Global against divided optimization for the participation of an EV aggregator in the day-ahead electricity market. Part I: Theory

R.J. Bessa *, M.A. Matos

INESC TEC – INESC Technology and Science (formerly INESC Porto), FEUP – Faculty of Engineering, University of Porto, Portugal

A R T I C L E   I N F O

Article history:
Received 31 January 2012
Received in revised form 21 May 2012
Accepted 15 August 2012
Available online 7 September 2012

Keywords:
Electric vehicles
Aggregator
Electricity market
Forecasting
Optimization
Operational management

A B S T R A C T

This paper addresses the bidding problem faced by an electric vehicles (EV) aggregation agent when participating in the day-ahead electrical energy market. Two alternative optimization approaches, global and divided, with the same goal (i.e. solve the same problem) are described. The difference is on how information about EV is modeled. The global approach uses aggregated values of the EV variables and the optimization model determines the bids exclusively based on total values. The divided approach uses individual information from each EV. In both approaches, statistical forecasting methods are formulated for the EV variables. After the day-ahead bidding, a second phase (named operational management) is required for mitigating the deviation between accepted bids and consumed electrical energy for EV charging. A sequential linear optimization problem is formulated for minimizing the deviation costs. This chain of algorithms provides to the EV aggregation agent a pathway to move to the smart-grid paradigm where load dispatch is a possibility.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

In a forthcoming scenario with a significant penetration of electric vehicles (EV) in the power system, aggregation agents (or aggregators as an abbreviated form) will emerge as intermediary between EV drivers, electricity market, distribution system operator (DSO) and transmission system operator (TSO)[1,2].

The EV aggregation agent is a concept already adopted in business models for electrical mobility. For instance, in Portugal the industrial network MOBLE is implementing a charging network accessible to all users [3]. The business model includes aggregation agents that users may liberally choose. In this model, the aggregator is a simple electricity retailer for electrical mobility. A similar business model, from Better Place, is described in [4]. In both business models, the aggregator is a common load aggregator that buys electrical energy in the market for its clients and does not have any direct control over EV charging rates. A survey of technical and economic issues related with EV, and possible business models for the aggregated, can be found in [5].

A more elaborated aggregator model, comprising the possibility of controlling directly the charging process of each EV, is being explored in several research projects, such the EU project MERGE [6] and the Danish Edison project [7]. This type of EV aggregator enables the smart-charging approach. Moreover, the EV aggregator may also present offers for ancillary services (namely reserve), which will increase its retailing profit and decrease the charging costs for the EV drivers [8].

Many algorithms were developed for supporting the EV aggregator participation in the day-ahead electrical energy and reserve market. Kristoffersen et al. [9] developed a linear programming model for defining the optimal charging plan for EV fleets with vehicle-to-grid (V2G) by minimizing costs (electricity and battery wear) for a fleet operator in the day-ahead electrical energy market. Sundstrom and Binding [10] presented an optimization problem for minimizing the cost of charging EV constrained by the distribution network branch limits. Cao et al. [11] described a heuristic algorithm for controlling the EV charging in response to a time-of-use price in a regulated electricity market. Rotering and Ilic [12] described two optimization algorithms for an optimal controller installed in an EV: (a) optimization of charging rates and periods for minimizing the cost; (b) profit maximization from selling regulation power. Han et al. [13] proposed a dynamic programming model for exploring the possibility of offering regulation power from EV. Sortome and El-Sharkawi [14] presented three heuristic strategies and corresponding optimal analogues for exploring the EV operating as flexible load (i.e. without the need of V2G mode) for providing regulation service.

The major shortcoming of all these studies is the assumption that there is complete knowledge of all the EV variables (e.g. EV driving profiles) involved in the problem. In fact, it is necessary to
forecast these variables. With this objective, Bessa et al. [15] used a naive forecasting approach for an optimization algorithm that determines optimal bids for the day-ahead electrical energy and reserve markets. This methodology differed from other approaches because the optimization algorithm takes as input total (or aggregated) values for the EV variables, instead of optimizing the EV charging individually. Furthermore, there was an emphasis on the importance of having an operational management phase where the aggregator manages (or dispatch) the EV charging, in order to avoid penalizations due to deviations from the bid values. Otherwise, it may be unmanageable to comply with the market commitments, such as providing reserve. Wu et al. [16] also addressed this requirement by proposing a heuristic algorithm that distributes the purchased electricity by EV, with as low deviation from the schedule as possible.

The present paper covers the gap between the optimization and forecasting phases, and studies the impact of EV information modeling in the optimization algorithms. The optimization model presented in [15] (named global approach) is revised and enhanced. An alternative optimization model, named divided approach, which uses individual information from each EV, is formulated. Both approaches have the same goal (i.e. solve the same problem), the difference being on how information about the EV is modeled in the optimization phase. The participation in the reserve market is not addressed in this paper, because the primary objective is to evaluate the impact of modeling EV information with two different approaches.

