



Missouri University of Science and Technology
Scholars' Mine

Mathematics and Statistics Faculty Research &
Creative Works

Mathematics and Statistics

01 Aug 2019

Joint Manufacturing and Onsite Microgrid System Control using Markov Decision Process and Neural Network Integrated Reinforcement Learning

Wenqing Hu

Missouri University of Science and Technology, huwen@mst.edu


Zeyi Sun

Missouri University of Science and Technology, sunze@mst.edu

Y. Zhang

Y. Li

Follow this and additional works at: https://scholarsmine.mst.edu/math_stat_facwork

 Part of the [Mathematics Commons](#), and the [Operations Research, Systems Engineering and Industrial Engineering Commons](#)

Recommended Citation

W. Hu et al., "Joint Manufacturing and Onsite Microgrid System Control using Markov Decision Process and Neural Network Integrated Reinforcement Learning," *Procedia Manufacturing*, vol. 39, pp. 1242-1249, Elsevier B.V., Aug 2019.

The definitive version is available at <https://doi.org/10.1016/j.promfg.2020.01.345>



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License](#).

This Article - Conference proceedings is brought to you for free and open access by Scholars' Mine. It has been accepted for inclusion in Mathematics and Statistics Faculty Research & Creative Works by an authorized administrator of Scholars' Mine. This work is protected by U. S. Copyright Law. Unauthorized use including reproduction for redistribution requires the permission of the copyright holder. For more information, please contact scholarsmine@mst.edu.



25th International Conference on Production Research Manufacturing Innovation:
Cyber Physical Manufacturing
August 9-14, 2019 | Chicago, Illinois (USA)

Joint Manufacturing and Onsite Microgrid System Control Using Markov Decision Process and Neural Network Integrated Reinforcement Learning

Wenqing Hu^a, Zeyi Sun^{b*}, Yunchao Zhang^c, Yu Li^b

^aDepartment of Mathematics and Statistics, Missouri University of Science and Technology, Rolla MO 65409, USA

^bDepartment of Engineering Management and Systems Engineering, Missouri University of Science and Technology, Rolla MO 65409, USA

^cDepartment of Computer Science, Missouri University of Science and Technology, Rolla MO 65409, USA

Abstract

Onsite microgrid generation systems with renewable sources are considered a promising complementary energy supply system for manufacturing plant, especially when outage occurs during which the energy supplied from the grid is not available. Compared to the widely recognized benefits in terms of the resilience improvement when it is used as a backup energy system, the operation along with the electricity grid to support the manufacturing operations in non-emergent mode has been less investigated. In this paper, we propose a joint dynamic decision-making model for the optimal control for both manufacturing system and onsite generation system. Markov Decision Process (MDP) is used to formulate the decision-making model. A neural network integrated reinforcement learning algorithm is proposed to approximately estimate the value function given policy of MDP. A case study based on a manufacturing system as well as a typical onsite microgrid generation system is conducted to validate the proposed MDP model as well as the solution strategy.

© 2019 The Authors. Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Peer-review under responsibility of the scientific committee of the ICPR25 International Scientific & Advisory and Organizing committee members

* Corresponding author. Tel.: +1-573-341-7745; fax: +1-573-341-6567.

E-mail address: sunze@mst.edu

Keywords: Onsite generation system; Manufacturing system; Markov Decision Process; Neural network; Reinforcement learning.

1. Introduction

A microgrid is a localized autonomous energy system that consists of distributed energy sources and load, which can operate either separate from, or connected to, the existing utility power grid [1-3]. It has been widely recognized as a promising solution to meet the increasing electricity demand by relieving the disturbances and enhancing the resilience for the utility grid [4- 6]. The microgrid system typically works in a non-centralized mode, and thus, it is expected to be able to mitigate the complexity of traditional energy dispatch mechanisms in centralized grid operation through utilizing the distributed sources such as solar energy, wind energy, geothermal energy, hydro-thermal energy, etc. to serve local loads [7]. It can enhance the efficiency improvement in energy production [8-9] and accelerate the replacement for fossil fuels by various sources of renewable energy [10-11].

