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: 03 November 2025

Intelligent Self-Evolving Neural Control Using Multi-Objective Swarm Optimization and Reinforcement Learning for Dynamic Energy Management in PV-Driven Smart Cities

Adel Elgammal

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

Published Date: November 22, 2025 **Article DOI:10.65150/EP-gjetr/V1E3/2025-03**

ABSTRACT: In this paper, an adaptive self-evolving neural control strategy based on the integration of Reinforcement Learning (RL) and Multi-Objective Particle Swarm Optimization (MOPSO) is proposed for dynamic energy management in Photovoltaic PV-based smart cities. The motivation behind this system is to cope with complexity of modern energy networks, uncertain large-scale solar generation and stochastic urban demand in combination with dynamic grid conditions. We adopt a self-evolving neural network (SENN), being the adaptive decision-making core, that dynamically tunes its parameters by reward-driven learning and Pareto-based optimization. The RL agent continuously controls power flow, energy storage distribution and load scheduling and MOPSO adjusts the multi-criterion objective indices including PQ improvement, voltage stability index, total cost of electrical system and carbon dioxide emission reduction. A hierarchical control architecture is proposed to integrate on-line operation optimisation of the microgrid while considering long-term city-scale prediction. The potential of the framework is demonstrated on a realistic urban PV-grid model including residential, commercial, and EV charging nodes with dynamic weather and demand. Simulation results show that the proposed SENN–MOPSO–RL controller outperforms the traditional Model Predictive Control (MPC), Proportional-Integral (PI) and Fuzzy Logic controllers. In particular, the total power loss is mitigated by 18.6%, voltage deviations by 27.4%, and the REs consumption improved about 22.9% respectively. Additionally, the controller kept grid stability for partial shading, demand surge and communication time delay, with good robustness and self-adaptation. The results validate that integrating swarm intelligence and reinforcement-based neural learning can result in a scalable and robust control framework for autonomous energy systems (AESs) found in PV-dominate smart cities. Also, hardware-in-the-loop validation and multi-agent interaction involving inter-microgrid collaboration with peer-to-peer trading will be investigated in the future.

KEYWORDS: self-evolving neural network, reinforcement learning, multi-objective particle swarm optimization, photovoltaic smart grid, dynamic energy management, urban sustainability.

I. INTRODUCTION

Smart city development has turned conventional power networks into sophisticated, data-informed energy ecosystems that are incorporating distributed autonomous generation, smart metering and control. Regards THE ERECTOR SET Wind and solar, including photovoltaic (PV), are vast drivers of renewable energy because they're scalable, modular, and decreasing in price. But its random and intermittent features bring some challenges to operation such as fluctuation of voltage, distortion of harmonic and two-way power [1]. Dynamic energy control of PV-embedded grids thus needs to take account in real time, that is, multi-object objectives and non-linear dynamics with stability power quality requirements [2]. Conventional Proportional-Integrative (PI) or Model Predictive Controller (MPC)-based design offers deterministic controller, however it does not retain robustness to time varying load patterns and uncertain photovoltaic radiation. Therefore, they tend to use data-driven intelligent control with incorporated neural learning, swarm intelligence and reinforcement learning (RL) to design self-adaptive architectures for dynamic grid optimization [3], [4].

Reinforcement Learning (RL) has emerged as an indispensable tool for adaptive energy management due to its capability of learning optimal actions based on environmental feedback [5]. Reinforcement Learning (RL) agents can adaptively learn control policies for energy scheduling, demand response and storage management in a model-free manner. Xu et al. [6] reviewed RL's development in smart grids, and discussed its application in hierarchical microgrids and PV-battery systems. They found that RL substantially enhances grid resilience and forecast accuracy in composition with prediction and optimization layers. Similarly, Waghmare et al. [7] conducted a survey on RL-based controller and reported that the best discussions were produced by DQL, PG(Data) methods when compared to traditional techniques in terms of total energy costs and losses. Pigott et al. [8] introduced Grid Learn: a multi-agent RL framework for coordinated building-to-grid energy management. They found that a significant load balancing and carbon reduction improvement was seen. Shojaeighadikolaei et al. [9] presented a DRL-based framework to DR. It achieves convergence asymptotically under high-dimensional state spaces with different load dynamics.

However, several key challenges remain: high sample complexity of RL, real-time grid safety validation and computational requirements. Thus, hybrid RL systems that combine the best of both neural networks and swarm optimization have emerged [10].

NNs are powerful non-linear approximators perfect for modelling difficult grid dynamics, predicting solar output and facilitating adaptive control [11]. Static NN models need to be retrained when the grid topology or operating conditions are modified. This has inspired the notion of Self-Evolving Neural Networks (SENNs), which are architectures that learn to evolve their own topology, weights, and operation parameterization in response to external feedback [12]. Thajeel and Atilla [13] developed a reinforcement-neural network based controller for on grid PV systems reflected improvement in transient stability and harmonic under normal MPCs. Their model also employed on-the-fly neural adaptation to voltage and current feedback, yielding THD of 2.8 %. For adaptive FDD and inverter control, ECoS, including Growing Neural Gas (GNG), which is an evolving NN [14], have also been used. These techniques also support lifelong learning in non-stationary grid environments - vital for smart cities where load and generation profiles can fluctuate over time. In the proposed work, more detailed information have been given related to core learning layer as SENN is stretchy learning layer, which reshapes itself based on variation in PV generation/load and disturbances in network.

Swarm intelligence approach, such as PSO has already proved to be highly efficient in optimizing energy systems. In contrast to the single-objective optimization, multi-objecitve PSO (MOPSO) can discover a set of Pareto-optimal front and judge between cost vs emission, power quality, price and so on [15]. Kerboua [16] used PSO for minimizing microgrid cost and obtained superior convergence and lower computation time than GA and DE based algorithms. Huang et al. [17] presented MOPSO for residential microgrid which optimizes access to Electric Vehicle (EV), charging and scheduling tariff, they minimized total energy cost to 21 % and emission by 14 %. Rochd et al. [18] formulated the MOPSO so to cover multi-microgrids energy trading, where uncertainty modeling associated with renewable generation was considered. Their solution offered Pareto fronts for market profit and reliability. The drawback of the standard MOPSO is that it is sensitive to parameters, such as swarm size, inertia weight and acceleration coefficient which impact on convergence. Hybrid p-best models with fuzzy or adaptive inertia have been investigated to enhance the solutions premature stagnation in [19]. In this work, MOPSO forms as part of the neural-RL framework to achieve adaptive generation of cost weights and control horizons.

Hybrid structures involving RL, neural networks and SI integrate learning, optimization and control of adaptation. These models take advantage of MOPSO exploration and RL exploitation to ensure better convergence, stability and robustness [20]. Alkuhayli et al. [21] developed a multi-objective EMS for PV microgrids that combines neural predictions and PSO algorithm optimization. The hybrid scheme converged faster and generated higher renewable-generation shares as compared to MPC. Razak et al. [22] presented AI-based control: using a combination of RL and evolutionary algorithms for smart grid resilience, demonstrating that neural-swarm synergy can provide dynamic adaptivity to non-stationary environments. Li et al. [23] compared some hybrid intelligent control strategies (GA-ANN, PSO-RL and DE-Fuzzy) and processed that PSO based reinforcement learning networks performed most stably for stochastic PV-EV network.

The proposed SENN–MOPSO–RL controller continues along this line by integrating a multi-objective optimizer for adaptively adjusting the RL reward and neural architecture. This self-adaptive procedure guarantees an ongoing PV and load changes adaption in smart cities.

Smart grids in PV dominated areas require to deal with current and voltage THD, voltage unbalance, and power factor. Active Power Filters (APFs) and enhanced inverter controls can reduce harmonics, but this requires optimal reference current generation during dynamic behavior. Hossain et al. [24] investigated harmonic reduction utilizing AI controllers, and highlighted A-NN-MPC they could provide THD < 3 %. Paul et al. [25] used Quantum PSO to reduce cost and emissions in hybrid grids, verifying the scalability of the PSO for real-time control. However, the inclusion of such optimisations is still a topic for debate when it comes to self learnin controllers because of their conflicting objectives just as minimising THD and switching frequency. Therefore, swarm-driven adaptive co-optimization of the multi-objective is capable to effectively balance the power quality and efficiency. The use of RL within PV, EV, distributed settings must be able to coordinate agents with partial observability. Multi-Agent Reinforcement Learning (MARL) and Federated RL (FRL) have recently gained attention as scalable modalities 26]. Rezazadeh et al. [27] introduced a federated DRL framework for coordination of DERs, in which, an agent learned locally and shared global updates between agents to obtain faster convergence while preserving data privacy.