Statistical forecasting methods are proposed for the EV variables, which are addressed as time series. The comparison with other forecasting algorithms for time series forecasting is out of the paper’s scope, since the aim is just to describe how these variables can be forecasted and assess the impact of forecast errors in the optimization phase. Furthermore, a new sequential optimization algorithm for minimizing the deviation between the day-ahead market bids and the actual charging is described. Contrasting to the algorithm presented by Wu et al. [16], this operational management algorithm is an optimization problem and does not use as input the price ranking from the day-ahead electrical energy market.

This paper is organized as follows: Section 2 describes the aggregator architecture and the model chain for participating in the electricity market; Section 3 presents the global optimization problem and forecasting algorithms; Section 4 presents the divided optimization problem and forecasting algorithms; Section 5 describes the operational management algorithm; Section 6 presents the conclusions.

A companion paper [17] presents numerical analysis for comparing the two alternative optimization approaches.

2. EV aggregation agent: architecture and model chain

2.1. EV aggregation agent architecture

This paper adopts a hierarchical direct control [1,14] where a single entity (an aggregator) directly controls the charging of a group of EV. The aggregator may receive signals from the DSO, and take appropriate actions to avoid violation of network operational constraints. In case of abnormal operating conditions, the DSO can request load curtailment to the aggregator.

This architecture, adopted in EU project MERGE [6], allows the provision of ancillary services and supports the DSO in managing the distribution network with a high penetration of EV without the need to reinforce the network (cf. [1]). Fig. 1 depicts the adopted architecture.

Two different groups of clients are foreseen for the EV aggregator:

- **inflexible EV load**: a client who does not allow the aggregator to control the charging process, the aggregator being just an electricity provider;
- **flexible EV load**: a client who allows the aggregator to control the charging process (bidirectional communication), which means that its charging requirement must be satisfied, but a degree of freedom exists regarding the supply periods.

This paper assumes that only the clients that charge in residential installations (i.e. slow charging mode) are flexible EV load, while clients charging in commercial, public and fast charging stations are inflexible EV load. This is a probable scenario since clients choosing normal and fast charging modes are not interested in having the vehicle plugged-in for long periods.

The aggregator represents the EV drivers in the electricity market (purchases electrical energy for EV charging) and retains a profit that depends on its bidding strategy and charging control strategy. However, it does not have any control over the individual EV driving behavior, so the driver needs must still be respected and are the main priority. The benefits for the aggregator are the possibility of increasing its retailing profit by minimizing the cost of purchased electrical energy. In exchange, the aggregator offers cheap retailing prices or a discount in the monthly electricity bill, in particular for flexible loads.

For the inflexible EV clients, the aggregator only buys electrical energy for charging these clients. The interaction is unidirectional and just for billing purposes.

The interaction with the DSO is important for solving congestions in the distribution network. In [6,18] different solutions for mitigating congestion are discussed. The typical solution, that fits the current electricity market rules without need for changes, is to have the DSO making an ex-ante validation (and bid correction if necessary) of the aggregators’ bids. For example, the DSO performs power flow calculations with the aggregator bids and determines the consumption reduction when there is congestion in the network branches. The DSO, during the operational charging management, may also send signals to the aggregator requesting consumption’s reduction due to voltage limits or branch load limits violation. The management procedures for solving congestion situations are not addressed in this paper, but can be found in [19].
Thus, the bids obtained with the global and divided approaches are not limited by possible congestion in the distribution network.

2.2. Model chain and information representation

For the short-term time horizon (up to 48 h ahead), the aggregator participates with buying bids in the electrical energy market. The aggregator for each hour of the next day, based on forecasted variables, defines the bids for the day-ahead electrical energy market. The optimized decisions related with these markets are performed on a daily basis and the bids are not discriminated by EV.

The EV information can be modeled in the optimization process by two alternative approaches (depicted in Fig. 2):

- **global approach**: the variables related with each EV are aggregated (summed) and the optimization model determines the “optimal” bids entirely based on summed values of EV availability and consumption. EV individual information is not included in the bidding phase;
- **divided approach**: the variables related with EV behavior are forecasted for each EV and the optimization model based on this information computes the optimal charging for each EV. The bid is equal to the sum of the optimized individual charging.

The bids resulting from the day-ahead optimization phase are used as an input of an operational management algorithm. The operational management is discriminated by EV and takes advantage of the EV fleet flexibility. This flexibility allows different combinations of charging profiles for achieving a matching between consumed electrical energy and accepted bids. The inputs are the following: accepted bids from the day-ahead market session; expected end of charge time interval and charging requirement of the EV, communicated by the drivers that are plugged-in. Based on this information, the algorithm optimizes the individual EV charging to comply with the market commitments and satisfy the EV drivers charging requirement.

An important part of this model chain consists in forecasting algorithms for the electric energy consumption from EV and availability periods for charging. The load forecasting task is common in problems related with power systems and electricity markets. However, this problem is different because the aggregator controls EV consumption, which means that the classical approach of forecasting the consumption in each time interval cannot be strictly followed. The approach proposed in this paper is to forecast the charging requirement. The charging requirement of each EV is the electrical energy (including the losses in the charger) needed to get from the initial state of charge (SOC) (i.e. when the EV plugs-in for charging) to the target SOC defined by the EV driver for the next trip. The aggregator then distributes this quantity, according to its optimization strategy, by the time intervals of the corresponding plugged-in period. Note that this approach does not require personal information such as the driver route (historical and planned) or the number of traveled kilometers.

