Electricity Price Forecasting for Operational Scheduling of Behind-the-Meter Storage Systems

Hamed Chitsaz, *Student Member, IEEE*, Payam Zamani-Dehkordi, *Student Member, IEEE*, Hamidreza Zareipour, *Senior Member, IEEE*, and Palak P. Parikh, *Member, IEEE*

***Abstract*—Electricity price forecast plays a key role in strategic behavior of participants in competitive electricity markets. With the growth of behind-the-meter energy storage, price forecasting becomes important in energy management and control of such small-scale storage systems. In this paper, a forecasting strategy is proposed for real-time electricity markets using publicly available market data. The proposed strategy uses high-resolution data along with hourly data as inputs of two separate forecasting models with different forecast horizons. Moreover, an intra-hour rolling horizon framework is proposed to provide accurate updates on price predictions. The proposed forecasting strategy has the capability to detect price spikes and capture severe price variations. The real data from Ontario's electricity market is used to evaluate the performance of the proposed forecasting strategy from the statistical point of view. The generated price forecasts are also applied to an optimization platform for operation scheduling of a battery energy storage system within a grid-connected micro-grid in Ontario to show the value of the proposed strategy from an economic perspective.**

***Index Terms*—Price forecasting, high-resolution data, price spikes, energy storage systems, micro-grids.**

I. Introduction

THE MARKET for behind-the-meter (BTM) battery energy storage is growing rapidly. For example, in the third quarter of 2015, more than 13 MW of such units were installed, which indicates a 15 times year-over-year growth [1]. Behind-the-meter storage enables consumers to have more control over their own energy usage [2]. Depending on the market structure, the use of BTM storage may vary. Where flat rates are charged by utilities, a common use of BTM storage is for demand charge management [3]. In competitive markets where large customers are charged at wholesale pool price rates, BTM storage could be employed for avoiding peak prices. In such cases, an insight into price fluctuations is essential for the optimal use of the storage system. The focus of this work is the operation of a storage system owned by a large consumer that purchases electricity from the wholesale market, and thus, needs short-term price forecasting (STPF) for operation scheduling.

Most electricity price forecasting studies in the literature have considered short horizons because of the proximity to the real-time operation [4]. Different methodologies have been proposed for point, probabilistic, and threshold forecasting of the electricity market price in [4]. While the point forecasting approaches provide single-valued predictions [5], the probabilistic forecasting models can quantify the uncertainties associated to point forecasts using prediction intervals [6]–[8]. Alternatively, certain applications, e.g., demand-side management, do not require exact values for future prices but apply pre-specified price thresholds as the basis for the decision-making process [9]. Therefore, the threshold forecasting models work as classification problems and simply provide different classes for future prices [10].

Moreover, a few research works have presented methodologies aiming to detect price spikes using statistical [11]–[13] or data mining approaches [14]–[16]. In [11], a recursive dynamic factor analysis (RDFA) is combined with a Kalman Filter model for electricity price forecast. It is shown that the proposed model has better forecasting accuracy than three other approaches under the presence of price spikes, however, a solid price spike detection analysis has not been provided. Amjady and Keynia [12] propose a closed loop prediction mechanism including neural networks and feature selection techniques to predict both price spike occurrences and values. An autoregressive approach has been used to model the time series of price spikes in Australian electricity market in [13]. Classification techniques along with feature selections [14], [16] and similarity searching methods [15] have also been applied to detect price spikes. However, these proposed methodologies are based on analyzing the historical data. Such analyses might not be reliable in practice because the most influential factors creating price spikes are unplanned generation or transmission line outages, which are neither predictable nor possible to model [9]. In addition, price spikes are distinguished by defining fixed price thresholds in presented works. Despite simple implementation, it may not be efficient as price statistics vary significantly from month to month.

Most STPF models developed in the literature use hourly historical/forecast data of different explanatory variables [17] to predict hourly prices. Maciejowska and Weron [18]

Manuscript received August 24, 2016; revised November 30, 2016, February 26, 2017, and April 11, 2017; accepted May 16, 2017. Date of publication June 19, 2017; date of current version October 19, 2018. This work was supported in part by GE Digital Energy and in part by the National Science and Engineering Research Council of Canada. Paper no. TSG-01129-2016. *(Corresponding author: Hamidreza Zareipour.)*

H. Chitsaz, P. Zamani-Dehkordi, and H. Zareipour are with the Department of Electrical and Computer Engineering, University of Calgary, Calgary, AB T2N1N4, Canada (e-mail: hchitsaz@ucalgary.ca; pzamanid@ucalgary.ca; h.zareipour@ucalgary.ca).

P. P. Parikh is with the Department of Product Research and Development, General Electric Digital Energy, Markham, ON L6C 0M1, Canada (e-mail: palak.parikh@ge.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSG.2017.2717282

investigated that using intra-day prices with 30-min resolution could improve the short-term forecasts of base-load electricity prices in the U.K. market. However, the price settlement process in real-time markets is with a higher resolution than an hourly basis. For instance, Market Clearing Prices (MCPs) are set every five minutes in Ontario's [19], California's [20], Texas' [21] and New York's electricity markets [22], and every one-minute in Alberta's electricity market [23]. Such high-resolution market data contains recent updates on the electricity market conditions, and can be used along with hourly market data for predicting the hourly prices. The major benefit from utilizing higher resolution data is to capture price variations and detect price spikes efficiently.

In our previous works (e.g., [10], [24], and [25]), we demonstrated that unlike load forecasting where predicting the absolute value of demand is critical in operation scheduling, price forecasting accuracy should be measured in terms of economic savings/gains. Thus, the main contribution of the present paper is to propose an intra-hour rolling-horizon electricity price forecasting strategy that is specifically developed to optimize the operation of behind-the-meter Battery Energy Storage Systems (BESS). Hence, two objectives are of interest in this paper: an efficient price forecasting tool as well as optimal operation of the BESS using generated price forecasts. The proposed strategy takes advantage of high-resolution five-minute market clearing price data in the real-time market to generate hourly market price forecasts. Using intra-hour market clearing price information enables the method to detect if the current hour is going to have a high price. With the capability to capture high prices, the battery is discharged accordingly to offset energy purchase from the grid, and thus save on energy costs. A simple regression-based forecasting model is developed that is computationally very light and can be effectively used on micro-grid management firmware systems.

The remaining parts of the paper are organized as follows. The operation scheduling of a BESS and the proposed price forecasting strategy are described in Section II. In Section III, Ontario's electricity market is briefly presented as the case study in this paper. Afterwards, the proposed forecasting strategy is evaluated from both statistical and economic perspectives. Finally, Section IV concludes the paper.

II. Methodology

In this section, the operation of a BESS is outlined considering different forecasting perspectives. The components of the proposed forecasting strategy are then described.

A. Operation of a BESS

With operating a behind-the-meter BESS, the energy can be stored during off-peak hours when prices are low and then injected back to load/grid during peak periods to reduce the amount of energy purchased from the wholesale market at high prices. The price-based optimal operation scheduling of such a system can be presented as:

$$\mathop{\mathbf{Max}}_{P_t} \quad S = \sum_{t=1}^{T} P_t \hat{\lambda}_t$$
$$\text{Subject to } \Phi \tag{1}$$

where the objective function is to maximize the net energy arbitrage saving, denoted by S. T represents the scheduling horizon (e.g., $T = 24$ for day-ahead scheduling), $\hat{\lambda}_t$ (\$/MWh) is the price forecast for hour t, and P_t (MW) is the variable showing the power consumed ($P_t < 0$) from the grid or injected ($P_t > 0$) back to the micro-grid by the BESS at hour t. The objective function is subject to a set of battery operation and technical constraints, denoted by Φ, such as, charging/discharging rates, rated energy capacity, battery state of charge and number of cycles per day [26]–[28], as follows:

$$E_t = E^{ini} + \sum_{k=1}^{t} P_k \tag{2}$$
$$E^{emg} + (1 - DOD).E^{max} \le E_t \le E^{max} \tag{3}$$
$$P^{ch,max} \le P_t \le P^{dis,max} \tag{4}$$
$$M.u_t \le P_t \le M.(u_t - 1) \tag{5}$$
$$\sum_{t=1}^{T} u_t \le N_{Cycle} \tag{6}$$

Equations (2) and (3) ensure that the energy stored in the BESS at hour ending t, denoted by E_t, is within the allowable range. P_t is the amount of scheduled power for charging ($P_t < 0$) and discharging ($P_t > 0$). E^{ini} is the initial stored energy in the battery, E^{emg} is the energy related to the emergency load (150 kWh), and E^{max} is the maximum capacity for the battery (500 kWh). $P^{ch,max}$ and $P^{dis,max}$ are maximum charging and discharging rates. Equations (5) and (6) are auxiliary constraints to count the number of full cycles per day, denoted by N_{Cycle}. M is a big positive number and u_t is defined as a flag variable that is set to $u_t = 1$ when the battery is in the charging mode. It should be noted that this formulation only includes the basic, most dominant features of a storage unit in order to show the impact of price forecasts on the BESS's operation.

