Optimal Operation of Community Energy Storage using Stochastic Gradient Boosting Trees

Juan S. Giraldo*, Mauricio Salazar*, Pedro P. Vergara[†], Georgios Tsaousoglou*, J.G. Slootweg^{*‡},
and Nikolaos G. Paterakis*

*Dep. of Electrical Eng., Eindhoven University of Technology, Eindhoven, the Netherlands

[‡]Enexis, 's-Hertogenbosch 5223MB, the Netherlands

Email: {j.s.giraldo.chavarriaga, e.m.salazar.duque, g.tsaousoglou, j.g.slootweg, n.paterakis}@tue.nl

[†]Faculty of Electrical Eng., Math. & Comp. Sc., Delft University of Technology, the Netherlands
Email: p.p.vergarabarríos@tudelft.nl

Abstract—This paper proposes an algorithm for the optimal operation of community energy storage systems (ESSs) using a machine learning (ML) model, by solving a nonlinear programming (NLP) problem iteratively to obtain synthetic data. The NLP model minimizes the network's total energy losses by setting the operation points of a community ESS. The optimization model is solved recursively by Monte Carlo simulations in a distribution system with high PV penetration, considering uncertainty in exogenous parameters. Obtained optimal solutions provide the training dataset for a stochastic gradient boosting trees (SGBT) ML algorithm. The predictions obtained from the ML model have been compared to the optimal ESS operation to assess the model's accuracy. Furthermore, the sensitivity of the ML model has been tested considering the sampling size and the number of predictors. Results showed an accuracy of 98% for the SGBT model compared to optimal solutions, even after a reduction of 83% in the number of predictors.

Index Terms—Community energy storage, optimal energy management, stochastic gradient boosting trees, machine learning, renewable energy integration.

I. INTRODUCTION

Installed capacity of photovoltaic (PV) generation has been continuously increasing over the last few decades [1]. However, most of the low-voltage (LV) distribution networks are not prepared for the massive integration of prosumers since they were designed decades ago, leading to possible nodal overvoltages, reverse power flow problems, or overloading of distribution transformers [2].

There are several approaches towards managing the aforementioned problems, such as controlling the amount of power injected from PV units to the grid using inverters with droop control for overvoltage mitigation [3]. Another suitable technical solution for distribution networks with high PV penetration are energy storage systems (ESSs). The main idea behind it is to reduce the curtailment of PV power by storing excess generation at some periods and using that stored energy to provide ancillary services like voltage regulation or congestion management [4]. There are two main

approaches for coordinating the operation of ESSs in residential distribution networks, namely, household energy storage management and community energy storage management [5]. The former involves the individual control of several ESSs based on local user information, such as voltage measurements at the interconnection point, household consumption, and PV generation. The latter, is perceived as an intermediate solution between household and grid-scale ESS, in which the control of a single energy storage asset is required to achieve the desired objectives in residential areas within the kWh-MWh scale [6]. The main objective of a community ESS is to generate collective socioeconomic benefits, including the support of higher penetration of renewable energy sources, reduction in energy bills, reduction in technical power losses, and even rising revenue from offering demand flexibility [7]. Moreover, community ESS presents several advantages over household storage, such as the diminution on the size of the total required storage capacity, the amount of inverted power flow, while enhancing the self-sufficiency level of local smart energy communities [8].

An important aspect of decision making in ESS management problems is the way in which uncertainty is treated. A part of the literature disregards the uncertainty over system parameters (e.g. PV generation) by solving deterministic optimization problems assuming perfect forecasts. Closer to this work, another family of studies proposes end-to-end learning methods, i.e., algorithms trained to learn optimal decisions directly (in contrast to learning to predict system parameters). In [9], an iterative Q-learning method was proposed to solve the optimal battery management and control problem in smart residential environments. A reinforcement learning-based algorithm is used in [10] for the optimal coordination of different ESSs in microgrids. In [4], a predictive supervised machine learning (ML) using neural networks and random forests is applied for scheduling the real-time operation mode of the next operation interval for residential PV/battery systems controlled by mode-based controllers. Finally, a methodology for the home energy management of different appliances is proposed in [11] based on deep reinforcement learning and proximal policy optimization. However, it should be noted that these

G. Tsaousoglou received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No.754462.

Fig. 1. Proposed optimal operation of community ESS using ML.

applications are focused on solving the household energy storage management problem.