3. Global approach for bidding in the day-ahead market

3.1. Representation of the EV information

The EV information in the global approach is represented by aggregated values of three variables. The first variable is the total maximum available power for charging, and the aggregated value is the sum of the available power of each EV in a specific time interval where the EV is plugged-in. For example, if in a specific time interval, 10 EV are plugged-in with a maximum charging power of 3 kW, the total maximum available power is 30 kW. The second variable is the total charging requirement. The charging requirement of each EV is associated to a specific availability to charge period, and it is placed in the last time interval before departure. For example, consider an EV (with battery size 30 kWh) that plugs-in for charging at interval H1 with a SOC of 50% and the expected departure time interval is H9 with a target SOC of 100%. The charging requirement of this EV is 16.5 kWh (considering a charging efficiency of 90%), and this value is placed in time interval H8 of the charging requirement time series. This variable means that 16.5 kWh must be supplied to the EV by the end of time interval H8. The total charging requirement is the sum of the individual values of the EV fleet.

The third variable is the total charging requirement distribution, and the only difference is that the charging requirement is placed in all the time intervals of the availability period. For example, for the aforementioned EV, the 16.5 kWh are placed in all the time intervals between H1 and H8. This variable means that in each time interval
of the period between H1 and H8 there is one EV that requires 16.5 kWh for reaching its target SOC.

The measured values of these three variables can be collected with Information and Communications Technology (ICT) already available for industrial use [20]. The forecasted values can be obtained with the statistical algorithms described in Section 3.4, using historical time series data.

3.2. Advantages and limitations

The main advantage of the global approach is that the aggregated values present less variability and a more pronounced periodic behavior. Fig. 3 depicts a seasonal plot [21] for a one-year time series of the number of plugged-in EV from one fleet with 1500 EV (left-hand side plot) and from a single EV (right-hand side plot). The time series are synthetic and generated by the method described in [22]. The plot shows the complete time series (one year of data) grouped by the individual seasons (daily pattern) in which the data were observed. Each line in the plot, with 48 half-hours, is one day from the whole time series; thus, each plot has 365 lines. The time series of one EV is binary and shows a high variability from day-to-day. The aggregated time series does not show a high daily variability and depicts two clear seasonal patterns: one for week days where the number of plugged-in EV in residential areas after 10 AM is low, and another for weekend days where the number of plugged-in EV is higher.

The dimension of the optimization problem is also low, i.e., low number of decision variables and constraints. The main disadvantage is that this approach does not fully capture the impact of the charging process in the total maximum available power for charging in each hour. For illustrating this statement, an example with three EV plugged-in for 6 h, and with maximum charging power of 3 kW, is given in Table 1.

The global approach uses as input the total maximum available power for charging. In this example, the total value is 9 kW in each hour because the three EV are plugged-in in all hours and with a maximum charging power of 3 kW. As the charging progresses in time, this total maximum power must be corrected by discounting the EV with full battery (or almost) but that remain plugged-in. However, the global approach does not use individual information from each EV. The charging dispatch is not performed individually for each EV; instead, is made for the aggregated values of the EV fleet. If individual information was used (matter discussed in Section 4), the aggregator could distribute the charging of each EV by the 6 h period, for satisfying the drivers’ charging requirement, and respecting the maximum charging power in each hour. Using this individual information, it is possible to see that in hour H2 the maximum charging power is not 9 kW, but 8 kW because EV2 can only charge 2 kW. In hour H4 the maximum charge is decreased to 4 kW because the EV2 is already full, but the EV remains plugged-in. Finally, note that an adjustment can only be performed with perfect accuracy if perfect information is known for each EV.

3.3. Formulation of the optimization problem

This section presents an enhanced version of the optimization model described in a previous paper from the authors [15]. The model in [15] has the flexible charging control confined to predefined and rigid periods. Furthermore, even with these predefined periods, the limitation previously described in Table 1 may lead to a high deviation from the quantities offered in the market. Therefore, the model was revised to overcome these two limitations. The enhanced version does not require the definition of a flexible period (i.e., optimizes EV charging along the whole day) and mitigates the impact of the charging process on the maximum charging power.

In this paper, and for participating in the day-ahead market, the aggregator is assumed to be a price-taker that only presents bid for energy quantities.

The mathematical formulation is the following:

\[
\min \sum_{t \in H} (\hat{p}_t - E_t) \\
\text{subject to:}
\]

\[
\sum_{j=1}^{t} (E_j) \geq \sum_{j=1}^{t} (\hat{R}_j), \quad \forall t \in H
\]

\[
\sum_{j=1}^{t} (E_j) - \sum_{j=1}^{t} (\hat{R}_j) \leq \hat{R}_j^0, \quad \forall t \in H
\]

\[
\frac{E_t}{\Delta t} \leq \hat{p}_t^\text{max} (1 - \alpha_t), \quad \forall t \in H
\]

where \( H \) is a set of time intervals from the programming period, \( \hat{p}_t \) is the price for time interval \( t \), \( E_t \) is the electrical energy bid, \( \hat{R}_j \) is the forecasted total charging requirement, \( \hat{R}_j^0 \) is the forecasted
Table 1
Illustrative example of three EV with charging process controlled by the aggregator.

