Learning Dynamical Demand Response Model in Real-Time Pricing Program

Xu, H.; Sun, H.; Nikovski, D.N.; Shoichi, K.; Mori, K.

Abstract

Price responsiveness is a major feature of end use customers (EUCs) that participate in demand response (DR) programs, and has been conventionally modeled with static demand functions, which take the electricity price as the input and the aggregate energy consumption as the output. This, however, neglects the inherent temporal correlation of the EUC behaviors, and may result in large errors when predicting the actual responses of EUCs in real-time pricing (RTP) programs. In this paper, we propose a dynamical DR model so as to capture the temporal behavior of the EUCs. The states in the proposed dynamical DR model can be explicitly chosen, in which case the model can be represented by a linear function or a multi-layer feedforward neural network, or implicitly chosen, in which case the model can be represented by a recurrent neural network or a long short-term memory unit network. In both cases, the dynamical DR model can be learned from historical price and energy consumption data. Numerical simulation illustrated how the states are chosen and also showed the proposed dynamical DR model significantly outperforms the static ones.

Learning Dynamical Demand Response Model in Real-Time Pricing Program

Hanchen Xu[*†], Hongbo Sun[†], Daniel Nikovski[†], Kitamura Shoichi [‡], and Kazuyuki Mori [‡]

[*]Department of Electrical and Computer Engineering, University of Illinois, Urbana, IL 61801, USA
[†]Mitsubishi Electric Research Laboratories, 201 Broadway, Cambridge, MA 02139, USA
[‡]Advanced Technology R&D Center, Mitsubishi Electric Corporation, Hyogo 661-8661, Japan
Email: hxu45@illinois.edu, hongbosun@merl.com, nikovski@merl.com

Abstract—Price responsiveness is a major feature of end use customers (EUCs) that participate in demand response (DR) programs, and has been conventionally modeled with static demand functions, which take the electricity price as the input and the aggregate energy consumption as the output. This, however, neglects the inherent temporal correlation of the EUC behaviors, and may result in large errors when predicting the actual responses of EUCs in real-time pricing (RTP) programs. In this paper, we propose a dynamical DR model so as to capture the temporal behavior of the EUCs. The states in the proposed dynamical DR model can be explicitly chosen, in which case the model can be represented by a linear function or a multi-layer feedforward neural network, or implicitly chosen, in which case the model can be represented by a recurrent neural network or a long short-term memory unit network. In both cases, the dynamical DR model can be learned from historical price and energy consumption data. Numerical simulation illustrated how the states are chosen and also showed the proposed dynamical DR model significantly outperforms the static ones.

I. Introduction

Demand response (DR) has proven to be a successful paradigm that emerges in the context of smart grid [1]–[3]. The DR enables the participation of demand-side resources, including commercial buildings [4], [5], industrial loads [6], electric vehicles [7], in maintaining the supply-demand-balance in power systems through some economic means such as incentives or price signals [1]. A variety of DR programs have been developed so far on both the wholesale market level and the retail market level. On the wholesale level, qualified demand-side resources can offer in almost all markets including capacity market, energy market, as well as ancillary markets, and will be cleared by the independent system operator (ISO) in the same way as the conventional generating resources. DR programs on the retail level, however, have more variety than those on the wholesale level. Yet, the retail DR programs can, by-and-large, be categorized into two classes—incentive based and price based. The former includes, e.g, direct load control and some interruptible programs, while the latter includes the real-time pricing (RTP), time-of-use-pricing, critical peaking pricing.

In price-based DR programs, the load serving entity (LSE)—the operator of the retail market—needs to design effective pricing schemes such that certain objective such as minimizing load fluctuation [8], maximizing profit [9] or utility [10], can be achieved. Particularly in RTP, the LSE sends out a price signal for every time interval, and the end use customers (EUCs) then respond to the price signal by adjusting their energy consumptions. It is therefore critical for the LSE to know the price response characteristics of the EUCs in order to make informed decisions on the price signal sent to the EUCs. Conventionally, price response characteristics are modeled using a static function such as linear, exponential, logarithmic, and potential functions [11]. These static functions, however, neglects the inherent temporal correlation of the EUC behaviors, and as we will see in the later part of the paper, may result in large errors when predicting the actual responses of EUCs in RTP programs.