The solution of this optimization problem is the scheduled power of the BESS for each hour, denoted by P_t^*. It is noted in our work, we used Mixed-Integer Linear Programming (MILP) solver of the optimization toolbox in MATLAB to solve the optimization problem of battery operation. When the actual prices, λ_t, are realized, the after-the-fact value of the net energy arbitrage saving (S^*) can be calculated as follows:

$$S^* = \sum_{t=1}^{T} P_t^* \lambda_t \tag{7}$$

In this approach, the schedules are set once for the whole scheduling horizon with no changes. As a common practice to deal with uncertainties of forecasts, they are updated at each forecasting step (e.g., each hour). This is called the Rolling Horizon (RH) approach, which has successfully been applied to a few scheduling and energy management problems under load and wind uncertainties in [29] and [30]. Basically, the optimization problem is run on an hourly basis using the updates on price forecasts. As a result, schedules of the BESS can efficiently be updated once price forecasts are updated.

This approach is presented in Algorithm (1), where h represents the forecasting origin; that is, the hour at which the forecasts are generated and used in the optimization platform.

Algorithm 1 Scheduling Based on the Conventional RH

1: $h = 0$
2: **while** $h \leq T - 1$ **do**
3: Solve the optimization problem:
$$\underset{P_t}{\textbf{Max}} \;\; S = \sum_{t=h+1}^{T} P_t^{(h)} \hat{\lambda}_t^{(h)}$$
Subject to Φ
4: $h = h + 1$
5: **end while**

Algorithm 2 Scheduling Based on the Proposed IRH

1: $h = 1$
2: **while** $h \leq T$ **do**
3: Solve the optimization problem:
$$\underset{P_t}{\textbf{Max}} \;\; (1-\gamma)(P_t^{(h)} \hat{\lambda}_t^{(h)})\Big|_{t=h} + \sum_{t=h+1}^{T} P_t^{(h)} \hat{\lambda}_t^{(h)}$$
Subject to Φ
4: $h = h + 1$
5: **end while**

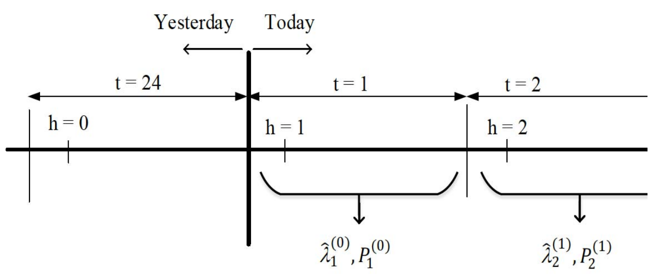


Fig. 1. Graphical representaion of Algorithm 1.

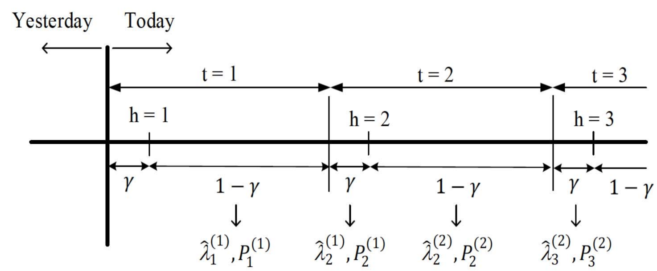


Fig. 2. Graphical representaion of Algorithm 2.

$h = 0$ is the last hour of day $D - 1$ when the first set of forecasts is generated for day D. $\hat{\lambda}_t^{(h)}$ is the price forecast for hour t generated at hour h, and $P_t^{(h)}$ is the power of the BESS for hour t scheduled at forecasting origin h. According to Algorithm (1), for scheduling the BESS from hour t up to hour T, the optimization problem is run one hour in advance, hour $t - 1$. For instance, forecasts are updated at hour 1, i.e., forecast origin $h = 1$, and fed into the optimization problem at the same hour to solve for the optimal scheduling of the BESS for hour 2 up to hour T $(t = 2, \ldots, T)$.

With this approach, updates on scheduling of the BESS leads to a more adaptive and efficient operation. The after-the-fact value of the net energy arbitrage saving is:

$$S^* = \sum_{t=1}^{T} P_t^{*(h)}\Big|_{h=t-1} \lambda_t \tag{8}$$

$P_t^{*(h=t-1)}$ is power of the BESS for hour t that has been scheduled at forecast origin h $(h = t - 1)$. In spite of successful application of this conventional RH approach in energy management problems, it may not be the most effective approach for operating a behind-the-meter BESS. The reason is that such a RH approach applies the market data of the previous hour to generate the forecast updates for the current hour. As illustrated in Fig. 1, the conventional RH provides price forecast $(\hat{\lambda}_1^{(0)})$ and the corresponding battery schedule $(P_1^{(0)})$ for the first hour of the current day $(t = 1)$ sometime in the last hour of the previous day $(t = 24)$ at forecasting origin $h = 0$. The battery schedule is fixed for the whole first hour. It is also noted that the schedules are provided for the whole operation horizon (e.g., up to $T = 24$), however, the schedules for $t = 2, \ldots, T$ will be updated once the next forecasts are available $(h = 1)$. This approach may not be able to efficiently capture electricity price spikes, as the most recent market data is not incorporated. Consequently, the operational scheduling of the BESS is negatively affected missing severe price variations.

In this paper, a forecasting strategy is proposed to provide more informative forecasts to enhance the operation of a BESS. In this strategy, severe variations in electricity price can potentially be captured by predicting the price for any hour during the same hour. The proposed rolling horizon framework is called Intra-hour Rolling Horizon (IRH). The operation scheduling of the BESS using the proposed time-frame is presented in Algorithm (2) and also illustrated in Fig. 2. Accordingly, the first updates on price forecasts are generated at forecast origin $h = 1$, which is the first hour of the current day (day D) as opposed to $h = 0$ (the last hour of day $D-1$) presented in Algorithm (1). As seen in Step 3 of Algorithm (2), the first hour is separated from the remaining hours of the scheduling horizon as the optimization problem is performed. γ depicts the fraction of time that the forecasting model requires to generate $\hat{\lambda}_t^{(h)}$ in practice. For this fraction of the current hour, the BESS follows the operation that has been scheduled in the previous update. Therefore, only the second fraction of the current hour proportional to $(1 - \gamma)$ is considered in the optimization problem for the current hour.

