Two EVs are selected to emulate the flexibility of the entire portfolio. The selected EVs correspond to the centroids of the cluster. The weights of the selected red and grey EVs are 4 and 3. The weight is given by the size of the cluster.

5. Case study

5.1. Main description

The aggregator of prosumers participates in the DA energy market of MIBEL during the first week of December 2015 (from November 30th to December 6th). The aggregator manages 10,000 prosumers from Porto (Portugal), each one has 1 PV system, 1 TCL, 1 SL, and 1 EV.

5.2. Prosumers information

Five types of prosumers information are considered: TCL parameters; EV parameters; SL parameters; inflexible load; and PV generation. The prosumers information can be fixed or variable. The variable

Table 5
Variable parameters and sets of the shiftable load.

Symbol	Unit	Name	Description
p^{SL}	kW	Load profile	The load profile depends on the program selected by the prosumer.
D	–	Duration of the program	–
\hat{T}^{SL}	–	Availability to complete the program	\hat{T}^{SL} is set by the prosumer.
\underline{T}^{SL}	–	Availability to start the program	$\underline{T}^{SL} = \hat{T}^{SL} \setminus \{\hat{T}_w^{SL}, \forall w = -D, \dots, -1\}$

	EV	TCL	SL
Physical parameters	$\frac{\bar{P}_i}{\underline{SOC}_i}$	$EE R_i \cdot \bar{P}_i^{CO}$ $CO P_i \cdot \bar{P}_i^{HE}$ $R_i C_i$	$\operatorname{avg}(\sum_w \bar{P}_{j,i,w}^{SL})$
Behavioral parameters	$\operatorname{avg}(SOC_{j,i}^{AR})$ $\operatorname{avg}(t_{j,i}^{DE} - t_{j,i}^{AR})$	$\sum_j O_{j,i}$ $(\theta_i - \underline{\theta}_i)$	$\operatorname{avg}(\hat{T}_{j,i,-1}^{SL} - \underline{T}_{j,i,0}^{SL})$ $\operatorname{avg}(D_{j,i})$

Fig. 4. Inputs of the k-means per type of flexible load.

information can assume the form of point forecasts or scenarios depending on the bidding optimization adopted. The two-stage stochastic model uses scenarios, while the deterministic model uses point forecasts.

5.2.1. Parameters of the thermostatically controlled loads

5.2.1.1. Fixed parameters. The fixe parameters include the physical characteristics of the rooms, temperature ranges, technical parameters of the TCLs. Fig. 6 shows the distribution of the rooms' characteristics of 10,000 prosumers (each TCL heats one room).

The technical parameters of the split inverter heat pumps adopted in this paper are described in Table 6. Four temperature ranges were considered: [19,24]; [20,22]; [20,23]; [19,22] °C.

5.2.1.2. Variable parameters. The variable parameters include scenarios of outdoor temperatures, house occupancy, and heat gains and losses. The scenarios of outdoor temperature were computed by a Gaussian copula method [28] using numerical weather predictions collected from the MeteoGalicia website [29]. Fig. 7 shows twenty scenarios of outdoor temperature.

The scenarios and point forecasts of house occupancy were computed by a seasonal naïve forecasting algorithm. The historical data includes annual time series of house occupancy of Portuguese households. Fig. 8 shows twenty scenarios of house occupancy.

A Gaussian white noise process was used to generate scenarios and point forecasts of heat gains and losses $\vartheta \sim N(0, 1 \times 10^{-6})^\circ \text{Cs}^{-1/2}$.

5.2.2. Parameters of the electric vehicles

5.2.2.1. Fixed parameters. The maximum charging and discharging electric power was selected by choosing one of two values: 3.7 or 7 kW [30]. The battery capacity was defined according to a truncated normal distribution described in Table 7.

5.2.2.2. Variable parameters. The scenarios and point forecasts of plug-

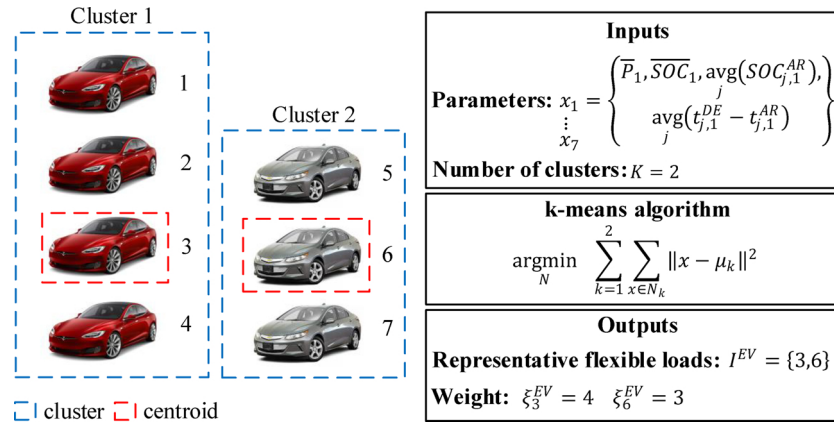


Fig. 5. Application of the k-means algorithm to a portfolio of 7 electric vehicles.

in availability, state-of-charge at departure and arrival times were generated by a seasonal naïve forecasting algorithm [3]. The historical data includes annual synthetic time series of EV mobility patterns [31]. Figs. 9 and 10 show twenty scenarios of plug-in availability, state-of-charge at arrival and departure times. The state-of-charge at departure time is equal to the battery capacity of the EV.

