ISSN 1330-3651 (Print), ISSN 1848-6339 (Online)
https://doi.org/10.17559/TV-20241029002101
Original scientific paper

Reinforcement Learning Techniques for Real Time Battery Cell Balancing Using Modified Cascaded H-Bridge Converters

M. HEMALATHA*, K. P. SAMPOORNAM

Abstract: This study presents an innovative approach to individual battery cell balancing, focusing on the implementation of a modified cascaded H-Bridge (MCHB) multi-level converter utilizing a reinforcement learning (RL) protocol. In battery systems, particularly those employing extended-lifecycle cells that exhibit intrinsic mismatches, achieving effective cell balancing is crucial for enhancing overall performance and lifespan. The proposed design connects each converter module directly to individual battery cells, facilitating precise balancing without the need for additional complex circuitry. This approach is especially beneficial for applications involving elongated facility-lifecycle grid storage, where optimal performance and reliability are paramount. The research employs an advanced control technique driven by reinforcement learning algorithms to dynamically regulate the interaction between the Battery State of Charge (SoC) and the outputs of the modified cascaded H-Bridge. By continuously adapting to varying operational conditions, the RL protocol ensures that each cell is charged and discharged optimally, minimizing the effects of imbalance and maximizing the overall efficiency of the battery system. The implementation of this technique allows for real-time monitoring and adjustment, significantly improving the response time and accuracy of the balancing process. In the context of grid storage applications, this novel methodology not only enhances the longevity of battery cells but also contributes to more sustainable energy management practices. The experimental results demonstrate that the RL-based MCHB converter significantly outperforms traditional balancing methods, offering a promising solution for future battery management systems. This work paves the way for the development of more efficient energy storage solutions, ultimately supporting the transition to cleaner energy sources and advancing the field of energy technology. Further research will focus on optimizing the RL algorithms and exploring their applicability in various energy storage scenarios.

Keywords: battery cell balancing; cascaded H-bridge; energy storage systems; reinforcement learning; state of charge (SoC)

1 INTRODUCTION

Large quantity storage systems of battery energy (BatESS) facilitate the combination of renewable energy sources by enabling exchange and providing other essential services, including voltage and frequency management [1]. Traditionally, the storage systems of battery energy (BatESS) are connected to the alternating current (AC) grid through distinct platform otherwise double platform direct current to alternating current (DC-AC) converter configurations. To achieve the prime value of the voltage and the current from the battery, the battery system utilizes the series and parallel connecting topology of battery cells [3]. Variations in battery cell properties within a pack are anticipated, even among ostensibly alike cells, due to discrepancies in the engineering process or operational conditions (such as pack temperature fluctuations and age effects) [4].

The battery cell with the smallest amount will be the first to reach the maximum voltage threshold, so terminating the charging process when a series-connected pack of cells is charged. This is because the remaining battery cells will not be fully charged until the battery cell with the bottommost measurements reaches the maximum voltage limit. On the other hand, while the pack is being discharged, the battery cell that has the bottommost measurement will be the one to reach the cutoff voltage first; this will cause the discharge to stop, even if there is still a charge left in the pack. To ensure that the battery pack's available ability is utilized to its fullest potential, a balancing circuit is necessary [5].

In the passive battery cell balancing mode, the resistor network drives away the excess energy in the battery cells; in the active mode, supplementary storage modules such as capacitors or inductors perform the battery energy balancing process, the unit distributes the surplus energy across the battery cells [6]. The battery cell valuation process has been utilizing two methods: the first method will use the battery cell terminal voltages, and the second method will use the state of charging (SoC).

It is important for the cascaded H- Bridge (Cas-HB) converter to be able to balance the voltage of the submodule battery cells because the architecture has many sub module cells whose DC battery voltage changes instantly depending on how the Cas-HB converter is being used. We must equilibrate the submodule cell voltage to effectively regulate the voltage values in the middle of unbalanced grid and load conditions, as well as within a stable system. Consequently, researchers have examined numerous techniques for sub module cell energy balancing under both imbalanced grid and load conditions, as well as in stable systems [7-9]. The delta-connected CasHB converter frequently uses a zero-sequence circulating current for battery voltage and leg energy equilibrium.

Additionally, our aim is to propose an innovative solution to the problem of uneven state of charge (SOC) values among the H-Bridge cell battery units. Machine learning approaches can be employed to balance the output line-to-line voltage of the converter and the state of charge (SOC) of the batteries.The Deep Q-reinforcement learning algorithm is employed to acquire an optimal balancing policy. A penalty is imposed for excessive switching, and the proposed reinforcement learning (RL) method effectively attains equilibrium between balancing and switching. This degree of intricacy cannot be attained using rule-based balancing, as the rules cannot be readily modified to accommodate such intricacies. Our developed reinforcement learning model employs a trial-and-error learning strategy, facilitating the exploration of various strategies and the identification of the ideal course of action in dynamic situations, particularly in light of the evolving optimal strategies influenced by factors such as battery deterioration.