<table>
<thead>
<tr>
<th>Information used in the global approach</th>
<th>H1</th>
<th>H2</th>
<th>H3</th>
<th>H4</th>
<th>H5</th>
<th>H6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total (sum of) max charging power [kW]</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Individual information from each EV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EV 1 (needs 18 kWh for SOC = 100%)</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>EV 2 (needs 5 kWh for SOC = 100%)</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>EV 3 (needs 7 kWh for SOC = 100%)</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total adjusted max charging power [kW]</td>
<td>9</td>
<td>8</td>
<td>8</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 2
Illustrative example of the charging requirement distribution of three EV.

<table>
<thead>
<tr>
<th>Information used in the global approach</th>
<th>H1</th>
<th>H2</th>
<th>H3</th>
<th>H4</th>
<th>H5</th>
<th>H6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual information from each EV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EV 1 (needs 9 kWh for SOC = 100%)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>EV 2 (needs 8 kWh for SOC = 100%)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>EV 3 (needs 3 kWh for SOC = 100%)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Total charging requirement [kWh]: (\bar{R}_t)</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Total charging requirement dist. [kWh]: (R_t^D)</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Bid [kWh]: (E_t)</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>LHS of Eq. (3): (\sum_{t=1}^{n} (E_t - \bar{R}_t))</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>9</td>
<td>15</td>
<td>20</td>
</tr>
</tbody>
</table>

The total charging requirement distribution, \(\bar{p}_{\text{max}}\) is the forecasted total available power for charging. \(\alpha\) is a factor that relates the available power for charging with the percentage of satisfied charging requirement, \(\Delta t\) is the time step (i.e., length of the time interval).

If the time step \(\Delta t\) is less than the market time step (typically 1 h), the bid’s quantity is the sum of each \(E_t\) contained in the market time step. In this case, the market price is also the same in each time interval contained in the market time step.

The objective function (Eq. (1)) consists in minimizing the cost of buying electrical energy \(E_t\) in the market for charging the EV fleet. The model has three constraints. The first constraint (Eq. (2)) assures that the total charging requirement is satisfied with the energy purchased in the market at each hour \(t\). This constraint guarantees that when the \(j\)th EV departs, the electrical energy required for satisfying the target SOC was purchased in the electrical energy market.

The second constraint in Eq. (3) guarantees that the charging requirement is only allocated when there is sufficient EV plugged-in for consuming the corresponding quantity. This constraint is explained with a small illustrative example for three EV, as presented in Table 2.

In this example, EV1 and EV2 are connected between H4 and H6 and have charging requirement equal to 9 kW and 8 kW correspondingly; EV3 is plugged-in during 6 h and has charging requirement equal to 3 kWh. The maximum available power for charging is 3 kW. The charging requirement is placed in the last time interval before departure, which is H6 for all EV. The sum of the individual values gives the total charging requirement (\(R_t\)). In the charging requirement distribution variable, the values are placed in the time intervals where the EV is plugged-in. For example, 9 kW is placed in hours H4–H6 of EV1. The sum for each hour is the total charging requirement distribution (\(R_t^D\)). A possible bid \(E_t\) is also illustrated in the table. Without the constraints of Eq. (3) it would be possible to present a bid value greater than 0 at intervals H2 and H3 if these were hours with a price lower than at H4–H6. As shown in Table 2, this constraint limits to zero the bid at H2 and H3 (otherwise the LHS would be greater than the \(R_t^D\)), since there is no EV plugged-in for consuming more than 3 kWh.

The last constraint in Eq. (4) guarantees an electrical energy bid below or equal to the forecasted total maximum available power (\(p_{\text{max}}\)) for charging at time step \(t\). The factor \(\alpha\) was introduced for adjusting the maximum charging power as the charging process evolves (as explained in Section 3.1), and is a linear function of the percentage of satisfied charging requirement, which is the ratio \((E_t - \bar{R}_t)/R_t^D\), at the beginning of period \(t\):

\[
\alpha_t = \beta \cdot \left( \frac{\sum_{t=1}^{n-1} (E_t - \bar{R}_t)}{R_t^D} \right), \quad \forall t \in H
\]

The coefficient \(\beta\) that minimizes the deviation between bids and actual consumption is estimated from historical data (e.g., EV availability, charging requirement). The estimation process has two steps: (i) run the global optimization model for different values of \(\beta\); (ii) select the \(\beta\) that leads to the lowest deviation. Historical data are available, unless the aggregator is starting its business; in this case the aggregator should start with a high \(\beta\) value (above 0.5) to be conservative. In the companion paper [17] a sensitivity analysis of different \(\beta\) values between 0 and 1 is presented for two different EV fleets.

It is important to stress that Eq. (2) mitigates the problem, but does not solve it totally. The problem can only be solved with information from individual EV, like in the divided approach that will be described in Section 4.