5.2.3. Parameters of the shiftable loads

5.2.3.1. Variable parameters. The scenarios and point forecasts of load profile, duration, availability to start and complete the program were generated by a seasonal naïve forecasting algorithm. The historical data includes annual synthetic time series of SL usage patterns [2].

Examples of scenarios are described in Fig. 11. The duration and load profiles are adjusted to the time unit of the energy market (i.e., 1 h). The availability to start a program is computed based on the availability to complete the program and duration of the program.

5.2.4. Inflexible load and photovoltaic generation

The inflexible load and PV generation can assume the form of scenarios and point forecasts, as described in Figs. 12 and 13. The point forecasts were computed by the gradient boosting algorithm [32] from the python package “scikit-learn” [33]. The scenarios were generated by the Gaussian copula method. The historical data of inflexible load consists of hourly load profiles of low-voltage consumers from Portugal for the period of two years (2014–2015). The input data used to forecast PV generation includes hourly measurements of PV generation and numerical weather predictions collected from the MeteoGalicía website for the period of one year (2015).

5.3. Electricity prices

The energy, positive and negative imbalance prices were forecasted by the gradient boosting algorithm. The input data was collected from the ENTSO-E transparency platform for the period of one year (2015) [2,34]. The data includes electricity prices, timestamp, forecasted demand, forecasted wind generation and forecasted solar generation of Portugal and Spain.

Table 8 presents the performance metrics for the forecasted

electricity prices. The mean absolute error (MAE) measures the accuracy of the forecasted prices and the Spearman’s rank correlation measures the prices ranking quality [35]. The forecasting performance of the energy price is good. The imbalance prices present a low performance. There is room for improving the forecasting of the imbalance prices. One of the possible paths is to use better forecasts of load, wind and solar generation, as explanatory variables.

6. Results

The results are divided into four main sections. Section 6.1 compares the three bidding strategies without considering pre-clustering. The three bidding strategies are the two-stage stochastic optimization, deterministic optimization and inflexible. The inflexible strategy considers that the load and generation of the prosumers are inflexible. Section 6.2 discusses the results of the centroid-based clustering algorithm. Section 6.3 discusses the impact of the centroid-based clustering algorithm in the bidding optimization strategies. Finally, Section 6.4 presents and discusses the computation performance of the bidding approaches with and without pre-clustering for portfolio sizes.

6.1. Comparison between the day-ahead bidding strategies without considering pre-clustering

6.1.1. Visual comparison

Fig. 14 shows the demand and supply bids of deterministic optimization, two-stage stochastic optimization and inflexible strategy for December 1st (Tuesday). Positive values are demand bids and negative values are supply bids.

The inflexible strategy places the demand bids in the periods of high energy prices. The optimized strategies (i.e., deterministic and two-stage stochastic optimizations) place most of the demand bids in the periods of low energy prices and the supply bids in the periods of high energy prices. The supply bids during the night period result from the V2G functionality.

The two-stage stochastic optimization introduces one additional term in the objective function compared to the deterministic optimization, the sensitivity to imbalance prices. The additional term affects

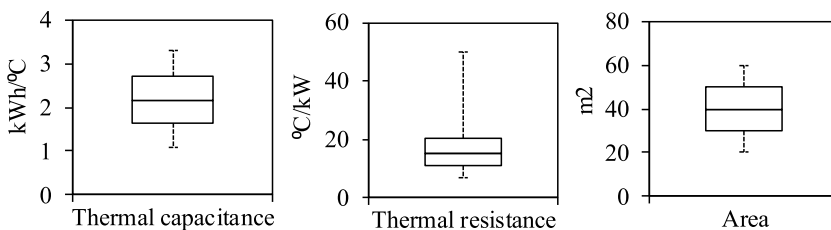


Fig. 6. Physical characteristics of the rooms.

Table 6
Parameters of split inverter heat pumps.

Model	1	2	3	4	5	6	7	8
COP	4.4	4.6	4.7	4.9	4.3	4.4	4.4	3.9
EER	4.3	4.4	4.5	5.0	3.8	3.9	4.1	3.4
\bar{P}^{HE} (kW)	0.9	0.9	0.9	1.1	1.3	1.3	1.5	1.5
\bar{P}^{CO} (kW)	0.7	0.7	0.7	0.7	1.0	1.0	1.0	1.5
A (m ²)	[20,30[[20,30[[30,40[[30,40[[40,50[[40,50[[50,60[[50,60[

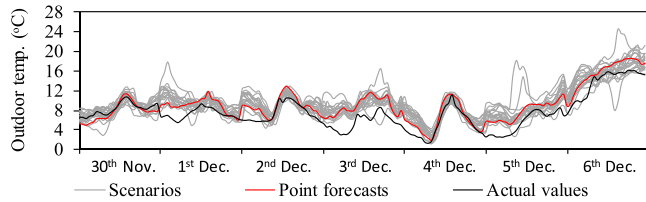


Fig. 7. Outdoor temperatures for the first week of December (20 scenarios, point forecasts, and actual values).

