Journal of Power and Energy Engineering, 2022, 10, 1-13
https://www.scirp.org/journal/jpee
ISSN Online: 2327-5901
ISSN Print: 2327-588X

Reinforcement Learning-Based Control for Resilient Community Microgrid Applications

Md Mahmudul Hasan, Ishtiaque Zaman, Miao He, Michael Giesselmann

Electrical and Computer Engineering, Texas Tech University, Lubbock, USA
Email: md-mahmudul.hasan@ttu.edu, ishtiaque.zaman@ttu.edu, miao.he@ttu.edu, michael.giesselmann@ttu.edu

How to cite this paper: Hasan, M.M., Zaman, I., He, M. and Giesselmann, M. (2022) Reinforcement Learning-Based Control for Resilient Community Microgrid Applications. *Journal of Power and Energy Engineering*, **10**, 1-13.
https://doi.org/10.4236/jpee.2022.109001

Received: July 26, 2022
Accepted: September 27, 2022
Published: September 30, 2022

Abstract

A novel microgrid control strategy is presented in this paper. A resilient community microgrid model, which is equipped with solar PV generation and electric vehicles (EVs) and an improved inverter control system, is considered. To fully exploit the capability of the community microgrid to operate in either grid-connected mode or islanded mode, as well as to achieve improved stability of the microgrid system, universal droop control, virtual inertia control, and a reinforcement learning-based control mechanism are combined in a cohesive manner, in which adaptive control parameters are determined online to tune the influence of the controllers. The microgrid model and control mechanisms are implemented in MATLAB/Simulink and set up in real-time simulation to test the feasibility and effectiveness of the proposed model. Experiment results reveal the effectiveness of regulating the controller's frequency and voltage for various operating conditions and scenarios of a microgrid.

Keywords

Microgrid, Reinforcement Learning, Q-Learning Algorithm, Vehicle-to-Grid (V2G)

1. Introduction

Microgrid systems refer to an interconnected set of distributed energy sources (traditional and/or renewable) and controllable loads with clear boundaries to the main grid. Microgrid systems are typically equipped with islanding capability that enables microgrids to operate as a standalone entity with respect to the main grid [1] [2]. For decades, microgrid systems have been deployed and used in rural areas by numerous projects and agencies to supply power under the condition that support from the traditional power system is lacking or cost-ineffective

[3]. Microgrid systems have also been used in urban developments (e.g., community energy projects and initiatives) aiming to decentralize and localize energy generation to increase production from distributed renewable sources. In the following three decades, electrical energy demand is expected to increase by nearly 50% [4]. With the ramping penetration of renewable power generation, microgrid systems have gained increased attention in energy infrastructure development and enhancement [5]. Microgrid systems are also promising to optimize the energy portfolio and operating costs to achieve certain reliability of power supply [6].

On the flip side of the coin, significant challenges and obstacles arise toward efficiently utilizing renewable energy in microgrid systems. This is mainly due to the intermittency and stochastic nature of renewable energy sources, the lack of inertia of DC power sources (solar PV, battery, etc.), and the lack of efficient and robust hierarchical controllers (the capital cost of the distributed energy sources and microgrid systems is another contributing factor). Due to limited grid-forming capability, hierarchical control capabilities, and lack of inertia, existing microgrid systems can underperform in islanding operations [7] [8]. Maintaining the reliability of microgrid systems is of paramount importance yet challenging when it comes to intermittent and stochastic renewable energy sources such as solar PV.

With the advancement of smart controllers for power grid applications, microgrid systems can be enhanced in terms of improved controllability and stability [1]. Microgrid systems have been utilized as a cost-effective solution to circumvent the technical difficulties of integrating distributed renewable energy sources into residential and community energy systems [9] [10]. On the condition that the main grid fails due to rare yet disastrous events within the local grid, microgrid systems can keep the local power operation alive throughout the outage [11]. Therefore, modification, improvement, and adaptation of the control mechanism of microgrid systems to these rare operating environments under special circumstances are pivotal issues for community energy resilience. Under these circumstances, microgrid systems can tap the potential of EVs as a battery energy storage system (BESS) [12] to improve the reliability and power quality of microgrid systems with only renewable energy sources.

Nowadays, transportation electrification, to a greater extent, has changed the point of view toward EVs, which used to be considered part of the netload. As light-duty EVs are becoming widely available, the high-capacity batteries in the EV fleet could form an invaluable energy asset in community microgrid systems [13]. V2G technology could be applied to the EVs and plug-in hybrid electric vehicles (PHEVs) that are fitted with charge de move (CHAdeMO) or combined charging system chargers [11] [14]. Therefore, during a power interruption, the EV's battery can be used as a source for local loads in the microgrid. EVs with V2G capabilities and flexibilities can charge and discharge electrical energy from and to microgrid systems are feasible and cost-effective options to balance the

electrical power demand and supply, which can significantly contribute to the energy reliability and resilience of communities [15] [16].