The implementation of a community ESS is discussed in [12], where a total of 2 MWh was connected to a 13.2 kV distribution feeder providing load leveling, reactive power compensation, and reliability enhancement. In [13], authors describe the development of a control algorithm for a community ESS of 400 kVA/232 kWh in an LV distribution feeder using forecasting techniques. The optimal operation of a community ESS to perform time-shifting on PV generation and demand load simultaneously is presented in [14], while authors in [15] propose a control framework for a community battery for grid congestion reduction, presenting experimental results with a battery of 55 kW and a capacity of 126 kWh. However, it should be noted that the cited works use traditional regression or forecasting techniques to predict the community ESS operation.

This paper proposes an algorithm for the near-optimal operation of community ESS under uncertainty, by iteratively solving different instances of an optimization problem and using the input-output data to perform supervised learning with a stochastic gradient boosting trees (SGBT) learning method [16]. Although there is plenty of scientific literature on household ESS applications using ML algorithms and on community ESS using traditional techniques, to the best of the authors knowledge, there is a gap regarding the operation of a community ESS considering uncertainty using ML. A nonlinear programming problem (NLP), modelling the optimal energy management (OEM) of the community ESS, has been implemented in Pyomo and solved using the open source solver Ipopt. The OEM has been solved recursively using Monte Carlo simulations (MCS), providing the training and test sets for the ML algorithm. The accuracy of the SGBT learning method has been assessed regarding its sensitivity to the size of the training dataset, the accuracy of the predictions, and its sensitivity to the number of predictors.

II. OPTIMAL OPERATION OF COMMUNITY ENERGY STORAGE SYSTEMS USING MACHINE LEARNING

The methodology proposed in this paper is based on the use of recursive solutions of the OEM optimization problem in Section II-A as training data for ML algorithms to obtain fast, reliable, and optimal operation of community ESS. The proposed approach is conformed by four main stages: System data, dataset, supervised ML modeling, and results assessment; as depicted in Fig. 1. The system data stage comprises the information of the network and needed forecasts, e.g., topology, operational limits, power demand, PV generation.

Dataset corresponds to the stage involving the acquisition of the required data, which can be obtained either by real measurements or via synthetic data (as performed in the presented approach). The output of the dataset, represented by z_n , is then used as training and test sets for the ML algorithm. The nested cross-validation [17] method is proposed to obtain the appropriate values for the hyperparameters with low bias. The last section of the methodology will provide a main predictor analysis, where the sensitivity of the ML algorithm will be assessed using complete and reduced models according to the number of predictors, represented in Fig. 1 as $\hat{Y}_c$ and $\hat{Y}_r$, respectively.

A. Optimal Energy Management Model (OEM)

The objective function for the community ESS has been defined as minimizing the total energy losses of the network in (1) over the state variables $z^{(x)} = \{P, Q, I, V, SOC\}$. The problem constraints are represented by (2)–(13):