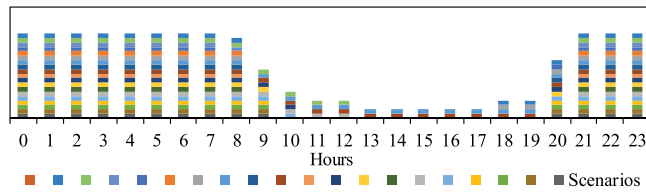


Fig. 8. Twenty scenarios of house occupancy for one day. A bar means that the household is occupied.

Table 7
Technical parameters of the electric vehicles.

η	\bar{P} (kW)	\underline{SOC} (kWh)	\overline{SOC} (kWh)
0.93	3.7 or 7	$0.1 \cdot \overline{SOC}$	$\overline{SOC} \sim N(33.3, 146.5)15 \leq \overline{SOC} \leq 70.4$

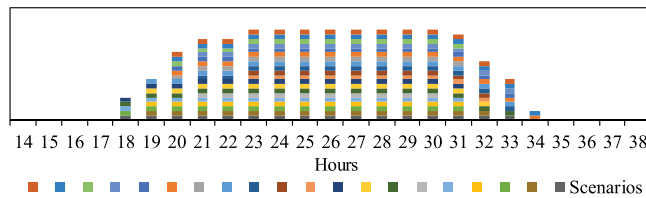


Fig. 9. Twenty scenarios of availability. A bar means that the EV is available and plugged-in.

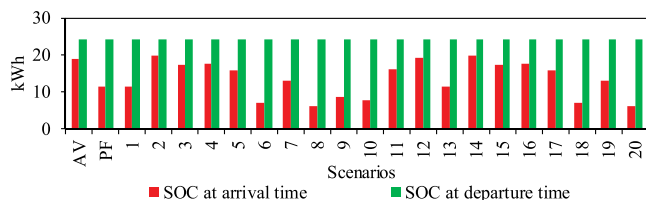


Fig. 10. Twenty scenarios of state-of-charge at arrival and departure times (20 scenarios, point forecasts [PF] and actual values [AV]).

the definition of the energy bids since it weighs the expected cost of the load and generation uncertainties. In addition, the two-stage stochastic optimization models the uncertainties of the prosumers through a set of scenarios, while the deterministic optimization considers point

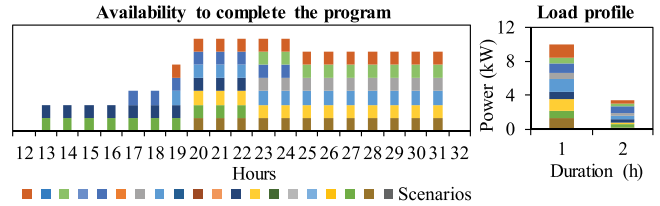


Fig. 11. Twenty scenarios of load profiles, duration, and availability to complete the program (a bar means that the SL is available).

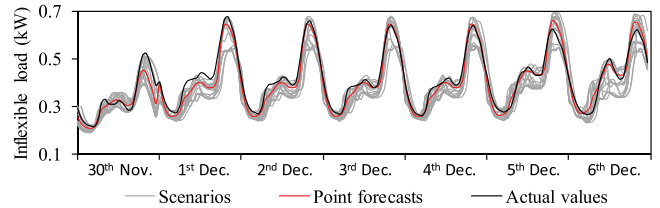


Fig. 12. Inflexible load of one prosumer in the first week of December 2015 (20 scenarios, point forecasts, and actual values).

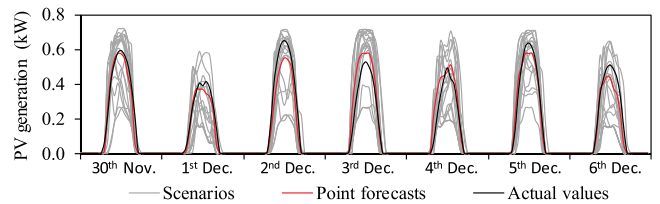


Fig. 13. Photovoltaic generation of a PV unit of 1 kWp for the first week of December (25 scenarios, point forecasts, and actual values).

Table 8
Performance metrics of the forecasted energy and imbalance prices.