In this paper, a novel control mechanism based on the previous work [17] is developed. The developed microgrid V2G control strategy utilizes droop control and virtual inertia control to regulate the frequency and voltage. Different from the household microgrid that contains only one PV system and on EV in the previous work [17], multiple solar PV systems and EVs are considered in the community microgrid model, and the controller is designed to achieve accurate real and reactive power-sharing among the multiple solar PV systems and EVs. Further, the controller in the previous work [17] assigns fixed weights to droop and virtual inertia controls, which may not effectively manage the variations in frequency and voltage caused by volatile solar PV power production. Therefore, in this paper, by proactively considering the stochastic nature of solar radiation, a reinforcement learning-based control mechanism is incorporated into the microgrid controller to adapt the weights to droop dynamically, and virtual inertia controls overall stability and power qualify could be improved. The rest of the paper is organized as follows. Section 2 presents the model for the community microgrid with multiple parallel solar PV systems and EVs and the control strategy. Experiment results are presented in Section 3. Conclusions are provided in Section 4.

2. Proposed Microgrid V2G Control Strategy

2.1. Community Microgrid System Model

The model for the community microgrid system considered in this study is illustrated in **Figure 1**. The microgrid system contains two parallel sets of distributed energy resources. The energy resource is comprised of a solar PV system and a BESS (as the battery of EVs). The solar PV system is connected to the DC link through a boost converter, for which a maximum power point tracking (MPPT) controller is utilized to adjust the duty cycle in an online manner. The DC link voltage is set b the BESS and is regulated by the DC link capacitor. The

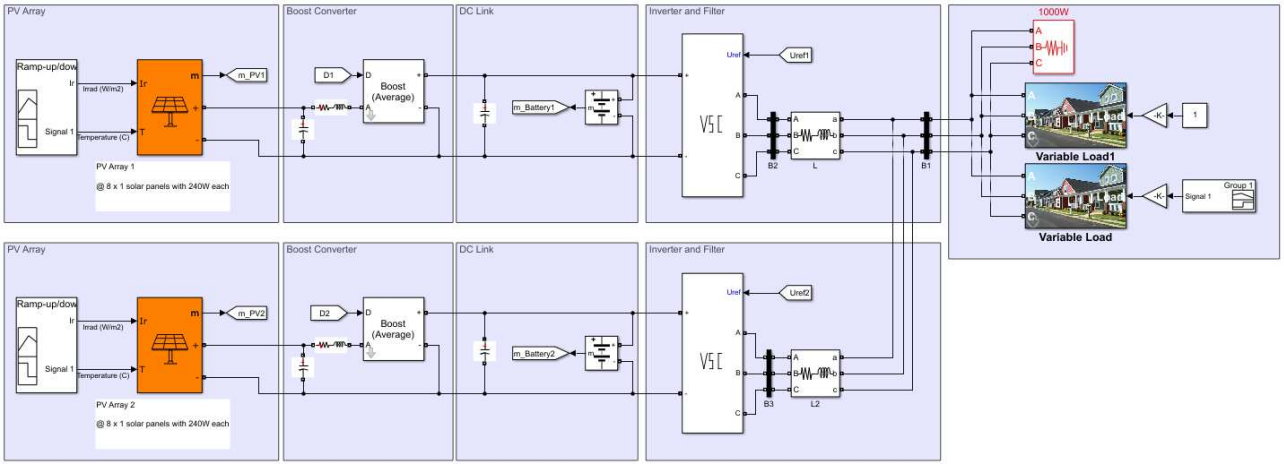

Figure 1. A community microgrid with parallel distributed energy resources.

DC link is then fed to the voltage source converter (VSC) to produce an AC output. Through series filters and lines, The VSCs are both connected to the point of common coupling (PCC) of the microgrid systems, where a set of loads are connected. The loads include a constant baseload, a ramping load with slow variations, and a variable load with fast variations.

In the above community microgrid system model, note that the controller for the grid-forming inverters of the community microgrid system is designed so that all distributed energy resources could accurately share the load according to their capacity. Thus, the microgrid systems could quickly scale up to incorporate more distributed energy sources. Also, the BESS is deployed between the solar PV system and the VSC so that the solar PV system can continuously operate in MPPT mode to produce maximum power output without the need to match the variable power output of VSC as required by the load, which is advantageous compared to the case of separate solar PV and BESS.

2.2. Controller of Distributed Energy Sources

To achieve accurate real and reactive sharing between the parallel distributed energy sources under steady-state operating conditions, the universal droop control [18] is adopted. Further, due to the lack of inertia of the DC energy sources, the community microgrid system can deliver poor transient response during rapidly changing operating conditions, e.g., during the fluctuation of solar PV power generation (as a very likely result of fog shading) or switching of loads. Therefore, the virtual inertia control from the previous work [17] is also adopted with a key modification. Specifically, combining the droop control and virtual inertia control, the real power to frequency (*p*-*f*) characteristic of the proposed controller is described below:

$$2H\dot{\omega} = \left(p^* - p\right)/\omega - \alpha D_\omega \left(\omega - \omega^*\right) - \beta K_d \left(\omega - \omega_g\right) \tag{1}$$