Various studies on microgrids have been conducted from the perspectives of system design and control scheme considering different components within the microgrid for the customers in various end-use sectors [12-14]. While, since manufacturing is traditionally not considered a critical facility, the research on the application of microgrid in manufacturing has been less reported. Recently, the investigation has been implemented on the optimal design and component sizing of the microgrid for manufacturing plant [15-17], as well as the optimal energy control from the manufacturing side [18-20] towards sustainability. However, the topic of joint energy control and management for both on-site microgrid generation system and manufacturing plant simultaneously has not been touched. Actually, manufacturing activities dominate the energy consumption and GHG emissions in the industrial sector [21] that accounts for one third of the total energy consumption in the U.S. [22]. In addition, at an age when it is impossible to conduct manufacturing activities in the absence of electricity, even a short power outage can cause detrimental impacts on manufacturing enterprises [23-26].

Therefore, there is an urgent need to extend the research on microgrid technology from traditional residential sector, commercial sector, critical facilities, etc., to the end-use customers in manufacturing sector. Specifically, the state-of-the-art that 1) microgrid utilization is less commonly studied for manufacturing plant, and 2) most existing literature with respect to control schemes focuses on a single side, either the microgrid, or the target customer load, needs to be advanced considering the joint energy control for both energy supply from the microgrid and energy load from the manufacturing plant.

In this paper, we propose to establish a systems and theoretical approach for joint energy management of manufacturing plant and on-site microgrid to achieve highly resilient manufacturing in a cost-effective and environmentally sustainable manner. A joint optimal control problem to coordinate the energy supply of the microgrid and the load of the manufacturing plant is formulated using Markov Decision Process (MDP). A neural network based reinforcement learning (NNRL) algorithm to explore the solution efficiency to the formulated MDP is proposed. A numerical case study for a small size manufacturing system is conducted to validate the proposed MDP model and discuss the effectiveness of the proposed solution strategy. The rest of the paper is organized as follows. Section 2 lays out the details of the formulation of the energy joint control model. Section 3 introduces the NNRL algorithm in detail. Section 4 introduces the numerical case study. Section 5 concludes the paper and discusses the future work.

2. Joint Energy Control Model Using Markov Decision Process

In this section, an MDP model is proposed to model the decision-making of the joint control of both manufacturing system and onsite distributed generation system. The manufacturing system modeled is a typical serial production line with I machines and $I-1$ buffers as shown in Fig.1 where rectangles denote machines and circles denote buffers. The time horizon is discretized into a set of discrete intervals. At the beginning of each interval, the control actions identified based on the optimal policy and the given state can be implemented. The state, policy, state transition, objective function, and constraint are introduced as follows.

System State. The system state at time decision epoch t is denoted by \mathbf{S}_t . It includes the states of manufacturing system (\mathbf{S}_t^{mfg}), onsite generations system (\mathbf{S}_t^{ong}), and surrounding environment (\mathbf{S}_t^{env}), which can be formulated by $\mathbf{S}_t = (\mathbf{S}_t^{mfg}, \mathbf{S}_t^{ong}, \mathbf{S}_t^{env})$. \mathbf{S}_t^{mfg} can be denoted by $\mathbf{S}_t^{mfg} = (M_t^1, M_t^2, \dots, M_t^I, B_t^1, B_t^2, \dots, B_t^{I-1})$, where M_t^i ($i=1, 2, \dots, I$)

denotes the state of machine i in the manufacturing system at decision epoch t . M_i^t can represent the status of the machine, denoting that if it is turned off, or kept working, or failure. B_i^t ($i=1, 2, \dots, I-1$) denotes the state of the buffer i in the manufacturing system at decision epoch t . B_i^t reflects the status of the buffer by specifying the number of Work-in-Progress (WIP) stored in the buffer. S_t^{ong} can be denoted by $S_t^{ong} = (S_t, W_t, G_t, SOC_t)$, where S_t, W_t and G_t denote the working status of solar PV, wind turbine, and generator, respectively, of the onsite generation system at decision epoch t . SOC_t denotes the state of charge of the battery system at decision epoch t . S_t^{env} can be denoted by $S_t^{env} = (I_t, F_t)$, where I_t denotes the solar irradiance at decision epoch t . F_t denotes the wind speed at



Fig. 1. A typical manufacturing system with I machines and I-1 buffers

decision epoch t .

