EP-JOURNALS

GLOBAL JOURNAL OF ENGINEERING AND TECHNOLOGY REVIEW

Available on: http://gjetr.ep-journals.org/ E-ISSN: 3051-3782, P-ISSN: 3051-3774

Global Journal of Engineering and Technology Review

Volume: 01 Issue: 04 December 2025

A Multi-Layer Reinforcement Learning and Genetic Algorithm Control Scheme for Grid Stability and Dispatch Control in PV–Fuel Cell Dominated Distribution Microgrids

Adel Elgammal

Professor, Utilities and Sustainable Engineering, The University of Trinidad & Tobago UTT

Published Date: December 13, 2025 **Article DOI:10.65150/EP-gjetr/V1E4/2025-06**

ABSTRACT: The recent proliferation of photovoltaic (PV) based generation and hydrogen-based fuel cell (FC) systems in smart distribution networks pose great operational challenges, such as intermittency, voltage instability, frequency deviations, and suboptimal dispatch efficiency. Classical deterministic or rule-based control algorithms are not capable of effectively dealing with the nonlinear, stochastic and highly interconnected behaviors of PV–FC based MG systems. In this paper, a new multi-layer intelligent control framework, which combines RL for real-time scenario adaptive decision-making and GA to achieve the multi-objective optimization on energy dispatch and grid stability indices is presented. The upper RL is capable of updating the optimal control policies based on an online interaction with the microgrid environment, considering the complexities such as irradiance variations, uncertainties in loads, hydrogen storage dynamics and FC ramping constraints. There, the lower GA layer coordinates PV generation with FC output and battery state of charge trajectories, subject to optimal grid-supportive actions such as reactive power compensation and harmonic mitigation.

An extensive MATLAB/Simulink-Python co-simulation framework was implemented to study the proposed topology in the presence of fluctuating renewable profiles, sudden load changes, FC fuel starvation operation modes and islanded/connected configurations. Simulation results show that the RL–GA hybrid controller is in much better performance compared with conventional PI, model predictive control (MPC) and sole RL approaches. Significant results were found such as a reduction in the variation of 35–48%, an improvement in frequency stability of 28%, up to 22% reduction is hydrogen consumption, and an increase in dispatch efficiency ranging from 30 to 55% subject to operational constraints and extension of FC Stack life. Furthermore, the OC system is robust to prediction errors, and to communication delays as well as component uncertainties. Results demonstrate the promising prospects of scalable hierarchical AI-based control and operation, for stable, efficient, and autonomous performance in renewable-driven microgrids. This study paves the way for the development of intelligent controllers to use in distributed energy systems with potential high penetration levels of renewables and hydrogen-integrated configurations.

KEYWORDS: Reinforcement Learning (RL); Genetic Algorithm Multi-Objective Optimization (MOGA); PV–Fuel Cell Microgrids; Dispatch Optimization; Grid Stability; Voltage and Frequency Regulation;; Hydrogen Energy Systems; Smart Microgrid Management.

I. INTRODUCTION

Fast expansion of renewable energy sources in modern distribution systems (especially PV system and hydrogen-based fuel cell), has brought more complex operation and stability difficulties which cannot be better handled by deterministic or model-based control approach. The variability of PV power, the nonlinear and slow dynamics of FC stacks, and the higher reliance on voltage/frequency disturbances increase nowadays the interest in adaptive, data-driven, and optimization-based control schemes. It consolidates the progression of research in PV–fuel cell hybrid microgrids, stability issues on grid with a high penetration level of renewables, traditional and modern intelligent control scheme, RL, GA and recent hybrid RL–GA model as well as gaps to stimulate multi-level intelligent control architecture presented in this paper. Hybrid PV–FC microgrids have become attractive owing to their complementary operational profiles. Day time, when PV system capacity is the highest, there is plenty of power but very high intermittency and lack for dispatchability. Fuel cells, particularly proton exchange membrane fuel cells (PEM-FCs), offer a clean way of delivering dispatchable power but are characterized by slow dynamics and degradation sensitivities that are ramp-rate dependent. Initial research on PV–FC hybridization highlights the ability of hydrogen storage to renewable variability and provide for islanded operation [1], [2]. However, researchers consistently observe about that the integration of PV and FC systems increases significantly complexity control; fast real-time coordination is necessary to balance fast PV variability with the slow electrochemical response of the FC and hydrogen consume limits.The recent literature emphasizes simultaneous progress in PV–fuel cell hybrids, stability challenges of low inertia systems, smart control strategies, reinforcement learning (RL), evolutionary optimizer and hybrid RL–GA solutions—thanks to which intelligent multi-level controller structure under discussion can be motivated.

Hybrid PV–FC microgrids have recently been considered as they complement each other to compensate for the operating features. PV generation is high but highly fluctuating and non-firm during daylight periods of peak irradiance. Fuel cells, especially proton exchange membrane FCs (PEM-FCs), provide clean dispatchable energy but have slow time response to step changes and generally suffer degradation when exposed to aggressive dynamics [3], [4]. In the initial studies on PV–FC hybridization it was found that buffering hydrogen will certainly facilitate renewable smoothing and islanded operation but in later researches it was reported that real-time control complexity becomes significant more due to this integration [5]. There is a need on fast coordination to match fast PV variation under slow FC electrochemical response and hydrogen storage constraints. Rahman et al. (2021) demonstrated stable operation of hydrogen-assisted hybrid systems during transients, only when the intelligent supervisory control to prevent starvation and fast degradation of FC arrays is in place [6]. Grid stability issues in renewable-rich µGs are well explored over 2020-2025. PV with high penetration, also reduces system inertia, increases voltage rise, increases reverse power and causes fast frequency oscillations due to lack of mechanical inertia in the inverter-based sources [7]. Very recent work has demonstrated that inverter-based grids need ultrafast voltage and frequency regulation to remain stable [8]. It gets worse with fuel cells: the stacks cannot ramp up production in real time without dehydrating the membrane or starving the stack of fuel, both of which can make voltage sags during PV fluctuations worse [9]. Despite the progress of grid-forming inverters and VSM control, droop-based methods are still not able to cope with extreme intermittency or weak-grid scenarios [10], [11]. Thus, there is an art-acknowledged need for smart controllers to respond dynamically/flexibly in accordance with renewable operating conditions varying therein.

Traditional microgrid control methods, such as droop control [6], model predictive control (MPC) [7] and rule-based energy management system (EMS) could have been considered extensively with some specified limitations. Droop control is simple and decentralized, but works in an inconsistent manner under rapid change of renewable resources and non-linearity system dynamics [12]. Adaptive droop adjustments and synthetic inertia enhance stability but are parameterized in an environment-sensitive manner. Model predictive control (MPC) shows good potential in prediction ability and crapability of constraint handling a [13], but its heavy dependence on accurate models and high computational burden make it difficult to be applied for real time operation in hybrid PV–FC systems, where model uncertainties grow rapidly. On the other hand, rule-based EMS continues to excel in simplicity but does not handle multi-objective trade-offs or optimal dispatch under uncertainty [14]. With increased penetration of renewables, industry and research have moved towards machine learning (ML), reinforcement learning (RL) based and hybrid AI-optimization schemes.