In the above equation, H is the inertia constant, K_d is the damping factor, D_ω is the *p*-*f* droop coefficient, p^* is the set point for real power p, ω^* is the set point for angular frequency ω, and ω_g is grid frequency. Note that under a quasi-steady state, ω is maintained at ω_g and thus $\dot{\omega} = 0$, then "Equation (1)" reduces to the conventional *p*-*f* droop characteristic equation. Otherwise, for transient conditions, the controller would direct the VSC and distributed energy sources to act as a spinning generator with inertia quantified by H and D_ω. Further, the coefficients, α and β, in "Equation (1)" refer to the weights assigned by the controller to droop control and virtual inertia control, respectively. Specifically, in the previous work [17], α and β are predetermined constants. When α is greater, the controller is more prone to perform droop control. However, when β is greater, the controller will assume more virtual inertia to improve transient response. Indeed, there is a trade-off between the two control modes. When α is greater, the frequency and voltage can fluctuate more than acceptable to the load, causing power quality issues. However, when β is greater, the controller would respond slower to compensate for the change in net load.

Therefore, in an online operating condition, the weights should be adjusted according to the net load, especially to the solar PV power production (as well as solar radiation).

A new contribution made by this paper is that a dynamic weight scheme is developed to address the issues described above. Based on the control strategy illustrated in **Figure 1** of the previous work [17], the dynamic weight scheme is incorporated, as shown in **Figure 2**. Specifically, the weights for the droop and the damping terms in "Equation (1)" are dynamically determined according to a look-up table that takes the grid condition (voltage magnitude and frequency) and the solar radiation as the input (note that in a realistic setting, this input should be the solar PV power calculated in real-time). The look-up table contains an off-policy which is learned offline with a discrete state and action space. More technical details of the off-policy obtained from reinforcement learning are provided in Section 2.3.

Further, built on the previous work [17], this paper incorporates cascaded voltage and current regulators following the improved droop and virtual inertia control (**Figure 3**). The voltage regulator accepts the reference voltage produced

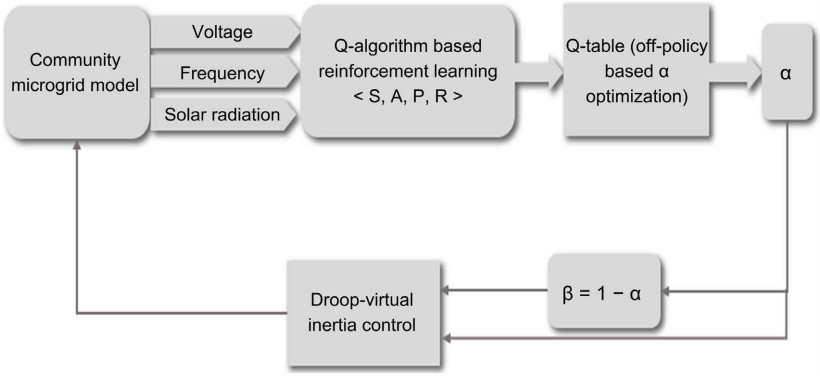

Figure 2. The proposed control strategy with dynamic weights for droop-virtual inertia control.

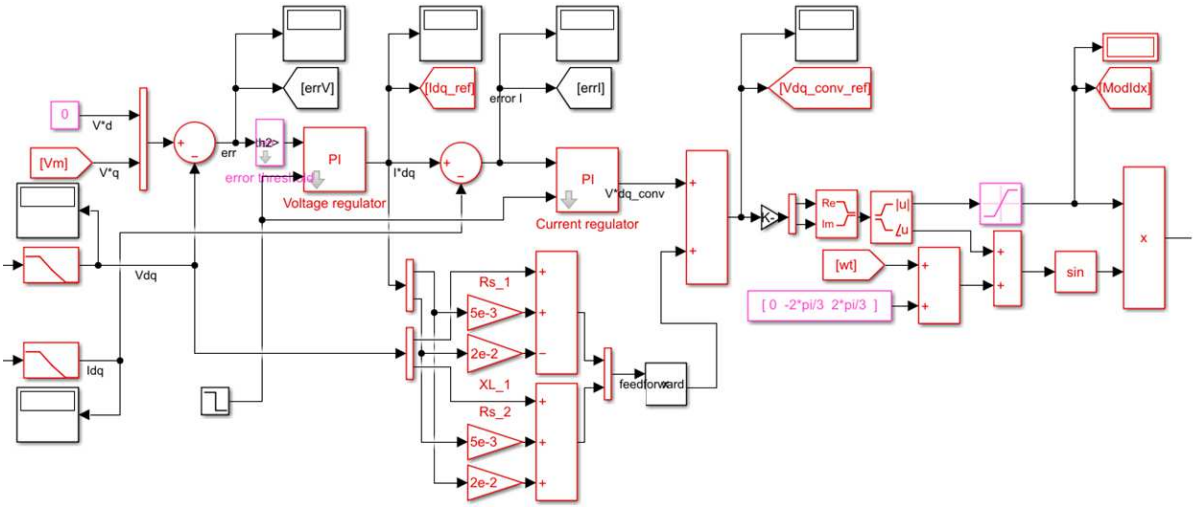

Figure 3. Voltage and current regulators of the community microgrid system controller.

by droop control and regulates the d/q components of the voltage at PCC. The voltage regulator produces the reference value for the d/q components of the VSC's output current, which the current regulator then uses. Particularly, the current regulator adopts a feedforward loop to account for the voltage drop across the filter and series impedance between VSC and PCC. Finally, the current regulator produces the VSC's output voltage to compute the control reference to the VSC.