	Energy	Positive imbalance	Negative imbalance
MAE (€/MWh)	3.0	8.7	8.6
Spearman's rank corr. (0–1)	0.91	0.55	0.44

forecasted information. These differences between the stochastic and deterministic models lead to the placement of different quantities of energy bids. For instance, the deterministic optimization in the 3rd hour places 57.6 MWh, while the two-stage stochastic optimization places 53.8 MWh

6.1.2. Cumulative bidding results

Table 9 compares the cumulative bidding results of the three bidding strategies for the first week of December 2015. The two-stage stochastic optimization presents the lowest bidding net cost (57.0 k€) followed by deterministic optimization (58.0 k€) and inflexible strategy (73.2 k€). The two-stage stochastic optimization places lower quantities of demand and supply bids than the deterministic optimization due to its conservative behavior. The inflexible strategy places low quantities

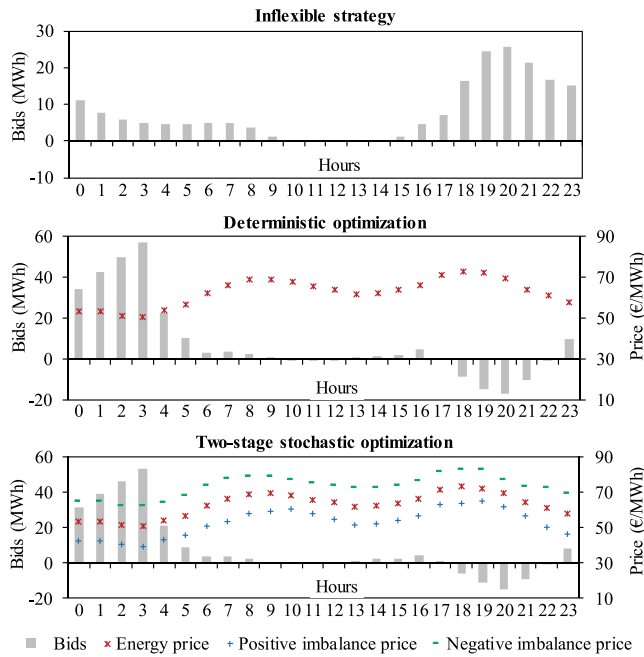


Fig. 14. Energy bids for December 1st. Positive values are demand bids and negative values are supply bids.

Table 9
Cumulative bidding results for the first week of December 2015.

	Two-stage stochastic	Deterministic	Inflexible
Demand bids (MWh)	1635	1754	1244
Supply bids (MWh)	400	460	36
Expected imbalances (MWh)	−4	0	0
Expected net load (MWh)	1231	1293	1208
Energy cost (k€)	82.1	87.6	75.4
Energy revenue (k€)	25.8	29.6	2.3
Expected imbalance cost (k€)	0.7	0.0	0.0
Total net cost (k€)	57.0	58.0	73.2

of supply bids since it does not exploit the flexibility of the EVs to perform energy arbitrage. The charging and discharging cycles of the EVs generate energy losses that increase the expected net load of the optimized approaches compared to the inflexible strategy. The difference in expected net load between the optimized approaches is due to the consideration of different types of input data. As previously referred, the two-stage stochastic optimization considers stochastic scenarios, while the deterministic approach considers point forecasted information.

The optimized bidding models outperform the inflexible strategy by more than 20% of cost savings. The two-stage stochastic optimization computes more robust solutions than the deterministic option since it models the uncertainty of the prosumers information through a set of scenarios. The performance of both approaches is similar. However, the complete performance of the optimization strategies can only be effectively evaluated in RT due to the uncertainties of the prices, generation, and consumption. Therefore, the stochastic approach should be adopted by conservative aggregators and in scenarios of high uncertainty. The deterministic option may present a better performance in scenarios of low uncertainty.

6.1.3. Distribution of the bidding net cost of the aggregator by the prosumers

Fig. 15 presents the distribution of the bidding net cost of the aggregator by the prosumers for the first week of December 2015. The two-stage stochastic optimization presents the lowest median 5.6 €

followed by deterministic optimization 5.7 € and inflexible strategy 7.2 €. This proves that the optimized strategies reduce the disaggregated and aggregated net costs.

As mentioned before, the remuneration and engagement mechanisms are outside the scope of this paper. However, Fig. 14 suggests that the participation of the prosumers in this flexibility service may contribute to significantly reduce their electricity costs. For example, if the inflexible strategy is compared to the two-stage stochastic optimization, the average savings per prosumer is 1.6 €/week (21% reduction).

6.2. Centroid-based clustering algorithm

The centroid-based clustering algorithm includes two phases. The first phase consists in selecting the number of clusters. The second phase consists in defining the representative flexible loads and their respective weights. The centroid-based clustering algorithm is run every day before the definition of the DA energy bids.

6.2.1. Number of clusters

One of the inputs of the k-means algorithm is the number of clusters per type of flexible load. The number of clusters is selected by a two-step heuristic. The first step consists in running the k-means algorithm for a series of number of clusters. The second step consists in selecting the number of clusters based on the silhouette score [36]. The silhouette score measures how similar an object (i.e., flexible load) is to its own cluster compared to other clusters. A high score suggests cluster cohesion, while a low score indicates cluster separation.