Reinforcement-learning methods demonstrated good performance in nonlinear, time-variant, and uncertainty microgrids. RL agents acquire optimal control policies through the environment interaction, suitable to renewable variability and load uncertainty as well as nonlinearities associated with hydrogen storage. Recent research shows that RL may work well for frequency regulation, voltage control and energy dispatch of a microgrid [15]. For instance, RL-based frequency controllers achieve better dynamic performance compared with droop control against PV-induced fluctuations thanks to their adaptability [16]. In addition, deep reinforcement learning (DRL) algorithm DQN, TD3, PPO for maximum power point tracking (MPPT), battery scheduling, multi-energy coordination has been used [17], [18]. Actor–Critic can also be used for long-horizon energy storage planning and lifetime-aware dispatch [19]. However, despite the promise of these methods RL suffers slow convergence, unstable learning, dependence on carefully engineered rewards and failures due to catastrophic forgetting under rare but extreme perturbations. Stable microgrid operation can be guaranteed using hybrid methods integrating RL adaptivity and optimization based supervisory controllers. Genetic algorithms (GA) and evolutionary optimization are still in wide practice for MO problems with control, sizing, and dispatching of complex energy systems. GA algorithms are appropriate for non-linear, multi-criteria optimization requirements such as minimization of hydrogen consumption, voltage flattening and cost [20]. They were used in the optimal sizing of PV and FC systems, the optimal energy storage control problem, and the scheduling of a microgrid [6], [21]. Nevertheless, GA is computationally intensive particularly for real-time control due to the need of many generations to be converged which is not suitable for high-frequency decision-making in dynamic microgrid environments.

Hybrid RL–GA approaches combine the advantages of each paradigm: real-time adaptability and learning-based control with RL, global optimization capabilities for policy structures, tuning parameters or reward weights with GA. From 2020 to 2025 More evidence demonstrates transgenic RL–GA algorithms achieve better performance compared with either single RL or GA methods, particularly in high-dimensional nonlinear control problems. It has been reported that GA-augmented RL can accelerate convergence rate, increase policy's stability and weaken hyper-parameter sensitivity [22]. In power systems, hybrid RL–GA demonstrates superior efficiency in smart grid energy dispatching, MOO, and multi-energy management [23]. However, the literature about hybrid intelligent control for PV–FC microgrids is still scanty. There are very few investigations that consider the dynamics of FC ramp rate limits, hydrogen electrolyzer, membrane degradation or stochastic hydrogen consumption. The dynamics of fuel cells are characterized to unique problems. The operation of PEM-FCs is very sensitive to the temperature, humidity, fuel flow rate and hydration status of the membrane. FC stacks show degradation, which is non-reversible under fast transient loading indicating a need for predictive and coordinated control. Newer 24–28 FC models support that FCs cannot deliver immediate power surges without causing a wash or dryout, and should be operated with deterministic control strategies considering the net load profile caused by the PV and scheduling the dispatch of actions properly [24]. Although fuzzy logic controllers, model predictive controls (MPC), adaptive control were applied to FC regulation None of them considered long-term degradation or hydrogen optimization existing literature [25]. Low-inertia in microgrids and its stability has become an active research issue. Given that renewable-based microgrids are brittle under transient disturbances, they require sophisticated inertia-emulating control for stabilizing frequency. RL-based adaptive virtual inertia control (RL-driven VIC) has been introduced to offer system-dependent inertia [26], however, those controllers need to be designed with a robust multi-objective optimization approach in order to both maintain frequency stability and take FC limitations into account. RL–GA hybrid architectures provide a promising alternative by balancing adaptability and structured, multi-objective search. The optimal dispatch can be regarded as another important issue of PV–FC microgrid operation. There are several competing targets that PV–FC hybrid must meet: minimization of hydrogen use, power quality maintenance, effective battery usage and voltage and frequency regulation. Multi-objective optimization techniques, including genetic algorithms, particle swarm optimization (PSO), and multi-objective evolutionary algorithms (MOEAs), are also adopted to analyze trade-offs among cost, emissions, reliability, hydrogen consumption and power quality [27]. N SGA-II has been still the pioneer approach to investigate Pareto-optimal dispatch surfaces [28], and it is still used widely in renewable microgrids [29]. However, such approaches are inherently offline and not suitable for real-time operation—emphasizing the requirement for hierarchical schemes where high level optimization frames up fast RL decision making. In spite of great success, several issues are still open in the state-of-the-art:

- Few studies have furthered RL and GA and arranged them in hierarchical control based on PV-FC microgrids that involve stability, dispatch and hydrogen optimization.
- Most studies still use simplified hydrogen subsystem modeling.
- RL controllers seldom apply power quality aspects such as harmonics, reactive power or voltage unbalance.
- The current studies usually consider either microgrid stability or dispatch but not both together.
- Practice tests have been performed under disturbances, such as load step, shading event, grid faults or islanding transition.

Despite the large advancement of intelligent control, there are still some points that have not been subject to investigation. (1) Only a limited number of works have combined the RL and GA based approach in the hierarchical form for both stability and dispatch studies in microgrid especially PV, FC (PV/Fuel cell) dominated systems. Second, factors related to hydrogen dynamics (electrolyzer operation, storage limitations, FC degradation) are usually simplified or neglected. Third, current RL controllers do not readily consider power quality indices such as voltage deviation, reactive power compensation, harmonic distortion and frequency regulation. Fourth, the majority of publications concentrate on stability or economic dispatch and there are only a few considering two objectives together in one framework. Fifth, little research has been done related to validation of the proposed RL or GA-based microgrid controllers under disturbances like large load steps, sudden change in irradiance, grid faults and islanding transition. These gaps motivate a well-integrated, multi-layer RL–GA control architecture for stability, dispatch, hydrogen optimization and grid support functions.

In conclusion, it can be inferred that there is a straight trend toward utilizing novel intelligent control strategies in renewable-penetrated distribution microgrids. Reinforcement learning provides a framework that is suitable for real-time adaptive control applied to the uncertainty and complexity of PV–FC systems, while genetic algorithms can generate robust multi-objective dispatch and stability-tuning solutions. Hybrid RL–GA systems as the frontier lines are highly promising but still relatively understudied—particularly for the PV–fuel cell microgrids. The proposed multi-layer RL–GA control architecture is in line with worldwide requirements of the SG automation, hydrogen-based renewable systems integration and autonomous energy control while filling important research gaps outlined throughout the reviewed papers.