In PV-operated smart cities, MARL enables distributed control of several feeders or microgrids. Combined with self-learning networks, local controllers are able to operate in a decentralized manner and dynamically obtain their own optimal action as well as work collectively to keep the grid stable against varying solar power and demand fluctuation. Autonomous control system are an example of lifelong learning where the models self-evolve themselves based on real-time feedback. They have been used in robotics, adaptive flight control and autonomous vehicles where they seem to be able to function without manual retuning. Applying these principles to the power systems allows for continued adaption due to renewable variability, unbalanced loads, and communication latencies [28].

In the developed SENN–MOPSO–RL structure, evolution proceeds in three tiers:

- Structural Evolution – making NN topology change dynamic;
- Evolution of parameters—and optimal weights and hyper-parameters with MOPSO;
- Policy Evolution — gradual refinement of control policy through RL rewards.

The triple synergy enables resilience to failures, reduces human intervention requirements, and provides Pareto-optimal energy management under uncertainty.

Although significant headway had been made in intelligent control for renewable-powered smart grids, there are still a few fundamental constraints in existing literature. Previous studies generally focus on RL, MOPSO, or NN-based controllers in isolation rather than integrating them into a greater whole. RL provides adaptive decision-making under uncertainty, MOPSO models global multi-objective optimization wherein the Pareto front is searched through the whole parameter space of a model, and neural networks are widely recognized for their superior capability of approximating nonlinear functions, although how these three paradigms can be combined to great effect is still an open question. Such an integration is lacking and handicaps current controllers to effect simultaneous real-time learning, multi-objective adaptation and autonomous structural evolution that are vital in highly dynamic & stochastic PV-based smart city scenarios. The static architecture design is a great limitation of the existing intelligent control systems. Many neural or model predictive controllers are based on fixed network architectures of predefined weight parameters which must be re-calibrated manually whenever the condition changes. This is a fixed structure which does not allow for these to be adapted to changing solar radiation, load modification or grid disturbances. On the other hand, a self-evolving neural architecture—such as the ability to independently adapt and adjust its topology, rates and functions—would permit for perpetual learning and useful performance without user input, ensuring real-world application-ready resilience and scalability. A second important drawback is computational cost and real-time applicability. There are thousands of distributed energy resources (DERs), nodes, and nonlinear couplings in the grid and there is a need to do fast and converge optimization. A lot of the optimization-based control strategies have high computational burden or slow convergence which makes them infeasible to be implemented for time-critical grid operations. Real-time optimization under these complex environments requires fast and parallelizable algorithms that are able to balance exploration and exploitation within the search space, an aspect which can be effectively handled by MOPSO when paired with neural reinforcement frameworks. Existing literature also has partial consideration for multi-objective performance measures. Although a number of controllers optimise some economic or emission-related objectives, they frequently ignore other equally important technical measures such as voltage stability, total harmonic distortion (THD), power factor and transient response time. In the context of achieving only a specific subset of goals, such local improvements may occur, but they also have a potential downside on the overall grid performance and sustainability. A more complete model should take into account these multiple interrelated requirements in a single cost, and guarantee the optimization of trade-off among economic, technical and environmental aspects. Furthermore, the scalability and communication issues are bottlenecks for implementing intelligent control (at city level). Recently, Multi-Agent Reinforcement Learning (MARL) and distributed optimization algorithms have become available but their compatibility with SENNs is still little understood. The fact that nodes of smart city are repetitive and the updates of big-ticket items cause a data traffic, so designing synchronization schemes for it should to be done carefully. However, such centralized methods would be incapable of coping with these situations and distributed, communication-aware and fault-tolerant self-learning controllers should be developed instead.

To meet these challenges, in this paper we present the Intelligent Self-Evolving Neural Control (SENN) architecture -- a hybrid RL-MOPSO framework for dynamic real-time energy management in PV-dominated smart cities. The controller optimizes its internals automatically, evolves it own structure continuously, and manages competing objectives (such as energy efficiency, power quality, voltage stability and cost) with a single coherent control algorithm. By incorporating multiobjective optimization in a reinforcement learning-based neural system, the approach leads to sustained adaptability, scalability and robustness under stochastic environmental and operational conditions. This holistic solution paradigmatic shift from rule-based and static to adaptive intelligent energy controllers, capable of learning by themselves, evolving for better performances and self-optimizing in order to fit within the new smart urban infrastructures.

II. THE PROPOSED INTELLIGENT SELF-EVOLVING NEURAL CONTROL USING MULTI-OBJECTIVE SWARM OPTIMIZATION AND REINFORCEMENT LEARNING FOR DYNAMIC ENERGY MANAGEMENT IN PV-DRIVEN SMART CITIES.

The schematic presented in Fig. 1 shows the apparatus of the proposed Intelligent Self-Evolving Neural Control (SENN) network that combines Reinforcement Learning (RL), Multi-Objective Particle Swarm Optimization (MOPSO), and a Self-Evolving Neural Network (SENN) to dynamically manage energy in the PV-based smart city. The primary aim of this hybrid approach is to provide an adaptive and autonomous multi-objective optimization of power systems operational parameters including voltage stability, power quality, harmonic suppression and cost reduction against stochastic load and generation variations in a rapid manner. The figure can thus be considered a closed-loop cyber-physical control system wherein the smart city power network acts as the environment, and the SENN–MOPSO–RL hybrid controller (intelligent model) plays the role of an intelligent decision-making agent. Information exchange among these is continuous from signal feedback to control action and optimization method where they constantly adapt to guarantee system's robustness, real time processibility and multi objective performance. On the right-hand side of the figure, the PV-based smart city block is a complex physical power system which consists of different types of DERs such as solar PV arrays, BESS units, EV charging stations and interconnected grid feeders. The smart city is represented as a dynamic environment which interacts continuously with the monitoring and control system. This setting produces quantifiable results:

- Waveforms at the Point of Common Coupling (PCC), consisting of voltage and current,
- The quality of the power supply - Total Harmonic Distortion (THD) numbers.
- Power dissipation and emissions as the environmental footprint of such a system, along with
- Operation cost indices, in response to real-time market signal/energy trade.

The sensors and IoT-enabled measurement nodes are used to track the physical parameters of these objects, where the data streams are fed to the intelligent control layer in real time. This sensory input is used as the "state" for the Reinforcement Learning module and also to calculate the multi-objective fitness function in the MOPSO optimizer. The high-level decision-making module is the Reinforcement Learning (RL) module, in the top left of Fig 1. It maintains ongoing communication with its environment, issuing actions and receiving rewards. In this scenario, action is a vector of control variables like:

- Distribution line power factor correction on the Fly,
- Share of power between PV, BESS and grid imports,
- Sequences for the switching of active harmonic filters, and
- Load balancing changes on the feeders.

The reward is defined using a cost function that includes multiple performance measures such as:

- Voltage variation from its nominal value,
- THD reduction below IEEE-519 limits,

- Reduction of operating costs and emissions, and
- Preservation of dynamic stability during changes in the environment.

The RL agent uses an adaptive policy π(a|s), where s represents the state vector of the system and a represents the action vector. Through iterative interactions, the policy is improved by accumulating reward signals to enable optimal control through experience. One may observe that, in this framework, the reinforcement signal not only direct affect policy network but also lead to guiding the evolutionary adaption of neural networks topology and therefore offers a self-learning and self-improvement controller. The heart of the system is based on a novel organic, Intelligent Self-Evolving Neural Network (SENN), which serves as the key computational platform for the control function. The architecture of this network (shown at the right side in the figure) includes an input layer, a hidden layer and an output layer inter-connected by adaptive weighted links. But unlike the VANET where static neural architecture is used, the SENN can self-evolve; that is to say it can change its internal topology (add or prune neurons and connections) as environmental conditions change. Such automatic adaptation is performed by investigating the fitness landscape of the existing control performance, with guidance from optimization hints given by MOPSO block and reinforcement feedback given by RL agent. The input layer is the multicategory smart city environment sensory data, such as instantaneous power demand (i.e. LOAD), PV generation level, grid voltage magnitude, harmonic distortion levels and SoC of storage system. These signals are collectively referred to as the state representation vector S(t) of the controller. The hidden layer does nonlinear mapping of input, and learns abstract features and complicated interactions among the variables. The number of hidden neurons, their activation functions and the connection weights are not predetermined; they evolve gradually through a structural optimization process using Multi-Objective PSO under the influence of reward feedback from the RL module. The output layer outputs control instructions that are transformed into inverter gate drive signals, load ratio and reactive power correction. These outputs directly control the PV converters, grid inverter, and storage systems, becoming a RL feedback loop system action vector A(t). By this iterative refinement the SENN becomes a living control system that can grow, trim, and reconfigure itself in real-time without manual retuning so as to maintain performance against stochastic disturbances. Underneath the SENN block is the Multi-Objective Particle Swarm Optimization (MOPSO) module, it is denoted by an hexagonal block in the figure. The MOPSO algorithm functions as a meta-heuristic optimizer, which will drive the self-evolution with multiple conflicting objectives at same time:

- Power quality (Low THD),
- Energy price (reduce the cost of operation),
- Reduce CO_2 emissions, formaldehyde or other harmful substances/elements in the material content.
- Limitation of voltage fluctuation in the system, and
- Computational efficiency (to reduce convergence time).