As shown in Fig. 2, at the forecasting origin $h = 1$ that corresponds to the first hour $(t = 1)$ of the current day, the first set of price forecasts are generated, i.e., $\hat{\lambda}_1^{(1)}, \hat{\lambda}_2^{(1)}, \ldots, \hat{\lambda}_T^{(1)}$. In a practical live forecasting system, it takes a fraction of the first hour, i.e., γ, before the forecasts become available. This is the time needed for the forecasting tool to fetch new data, process them and generate new forecasts and communicate them to the optimization platform. All communication delays and latencies are included in this time. Afterwards, the price forecasts are fed into the optimization problem to schedule the BESS for the $(1 - \gamma)$ fraction of the first hour, shown in the first term of step 3, and all remaining hours of the scheduling horizon, shown in the second term of step 3. The optimization platform applies the prices forecasts to provide the most economic hourly schedules for the BESS operation up to T. In the first fraction of the second hour ($t = 2$) associated to γ, the battery follows the operation instruction that was scheduled in the first hour for the second hour, i.e., $P_2^{(1)}$. Once the second updates on price forecasts become available (at forecast origin $h = 2$), the optimization problem apply them to update battery schedules for the remaining fraction of the current hour, $P_2^{(2)}$, and the remaining hours, i.e., $P_2^{(2)}, \ldots, P_T^{(2)}$. This process is repeated until the last hour of the operating horizon, T. As a result, the proposed IRH can update the price forecasts and consequently operation schedules of the battery for the current hour up to T while standing in the current hour. Whereas in the conventional RH approach, the latest update for the operation of battery for any given hour is prepared one hour in advance. The net energy arbitrage saving is calculated as follows:

$$S^* = (1-\gamma)P_1^{*(1)}\lambda_1 + \sum_{t=2}^{T}\left[\gamma P_t^{*(h)}\Big|_{h=t-1} + (1-\gamma)P_t^{*(h)}\Big|_{h=t}\right]\lambda_t \quad (9)$$

Equation (9) suggests that the saving value associated to the first hour of the scheduling period is proportional to $(1-\gamma)P_1^{*(1)}$. In other words, the first fraction of only the first hour of the scheduling horizon is not considered because the algorithm does not take into account the price forecast generated in the last hour of the previous day. For any remaining hour of the scheduling period, the first part of the saving comes from $\gamma P_t^{*(h)}|_{h=t-1}$ that includes the operation scheduled in the previous hour. The second part is related to the operation scheduled in the same hour, $(1-\gamma)P_t^{*(h)}|_{h=t}$. In this way, the optimization problem can benefit from more recent forecast updates generated during the operation hour. The operation scheduling of the BESS with the proposed strategy is evaluated and compared with the conventional RH approach in Section III.

B. Forecasting Strategy

In general, prediction models consist of three main components: 1) data pre-processing, 2) feature selection, and 3) model selection [9]. In this section, a brief background regarding each of these components are provided, and the methods developed in this work are described as well.

1) Data Pre-Processing: Also known as data cleaning, this component performs preliminary analyses on the raw data gathered for the prediction purpose, such as dealing with missing values, removing the outliers, and normalizing the data [31]. In this work, outliers such as price spikes are limited by defining a threshold calculated by the mean and variance of the times series. Hence, outliers remain in the time series as they carry important information about the nature of the time series, while their abnormal values are reduced to not negatively affect the learning capability of the forecasting model.

2) Feature Selection: This component selects a subset of features among all candidate features from the original dataset according to a feature goodness criterion [32]. There are several features influencing the electricity spot price, including historical load and price, imports/exports, capacity excess/shortfall, historical reserves, and generation types as well as calendar effects, e.g., day of week, month of year, seasonal and holiday effects [17]. In this paper, Mutual Information (MI) [33] technique was applied to the original database with hourly resolution to select the best subset that contains the least number of key features contributing to forecast accuracy, while discarding the remaining insignificant features. This feature selection technique is briefly described in the Appendix, and the detailed formulations can be found in [33].

3) Model Selection: The forecasting model constructs the input/output mapping function, where inputs come from the feature selection and the output is the forecast value. A number of approaches have been presented in the literature for electricity price prediction for various prediction horizons, prediction steps and applications. In general, the point forecasting engines fall into three main categories of statistical time series models, computational intelligence models, and hybrid models [17]. For instance, ensemble price forecasts from individual models are generated using linear regression models for price-directed demand management in smart grids [34]. In [35], an adaptive Wavelet Neural Network (WNN) is proposed for short-term price forecasting that outperforms a number of prediction approaches such as, statistical models (Auto-regressive Integrated Moving Average - ARIMA), Multi-Layer Perceptron (MLP) and Radial Basis Function (RBF) neural networks, and fuzzy neural network (FNN). Each group of models has its own strengths and weaknesses. Selecting a model depends on the nature of the data and the application. Weron [4], has provided a comprehensive review on state-of-the-art forecasting models for the electricity price.

In the present work, the forecasting model itself is not the main focus, and therefore, an autoregressive model with exogenous variables (ARX) is implemented, as follows:

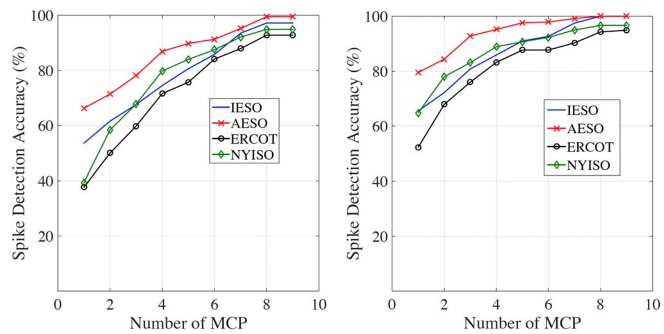
$$y_t = \sum_{i=1}^{p} a_i.y_{t-i} + \sum_{j=1}^{l} b_j.x_{t-j} + \epsilon_t \quad (10)$$

where y_t is the output of the model (the price forecast). The first term on the right-hand side of the formulation shows the auto-regressive part with lagged values of the target variable and the corresponding parameters, i.e., a_i. The second term represents the exogenous variables and their associated parameters, i.e., b_i. The parameters of the model are determined in the training phase. ϵ_t is a normal white noise process with zero mean and variance. The input features of the forecasting model are discussed later in this section. Despite the simplicity of the regression model, it is specifically tailored in the proposed forecasting strategy for the application of the operation of a storage system. In other words, this work aims to build a forecasting strategy to provide informative forecasts such that the operation of a BESS is optimized. In particular, the forecasting strategy should be capable of capturing high prices and severe variations in real-time prices as much as possible. The better price variations can be captured, the better an energy storage control system can adapt its scheduling strategy. Accordingly, high-resolution market information is used along with low-resolution market data in order to enable the model to detect sudden price fluctuations. The low-resolution data includes the hourly market data selected by the feature selection. The high-resolution data could be from any informative market data with higher resolution than one hour that is publicly available.

Our studies showed that market clearing prices carry significant information regarding the price variations in electricity markets. MCPs are set every five minutes in most electricity markets [19]–[22], and the average of twelve MCPs in an hour represents the hourly price. MCPs contain the information about the most recent state of the electricity market. For instance, in the event of a contingency in power systems, e.g., generation outages, the consequent power imbalance is reflected as a change of MCP in supply-demand curve. Here, four North American electricity markets are studied to demonstrate the effectiveness of MCPs in detecting price variations. MCP values over the year 2015 are considered for Ontario's (Independent Electric System Operator-IESO), Alberta's (Alberta Electric System Operator-AESO), Texas' (Electric Reliability Council of Texas-ERCOT), and New York's (New York Independent System Operator-NYISO) electricity markets. It is noted that MCPs are set every minute in Alberta's market and hence, the average of five one-minute MCPs are calculated to form MCPs with five-minute resolution in this study.

Being able to capture high price hours is important in order to increase the profit gained by an energy storage system. High prices, referred to as price spikes, are abnormal high prices that can be distinguished by statistical methods based on the historical data. In [15], a price spike threshold is defined using the mean (μ) and the standard deviation (σ) of historical prices as, $T_{Spike} = \mu + 2\sigma$. Electricity prices greater than $\mathrm{T_{Spike}}$ are considered as spikes. Thresholds are calculated for each month, as the electricity price has seasonal trends. Once the price spikes for each month are distinguished, the capability of MCPs to potentially detect those high price spikes is of interest in this experiment. Defining a threshold for MCPs, $\mathrm{T_{MCP}}$, it is investigated if MCPs can potentially detect a distinguished price spike. To do so, for any hour at which a price spike has been detected, if MCPs in that hour are greater than $\mathrm{T_{MCP}}$, then they are likely to reflect the price spike in that given hour.

Here, two important factors should be considered: 1) the value of $\mathrm{T_{MCP}}$ and 2) the number of MCPs in the given hour. Regarding $\mathrm{T_{MCP}}$, two values of $\mathrm{T_{MCP,1}} = \mu + 2\sigma$ and $\mathrm{T_{MCP,2}} = \mu + \sigma$ are considered. The reason for considering two values is that a MCP with the value greater than $\mathrm{T_{MCP,1}} = \mathrm{T_{Spike}}$ can clearly detect a price spike, while a distinguished price spike can also be potentially detected having a lower MCP value, i.e., $\mathrm{T_{MCP,2}}$. With regard to the number of MCPs, it is evident that the more MCPs are considered during an hour, the higher is the chance of detecting the price spike in the same hour. On the contrary, the cost of considering more MCPs is the time that has been lost during an hour waiting for new MCPs to be released. In other words, the more MCPs are used, the longer the forecasting model has to wait to generate the forecast for a given hour and therefore, such a late forecast may not be as useful for decision-making process. In this experiment, up to 9 MCPs values are considered for price spike detection.