The silhouette scores were computed for a series of clusters from 50 to 500. Fig. 16 shows the silhouette scores for December 1st. The best maximum numbers of clusters for the TCLs, EVs, and SLs are 400, 450 and 350 since they present the highest silhouette scores.

Table 10 presents the best numbers of clusters and respective silhouette scores for all the days of the week. The most frequent numbers of clusters are 450 for TCLs and EVs, and 500 for SLs. These numbers of clusters present the highest silhouette scores, i.e. the highest cohesion.

6.2.2. Representative flexible loads and weights

The representative flexible loads $i \in L^v$ and their respective weights ξ_i^v are computed by the k-means algorithm based on the number of clusters previously selected. The representative flexible load corresponds to the center of the cluster. The weight of the representative flexible load is defined by the size of the cluster. Fig. 17 presents the number of clusters per size for December 1st. The initial 30,000 flexible loads were reduced to 1,200 representative flexible loads. Most of the clusters present a size between 1 and 40 flexible loads. The largest cluster has a size of 140 SLs.

6.3. Impact of pre-clustering in the day-ahead bidding strategies

6.3.1. Visual comparison

The energy bids computed by both optimization models with and without pre-clustering present similar bidding behaviors, as shown in Fig. 18. The results show that the 1176 representative flexible loads emulate effectively the aggregated flexibility of the 30,000 loads. The MAEs are 0.21 and 0.22 MWh for deterministic optimization and two-

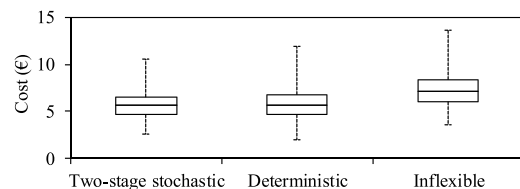


Fig. 15. Distribution of the cumulative costs of the aggregator by the prosumers.

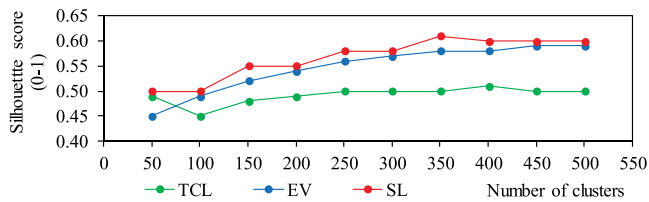


Fig. 16. Silhouette scores for December 1st (Tuesday).

Table 10

Best numbers of clusters and silhouette scores for the first week of December 2015.

	30th Nov.	1st Dec.	2nd Dec.	3rd Dec.	4th Dec.	5th Dec.	6th Dec.
TCL	350	400	450	400	450	450	500
EV	500	450	450	450	400	450	450
SL	500	350	450	400	400	500	500

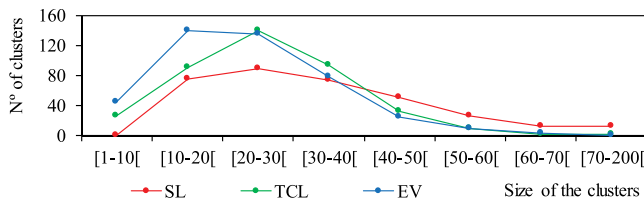


Fig. 17. Number of clusters per size for December 1st.

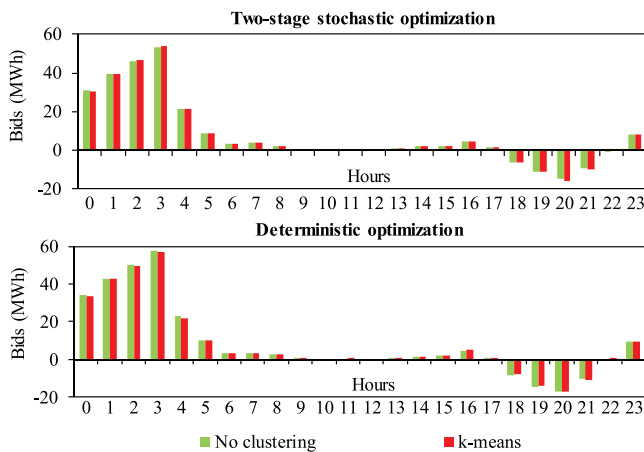


Fig. 18. Comparison of the energy bids with and without pre-clustering for December 1st. Positive values are demand bids and negative values are supply bids.

stage stochastic optimization, respectively. The deviations are small compared to the average bid size of 13 MW/h.

6.3.2. Cumulative bidding results

Table 11 displays the metrics used to compare the bids computed with and without pre-clustering for one week. Both optimizations present an almost perfect Spearman's rank correlation (i.e., +1) and a low MAE. The two-stage stochastic optimization presents better metrics due to its conservative nature. The weekly costs considering k-means are 57.5 k€ and 58.6 k€ for the two-stage stochastic optimization and deterministic optimization. The cost deviations are very small, around 1% when the results with and without pre-clustering are compared. These results demonstrate that pre-aggregation based on clustering is applicable to deterministic and stochastic approaches without significant loss of detail.