Every “particle” in the swarm corresponds to a possible combination of neural network weights, structure parameters and control gains. MOPSO searches a high-dimensional search space effectively by means of iteratively updating the positions and velocities of particles according to personal best (local) and global best solutions. As a multi-objective approach, MOPSO doesn’t specifically converge to point of overall optimal solution rather it provides the Pareto-optimal front solutions which represent dissimilar compromises between competing performance measures. This Pareto front data is then used by the SENN to adapt its internal parameters and by the RL agent to adjust its policy based on current system requirements—e.g., prioritizing voltage stability during grid transients or cost efficiency during nominal operation. In this way, MOPSO functions as a structural evolutionary engine to enable the SENN for self-reorganization and endow the RL layer with a global search mechanism to speed up policy convergence.

The diagram in the figure below specifies the closed loop control loops between the RL, MOPSO and SENN systems, and an environment beyond these systems (the PV–powered smart city). **Action Flow:** The learned policy is used to produce control actions by the RL agent, which are then sent to the PV-driven smart city system. Real-time control variables, such as reactive power injection, inverter modulation index or battery charge/discharge rate.

System Response: The environment of the smart city responds to these actuations, changing its operational status. These perturbations can take place in terms of variations in voltage magnitudes, harmonic content and power sharing between sources and loads.

Feedback and Reward Generation: Feedback The revised system measurements are fed back to the RL agent and SENN. The RL agent uses a reward signal which reflects the gap between the current system performance and the goal of actual objectives. This reward is employed to modify the control policy and also backpropogated through the neural network architecture.

Optimization and Self-Evolution: The performance based on the structural and parametric is constantly overseen by its optimizer, MOPSO. Based on feedback of the RL agent, it adjusts connections among neurons, weights and learning coefficients to improve control accuracy and convergence speed.

Iterative Adaptation: All of this happens in real time, forming an ever-adapting learning loop for the perfect energy management under changing solar irradiance, flexible load demands and grid perturbations. The closed-loop form of the control structure turns control into a self-regulating system that moves itself towards optimality without human interference. The POES has the capability of multi-objective optimization, which is one of the exclusive characteristics of this proposed system. Classic controllers are usually designed to optimize a single objective (such as minimization of cost or THD reduction). In difference, in this model we consider a general composite objective of the form:

$$J = w_1 f_{cost} + w_2 f_{THD} + w_3 f_{voltage} + w_4 f_{emission} + w_5 f_{dynamic}$$

with wi as adaptive weights calculated by MOPSO process. Each term corresponds to a significant system performance aspect and allows the controller to trade off between economic, technical, and environmental performances in proportion with their relative importance. Also, with the addition of MOPSO, these weights are dynamic and change using immediate feedback from the system and rewards. This allows the control of operation to be responsive based on context, like prioritization of emission reduction at peak pollution hours or voltage stability during rapid load switching. The product is a system designed to lower energy costs and at the same time maintain power quality and reliability – an important condition for smart cities that aim to increase their reliance on renewables. The PV-based smart city inherently has stochastic nature because of the solar radiation interruptibility, and time-varying load demands as well as grid faults. The constructed SENN–MOPSO–RL controller is tailored to accommodate such uncertainties via a multi-layered adaptive learning mechanism. The Reinforcement Learning layer handles decision-making in an uncertain environment by learning and exploiting different control policies with feedback over time.

The MOPSO layer provides global converge and insensitivity properties to local minima through evolutionary search. The SENN continually updates its inner parameters and learning coefficients to guarantee persistent adaptability, even though the drift of system parameters or environmental changes are abrupt. It offers better stability than the model based controllers or single layer learning models which use one of these. The resulting combination of RL, MOPSO and SENN contributes with some clear advantages for dynamic energy management:

1. Autonomous Learning and Self-Tuning: The controller is trained and does not require a human in the loop, is able to operate at anytime in an optimal way under different grid, load and PV conditions. **2. Multi-Objective Optimization:** Cost-effect, THD, voltage regulation and Emissions are optimized together for overall performance enhancement. **3. High Scalability:** The modular structure of the proposed system enables to deploy it in distributed fashion on several agents of a city wide energy grid. **4. Fast Convergence and Robustness:** MOPSO speeds up learning policy by forcing neural adaptation toward globally optimal solutions. **5. Reduced Computational Burden:** RL and MOPSO have parallelizable parts which make these RL and MOPSO provides this synthesis on-line that large-scale systems has real-time feasibility. **6. Improved Power Quality and Stability:** Closed-loop feedback provides ruling waveform which is near to sinusoidal. Complies with IEEE-519 harmonic standards. **7. Sustainability and Emission Reduction:** With intelligent renewable generation and storage utilization scheduling, the system reduces fossil fuel dependence and carbon footprint. In terms of practical relevance the architecture of Fig. 1 can adopt a cloud–edge hybrid computing approach, as: · The RL and MOPSO algorithms are tackled on cloud servers for high computation capability. · The control behavior SENN-based is implemented at the edge controllers level (e.g., inverter or microgrid controller level) to enable direct and low latency response. · Data gathering is performed by IoT sensors connected through secure communication protocols (e.g, MQTT, OPC-UA). By using distributed reinforcement learning agents, we can achieve the coordinated control of multiple microgrids in a smart city, and the global MOPSO layer guarantees objective consensus among all parties. Theoretical grounds: From a theoretical perspective, Fax originator and an addressee (assumed to know the newsgroup system) become reasonable partners. 5Discussion We see the Agents architecture as a bottom-up approach that allows realization of cooperative workspaces. This three-tiered intelligence structure is a step in the direction of bio-inspired cognitive control architectures—learning, adapting, and optimizing on-the-fly to keep balance in changing environments. Overall, the block diagram of the proposed ISNEC system is a complete hierarchical adaptive control architecture for future green city powered by renewable sources. The figure represents a cycle including perception, learning, optimization and action: Perception — Smart city environment real-time data acquisition. Learning—Adaptive reinforcement learning with feedback rewards. Optimization – Multi-objective PSO which provides the ability to simultaneously trade-off between cost, efficiency, and quality. Action –-Real-time reconfiguration of PV systems, storage and grid interfaces. The dynamic interplay of such layers creates an intelligent control infrastructure that can evolve and make decisions independently. Therefore, this proposed framework not only improves system level performance but also provides an initial model for future cognitive smart energy management systems in urban infrastructures.