2.3. Reinforcement Learning

In what follows, the proposed reinforcement learning-based algorithm for adaptive weights for droop and inertia control is described. Generally, in reinforcement learning, an agent is a function unit or computation unit that can automatically perceive information, generate the corresponding actions through decision-making and reasoning, and respond to the environment (**Figure 4**). Cooperative agents can work together in hybrid microgrid systems such as wind/solar systems in so-called cooperative reinforcement learning. The basic principle of reinforcement learning is that if one action could make the environment give the system an incremental reward, the trend of this action produced by the system will be either strengthened or weakened [19].

Inspired by the basic principles of reinforcement learning, the policy for finding the optimal dynamic weight of droop and virtual inertia control could be formulated as a sequential behavior decision process like the Markov decision process. Under quasi-steady states (10 s - 100 s), the system operating condition change mainly due to random load variation and fluctuation of solar PV power production, and thus system states assume Markovian properties. Considering this, Markov's decision process-based formulation is plausible. In what follows, the main elements of the Markov decision process-based formulation are presented [20].

State: the system states include microgrid and solar PV power states. More specifically, the voltage magnitude (V_{magt}) and frequency (f_t) at the PCC constitute the microgrid states (as the load's real and reactive power consumption

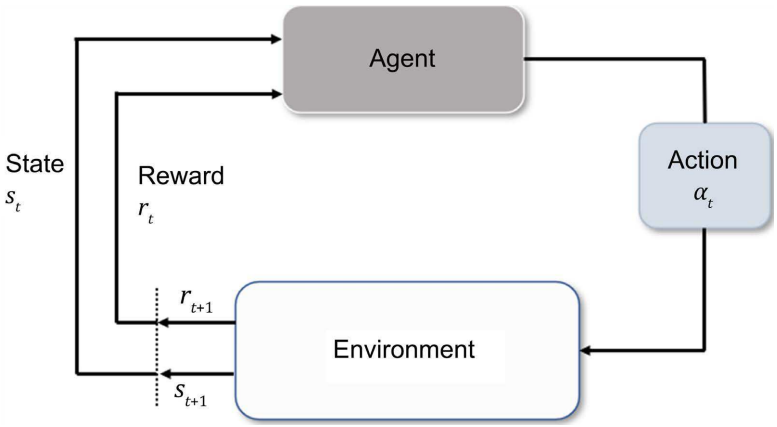

Figure 4. Reinforcement learning algorithm.

depends on these two variables under a dynamic load model). Further, solar radiation (Ir_t) which dictates solar power production should also be included in the state. Note that the ambient temperature is typically constant at the time-scales of quasi-steady states. Thus, it should not be included in the state, even though ambient temperature is a critical physical parameter for the PV panel's power output. It can also be seen from **Figure 2** that solar radiation, not the ambient temperature, is considered as the input to the off-policy.

Action: the action includes the values for the weights of droop and virtual inertia control. For computational efficiency, the action space is discretized, such that the weights take values from 0 to 1, with a step of 0.01. As discussed in Section 2.1, a trade-off exists between the two controls, and thus the weights are set to be complementary, as seen in **Figure 2**.

Transition probability: because the state space is continuous, rigorous characterization of the transition probability is intractable, and thus a model-free approach is adopted in this study.

Reward: the immediate reward of the system is comprised of three components, including frequency regulation, voltage regulation, and the electrical load satisfied. With this insight, the system's immediate reward at time *t* is given by:

$$R_t = -\left\{k_f \left|f - f^*\right| + k_v \left|V - V^*\right| + k_L \left|S\left(f,v\right) - S^*\right|\right\} \tag{2}$$

In "Equation (2)", f^*, V^*, and S^* denote the nominal values for the microgrid system frequency, the voltage at PCC, and the nominal load. k_f, k_v, and k_L are the weights for the three components. The dynamic load model in [19] is used for $S\left(f,v\right)$. Where $S\left(f,v\right)$ represents the complex power of the exponential load model and is a function of voltage and frequency. The mathematical relation between complex load power, voltage and frequency can be written as "Equation (3)". Where, S^*, V^*, and f^* are the nominal load power, nominal voltage, and nominal frequency, respectively. The exponential factor *a* and *b* are the load parameters and varies between 0 to 2.

$$S\left(f,v\right) = S^*\left(v/v^*\right)^a \left(f/f^*\right)^b \tag{3}$$

Solution: Q-learning algorithm is used to optimize the actions based on the state of the microgrid system. Q-table is the data structure used to calculate the maximum expected future rewards for action at each state. Basically, this table will guide us to the best action in each state. To find the best possible action value in every state.