(a) $\mathrm{T_{MCP,1}} = \mu + 2 \times \sigma$ (b) $\mathrm{T_{MCP,2}} = \mu + \sigma$

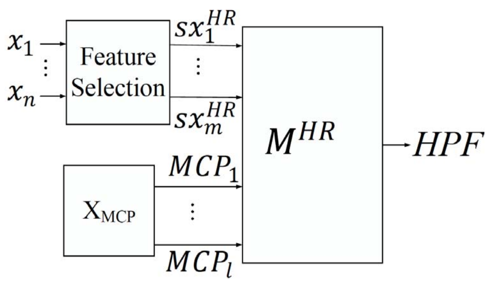
Fig. 3. Potential capability of MCPs in spike detection.

Spike Prediction Accuracy (SPA) criterion measures the ability of correctly predicting spike occurrences, defined as follows:

$$\mathrm{SPA}(\%) = \frac{\text{Number of correctly predicted spikes}}{\text{Number of spikes}} \quad (11)$$

Fig. 3 illustrates the value of SPA versus the number of MCPs considered for detecting price spike in four electricity markets. SPA is calculated for each month of the year 2015, and the average over the year is reported. Fig. 3(a) takes into account $\mathrm{T_{MCP,1}}$ for evaluating the capability of MCPs in spike detection. Evidently, the more MCPs are included, the higher SPA value would be. For instance, the chance of detecting a price spike is over 90% considering nine MCPs. However, even considering the first MCP in each hour results in at least around 40% chance of detecting a spike (ERCOT). Fig. 3(b) shows the same experiment by considering $\mathrm{T_{MCP,2}}$. In fact, the values for MCPs do not need to be as high as $\mathrm{T_{MCP,1}} = \mathrm{T_{Spike}}$ to be able to detect the spike. A high MCP with the value lower than $\mathrm{T_{Spike}}$ could also be capable of detecting a high hourly price. SPA values considering $\mathrm{T_{MCP,2}}$ are expected to be higher than those with $\mathrm{T_{MCP,1}}$ as the threshold. For instance, there is at least 50% chance of detecting a spike using only the first MCP in an hour. Thus, this experiment shows the effectiveness of MCP values as high-resolution market data for capturing high prices. Inspired by this high capability, a forecasting strategy is proposed in this work to include MCP values as high-resolution market data for price prediction.

The proposed forecasting strategy consists of two separate forecasting models, high-resolution model (M^{HR}) and low-resolution model (M^{LR}), with different purposes, illustrated in Fig. 4. The forecasting models provide day-ahead price predictions with an hourly resolution. M^{HR} generates the price forecast for the first step of the forecasting horizon, i.e., 1-hour-ahead forecast. Because of this very short prediction horizon, M^{HR} can take advantage of the high-resolution data available in the current hour ending. Note that Hour Ending (HE) denotes the time when an hour ends, e.g., HE 23 represents the time period 22:00-23:00. Thus, hourly inputs selected by the feature selection, $SX^{HR} = \{sx_1^{HR}, \ldots, sx_m^{HR}\}$,



(a) High-resolution model (M^{HR})

(b) Low-resolution model (M^{LR})

Fig. 4. Structure of the models.

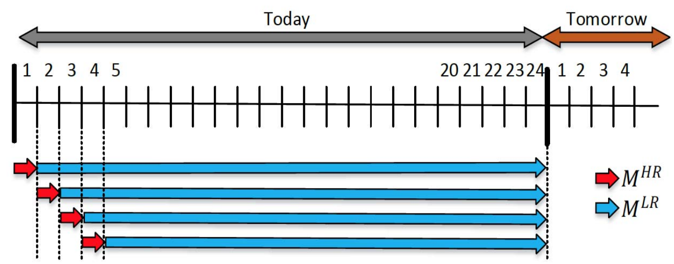


Fig. 5. IRH framework.

as well as a number of MCPs in the current HE are fed to M^{HR} to predict the hourly electricity price of the same HE. A vector of high-resolution inputs is denoted by $X_{MCP} = \{MCP_1, MCP_2, \ldots, MCP_l\}$, where MCP$_1$ corresponds to the MCP that is set for the first five-minute and so forth. l represents the number of MCPs fed to the high-resolution model. As mentioned, M^{HR} provides the price prediction for the first step of the forecasting horizon, i.e., current HE. If the after-the-fact value of the current hour turns out to be a price spike, there is a high chance that any of the included MCPs have an unusual high value leading to this hourly price spike. Hence, the inclusion of these MCPs in the model increases the value of the output of the model, i.e., 1-hour-ahead price spike prediction. Obviously, the more MCP values are included in the model, the higher the chance would be to detect the hourly spike by any of those high MCP values. Once the first step is predicted, the forecast is used as an input in M^{LR}, which provides multi-step-ahead price predictions for the remaining hours of a day recursively. M^{LR} is fed by hourly input features selected by the feature selection, i.e., $SX^{LR} = \{sx_1^{LR}, \ldots, sx_k^{LR}\}$. Fig. 4 illustrates the structure of both models. Hourly Price Forecast (HPF) denotes the output for both models.

The forecasting strategy applies the two forecasting models in the IRH framework. This approach can mitigate the impacts of uncertainties associated to forecasts by updating the forecasts at each forecasting step over the forecasting horizon. In this work, updates of the price forecasts are generated once new observations of the required market data are available, i.e., every hour. Fig. 5 demonstrates how the IRH performs to update the forecasts employing M^{HR} and M^{LR}. As seen, the first set of forecasts is generated in HE 01 for the whole hours of the current day. M^{HR} generates the price forecast for the first hour, i.e., HPF 1, while M^{LR} provides the forecasts for the remaining hours of the day, i.e., HPF 2 to HPF 24. Forecasting horizons for M^{HR} and M^{LR} are shown in red and blue arrows in the figure, respectively. When the first hour is past, both models are used to update the forecasts up to HE 24 of the current day. M^{HR} predicts the price for the second HE, HPF 2, and M^{LR} provides HPF 3 to HPF 24. This process is repeated every hour to update the forecasts up to the last hour of the current day. Updates can go beyond the current day if the optimization problem of the storage system is designed for multiple days. Once the updates are available, they can be applied to the optimization platform of a behind-the-meter BESS to adjust charging/discharging strategies.

III. CASE STUDY

In this section, first, Ontario's electricity market is briefly described as a case study. The proposed price forecasting strategy is then evaluated from both statistical and economic points of view.

A. Ontario's Electricity Market

The Ontario's wholesale electricity market is real-time energy and operating reserves markets. Ontario demand is supplied by local installed capacity within the province and imports from neighboring power systems. The IESO accepts the lowest-cost offers to supply electricity until sufficient power generations are available to meet the demand. MCPs are set every five minutes and the Hourly Ontario Electric Price (HOEP), which is the average of the twelve MCPs in each hour, is published 5 minutes past each hour.

Pre-Dispatch Price (PDP) is the publicly available price prediction published by the IESO. PDPs are updated hourly and provided according to the conventional rolling horizon approach explained in Section II-A. For instance, the 1-hour-ahead PDP for HE 2 is published in HE 1. This is the most recent update on future electricity prices available in this market. A data analysis over a two-year period (January 2014 to December 2015) reveals a significant deviation of 1-hour-ahead PDP from HOEP values, i.e., over 40% error on average. For optimizing the operation of a behind-the-meter storage system, this source of price forecasts may not be helpful because high prices are likely to be missed. For this reason, the proposed price forecasting strategy is applied to Ontario's market to improve the forecasting accuracy, and consequently operation of such participants.