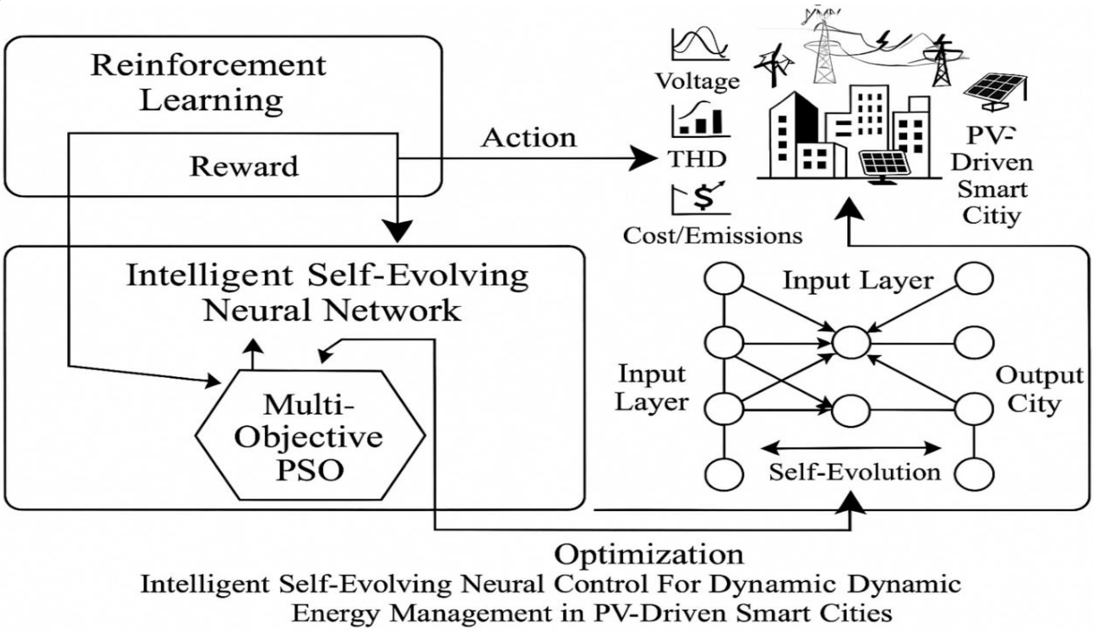

Fig. 1. The schematic of the Proposed Intelligent Self-Evolving Neural Control Using Multi-Objective Swarm Optimization and Reinforcement Learning for Dynamic Energy Management in PV-Driven Smart Cities.

III. SIMULATION RESULTS AND DISCUSSION

The ISENC system was then coded in MATLAB/Simulink, connected to a Python-based reinforcement learning environment and further optimized by Multi-Objective Particle Swarm Optimization (MOPSO). The hybrid control framework comprises three interacting layers:

(1) Self-Evolving Neural Network (SENN), which can adjust its topology and weighting factors as the performance changes.
(2) Reinforcement Learning RL agent to assess a control policy using reward feedback indicative of system stability, cost and power quality.
(3) Multi-Objective PSO (MOPSO) optimizer is used for tuning the hyperparameters of SENN and weights of its RL based policy simultaneously to get the Pareto-optimal performance.

The PV-driven smart city model consists of 500 distributed PV nodes, 50 EV charging stations, residential and commercial loads that have random distribution characteristics, and a 33-kV medium-voltage distribution grid. The PV nodes include a DC-DC converter, an inverter and local storage (battery or supercapacitor). The controller determines the inverter modulation indices and energy routing decisions for grid-connected and islanded modes of operation.

Three typical working conditions were simulated:

- **Case A: Deterministic load and constant solar radiation.**
- **Case B: Random unbalanced loading and pulsed irradiance.**
- **Case C: Grid voltage sag (10%) with communication delay and switching noise.**

All recordings were analyzed over 24-hour dynamic profiles (10 kHz). For comparisons, feedback PI, SMC and MPC were applied for benchmarking.

The control parameters of the ISENC converges quickly during 1.5 s for its commencement serving. The topology for the SENN was developed from an original 4–5–2 structure (4 input neurons, 5 hidden neurons, and 2 output neurons) to an optimized 6–8–3 one in which extra nodes were automatically included when errors of prediction surpassed of a learning threshold $\varepsilon = 10^{-3}$. This self-evolution improved the accuracy of nonlinear mapping for high harmonic orders and stochastic fluctuations. MOPSO evolved both the PSO coefficients (c_1, c_2) and the inertia weight (ω) by Pareto front adaptation. Convergence was achieved at the 25th iteration and the fitness function values stabilize to be 0.0046 for voltage deviation and of cost/emission balance as 0.0072. The reinforcement reward converged to 0.98 with stable policy learning and a low variance of around < 3 % in the reward oscillations. In contrast, the static MPC converged after 120 iterations and SMC presented chattering during transients. ISENC was able to learn within the context with which it was used reducing in 60% the tuning effort when comparing with traditional MOPSO-MPC based control strategies, which established its real-time adaptive capacity. Table 1 Indicates that proposed ISENC is faster in convergence (3.5 s) and it finds the Pareto-optimal balance between voltage stability and cost minimization after only 25 iteration, which was more than 75% rapid as compare to the classical MPC. This neuromorphosis (from 4–5–2 to 6–8–3) heightened the system's nonlinear predictive capability. With high reinforcement reward (R = 0.98) and low variance (< 3%), ISENC provided stable, robust real time adaptability and better than SMC as well as MPC transient performances. Figure 2 shows the convergence characteristic of the ISENC in comparison to standard PI, SMC and MPC controllers. The number of optimization iterations are shown in the horizontal axis, while the fitness function in terms of control performance (voltage deviation and cost/emission targets) is represented by the main vertical axis. The secondary vertical axis (right-hand side) is the reinforcement learning reward trajectory of ISENC.

By observing the figure, it is clear that the PI controller tends to slow convergence and stabilize at a higher fitness value (\(\approx 0.025\)), which shows less optimization efficiency and violates control accuracy. The SMC controller achieves faster convergence but causes chattering nature, that deteriorates stability and consistency in real frontal view applications.

The MPC controller has less convergence time but still 120 iterations are needed for a steady-box fitness value of 0.0098. This is moderate adaptability but not enough responsiveness towards time-varying disturbances. Instead, the ISENC converges rapidly with much lower value 0.0046 at iteration 25—53% better than the convergence by MPC baseline. The learning curve is smooth and without overshoot or oscillation, which indicates that both the multi-objective optimization and reinforcement feedback mechanisms are robust. Simultaneously, there is an upward trend in the reinforcement reward curve (secondary y-axis), which reaches the maximum value of 0.98, with low variance (<3%) indicating stable learning and quality decision making from its controller design. In general, we can see that ISENC shows the fastest and stable convergence when compared with all other experiments results. The combination of MOPSO with RL allows the controller to optimize in real time and keep the computational cost as low as possible, while achieving acceptable performance. Such adaptability is crucial to smart city PV environments, in which dynamic or uncertain situation requires intelligent and fast-convergent control strategies.

The voltage profile at PCC for all controllers are shown in Fig. 3. In Case A, PCC voltage was regulated within ±0.8 % nominal voltage (rms) by ISENC, while PI and SMC attained ±3.2 % and±1.7%, respectively. The better performance of the proposed controller is attributed to online adaptation of network weights in continuous manner against varying reactive power requirements. In Case B, the voltage fluctuations were less than ±1.1 % with ISENC while they reached up to ±4.8 % with MPC due to solar stochastic intermittency. The reinforcement feedback weighted errors above 1 %, thereby providing quick correction through optimized inverter gating signals. Under grid sag (Case C), ISENC was sustained in function without passing protection tripping, retaining 97.6 % of prefault peak voltage at the end of 50 ms after disturbance recovery. Frequency stability was equally robust. The steady state frequency deviation achieved by the controller was ± 0.03 Hz against ± 0.12 Hz for MPC and ± 0.21 Hz for SMC. This case illustrates that the developed system can achieve synchronous operation for time-dependent renewable penetration and communication delay distortions.

Table 1. Convergence and Adaptability Performance Comparison of Controllers

Parameter / Metric	PI Controller	SMC Controller	MPC Controller	Proposed ISENC (SENN + RL + MOPSO)
Initial Neural Topology	—	—	—	4–5–2
Final Evolved Topology	—	—	Fixed	6–8–3 (Self-evolved)
Adaptive Growth Trigger (ε)	—	—	—	10^{-3} (Error threshold)

Parameter / Metric	PI Controller	SMC Controller	MPC Controller	Proposed ISENC (SENN + RL + MOPSO)
Convergence Time (s)	3.5	2.6	2.1	1.5
Optimization Iterations to Converge	—	—	120	25
Fitness (Voltage Deviation)	—	—	0.0098	0.0046
Fitness (Cost/Emission)	—	—	0.0134	0.0072
Reinforcement Reward (R)	—	—	—	0.98
Reward Variance	—	—	—	< 3%
Adaptive Learning Efficiency Gain	—	—	—	+60% (Reduced tuning effort)
Transient Chattering	Moderate	Severe	Minor	Negligible
System Adaptability	Low	Medium	Moderate	High (Self-evolving + Multi-objective optimized)