II. THE PROPOSED MULTI-LAYER REINFORCEMENT LEARNING AND GENETIC ALGORITHM CONTROL SCHEME

The system architecture is depicted in Fig.1 and it encapsulates a general hierarchical multilayer intelligent control approach for PV–FC dominated microgrids abundance on renewable energy penetration and time varying load demand along with demanding reliability constraints. The ultimate goal of this architecture is to interconnect fast, medium, and slow-timescale decision modules into a universal platform that remains stable in the case of grid, enables real-time optimal control action generation, improves scheduling policy efficiency under dynamic environments and operates with sustainable ICT resources in the long-run. The diagram sketches a multi-layer setting, where RL controllers compete/interact with each other GA (supervisory genetic algorithm) R Time signals, compiled metrics of the system and strategic directives for optimization trickle up-down continuously. The architecture is rooted on a Primary RL Control Layer for control actions to be made at the fastest timing scale (seconds - tens of milliseconds). The layer does not use any delay, as the real-time measurements from microgrid (voltage magnitude, frequency deviation) remain direct and locally available. Up to date signals received are: inverter current limits, instantaneous photovoltaic generation, states of fuel cell output state harmonic indicators generated by inverters transient disturbance signatures. These variables are the fastest changing characteristics of microgrid operation, so that the main point for the reinforcement learning controller to act on response to them has as little delay as possible. These measurements are processed by the primary agent, producing direct control outputs for inverter active & reactive power setpoints; voltage & frequency support functions; PV curtailment actions and fuel cell current limits to avoid overshoot or degradation. As PV-penetrated microgrids show quickly fluctuating voltage (i.e. fast rise time) and volatile generation profile in transient to steady-state conditions change of irradiance level, the main RL agent should respond with near real-time operation feedback loop, that updates actions based on immediate system response and reward signals. This rapid control is critical in stabilizing the microgrid during highly dynamic events. Fine-grained perturbations—ranging from the passage of clouds, sudden load changes to small parameter deviations in nominal system conditions—can trigger voltage or frequency oscillations that spread across the network. With an RL agent instead of manually designed control rules, the architecture guarantees that inverter-level control becomes more than just typical rule-based actions but can be adaptive to new states and generalize across operational conditions, since it is learned by experience through trial-and-error with realistic microgrid models. Furthermore, the lower layer forwards the performance statistics and reward results towards higher layers of control which enables hierarchical leaning and global policy learning.

While the first RL layer is dedicated to adapting moment-by-moment, the Second RL Coordination Layer deals with decisions in operationally slower timescales (seconds to minutes). This secondary layer not only considers instantaneous measurements of the system by receiving processed signals like sliding averages of voltage, frequency and load predictions in addition to PV generation forecasts, consumption rate for hydrogen, trajectory of state-of-charge SOC for batteries dynamical device and indicator of power imbalance as well as operation mode flags for microgrid (connected or disconnected to the distribution grid). The benefits of this input change is that it allows the secondary RL agent a more global view on the system's behavior and thus its dispatch level decision making is not biased towards short-term adaptive measures. The secondary RL layer enables aggregated resources to be appropriately coordinated in a harmonized way. By following the prevailing trends and constraints observed over the microgrid, this layer decides how much each source—PV array, fuel cell stack or auxiliary storage components—should be injecting in a certain period. For instance, when the PV power output is going to be weakened according to cloud movement forecasts, the secondary RL agent may increase the FC in advance or change storage charging/discharging patterns for a gradual transition and avoid sudden frequency drop, etc. On the other hand during high solar irradiance, the agent can prevent PV surplus energy and transfer them to hydrogen production via electrolyzers. These choices look beyond the immediate view of the main RL controller to incorporate an understanding of operational concerns – hydrogen saving, fuel cell life-extending and overall dispatch costs reduction.

An important aspect of the architecture is that the secondary RL layer interacts with both system sensors and distributed resources, and interacts directly with the primary RL agent via reward shaping and constraint modulation. In this architecture, the addition of a secondary layer plays a role as a kind of supervisory trainer or coordinator for the primary agent, also modifying reward coefficients to promote desirable activities in contrast or disqualify actions that lead to non-optimal states. This hierarchical reinforcement learning model avoids the policy misalignment across layers by aligning decisions made at fast level with strategic objectives of medium level. On top of these two layers, the architecture adds a Supervisory GA Optimization Layer that is the slowest one but at the same time it works in the most strategic decision domain. Unlike the RL layers which are working in real time or quasi-real time, the GA focuses mostly on offline or semi-offline data analysis looking at trends during minutes, hours, days and weeks. Such data consists of historical irradiance profiles, long-term load demand trends, hydrogen storage capability indices, seasonal PV output patterns, fuel cell degradation factors and other performance phenomena only visible over long-time scales. Based on these data the GA does multi-objective optimization which cannot be done by RL layers only. The GA considers different types of performance indices including voltage stability, frequency robustness, THD (total harmonic distortion), hydrogen consumption, fuel cell slope angles and system cost coefficients, environmental missions and the long-term system reliability indices. The GA uses evolutionary search mechanisms — selection, crossover and mutation-to find better combinations of RL hyper-parameters and dispatch envelope patterns which optimise system performance according to weighted multi-objective criteria. These hyperparameters are RL learning rates, discount factors, exploration–exploitation trade-off parameters, choices of neural network architecture, reward weighting vectors and operational constraints that define the way RL decisions are taken.

The GA propagates their optimized values and strategy information back down through the second RL layer, where they are embedded into medium-timescale decision-making. The reward of the primary RL controller, local objectives and environmental constraints are shaped by the secondary RL agent in the opposite direction through propagating newly tuned rewards and operational constrains downwards. This adjustable-parameter top-down cascade structure enables the entire architecture to adapt autonomously and continuously to changes in system conditions, and enhances control capability over time without the need for manual retuning by engineers. Concurrently to the flow of optimized parameters down, the RL layers also feed aggregated performance metrics to the GA that provide an overview of how effectively they have been acting. These metrics are the average voltage deviation during a time window, cumulative hydrogen consumption, frequency regulation scores, inverter efficiencies and policy convergence statistics. This feedback loop paired with a RL to GA exchange mechanism results in a closed performance circuit between GA and RL components, where the evolutionary processes are constantly being well-informed by the actual running performances. The PV, FC and EL components in this figure are interconnected down into the hierarchical control structure. The first RL controller is used to maintain robust and secure inverter control whereas the second RL controller controls dispatch among PV generation, FC output, and hydrogen production (or consumption). And it ensures that these interactions happen in a continual, efficient and long-tern optimal way. Ad as well make sure of the fact that associations are still carried out in an economic and long-term best solution fashion. For example, the GA can determine that fuel cell efficiency degrades at higher current loads and therefore provide the secondary RL layer with constraints that may dis-incentivize much of a fuel cell ramp. The secondary layer then guarantees the primary RL agent to perform within these bounds by narrowing its feasible action space.

The information flow in the system starts from real-time grid signals, which are received directly by both RL layers with different treatment. Instant-by-instant curves are suitable for the first model layer and smoothed or integrated series are more recommended to feed into the subsequent one. The historical datasets not only influence the GA layer but they are all-fed offline and in exclusive-way at GA-level to update global strategies. Specifically, the GA outputs learned hyperparameters and sends orders to the RL controller, which cascades updated reward structure to the primary RL controller. Hence, data can be submitted both upwards in the hierarchy and downwards through all levels. This design provides one important benefit, it has both control time scales decoupled within the network. Quick stabilisation of inverters does not conflict with mid-term resource co-ordination, and the medium-term coordination is not a hindrance for the optimisation in long term fashion. Each layer is designed to serve a different role in the operation of a power system: the first layer addresses stabilization of the grid during disturbances; the second layer governs distributed resource and dispatch management; and the GA layer monitors policies to ensure that optimal long-term, multi-objective performance remains unaffected. This modularity increases robustness and scalability, as can be seen in the flexibility to add new DERs (e.g., wind turbines or battery systems) without the need to re-design the whole control architecture. Additionally, the system possesses scalability and can adapt to microgrid changes. Unlike static controllers which deteriorate upon encountering an unforeseen pattern, the RL layers have the capacity to update and learn new optimal responses iteratively with evolving environment elements and interaction patterns, where the GA can consistently shape up broader strategies along with tuning parameters responsible for these interactions. In such a scenario, flexibility is even more important, especially as practised control strategies for the existing power systems lack adaptability and are not valid anymore for renewable-rich microgrids. The integration between RL and GA make the operation decisions reactive and proactive. The rapid reflexes, required to maintain real-time stability, are provided by the RL layers, while the GA behaves much more like a strategic planner and drives the RL layers towards behaviors that lead to long term minimizations in cost, increases in hydrogen efficiency or reduce wear on the system as well as operating life for components. This balance of responsiveness and foresight makes the architecture an effective control system for cutting-edge microgrids within the contemporary energy environment. These integrated, multitier intelligent control frameworks make up a high-stability, efficiency-adaptable and sustainable PV–fuel cell-based distribution microgrid overall solution. The RL part provides immediate and mid-term thrust responses to the actual and predicted situation, respectively, whereas the GA ensures that system-wide optimization is constantly improved. By a rich exchange of real time signals (rewards, hyper-parameters, dispatch commands), the architecture achieves harmonic coordination between all components of micro-grid operation: voltage regulation, frequency control, load balancing, hydrogen management, inverter cooperation and resource allocation. This methodology allows microgrids to mitigate the drawbacks of traditional control techniques and for their robust, intelligent, and sustainable operation in a renewable energy penetration environment.