Control Actions and Policy. Let π be the policy that maps from different states S to the actions A . The control actions adopted at decision epoch t can be denoted by A_t . It includes the control actions for the manufacturing system (A_t^{mfg}) and the onsite generation system (A_t^{ong}), which can be denoted by $A_t = (A_t^{mfg}, A_t^{ong})$. A_t^{mfg} can be denoted by $A_t^{mfg} = (a_t^1, a_t^2, \dots, a_t^I)$, where a_t^i ($i=1, 2, \dots, I$) is the control actions (on/off state adjustment) for machine i at the decision epoch t . A_t^{ong} can be denoted by $A_t^{ong} = (c_t, d_t, s_t, p_t, g_t)$, where c_t, d_t, s_t, p_t , and g_t denote the energy charging to the battery, the energy discharging from the battery, the energy sold back to the grid, the energy purchased from the grid, and the power generated by the generator, respectively.

State Transition. Let the function $P: S \times A \times S \rightarrow [0, 1]$ be the transition probability function, so that $P(S', S, A) \equiv P(S' | S, A)$ is the probability of transition to state S' given that the previous state was S and action A was taken at state S . The state transition given state S and adopted action A at decision epoch t is partially deterministic and partially stochastic, although they can all be put in the transition probability function P . For example, the states of the environmental factors like solar irradiance, and wind speed are random variables. For the manufacturing plant, when the “off” action is adopted for a certain machine, the state of such a machine at the next decision epoch t is deterministic. While, the “on” action adopted for a certain machine at the decision epoch t cannot necessarily guarantee the working state of the machine at decision epoch $t+1$ when the possibility of machine random failures is not excluded. The buffer state transfer is a stochastic process since the variation of the number of WIP in buffer depends on the working states of both upstream and downstream machines. The state transition of the throughput of the manufacturing system at decision epoch $t+1$ given state S and adopted action A at time t is also a random process.

Objective Function. At decision epoch t , a transition from state S_t to state S_{t+1} under action A_t results in an incurred cost $E(S_{t+1}, A_t, S_t)$, which includes the energy cost considering both microgrid and utility grid and the throughput reward of the manufacturing system. The total incurred cost from the beginning to the end of planning horizon, starting from state S and under policy π , is given by

$$C(S, \pi) = \mathbf{E} \left[\sum_{t=0}^{\infty} \lambda^t E(S_{t+1}, \pi(S_t), S_t \mid S_0 = S) \right] \tag{1}$$

where $\lambda \in [0, 1)$ is the discount factor. The objective is to identify an optimal policy $\pi^* = \underset{\pi \in \Pi}{\operatorname{argmin}} C(S, \pi)$ that can guide the decision maker to find appropriate actions based on the given system state to minimize the total incurred cost $C(S, \pi)$ in (1).

3. Neural Network Integrated Reinforcement Learning

The objective is to identify an optimal policy π^* and its corresponding total cost $C^*(S)$ such that

$$C^*(\mathbf{S}) \equiv C(\mathbf{S}, \pi^*) = \min_{\pi \in \Pi} C(\mathbf{S}, \pi) \quad (2)$$

Note that $C^*(\mathbf{S})$ includes both energy consumption cost and production throughput reward, which can be calculated by

$$C^*(\mathbf{S}) = F(\mathbf{S}) - TP(\mathbf{S}) \quad (3)$$

where $F(\mathbf{S})$ is the energy consumption cost, and $TP(\mathbf{S})$ is the production throughput of the manufacturing system.