$$\begin{aligned}
 & \min_{z^{(x)} = \{P, Q, I, V, SOC\}} \sum_{t \in \Omega_T} \Delta_t \sum_{ki \in \Omega_L} R_{ij} I_{ij,t}^{sqr} \quad (1) \\
 & \text{s.t.} \\
 & \sum_{ki \in \Omega_L} P_{ki,t} - \sum_{ij \in \Omega_L} (P_{ij,t} + R_{ij} I_{ij,t}^{sqr}) + P_{i,t}^S + \\
 & + \sum_{e \in \Omega_E: e=i} P_{e,t}^{ess} = P_{i,t}^D - P_{i,t}^{PV}, \quad \forall i \in \Omega_B, t \in \Omega_T \quad (2) \\
 & \sum_{ki \in \Omega_L} Q_{ki,t} - \sum_{ij \in \Omega_L} (Q_{ij,t} + X_{ij} I_{ij,t}^{sqr}) + Q_{i,t}^S + \\
 & + \sum_{e \in \Omega_E: e=i} Q_{e,t}^{ess} = Q_{i,t}^D - Q_{i,t}^{PV}, \quad \forall i \in \Omega_B, t \in \Omega_T \quad (3) \\
 & V_{i,t}^{sqr} - 2(R_{ij} P_{ij,t} + X_{ij} Q_{ij,t}) - (R_{ij}^2 + X_{ij}^2) I_{ij,t}^{sqr} = V_{j,t}^{sqr}, \\
 & \quad \forall ij \in \Omega_L, t \in \Omega_T \quad (4) \\
 & V_{j,t}^{sqr} I_{ij,t}^{sqr} = P_{ij,t}^2 + Q_{ij,t}^2, \quad \forall ij \in \Omega_L, t \in \Omega_T \quad (5) \\
 & \underline{V}^2 \leq V_{i,t}^{sqr} \leq \bar{V}^2, \quad \forall i \in \Omega_B, t \in \Omega_T \quad (6) \\
 & 0 \leq I_{ij,t}^{sqr} \leq \bar{I}^2, \quad \forall ij \in \Omega_L, t \in \Omega_T \quad (7) \\
 & SOC_{e,t} = SOC_{e,t-1} - \frac{\Delta_t}{EC_e} P_{e,t}^{ess}, \quad \forall e \in \Omega_E, t \in \Omega_T \quad (8) \\
 & -\bar{P}_e^{\text{ramp}} \leq P_{e,t}^{ess} - P_{e,t-1}^{ess} \leq \bar{P}_e^{\text{ramp}} \quad \forall e \in \Omega_E, t \in \Omega_T \quad (9) \\
 & -\bar{P}_e \leq P_{e,t}^{ess} \leq \bar{P}_e \quad \forall e \in \Omega_E, t \in \Omega_T \quad (10) \\
 & P_{e,t}^{ess} \tan(\cos^{-1}(\text{pf}_e)) \geq |Q_{e,t}^{ess}| \quad \forall e \in \Omega_E, t \in \Omega_T \quad (11) \\
 & \underline{SOC}_e \leq SOC_{e,t} \leq \overline{SOC}_e \quad \forall e \in \Omega_E, t \in \Omega_T \quad (12) \\
 & SOC_{e,0} \leq SOC_{e,t} \quad \forall e \in \Omega_E, t \in \Omega_T : t = |\Omega_T| \quad (13)
 \end{aligned}$$

where x represents the exogenous parameters, i.e., $\{P_{i,t}^{PV}, P_{i,t}^D\}_{i \in \Omega_B, t \in \Omega_T}, SOC_{e,0}\} = x$. $R_{ij} \forall ij \in \Omega_L$ represents the branch resistance between buses $i \in \Omega_B$ and $j \in \Omega_B$, while $I_{ij,t}^{sqr} \triangleq I_{ij,t}^2$ is the squared current flowing through the branch $ij \in \Omega_L$ at period $t \in \Omega_T$, where $I = \{I_{ij,t}^{sqr}, \forall ij \in \Omega_L\}_{t \in \Omega_T}$. Active and reactive power flows are represented in (2) and (3), respectively. Specifically,

the tuple of active power injections is represented by $\mathbf{P} = \{P_{ki,t} \forall k,j \in \Omega_L, P_{i,t}^S \forall i \in \Omega_B, P_{e,t}^{ess} \forall e \in \Omega_E\}_{t \in \Omega_T}$, which stands for the active power flowing through the branches, the active power injected by the substation ($i \in \Omega_B : i = ES$), and the active power from the ESS, usually connected the secondary side of the substation transformer. Similarly, the tuple of reactive power injections is defined as $\mathbf{Q} = \{Q_{ki,t} \forall k,j \in \Omega_L, Q_{i,t}^S \forall i \in \Omega_B, Q_{e,t}^{ess} \forall e \in \Omega_E\}_{t \in \Omega_T}$. The voltage magnitude drop between buses i and j is defined by (4), where $V_{i,t}^{sqr} \triangleq V_{i,t}^2$ is the squared voltage magnitude at bus i period t , and $\mathbf{V} = \{V_{i,t}^{sqr}, \forall i \in \Omega_B\}_{t \in \Omega_T}$. Branch current magnitudes are calculated in (5). Minimum and maximum nodal voltage magnitudes are limited by (6), while branch current magnitudes by (7). The tuple containing the state of charge of the community ESS is represented by $\mathbf{SOC} = \{SOC_{e,t}, \forall e \in \Omega_E\}_{t \in \Omega_T}$ and it is calculated in (8), where Δ_t is the length between two consecutive time periods and EC_e is the total storage capacity. Charging and discharging maximum power ramps are constrained in (9) by $\bar{P}_e^{\text{ramp}}$, while the rated power of the storage system is limited in (10) by $\bar{P}_e$. The reactive power injected by the ESS is limited in (11) representing the minimum power factor (pf_e) of a four-quadrant power converter. Finally, constraint (12) represents the maximum depth of discharge of the ESS and (13) guarantees the state of charge at the final time period to be at least equal to the initial value, $SOC_{e,0}$. Notice that the nonlinearity of the model is only introduced by (5) and it could be convexified by relaxing the model as in [18], obtaining a second-order cone programming problem.