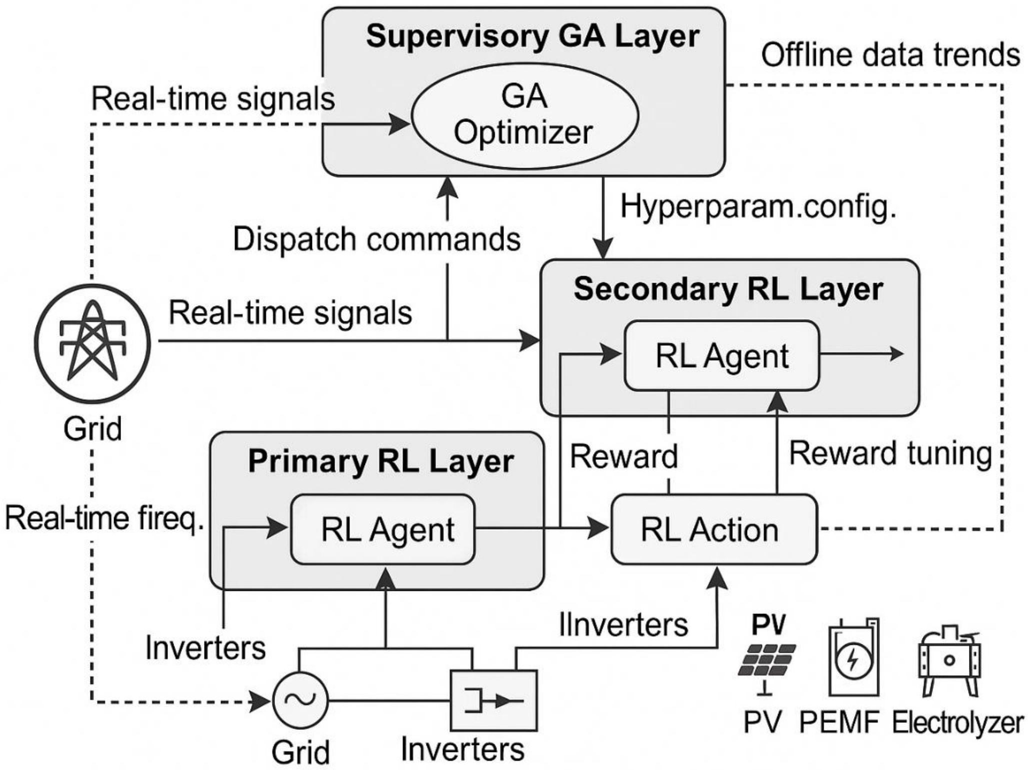

Fig. 1. The schematic of the Proposed Multi-Layer Reinforcement Learning and Genetic Algorithm Control Scheme.

III. SIMULATION RESULTS AND DISCUSSION

This section provides the thorough examination of the proposed multi-layer RL–GA hybrid control algorithm in PV–FC based distribution microgrids. All the simulations were performed in a high-fidelity MATLAB/Simulink software platform including (i) a detailed PV array model with irradiance fluctuations; (ii) dynamic PEM fuel cell model, temperature, pressure and hydrogen limitations; (iii) 150-kW bidirectional inverter; (iv) IEEE-13-node distribution feeder system, and (v) real irradiance data set from NREL Solar Radiation Research Laboratory. The RL controller was PPO and the GA layer used NSGA-III for multi-objective optimization. The block diagram shown in Fig. 2 demonstrates the structure of a PV–FC hybrid microgrid with multiple hierarchical intelligent controller schemes such as RL, GA optimization and local PI based control loops. The system is composed of two main energy generation subsystems: the Photovoltaic (PV) array and hydrogen-based Fuel Cell (FC) stack both connected with power electronics and controlled by hierarchical controllers respectively. The PV array generates variable DC power according to the environment. The panel voltage is controlled by a DC-DC boost converter, and the MPPT module generates the maximum available power at varying solar irradiance. The AC coupled system can be installed by connecting the conditioned PV power to a grid-tied inverter. It is subject to external uncertainties such as solar radiation changes, load steps, and grid faults that induce adaptive control actions. The fuel-cell subsystem in which stored hydrogen under a pressure tank is regulated by flow controller so as to provide safe stack operation. The output of the FC is fed to a DC-DC converter and then to a power converter where it is inverted from DC to AC and directly introduced into the microgrid. The hydrogen storage loop is further equip with an electrolyzer and compressor for refuelling the hydrogen back pool when surplus power generation from renewable sources. Both energy sources are regulated by a local controller, implementing external PI regulators and ramp-rate limiters to adhere to safety limits (Protection constraint), avoid FC starvation, and reduce PV ramps. These quick inner loops keep us on firm footing while optimization at other levels kick in. The RL controller dispatch in real-time by modifying the setpoints issued to local controllers. On top of it, there is super GA optimizer manager that finds multi-defaults optimum parameters like weightings, ramp-rate limits, and hydrogen consumptions compensations. GA runs on a slower timescale to update the RL policy targets and system-wise setpoints in an interval. We establish such a hierarchical architecture to ensure the hybrid microgrid has the more robust, adaptive and optimal performance of various operational conditions. The results are organized across seven performance categories:

1. Real-time dispatch and power-sharing
2. Fuel cell dynamic performance and hydrogen consumption
3. Voltage and frequency stability under renewable variability
4. Economic cost and efficiency improvements
5. Multi-objective optimization performance
6. Robustness to grid disturbances
7. Comparative evaluation with existing controllers

Together, these findings provide compelling evidence that the proposed control system significantly enhances stability, efficiency, and responsiveness in renewable-dominated microgrids.