"Equation (4)" is used [20].

$$J = \max\left\{Q\left(s_{t+1}, a\right)\right\} \tag{4}$$

where, s_{t+1} is the future state, a is the action by the agent. The Q-learning algorithm adopts the optimal action value in the state action procedure to update the Q-table in every step of the states. The state action value equation can be written as "Equation (5)" [20].

$$Q^{\text{new}}\left(s,a\right) = \left(1-\alpha\right)\cdot Q^{\text{old}}\left(s,a\right) + \alpha\left(R_t + \gamma \cdot J\right) \tag{5}$$

where, α is the learning rate (0, 1), γ is the discount factor and R_t is the reward. In Q-table, the action values are repeatedly updated by every step of the state until all the actions are optimized. If all the actions are optimized, the solution to the Markov decision process-based problem, as an off-policy, is then prepared into a look-up table to produce real-time decisions on the weights for droop and virtual inertia control, as shown in (**Figure 5**).

3. Experiment Results

In this section, the experiment results are presented. The system settings are the same as the previous work [17], except that the capacity of the second distributed energy source is doubled, and the total load is tripled according. In addition, real-world weather data, including solar irradiance and temperature at high time resolutions, which is collected from the West Texas Mesonet [21], is used as the data input to the Simulink model. The experiment results for the individual

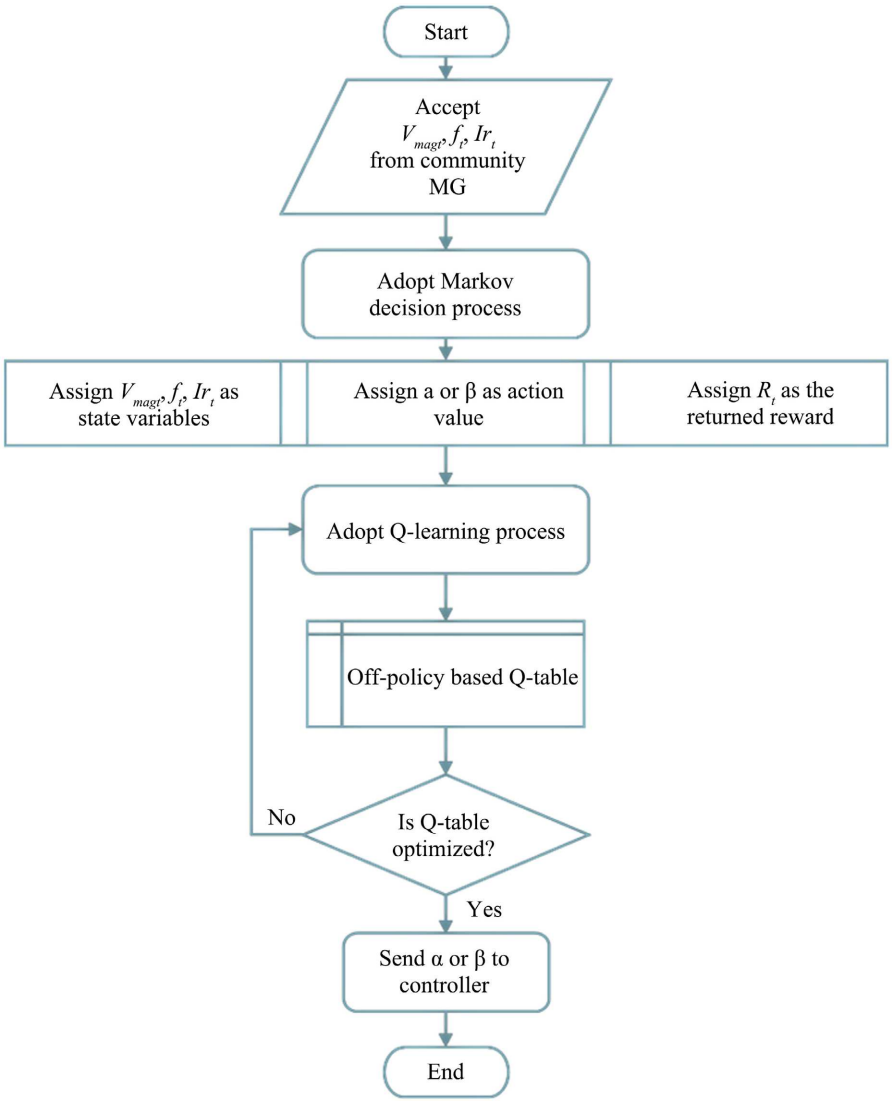

Figure 5. Reinforcement learning-based flowchart in MG control.

controller, the power-sharing, and the reinforcement learning-based control strategy are presented in what follows.