III. PROBABILISTIC OEM

A. OEM under Uncertainty

The OEM problem of Section II-A could be solved deterministically if all the system parameters were known beforehand. However, the PV generation $P_{i,t}^{\text{PV}}$, the electricity demand $P_{i,t}^{\text{D}}$, and the initial state of charge $SOC_{e,0}$ are probabilistic parameters and their values for timeslot t are uncertain. Thus, the ESS energy management problem has been formulated as a problem of decision making under uncertainty. It was assumed that the values of $P_{i,t}^{\text{PV}}$, $P_{i,t}^{\text{D}}$, and $SOC_{e,0}$ follow known probability distributions:

$$P_{i,t}^{\text{D}} \sim \text{Normal}(\mu_{i,t}^{\text{D}}, \sigma_{i,t}^{\text{D}}) \quad \forall i \in \Omega_B, t \in \Omega_T \quad (14)$$

$$P_{i,t}^{\text{PV}} \sim \text{Beta}(\mu_{i,t}^{\text{PV}}, \sigma_{i,t}^{\text{PV}}) \quad \forall i \in \Omega_B, t \in \Omega_T \quad (15)$$

$$SOC_{e,0} \sim \text{Uniform}(\underline{SOC}_e, \overline{SOC}_e) \quad \forall e \in \Omega_E \quad (16)$$

For the sake of simplicity, it has been assumed that $Q_{i,t}^{\text{PV}} = 0$ and that $Q_{i,t}^{\text{D}}$ follows the active power with a constant power factor. Furthermore, the values for parameters $\mu_{i,t}^{\text{D}}, \sigma_{i,t}^{\text{D}}, \mu_{i,t}^{\text{PV}}, \sigma_{i,t}^{\text{PV}}$ are assumed to be known from historical data. Under these rather mild assumptions, the problem of OEM under uncertainty could be formulated as a Markov decision process (MDP) $\mathcal{G}$. The MDP $\mathcal{G}$ features continuous state and action variables; thus, it cannot be tackled by standard MDP-solving frameworks (e.g. dynamic programming).

B. Monte Carlo Simulations

The MCS method has been used in this paper to consider different case scenarios. MCS are mathematical tools for inferring the outcome of physical stochastic processes by generating ($\mathbf{x}_n$) and evaluating a set $\mathcal{N}$ of random samples recursively for a number of simulations $N = |\mathcal{N}|$. Each set of random samples consists of one generated value for each random variable chosen accordingly based on its stochastic information, i.e., its cumulative distribution function (CDF), F . The static model representing the physical system is then evaluated, providing one set of output results correspondingly ($\mathbf{z}_n$), containing one set of output results for each set of random samples [19].

Let $\mathbf{p}$ denote the set of CDFs for the exogenous random variables $\{P_{i,t}^{\text{PV}}, P_{i,t}^{\text{D}}\}_{i \in \Omega_B, t \in \Omega_T}$ and for the initial state of charge $SOC_{e,0}$, i.e., $\mathbf{p} = \{F(P_{i,t}^{\text{PV}}), F(P_{i,t}^{\text{D}}), F(SOC_{e,0})\}$. Also, let $\mathbf{x}_n = \{\{P_{i,t}^{\text{PV},(n)}, P_{i,t}^{\text{D},(n)}\}_{i \in \Omega_B, t \in \Omega_T}, SOC_{e,0}^{(n)}\}_n$ denote an instantiation n of the realization trajectory of these random variables. A number N of such instantiations (scenarios) is generated using the inverse-transform method and evaluated by solving problem (1)-(13) for each scenario. The results from the MCS, represented by $\mathbf{z}_n^{(x_n)} = \{\mathbf{P}, \mathbf{Q}, \mathbf{I}, \mathbf{V}, \mathbf{SOC}\}_n$, $\forall n \in \mathcal{N}$, contains the active and reactive powers injected at each node and flowing through all branches, (including the power from/to the community ESS, $\mathbf{P}_{\text{ESS}}, \mathbf{Q}_{\text{ESS}}$), all nodal voltage magnitudes, all branch current magnitudes (including the loading percentage of the substation transformer $\mathbf{I}_{\text{ST}}$), and the state of charge of the community ESS for every tested scenario.