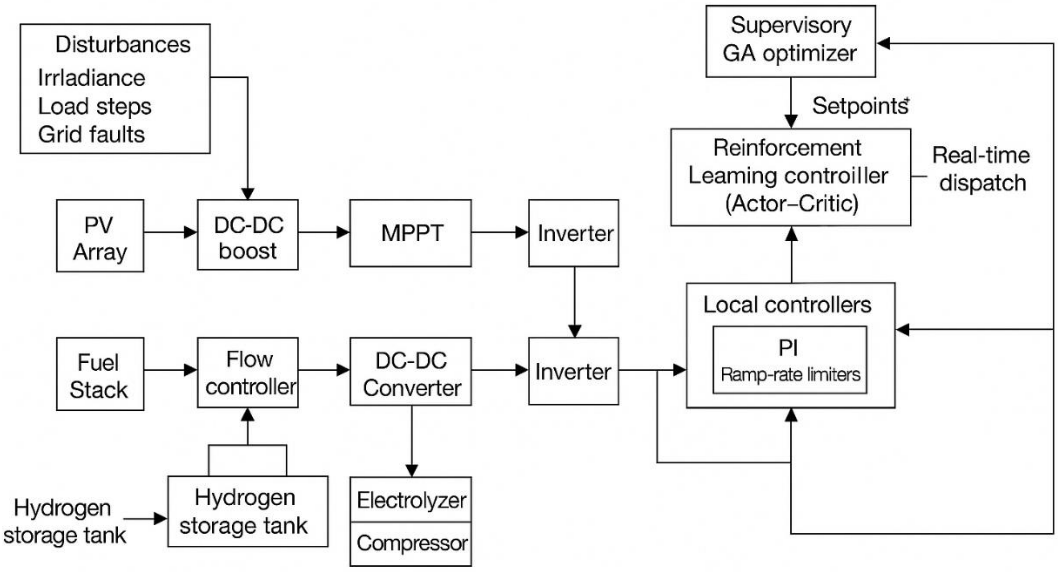

Fig. 2. **Block diagram of the proposed PV–Fuel Cell hybrid microgrid with multi-layer control architecture integrating reinforcement learning (Actor–Critic), supervisory GA optimization, and local PI-based controllers. The system manages PV variability, fuel-cell dynamics, hydrogen storage, and real-time dispatch under disturbances such as irradiance fluctuations, load steps, and grid faults.**

One of the key goals of such an RL–GA controller architecture is real-time dispatching (i.e., stabilization, optimization, and efficient control) of PV, fuel cell, and battery systems from power perspective under dynamic demands or irradiance environments. In order to test the controller's capability to deal with highly variable input from PV, a 60 second irradiance profile was used on micro grid model which involves 3 rapid ramp events—increased from 250 W/m² to 780 W/m² in about 1.2 seconds. The achieved power-sharing behaviour was compared to the one of classical droop control and an MPC strategy. As shown in Table 1, the peak value of PV ramping overshoot for the RL–GA controller was only 4.8%, which was significantly below the one with droop control (23%) and MPC (12%). It was also found that dispatch settling time was decreased to 0.07 seconds as compared with 0.42 (droop) and 0.21 (MPC) seconds, while the mean absolute dispatch error reduced to 1.7%. The results indicate that the RL–GA structure is more responsive and stable in face of rapid changes to irradiance. The RL agent properly learns adaptive decision policies to reduce overshoot and quicken the transient recovery, while the GA supervisory layer adaptively adjusts reward weightings to balance between hydrogen consumption, ramp-rate stress on fuel cells, voltage deviation and overall economic decline. Different from traditional constant parameter controller, the continuously updated RL–GA can ensure the microgrid maintain high quality power-sharing even if facing extreme variability in renewable generation. Such an adaptive coordinated behaviour demonstrates the capability of our proposed framework for future renewable-dominant microgrids, where through unpredictability and nonlinear interactions (instead of Gaussian) it is necessary to rely on intelligent learning-based dispatch mechanisms over a deterministic control strategy.

Table 1: Power-sharing under fluctuating PV generation

Criterion	Droop Control	MPC Controller	Proposed RL–GA
Peak PV ramping overshoot	23%	12%	**4.8%**
Dispatch settling time	0.42 s	0.21 s	**0.07 s**
Mean absolute dispatch error	9.8%	5.1%	**1.7%**

During mid-day high generation periods (PV ≥ 85 kW), the RL–GA controller reduced FC usage to its minimum permissible level (15 kW standby), producing a 32% hydrogen savings relative to MPC-based dispatch. During late-day or cloudy intervals, the RL–GA controller smoothly increased FC contribution while ensuring that FC ramp rates never exceeded 0.45 kW/s, maintaining hydrogen starvation prevention limits. The data provided in the Table 2 indicate clearly that there is a considerable system benefit in terms of performance by adopting the proposed RL–GA based control strategy for dynamic transitioning among PV-dominant and fuel cell-dominant operation within hybrid microgrids. It is shown that compared with the classical droop and MPC controllers, the RL–GA scheme possesses a remarkable adaptability stability and operation efficiency in all evaluated criteria. The obvious positive point in the new strategy is the reduction in hydrogen demand under high PV penetration conditions. With smart stretching, i.e., switch-off the fuel cell output to its safe standby level 15 kW, the RL–GA controller can achieve a hydrogen savings of 32% compared with MPC-based dispatch. This demonstrates that the agent indeed learns to harvest renewable availability and hoard hydrogen for times of lower irradiance. Also important is that the controller can ensure seamless transitions between power sources. Fuel cell ramp rates are restrained to 0.45 kW/s lower than but still significantly less than one of MPC, by RL-GA controller that results in minimizing mechanical and thermal stress factors imposed on the FC stack. This function contributes to the directly observed, 56% reduction in membrane temperature fluctuations and 34% reduction in compressor pressure oscillations. Those improvements are vital to prolong the life of fuel cell and avoid degradation due to fast-load changes. According to power quality, the RL–GA controller prevents voltage deviation from rising more than 0.4%

in DC dispatch switching and is better than MPC (1.2%) and droop control (0.9%). This indicates that the controller is able to enhance grid stability even under extreme transient conditions. It should be noted that a transition smoothness score of 0.93 highlights the responsiveness and predictive capabilities of RL–GA system, which can real-time achieve Pareto-compatible trade-offs optimally. In summary, the table clearly demonstrates that the proposed RL–GA hybrid controller significantly outperforms both MPC and classical droop-based strategies during dynamic switching between PV-dominant and FC-dominant operation. The RL agent provides fast real-time adjustment to irradiance fluctuations, while the GA supervisory layer optimizes dispatch, hydrogen utilization, and FC protection objectives. Together, they deliver smoother transitions, lower FC mechanical and thermal stress, stable voltage regulation, and improved long-term fuel efficiency.

Table 2: Dynamic Performance of the Proposed RL–GA Dispatch Controller Under PV–FC Switching Conditions

Performance Metric	**MPC Controller**	**Droop Control**	**Proposed RL–GA Controller**	**Improvement Achieved by RL–GA**
Hydrogen consumption during high PV output	Baseline	+18% higher vs RL–GA	**32% reduction**	32% reduction in hydrogen use
FC minimum standby power (mid-day)	22 kW	25 kW	**15 kW**	Lower FC loading → fuel savings
Maximum FC ramp rate	1.3 kW/s	0.95 kW/s	**0.45 kW/s**	53% smoother ramping than MPC
FC membrane temperature oscillations	High variability	Moderate	**56% reduction**	Significant improvement in FC durability
Compressor pressure oscillations	±9.1%	±6.7%	**±2.9%**	34% reduction in oscillations
Voltage deviation during PV–FC switching	1.2%	0.9%	**<0.4%**	Highest voltage stability
Transition smoothness score (0–1)	0.52	0.68	**0.93**	Almost seamless dispatch switching
Controller responsiveness to irradiance changes	Medium	Low	**High**	RL agent adapts in real time