To explicitly capture the temporal behavior of the EUCs, we propose a dynamical DR model which has states that evolve with time. The states in the proposed dynamical DR model keep necessary information from the past time interval, and allows more accurate prediction of the EUCs' response. These states can be explicitly chosen, in which case the model can be represented by a linear function or a multi-layer feedforward neural network (FNN), or implicitly chosen, in which case the model can be represented by a recurrent neural network (RNN) or a long short-term memory (LSTM) unit network. The proposed dynamical DR model can be learned from historical price and energy consumption data. In particular, the states can be determined empirically from the historical data. We will also show through numerical simulation that the proposed dynamical DR model significantly outperforms the conventional static models.

The remaining part of this paper is organized as follows. Section II reviews decision process of EUCs and proposes a dynamical DR model that is based on available data. The neural network representations of the dynamical DR model are developed in Section III, and results from numerical simulation are presented in Section IV. Some concluding remarks are made in Section V.

II. Demand Response Model

In this section, we first briefly review how a myopic EUC may determine its energy consumption at a given electricity price. In light of the deficiencies in static DR models, we propose a dynamical DR model in the end. Throughout this paper, we use a subscript t to denote the time interval, and assume one day is decomposed into T segments.

A. Decision Process of End Use Customers

In a RTP program, the LSE sends out a energy price for time interval t, denoted by p_t, before the beginning of that time interval. Each EUC then responds to the price signal by optimizing its energy consumption in a way that maximizes its overall benefit or equivalently, minimizes its overall cost. While there exist many detailed models for EUCs', a generic model that is agnostic to the underlying components is to model the *energy demand* and *energy consumption* separately, where the energy demand is the needed energy while the energy consumption is the actual consumed energy [10], [12].

Let $\mathcal{K} = \{1, \cdots, K\}$ denote the set of EUCs served by the LSE. Let $e_t^{k,d}$ and $e_t^{k,c}$ denote the energy demand and energy consumption of EUC $k \in \mathcal{K}$ at time interval t, respectively. A myopic EUC finds its optimal energy consumption by solving the following utility maximization problem [10]:

$$\underset{e_t^{k,c}}{\text{maximize}} \ \beta_t^k(e_t^{k,d}, e_t^{k,c}) - p_t e_t^{k,c}, \tag{1a}$$

subject to

$$e_{t+1}^{k,d} = \alpha_t^k(e_t^{k,d} - e_t^{k,c}) + \xi_{t+1}^k, \tag{1b}$$

$$e_t^{k,c} \in \mathcal{E}_t^{k,c}, \tag{1c}$$

where $\beta_t^k(\cdot)$ is the benefit function, which gives the benefit of the EUC at certain energy demand and energy consumption, α_t^k is the backlog rate that represents the percentage of unmet energy demand that is carried over to the next time interval, ξ_{t+1}^k is a random variable that models the new demand, $\mathcal{E}_t^{k,c}$ is the feasible set of energy consumption.

From the perspective of the LSE, the collective behavior of all EUCs are of more interest since its total profit demands on the aggregate energy consumption. The DR model that is of interest to the LSE is essentially a mapping from the price to the aggregate energy consumption as follow:

$$p_t \longmapsto e_t^c := \sum_{k \in \mathcal{K}} e_t^{k,c}. \tag{2}$$

B. Dynamical Demand Response Model

The mapping from price to energy consumption is conventionally modeled using a static function such as linear, exponential, logarithmic, and potential functions [11], which can be conceptually expressed as:

$$e_t^c = \psi(p_t). \tag{3}$$

These static functions generally work well on the whole level since the price response characteristics of DR resources in the wholesale market are mainly determined by their offers. Once the energy demand are cleared for each time interval, the DR resources will be controlled such that their energy consumptions follow the schedule, since otherwise they will incur costs. Therefore, given a certain price, the energy consumption can be determined from the relative static offer stacks directly, i.e., the mapping is relatively static.

However, in a RTP problem, the EUCs do not have the responsibility to follow any energy consumption profile. Indeed, it is obvious from (1) that given a price, the energy consumption of an EUC will depend on its current energy demand, which itself depends on the previous energy consumption. Therefore, the mapping in (2) is dynamical.