The bellman equation of $C^*(\mathbf{S})$ is

$$C^*(\mathbf{S}, \pi) = \min \left[\bar{E}(\mathbf{S}, \mathbf{A}) + \lambda \sum_{t=0}^{\infty} P(\mathbf{S}', \mathbf{A}, \mathbf{S}_t) C^*(\mathbf{S}') \right], \text{ for } \mathbf{S} \in S \quad (4)$$

where $\bar{E}(\mathbf{S}, \mathbf{A})$ is the average total cost when action \mathbf{A} is taken at state \mathbf{S} , which can be calculated by

$$\bar{E}(\mathbf{S}, \mathbf{A}) = \sum_{S' \in S} P(\mathbf{S}', \mathbf{A}, \mathbf{S}_t) E(\mathbf{S}', \mathbf{A}, \mathbf{S}_t) \quad (5)$$

To solve the above Bellman equation without a-priori knowledge of the transition probability $P(\mathbf{S}', \mathbf{A}, \mathbf{S})$, we have to resort to *simulation based methods*, such as Reinforcement Learning [27-28]. In particular, we propose to use an algorithm called *Q-learning*. To perform *Q-learning*, let us first introduce the *Q-factor* $Q(\mathbf{S}, \mathbf{A})$, $\mathbf{S} \in S$, $\mathbf{A} \in \Pi(\mathbf{S})$ such that

$$C^*(\mathbf{S}) = \min_{\mathbf{A} \in \Pi(\mathbf{S})} Q(\mathbf{S}, \mathbf{A}) \quad (6)$$

Then, the *Q-factor* version of the Bellman equation takes the form

$$Q(\mathbf{S}, \mathbf{A}) = \sum_{S' \in S} P(\mathbf{S}', \mathbf{A}, \mathbf{S}) \left[E(\mathbf{S}', \mathbf{A}, \mathbf{S}) + \lambda \min_{\mathbf{B} \in \Pi(\mathbf{S}')} Q(\mathbf{S}', \mathbf{B}) \right] \quad (7)$$

Once we solve the above *Q-factor* version of the Bellman equation, we can then calculate $C^*(\mathbf{S})$ from (6). The equation (7) can be solved via the Robbins-Monro iteration

$$Q^{n+1}(\mathbf{S}, \mathbf{A}) \leftarrow (1 - \alpha^{n+1}) Q^n(\mathbf{S}, \mathbf{A}) + \alpha^{n+1} \left[E(\mathbf{S}', \mathbf{A}, \mathbf{S}) + \lambda \min_{\mathbf{B} \in \Pi(\mathbf{S}')} Q^n(\mathbf{S}', \mathbf{B}) \right] \quad (8)$$

Here \mathbf{S}' is the state next to \mathbf{S} under action \mathbf{A} . The iteration (8) can be performed by simulating the MDP for sufficiently long time, while updating the $Q^n(\mathbf{S}, \mathbf{A})$. We can pick $\alpha^n = A/(B+n)$ for some $A, B > 0$. A possible choice is to take $A = 1$ and $B = 0$, so that $\alpha^n = 1/n$.

This is the *Q-learning* algorithm for the look-up table case, and such algorithm has guaranteed convergence ([28, Chapter 5], [27, Chapter 6]). In practice, due to the possibly very large size of the state space S and the action space A , one can consider, instead of solving for $Q(\mathbf{S}, \mathbf{A})$, an approximate solution of the dynamic programming problems (see [28, Chapter 6], [27, Part II]). We will store the information of the values of the *Q-function* in a neural-network

with weight vector w . In this case, our approximate Q -function is given by the neural-network output $Q(S,A;w)$, where w is a parameter vector, and (S,A) are the inputs of the neural-network. Given any $S \in S$ and $A \in A$, the function $Q(S,A;w)$ will return an approximation of the value $Q(S,A)$.

The proposed neural-network has a fixed architecture, with two input neurons standing for both S and A and one output neuron standing for $Q(S,A;w)$. Various methods in the training process of neural network have been proposed and implemented. For example, some recent works have reported the training process via stochastic gradient descent [29, 30]. In this paper, the training is the iteration about the weight vector w . At iteration n , we replace the right hand side of the iteration scheme (8) by the $Q(S,A;w^n)$, where w^n is the weight vector at the n -th iteration. This produces a new Q -value that is used as a training data to feed the neural-network. We can obtain an updated w^{n+1} via an online training of the neural-network with this Q -value.