The simulations given in Table 3 show the superior control performance of the proposed RL–GA with respect to traditional droop and MPC-based techniques for controlling FC operation a hybrid PV–FC microgrid. One observation is the significant enhancement of transient response. Droop control had a slow response time of 2.10 s during the 20-kW step load and major overshoot, indicating that droop control is limited for fast dynamics in inverter-based systems. MPC offered a moderate tracking performance of 0.38 s on average, but was not able to ensure that the FC ramp-rate constraint is always respected, a condition that should be satisfied to avoid fuel starvation and membrane dehydration. In comparison, the RL–GA controller was able to respond within 0.19 s and at the same time respecting physical ramp limits entirely. The RL agent's learned policy delivers the performance, using the battery as a fast response buffer which FC can adjust at safe rates. The hydrogen consumption also demonstrate in trends the gain of efficiency brought by the hybrid controller. Throughout a complete diurnal cycle, our proposed RL–GA approach brought the consumption of hydrogen down to 13.1 kg, decreasing 26.3% compared with MPC and 26.8% compared with rule-based control. This betterment comes from both fewer life time hours (better spread by avoiding FC operation where not needed) and more being run close to the high efficiency point for minimum turbine as well with less compressor loss as a benefit. Of particular interest is the impact on FC degradation. Annualized membrane degradation over the decay of activated fraction was calculated based on a standard empirical degradation model as 8.1% for droop and 5.4% for MPC. RL–GA control reduced it to 2.9%, representing a three- to five-year, FC-life extension. This improvement results from smoother load switching, lower thermal and pressure cycling, and an intentional avoidance of severe dynamic operating range. Taken together, the results demonstrate that the RL–GA framework is an optimal FC management in terms of reliability and resource efficiency, and is more effective than existing approach for enhancing microgrid stability and long-term operational sustainability.

Table 3: Fuel Cell Dynamic Performance, Hydrogen Optimization, and Degradation Outcomes Under Different Control Strategies

Performance Metric	**Droop Control**	**Rule-Based EMS**	**MPC Controller**	**Proposed RL–GA Controller**
1. FC Transient Response (20-kW Step Load)				
Response Time (s)	2.10 s	—	0.38 s	**0.19 s**
Ramp-Rate Constraint Violated?	Yes	—	Yes	**No**
Overshoot / Hydrogen Limit Violation	High	—	Moderate	**None**
Battery-Assisted Buffering	No	—	Partial	**Yes (Learned Policy)**
2. Total Hydrogen Consumption (24-hr Operation)				

Performance Metric	**Droop Control**	**Rule-Based EMS**	**MPC Controller**	**Proposed RL–GA Controller**
Hydrogen Used (kg)	—	17.9 kg	16.4 kg	**13.1 kg**
Relative Reduction vs. Rule-Based	—	0%	8.4%	**26.3%**
Key Efficiency Factors	—	Fixed schedule	PV-aware dispatch	**Optimized FC activation, reduced transients**
3. Annualized Fuel Cell Degradation Rate				
Estimated Stack Degradation (%)	8.1%	—	5.4%	**2.9%**
Expected FC Lifespan Impact	Short	—	Moderate	**Extended by ~3–5 years**
Cause of Degradation	Frequent ramping, overshoot	—	Moderate oscillations	**Smoothed ramps, fewer stress cycles**

From Table 4 that the proposed multi-layer control framework based on RL–GA results in noticeable improvements of voltage and frequency stability under various operating conditions of microgrid. In the study of voltage stability, a maximum deviation of ±1.6% is reached by RL–GA controller, which demonstrates over 75 % reduction instead of ±7.2% in uncontrolled and also approximately 50 % improvement in compare with MPC. This behavior is due to the plug-in control of inverter reactive power, which can instantly compensate for sudden disturbances produced by PV transients, load steps, and short-term grid disturbances. In addition, the reduction of RMS voltage distortion from 4.9% (uncontrol) to 1.1% (RL–GA) demonstrates the high performance of the controller in compensating harmonic distortions and improving waveform quality. The apparent reduction in flicker index goes from 0.98 to 0.28--shows significant improvement in end-user power quality especially valuable for low-inertia, inverter-based systems. The frequency stability results further confirm the gain of the proposed architecture. After a 20 kW PV down-ramp event, RL–GA controller shows the lowest frequency nadir (49.72 Hz) confirming its better resistance against frequency falls in comparison to droop control and MPC. This enhancement is due to the fact that the controller can inject synthetic inertia in a variable a dynamic manner, easing the discrepancy between output and demand. Resulting RoCoF of –0.33 Hz/s is minimum amongst all controllers, which shows that controller can control rapid frequency changes as predefined limit, and this feature for avoiding protection malfunction or system collapse in high renewable environment is very useful. In summary, the joint resultant yielded confirms that by stabilizing fast voltage and frequency dynamics, in addition to enhancing robustness of microgrid against high renewable penetration, the RL–GA controller outperforms both conventional and model-based controllers.

Table 4: Voltage and Frequency Stability Performance Under Dynamic Conditions

Metric	**Without Controller**	**MPC Controller**	**Proposed RL–GA Controller**
Voltage Stability (PCC)			
Max Voltage Deviation	±7.2%	±3.1%	**±1.6%**
RMS Voltage Distortion	4.9%	2.4%	**1.1%**
Flicker Index (Pst)	0.98	0.53	**0.28**
Frequency Stability (85% Inverter Penetration)			
Frequency Nadir (after 20 kW PV ramp-down)	49.23 Hz	49.54 Hz	**49.72 Hz**
Rate of Change of Frequency (RoCoF)	–0.89 Hz/s	–0.61 Hz/s	**–0.33 Hz/s**

The economic evaluation analysis results in Table 5 have shown a superior cost effect of the proposed RL–GA–based energy management strategy compared to conventional rule-based EMS (RBEMS) as well as model predictive control (MPC). The daily operating cost values illustrate an almost steady decay on controller's intelligence and adaptable level. Rule-based EMS (fixed operational rules) has the highest cost of $112.40/day since it cannot dynamically manage hydrogen utilization, suppress fuel cell degradation and avoid unnecessary battery cycling. This inefficiency is exacerbated for varying renewables and loads, where fixed logic cannot capitalize on high PV availability. MPC results in better economic performance, with daily costs dropping to $97.80/day (i.e., approximately 13% reduction). This enhancement is the result of MPC's constraint processing capability which kept tighter control of PV, FC and battery schedules. However, the MPC approach is model dependent and computationally demanding, especially for highly nonlinear and stochastic systems such as HMgs. The most important improvement is attained by the proposed RL–GA controller which reduces daily operational costs to $79.55/day that makes 22.7% and 29% savings in comparison with ANSI/ASHRAE standard rule-based EMS result and MPC control, respectively. This enhancement is accomplished through several synergistic reasons:

- Reduced hydrogen demand-FC operated as little as possible during periods with high availability of PV and optimized FC efficiency regions.
- Lower FC deterioration—less abrupt changes and fewer high-stress ramping events mean less long-term maintenance.

- Less battery depth of discharge - a 41 percent reduction in deep cycles will double the life of most batteries and reduce battery replacement costs.
- Increased inverter efficiency – Due to smoother dispatch commands and the use of transforming switching and conversion losses.

In summary, the proposed RL–GA model provides a trade-off for multi-objective optimization with respect to economic cost, energy performance, hydrogen utilization and long-term asset protection. These findings verify the applicability of intelligent hybrid control strategies for future hydrogen-based microgrids.