In light of this observation, we propose a dynamical model for the DR model for the EUCs in a RTP program as follows:

$$e_t^c = \varphi(\boldsymbol{s}_t, p_t), \tag{4}$$

where $\boldsymbol{s}_t$ is a state vector that captures all the factors besides the price that impact the aggregate energy consumption, and evolves along time given p_t.

While it is difficult to identify the "perfect" state vector, we may still be able to construct one that is good enough in the sense it gives good prediction accuracy of the aggregate demand. From a practical view, the state vector can be compromised of elements from the set of information that is available to the LSE at time interval t, denoted by $\mathcal{I}_t = \{(p_\tau, e_\tau^c) \ \forall \tau < t, p_t, t\}$. There are two potential approaches to construct such a state vector, a *direct* approach, and an *indirect* (implicit) approach. In the direct approach, we select the state vector to be

$$\boldsymbol{s}_t = (p_{t-n}, e_{t-n}^c, \cdots, p_{t-1}, e_{t-1}^c, t \bmod T), \tag{5}$$

where n—referred to as the *order* of the DR model—is a parameter that can be determined from the historical data. We adopt the convention that $\boldsymbol{s}_t = \{t \bmod T\}$ when $n = 0$. In the indirect approach, the state vector in (5) is further transformed in a series of nonlinear operations. With the state vector and the price information, we can predict the aggregate energy consumption using supervised learning techniques.

III. NEURAL NETWORK REPRESENTATION OF DYNAMICAL DEMAND RESPONSE MODEL

In this section, we develop neural network representations for the dynamical DR model. A multi-layer FNN is applied in the direct approach, and a multi-layer RNN is applied in the indirect approach.

A. Direct Approach Using Feedforward Neural Network

When using the direct approach, the state vector is manually selected from the set of available information. Therefore, the input vector and the output of the DR model are both readily available. We can thus fit the dynamical DR model in (4) using, for example, a linear function as follows:

$$e_t^c = \boldsymbol{w}^\top[\boldsymbol{s}_t^\top, p_t] + b, \tag{6}$$

where $\boldsymbol{w}$ is a weight vector, and b is the bias, the superscript $\top$ denotes the transpose of a vector or matrix.

Also, we may represent φ with a nonlinear function such as a multi-layer FNN, which consists of one input layer, L hidden layer, and one output layer as illustrated in the left part in Fig. 1. The hidden layer l takes an input vector $\boldsymbol{x}_t^{[l]}$, and computes a (hidden) output vector $\boldsymbol{h}_t^{[l]}$ according to

$$\boldsymbol{h}_t^{[l]} = \text{relu}(\boldsymbol{W}^{[l]}\boldsymbol{x}_t^{[l]} + \boldsymbol{b}^{[l]}), \tag{7}$$
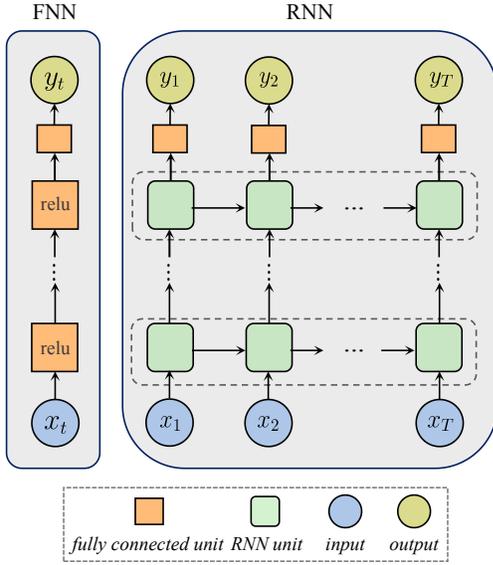
Fig. 1. Neural networks for dynamical demand response model.

Fig. 2. Structure of an LSTM unit [15].