Finally, for inflexible EV loads an optimization model is not necessary. The aggregator only needs to forecast the total consumption in each hour and purchase in the market the forecasted quantity.
3.4. Forecasting methodology

It is assumed that the EV driver (acting as a flexible EV load) defines, when plugs-in the EV for charging, the desired battery SOC for the next trip and expected end of charge hour. Since this information may be provided only after the bidding in the day-ahead market, the aggregator needs to forecast the following variables for the next day: total charging requirement \( R_t \), total charging requirement distribution \( R_t^0 \) and the total maximum available power for charging \( P_{\text{max}} \). For the EV acting as an inflexible load, a forecast model is also necessary.

For the three EV variables, the proposed forecast algorithm consisted in a linear model with lagged variables and covariates. The model can be written as:

\[
y_t = \phi_0 + \phi_1 \cdot y_{t-1} + \phi_2 \cdot y_{t-2} + \cdots + \phi_l \cdot y_{t-l} + H_t + D_t + \varepsilon_t
\]

where \( y \) is the response variable \( (P_{\text{max}}, \text{or } R_t, \text{or } R_t^0) \), \( \phi \) are the model’s coefficients, \( y_{t-j} \) is the \( j \)th lag of the response variable \( y \), \( I \) is the maximum order of lagged variables, \( H_t \) is a seasonal index that takes a different value for each hour of the day, \( D_t \) is also a seasonal index that takes a different value for each day of the week, and \( \varepsilon \) is the noise term.

This model can be fitted on historical time series data using the generalized least squares since the residuals are autocorrelated [24]. If the number of EV drivers under contract with the aggregator changes, the model parameters need to be re-estimated. Recursive least squares with forgetting can be used for estimating the parameters in time varying conditions [25].

For selecting the lagged variables of Eq. (6), the first step is to check whether or not the time series is stationary, using a unit-roots test [26]. Then, the visual analysis of the autocorrelation diagram together with the Akaike information criterion (AIC) [27] is used for selecting the lagged variables.

For the objective function in Eq. (1), it is necessary to have forecasts of the day-ahead electrical energy price. For this purpose, a forecast algorithm based on an additive model with nonlinear relations modeled by smoothing splines is proposed [28]. The model is a linear predictor involving a sum of smooth functions of covariates, and it is written as:

\[
\hat{p}_t = \phi_0 + \phi_1 \cdot p_{t-1} + \phi_2 \cdot p_{t-2} + \cdots + \phi_l \cdot p_{t-l} + g(w_{p_t}) + H_t + D_t + \varepsilon_t
\]

where \( g \) is a smooth function estimated using cubic basis splines, and \( w_{p_t} \) is the forecasted wind power penetration level.

For feature selection, the same procedure of the EV variables can be followed. Based on recent price forecasting literature, such as [29], it was decided to include as covariate the forecasted penetration of wind power (forecasted total load divided by the forecasted total generation). The smooth function \( g \) captures a nonlinear relationship between the electrical energy price and forecasted wind power penetration.

Finally, multi-step ahead forecasts are necessary for the optimization problem. For the models in Eqs. (6) and (7) two alternative multi-step forecasting strategies are: iterated and direct approach [30]. In the iterated approach the one-step-ahead forecast is computed and used as the real value for producing the second step-ahead forecast, and the process is repeated iteratively and always using the same model. This approach has two disadvantages: (a) since the measured value is replaced by the forecasted value, the error is propagated through the time steps; (b) the model is fitted only for one-step ahead forecasts. In the direct approach, \( k \) models are fitted for each look-ahead time step. The inputs are always the same in each model, while the response variable is the \( k \) step-ahead value. This model solves the two disadvantages of the iterated approach. Nevertheless, the \( k \) models are learned independently which induces a conditional independence, and inhibits the modeling of dependencies between the variables.

Initial tests in the EV and day-ahead price time series showed that the iterated approach leads to a better performance.

4. Divided approach for bidding in the day-ahead market

4.1. Representation of the EV information

In the divided approach, EV information is disaggregated by EV and represented by two variables. The first variable is the availability period, which is a binary variable indicating the time intervals where the EV is plugged-in and available for charging. The second variable is the charging requirement of each EV. For forecasting these variables, a new forecasting process based on statistical algorithms is described in Section 4.4.

4.2. Advantages and limitations

The divided approach uses the individual information from each EV, which prevents the problem with the maximum available power for charging, in contrast to the global approach. However, the main disadvantage is that the forecasts for each EV present a high variability (cf. variability of a single EV in the right-hand side of Fig. 3), which is reduced by the aggregation effect. Nevertheless, it remains necessary to study the influence of the individual forecast errors in the final solution. Another disadvantage is the very high dimension of the optimization problem, which may difficult the inclusion of information about uncertainties of the input variables.