Rodriguez and Anders [36] evaluated several potential input variables for predicting HOEP. The results reveal that none of temperature, predicted shortfalls and predicted transmission constraints can improve the forecast accuracy of hourly electricity prices in Ontario. The feature selection selects hourly

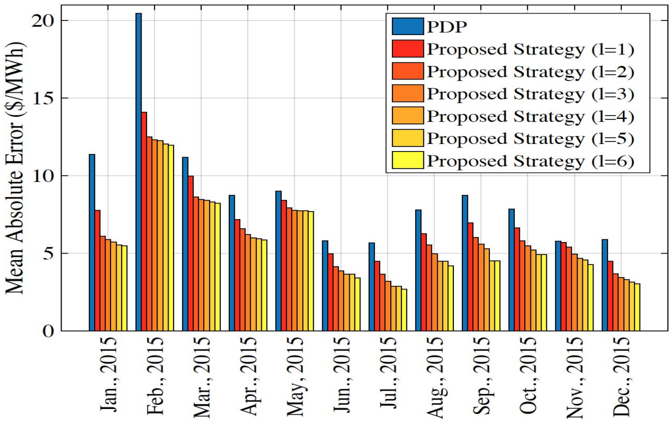


Fig. 6. Forecasting errors in different months.

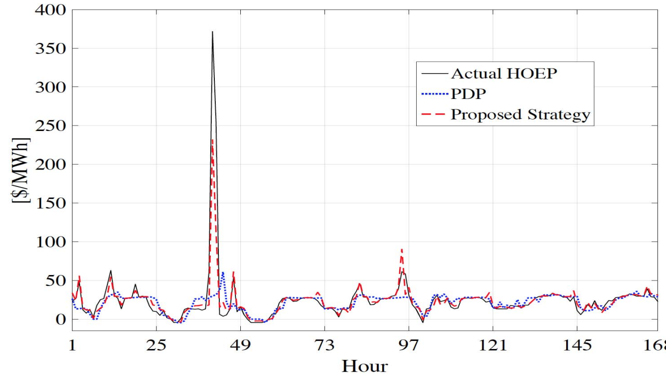


Fig. 7. Price values for the first week of August 2015.

features including PDP, Ontario's demand forecast, supply cushion, lagged values of HOEP as inputs of the two forecasting models. Although PDP is likely to have an error with respect to electricity prices in real time, it is regarded as an initial price forecast for the two forecasting models. In [37], it is stated that price spikes are usually observed when the value of supply cushion is very low, e.g., 10%. Supply cushion is an important publicly available variable that includes the information about forecasts of demand and non-dispatchable generations (e.g., wind and solar generations), and import schedules. The number of MCPs considered for the high-resolution model can be influential on the performance of the forecasting strategy from statistical and economic aspects. This is studied in the following sections. Note that all the required market data for this study are publicly available on IESO's website [38].

B. Statistical Analysis

In this section, the proposed price forecasting strategy is evaluated from the statistical perspective, i.e., the prediction errors of the generated price forecasts are analyzed. For this, the proposed forecasting strategy is tested on publicly available data of Ontario's electricity market in year 2015. An error measure of Mean Absolute Error (MAE) is considered to evaluate the forecasting errors for each month, defined as:

$$\text{MAE (\$/MWh)} = \frac{1}{\text{N}} \sum_{t=1}^{N} \left| \text{P}_{\text{ACT(t)}} - \text{P}_{\text{FOR(t)}} \right| \qquad (12)$$

where N denotes the number of hours in a month. $\text{P}_{\text{ACT}(t)}$ and $\text{P}_{\text{FOR}(t)}$ represent the actual and forecast values of price at hour t, respectively. Monthly forecasting errors in terms of MAE are shown in Fig. 6. In this figure, the proposed strategy is applied with different number of MCPs fed to M^{HR}, e.g., $l = 1, 2, \ldots, 6$. As expected, the more MCP values are used, the lower prediction error is achieved, because of a higher chance of detecting price spikes during the predicting hour. The average of forecasting errors in terms of MAE are respectively \$7.24/MWh, \$6.33/MWh, \$5.98/MWh, \$5.80/MWh, \$5.65/MWh, \$5.56/MWh for $l = 1$ to $l = 6$. For the sake of a comparison, 1-hour-ahead PDP published by the IESO is also shown in blue for all months. The proposed forecasting strategy, even with $l = 1$, results in lower forecasting errors in all test months compared to PDP values. The average of forecasting error for PDP is \$9.02/MWh. Therefore, using only one MCP in the proposed strategy results in 20% improvement in forecast accuracy in comparison with PDP.

TABLE I
CONFUSION MATRICES

No. MCP	Actual	Predicted Spike	Predicted Normal	Predicted Accuracy
l=1	Spike	102	99	50.7%
	Normal	15	8544	99.8%
l=2	Spike	126	75	62.7%
	Normal	8	8551	99.9%
l=3	Spike	133	68	66.2%
	Normal	11	8548	99.9%
l=4	Spike	140	61	69.7%
	Normal	10	8549	99.9%
l=5	Spike	147	54	73.1%
	Normal	8	8551	99.9%
l=6	Spike	153	48	76.1%
	Normal	7	8552	99.9%

Fig. 7 displays hourly actual prices of Ontario's market against 1-hour-ahead forecasts from the proposed strategy and PDP values during the first week of August 2015. As seen, the proposed model can satisfactorily track the price fluctuations and spikes. It should be noted that in many scheduling and operational optimization problems, the exact value of the price peaks is not of interest but the occurrence. This is because the control system of a BESS only requires a precise occurrence of the price spike rather than the magnitude. For instance, there was a price spike of \$371/MWh at HE 41 of this test week, where the proposed model effectively detected this price spike although with the value of \$232/MWh.

To evaluate the performance of the proposed strategy in price spike detection, confusion matrices for different number of MCPs are presented in Table I. Here, the confusion matrix is a 2×2 square matrix indicating two classes of prices, i.e., price spikes and normal prices, for actual and predicted price values. In this experiment, the total 201 spikes are distinguished from normal prices using T_{Spike}, defined in Section II-B, for each month in 2015. For instance, including one MCP ($l = 1$), 102 spikes have been correctly detected, while 99 spikes have been missed. This results in 50.7% accuracy in spike detection shown in the last column of the table. Whereas, only 15 normal prices have incorrectly been detected as price spikes among 8559 normal prices, which leads to 99.8% accuracy in this class. As the number of MCP values used in the model is increased, the spike detection accuracy is enhanced significantly, e.g., 76.1% for $l = 6$, which is expected according to Fig. 3. Although considering more MCPs results in more accurate forecasts, an economic aspect should also be considered to assess whether more MCPs lead to higher values for a storage system.

C. Economic Analysis

In this section, the performance of the proposed forecasting strategy is evaluated from an economic point of view. Thus, the generated price forecasts from the proposed strategy are applied to operation scheduling of a behind-the-meter BESS within a micro-grid facility in Ontario. The micro-grid is designed to operate as backup power during a power outage of the main grid. A 500kW Li-ion BESS should provide emergency power for critical loads in a building of this micro-grid. The remaining capacity of the BESS can be utilized to trade energy with the main grid.

Real-time electricity prices can provide end-users in electricity markets with the opportunity to reduce their electricity costs by strategically responding to prices that varies with different times of the day [39]. The size of the battery is very small compared to the total load of the micro-grid, and thus its operation does not cause any major issues in power flows. Hence, the BESS in the micro-grid is modeled as an individual agent with the objective of maximizing its profit (or the net energy arbitrage saving). This can also be interpreted as reducing the amount of energy purchased from the wholesale market during peak hours when the electricity prices are high. The energy is stored during off-peak hours where prices are low and injected back to micro-gird/grid during peak periods. It should be noted that other factors may impact the operation of a battery system inside a microgrid (e.g., renewable energy fluctuations or load balance). The microgrid operator may consider those factors, in addition to prices, when making charging/discharging decisions.

A significant input for this optimization problem is electricity price forecasts of Ontario's market, which are provided by the proposed forecasting strategy. The formulation for this optimization problem is provided in Algorithm (2). Accordingly, the optimization problem is run every hour up to the last hour of the day. The value of γ is calculated using the number of MCPs (l), $\gamma = (l + 1)/12$.