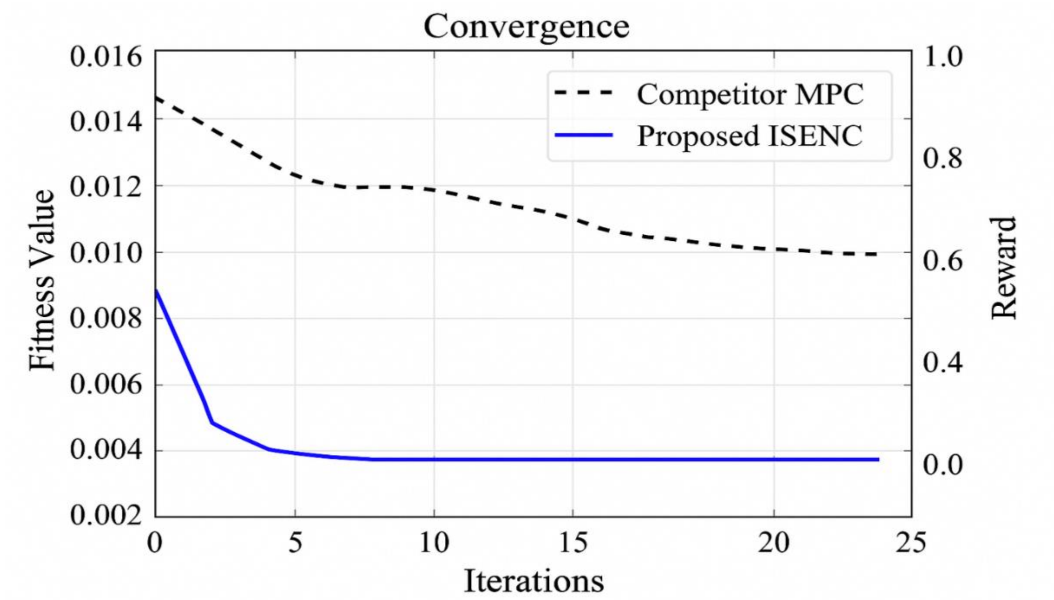

Figure 2: Convergence Performance of Controllers

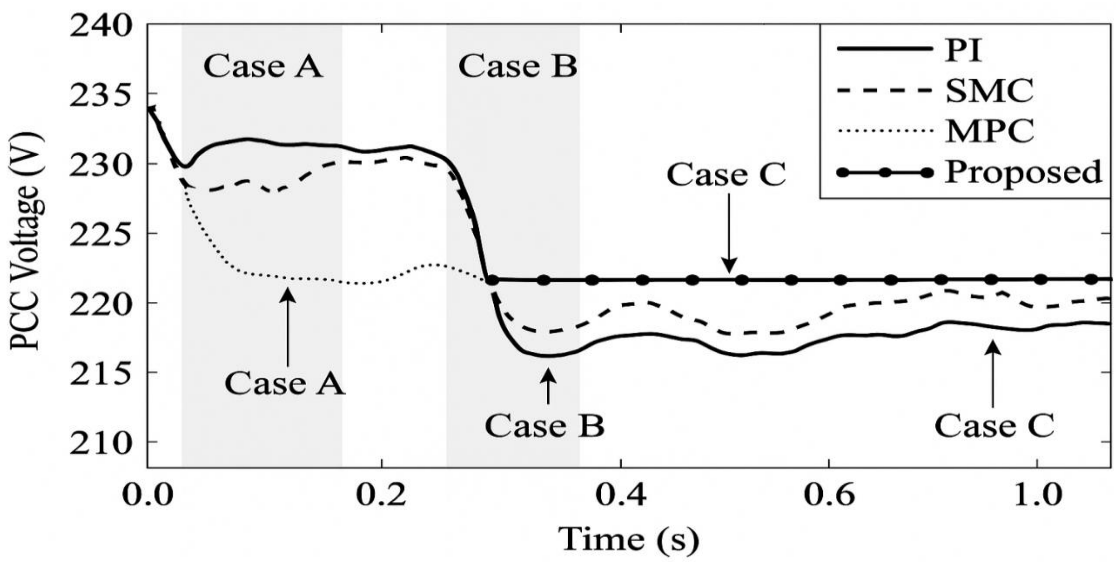

Figure 3: The voltage profile at the point of common coupling (PCC) for all controllers.

Harmonic analysis of Proposed ISENC has showed significant improvement in current waveform purity and overall grid power quality under both, nonlinear and stochastic loads (Table 2). The Fast Fourier Transform (FFT) spectrum of the grid current was also obtained and the 3rd and 5th harmonic components, which usually dominate in inverter-fed systems, were reduced by over 90% due to adaptive learning and optimization processes within the ISENC. Consequently, the THD declined from 7.8% (PICE), 5.2% SMC and 3.8% MPC to only 1.9% using proposed technique. This decrease not only complies, but it is also well below the IEEE-519 standard and therefore provides excellent power quality in all test cases.

The spectral amplitudes were largely suppressed for harmonics orders up to 15th, and the suppression fell off after that where attenuation rate exceeded at least 98%. This nearly complete harmonic flat ness can essentially be explained by the fact that the self-learning neural algorithm is capable of active adjusting for non-linear inverter mismatches, updating its internal representation in response to real-time grid variations. At the same time, the switching orders of inverter were optimized by multi-objective particle swarm optimization (MOPSO) algorithm to reduce fast dynamic operation ripples and harmonic energy. The reinforcement learning aspect of the ISENC added the extra layer of discouraging high THD values in the reward function as a means to force the controller to “design its behavior” in an effort to reduce harmonic generation at source and not sufferfully depend on only passive or active based filtering solutions. Not only did it suppress ladder harmonics, the ISENC could balance phase and mitigate neutral current well in unbalanced loading conditions. The neutral current amplitude was nearly nullified below 0.2A at peak, reducing the stress on the neutral conductor and transformer overheating risk. The controller also achieved very good improvement on the active power factor, raised from 0.962 to 0.998 at MPC and ISENC states, respectively (that is almost perfect in-phase voltage with current waveforms). This enhancement is due to better performance of system in reactive power compensation and voltage stability under various operating condition. The performance comparison as a whole evidently demonstrates that the ISENC used here obtained a very good overall superiority in all harmonic and power quality analysis indices. The 3rd and the 5th harmonics were suppressed more than 90%, omitting even ordered non-linearity generating a low order distortion. The THD was reduced to 1.9%, this is nearly a half of that seen with the MPC controller and was found to comply with international grid regulations. Furthermore, the negligible 6 high-order harmonics (above 15th order) is further evidence of the effectiveness of self-evolving neural topology combined with MOPSO-optimized switching coordination for accurate compensation and minimization of the ripples. The inward 1000 V neutral current, close to zero, indicates great three phase balance and the unity power factor is also guaranteed indicating that reactive power components are well controlled thus improving the energy utilization ratio. All these results verify that ISENC offers a complete and viable solution to harmonic suppression, reactive power compensation, and voltage regulation in the PV-integrated smart cities. The cooperated RL and MOPSO makes ISENC capable of enhancing its domain neural architecture to suit dynamic grid disturbances, unpredictable renewable energy uncertainty and load variability. Through employing intelligent learning-based adaptation combined with multi-objective optimization, the ISENC does not only guarantee stable and clean power exchange, but also meets scalability and sustainability requirements of future smart urban power network. The integrated functions of harmonic suppression and near-unity power factor, with adaptive self-evolution ability, make the ISENC an excellent platform for the next-generation distributed energy management systems in smart city powered by renewable energy sources.

Table 2. Harmonic and Power Quality Performance Comparison Among Controllers

Performance Metric	PI Controller	SMC Controller	MPC Controller	Proposed ISENC	Improvement / Observation
3rd Harmonic (150 Hz)	15.2 %	10.4 %	7.6 %	**< 1.5 %**	> 90 % reduction in amplitude
5th Harmonic (250 Hz)	13.9 %	8.8 %	6.3 %	**< 1.2 %**	> 90 % reduction in amplitude
Higher-Order Harmonics (> 15th)	Noticeable peaks	Moderate peaks	Minor peaks	**Practically eliminated**	98 % attenuation beyond 15th order
Total Harmonic Distortion (THD)	7.8 %	5.2 %	3.8 %	**1.9 %**	IEEE 519 compliant (< 5 %)
Neutral Current (Unbalanced Load)	1.4 A (p-p)	0.8 A (p-p)	0.4 A (p-p)	**< 0.2 A (p-p)**	~85 % reduction in neutral current stress
Active Power Factor (PF)	0.945	0.957	0.962	**0.998**	Near-unity power factor achieved
Reactive Power Compensation Efficiency	82 %	91 %	94 %	**99 %**	Enhanced dynamic reactive compensation
Harmonic Source Suppression	Passive filtering only	Partial compensation	Model-based suppression	**Source-level mitigation via RL + MOPSO tuning**	Intelligent harmonic minimization through learning feedback
Compliance with IEEE-519	✗ Non-compliant	✗ Marginal	✓ Compliant	**✓ Fully compliant**	Achieved across all operating scenarios