IV. SUPERVISED MACHINE LEARNING MODELING

A. Prediction of the Operating Set Points

The main goal of the supervised learning model is to predict the active and reactive power set points of the community ESS for the different operating modes i.e., charge, discharge, or idle. The model is defined as a regression problem in which the set of predictor variables is a subset of the output from the OEM, i.e., $\mathbf{X} \subset \mathbf{z}^*$, where $\mathbf{z}^* = \mathbf{z}_n^{(x_n)}, \forall n \in \mathcal{N}$. Specifically, the dataset corresponds to the optimal solutions for the transformer loading percentage, the nodal voltage magnitudes, the state of charge of the community ESS, and the time step of the day t . Thus, the input dataset corresponds to the tuples, $\mathbf{X} = \{t, \mathbf{I}_{\text{ST}}^*, \mathbf{V}^*, \mathbf{SOC}^*\}_{k=1}^K$ for $K = |\Omega_T| |\mathcal{N}|$. The predicted variables corresponds to the active and reactive power injections from the community ESS, which are also contained in $\mathbf{z}^*$, namely, $\mathbf{Y} \subset \mathbf{z}^*$, where the predicted dataset is $\mathbf{Y} = \{\mathbf{P}_{\text{ESS}}^*, \mathbf{Q}_{\text{ESS}}^*\}_{k=1}^K$. The problem is simplified by training two separated regression models $\mathbf{Y}_P = f_P(\mathbf{X})$ and $\mathbf{Y}_Q = f_Q(\mathbf{X})$, responsible for predicting the set points of the active and reactive power, respectively. The objective is to obtain a trained model able to generalize over new states of the grid, and provide optimal predictions, $\hat{\mathbf{Y}}$, close to the optimal solution for the ESS operation in the distribution grid i.e., $\mathbf{P}_{\text{ESS}}^* \cong \hat{\mathbf{Y}}_P$ and $\mathbf{Q}_{\text{ESS}}^* \cong \hat{\mathbf{Y}}_Q$.

We focused on the well-known SGBT algorithm [16], which is known for being one of the best off-the-shelf supervised learning algorithms [20]. The SGBT algorithm creates multiple decision tree models, called weak learners, which are then combined to make a single regression estimate. The creation of the decision trees is iterative. In each iteration, the sampling on the training samples to build new trees is focused on the observations that were not accurately predicted by previous weak learners; enhancing the SGBT model to learn multiple sets of rules. The SGBT is classified as an embedded algorithm for predictors selection, meaning that the most important predictors are implicitly selected in the algorithm, providing a predictor relevance measure.

B. Model selection and reduction

The performance metric to be minimized in the SGBT regression model is the root mean squared error (RMSE). The nested cross-validation method [17] is used with a three fold purpose in our modeling: (i) calculate a lower bias metric of the performance of the algorithm in the dataset $\mathbf{X}$, (ii) analyze the predictor importance based on the best models of each outer fold, and (iii) create a subset of hyperparameters to perform a model assessment for the final model. The inner cross-validation is used for hyperparameter tuning, which helps to discard models that are overfitting the dataset. The hyperparameter tuning is done using a random search method over a predefined range of values for each hyperparameter. To estimate the generalization error and select the best final model, the outer cross-validation is used to compute the performance metrics of the algorithm and to create a subset of models that are then evaluated in one single cross-validation. The samples for the inner and outer folds are selected using a stratified approach. The scenario number, n , of the dataset $\mathbf{X}$ is used for the stratification. The purpose of this is to avoid imbalance cases with folds having more samples from one specific time steps, t , than others.

As mentioned in Section IV-A, SGBT provides a predictor relevance measure by calculating the gain of each tree over a given predictor. Nevertheless, in our case, we expect that most predictors, i.e., voltage magnitudes in neighbouring nodes, would have high correlations due to their electrical interdependence. Predictors with high correlation tend to reduce the reliability of the predictor relevance measure. Hence, a permutation-based feature/predictor importance (PFI), as in [21], is tested in our case study to perform the predictor rank and selection of the reduced models.