Thus, we have the neural-network based Q learning algorithm as shown in Fig. 2. It can be clearly seen that in this algorithm, the iterations are with respect to the weights of the neural-network and information regarding the total cost at each step is obtained by running the MDP and they are fed into the iteration scheme. Thus, essentially it is the MDP updating the neural-network.

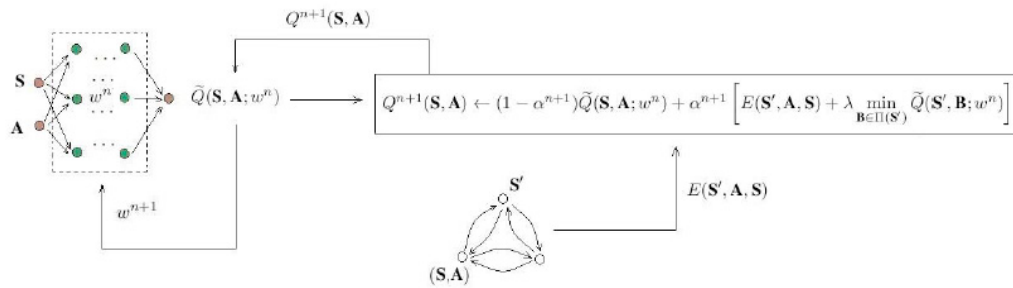


Fig. 2. Q-learning and approximate dynamic programming via neural network

4. Case Study

The manufacturing system used in the case study is composed of two machines and one buffer as shown in Fig.

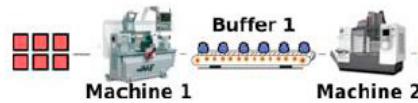


Fig. 3. A two-machine-and-one-buffer manufacturing system

3.

Each machine can be in one of four different states, i.e., off, idle, failure, and on at each time interval and the power consumption is different for each state. The mean time between failure (MTBF) and mean time to repair (MTTR) of each machine as well as the power at different states are shown in Table 1. The buffer has a limited capacity and will be initialized with certain quantity as presented in Table 2.

Table 1. Parameters of manufacturing machines

	MTBF (min)	MTTR (min)	Throughput	Power Consumption (kW)		
				Off/Failure	Idle	ON
Machine 1	95	10	1	0	600	1,000
Machine 2	105	15	1	0	600	1,000

Table 2. Parameters of buffers

	Initial Amount	Capacity
Buffer 1	10	20

The proposed microgrid system consists of batteries, onsite generators, solar panels, wind turbines. Batteries can be discharged for manufacturing system use and charged by solar panels, wind turbines, and utility grid. Onsite generators will only generate energy for manufacturing system. Solar panels and wind turbines can be used to generate energy for either manufacturing system, battery charging, or sold back. While the remaining need will be satisfied by utility purchasing from utility grid. The configurations of the proposed microgrid are shown in Table 3 and utility price is shown in Table 4.

Table 3. Parameters of the onsite microgrid system

Component of Microgrid	Parameters	Values
Solar PV	Area of solar PV (m ²)	300
	Efficiency of the solar PV	0.2
	O&M cost (\$/kWh)	0.02
Wind Turbine	Density of air (kg/m ³)	1.225
	Power coefficient	0.593
	Gearbox transmission efficiency	0.9
	Electrical generator efficiency	0.9
	Blade Radius (m)	20
Battery	O&M cost (\$/kWh)	0.03
	Charging Efficiency	0.9
	Discharging Efficiency	0.9
	O&M cost (\$/kWh)	0.1
	Capacity (kWh)	600
Generator	O&M cost (\$/kWh)	0.2
	Capacity (kWh)	400

Table 4. Electricity price

Season	Period	Time	Consumption charge (\$/kWh)	Sold back price (\$/kWh)
Summer	On-Peak	1:00 pm-6:00 pm	0.35	0.17
	Mid-Peak	10:00 am-1:00 pm	0.19	0.07
	Off-Peak	6:00 pm-9:00 pm	0.06	0

The state space and action space of the MDP model depend on the capacity of the microgrid system, the size of the manufacturing system, as well as the length of the planning horizon. For state space, three possible states (“ON”, “IDLE”, “OFF”) of each machine are represented by three binary inputs, respectively. We also have the amount of each intermediate buffer, solar irradiance used for generation, wind source used for generation, generator output, state of charge of the battery, soldback rate, electricity amount purchased from utility, and time of period as discrete inputs.