4.3. Formulation of the optimization Problem

The basic idea of the divided approach is to determine the market bids considering the individual information of each EV. The mathematical formulation is as follows:

\[
\min \sum_{t \in H} \left( \hat{p}_t \sum_{j=1}^{M_t} (E_{t,j}) \right)
\]

subject to:

\[
\frac{E_{t,j}}{\Delta t} \leq P_{\text{max}}, \quad \forall t \in H, \quad \forall j \in \{1, \ldots, M_t\}
\]

\[
\sum_{t \in \Delta t} (E_{t,j}) = \hat{R}_j, \quad \forall j \in \{1, \ldots, M_t\}
\]

where \( H \) is a set of time intervals from the programming period, \( \hat{p}_t \) is the price forecast for time interval \( t \), \( E_{t,j} \) is the electrical energy for charging the \( j \)th EV in time interval \( t \), \( \hat{R}_j \) is the forecasted charging requirement, \( P_{\text{max}} \) is the maximum available power for charging, \( \hat{R}_j \) is the forecasted plugged-in period of the \( j \)th EV, \( M_t \) is the total number of EV plugged-in at time interval \( t \), \( \Delta t \) is time step.

The objective function minimizes the total cost of purchased electrical energy. The constraint of Eq. (9) limits the electrical energy purchased for each EV by the maximum available power for charging the EV. The constraint of Eq. (10) ensures that the electrical energy purchased for each plug-in period \( P_{\text{plug}} \) of each EV matches the charging requirement defined by the EV driver for that period.

In the divided approach, in addition to the day-ahead electrical energy price, there are two variables that need to be forecasted: the EV availability \( P_{\text{plug}} \) (i.e. period where the EV is plugged-in for charging) and the charging requirement \( R \). The following section formulates a forecasting methodology for these variables.
4.4. Forecasting methodology

The forecasting algorithm for the EV availability and charging requirement is divided in two phases. First, a binary variable for the EV availability is forecasted. Then, non-parametric bootstrapping is used to forecast the charging requirement for the plugged-in period. This approach is inspired on the work of Willemain et al. [31] for estimating the entire distribution of the sum of the demands for service parts inventories over a fixed lead-time.

For the binary forecasting phase, the generalized linear models (GLM) theory is used [32]. Compared to the classical linear models, GLM models are for non-Gaussian response variables, such as count and binary data. The basic idea is to express linear models for a transformation of the mean value (link function), and keep the observations untransformed, which preserves the distributional properties of the observations. The link function is any monotone mapping of the mean value space to the real line used to form the linear predictor. The log it function \( \ln(a/(1-a)) \) was adopted.

In this problem, the response variable \( y \) is 1 if the EV is plugged-in or 0 otherwise. A natural distributional assumption is the Bernoulli distribution, \( y \sim Bernoulli(p) \). Note that the quantity modeled by the GLM is the posterior probability \( p(y = 1|x) \), where \( x \) is a set of covariates.

Let \( y_i \) be the response variable, the GLM model for the EV availability can be written as:

\[
p(y_i = 1|y_{i-1} \ldots y_{i-\lambda}) = \frac{1}{1 + \exp(-(\phi_0 + \phi_1 \cdot y_{i-1} + \phi_2 \cdot y_{i-2} + \cdots + \phi_i \cdot y_{i-\lambda}))} \tag{11}
\]

where \( 1/(1 - \exp(-a)) \) is the inverse of the link function, \( \phi \) are the model’s coefficients, \( y_{i-j} \) are lagged values of the response variable and \( \lambda \) is the maximum order of lagged terms.

The model is a binary regression model with lagged values of the response variable and the coefficients \( \phi \) can be estimated with the iteratively reweighted least squares, using the function glm from the R base distribution [33].

Multi-step ahead forecasts are necessary for the EV availability. In the iterated approach the forecasting model is fitted for one-step ahead forecasts, and because of this, the probability of having a 0 following a 1 (i.e. EV departing) would be very low for any look-ahead time step. For example, the model’s coefficients from fitting Eq. (11) to a synthetic time series of one EV could be the following:

\[
p(y_i = 1|y_{i-1} \ldots y_{i-\lambda}) = \frac{1}{1 + \exp(-(4.03 + 6.276 \cdot y_{i-1} + 0.6689 \cdot y_{i-48} + 1.05 \cdot y_{i-336}))} \tag{12}
\]

where the lags 48 and 336 are for modeling the daily and weekly seasonal pattern (with half-hour time steps). With the model of Eq. (12), the posterior probability in time interval \( t \) of the forecast horizon is equal to 0.98 when \( y_{t-1} \cdot y_{t-48} = y_{t-336} = 1 \), and equal to 0.9 when \( y_{t-1} = 1 \) and \( y_{t-48} = y_{t-336} = 0 \). Moreover, the subsequent look-ahead time steps \( t+1, t+2, \ldots \), will always have a posterior probability greater or equal to 0.9, even when \( y_{t-48} = y_{t-336} = 0 \).

Therefore, the direct approach seems to be more appropriate; however a modification is necessary to include the two seasonal patterns. The modified direct approach is:

\[
p(y_i = 1|y_{i-1} \ldots y_{i-\lambda}) = \frac{1}{1 + \exp(-(\phi_0 + \phi_1 \cdot y_{i-1} + \phi_2 \cdot y_{i-2} + \phi_3 \cdot y_{i-3} + \phi_4 \cdot y_{i-48} + \phi_5 \cdot y_{i-336}))} \tag{13}
\]

\[
p(y_{i+1} = 1|y_{i-1} \ldots y_{i-\lambda-1}) = \frac{1}{1 + \exp(-(\phi_0 + \phi_1 \cdot y_{i-1} + \phi_2 \cdot y_{i-2} + \phi_3 \cdot y_{i-3} + \phi_4 \cdot y_{i-47} + \phi_5 \cdot y_{i-335}))} \tag{14}
\]

where each model is fitted individually for each look-ahead time step. The difference for the direct approach is that the lagged variables related with the seasonal pattern are not fixed and change with the look-ahead time step.