The result for the individual controllers is illustrated in Figure 6. A few critical quantities in Figure 3 are plotted here, which include (from top-left to bottom right of Figure 6 the *d*/*q* components of the VSC's output current (the reference produced by the voltage regulator vs. the actual), the modulation index for the VSC, the VSC's output voltage (before the feedforward term) produced by the current regulator, the frequency at the PCC, the *d*/*q* components of the voltage at the PCC, and the reference voltage produced by droop control. Note that once the developed control mechanism is activated at 4 s, the VSC's output current is well regulated around its reference value (*q* components have a slight deviation). The modulation index and the VSC's output voltage are maintained. The frequency at the PCC varies according to the net load profile (load minus solar power). Further, the *d*/*q* components of the voltage at the PCC are also well regulated. The reference voltage produced by droop control dropped due to the large reactive power consumption of the load.

The result for load sharing of the two distributed energy sources is illustrated in Figure 7. Note that the power output is in p.u., *i.e.*, normalized to the rated MVA of the solar PV systems. Once the developed control mechanism is activated at 4 s, both the real and reactive power production of the two distributed energy sources converges to the same p.u. values, indicating that the two distributed energy sources indeed supply real and reactive power to the load proportionally to their rated capacity. The tight convergence of these curves in Figure 7 corroborates that accurate load sharing is achieved.

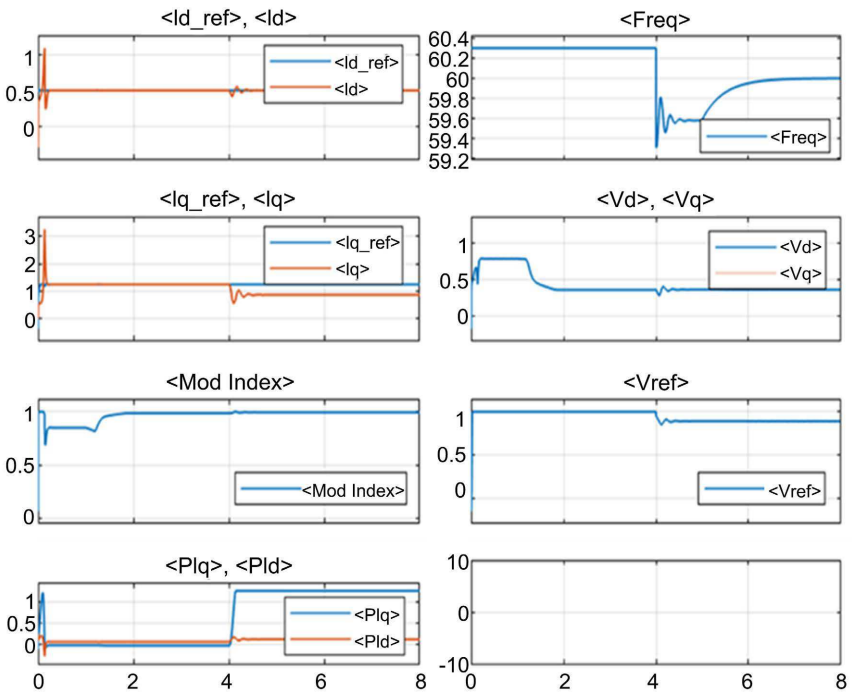

Figure 6. Variables for droop, virtual inertia, voltage/current regulator of the controller for distributed renewable source 1.

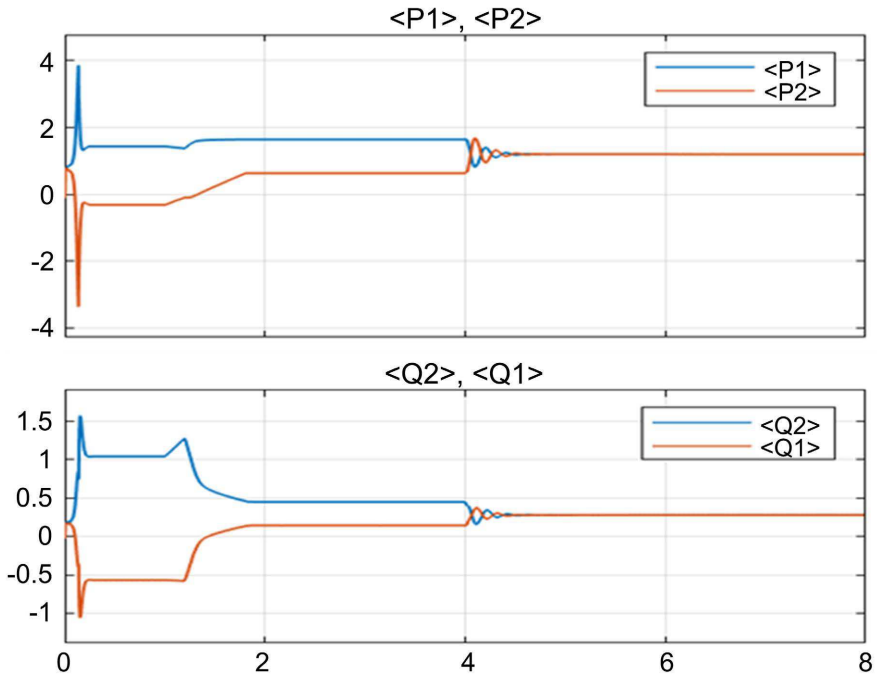

Figure 7. Load sharing (power in p.u.) between the two distributed energy sources.