For action space, control action of each machine is still binary, either “Turn on” or “Turn off”. In order to keep the action space at a reasonable size, we discretize solar energy and wind energy directly used to power manufacturing system, charged to battery, and sold back into five intervals. Generator output and battery discharging rate are also discretized and they will only be considered if they are cheaper alternatives of purchasing from utility directly. After solar energy, wind energy, generator and battery are all considered, we set utility purchasing amount to any leftover energy demand. With smaller microgrid capacity and shorter planning horizon, the state space and action space will also be smaller and vice versa.

In the experiment, the neural network is designed to have two fully connected hidden layers with sigmoid activation functions. As there are variables for three states for each of two machines, one intermediate buffer, solar charging rate, wind energy charging rate, generator output, SOC of battery, sold back rate, utility purchasing, current solar irradiance, current wind speed, and time of period, we have to consider the state with a dimension of 16 in total. Since we have two machine operation actions, three solar energy actions (for support mfg, charging battery, sold back), three wind energy actions, one generator action, one battery discharging action, and one utility purchasing action, the dimension of the action is 11. Thus, the input layer has a dimension of 27. Each of the two hidden layers has 32 neurons. Our reward function incorporates two components, energy cost and total final throughput. The weights in the model are initialized randomly following normal distribution. The discount factor λ in Q -Learning is set as 0.1.

The convergence of the weight of the neural network is shown in Fig. 4. The (a-1), (a-2), (a-3) are for a relatively small state space of size 3.8×10^3 and action space of size 2.6×10^4 ; (b-1), (b-2), (b-3) are for a larger state space of

size 2.5×10^{19} and action space of size 1.0×10^7 .

We used Adam as our optimizer and learning rates are set differently: (a-1) and (b-1) with learning rate 0.001, while (a-2), (a-3), (b-2), and (b-3) with learning rate 0.00001. In Fig. 4, we plotted the square norms of the differences of the neural-network weight vectors for each two consecutive algorithm iterations: the horizontal axis is the number of iterations and the vertical axis represents the progression of the square norms of the differences of the weight vectors. The number of iterations for all six pictures are 10^5 . However, the vertical axes are of different scales. Pictures (a-1) and (a-2) have vertical axis of same size between 0 and 0.035 while (a-3) is a rescaled version of (a-2) with vertical axis of size 0.0009; (b-1) and (b-2) have vertical axis of same size 0.2 while (b-3) is a rescaled version of (b-2) with vertical axis of size 0.0009. A comparison of (a-1) and (a-2), as well as a comparison of (b-1) and (b-2) shows clearly that as the learning rate decays, the weights of the neural network get more stabilized, implying a faster convergence. In fact, with the same choice of scales of the vertical axis, in (a-2) and (b-2) the square norms of the weight differences are almost zero and hardly recognized in the pictures. When we look at a comparison of (a-1) and (b-1), as well as (a-3) and (b-3) (which are rescaled versions of (a-2) and (b-2)), we also see that the convergence speed depends on the size of the state and action spaces: the smaller these spaces are, the better convergence is guaranteed. In fact, we can see in (a-1) and (a-3) that the weight differences increase a little at the beginning and then gradually start to decay. However, in (b-1) and (b-3) we do not see a significant decay of the weight differences due to the huge size of state and action spaces, indicating that more iterations are needed. The increase of the weight differences at the beginning epochs of iterations in (a-1) and (a-3) can be explained by the fact that the MDP needs to first explore the state space and update the learning network. This results in an unstable behaviour of the network weights during these beginning epochs.