2 RELATED WORKS

A one-phase CasHB hierarchical converter incorporating battery balancing is introduced in [10]. Each four leg converter in the cascaded architecture links a battery pack, and it achieves balance by regulating the duty cycle of each module. The balance relies on observations of terminal voltage. This method has the limitation that voltage measurements fail to yield an accurate calculation of state of charge (SoC) due to voltage errors caused by internal resistance and transient electrochemical processes. Furthermore, the equilibrium within each module's battery pack is not considered.

In [11], the author presents an active power control strategy for individual converter cells in a battery energy storage system using a cascaded H-bridge converter. For dynamically adjusting the zero voltage measures are used to achieve the efficient three phases with the reference methods of balancing, also every battery cell connected to every convert cell. Each module's respective state of charge (SoC) determines its power output. To balance the dynamic changes of the battery cell requires supplementary battery management system.

The different level battery monitoring for Cas-HB was proposed in [12]. Prior instances demonstrate the utilization of the multilevel converter topology to accomplish phase balancing and module balancing. Additional dc-ac inverters achieve pack-level balancing, while a multi-winding transformer achieves whole cell-level balancing. In this the extended kalman filter method based circuitry design was used to measure the battery SoC values.

Several applications, including active battery cell balancing, have extensively researched and effectively used Model Predictive Control (Mod-PC), a sophisticated control methodology [13, 14]. Mod-PC figures out what control actions to take by using a system circuitry with a cost function utilized to find the best order of switch actions that will minimize the effect of cost function over a certain amount of period in the future, substantial computational demands necessary to resolve the Opt-Control-Issues at each period step continue to be a challenge for Mod-PC operation on embedded circuitry. Researchers have explored methods such as explicit Model Predictive Control (MPC) and online fast MPC to reduce the processing time required [15].

The mod-PC based occurrence enabling protocol activates (Omod-PC) the mod-PC program to compute an innovative categorizations of control actions upon fulfilling specified occurrence circumstances; otherwise, it adheres to the anticipated optimal control sequence [11]. In [12], we thoroughly examine the stability of Omod-PC, providing critical insights and approaches to ensure its stability. Lyapunov stability theory establishes the stability of Omod-PC, using the cost function. However, due to the complexity of the trigger's design and calibration, we have constructed and trained a Reinforcement Learning (RL) agent to initiate the mod-PC control action, maximizing its efficacy. This methodology is referred to as RLemod-PC and has been utilized in self-directed automobile path-following challenges as indicated in [18].

In [19], the researcher presents a control approach to equilibrate the state of charge (SOC) of all modules, irrespective of the intermittent conditions of each module's photovoltaic (PV) power. The approach relies on battery state of charge (SOCs), each module's SOC limits, and the overall power fed into the power grid. The battery energy stored quasi-Z source cascaded multilevel inverter (qZS-CMI) photovoltaic power generation system combines the advantages of the qZS inverter, the cascaded multilevel inverter (CMI), and the battery energy storage system. Nonetheless, an imbalanced battery state-of-charge (SOC) across cascaded H-bridge inverter modules will impair the overall system performance and reduce the battery lifespan. The proposed control approach is supported by simulations and experiments. It ensures that the battery energy storage in the qZS-CMI photovoltaic system has a uniform state of charge (SOC).

Recently, machine learning (ML) methodologies have been suggested for multiple facets of energy storage systems, including battery systems [20, 21]. Generally, machine learning is employed for either predictive analytics or prescriptive analytics. Predictive analytics seeks to anticipate future outcomes or patterns utilizing past data, whereas prescriptive analytics endeavors to recommend ideal actions or decisions through data and optimization methodologies. Within battery management systems, the predominant uses of machine learning are predictive, including the forecasting of state of charge (SoC) [22], state of health (SoH) [23], battery safety [24], and load. Reinforcement learning in predictive analysis has demonstrated encouraging outcomes in enhancing the performance, dependability, and safety of battery systems [25].

To summarize, additional circuitry and novel control methodologies are always required in order to provide monitoring and dynamic adjusting at the distinct cell level. Applications must weigh the supplementary dimensions utilization and reliability enhancements expanded from this approach against the cost of the additional circuitry. One of the benefits of the modified Cas-HB different level converter offers comparable controllability, in addition to having the capability to convert DC to AC built right in. Employing larger, high-value cells generally enhances the appeal of all these technologies.

3 PROPOSED WORK

We present an illustration of the conventional topology of power loading models here. Fig. 1 depicts the topology of the three-phase, two-level converter. It is fairly common to employ this topology, and the control algorithm has reached a very advanced stage. However, connecting the system to the intermediate or high-voltage power grid necessitates the inclusion of a massive transformer. Not enough advanced output voltage from the converter causes this. Fig. 2 and Fig. 3 depict two additional multilevel converter topologies that are frequently used. When the output voltage is the same, as shown in Fig. 1, these two topologies have the potential to diminish the voltage strain that a changing device experiences. However, these two topologies also encounter the previously discussed challenge; they typically require a high-voltage power grid to connect to.