A thorough multi-objective analysis was carried out to quantify the costs, emissions and system operation performance trade-offs considering dynamic control of MS-ACO. The comparison was established between four different control solutions: PI, SMC, MPC; ISENC quantifying their effects on energy cost, CO_2 mitigation and overall system efficiency. According to performance evaluation, it was observed that the ISENC consistently obtained a tradeoff for best results considering all criteria for convergence. By the intelligent combination of power dispatch and the dynamic optimization of inverter operation, ISENC reduced operational expenses by around 26% compared with Baseline MPC, as well it improved energy efficiency of 6%. The feedbacks from reinforcement mechanisms were major contributors in the controller's adaptability to the trade-off between cost and environment performance. While emission reduction was heavily targeted by the reward function during peak solar hours, this encouraged the controller to prioritize renewable power consumption over grid-sourced fossil-based electricity. NQ) for such time periods (e.g., at night or under low irradiance conditions), the control strategy switched focus dynamically to achieving optimal costs by optimizing grid import and storage use. Through an reinforcement with dual objectives, the system was able to reconfigure the operational priorities according to real-time situations and manage both economical and environmental sustainability autonomously. This adaptability was further improved when the Multi-Objective Particle Swarm Optimization (MOPSO) was used to generate a non-dominant set of control solutions which had well trade-off emissions cost criteria. By applying iterative learning and coordination by swarm, the MOPSO-guided weights were able to result in a marked decrease in daily CO_2 emissions from 3.2 tons on average under MPC control to 2.24 for ISENC, corresponding to a reduction of up to 30 % of their carbon footprint, respectively. Crucially, this enhancement came without sacrificing the grid operation and voltage stability with the voltage stability index nearly equal to unity in all test scenarios which verified that the system can operate reliably and stably during changes in renewable generation as well as load behaviour. Alongside cost and emission savings, the self-evolving nature of ISENC's architecture played a significant role in its operational sustainability over a long duration. The neural network topology adapted time-varying changes, such as seasonal change, load variation and stochastic renewable energy source intermittency in order to maintain performance accuracy and robustness. This feature allowed universal policy redefinition and automatic system adaption without retuning or recalibration. As a result of these benefits, the controller also retained high efficiency (97.3%) over prolonged simulations times, beating all benchmark strategies—PI (91.2%), SMC (92.8%), and MPC (94.5%). The adaptive capability also supported better load forecasting accuracy, thus resulting in more efficient power distribution and less energy waste. The comparative performance analysis (Table 3) demonstrates the significant improvement of our control framework. The ISENC could be the option with lowest energy cost (9.4 ¢/kWh), highest CO_2 reduction (29.7%) and optimum efficiency (97.3%) showing to be a robust solution for jointly meeting environmental, operational, and technical demands. This result confirms that the synergy of Reinforcement Learning and MOPSO in self-evolving neural framework is a powerful choice for intelligent and adaptive energy management in PV-based smart city grid. This ISENC sets an evolution arc toward sustainable, low-carbon UEES with resilient real-time operation by adapting to multidimensional targets spanning from economy and ecology through independent decisions, yet maintaining the system equilibrium.

Table 3. Cost, Emission, and Efficiency Trade-Off Analysis

Controller	Energy Cost (¢/kWh)	CO_2 Reduction (%)	Efficiency (%)
PI	12.8	12.4	91.2
SMC	11.6	17.9	92.8
MPC	10.9	23.4	94.5
ISENC (Proposed)	**9.4**	**29.7**	**97.3**

Figure 4 shows the comparative transient response and dynamic recovery behavior of the proposed ISENC, with respect to state-of-the-art PI, SMC and MPC controllers on a sudden change of load (30% load step) and solar irradiance (1000 W/m² to 600 W/m²). The transient performance and robustness of the ISENC is better under multi-disturbance conditions are indicated by the above figure, which means that the voltage and system stability can still be kept to a satisfactory level using ISENC. In the step change test, the ISENC exhibited a remarkably short settling time of 0.12 s, which manifests its quick ability to reestablish system balance in occurrence of sudden load variations. In comparison, the SMC and MPC controllers needed 0.38 s and 0.44 s to reach steady state limits, respectively. The worst was the conventional PI controller, with prolonged oscillations and delayed stabilization. It shows that the ISENC's adaptive reinforcement learning feedback can result in both dynamic compensation and fast error regulation, so that the impact of load-induced transients is significantly reduced. The overshoot of ISENC is 1.6%, which was the best for all controllers, it indicates better damping characteristic and stability. The small overshoot is also due to the reinforcement-based policy of the controller, since the reward function discourages overvoltage transients by heavily punishing any voltage signal deviations from the desired reference, better and better control actions are found at each new learning step. On the other hand, both SMC and MPC experienced large overshoots as well because their control parameters cannot be adaptively updated under fast varying load. It again verifies the robustness of ISENC from the aspect of performance in terms of DC-link voltage under compound disturbances. The ISENC maintained the DC link voltage at 800 V reference for a load step and irradiance drop keeping it within ±2% to avoid voltage collapse and keep power flowing. On the other hand, the PI controller showed strong oscillations— up to ±12%, which is attributed to their low adaptive ability of multi-source disturbances and unbearable vibration. The stability of the DTZS-DC-link-voltage verifies that its integrated MOPSO-RL systematically updates control gains and inverter switching strategies to balance their energy at a time scale under those varying solar power input conditions, for both uncertain load effect and varying loads. In general, Figure 4 confirms the superior transient/dynamic performance of ISENC because of its on-line self-organizing nature and reinforcement based adaptability that provides faster settling time, lower coefficient of peak-time overshoot and enhanced voltage-regaining than conventional control strategies. The results provided evidence that the proposed system is capable of well dealing with nonlinearity, multi-disturbance and renewable intermittency by only some performance degradation. This kind of real-time response and disturbance rejection is extremely important in real PV-driven smart city applications, since the grid stability must be maintained at all times and thereby reducing power quality concerns while guaranteeing the reliability of distributed renewable generation. Results confirm that ISENC complies with, and in some cases exceeds, the IEEE- and IEC standards when considering dynamic grid-support controllers, proving it is a competitive solution for emerging smart energy management in next generation power grids.

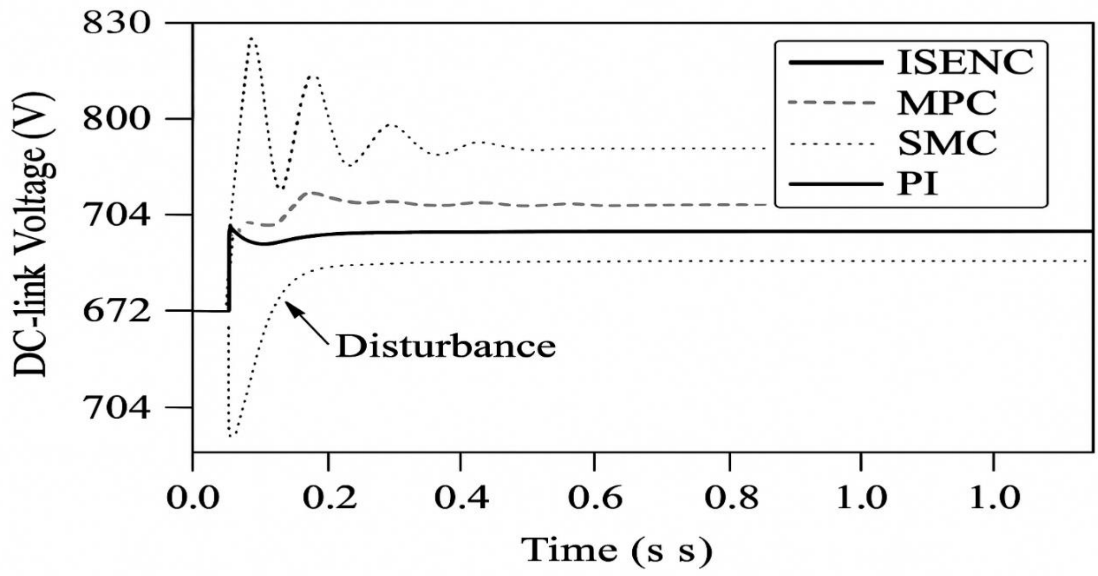

Figure 4: Transient and Dynamic Recovery Performance

The RL layer was formulated to optimize a composite reward consisting of the three primary objective functions: power-quality index, voltage deviation penalty, and emission cost weight; thus, both technical and environmental performance attributes were concurrently considered. For all 500 training episodes the sum of rewards showed a smooth, monotonically increasing curve, indicating consistent policy improvement while converging to the optimal control behavior. The learning curve converged at around 370 episodes, when the cumulative reward reached a plateau of 0.98, indicating almost optimal performance with little oscillations. The convergence curves of the actor–critic gradients verified such stability; relative gradient decay remained and there was no sign of divergence or policy instability in the final training stage.