V. TESTS AND ASSESSMENT

The proposed methodology has been tested in a radial, 11kV distribution system with 34 buses, shown in Fig. 2, with a peak total nominal power of 6.2MW and 4.1Mvar [22]. A community ESS has been added to the LV side of the substation transformer with a total capacity of $EC_e = 6$ MWh, a maximum power and maximum power ramp of $\bar{P}_e^{\text{ramp}} = \bar{P}_e = 3$ MW, state of charge limits of $\text{SOC}_e = 30\%$ and $\text{SOC}_e = 100\%$, and a minimum power

Fig. 2. 34-bus test system.

Fig. 3. Loading percentage and voltage magnitudes.

factor of $\text{pf}_e = 0.9$. Voltage magnitude limits have been set to $\underline{V} = 0.9$ p.u. and $\bar{V} = 1.1$ p.u., while the maximum current of the transformer has been set to 700 A. All simulations were performed for a 24 hour period in intervals of one hour. Finally, PV units have been added to all load buses with their nominal capacity equal to twice the bus load; hence, the base case represents 200% penetration. Four tests have been performed to assess the performance of the methodology:

- *T1*: MCS using conventional power flows without the inclusion of the community ESS to provide a base case (no enforcement of operational constraints).
- *T2*: MCS of the OEM as in Section III-B to analyze the effect of adding a community ESS.
- *T3*: Hyperparameter tuning and accuracy assessment of the SGBT using the normalized RMSE (NRMSE) to analyze the outputs of the ML model. Then, analyze the improvement of the best model as a function of the size of the dataset provided to the SGBT algorithm.
- *T4*: Predictor importance analysis using predictor information gain and PFI. Accuracy assessment of the results using the NRMSE as a function of the number of predictors, analyzing the error obtained by the reduced models.

The first test, *T1*, has been performed for $N = 500$ random scenarios defining the base state of the system. The loading percentage of the distribution transformer and the maximum and minimum voltage magnitudes are depicted in Fig. 3. It can be seen that the base case presents overloading scenarios

Fig. 4. Normalized prediction error $f_P(\cdot)$ as a function of the number of scenarios.

of almost 60% (mostly due to reverse power flows) and over/under voltage problems during some periods (e.g., around 12h and 19h). In the same figure, it is evident that the optimal operation of the community ESS, tested in T_2 for $N = 500$ as well, provides flexibility and presents a solution to the congestion management problem guaranteeing minimal power losses and operation within the predefined limits.

In the test T_3 , a number of $N=200$ scenarios were used to assess the performance of the SGBT algorithm. A nested, 50-outer and 5-inner, cross-validation folds were carried out, obtaining an NRMSE of $2.30\% \pm 0.33\%$ which is the mean of the outer folds of the 50 best models. In order to select the final models, a 5-fold cross-validations is done for each one of the best models. The hyperparameters for the final models of $f_P(\cdot)$, $f_Q(\cdot)$, and its metrics, are summarized in Table I.

A series of tests were performed to assess the sensitivity of the learning model with respect to the size of the training set (scenarios), as explained in T_3 . The NRMSE of the model $f_P(\cdot)$ as a function of the used scenarios is depicted in Fig. 4 from 10–1000 scenarios. It can be seen that the NRMSE of the predictions provided by the model decreases considerably within the first 100 scenarios, and continues decreasing after that with a lower slope. Specifically, it was found that the NRMSE is $3.38\% \pm 0.81\%$ for 10 scenarios, $1.92\% \pm 0.10\%$ for 200 scenarios and $1.75\% \pm 0.04\%$ for 1000 scenarios. It should be noted that increasing the dataset size for training the ML model increases the computational burden without a meaningful reduction on the NRMSE. Hence, 200 scenarios are enough to provide good quality results for this test case. A similar conclusion can be drawn for the $f_Q(\cdot)$ model.

Figure 5 shows a comparison between the power injection from/to (positive/negative) the community ESS using the optimization model (OEM), and the predictions obtained with

Fig. 5. Active and reactive power set points for the community ESS obtained with the OEM and with the SGBT.