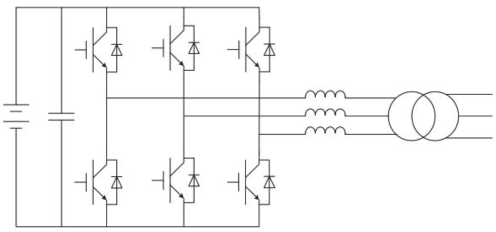

Figure 1 Three phase two level converter topology

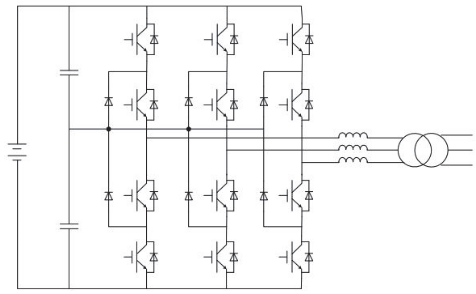

Figure 2 Neutral point clamped multi-level converter topology

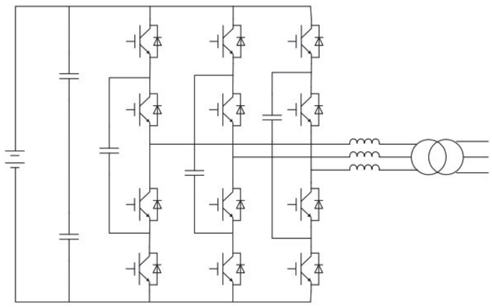

Figure 3 Flying capacitor multi-level converter topology

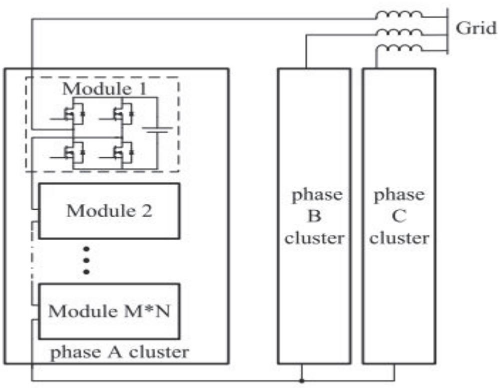

Figure 4 Cascaded H-bridge multilevel converter BESS

Fig. 4 illustrates a BESS approach that utilizes a Cas-HB multilevel converter. This methodology utilize the independent connection of every converter to individual battery cell. This architecture allows for discrete monitoring and watch dog of each electrochemical module, eliminating the need for additional balancing circuits. By replacement of entire cell system to single cell circuitry, we can control cell failure online, significantly reducing the system's dependence on cell reliability. In conclusion, we leverage the flexibility of the multilevel architecture to perform cell balancing through the use of RL based interplay with the output voltage and battery cell SoC.

The main focus of the Bat-ESS strategy is to measure close to idle characteristics of the battery module with the references of SoC and heath of the battery parameters. The SoC value validating the battery module ability, although the health of the battery parameter (SOH) indicates the maximum value of battery charge towards the mentioned capacity. Fig. 5 illustrates the associations in the middle of SoC and SoH. A battery cell cannot directly measure SoC and SoH parameters. Instead, we estimate those using measurable parameters like voltage, current, and cell temperature. We schedule the Enable/Disable modes for the batteries based on the conditions of the battery cells to balance the health of all batteries. This results in an extension of the battery pack's lifetime. As a result, the truthful measurements of SoC and SoH, in conjunction with the forecasting of battery module replacements, is required to augment a BESS's recital.

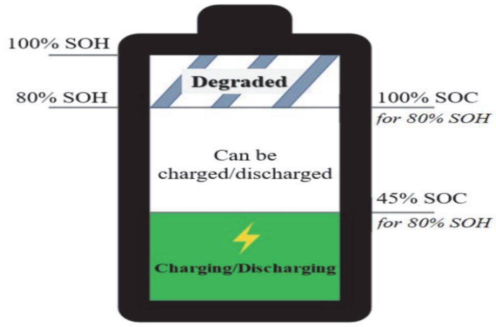

Figure 5 The relationship between SoC and SoH

The author of [23] presented a deep learning technique for virtual state of charge (SoC) measurements, but it ignores health of the battery module. The authors employed a neural network to assess the State of Health (SoH) using empirical datasets [24]. Data on both the state of charge (SoC) and state of health (SoH) of battery is essential for effective battery pack scheduling to enhance Battery Energy Storage System (BESS) performance. In [25], researchers used the reduction in State of Health (SOH) to determine Electrochemical Impedance Spectroscopy (EIS) parameters; however, they only made offline revisions to the SOH. The authors of [26] used fuzzy logic control to change the number of battery cells in a design based on load demand. They did this to lower the loop current and keep the battery from being inconsistent. In [27], researchers observed a decrease in battery resistance to identify feeble cells and isolate them from the battery set. This method addressed the problem of inconsistent characteristics, but it necessitates a sophisticated measurement device or imposes a significant processing load. The introduction of the weighted-k round-robin (kRR) scheduling scheme in [28, 29] aimed to extend the battery pack's lifespan by considering load demand and state of health (SOH) degradation. However, one can only implement kRR-based scheduling for a static model, which requires a constant number of cells or battery types in the battery pack.

A forecast method is essential to optimize the lifespan of a battery system, which consists of parallel or serially connected battery cells with either heterogeneous or homogeneous states of health. Focus is on maximizing battery lifetime by mitigating the reduction in the State of Health (SoH) of a battery pack through the minimization factor of SoC imbalances among individual battery cells.