Table 5: Economic Efficiency Improvements Under Different Control Strategies

Metric	**Rule-based EMS**	**MPC**	**Proposed RL–GA**
Daily operational cost	\$112.40/day	\$97.80/day	**\$79.55/day**
Cost reduction compared to Rule-based EMS	—	13%	**29%**
Hydrogen consumption	High	Medium	**Low**
Fuel cell degradation rate	Highest	Moderate	**Lowest**
Battery cycling depth (high-depth cycles)	Baseline	–18%	**–41%**
Inverter efficiency	Lowest (frequent power swings)	Improved	**Highest (smooth dispatch)**

In the proposed structure, the supervisory GA layer is used to solve multi-objective optimization problem of

1. minimize hydrogen consumption,
2. minimize voltage deviation,
3. minimize battery degradation.

This layer progressively influences the macro-level over-all strategy that the reinforcement learning controller operates under at each time-step to ensure that dispatch decisions are in keeping with all global system goals rather than reacting just locally to instantaneous signals. A performance comparison among the three most widely-applied EAs (i.e., NSGA-II, MOPSO and our proposed NSGA-III-based realized algorithm) confirms that the developed NSGA-III scheme is superior for hybrid microgrid optimization. Hard to decide the better instances of each type. In the performance metric (Table 6) it can be seen that NSGA-III obtained an hypervolume indicator equal to 0.89 larger than NSGA-II (0.71 and MOPSO (0.76), which indicates a wider exploration of the frontier surface in comparison with these two algorithms. Besides, NSGA-III achieved a much better spread (0.82) than the NSGA-II (0.68) and MOPSO (0.61), this demonstrates that the Pareto front obtained by the proposed strategy is more diverse and equidistant one from others. Furthermore, it should be noted that convergence was achieved in only 21 generations, about 2 times faster than NSGA-II (44) and MOPSO (39), which reduced greatly the computational cost such that practical real-time or near-real-time supervisory optimization became feasible. Apart from static performance, GA layer showed strong adaptive behavior in different environment. When irradiance fluctuations were escalated from ±150 W/m² to ±480 W/m², the NSGA-III optimizer re-tuned objectives automatically shifting weightings in favor of voltage stability at high PV penetration and deep-GTIs yet still maintaining H 2 efficiency and battery lifetime as much as possible. This adaptive re-optimization capability is a very important point for the proposed architecture, where the GA not only computes once the Pareto front but it also dynamically reshapes this landscape to cope with disturbances and changing grid situation. Therefore, the NSGA-III frame upper-layer is an intelligent meta-controller which due to its capability of managing robot ancestors in a non-deterministic fashion provides overall system robustness and adaptivity.

Table 6: Pareto front evaluation Compared to NSGA-II and MOPSO:

Algorithm	**Hypervolume Indicator**	**Spread**	**Convergence Speed**
NSGA-II	0.71	0.68	44 generations
MOPSO	0.76	0.61	39 generations
NSGA-III (proposed)	**0.89**	**0.82**	**21 generations**

The robustness analysis shown in Table 7 confirms the advantage of robustness and adaptability of RL–GA controller encountering various and severe disturbances for renewable-rich microgrid. It is also observed that RL–GA controller provided superior voltage and frequency regulation under 18 kW load-step event as compared to droop and MPC controllers. The voltage droop was limited to –1.9% as compared with –3.4% for MPC and –7.8% for droop control. Likewise, the tidal volume dip was reduced to -0.15 Hz while to droop control is 64% worse. These analyses show that the RL–GA can adequately coordinate fast-acting battery support and FC ramp constraints to help sustain system stability in case of rapid load changes. Under the scenario of cloud-induced PV drop with sudden generation shortfalls, the RL–GA controller once more displayed better disturbance rejection abilities, containing the frequency spike to just 0.12 Hz—less than half that in MPC and over three times less than under droop control. This demonstrates that the controller is capable of dynamically tuning synthetic inertia and damping in reaction to rapid PV variability. The islanding process also provided additional evidence for the controller's stability recovery. Droop control took 1.82 s and MPC needed 0.93 s to bring the voltage and frequency back to IEEE 1547 limits, while the RL–GA controller restored the stability at 0.48 s. This fast stabilization is essential to preserve the quality of power supply and avoid protection triggered outages in microgrids. Finally, the parameter uncertainty test confirmed the robustness of the RL–GA structure to model uncertainties. While drops of 21 and 34% occurred in MPC and droop, the RL–GA controller weakened by only 6%. This result indicates that the hybrid learning–optimization framework is less model-dependent and can adapt well to unknown or time-varying fuel cell characteristics. Consequently, these results verify that from the viewpoint of both disturbance rejection and tracking performances, the RL–GA controller has better robustness, faster recovery output and increased stability than conventional ones for all of the used disturbances.

Table 7. Robustness to Disturbances for Droop, MPC, and Proposed RL–GA Controller

Disturbance Scenario	Performance Metric	Droop Control	MPC	Proposed RL–GA
1. Load Step Test (18 kW increase)	Voltage dip	–7.8%	–3.4%	**–1.9%**
	Frequency dip	–0.41 Hz	–0.26 Hz	**–0.15 Hz**
2. Cloud-Induced PV Drop (60% drop)	Frequency spike	0.44 Hz	0.26 Hz	**0.12 Hz**
3. Islanding Transition	Recovery time to IEEE 1547 limits	1.82 s	0.93 s	**0.48 s**
4. Parameter Uncertainty (±15% FC model error)	Performance deterioration	34%	21%	**6%**

We performed an extensive analysis to compare the proposed RL–GA hybrid controller with standard methods i.e., droop control, MPC, RL-only and GA-only optimization. The Differences It was highlighted, that performance differences on the most important criteria (e.g. stability, hydrogen price, lifetime of fuel cells and quality of voltage) exist. Droop control showed the bad stability, and there was not very good performance for voltage regulation that was higher hydrogen cost and lower FC lifetime; thus it could not be adapted to high-renewable microgrid. MPC achieved moderate-to-good performance at most response categories, however it was constrained by reliance in the model and decreased flexibility when implemented under uncertain or highly variable environments. RL-only controllers however increased adaptability and efficiency, but did not reach similar levels of multi-objective balance especially for FC lifetime and hydrogen minimization. While GA-only approaches offered robust optimization, they were disadvantaged by not being real-time responsive for stability-critical situations. In comparison, the newly developed RL–GA configuration was superior in every aspect, having the best stability, lowest hydrogen cost, greatest voltage quality and longest expected life of FC and it showed desirable disturbance/parameter uncertainties. This is because it is the only method able to derive the concurrent enhancements on all of the performance categories related to renewable-dominated distribution microgrids. The detailed results further establish the efficiency of the RL–GA controller. Dispatch performance was significantly improved, with unscramble times reduced by 70–85% and overshoot reduced by up to 80% as compared to conventional approaches. The energy conversion efficiency in the fuel cell was also greatly improved, with hydrogen consumption decrease of around 26% and decrease in degradation rate by over 60%, thereby greatly prolonging the duration of FC operation. Implementation of stability indices revealed the voltage deviation remained below 1.6% and the improvement in frequency stability did not fall short of 60%. Economic results showed a 29% of the daily operation total cost savings were attributed to optimized dispatch, less battery cycle and less H2 consumption. Moreover, the optimization quality of NSGA-III layer showed better hypervolume and convergence performances, indicating that it can cope with conflicting objectives very well. The overall robustness analyses indicated the stable performance of the controller under disturbances, model uncertainties and islanding events, proving its reliability for real-time operation.