In Ontario's market, MCP values are published 3 minutes past each five-minute interval, and thus, it should be taken into account in the optimization problem. For instance, given that only the first MCP is used, the forecasting strategy generates the forecasts eight minutes past each hour. Then, the forecasts are used in the BESS optimization platform, and hence, 10 minutes is expected to be past until the scheduling updates of the BESS is available. The emergency load is 150 (kW) and the Depth of Discharge (DOD) for this battery is 70%.

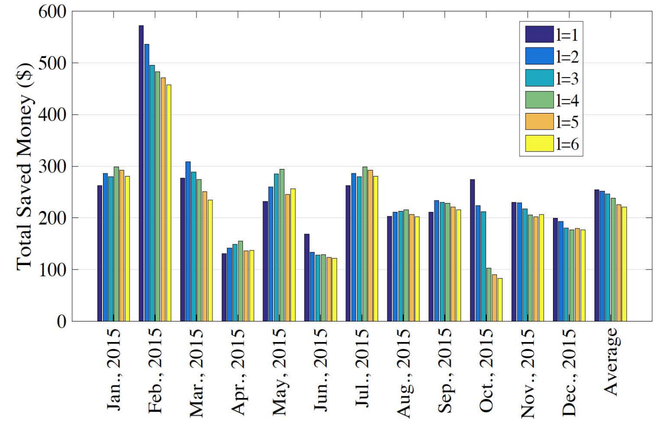


Fig. 8. Total money saved for different number of MCPs.

This leaves 200 (kW) available capacity for the operation of this battery in energy arbitrage. Let assume that the BESS can only have one full charge-discharge cycle per day. Obviously, the more cycles the battery can be operated per day, the more profit can be gained; however, it may decrease the battery's life-time.

In the first experiment, the proposed strategy generates forecasts considering different number of MCPs as inputs ($l = 1, \ldots, 6$). The aforementioned operation scheduling is then applied to calculate the amount of money saved in energy cost. The goal is to find an optimal value for l (or γ). The total saved money for each month is demonstrated in Fig. 8. Interestingly, this figure shows that more accurate forecasts generated with higher number of MCPs do not necessarily lead to higher profits gained by the BESS. The monthly average of saved money shown in the last column for $l = 1$ to $l = 6$ are \$254.1, \$251.7, \$246.5, \$238.5, \$225.9, \$221.1, respectively. The relationship between the forecasting performance and corresponding economic values can be interpreted using the parameter γ. An increase in the value of l means that the forecasting strategy has to wait longer to use more MCPs as inputs of the model. This consequently leads to a higher value for γ in the optimization problem. With larger γ, although the generated forecasts for the current hour are more accurate, a smaller portion of the current hour associated to $(1 - \gamma)$ uses such accurate forecasts for scheduling the BESS. This can be seen form Fig. 2 when the value of γ increases and $(1 - \gamma)$ decreases. According to the same figure, the first portion of the current hour associated to γ has already been scheduled in the previous hour using the forecasts generated in the same hour, which is not the latest forecast update for the current hour. As a result, more accurate predictions from a higher value of l could not be necessarily effective for the BESS from the economic point of view. Thus, the optimal value of γ that results in the highest economic value is $\gamma = 1/6$ ($\gamma = (l + 1)/12$). This means that only the first MCP of the current hour should be fed into the model. Thus, the forecasting strategy can be tuned including only the first MCP in its high-resolution model, because it leads to the highest economic benefits although the forecasting accuracy is slightly compromised.

For the sake of comparisons, four operational strategies are considered based on available price forecasts, generated price forecasts, and the historical data as follows:

1. PDP Scheduling: This scheme considers PDP values with hourly updates published by the IESO as available price forecasts. PDPs are applied to the optimization problem in accordance with the formulation presented in Algorithm (1).

2. Proposed Strategy Scheduling: In this scheme, price forecasts with an hourly updates generated by the proposed forecasting strategy (with $l = 1$) are fed into the optimization platform (Algorithm (2)). The operational schedules of the BESS are updated every hour as the updates of price forecasts are generated. The scheduling horizon is kept the same, up to HE 24 of the day.

3. Ad-hoc Strategies: It is considered that the BESS has an unchanging operation strategy rather than finding the optimal operation scheduling using price forecasts fed into the optimization problem. Two different ad-hoc strategies for the operation of the BESS are considered as below:

3.a. Ad-hoc #1: Weighted averages of electricity prices from 2002 to 2013 are calculated for each month separately. Hours with the lowest and the highest electricity prices are correspondingly considered for charging and discharging. For instance, the charging hour is HE 4 and the discharging hours are HE 18 and HE 19 for the month October.

3.b. Ad-hoc #2: Charging and discharging decisions for the current day are made based on the profile of electricity price in the previous day. Hours with the lowest and the highest electricity prices in the previous day are considered for charging and discharging in the current day, respectively.

In this experiment, the potential profit that could be achieved by operating the BESS based on perfect forecasts is calculated. Having the perfect price forecasts, the micro-grid could have potentially saved \$4,688 in total in energy cost by operating the BESS over the year 2015. Applying the proposed forecasting strategy into the developed optimization platform, 62% of the potential saving could be captured (total \$2,937). Scheduling based on available PDP could only capture 43% of this potential (total \$2,019). Hence, applying the proposed forecasting strategy could result in around 50% improvement over the use of available PDP in the profit gained by operating the BESS. The results indicate that the total profits based on two ad-hoc strategies are only \$1,407 and \$1,120, respectively. Fig. 9 illustrates the total amount of money saved by operating the BESS for each month by adopting different strategies compared with the potential profit. It is also noted that the size of the battery can affect the total energy arbitrage saving, while this is not the focus of the present paper.

An example of the operation of the BESS on October 2^{nd}, 2015 is presented for different strategies in Fig. 10. Observe that the proposed strategy could effectively detect the price spike at HE 10, while PDP failed to do so. Therefore, the BESS is scheduled to discharge when the price is at its highest value. For both ad-hoc strategies, equal discharging rates are considered. It is also noted that the charging hour for ad-hoc #2 is at HE 6, which has been covered with the one for PDP scheduling. The amount of money saved by applying the proposed strategy is \$140.58 for this particular day. This

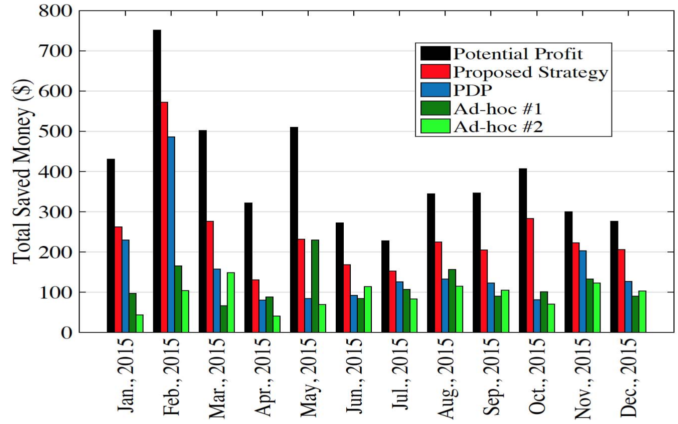


Fig. 9. Monthly total money saved by operating the BESS.

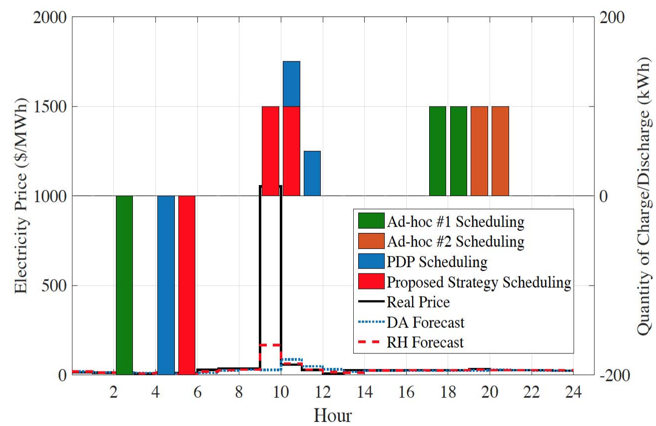


Fig. 10. BESS schedules based on different strategies for October 2nd, 2015.

amount is 52 times greater than the corresponding amount obtained from scheduling based on PDP values, i.e., \$2.67. The amounts of saved money are respectively \$3.03 and \$3.59 for the first and the second adopted ad-hoc strategies. This figure clearly depicts the effect of accurately detecting price spike occurrences in the operation of a storage system. In this practical application, it is of high importance to predict the exact timing of price spikes in order to discharge the battery accordingly. It is noted that the higher the difference between electricity prices at charging and discharging hours, the higher would be the saved money by operating the BESS.