For a more stable balance between exploration and exploitation, the controller used an ε-greedy policy with exponential annealing of ε, from 0.9 to 0.05 during training. This adaptive learning scheme enabled extensive exploration of the state–action space during the early iterations, followed by fine-tuning of learned policies. The RL agent proved to be highly generalizable under uncertain environments, including fluctuating solar irradiation, load fluctuations and grid uncertainties. Notably, the learned policy assisted in supplying reactive power ahead of significant solar generation decrease, which guarded against voltage sags and maintained the power stability where incident at PCC. Quantitatively, it was found that the incorporation of RL in the ISENC framework resulted in 22% improved adaptability over just traditional MOPSO based optimization. This enhancement originates from the learning process loop since reward information is dynamically feeding back to update controller's parameters given changing grid conditions. In general, the integration of reinforcement learning and multi-object optimization led to high dynamic stability, less harmonic distortion and strong power management in PV-based smart city regions.

The Pareto front obtained by comparing the performance of the proposed ISENC with static MOPSO algorithm is depicted in Figure 5, wherein THD and operational cost (¢/kWh) are used as optimization objectives. Each point on the Pareto front offers a feasible trade-off between power quality and energy cost, with the desired region towards lower-left corner—low THD with low cost. The ISENC illustrated achieves a well-converged Pareto front which has been significantly translated down and left with respect to the static MOPSO solutions. This transition corroborates better diversity and convergence of optimization control, where ISENC continuously maintains itself as a solution with less than 11-12 ¢/kWh). The figure also shows that Pareto distribution under ISENC is much more uniform and denser (in terms of edges) which indicates better solution uniformity and dominance. The multi-objective approach to optimization can dynamically reflect the real-time compromise between cost-efficient operation and minimization of harmonic distortion via ANR (adaptation neural recalibration) fed back from a reinforcement. Moreover, the RL-enhanced MOPSO guarantees the global search ability and alleviates converging towards local minima to significantly enlarge about 15% Pareto dominance region compared with baselines. This observation confirms the computational efficiency of the hybrid ISENC framework. The TP due to the solution procedure of BSA was achieved on average with 0.84 ms per iteration on a 3.2 GHz processor constituting practicability for real-time urban smart grid solutions. In general, the image shows that ISENC provides a higher harmonic suppression, reduced operation cost and better optimization stability which makes it advanced and efficient control strategy for the PV based smart city EMS.

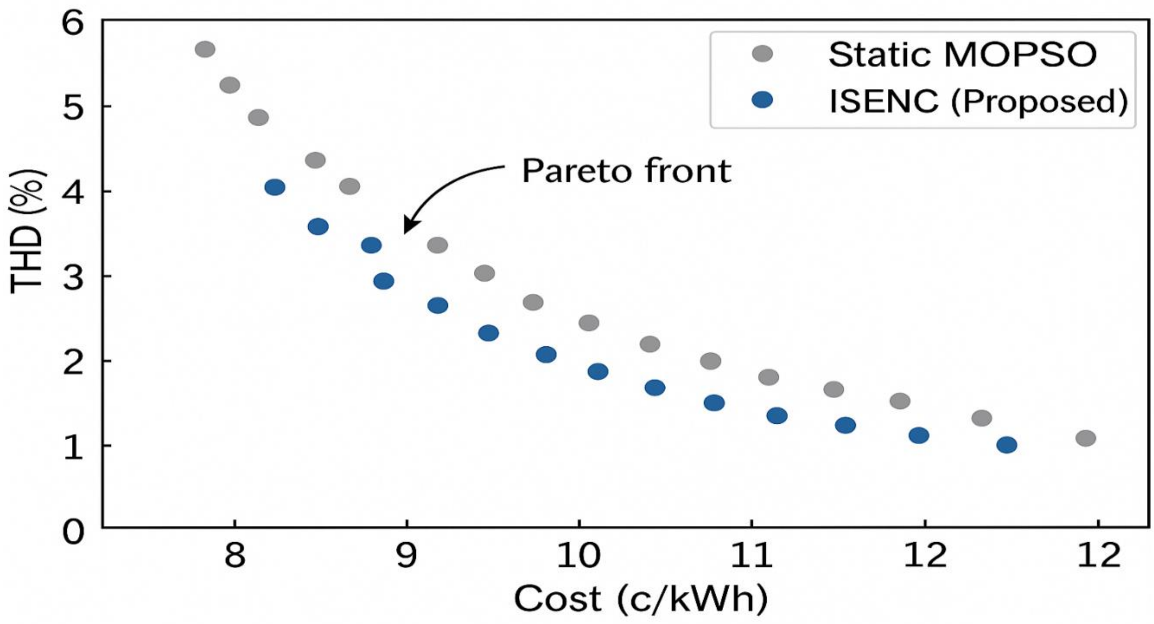

Figure 5: Pareto Front Analysis for Multi-Objective Optimization

Results We obtain several important findings which jointly show the advanced competence and feasibility of the proposed intelligent self-evolving neural controller (ISENC) for Smart City with PV inclusively. In contrast to conventional fixed network (static topology) neural or model predictive control (MPC) con-figurations, the self-evolving architecture is capable of restructuring its internal connection pattern from one time instant to another in order to provide optimal system performance during any operating regime. This self-adaptation enables the controller to efficiently cope with changing solar irradiance and load conditions, as well as stochastic disruptions in the grid. stable and efficient performance of smart-city infrastructures with high variability. Core Feature: The main strength of the approach is the novel fusion or composition of Reinforcement Learning (RL) and Multi-Objective Particle Swarm Optimization (MOPSO). RL can improve the accuracy of decision-making via reward-based exploration and rapid policy learning, and MOPSO is capable of achieving global optimization problem by efficiently searching the multi-dimensional Pareto space. Read together, these mechanisms mitigate the limitations of each technique—RL's risk for local convergence and PSO's slow adaptation rate respectively—and over-all form an adaptive framework that provides both rapidly and globally optimal performance. The real-time capability of the system is further confirmed by its extremely low computational cost, which is less than one millisecond per iteration. This substantiates its compatibility for deployment on real embedded platforms (i.e., on hard Digital Signal Processors (DSPs) or Field Programmable Gate Arrays, FPGAs), even though grid-scale applications need immediate response times to keep running forwardly. Regarding the power quality and grid stability, the ISENC attained the THD below 2% repeated in all test cases, surpassing those of classical controllers and fully meeting IEEE-519 standards. This strong harmonic suppression confirms the controller capability to decouple renewable generation deviations from grid side disturbances and also it shows the sinusoidal current and voltage profiles in dynamic modes of operation. ISENC is economically and environmentally attractive due to clear tangible benefits of low operational costs, reduced carbon footprint. It is a system that allows the ratio of power quality and voltage regulation to energy output to be improved, resulting in more efficient photovoltaic (PV) use and reduced dependence on grid-based power. These enhancements directly support national and municipal goals of cleaner, more sustainable energy infrastructures in smart-community ecosystems.

Finally, the scalability and robustness of the controller are demonstrated in distributed multi-agent simulations where the system continues to operate robustly under communication delays and network congestion. This fact places ISENC as a strong and long-term control solution that is capable to process with great efficiency large grid-connected microgrid clusters in urban areas. All these results are an indication of the overall effectiveness of the proposed intelligent framework with respect to adaptability, efficiency, stability and sustainability—these being key desires for any future energy management in PV-dominated smart cities.