I. Preliminaries on Reinforcement Learning (RL)

The reinforcement learning paradigm [29] functions inside a state action framework, wherein at time step t, the agent observes the environment's state, s_t and then selects an action, at, to execute. Upon the execution of the action, the agent is getting the reward multiplied by scalar value with respect to the reward function, represented as $1(s_t, a_t)$. The focus of this agent work is to attain the finest policy $\pi_p^*: s_1 \rightarrow a_1$, this process maximizes the foreseen of aggregate future rewards; this can be articulated as:

$$G_r = E_{\pi p}\left[\sum_{t=0}^{\infty} \gamma^t r_t\right] \tag{1}$$

where $r_{tp} = r_1\left(s_{t_1}, a_{t_1}\right)$, and the value of scalar $\gamma_s \in [0, 1]$ represents the cut rate factor that diminishes the significance of anticipated forthcoming rewards. To ascertain the finest policy πp^*, the agent engages with the environs to determine the appropriate activities to do at a specific state s_t, aiming to maximize the anticipated aggregate reward G.

The state value function $V^\pi(s)$ quantifies the worth of the agent in state s under policy π, representing the anticipated return from state s while adhering to π. This can be articulated as:

$$V^\pi(s) = E_\pi\left[\sum_{t=0}^{\infty} \gamma^t r_{t+k} \,|s_t = s\right] \tag{2}$$

The value function of state and action $Q^{\pi p}(s_1, a_1)$ for a state and action pair (s_1, a_1) is defined as the anticipated profit commencing from s_1, followed by action a_1 and subsequently policy πp. This can be articulated as:

$$Q^{\pi p}(s_1, a_1) = E_{\pi p}\left[\sum_{t=0}^{\infty} \gamma^t r_{t+k} \,|s_{t1} = s_1, a_{t1} = a_1\right] \tag{3}$$

The ideal Q function $Q_p^*(s_1, a_1) = \max_{\pi p} Q^{\pi p}(s_1, a_1)$ denotes the Q function under a strategy that aims to optimize the Q value. This policy π^* may be delineated as:

$$\pi p^*(s_1) = arg \max_{a_1} Q_p^*(s_1, a_1) \tag{4}$$

The finest policy πp^* can be determined by learning the optimal Q function $Q_p^*(s_1, a_1)$. The aim of reinforcement learning is to acquire $Q_p^*(s_1, a_1)$, commonly known as Q learning. This strategy enables the Ri-L agent to learn the ideal Q function solely through interaction with the environs. Significantly, the agent can acquire the optimal policy without necessitating a model of the system dynamics.

II. Deep Neural Network based – RL.

The Ri-L agent necessitates a mechanism to simplify the state and action significance function across all states and actions to determine an action. This application of reinforcement learning is termed Deep Q Learning due to the deep neural network modeling the state and action significance function, $Q_p(s_1, a_1)$. Thus, the Deep-NN can be more accurately characterized as a Deep Q Network (DeQNe).

The DeQNe was trained through twofold Q learning and batch update techniques. In every update value, a group of proficiencies $\left(s_{i1}, a_{i1}, r_{i1}, s_{i1}'\right)$ of size M_p is extracted commencing the memory for use. Twofold Q learning employs 2 distinct collections of linkage weights and biases, represented by ϕ_q for the critic linkage parameters and ϕ_{tq} for the goal linkage parameters. Reliant on the situation, ϕ_q is utilized to ascertain the subsequent action with a chance of value $1- \in 1$. Alternatively, a random action is chosen. Subsequent to the execution of the action, ϕ is revised in accordance with the target values y_i. The objective is established using double Q-learning and batch techniques.

$$a_{\max} = arg \max_{a'} Q(s'_i, a', \phi) \tag{5}$$

$$\gamma_i = r_i + \gamma Q_t(s'_i, a_{\max}, \phi_t) \tag{6}$$

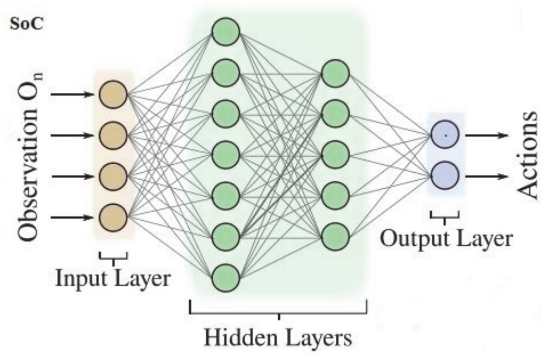

Figure 6 Deep reinforcement learning basic layers

Fig. 6 comprises the present topology, with the total number of sixteen cell state-of-charge (SoC) values at the closing value of the preceding control period, and maximum number of statistical characteristics of the cells' SoC. In every hidden layer that is utilizing 10% dropout acceptable scenario, indicating that each node has a 0.1 probability of being omitted during each training period. The current ReLe model has a double layer of hidden nodes, these nodes utilizing the 64 and 32 neurons, also the metrics of loss function, cross entropy and assigning the learning rate value of 0.00017 at that time of training process.