In the last experiment, we evaluate the potential for gaining higher economic values by increasing the number of cycles in the operation of the BESS. It is noted that frequent and deep cycles accelerate cyclic aging and consequently reduce the life of the battery. For any type of battery, number of cycles is usually defined as a function of depth of discharge, which can be obtained by a fitting technique using detailed experimental data provided by manufacturers [40]. However, since we do not have access to such data, we assume a constant depth of charge for analyzing the economic impact of higher number of cycles. Fig. 11 shows the total saved money obtained by increasing the number of cycles. As expected, the higher the number of cycles, the higher the potential profit would be if perfect price forecasts were available. For instance, the potential profit increases from \$4,688 with one cycle per day

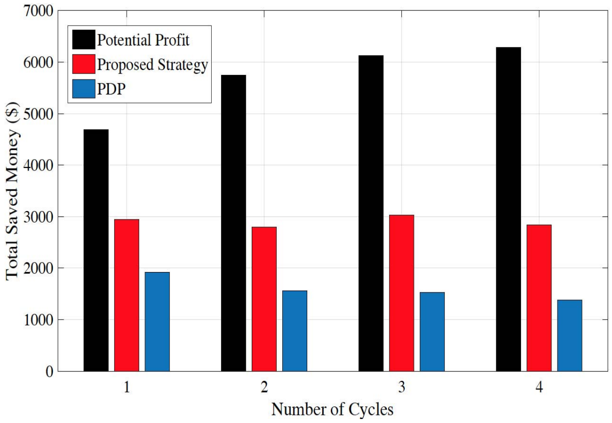


Fig. 11. Economic effect of number of cycles for the BESS.

to \$6,283 with 4 cycles per day. It is observed from this figure that the total saved money starts to saturate once the battery is operated with more than two cycles per day. This is because the opportunity for arbitrage is limited during a day depending of the price profile, and therefore, an increase in the number of cycles cannot necessarily lead to a significant increase in the total saved money. Higher cycles for the proposed strategy, on the other hand, does not result in any noticeable changes in the total saved money. This can be justified by the effect of forecasting errors along with the limited arbitrage capacity during a day. Finally, high forecasting errors associated with PDP even leads to lower total saved money as the number of cycles increases. As a result, this BTM BESS is economically better off being operated with one cycle per day using the proposed strategy.

IV. Conclusion

In this paper, a forecasting strategy is proposed to provide accurate price forecasts for operation of behind-the-meter storage systems. This strategy includes two separate forecasting models to take advantage of high-resolution market data along with hourly data in order to capture price spikes as much as possible. The proposed inta-hour rolling horizon framework is applied to update the forecasts on an hourly basis. From statistical analysis, the proposed strategy results in 20% improvement in forecast accuracy compared to available PDPs, and has a high capability of detecting price spikes. The generated forecasts are fed to an optimization platform for operation scheduling of the BESS within a micro-grid facility. It is concluded that the BESS can bring more economic values using the price predictions generated by the proposed forecasting strategy, 62% of the potential saving, in comparison with a number of other strategies, e.g., 43% of the potential saving using PDPs.

Appendix

Mutual-Information Feature Selection

In this paper, we implemented the feature selection technique proposed in [33] for electricity price prediction. Constructed on the information theoretic criterion of mutual information, this method selects the most informative input features for the forecast process by filtering out the irrelevant and redundant candidate features through two stages. In this work, a vector of candidate features is formed including market clearing prices, hourly Ontario electric prices, pre-dispatch prices, supply cushion and Ontario demand forecast.

In the first stage, called irrelevancy filter, mutual information between each candidate feature, i.e., $f_i(t)$, and the target variable is calculated. The higher value of mutual information for $f_i(t)$ means the more common information content of this feature with the target variable. Defining a relevancy threshold denoted by T_{Rel}, the candidate inputs with calculated mutual information value greater than T_{Rel} are considered as the relevant features of the forecast process. These features are retained for the second stage, while other candidate features whose mutual information values are lower than T_{Rel} are considered as irrelevant features, which are filtered out.

In the second stage, redundant features among the candidate features selected by the relevancy filter are detected and filtered out. This stage is called redundancy filter. Two selected candidates from the first stage, e.g., $f_k(t)$ and $f_l(t)$, with high value of mutual information have more common information, meaning high level of redundancy. Therefore, the redundancy of each selected feature $f_k(t)$ with the other candidate inputs is first calculated. Defining a redundancy threshold denoted by T_{Red}, if the measured redundancy becomes greater than T_{Red}, $f_k(t)$ is then considered as a redundant candidate feature. Thus, between this candidate and the feature that has the maximum redundancy with $f_k(t)$, one with lower relevancy should be filtered out [41].

The selected candidate features in redundancy filter are considered as the inputs of the price forecasting model. It is noted that cross validation technique is used for fine-tuning the values of the thresholds T_{Rel} and T_{Red}. Since this method is not the focus of this paper, the interested reader can refer to [42] for detailed formulation of mutual information criterion.

Acknowledgment

The authors would like to thank NRGStream Inc. for providing a complimentary license to use their market data collection platform. Finally, they want to thank Dr. Mostafa Kazemi for his comments on operation of energy storage systems.

References

[1] (2016). *IEEE Smart Grid Newsletter*. [Online]. Available: http://smartgrid.ieee.org/newsletters/january-2016/emergence-of-the-behind-the-meter-energy-storage-market

[2] Y.-J. Kim, G. Del-Rosario-Calaf, and L. K. Norford, "Analysis and experimental implementation of grid frequency regulation using behind-the-meter batteries compensating for fast load demand variations," *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 484–498, Jan. 2017.

[3] J. Neubauer and M. Simpson, "Deployment of behind-the-meter energy storage for demand charge reduction," Nat. Renew. Energy Lab., U.S. Dept. Energy, Washington, DC, USA, Tech. Rep. NREL/TP-5400-63162, Jan. 2015. [Online]. Available: http://www.nrel.gov/docs/fy15osti/63162.pdf

[4] R. Weron, "Electricity price forecasting: A review of the state-of-the-art with a look into the future," *Int. J. Forecasting*, vol. 30, no. 4, pp. 1030–1081, Oct./Dec. 2014.

[5] P. Mandal, A. U. Haque, J. Meng, A. K. Srivastava, and R. Martinez, "A novel hybrid approach using wavelet, firefly algorithm, and fuzzy artmap for day-ahead electricity price forecasting," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1041–1051, May 2013.

[6] C. Wan, Z. Xu, Y. Wang, Z. Y. Dong, and K. P. Wong, “A hybrid approach for probabilistic forecasting of electricity price,” *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 463–470, Jan. 2014.

[7] T. Hong *et al.*, “Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond,” *Int. J. Forecasting*, vol. 32, no. 3, pp. 896–913, Jul./Sep. 2016.

[8] C. Wan, M. Niu, Y. Song, and Z. Xu, “Pareto optimal prediction intervals of electricity price,” *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 817–819, Jan. 2017.

[9] H. Zareipour, A. Janjani, H. Leung, A. Motamedi, and A. Schellenberg, “Classification of future electricity market prices,” *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 165–173, Feb. 2011.

[10] D. Huang, H. Zareipour, W. D. Rosehart, and N. Amjady, “Data mining for electricity price classification and the application to demand-side management,” *IEEE Trans. Smart Grid*, vol. 3, no. 2, pp. 808–817, Jun. 2012.

[11] H. C. Wu, S. C. Chan, K. M. Tsui, and Y. Hou, “A new recursive dynamic factor analysis for point and interval forecast of electricity price,” *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2352–2365, Aug. 2013.

[12] N. Amjady and F. Keynia, “A new prediction strategy for price spike forecasting of day-ahead electricity markets,” *Appl. Soft Comput.*, vol. 11, no. 6, pp. 4246–4256, Apr. 2011.