where relu$(\cdot)$ is a rectified linear unit function that is applied element-wise, $\boldsymbol{W}^{[l]}$ is a weight matrix, and $\boldsymbol{b}^{[l]}$ is a bias vector. Note that the output vector of one hidden layer is the input vector for the next hidden layer, i.e., $\boldsymbol{x}^{[l+1]} = \boldsymbol{h}^{[l]}$, except the last hidden layer, the output of which is mapped to the output through a fully connected unit as follows:

$$y_t = \boldsymbol{W}\boldsymbol{h}^{[L]} + \boldsymbol{b}, \tag{8}$$

where $\boldsymbol{W}$ is a weight matrix, and $\boldsymbol{b}$ is a bias vector. Note that $\boldsymbol{x}_t^{[1]} = (\boldsymbol{s}_t, p_t)$. The multi-layer FNN can be trained using back-propagation such that the mean squared error between the predicted output y_t and the true value e_t^c is minimized, i.e., by minimizing the following loss function:

$$\ell = \frac{1}{m} \sum_{l=1}^{m} (y_t - e_t^c)^2. \tag{9}$$

B. Indirect Approach Using Recurrent Neural Network

Alternatively, the states can be implicitly constructed within the neural network, which leads to RNNs [13]. The right part in Fig. 1 illustrate a multi-layer RNN with one input layer, L hidden layers, and one output layer. The hidden state of the RNN unit in layer l, denoted by $\boldsymbol{h}_t^{[l]}$, is the input for RNN unit in the next layer as well as the input for itself at the next time step, as indicated by the arrows in Fig.1. This RNN takes a sequence $\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_T\}$ as the input, and outputs a sequence $\{y_1, \cdots, y_T\}$. Meanwhile, L sequences of hidden states $\{\boldsymbol{h}_t^{[l]}\}$, $l = 1, \cdots, L$, are generated along the trajectory, based on the following equations:

$$\boldsymbol{h}_t^{[l]} = \tanh(\boldsymbol{W}_h^{[l]} \boldsymbol{h}_{t-1}^{[l]} + \boldsymbol{W}_x^{[l]} \boldsymbol{x}_t^{[l]} + \boldsymbol{b}^{[l]}), \tag{10}$$

where $\tanh(\cdot)$ is applied element-wise, $\boldsymbol{W}_h$ and $\boldsymbol{W}_x$ are weight matrices, $\boldsymbol{b}$ is a bias vector. Note that $\boldsymbol{h}_{-1}^{[l]}$ are initialized to zeros, $\boldsymbol{x}_t^{[l]} = \boldsymbol{h}_t^{[l-1]}$ for $l = 2, \cdots, L$, and

$\boldsymbol{x}_t^{[1]} = (\boldsymbol{s}_t, p_t)$. The hidden states in RNN are dynamical since their values also depend on their previous values, while those in the FNN are static since their values purely depend on the inputs. The output of the last hidden state vector is mapped to the output through a fully connected unit as in the case of multi-layer FNN. The RNN can be trained by minimizing the same loss function as in (9) using backpropagation through time technique (see [14]). The input vector only has to include the most-recent information, i.e., when RNN is used, $n = 1$ and $\mathrm{mod}\, s_t = \{p_{t-1}, e_{t-1}^c, t \bmod T\}$.

The key difference between the RNN and FNN in representing the dynamical DR model is that the FNN captures the temporal impacts by explicitly specifying a set of historical data as inputs, while the RNN keeps the temporal impacts by implicitly computing a dynamical hidden state.

One of the deficiency of the basic RNN unit is the lack of ability to model long-term dependencies. As a significant improvement over the basic RNN unit, the LSTM is proposed by Hochreiter and Schmidhuber in [16]. The structure of an LSTM unit is illustrated in Fig. 2, in which $\sigma(\cdot)$ indicates a sigmoid function. For the purpose of simplicity, we drop the superscript that indicates the layer and focus on structure inside one LSTM unit. The LSTM unit introduces a new hidden state vector $\boldsymbol{C}_t$, which is used to keep long-term memories. The LSTM unit works as follows. First, a forget gate vector $\boldsymbol{f}_t$, an information gate vector $\boldsymbol{i}_t$, and an output gate vector $\boldsymbol{o}_t$ is computed from previous hidden state $\boldsymbol{h}_{t-1}$ and new input vector $\boldsymbol{x}_t$ as follows:

$$\boldsymbol{f}_t = \sigma(\boldsymbol{W}_{fh}\boldsymbol{h}_{t-1} + \boldsymbol{W}_{fx}\boldsymbol{x}_t + \boldsymbol{b}_f), \tag{11}$$
$$\boldsymbol{i}_t = \sigma(\boldsymbol{W}_{ih}\boldsymbol{h}_{t-1} + \boldsymbol{W}_{ix}\boldsymbol{x}_t + \boldsymbol{b}_i), \tag{12}$$
$$\boldsymbol{o}_t = \sigma(\boldsymbol{W}_{oh}\boldsymbol{h}_{t-1} + \boldsymbol{W}_{ox}\boldsymbol{x}_t + \boldsymbol{b}_o). \tag{13}$$