In twofold Q learning, ϕ_q is employed to select the action $a_{\max V}$ that ϕ_{tq} will be utilized to ascertain the goal. The goal values y_i are incorporated into a loss function that contrasts the observed target value with the current estimate of the detractor network.

$$L_q = \frac{1}{2M_q} \sum_{i=1}^{M_q} \left(y_{iq} - Q\left(s_{iq}, a_{iq}, \phi_q\right)\right)^2 \tag{7}$$

The updated weights values and biases values ϕ_p are subsequently configured in the behavior network for the agent to determine the subsequent action. Following a predetermined quantity of training time steps, the parameters of the target linkage ϕ_{tq} are aligned with the parameters of the behavior network ϕ_p to enhance learning reliability and mitigate Q value overestimation. Techniques like epsilon-greedy might compel the agent to explore prior to acquiring substantial experience in the environment, thereby facilitating a gradual shift towards exploitation once sufficient experience has been amassed to reliably get high rewards.

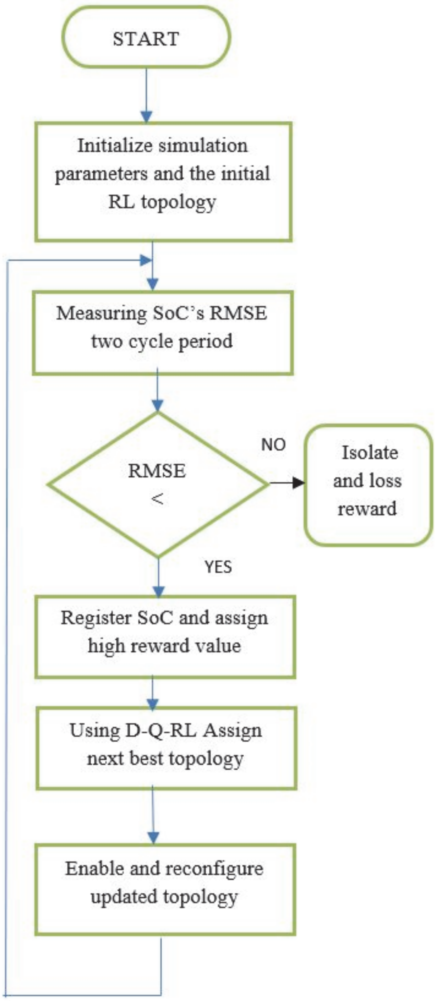

Figure 7 Proposed algorithm flow chat

Algorithm 1: D-Q-RL based Switches Scheduling

Input: Root RMS State Vector r_{tp} and η
Output: Optimal Action of Scheduling
Initialize the value function $V_\pi(s)$ under the policy π_p
Add $\left\langle r_{tb}\left\langle t_p-1\right\rangle, r_p\left\langle t_r-1\right\rangle, V_\pi\left(\pi_p-1\right)\right\rangle$ into η
Construct the Load value $\mathbb{Q}_L$ and source cell count $\mathbb{S}_t$
Initialize $\mathbb{Q}_L$ and $\mathbb{S}_t$ with random weights
Perform RL to minimize loss reward action
if $r_{tp} \in \eta$ then
    Select value action $V_\pi(s)$ using Eq. (2)
else
    Select value action $V_\pi(s)$ randomly
end if
Calculate instantaneous reward using Eq. (4)
Calculate accumulative reward using Eq. (6)

In summary, the design of traditional event-trigger policies is complex because it must satisfy both data throughput and controller performance criteria. To overcome this difficulty, the SoC parameter must be developed by training a deep Q linkage Reinforcement Learning representative to determine the appropriate SoC and current path parameter policy. We propose this control method mentioned in Fig. 7 and Algorithm 1 an active-cell balancing controller to extend the state of charge (SoC) of a battery cell.

4 RESULT AND DISCUSSION

Fig. 8 illustrates the Root-MSE with the reference in the middle of the observed terminal voltage and the predicted terminal voltage. The Root-MSE between the observed and estimated terminal voltage values for each cell is approximately 0.015 V and remains consistently diminished towards the time value. The minimal discrepancy in the middle of the measured and anticipated terminal voltages indicates that the algorithms effectively models terminal voltage, resulting in a more precise estimation of the state of charge (SOC).

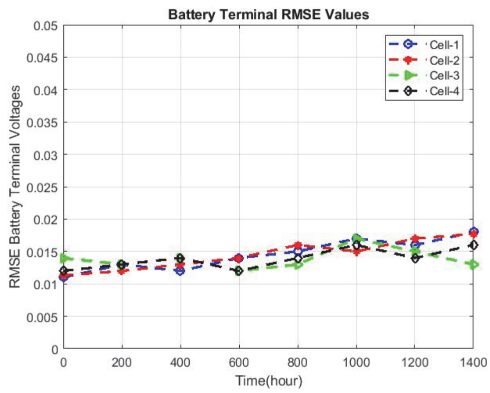

Figure 8 Measured and projected RMS terminal voltage values of different cells

The simulation results of without machine learning based balancing method are depicted in Fig. 9, where the x-axis denotes time in seconds. This duration represents the operational time of the battery cells prior to its overall state of charge (SoC) falling below a predetermined threshold of 10%.