Fig. 6. Sorted predictor importance by a) Predictor information gain. b) PFI

the final ML models. The predictions displayed in the figure are the results of one of the testing folds of the 5-fold cross-validation over the final ML models, i.e., the dataset was not part of the training set, meaning that these instances were unexplored by the model. From Fig. 5, it can be seen that both active and reactive power injections obtained with the ML models follows closely the actual injections found with the optimization problem with a maximum absolute error of 73.4 kW (2.44%) and an NRMSE lower than 2.00%, providing near-optimal operation.

A predictor importance analysis has been performed, as explained in T_4 , calculating the predictor relevance among the decision trees built by the SGBT algorithm. The SGBT is stochastic by design e.g., due to random sampling on the training datasets and predictors to build the trees. Subtle differences in predictor importance ranking can be seen between models even with the same hyperparameters values. Therefore, the analysis has been executed for the 50 best models from the nested cross-validation analysis in T_3 , to obtain the average and standard deviation of the predictors' gains, as shown in Fig. 6 and sorted by importance. Similarly, a predictor importance analysis has been performed based on predictor permutation as depicted in Fig. 6b. Notice that the order of the most important predictors is not strictly the same for both importance metrics; for example, take v_{28} and v_{29}

TABLE I
SUMMARY - HYPERPARAMETERS OF SGBT MODELS

Hyper parameter	Search space	Best $f_P(\cdot)$	Best $f_Q(\cdot)$
Num. of trees	(100-1000)	886	793
Shrinkage	U(0, 1)	0.04	0.01
Subsample	U(0.2 - 0.7)	0.59	0.50
Tree depth	(2 - 10)	8	9
Cross validation score (NRMSE)	-	$1.92\% \pm 0.10\%$	$1.94\% \pm 0.12\%$

Fig. 7. Sensitivity of SGBT to the number of predictors.

corresponding to the voltage magnitudes at buses 28 and 2, respectively.

Finally, the sensitivity of the SGBT model to the number of predictors has been tested. It has been done by calculating the NRMSE after adding predictors one by one in the order presented in Fig. 6. This is, v_{28} , v_6 , v_{29} , etc. for the predictor information gain; and Loading, time, v_2 , v_6 , etc. for the PFI. The results are shown in Fig. 7, where it can be seen that the results obtained with the PFI overperformed the information gain predictor importance metric. It can be seen that the inclusion of the first 6 predictors provided a NRMSE of approximately 2% with the PFI, whereas the information gain predictor was slightly higher. On the second metric, the information gain predictor importance flatlined after the 6th predictor and only achieved the best accuracy when 18th predictor was included; which, according to the order in Fig. 7, corresponds to v_2 . This indicates that the permutation-based technique was able to rank the importance of v_2 properly. This effect is expected when using highly correlated predictors carrying similar predictive information since the SGBT model selects the correlated predictors according to their frequency, which may mask their information gain importance metric.

VI. CONCLUSIONS

This paper proposed an ML approach for the optimal operation of community ESS under uncertainty using SGBT. The recursive solutions of NLP problems representing the OEM have been used as input/output data of a supervised learning model. The objective of the learning model is to predict the optimal active and reactive power set points of the ESS. The obtained SGBT learning model was assessed using nested cross-validation and predictor permutation, considering its sensitivity to the size of the training dataset, the accuracy of the obtained predictions, and its sensitivity to the number of predictors. Results showed that the complete model is able to predict the operation of the community ESS close to optimality ($\sim 98\%$) using a small training dataset (~ 200 samples). Moreover, the reduced models presented a similar accuracy even after removing 83% of the predictors.

These results are promising for real implementations with limited training datasets and measurement infrastructure. Future work will consider a flexibility market structure with different aggregators, a deeper analysis of the predictor importance including noise and erroneous data, and different ML algorithms for comparison.