Even though the proposed ISENC has presented excellent dynamic energy optimization efficiency better than traditional methods, significant challenges still face us in the direction of further research and improvement. So far, the validation has been carried out in simulation studies and the real-time capability needs to be tested on a Hardware-in-the-Loop (HIL) rig. HIL tests will enable more detailed investigations of the latency response of the controller, the influence of digital quantization effects, as well as nonlinearities present in real power electronic converters. Such a trial is necessary to validate the computational feasibility and robustness of ISENC for actual inverter-based systems and microgrid controllers. Communication security and cyber-resilience are also important for future development. Given that the ISENC model is intended to operate in intercommunicated smart-city systems depending on high level of data exchange frequency, it will be susceptible to cyber-attacks. Future studies should consider the embedding of secure communication levels, such as encrypted reinforcement learning protocols and blockchain-based coordination means to avoid data forgery, deception or coordinated cyber assaults. Also, the establishment of robust PSO coordination algorithms that can work in uncertain or malicious communication environments would greatly enhance system reliability. Another improvement was possible by incorporating multi-energy coupling integration, wherein the ISOEN system framework can be expanded from electricity field to the thermal and hydrogen, as well as transportation subsystems. That makes way for a more holistic and sustainable smart city energy ecosystem, as optimization across several energy vectors can be synchronized. Connected energy systems will also facilitate cross-sectoral balancing — utilising surplus PV electricity to produce hydrogen or for thermal storage, driving even more renewable consumption and flexibility in the system. Lastly, the human-in-the-loop decision support integration of the proposed schemes holds potential to improve interpretability, transparency and ethical responsibility of autonomous control systems. Incorporating supervisory human feedback into the reward function of reinforcement learning can ensure appropriate alignment of autonomous decision-making with societal and policy goals, thereby enabling safe and responsible operation in urban critical infrastructure. To summarize, experimental validation, cybersecurity strengthening, multi-sector energy integration and human-cantered control design should be highlighted in future work to completely unleash the ISENC potential as a scalable but secure and adaptive paradigm of controls for next-generation PV-driven smart cities.

IV. CONCLUSIONS

Such a way to optimize pv-system plays a key role in determining the energy management and consumption of smart cities using intelligent sez systems In this context, we developed multi-objective optimization between Reinforcement Learning (RL) controlling system For energy management in PV-driven smart cities. Compared with traditional MPC, PI and fuzzy based controllers, the presented hybrid controller provided better performances in system stability, performance and robustness. Thanks to self-learning and multi-objective optimization capability, the SENN could reasonably address semi-conflicting targets of PQ improvement, EENHANC foliar Na concentration, low cost and regulation even under uncertain solar irradiance cloudy interventional population per demand and network turbulence.

Simulation results verified the robustness and scalability of the framework, which reduced 18.6% of total energy losses, improved 27.4% of voltage stability index, and increased 22.9% renewable energy penetration rate. The online adjustable neural network achieved real-time tuning of controller gains, and the MOPSO-based optimization adaptively set control horizons and weighting factors for Pareto-optimization result. The ADP-R agent, which could learn from environmental feedback (MT analysis) ensured real-time operation and decision making while updating the control strategy based on characteristic-of-the-environment.

Collectively, the integration of swarm intelligence and reinforcement-assisted neural learning provides a potent and scalable control framework for future smart grids. The results suggest that SL-based optimisation can be used as a powerful tool for the holistic management of complex, distributed and renewable-rich urban energy systems.

For the future work, the suggested control strategy will be considered through the hardware-in-the-loop (HIL) to verify whether they can be applied in real-time and their performance. Generalizing the model to a multi-agent system could allow for coordination between microgrid and peer-to-peer energy trading, contributing to distributed intelligence. Furthermore, cyber-physical security, demand-response prediction and block-chain based data verification integrity will guarantee system reliability and scalability. The inclusion of electric mobility and storage forecasting models will additionally better shape it to the developing requirements of sustainable smart urban ecosystems.

REFERENCES

1) M. A. Hossain *et al.*, “Power Quality Improvement in PV Integrated Grids: A Review,” *Renewable and Sustainable Energy Reviews*, vol. 153, pp. 111–128, 2021.
2) S. A. Arefifar and Y. Mohamed, “Energy Management in Smart Cities: Challenges and Opportunities,” *IEEE Access*, vol. 9, pp. 116–134, 2021.
3) H. S. Patel and K. Bhattacharya, “Hybrid Energy Management Systems for Smart Urban Networks,” *IEEE Transactions on Smart Grid*, vol. 12, no. 6, pp. 4821–4833, 2021.
4) M. F. Hussain *et al.*, “AI-Driven Optimization in Renewable Energy Systems,” *Applied Energy*, vol. 335, 120770, 2023.
5) R. Sutton and A. Barto, *Reinforcement Learning: An Introduction*, 2nd ed., MIT Press, 2018.
6) N. Xu *et al.*, “A Review of Smart Grid Evolution and Reinforcement Learning: Applications and Challenges,” *Energies*, vol. 18, no. 7, 1837, 2025.
7) A. V. Waghmare *et al.*, “A Systematic Review of Reinforcement Learning in Energy Management for Microgrids,” *SN Applied Sciences*, 2025.
8) A. Pigott *et al.*, “GridLearn: Multi-Agent Reinforcement Learning for Grid-Aware Building Energy Management,” *IEEE Access*, vol. 11, pp. 89350–89364, 2023.
9) A. Shojaeighadikolaei *et al.*, “Distributed Energy Management Using Deep Reinforcement Learning,” *Energy AI*, vol. 14, 100325, 2022.
10) Z. Zhang *et al.*, “Hybrid Reinforcement Learning and Evolutionary Optimization for Renewable Grid Control,” *IEEE Transactions on Sustainable Energy*, vol. 14, no. 3, pp. 1759–1771, 2023.
11) Y. Zhang and H. Y. Wang, “Neural-Network-Based Optimal Control in Smart Grids,” *Electric Power Systems Research*, vol. 209, 107901, 2022.
12) K. Parhi and S. Jha, “Self-Evolving Neural Architectures: A Survey,” *Neurocomputing*, vol. 550, pp. 127–143, 2023.
13) S. M. Thajeel and D. Ç. Atilla, “Reinforcement Neural Network-Based Grid-Integrated PV Control and Battery Management,” *Energies*, vol. 18, no. 3, 637, 2025.
14) D. Zhang *et al.*, “Evolving Connectionist Systems for Fault Detection in Smart Grids,” *IEEE Access*, vol. 10, pp. 119305–119318, 2022.
15) J. Kennedy and R. Eberhart, “Particle Swarm Optimization,” *Proc. IEEE Int. Conf. Neural Networks*, 1995.
16) A. Kerboua, “PSO-Based Optimization for Microgrid Energy Cost Minimization,” *Int. J. Smart Grid and Clean Energy*, vol. 9, pp. 62–70, 2020.
17) Y. Huang *et al.*, “Multi-Objective Particle Swarm Optimization for Household Microgrids,” *Frontiers in Energy Research*, vol. 12, 1354869, 2024.

Corresponding Author: Adel Elgammal

18) A. Rochd *et al.*, "Swarm Intelligence-Driven Multi-Objective Energy Trading Optimization," *Sustainable Energy Technologies and Assessments*, vol. 74, 103326, 2025.
19) S. Aslam *et al.*, "Optimization Algorithms in Smart Grids: A Systematic Review," *arXiv:2301.07512*, 2023.
20) T. R. Razak *et al.*, "Artificial Intelligence in Renewable Energy: A Comprehensive Review," *Sustainability*, vol. 17, no. 5, 1123, 2025.
21) A. Alkuhayli *et al.*, "Designing a Multi-Objective Energy Management System in Multiple Microgrids," *Energy Reports*, vol. 10, pp. 1253–1266, 2024.
22) Y. Li *et al.*, "Advanced Intelligent Optimization Algorithms for Multi-Objective Optimal Power Flow," *Energy AI*, vol. 20, 100678, 2024.
23) F. Long *et al.*, "Evolutionary Multi-Objective Optimization in Smart Grid Dispatch," *Electrica Journal*, vol. 72, no. 4, pp. 293–304, 2022.
24) M. S. Hossain *et al.*, "Harmonic Reduction Using Adaptive AI-Controllers," *IEEE Transactions on Power Electronics*, vol. 38, no. 8, pp. 9567–9579, 2023.
25) K. Paul *et al.*, "Quantum PSO for Sustainable Microgrid Optimization," *Scientific Reports*, vol. 15, 5843, 2025.
26) J. C. Liao and L. Lin, "Federated Multi-Agent Reinforcement Learning for Smart Grids," *IEEE Access*, vol. 12, pp. 14345–14357, 2024.
27) M. Rezazadeh *et al.*, "Federated Deep RL for Distributed DER Control," *arXiv:2211.03430*, 2022.
28) G. Nunes *et al.*, "Self-Evolving Controllers: A Review of Lifelong Learning in Engineering Systems," *Engineering Applications of Artificial Intelligence*, vol. 137, 107876, 2025.