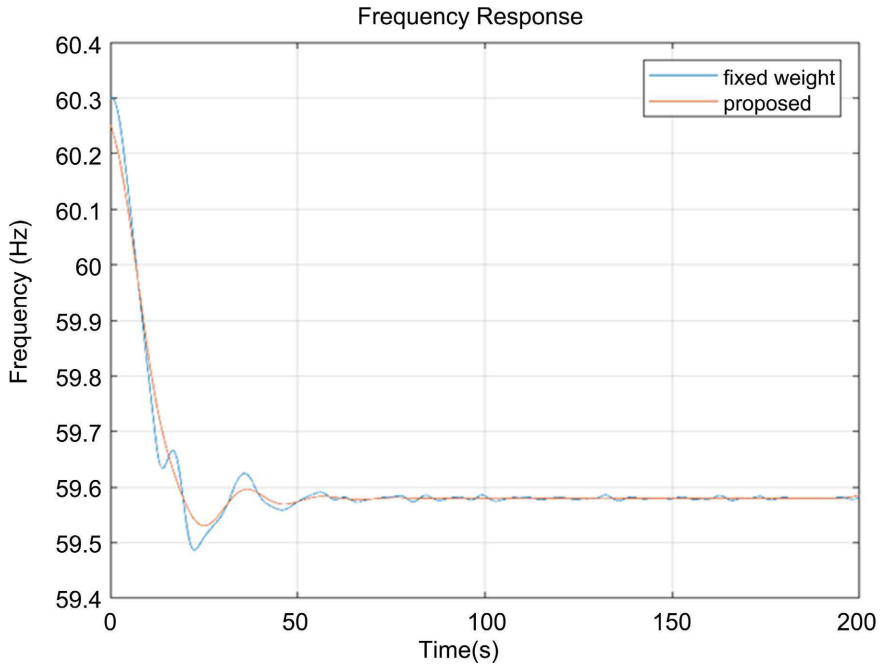

Figure 8. Frequency variation under volatile solar radiation (proposed vs. fixed weights for droop and virtual inertia control).

The result for the frequency response of the community microgrid system under volatile solar radiation is illustrated in **Figure 8**. The proposed control strategy is compared with that of [17], which used fixed equal weights for droop and virtual inertia control. It can be seen from the 200 s episode of the simulation results that the proposed control strategy incurs fewer frequency deviations

when there is a noticeable drop (which is caused by solar radiation ramp down) within the 20 s - 40 s, followed by minimal fluctuation in frequency, which is significantly less than that of fixed weights.

4. Conclusion

This paper develops a novel control strategy for a community microgrid system operating in an islanded mode. The community microgrid equips with two parallel solar PV sources as a power generator and an EV battery as storage. The EV battery serves two manifolds; it provides power for microgrid resiliency and inertia for virtual inertia emulation. The reinforcement learning-based weights adjustment scheme is utilized to adjust the impact of the virtual-inertia emulation and droop control online based on the active power to frequency characteristic. The microgrid model and control mechanisms are implemented in MATLAB/Simulink and set up in OPAL-RT real-time simulation to test the feasibility and effectiveness of the developed control strategy. The simulation results demonstrate robust voltage and frequency regulation, accurate power-sharing, and improved transient response of the community microgrid systems. After the MATALB simulation, the proposed microgrid architecture and the related control have undergone extensive testing in OPAL-RT to verify their validity, and the real-time simulation results indicate that they have many potential uses in the development of the smart grid.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Ton, D.T. and Smith, M.A. (2012) The US Department of Energy's Microgrid Initiative. *The Electricity Journal*, **25**, 84-94. https://doi.org/10.1016/j.tej.2012.09.013

[2] Hirsch, A., Parag, Y. and Guerrero, J. (2018) Microgrids: A Review of Technologies, Key Drivers, and Outstanding Issues. *Renewable and Sustainable Energy Reviews*, **90**, 402-411. https://doi.org/10.1016/j.rser.2018.03.040

[3] Li, Y., *et al.* (2014) Study and Application of Microgrid Energy Management System Based on the Four-Dimensional Energy Management Space. 2014 *International Conference on Power System Technology*, Chengdu, 20-22 October 2014, 3084-3089.