[13] T. M. Christensen, A. S. Hurn, and K. A. Lindsay, “Forecasting spikes in electricity prices,” *Int. J. Forecast.*, vol. 28, no. 2, pp. 400–411, Apr./Jun. 2012.

[14] J. H. Zhao, Z. Y. Dong, X. Li, and K. P. Wong, “A framework for electricity price spike analysis with advanced data mining methods,” *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 376–385, Feb. 2007.

[15] X. Lu, Z. Y. Dong, and X. Li, “Electricity market price spike forecast with data mining techniques,” *Elect. Power Syst. Res.*, vol. 73, no. 1, pp. 19–29, Jan. 2005.

[16] A. Fragkioudaki, A. Marinakis, and R. Cherkaoui, “Forecasting price spikes in European day-ahead electricity markets using decision trees,” in *Proc. 12th Int. Conf. Eur. Energy Market (EEM)*, Lisbon, Portugal, May 2015, pp. 1–5.

[17] S. K. Aggarwal, L. M. Saini, and A. Kumar, “Electricity price forecasting in deregulated markets: A review and evaluation,” *Int. J. Elect. Power Energy Syst.*, vol. 31, no. 1, pp. 13–22, Jan. 2009.

[18] K. Maciejowska and R. Weron, “Short-and mid-term forecasting of baseload electricity prices in the U.K.: The impact of intra-day price relationships and market fundamentals,” *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 994–1005, Mar. 2016.

[19] *Independent Electric System Operator (IESO)—Pricing*. Accessed on Jul. 5, 2017. [Online]. Available: http://www.ieso.ca/Pages/Ontario’s-Power-System/Electricity-Pricing-in-Ontario/How-Wholesale-Electricity-Price-is-Determined.aspx

[20] *California Independent System Operator (CAISO)—Pricing*. Accessed on Jul. 5, 2017. [Online]. Available: http://www.caiso.com/market/Pages/MarketProcesses.aspx

[21] *Electric Reliability Council of Texas (ERCOT)—Pricing*. Accessed on Jul. 5, 2017. [Online]. Available: http://www.ercot.com/mktinfo/prices

[22] *New York Independent System Operator (NYISO)—Pricing*. Accessed on Jul. 5, 2017. [Online]. Available: http://www.nyiso.com

[23] *Alberta Electric System Operator (AESO)—Pricing*. Accessed on Jul. 5, 2017. [Online]. Available: http://www.aeso.ca/rulesprocedures/18592.html

[24] H. Zareipour, C. A. Canizares, and K. Bhattacharya, “Economic impact of electricity market price forecasting errors: A demand-side analysis,” *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 254–262, Feb. 2010.

[25] B. Mohammadi-Ivatloo, H. Zareipour, M. Ehsan, and N. Amjady, “Economic impact of price forecasting inaccuracies on self-scheduling of generation companies,” *Elect. Power Syst. Res.*, vol. 81, no. 2, pp. 617–624, Feb. 2011.

[26] E. Nasrolahpour, S. J. Kazempour, H. Zareipour, and W. D. Rosehart, “Strategic sizing of energy storage facilities in electricity markets,” *IEEE Trans. Sustain. Energy*, vol. 7, no. 4, pp. 1462–1472, Oct. 2016.

[27] M. Kazemi, H. Zareipour, M. Ehsan, and W. D. Rosehart, “A robust linear approach for offering strategy of a hybrid electric energy company,” *IEEE Trans. Power Syst.*, vol. 32, no. 3, pp. 1949–1959, May 2017.

[28] S. Shafiee, P. Zamani-Dehkordi, H. Zareipour, and A. M. Knight, “Economic assessment of a price-maker energy storage facility in the Alberta electricity market,” *Energy*, vol. 111, pp. 537–547, Sep. 2016.

[29] R. Palma-Behnke *et al.*, “A microgrid energy management system based on the rolling horizon strategy,” *IEEE Trans. Smart Grid*, vol. 4, no. 2, pp. 996–1006, Jun. 2013.

[30] J. Silvente, G. M. Kopanos, E. N. Pistikopoulos, and A. Espuña, “A rolling horizon optimization framework for the simultaneous energy supply and demand planning in microgrids,” *Appl. Energy*, vol. 155, pp. 485–501, Oct. 2015.

[31] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*. San Francisco, CA, USA: Morgan Kaufmann, 2006.

[32] K. J. Cios, W. Pedrycz, R. W. Swiniarski, and L. A. Kurgan, *Data Mining: A Knowledge Discovery Approach*. New York, NY, USA: Springer, 2007.

[33] N. Amjady, F. Keynia, and H. Zareipour, “Wind power prediction by a new forecast engine composed of modified hybrid neural network and enhanced particle swarm optimization,” *IEEE Trans. Sustain. Energy*, vol. 2, no. 3, pp. 265–276, Jul. 2011.

[34] M. Alamaniotis, D. Bargiotas, N. G. Bourbakis, and L. H. Tsoukalas, “Genetic optimal regression of relevance vector machines for electricity pricing signal forecasting in smart grids,” *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 2997–3005, Nov. 2015.

[35] N. M. Pindoriya, S. N. Singh, and S. K. Singh, “An adaptive wavelet neural network-based energy price forecasting in electricity markets,” *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1423–1432, Aug. 2008.

[36] C. P. Rodriguez and G. J. Anders, “Energy price forecasting in the Ontario competitive power system market,” *IEEE Trans. Power Syst.*, vol. 19, no. 1, pp. 366–374, Feb. 2004.

[37] H. Zareipour, C. A. Cañizares, K. Bhattacharya, and J. Thomson, “Application of public-domain market information to forecast Ontario’s wholesale electricity prices,” *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1707–1717, Nov. 2006.

[38] (2016). *Independent Electric System Operator*. [Online]. Available: http://www.ieso.ca/Pages/Power-Data/default.aspx

[39] A.-H. Mohsenian-Rad and A. Leon-Garcia, “Optimal residential load control with price prediction in real-time electricity pricing environments,” *IEEE Trans. Smart Grid*, vol. 1, no. 2, pp. 120–133, Sep. 2010.

[40] G. He, Q. Chen, C. Kang, P. Pinson, and Q. Xia, “Optimal bidding strategy of battery storage in power markets considering performance-based regulation and battery cycle life,” *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2359–2367, Sep. 2016.

[41] H. Chitsaz, H. Shaker, H. Zareipour, D. Wood, and N. Amjady, “Short-term electricity load forecasting of buildings in microgrids,” *Energy Build.*, vol. 99, pp. 50–60, Jul. 2015.

[42] N. Amjady and F. Keynia, “Day-ahead price forecasting of electricity markets by mutual information technique and cascaded neuro-evolutionary algorithm,” *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 306–318, Feb. 2009.

Hamed Chitsaz (S’14) received the B.Sc. and M.Sc. degrees in electrical engineering from Semnan University, Semnan, Iran, in 2009 and 2012, respectively. He is currently pursuing the Ph.D. degree with the University of Calgary, Calgary, AB, Canada. His research interests include energy forecasting, microgrids, and electricity markets.

Payam Zamani-Dehkordi (S’11) received the B.Sc. degree from the University of Tehran, Tehran, Iran, in 2010, the M.Sc. degree from the Sharif University of Technology, Tehran, in 2012, and the Ph.D. degree from the University of Calgary, Calgary, AB, Canada, in 2017, all in electrical engineering. He is currently an Electrical Engineer with BBA, Calgary. His research interests include data mining applications in electricity markets, renewable energy, and energy storage systems.

Hamidreza Zareipour (S’03–M’07–SM’09) received the Ph.D. degree in electrical engineering from the University of Waterloo, Waterloo, ON, Canada, in 2006. He is currently a Professor with the Department of Electrical and Computer Engineering, University of Calgary, Calgary, AB, Canada. His research focuses on economics, planning, and management of power and energy systems in a deregulated electricity market environment.

Palak P. Parikh (S’08–M’11) received the bachelor’s and master’s degrees in electrical engineering from Sardar Patel University, India, and the Ph.D. degree from the University of Western Ontario, Canada. She has expertise in the area of power protection, distribution automation, and microgrid solutions. She has authored over 20 international journal and conference articles. She is currently with GE Digital Energy in product research and development application engineer group, and focusing on various product designs in the area of protection, distribution automation, and microgrid.