Fig. 10 shows the runtime of the battery pack for a ML based balancing has increased from 14,400 seconds to 17600 seconds, representing a 23.4% enhancement. Moreover, in comparison to conventional methods under without ML, the outcomes have improved by 11.3%. Tab. 1 shows the proposed comparision algorithm.

Table 1 Comparison of proposed design with convention methods

Algorithms	Battery run time efficiency improvements / %	Dynamic Workload condition	Mean usable capacity
OCV	19.80	Unstable	2015 mAh
ARS	20.14	Partially stable	2160 mAh
Proposed D-Q-RL	22.63	Stable	2316 mAh

OCV-Open Circuit Voltage, ARS- Augmented open search

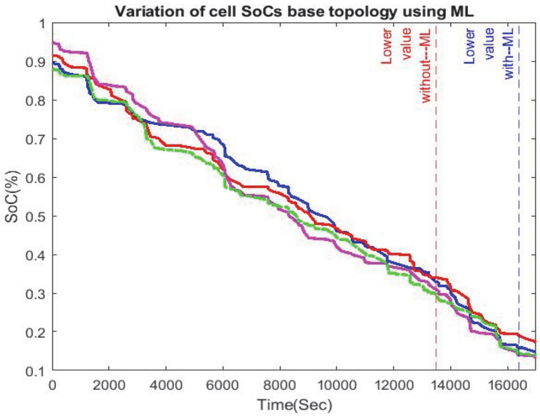

Figure 9 Disparities of cell SoC base topology without Machine Learning balancing

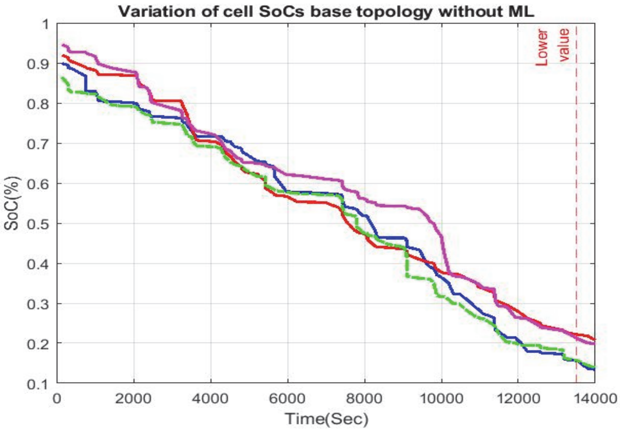

Figure 10 Battery cell SoC base topology using Machine Learning (ML) balancing

Fig. 11, Fig. 12 and Fig. 13 shows the battery charging and discharging in balanced and unbalanced conditions.

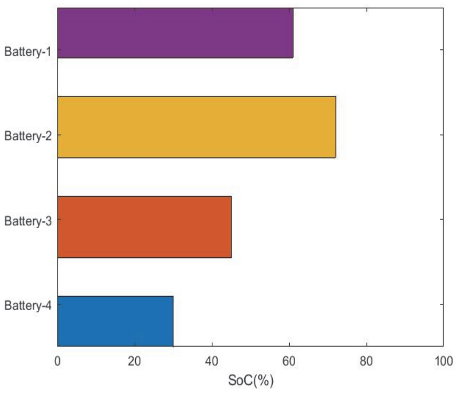

Figure 11 Initial state of the battery cells is unbalanced

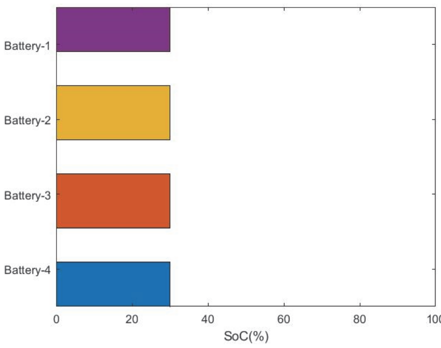

Figure 12 Discharging balanced using proposed algorithm

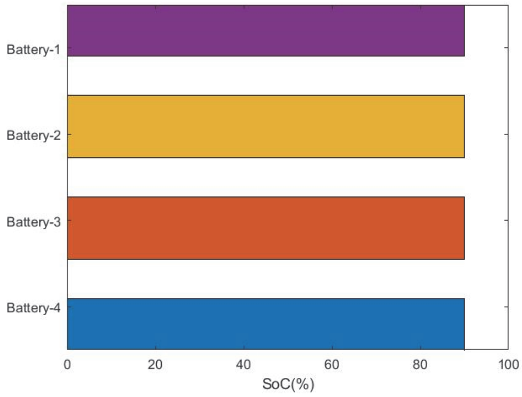

Figure 13 Battery charging balanced using proposed algorithm

5 CONCLUSIONS AND FUTURE ENHANCEMENTS

The primary objective of this research is to minimize the disparity in the state of charge (SoC) among individual battery cells within a network, thereby maintaining a uniform SoC across all cells at any given moment. This uniformity is crucial for maximizing the overall performance and lifespan of the battery energy storage system. Our preliminary studies demonstrate that the machine learning (ML) based topology switching algorithm enhances battery cell equalization significantly more than traditional methods that rely on a single optimal topology. In this effort, we will investigate various machine learning and reinforcement learning (RL) models to determine the most effective topology for the next control period. Our approach involves developing comprehensive datasets that include a wide range of input features and corresponding labels to facilitate the training of these models. By employing advanced ML and RL techniques, we aim to dynamically adapt the system to changing conditions, improving the responsiveness and efficiency of the cell balancing process.The implications of our research extend beyond battery management; they contribute to the broader field of renewable energy integration and grid stability. Enhanced cell balancing not only leads to more efficient energy utilization but also supports the reliable operation of energy storage systems, which are critical in facilitating the transition to cleaner energy sources.