6.4. Computational performance

The flexible loads were clustered by the k-means algorithm from the Python package “scikit-learn” [33]. The optimization problems were implemented in Python and solved by the mixed-integer linear programming library from the IBM CPLEX 12.7 solver on a machine with 256 GB RAM and an Intel(R) Xeon(R) CPU E7-4820 CPU clocked at 2.0 GHz.

Table 12 shows that clustering flexible loads through k-means reduces the size of the optimization problems, both in terms of variables and constraints, and consequently the execution times without compromising the quality of the energy bids. Although the analyzed case study, with 40,000 flexible resources, resulted on tractable optimization problems, a significant increase of flexible resources, e.g. to 1,000,000, increases the complexity of the problems and makes them computationally intractable, as shown in Table 13. An effective way of making the optimization problems computationally tractable is through pre-clustering, as the results in Table 13 demonstrate.

The flexible loads of the portfolios of 10,000 and 250,000 prosumers are clustered based on the same number of clusters (see Table 10). However, the k-means selects different representative flexible loads since the portfolios have different characteristics. This generates different optimization problems, both in terms of variables and constraints, and consequently different execution times, as shown in Tables 12 and 13. Nonetheless, the size of the optimization problems and execution times are very similar. For instance, the difference between the execution times of the deterministic optimizations with k-means for the two portfolios is around 0.1 s.

7. Conclusion

The foreseen transformation of passive consumers into active prosumers creates a business opportunity for aggregators trading the prosumers' flexibility in the electricity markets. Two new challenges arise from this new paradigm: (1) how to transform the load and generation flexibilities of the prosumers into market products; (2) how to manage and optimize thousands or millions of flexible resources in a timely and effective way. This paper proposes a two-step approach to support an aggregator in the definition of demand and supply bids for the DA energy market. In the first step, a clustering algorithm based on k-means was presented to compute the aggregated flexibility of the aggregator's portfolio. The aim was to reduce the size of the bidding optimization problems making them tractable and fast to solve. In the second step, the energy bids were defined by an optimization model that can assume the form of a deterministic or a two-stage stochastic problem.

The numerical results show that the optimized bidding strategies based on deterministic and stochastic models outperform an inflexible strategy by more than 20% of cost savings for the presented case study (different case studies are likely to produce different results). The stochastic approach slightly outperforms the deterministic model by 2%. The stochastic approach should be adopted by conservative aggregators in scenarios of high uncertainty. The deterministic option may present a better performance in scenarios of low uncertainty. The disaggregation results suggest that prosumers can reduce the electricity costs by 1.6 €/week with the two-stage stochastic optimization. This demonstrates that transforming prosumers' flexibility into demand and supply bids contributes to reducing the bidding cost of the aggregator and the

Table 11

Performance metrics of the energy bids for one week.

	Deterministic	Two-stage stochastic
MAE (MW/h)	0.35	0.28
Spearman's rank correlation	0.99	0.99

Table 12

Average size of the optimization problems and execution times for a portfolio of 40,000 flexible resources (10,000 prosumers).

	Deterministic		Two-stage stochastic	
	No clustering	k-means	No clustering	k-means
Continuous variables	1,175,836	48,388	22,980,482	950,075
Binary variables	25,666	923	517,231	19,098
Constraints	1,074,720	44,126	15,948,574	653,980
Execution time	41 s	0.5 s	7.8 h	2 min

Table 13

Average size of the optimization problems and execution times for a portfolio of 1,000,000 flexible resources (250,000 prosumers).

	Deterministic		Two-stage stochastic	
	No clustering	k-means	No clustering	k-means
Continuous variables	29,341,094	48,530	572,746,075	950,102
Binary variables	641,111	952	12,957,750	19,035
Constraints	27,129,286	44,181	404,320,058	655,187
Execution time	intractable	0.6 s	intractable	1.9 min

electricity costs of its clients.

The results achieved proved that the consideration of pre-clustering reduces the size of the optimization problems, both in terms of variables and constraints, and consequently the execution times without compromising the quality of the energy bids. In addition, it keeps the optimization problems tractable for millions of flexible resources. This demonstrates that reducing the information of the flexible resources via centroid-based clustering enables the optimization of millions of resources.

Future work will focus on studying different business models based on incentive mechanisms and innovative engagement techniques, such as gamification [37,38]. The mid-term (months or seasons) and long-term (one or more years) economic performances of the bidding strategies will also be addressed in future work.

Acknowledgments

The work of José Iria was supported by Fundação para a Ciência e Tecnologia (FCT) Ph.D. Scholarship PD/BD/113716/2015. This work is also financed by the GREsBAS project (SmartGP/0003/2015) under the framework of ERA-Net Smart Grids Plus.