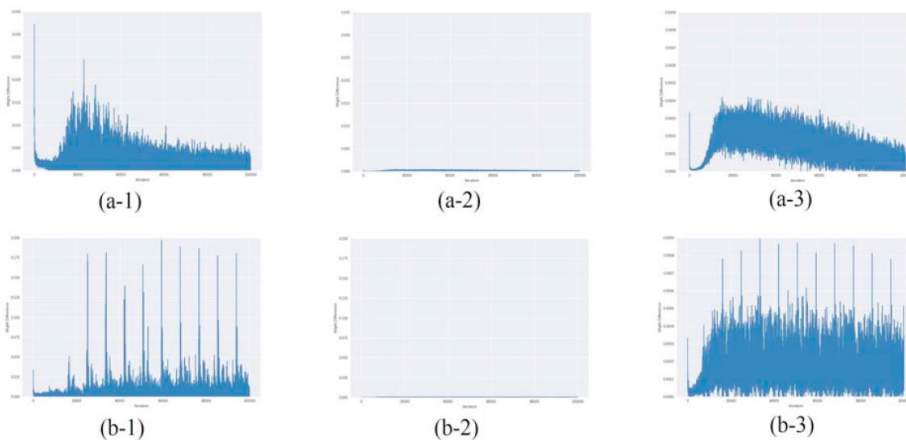


Fig. 4. Decay of the difference in the weights of the neural network approximating the MDP

5. Conclusions and Future Work

In this paper, we propose a joint energy control model for the optimal energy control for both the manufacturing system and onsite microgrid system towards sustainability. A neural network integrated reinforcement learning framework is proposed to explore the solvability of the proposed MDP formulation with high dimension and high complexity. The experimental results in the case study with different settings of neural network parameter and state space illustrate an appealing potentials of the convergence of the proposed algorithm.

For future work, more experiments can be implemented with larger state and action space. In addition, the proposed model can also be expanded to integrate the HVAC system in the manufacturing plant along with the manufacturing system since HVAC is widely considered the second top energy consumer, immediately following the manufacturing system, in a typical manufacturing plant.