[4] U.S. Energy Information Administration (2019) Annual Energy Outlook 2019 with Projections to 2050. https://www.eia.gov/pressroom/presentations/Capuano_01242019.pdf

[5] Shi, K., Ye, H., Song, W. and Zhou, G. (2018) Virtual Inertia Control Strategy in Microgrid Based on Virtual Synchronous Generator Technology. *IEEE Access*, **6**, 27949-27957. https://doi.org/10.1109/ACCESS.2018.2839737

[6] Carpintero-Rentería, M., Santos-Martín, D. and Guerrero, J.M. (2019) Microgrids Literature Review through a Layers Structure. *Energies*, **12**, Article No. 4381. https://doi.org/10.3390/en12224381

[7] Kafle, L., Zhen, N., Tonkoski, R. and Qiquan, Q. (2016) Frequency Control of Iso-

lated Micro-Grid Using a Droop Control Approach. *IEEE International Conference on Electro Information Technology* (*EIT*), Grand Forks, 19-21 May 2016, 771-775. https://doi.org/10.1109/EIT.2016.7535337

[8] Yuan, Y., Wu, L., Song, W. and Jiang, Z. (2009) Collaborative Control of Microgrid for Emergency Response and Disaster Relief. 2009 *International Conference on Sustainable Power Generation and Supply*, Nanjing, 6-7 April 2009, 1-5. https://doi.org/10.1109/SUPERGEN.2009.5348229

[9] Subburaj, A.S., Pushpakaran, B.N. and Bayne, S.B. (2015) Overview of Grid Connected Renewable Energy Based Battery Projects in USA. *Renewable and Sustainable Energy Reviews*, **45**, 139-234. https://doi.org/10.1016/j.rser.2015.01.052

[10] Koduri, N., Kumar, S. and Udaykumar, R.Y. (2015) On-Board Vehicle-to-Grid (V2G) Integrator for Power Transaction in Smart Grid Environment. 2014 *IEEE International Conference on Computational Intelligence and Computing Research*, Coimbatore, 18-20 December 2014, 1-4. https://doi.org/10.1109/ICCIC.2014.7238404

[11] Mouli, G.R.C., Kaptein, J., Bauer, P. and Zeman, M. (2016) Implementation of Dynamic Charging and V2G Using Chademo and CCS/Combo DC Charging Standard. 2016 *IEEE Transportation Electrification Conference and Expo* (*ITEC*), Dearborn, 27-29 June 2016, 1-6. https://doi.org/10.1109/ITEC.2016.7520271

[12] Li, S., He, H., Chen, Y., Huang, M. and Hu, C. (2015) Optimization between the PV and the Retired EV Battery for the Residential Microgrid Application. *Energy Procedia*, **75**, 1138-1146. https://doi.org/10.1016/j.egypro.2015.07.537

[13] Vasirani, M., Kota, R., Cavalcante, R.L.G., Ossowski, S. and Jennings, N.R. (2013) An Agent-Based Approach to Virtual Power Plants of Wind Power Generators and Electric Vehicles. *IEEE Transactions on Smart Grid*, **4**, 1314-1322. https://doi.org/10.1109/TSG.2013.2259270

[14] Ota, Y., Taniguchi, H., Suzuki, H., Nakajima, T., Baba, J. and Yokoyama, A. (2012) Implementation of Grid-Friendly Charging Scheme to Electric Vehicle Off-Board Charger for V2G. *IEEE PES Innovative Smart Grid Technologies Europe* (*ISGT Europe*), Berlin, 14-17 October 2012, 1-6. https://doi.org/10.1109/ISGTEurope.2012.6465880

[15] Xu, N.Z. and Chung, C.Y. (2016) Reliability Evaluation of Distribution Systems Including Vehicle-to-Home and Vehicle-to-Grid. *IEEE Transactions on Power Systems*, **31**, 759-768. https://doi.org/10.1109/TPWRS.2015.2396524

[16] Vaishnav, S.N. and Krishnaswami, H. (2011) Single-Stage Isolated Bi-Directional Converter Topology Using High Frequency AC Link for Charging and V2G Applications of PHEV. 2011 *IEEE Vehicle Power and Propulsion Conference*, Chicago, 6-9 September 2011, 1-4. https://doi.org/10.1109/VPPC.2011.6043138

[17] Dinkhah, S., Negri, C.A., He, M. and Bayne, S.B. (2019) V2G for Reliable Microgrid Operations: Voltage/Frequency Regulation with Virtual Inertia Emulation. 2019 *IEEE Transportation Electrification Conference and Expo* (*ITEC*), Detroit, 19-21 June 2019, 1-6. https://doi.org/10.1109/ITEC.2019.8790615

[18] Zhong, Q. (2013) Robust Droop Controller for Accurate Proportional Load Sharing Among Inverters Operated in Parallel. *IEEE Transactions on Industrial Electronics*, **60**, 1281-1290. https://doi.org/10.1109/TIE.2011.2146221

[19] Clifton, J. and Laber, E. (2020) Q-Learning: Theory and Applications. *Annual Review of Statistics and Its Application*, **7**, 279-301. https://doi.org/10.1146/annurev-statistics-031219-041220

[20] Jang, B., Kim, M., Harerimana, G. and Kim, J.W. (2019) Q-Learning Algorithms: A

Comprehensive Classification and Applications. *IEEE Access*, **7**, 133653-133667. https://doi.org/10.1109/ACCESS.2019.2941229

[21] West Texas Mesonet. https://www.depts.ttu.edu/nwi/research/facilities/wtm/index.php