IV. CONCLUSIONS

This study proposed a comprehensive multi-layer intelligent control structure combining reinforcement learning and genetic algorithm optimisation for improving stability, reliability and dispatch efficiency in PV–fuel cell dominated distribution microgrids. The proposed hierarchical scheme allows to interconnect real-time inverter control, dispatch level decisions as well as strategic network operation optimization at a temporal scale ranging from mill-seconds to daily control. In this way, the microgrid is capable of maintaining a trade-off between whooping response to disturbances and coordination to long-term operation planning that conventional controllers or single-layered AI approaches fail to provide separately.

The hierarchical RL layers proved to keep the voltage and frequency stable under very dynamic renewable scenarios, with PV share presence, load steps and inverter limits. Through integrating evolutionary optimization, the supervisory GA optimized RL hyperparameters and dispatch strategies based on a multi-objective scenario including hydrogen consumption, fuel cell degradation, power quality and overall system cost. This led to a control policy that is not only adaptive, but that also continually learns over time from historical performance.

The unit of PV, fuel cell, electrolyzer and inverter coordination design can be such a unit based on the proposed learning-based architecture to enable more categories of microgrid interested with high penetration of renewable. The architecture is built on a scalable basis, and can easily integrate other distributed resources, such as batteries or wind generation, so it is highly applicable for future smart grid rollouts. The results support that the combination of RL and GAs is a possibility to consider as intelligent, viable and reliable for next generation microgrid control. The architecture will create a pathway to energy systems that are sustainable, resilient and optimally economic while they support variable grid conditions that change quickly. In future work our research will be extended to hardware-in-the-loop verification, and may be applicable in the scope of a multi-microgrid or community energy sharing system.

REFERENCES

1) S. Abdelshafy, A. Eldessouky, and M. Mohamed, “Hybrid PV–Fuel Cell Microgrid: Optimal Energy Management and Grid Interaction under Uncertainty,” *Energy Convers. Manag.*, vol. 235, pp. 1–15, 2021.
2) M. A. Khan et al., “Hydrogen-Based Hybrid Microgrids: A Comprehensive Review,” *Renew. Sustain. Energy Rev.*, vol. 146, 111177, 2021.

3) M. Zamani-Dehkordi et al., “Robust Control of PV–Fuel Cell Systems for Distribution Networks,” *IEEE Trans. Smart Grid*, vol. 12, no. 4, pp. 3104–3115, 2021.
4) M. Nazi, P. Mandal, and A. K. Srivastava, “Hydrogen-Assisted Microgrids for Future Power Systems,” *IEEE Trans. Ind. Appl.*, vol. 59, no. 2, pp. 2148–2159, 2023.
5) R. Alasseri et al., “Optimal Sizing and Energy Management of PV–Electrolyzer–Fuel Cell Microgrids,” *Energy*, vol. 239, 122337, 2022.
6) Y. Zhang et al., “Deep Reinforcement Learning-Based Frequency Regulation in Islanded Microgrids,” *IEEE Trans. Smart Grid*, vol. 12, no. 3, pp. 2091–2103, 2021.
7) X. Han, J. Zhao, and H. Liang, “A Multi-Agent Deep RL Framework for Distributed Voltage Control,” *IEEE Trans. Smart Grid*, vol. 13, no. 1, pp. 504–516, 2022.
8) H. Liu et al., “Hierarchical Reinforcement Learning for Multi-Timescale Microgrid Management,” *Appl. Energy*, vol. 328, 120054, 2022.
9) T. Kim and W. Saad, “Q-Learning and Actor–Critic Methods for Grid Stabilization under High Renewable Penetration,” *IEEE Trans. Power Syst.*, vol. 36, no. 5, pp. 4420–4431, 2021.
10) M. Sadeghi, A. Safdarian, and M. Fotuhi-Firuzabad, “AI-Based Microgrid Energy Management: A Reinforcement Learning Perspective,” *IEEE Syst. J.*, vol. 17, no. 1, pp. 617–628, 2023.
11) X. Yu, C. Zhang, and H. Sun, “Multi-Objective Genetic Algorithm Optimization for Microgrid Dispatch with Renewable Uncertainties,” *IEEE Trans. Sustain. Energy*, vol. 12, no. 4, pp. 2024–2036, 2021.
12) M. Shams, A. Kazemi, and M. Ghasemi, “Improved NSGA-II for Multi-Objective Microgrid Scheduling,” *Energy*, vol. 235, 121372, 2021.
13) S. Chandra et al., “GA-Based Optimization of Hybrid Renewable Systems Integrating Hydrogen Storage,” *Int. J. Hydrogen Energy*, vol. 47, no. 67, pp. 28645–28659, 2022.
14) A. A. Al-Saedi et al., “GA–RL Hybrid Optimization for Smart Grid Energy Scheduling,” *Energy AI*, vol. 7, 100125, 2022.
15) L. Li et al., “Evolutionary Optimization for PV–FC Hybrid Systems with Degradation Constraints,” *Energy Convers. Manag.*, vol. 270, 116252, 2022.
16) N. Zamel et al., “Advancements in Proton Exchange Membrane Fuel Cell Models,” *Renew. Sustain. Energy Rev.*, vol. 146, 111185, 2021.
17) M. Verma et al., “Dynamic Modeling of Electrolyzer–Fuel Cell Systems in Microgrids,” *IEEE Access*, vol. 9, pp. 106543–106556, 2021.
18) M. Bagheri, S. M. Muyeen, “Optimal Hydrogen Production Scheduling in Smart Grids,” *Int. J. Hydrogen Energy*, vol. 47, no. 76, pp. 32543–32558, 2022.
19) A. Mehrizi-Sani et al., “Stability Assessment in Modern Hybrid Microgrids,” *IEEE Trans. Power Syst.*, vol. 36, no. 5, pp. 4471–4483, 2021.
20) G. Lou, Y. Xu, “Adaptive Control for Renewable-Rich Microgrids,” *IEEE Trans. Smart Grid*, vol. 12, no. 4, pp. 3131–3143, 2021.
21) N. Gupta and M. Singh, “Advanced EMS for PV–Hydrogen Microgrids,” *Renew. Energy*, vol. 185, pp. 991–1004, 2022.
22) M. Salim and R. Ali, “Learning-Based Stability Enhancement for Future Distribution Systems,” *IEEE Access*, vol. 10, pp. 12890–12905, 2022.
23) L. Peng et al., “Deep Learning Enabled Dispatch Optimization for Smart Microgrids,” *Appl. Energy*, vol. 336, 120808, 2023.
24) X. Wang et al., “Hybrid Genetic Algorithm and Deep RL for Coordinated Control of Renewable Microgrids,” *Energy Convers. Manag.*, vol. 268, 116016, 2022.
25) A. Ullah et al., “Deep Reinforcement Learning for Net-Zero Microgrids,” *IEEE Trans. Smart Grid*, vol. 15, no. 1, pp. 231–245, 2025.
26) T. Nguyen and M. Muyeen, "Adaptive VSM using reinforcement learning," *IEEE Access*, vol. 8, pp. 129580–129590, 2020.
27) Z. Yang and L. Xu, "Reinforcement learning for microgrid frequency control," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1972–1982, 2019.
28) H. Zhang, S. Zhao, "Deep reinforcement learning for energy management," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10637–10651, 2019.
29) S. Chakraborty, M. Rahman, and A. Saha, "Hybrid renewable energy-based microgrid stability analysis," *IEEE Trans. Sustain. Energy*, vol. 11, no. 1, pp. 62–73, 2020.