References

- [1] C. Mahieux, A. Oudalov, Microgrids Enter the Mainstream, <http://www.renewableenergyfocus.com/view/43345/microgrids-enter-the-mainstream/>, 2015.
- [2] U.S. Department of Energy, Technical Report: How Microgrids Work, <http://www.energy.gov/articles/howmicrogrids-work>, 2014.
- [3] Lawrence Berkeley National Laboratory, Technical Report: Microgrid Definitions, <https://buildingmicrogrid.lbl.gov/microgrid-definitions>, 2016.
- [4] B. Lasseter, Microgrids, In IEEE Power Engineering Society Winter Meeting, 1 (2001) 146–149.
- [5] B. Lasseter, Microgrid, In IEEE Power Engineering Society Winter Meeting, 1 (2002) 305–308.
- [6] D.E. Olivares, A. Mehrizi-Sani, A.H. Etemadi, C.A. Cañizares, R. Iravani, M. Kazerani, A.H. Hajimiragha, O. Gomis-Bellmunt, M. Saadifard, R. Palma-Behnke, G.A. Jimnez-Estvez, N.D. Hatziargyriou, Trends in microgrid control, IEEE Transactions on Smart Grid, 5(2014) 1905–1919.
- [7] D.T. Ton, M.A. Smith, The U.S. Department of Energy’s microgrid initiative, The Electricity Journal, 25 (2012) 84–94.
- [8] N. Lior, Advanced energy conversion to power, Energy Conversion and Management, 38 (1997) 941–955.
- [9] N. Lior, Thoughts about future power generation systems and the role of exergy analysis in their development, Energy Conversion and Management, 43 (2002) 1187–1198.
- [10] X. Li, Diversification and localization of energy systems for sustainable development and energy security, Energy Policy, 33 (2005) 2237–2243.
- [11] H. Lund, Renewable energy strategies for sustainable development, Energy, 32 (2007) 912–919.
- [12] W.T. Chong, M.S. Naghavi, S.C. Poh, T.M.I. Mahlia, K.C. Pan, Techno-economic analysis of a wind– solar hybrid renewable energy system with rainwater collection feature for urban high-rise application, Applied Energy, 88(2011) 4067–4077.
- [13] A.B. Kanase-Patil, R.P. Saini, M.P. Sharma, Sizing of integrated renewable energy system based on load profiles and reliability index for the state of Uttarakhand in India, Renewable Energy, 36 (2011) 2809–2821.
- [14] J. Jung, M. Villaran, Optimal planning and design of hybrid renewable energy systems for microgrids, Renewable and Sustainable Energy Reviews, 75 (2017) 180–191.
- [15] X. Zhong, M.M. Islam, H. Xiong, Z. Sun, Simulation-based investigation for the application of microgrid with renewable sources in manufacturing systems towards sustainability, Procedia Computer Science, 114 (2017) 433–440.
- [16] Y. Zhang, M.M. Islam, Z. Sun, S. Yang, C. Dagli, H. Xiong, Optimal sizing and planning of onsite generation system for manufacturing in critical peaking pricing demand response program, International Journal of Production Economics, 206(2018) 261–267.
- [17] M.M. Islam, Z. Sun, Onsite generation system sizing for manufacturing plant considering renewable sources towards sustainability. Sustainable Energy Technologies and Assessments, 32 (2019) 1–18.
- [18] Y. Levron, J.M. Guerrero, Y. Beck, Optimal power flow in microgrids with energy storage, IEEE Transactions on Power Systems, 28 (2013) 3226–3234.
- [19] S. Li, J. Proano, D. Zhang, Microgrid power flow study in grid-connected and islanding modes under different converter control strategies, In IEEE Power and Energy Society General Meeting, (2012), 1–8.
- [20] J.M. Guerrero, J.C. Vasquez, J. Matas, L.G. De Vicuña, M. Castilla, Hierarchical control of droop controlled AC and DC microgrids—A general approach toward standardization. IEEE Transactions on Industrial Electronics, 58(2011) 158–172.
- [21] J.R. Dufloy, J.W. Sutherland, D. Dornfeld, C. Herrmann, J. Jeswiet, S. Kara, M. Hauschild, K. Kellens. Towards energy and resource efficient manufacturing: A processes and systems approach. CIRP Annals- Manufacturing Technology, 61 (2012) 587–609.
- [22] U.S. Department of Energy. Annual Energy Review 2009. <ftp://ftp.eia.doe.gov/multifuel/038409.pdf>.
- [23] F. Katiraei, C. Abbey, S. Tang, M. Gauthier. Planned islanding on rural feeders utility perspective. In IEEE Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century, 2008, 1–6.
- [24] J. Turkewitz. Unemployment Deepens Storms Loss as Businesses Stay Closed. <http://www.nytimes.com/2012/12/28/nyregion/unemployment-deepens-the-loss-from-hurricanesandy.html? r=0>, 2012.
- [25] M. Garber, J. Wohlford, L. Unger, L. White, Hurricane Katrinas Effects on Industry Employment and Wages, Monthly Labor Review August 2006. <http://www.bls.gov/opub/mlr/2006/08/art3full.pdf>, 2006.
- [26] T. Loix, K.U. Leuven. The First Microgrid in the Netherlands: Bronsbergen. http://www.leonardoenergy.org/webfm_send/493, 2009.
- [27] R.S. Sutton, A.G. Barto. Reinforcement Learning: An Introduction. Second Edition, in progress, complete draft online. MIT Press, November 5, 2017.
- [28] D.P. Bertsekas, J.N. Tsitsiklis. Neuro-Dynamic Programming. Athena Scientific, 1996.
- [29] Hu, W., Li, C.J., A convergence analysis of the perturbed compositional gradient flow: averaging principle and normal deviations. Discrete and Continuous Dynamical Systems, Series A, 38(10) 2018.
- [30] Hu, W., Li, C.J., Li, L., Liu, J., On the diffusion approximation of nonconvex stochastic gradient descent. Annals of Mathematical Science and Applications, to appear. arXiv:1705.07562 [stat.ML]