Future work will focus on refining the selected models and validating their performance in real-world scenarios, ensuring that our approach is practical and scalable. Through these efforts, we aspire to contribute to the development of more resilient and efficient energy storage solutions, ultimately advancing the capabilities of battery systems in modern energy applications. Concerning future endeavors, the implementation of grid connected technologies necessitates the integration of energy storage systems comprising numerous battery cells arranged in parallel or series within a Battery-ESS. In these systems, D-Q-RL based battery management algorithms may exhibit constrained performance owing to the high-dimensional state space. We will explore a distributed reinforcement learning methodology to address the shortcomings of centralized methods for extensive energy storage systems. An experimental setup will be established to examine the effects of the battery management algorithm on actual systems.

6 REFERENCES

[1] Suruli, K. & Ila, V. (2020). Energy management strategy using ANFIS approach for hybrid power system. *Tehnički vjesnik, 27*(2), 567-575. https://doi.org/10.17559/TV-20190722152008

[2] Iqbal, M., Benmouna, A., Becherif, M., &Mekhilef, S. (2023). Survey on battery technologies a nd modeling methods for electric vehicles. *Batteries, 9*(3), 185. https://doi.org/10.3390/batteries9030185

[3] Tran, M. K., Panchal, S., Khang, T. D., Panchal, K., Fraser, R., & Fowler, M. (2022). Concept review of a cloud-based smart battery management system for lithium-ion batteries: feasibility, logistics, and functionality. *Batteries, 8*(2), 19. https://doi.org/10.3390/batteries8020019

[4] Fang, K., Li, C., Tang, Y., He, J., & Song, J. (2022). China's pathways to peak carbon emissions: new insights from various industrial sectors. *Applied Energy, 306*, 118027. https://doi.org/10.1016/j.apenergy.2021.118039

[5] Abhyankar, N., Mohanty, P., & Phadke, A. (2021). Illustrative Strategies for the United States to Achieve 50% Emissions Reduction by 2030. *Lawrence Berkeley National Laboratory.*

[6] Mock, P. & Díaz, S. (2021). Pathways to decarbonization: the European passenger car market in the years 2021-2035. *Communications, 49*, 105196. https://doi.org/10.1016/j.renene.2020.09.029

[7] Asoodar, M., Nahalparvari, M., Zhang, Y., Danielsson, C., Nee, H., & Blaabjerg, F. (2022). Accurate Condition Monitoring of Semiconductor Devices in Cascaded H-Bridge Modular Multilevel Converters. *IEEE Transactions on Power Electronics, 38*(3), 3870-3884. https://doi.org/10.1109/TPEL.2022.3221285

[8] Nami, A., Liang, J., Dijkhuizen, F., & Demetriades, G. D. (2014). Modular Multilevel Converters for HVDC Applications: Review on Converter Cells and Functionalities. *IEEE Transactions on Power Electronics, 30*(1), 18-36. https://doi.org/10.1109/TPEL.2014.2327641

[9] Choudhury, S., Bajaj, M., Dash, T., Kamel, S., & Jurado, F. (2021). Multilevel Inverter: A Survey on Classical and Advanced Topologies, Control Schemes, Applications to Power System and Future Prospects. *Energies, 14*(18), 5773. https://doi.org/10.3390/en14185773

[10] Ranjan, M. & Shankar, R. (2021). A literature survey on load frequency control considering renewable energy integration in power system: Recent trends and future prospects. *Journal of Energy Storage, 45*, 103717. https://doi.org/10.1016/j.est.2021.103717

[11] Adebanji, B., Atoki, O., Fasina, T., Adetan, O., & Abe, A. (2021). Comparative study of off-grid and grid-connected hybrid power system: issues, future prospects and policy framework. *Indonesian Journal of Electrical Engineering and Computer Science, 22*(2), 752. https://doi.org/10.11591/ijeecs.v22.i2.pp752-759

[12] Razmjooei, H., Palli, G., Abdi, E., Terzo, M., & Strano, S. (2023). Design and experimental validation of an adaptive fast-finite-time observer on uncertain electro-hydraulic systems. *Control Engineering Practice, 131*, 105391. https://doi.org/10.1016/j.conengprac.2022.105391

[13] Chen, J. & Behal, A. (2023). Active battery cell balancing by real-time model predictive control for extending electric vehicle driving range. *IEEE Transactions on Automation Science and Engineering.* https://doi.org/10.1109/TASE.2023.3291679