REFERENCES

- [1] "National Solar Trend Report 2020," Dutch New Energy Research, Tech. Rep., Jan. 2020. [Online]. Available: <https://bit.ly/2RS3upP>
- [2] H. Sadeghian and Z. Wang, "A novel impact-assessment framework for distributed PV installations in low-voltage secondary networks," *Renewable Energy*, vol. 147, pp. 2179–2194, Mar. 2020.
- [3] P. P. Vergara, M. Salazar, T. T. Mai, P. H. Nguyen, and H. Sloopweg, "A comprehensive assessment of PV inverters operating with droop control for overvoltage mitigation in LV distribution networks," *Renewable Energy*, vol. 159, pp. 172 – 183, 2020.
- [4] G. Henri and N. Lu, "A supervised machine learning approach to control energy storage devices," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 5910–5919, Nov. 2019.
- [5] B. P. Roberts and C. Sandberg, "The role of energy storage in development of smart grids," *Proceedings of the IEEE*, vol. 99, no. 6, pp. 1139–1144, 2011.
- [6] D. Parra, M. Swierczynski, D. I. Stroe, S. A. Norman, A. Abdon, J. Worlitschek, T. O'Doherty, L. Rodrigues, M. Gillott, X. Zhang *et al.*, "An interdisciplinary review of energy storage for communities: Challenges and perspectives," *Renewable and Sustainable Energy Reviews*, vol. 79, pp. 730–749, May 2017.
- [7] B. P. Koirala, E. van Oost, and H. van der Windt, "Community energy storage: A responsible innovation towards a sustainable energy system?" *Applied energy*, vol. 231, pp. 570–585, Sep. 2018.
- [8] E. Barbour, D. Parra, Z. Awwad, and M. C. González, "Community energy storage: A smart choice for the smart grid?" *Applied Energy*, vol. 212, pp. 489–497, Feb. 2018.
- [9] Q. Wei, D. Liu, and G. Shi, "A novel dual iterative Q-learning method for optimal battery management in smart residential environments," *IEEE Trans. Ind. Electron.*, vol. 62, no. 4, pp. 2509–2518, Apr. 2015.
- [10] X. Qiu, T. A. Nguyen, and M. L. Crow, "Heterogeneous energy storage optimization for microgrids," *IEEE Trans. Smart Grid*, vol. 7, no. 3, pp. 1453–1461, May 2016.
- [11] H. Li, Z. Wan, and H. He, "A Deep Reinforcement Learning Based Approach for Home Energy Management System," in *2020 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*. Washington, DC, USA: IEEE, 2020.
- [12] P. R. Thomas, T. J. Walker, and C. A. McCarthy, "Demonstration of Community Energy Storage fleet for load leveling, reactive power compensation, and reliability improvement," in *2012 IEEE Power and Energy Society General Meeting*, San Diego, CA, 2012.
- [13] J. Van Dun, R. de Groot, J. Morren, and J. Sloopweg, "Control of a battery energy storage system connected to a low voltage grid," in *2015 IEEE Eindhoven PowerTech*. IEEE, 2015, pp. 1–6.
- [14] D. Parra, S. A. Norman, G. S. Walker, and M. Gillott, "Optimum community energy storage for renewable energy and demand load management," *Applied energy*, vol. 200, pp. 358–369, 2017.
- [15] W. van Westering and H. Hellendoorn, "Low voltage power grid congestion reduction using a community battery: Design principles, control and experimental validation," *International Journal of Electrical Power & Energy Systems*, vol. 114, p. 105349, Jan. 2020.
- [16] J. H. Friedman, "Stochastic gradient boosting," *Computational Statistics and Data Analysis*, vol. 38, no. 4, pp. 367–378, 2002.
- [17] S. Varma and R. Simon, "Bias in error estimation when using cross-validation for model selection," *BMC bioinformatics*, 2006.
- [18] J. S. Giraldo, J. A. Castrillon, J. C. López, M. J. Rider, and C. A. Castro, "Microgrids energy management using robust convex programming," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4520–4530, Jul. 2019.
- [19] S. Raychaudhuri, "Introduction to Monte Carlo simulation," in *2008 Winter Simulation Conference*. IEEE, 2008, pp. 91–100.
- [20] R. Caruana, N. Karampatziakis, and A. Yessenalina, "An empirical evaluation of supervised learning in high dimensions," in *Proceedings of the 25th International Conference on Machine Learning*, 2008.
- [21] G. Casalicchio, C. Molnar, and B. Bischl, "Visualizing the feature importance for black box models," in *Machine Learning and Knowledge Discovery in Databases*. Springer, 2019, pp. 655–670.
- [22] LaPSEE Power System Test Cases Repository UNESP, "Downloads / sistemas testes," <https://bit.ly/2IdSn91>, 2020, [Online; accessed October-2020].