After producing forecasts for the EV availability periods (i.e. sequence of hours where the EV is plugged-in for charging), the second step is to use the non-parametric bootstrapping technique [34] for estimating the charging requirement of each plugged-in period. The samples for the bootstrapping approach are from an artificial time series created from historical charging values. This artificial time series consists in rearranging the historical charging data by removing the charging dependency from market prices inside each availability period; each EV starts charging when it plugs-in and until the charging requirement is satisfied. For example, an EV that needs 12 kWh for reaching full SOC will charge at 3 kW during the first 4h. With this reorganization the charging behavior only depends of the number of hours that the EV is plugged-in for charging (and of course from the SOC at arriving), and not from the market price.

The bootstrap samples are conditioned to the number of hours the EV is plugged-in. For example, for the first hour the bootstrapping technique resamples from the artificial time series, but only from historical data of consumption during the first hour of the availability period. The same process is followed for the subsequent hours. Summing the bootstrap samples of each time interval of the availability period gives the forecast for the charging requirement associated to a specific EV.

The bootstrapping technique is repeated \( N \) times and the result is a distribution of the charging requirement. The algorithm can be summarized as follows:

\[
\text{for } h \text{ in (1 to } H) \text{ do}
\]

\[
\text{use GLM to forecast } y_{h|1}\text{ }
\]

\[
\text{if } y_{h|1} = 1 \text{ then}
\]

\[
\text{plug.time}_{c,n} = \text{plug.time}_{c,n-1} + 1
\]

\[
\hat{y}_{i,n} = \text{ bootstrap sample plug.time}_{i,n}
\]

\[
\text{ else }
\]

\[
\text{plug.time}_{c,n} = 0
\]

\[
\hat{\bar{y}}_{n} = \sum_{s \in \text{ sample period}} (\hat{y}_{i,s})
\]

where \( H \) is the time horizon, \( h \) is the look-head time step, \( y \) is the availability variable, \( c \) is a sample from the artificial charging time series, \( R \) is the charging requirement of the bootstrap sample \( n \), \( plug.time \) is the plugged-in hour, and \( aval.period \) is the availability period.

The \( N \) bootstrapping samples create the charging requirement distribution, from which the expected value and other statistics can be computed.

Fig. 4 depicts an illustrative example of forecasted and realized values for the availability of an EV. The realized values are taken from a synthetic time series generated with the algorithm described in [22]. The forecasted value is obtained with the forecasting approach described in Eqs. (13) and (14), using the same time series, and for a time horizon of 100 look-ahead time steps. Based on the forecasted availability periods the bootstrapping approach estimated a charging requirement of 21.42 kWh for the period between intervals 20 and 50 (the realized value was 13.9 kWh), and for the period between intervals 66 and 92 the estimated charging requirement was 18.34 kWh (the realized value was 11.15 kWh).
In its mathematical form, the objective function for a complete day (with T time intervals of length $\Delta t$) it is given by:

$$\min \sum_{t=1}^{T} \left( \varphi \left( E_t - \sum_{j=1}^{M_t} (E_{t,j}^*) \right) \right)$$

(17)

where $E_t$ is the result from the global or divided optimization, $E_{t,j}^*$ is the electrical energy consumed by the jth EV in time interval t, $M_t$ is the number of EV plugged-in during time interval t, $\varphi$ is the loss function given by

$$\varphi(u) = \begin{cases} 
    u \cdot \bar{\pi}_t^+, & u \geq 0 \\
    -u \cdot \bar{\pi}_t^-, & u < 0 
\end{cases}$$

(18)

The piecewise linear convex function of Eq. (17) can be represented by:

$$\min \sum_{t=1}^{T} (\max(-u \cdot \bar{\pi}_t^+ , u \cdot \bar{\pi}_t^-))$$

(19)

where $\bar{\pi}_t^+$ is the forecasted price for positive deviations and $\bar{\pi}_t^-$ is the forecasted price for negative deviations. If $\bar{\pi}_t^+$ and $\bar{\pi}_t^-$ are symmetric, the objective function is the minimization of the deviations absolute value.

A convex objective function can be transformed into a linear objective function by expressing the formulation in its epigraph form [36]. The full linear programming problem at time interval t becomes:

$$\min \sum_{k=t_0}^{T} (v_k)$$

(20)

$$\left( E_k - \sum_{j=1}^{M_t} (E_{k,j}^*) \right) \cdot \bar{\pi}_k^+ \leq v_k, \quad \forall k \in \{t_0, \ldots, T\}$$

(21)

$$- \left( E_k - \sum_{j=1}^{M_t} (E_{k,j}^*) \right) \cdot \bar{\pi}_k^- \leq v_k, \quad \forall k \in \{t_0, \ldots, T\}$$

(22)

$$\frac{E_{k,j}^*}{\Delta t} \leq P_{k,j}^{\text{max}}, \quad \forall j \in \{1, \ldots, M_t\}, \quad \forall k \in H_j^{\text{plug}}$$

(23)

$$\sum_{k \in H_j^{\text{plug}}} (E_{k,j}^*) = R_{t_0,j}, \quad \forall j \in \{1, \ldots, M_t\}$$

(24)

$$v_k \geq 0$$

(25)

where $t_0$ is the first time interval, $v$ is a slack variable, $H_j^{\text{plug}}$ is the plugged-in period of the jth EV, $R_{t_0,j}$ is the residual charging requirement of the jth EV at beginning of time instant $t_0$.