[14] Zhang, Y., Huang, Y., Chen, Z., Li, G., & Liu, Y. (2021). A novel learning-based model predictive control strategy for plug-in hybrid electric vehicle. *IEEE Transactions on Transportation Electrification, 8*, 23-35. https://doi.org/10.1109/TTE.2021.3069924

[15] Badawi, R. & Chen, J. (2022). Performance evaluation of event-triggered model predictive control for boost converter. *Proceedings of the 2022 IEEE Vehicle Power and Propulsion Conference, Merced, CA, USA*, 1-4. https://doi.org/10.1109/VPPC55846.2022.10003282

[16] Daowd, M., Omar, N., Van Den Bossche, P., & Van Mierlo, J. (2011). Passive and active battery balancing comparison based on MATLAB simulation. *Proceedings of the 2011 IEEE Vehicle Power and Propulsion Conference, Chicago, IL, USA, 6-9*, 1-7. https://doi.org/10.1109/VPPC.2011.6043010

[17] Shang, Y., Cui, N., & Zhang, C. (2018). An optimized any-cell-to-any-cell equalizer based on coupled half-bridge converters for series connected battery strings. *IEEE Transactions on Power Electronics, 34*, 8831-8841. https://doi.org/10.1109/TPEL.2018.2888514

[18] Wang, L. Y., Wang, C., Yin, G., Lin, F., Polis, M. P., Zhang, C., & Jiang, J. (2015). Balanced control strategies for interconnected heterogeneous battery systems. *IEEE Transactions on Sustainable Energy, 7*, 189-199. https://doi.org/10.1109/TSTE.2015.2487223

[19] Ge, B., Liu, Y., Abu-Rub, H., & Peng, F. Z. (2017). State-of-Charge Balancing Control for a Battery-Energy-Stored Quasi-Z-Source Cascaded-Multilevel-Inverter-Based Photovoltaic Power System. *IEEE Transactions on Industrial Electronics, 65*(3), 2268-2279. https://doi.org/10.1109/TIE.2017.2745406

[20] Xu, J., Mei, X., Wang, H., Shi, H., Sun, Z., & Zou, Z. (2022). A model based balancing system for battery energy storage systems. *Journal of Energy Storage,* 49, 104114. https://doi.org/10.1016/j.est.2022.104114

[21] Gao, T. & Lu, W. (2021). Machine learning toward advanced energy storage devices and systems. *iScience, 24*(1), 101936. https://doi.org/10.1016/j.isci.2021.101936

[22] Elabbassi, I., Khala, M., El yanboiy, N., Eloutassi, O., & El hassouani, Y. (2024). Evaluating and comparing machine learning approaches for effective decision making in renewable microgrid systems. *Results Engineering, 21*, 101888. https://doi.org/10.1016/j.rineng.2023.101888

[23] Hassan, Q., Algburi, S., Sameen, A. Z., Salman, H. M., & Jaszczur, M. (2023). A review of hybrid renewable energy systems: solar and wind-powered solutions: challenges, opportunities, and policy implications. *Results Engineering*. https://doi.org/10.1016/j.rineng.2023.101621

[24] Abdullah, A. S., Joseph, A., Kandeal, A. W., Alawee, W. H., Peng, G., Thakur, A. K., & Sharshir, S. W. (2024). Application of machine learning modeling in prediction of solar still performance: a comprehensive survey. *Results Engineering.* https://doi.org/10.1016/j.rineng.2023.101800

[25] Lu, C., Chen, J., Chen, C., Huang, Y., & Xuan, D. (2023). An active equalization method for redundant battery based on deep reinforcement learning. *Measurement, 210*, 112507. https://doi.org/10.1016/j.measurement.2023.112507

[26] Yang, Y., Zhao, L., Yu, Q., Liu, S., Zhou, G., & Shen, W. (2023). State of charge estimation for lithium-ion batteries based on cross-domain transfer learning with feedback mechanism. *Journal of Energy Storage, 70*, 108037. https://doi.org/10.1016/j.est.2023.108037

[27] Xiong, X., Wang, Y., Li, K., & Chen, Z. (2023). State of health estimation for lithium-ion batteries using Gaussian process regression-based data reconstruction method during random charging process. *Journal of Energy Storage, 72*, 108390. https://doi.org/10.1016/j.est.2023.108390

[28] Leong, A. S., Ramaswamy, A., Quevedo, E. D., Karl, H., & Shi, L. (2020). Deep reinforcement learning for wireless sensor scheduling in cyber–physical systems. *Automatica, 113*, 108759. https://doi.org/10.1016/j.automatica.2020.108759

[29] Sutton, R. S. & Barto, A. G. (2018). Reinforcement Learning: An Introduction. *MIT Press*.

Contact information:

M. HEMALATHA, Research Scholar
(Corresponding Author)
Department of Electrical and Electronics Engineering,
Bannari Amman Institute of Technology, India
E-mail: engineerhemalatha@yahoo.com

K. P. SAMPOORNAM, PhD, Professor
Department of Electronics and Communication Engineering,
Bannari Amman Institute of Technology, India
E-mail: sampoornam@bitsathy.ac.in