References

- [1] J.P. Iria, F.J. Soares, M.A. Matos, Trading small prosumers flexibility in the energy and tertiary reserve markets, *IEEE Trans. Smart Grid.* (2018) 1, <https://doi.org/10.1109/TSG.2018.2797001>.
- [2] J. Iria, F. Soares, M. Matos, Optimal supply and demand bidding strategy for an aggregator of small prosumers, *Appl. Energy* 213 (2018) 658–669, <https://doi.org/10.1016/j.apenergy.2017.09.002>.
- [3] A. Alahäivälä, J. Corbushley, J. Ekström, J. Jokisalo, M. Lehtonen, A control framework for the utilization of heating load flexibility in a day-ahead market, *Electr. Power Syst. Res.* 145 (2017) 44–54, <https://doi.org/10.1016/j.epsr.2016.12.019>.
- [4] S. Chen, Q. Chen, Y. Xu, Strategic bidding and compensation mechanism for a load aggregator with direct thermostat control capabilities, *IEEE Trans. Smart Grid* 3053 (2016) 1–10, <https://doi.org/10.1109/TSG.2016.2611611>.
- [5] R.J. Bessa, M.A. Matos, Global against divided optimization for the participation of an EV aggregator in the day-ahead electricity market. Part II: numerical analysis, *Electr. Power Syst. Res.* 95 (2013) 309–318, <https://doi.org/10.1016/j.epsr.2012.08.013>.
- [6] R.J. Bessa, M.A. Matos, Global against divided optimization for the participation of an EV aggregator in the day-ahead electricity market. Part I: theory, *Electr. Power Syst. Res.* 95 (2013) 309–318, <https://doi.org/10.1016/j.epsr.2012.08.007>.
- [7] S.I. Vagropoulos, A.G. Bakirtzis, Optimal bidding strategy for electric vehicle aggregators in electricity markets, *IEEE Trans. Power Syst.* 28 (2013) 4031–4041, <https://doi.org/10.1109/TPWRS.2013.2274673>.
- [8] I. Momber, A. Siddiqui, T.G.S. Roman, L. Soder, Risk averse scheduling by a PEV aggregator under uncertainty, *IEEE Trans. Power Syst.* 30 (2015) 882–891, <https://doi.org/10.1109/TPWRS.2014.2330375>.
- [9] H. Mohsenian-Rad, Optimal demand bidding for time-shiftable loads, *IEEE Trans. Power Syst.* 30 (2015) 939–951, <https://doi.org/10.1109/TPWRS.2014.2338735>.
- [10] J.P. Iria, F.J. Soares, M.A. Matos, Trading small prosumers flexibility in the day-ahead energy market, 2017 IEEE Power & Energy Society General Meeting, IEEE, 2017, pp. 1–5, <https://doi.org/10.1109/PESGM.2017.8274488>.
- [11] Y. Luo, S. Itaya, S. Nakamura, P. Davis, Autonomous cooperative energy trading between prosumers for microgrid systems, *Proceedings - Conference on Local Computer Networks, LCN.* 2014–November, 2014, pp. 693–696, <https://doi.org/10.1109/LCNW.2014.6927722>.
- [12] D. Ilic, P.G. Da Silva, S. Karnouskos, M. Griesemer, An energy market for trading electricity in smart grid neighbourhoods, *IEEE International Conference on Digital Ecosystems and Technologies* (2012) 1–6, <https://doi.org/10.1109/DEST.2012.6227918>.
- [13] E. Mengelkamp, J. Gärtner, K. Rock, S. Kessler, L. Orsini, C. Weinhardt, Designing microgrid energy markets. A case study: the Brooklyn microgrid, *Appl. Energy* 210 (2018) 870–880, <https://doi.org/10.1016/j.apenergy.2017.06.054>.
- [14] S. Barot, J.A. Taylor, A concise, approximate representation of a collection of loads described by polytopes, *Int. J. Electr. Power Energy Syst.* 84 (2017) 55–63, <https://doi.org/10.1016/j.ijepes.2016.05.001>.
- [15] H. Hao, B.M. Sanandaji, K. Poolla, T.L. Vincent, A generalized battery model of a collection of Thermostatically Controlled Loads for providing ancillary service, 2013 51st Annual Allerton Conference on Communication, Control, and Computing, Allerton, B.M. (2013) 551–558, <https://doi.org/10.1109/Allerton.2013.6736573>.
- [16] H. Hao, B.M. Sanandaji, K. Poolla, T.L. Vincent, Aggregate flexibility of thermostatically controlled loads, *IEEE Trans. Power Syst.* 30 (2015) 189–198, <https://doi.org/10.1109/TPWRS.2014.2328865>.
- [17] M. Alizadeh, A. Scaglione, A. Goldsmith, G. Kesidis, Capturing aggregate flexibility in Demand Response, *Proceedings of the IEEE Conference on Decision and Control.* 2015–February (2014) 6439–6445, <https://doi.org/10.1109/CDC.2014.7040399>.
- [18] L. Zhao, H. Hao, W. Zhang, Extracting flexibility of heterogeneous deferrable loads via polytopic projection approximation, 2016 IEEE 55th Conference on Decision and Control, CDC 2016 (2016) 6651–6656, <https://doi.org/10.1109/CDC.2016.7799293>.