Then, the two hidden state vectors are updated as follows:

$$\tilde{\boldsymbol{C}}_t = \tanh(\boldsymbol{W}_{Ch}\boldsymbol{h}_{t-1} + \boldsymbol{W}_{Cx}\boldsymbol{x}_t + \boldsymbol{b}_C), \tag{14}$$
$$\boldsymbol{C}_t = \boldsymbol{f}_t \circ \boldsymbol{C}_{t-1} + \boldsymbol{i}_t \circ \tilde{\boldsymbol{C}}_t, \tag{15}$$
$$\boldsymbol{h}_t = \boldsymbol{o}_t \circ \tanh(\boldsymbol{C}_t), \tag{16}$$

where $\circ$ represents element-wise multiplication. This structure has proven to be very effective in capturing long-term temporal dependencies, and therefore, is expected to outperform the basic RNN unit when representing the dynamical DR model.
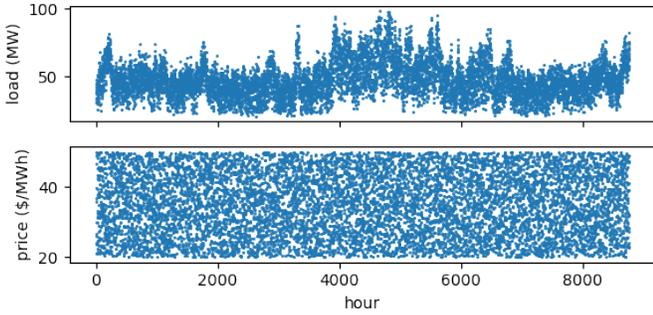
Fig. 3. Simulated historical energy consumption and price data.

	order n	0	1	2	3	4	5	6
Train	MAPE (%)	19.19	6.21	4.48	4.08	4.13	3.87	3.85
	SDAPE (%)	14.93	4.96	3.88	3.59	3.62	3.48	3.46
Test	MAPE (%)	14.41	5.57	4.08	3.91	4.13	3.84	3.82
	SDAPE (%)	10.84	4.49	3.22	3.13	3.27	3.04	3.03

	order n	0	1	2	3	4	5	6
Train	MAPE (%)	17.02	4.70	3.04	2.93	2.89	3.08	2.91
	SDAPE (%)	12.20	3.65	2.55	2.52	2.45	2.67	2.37
Test	MAPE (%)	13.64	4.88	3.37	3.32	3.31	3.28	3.25
	SDAPE (%)	9.85	3.48	2.58	2.52	2.49	2.53	2.60

The multi-layer LSTM network is similar to the RNN in Fig. 1. We simply replace the RNN unit with the LSTM unit.

IV. SIMULATION RESULTS

In this section, we numerically investigate the performance of the proposed dynamical DR model under the direct and indirect approaches.

A. Simulation Setup

Assume $T = 24$, i.e., each time interval covers one hour. We assume the benefit function has a quadratic form, the parameter in which is a constant, i.e.,

$$\beta_t^k(e_t^{k,d}, e_t^{k,c}) = \rho^k(e_t^{k,d} - e_t^{k,c})^2. \tag{17}$$

The feasible set is $\mathcal{E}_t^{k,c} = \{e_t^{k,c} \geq 0.5e_t^{k,d}\}$. The backlog rate is uniformly sampled from $[0, 1]$. The new demand is generated according the following procedures. First, the peak demand of each EUC is sampled uniformly from $[0.1, 2]$ MW, then the new demand of the EUC is computed as the product of the peak demand, the normalized annual load profile from one of the zones in PJM in 2017 [17], and a Gaussian random variable with a mean of 1 and a standard devision of 0.1. The value of ρ^k is taken as the ratio of -100 \$/MWh^2 and the EUC's peak demand. It is chosen in this way such that the responses from EUCs are reasonable. In reality, the value of ρ^k is a choice of the EUCs. The number of EUCs served by the LSE is 100. The historical data are simulated using the parameters described above, and a time series of prices that are sampled uniformly from $[20, 50]$ \$/MWh, which is shown in Fig. 3. The first 8016 sets of data (corresponding to January to October) are used for training, and the last 744 sets of data for testing (corresponding to November and December). The neural networks are implemented using Tensorflow [18], trained with the Adam optimizer with a learning rate of 0.001 and training step of 10000.