The constraints of Eqs. (21) and (22) result from the epigraph form and guarantee the deviations’ minimization. The constraint of Eq. (23) limits the charging by the maximum charging power of the EV. The constraint of Eq. (24) enforces the charging requirement communicated by the EV driver.

This optimization problem is solved for each time interval $k$ with the following sequential process:

1. in the beginning of time interval $t_0$, the new information (expected end of charge hour and target SOC) from the recently plugged EV is included in Eq. (24) of the optimization model;
2. the optimization problem is solved with this new information for a period between $t_0$ and the maximum departure hour of all the EV, $\max(H_j^{\text{plug}}) \forall j$; $\pi_{t_0}^+$ and $\pi_{t_0}^-$ are made equal to a large
number (e.g. 1000) in order to force the deviation to be zero at time interval \( t_0 \);
3. set points corresponding to the optimal charging levels for time interval \( t_0 \) are communicated to the plugged-in EV; only the dispatch for time interval \( t_0 \) remains unchanged, the charging levels for the subsequent time intervals can be modified in the next iteration (next time interval, \( t_0 + 1 \)). The charging requirement \( R_{t_0} \) is updated for the next period, \( R_{t_0+1} = R_{t_0} - E_{t_0} \).
4. this optimization process is repeated for the next time interval, \( t_0 + 1 \) (go to step 1).

5.2. Forecasting the deviation prices

For the constraints of Eqs. (21) and (22), it is necessary to forecast the deviation prices. Forecasting the prices is a complex task, since its variability is very high. In this paper, an approach based on additive models (similar to the model in Eq. (7)) is proposed:

\[
y_t = \phi_0 + \phi_1 \cdot y_{t-1} + \phi_2 \cdot y_{t-2} + \cdots + \phi_l \cdot y_{t-l} + g(w_p) + g(t_l^{\text{import}})
\]

(26)

where \( g \) is a smooth function estimated using cubic basis splines, \( p_t \) is the price from the day-ahead market, \( t_l^{\text{import}} \) and \( t_l^{\text{export}} \) are the interconnection exchanges (exported and imported electrical energy) of the bulk power system. The response variable \( y \) is the price of positive \((\pi_+^r)\) or negative deviation \((\pi_-^r)\).

It is important to stress that this model should be seen as a first approach to the problem, which will be a topic for future improvement.

6. Conclusions

In this paper, two alternative optimization approaches (with different representation of the information about EV) for supporting an EV aggregator participating in the day-ahead electrical energy market were presented. For each approach, statistical models were proposed for forecasting the required information about EV availability and consumption. Moreover, an operational management algorithm that extracts benefits from aggregating EV is described for minimizing the deviations between the bids and the consumed electrical energy.

In order to produce a complete approach, statistical forecasting algorithms were also proposed, as a first approach to the problem. It is plausible that these algorithms can be combined with “physical” algorithms, such as road traffic flow simulators. A statistical–physical hybrid method can certainly improve the results and should be a topic of future research. Furthermore, other future research topics consists in developing time-adaptive models capable of learning from new data without the need to re-train offline the statistical algorithm and capable of coping with changes in the EV fleet and drivers’ behavior.

The two proposed optimization approaches, global and divided, have both advantages and disadvantages. The major difference lies in the representation of the forecasted information for the EV fleet. In fact, the divided approach essentially consists in dispatching the EV individually based on the forecasted prices, and does not use the capability of combining the EV individual charging. Conversely, the global approach takes advantage of the aggregation capacity since it uses the aggregated variables related with the EV availability and consumption. Nevertheless, using the output of the divided approach, the operational management algorithm, explores the aggregation capacity (i.e. combines the individual EV charging) for minimizing the forecast error. In the companion paper [17], the optimization algorithms together with the operational management algorithm are evaluated and compared in a realistic case study.

These three layers of the model chain are crucial for the aggregator retaining activity. In fact, the models described in this paper can be used as the basis for building more complex approaches for bidding in the ancillary services market. Moreover, it is likely that these algorithms are adopted by aggregators of other types of flexible loads (e.g. battery chargers for consumer electronics, electric hot water heaters).

Acknowledgments

This work was supported in part by Fundação para a Ciência e Tecnologia (FCT) under Grant SFRH/BD/33738/2009, and by the European Union within the framework of the European Project MERGE-Mobile Energy Resources in Grids of Electricity, Contract 241399 (7th Framework Programme). The authors acknowledge Filipe J. Soares from INESC TEC for providing the electric vehicles synthetic time series data used in this paper.

References