- [19] F.L. Muller, J. Szabo, O. Sundstrom, J. Lygeros, Aggregation and disaggregation of energetic flexibility from distributed energy resources, *IEEE Trans. Smart Grid.* 3053 (2017) 1–10, <https://doi.org/10.1109/TSG.2017.2761439>.
- [20] EPEX, GME, N., Pool, OMIE, OPCOM, OTE, TGE, EUPHEMIA Public Description, 2015. <https://www.nordpoolspot.com/globalassets/download-center/pcr/euphemia-public-documentation.pdf>. (Accessed 5 July 2017).
- [22] Redes Energéticas Nacionais (REN), Settlement Manual for the Portuguese System, (2008) <http://www.mercado.ren.pt/EN/Electr/MarketInfo/Document/BibSubregula/MPACDezembro2008.pdf>.
- [23] E.R. Frederiks, K. Stenner, E.V. Hobman, Household energy use: applying behavioural economics to understand consumer decision-making and behaviour, *Renew. Sustain. Energy Rev.* 41 (2015) 1385–1394, <https://doi.org/10.1016/j.rser.2014.09.026>.
- [24] M. Maasoumy, A. Sangiovanni-Vincentelli, Total and peak energy consumption minimization of building HVAC systems using model predictive control, *IEEE Des. Test Comput.* 29 (2012) 26–35, <https://doi.org/10.1109/MDT.2012.2200871>.
- [25] K. Paridari, A. Parisio, H. Sandberg, K.H. Johansson, Robust scheduling of smart appliances in active apartments with user behavior uncertainty, *IEEE Trans. Autom. Sci. Eng.* 13 (2016) 247–259, <https://doi.org/10.1109/TASE.2015.2497300>.
- [26] A. Parisio, E. Rikos, L. Glielmo, A model predictive control approach to microgrid operation optimization, *IEEE Trans. Control Syst. Technol.* 22 (2014) 1813–1827, <https://doi.org/10.1109/TCSST.2013.2295737>.
- [27] J.A. Hartigan, M.A. Wong, Algorithm AS 136: a K-means clustering algorithm, *J. R. Stat. Soc. C.* 28 (1979) 100–108, <https://doi.org/10.2307/2346830>.
- [28] R. Pinto, R.J. Bessa, M.A. Matos, Multi-period flexibility forecast for low voltage prosumers, *Energy* 141 (2017) 2251–2263, <https://doi.org/10.1016/j.energy.2017.11.142>.
- [29] MeteoGalicia, (n.d.). <http://www.meteogalicia.gal/web/index.action>. (Accessed 26 July 2017).
- [30] J.P. Iria, F.J. Soares, I.G. Franchin, N. Silva, Development of a novel management system for Electric Vehicle charging, 2014 IEEE International Electric Vehicle Conference (IEVC), IEEE, 2014, pp. 1–7, <https://doi.org/10.1109/IEVC.2014.7056096>.
- [31] F.J. Soares, J.A.P. Lopes, P.M.R. Almeida, A stochastic model to simulate electric vehicles motion and quantify the energy required from the grid, 2011 17th Power Systems Computation Conference (PSCC) (2011) 22–26 http://www.pssc-central.org/uploads/tx_ethpublications/fp359.pdf.
- [32] T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning*, 2nd ed., Springer, 2009, <https://doi.org/10.1198/jasa.2004.s339>.
- [33] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay, Scikit-learn: machine learning in python, *J. Mach. Learn. Res.* 12 (2012) 2825–2830, <https://doi.org/10.1007/s13398-014-0173-7.2>.
- [34] L. Hirth, J. Mühlenpfordt, M. Bulkeley, The ENTSO-E transparency platform — a review of Europe's most ambitious electricity data platform, *Appl. Energy* 225 (2018) 1054–1067, <https://doi.org/10.1016/j.apenergy.2018.04.048>.
- [35] C. Cui, T. Wu, M. Hu, J.D. Weir, X. Li, Short-term building energy model

- recommendation system: a meta-learning approach, *Appl. Energy* 172 (2016) 251–263, <https://doi.org/10.1016/j.apenergy.2016.03.112>.
- [36] P.J. Rousseeuw, Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, *J. Comput. Appl. Math.* 20 (1987) 53–65 <http://www.sciencedirect.com/science/article/pii/0377042787901257>.
- [37] A. Barbosa, J. Iria, F. Cassola, A. Coelho, J. Portela, Ü. Küçük, A.G. Madureira, M.A. Zehir, A. Ozdemir, F.J. Soares, GReSBAS project : A gamified approach to promote more energy efficient behaviours in buildings, *Electrical and Electronics Engineering (ELECO)*, 2017 10th International Conference On, Bursa, Turkey, 2017, pp. 1258–1261 <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8266391>.
- [38] F. Cassola, J. Iria, H. Paredes, L. Morgado, A. Coelho, F. Soares, Using choreographies to support the gamification process on the development of an application to reduce electricity costs, *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Springer Verlag, 2017, pp. 256–257.