B. Performance Comparison

The mean absolute percentage error (MAPE), and the standard deviation of absolute error (SDAPE) are used to evaluate the performance of the learned model. Note that the units of MAPE and SDAPE are %.

1) Linear Function: We first test the performance of dynamical DR model under direct approach. Table I shows the results when the linear function is used in the direct approach. Note that when order $n = 0$, i.e., no information on previous time interval is utilized, both the training error and testing error are high, with MAPEs being 19.19% and 14.41%, respectively. The errors drop significantly when $n > 0$, which verifies the claim that the price responsive characteristics of the EUCs depend on their states in previous time intervals, rather than being static over time. Increasing the order significantly improves the model performance when $n \leq 2$, however, When $n \geq 2$, the improvement becomes negligible. Therefore, when the linear function is used, an appropriate order of the dynamical DR model would be 2. Intuitively, this means the response of the aggregate EUC demand to the price signal depends on the previous 2 hours, rather than being static.

2) FNN: A FNN with $L = 2$ hidden layers, each with 32 neurons, is also used to represent the dynamical DR model. Table II shows the results when the FNN is used in the direct approach. It is clear that the FNN outperforms the linear function when the order of the model is the same, which indicates that the linear function is underfitting the price and energy consumption data. In addtion, when the FNN is used in the direct approach, the number of order of the dynamical DR model can be much lower than the case with the linear function. For example, the testing MAPE of the FNN is 3.37% when $n = 2$, much better than that of the linear function when $n = 6$, which is 3.82%.

3) RNN/LSTM: In the indirect approach, a RNN/LSTM with $L = 1$ hidden layer and 32 neurons is adopted. The results are presented in Table II. In the indirect approach, only the most recent price and energy consumption is sent into the RNN/LSTM for prediction; The testing MAPEs of the indirect approach, which are 3.25% with RNN and 3.16% with LSTM, are comparable with that of the direct approach with FNN of order 6, which is 3.25%, while the LSTM performs better than the RNN by a small margin.

The testing MAPE obtained using linear function with order

TABLE III
PERFORMANCE OF DYNAMIC DR MODEL USING RNN/LSTM.

|		RNN	LSTM
Train	MAPE (%)	3.02	2.83
	SDAPE (%)	2.68	2.56
Test	MAPE (%)	3.25	3.16
	SDAPE (%)	2.55	2.47
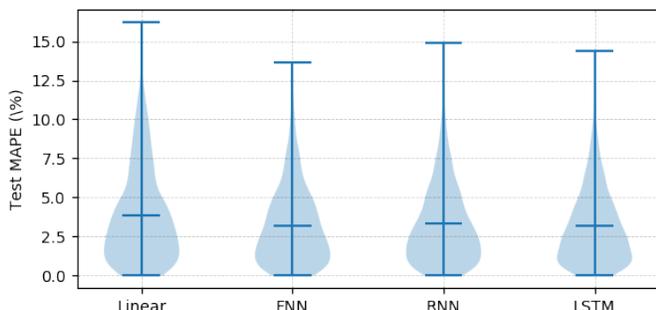
Fig. 4. Testing MAPE under different approaches.

6, FNN with order 6, RNN, and LSTM are presented in the violin plot in Fig. 4. To sum up, both the direct and indirect approach can be applied to learn a good dynamical DR model. FNN, RNN, and LSTM outperforms the linear function, at the cost of higher model complexity. In particular, when RNN or LSTM is used, no manual selection of states are required.

C. Discussion

When the neural networks are used to represent the dynamical DR model, we also investigated the impacts of hyperparameter such as the order the dynamical DR model, the number of layers, and the number of neurons in each layer. Among all factors, the order of the dynamical DR model has the most significant impact on the performance, which has already been discussed in details in the above section. The impacts of other parameters are relatively small.

Also, we would like to emphasize that the learned dynamical DR model can find its application in a variety of scenarios. For example, it can be used to predict the energy consumption profile over several time intervals, given the prices over the same intervals. It can also be used in an agent that simulates the collectively behavior of a set of EUCs without necessarily modeling all the underlying components; such an agent can be very helpful in developing certain algorithms.

V. CONCLUDING REMARKS

In this paper, we proposed a dynamical DR model that captures the temporal behavior of EUCs. A key element in the proposed model is the choice of state, which can be determined using either a direct approach that selects the state manually from the available information, or an indirect approach that computes the states from input information. The dynamical DR model can be represented by a linear function or FNN when

the direct approach is used, and a RNN or LSTM network when the indirect approach is used. Both these two approaches can achieve small MAPE when predicting the response of the aggregate energy consumption to the price. Numerical simulation results validated that the dynamical DR models are indeed necessary to model the price response characteristics of the EUCs, which are inherently temporally correlated.

Future research will utilize the proposed dynamical DR model to learn a pricing policy of the LSE purely from historical data using agent-based algorithms.

REFERENCES

[1] F. Wang, H. Xu, T. Xu, K. Li, M. Shafie-Khah, and J. P. Catalão, "The values of market-based demand response on improving power system reliability under extreme circumstances," *Applied energy*, vol. 193, pp. 220–231, 2017.

[2] F. Rahimi and A. Ipakchi, "Demand response as a market resource under the smart grid paradigm," *IEEE Transactions on smart grid*, vol. 1, no. 1, pp. 82–88, 2010.

[3] J. Medina, N. Muller, and I. Roytelman, "Demand response and distribution grid operations: Opportunities and challenges," *IEEE Transactions on Smart Grid*, vol. 1, no. 2, pp. 193–198, 2010.

[4] X. Zhang, M. Pipattanasomporn, M. Kuzlu, and S. Rahman, "Conceptual framework for a multi-building peak load management system," in *PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), 2016 IEEE*. IEEE, 2016, pp. 1–5.

[5] X. Zhang, M. Pipattanasomporn, and S. Rahman, "A self-learning algorithm for coordinated control of rooftop units in small-and medium-sized commercial buildings," *Applied Energy*, vol. 205, pp. 1034–1049, 2017.

[6] X. Zhang, G. Hug, J. Z. Kolter, and I. Harjunkoski, "Demand response of ancillary service from industrial loads coordinated with energy storage," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 951–961, 2018.

[7] K. Zhang, L. Lu, C. Lei, H. Zhu, and Y. Ouyang, "Dynamic operations and pricing of electric unmanned aerial vehicle systems and power networks," *Transportation Research Part C: Emerging Technologies*, vol. 92, pp. 472–485, 2018.

[8] H. Sun, A. Minot, D. Nikovski, H. Hashimoto, T. Takano, and Y. Takaguchi, "Mitigating substation demand fluctuations using decoupled price schemes for demand response," in *Innovative Smart Grid Technologies Conference (ISGT), 2016 IEEE Power & Energy Society*. IEEE, 2016, pp. 1–5.

[9] H. Xu, K. Zhang, and J. Zhang, "Optimal joint bidding and pricing of profit-seeking load serving entity," *IEEE Transactions on Power Systems*, 2018.

[10] B.-G. Kim, Y. Zhang, M. Van Der Schaar, and J.-W. Lee, "Dynamic pricing and energy consumption scheduling with reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 7, no. 5, pp. 2187–2198, 2016.

[11] S. Yousefi, M. P. Moghaddam, and V. J. Majd, "Optimal real time pricing in an agent-based retail market using a comprehensive demand response model," *Energy*, vol. 36, no. 9, pp. 5716–5727, 2011.

[12] A. J. Conejo, J. M. Morales, and L. Baringo, "Real-time demand response model," *IEEE Transactions on Smart Grid*, vol. 1, no. 3, pp. 236–242, 2010.

[13] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep learning*. MIT press Cambridge, 2016, vol. 1.

[14] P. J. Werbos, "Backpropagation through time: what it does and how to do it," *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1550–1560, 1990.

[15] "Understanding LSTM networks," http://colah.github.io/posts/2015-08-Understanding-LSTMs/, accessed: 2018-07-03.

[16] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[17] "PJM metered load data," http://www.pjm.com/markets-and-operations/ops-analysis/historical-load-data.aspx, accessed: 2018-07-23.

[18] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard *et al.*, "Tensorflow: a system for large-scale machine learning." in *OSDI*, vol. 16, 2016, pp. 265–283.