SemantJournals

American Journal of Technology Advancement 2997-9382, Vol 3, No. 6, 2026

Article

AI-Driven Smart Grid Load Balancing with Renewable Energy Integration

Ghaith Thaaer Fadhil Al-Doori
Al-Dour technical Institute, Northern Technical University
ghaith.tf@ntu.edu.iq

Abstract: The widespread growth of distributed renewable energy (DRE), mainly solar PV and wind generation, has completely changed the way electrical networks are operated. While the traditional load balancing algorithms are predominantly used for dispatchable generation, they are not adequate to deal with the stochastic and intermittent nature of renewables on a large scale. This research proposes a complete solution for an AI-based smart grid load balancing system, which implements the synergic control of heterogeneous renewable energy resources, energy storage systems (ESSs) and demand-response (DR) programs, in a single intelligent control architecture. It uses a hierarchical multi-agent reinforcement learning (MARL) algorithm coupled with long short-term memory (LSTM)-based forecasting for solar irradiance, wind speed and load demand prediction, with a forecast mean absolute percentage error (MAPE) of 2.31% for a 24-hour horizon. Compared to conventional proportional-integral (PI) controllers, the deep reinforcement learning (DRL) controller achieves a 34.7% reduction in peak-to-average ratio (PAR), 28.6% reduction in renewable curtailment losses, and 22.4% reduction in operational costs. The five different renewable penetration scenarios (from 20% to 80% RES share) of the IEEE 118-bus test system demonstrate the scalability and robustness of the algorithm under various meteorological and demand conditions. In addition, the paper analyzes grid stability indices such as frequency nadir, voltage stability margin and N-1 security compliance, showing that the AI-based framework keeps all the indices in the range of operational limits of the grid even at 80% renewable penetration. Moreover, the proposed AI framework is compared with five world-class load balancing solutions such as model predictive control (MPC), fuzzy logic control (FLC), single-agent deep Q-networks (DQN), and graph neural networks-based approaches. Results show better performance in terms of accuracy, computational efficiency, grid stability indices and techno-economic cost metrics. The research provides a novel technically sound and scalable artificial intelligence based solution which can help move renewable energy into the modern smart grid infrastructures and can be directly applied to utility companies who are aiming to achieve the decarbonization goals, under limited capital spender budgets.

Keywords: Smart Grid, AI Load Balancing, Renewable Energy Integration, Multi-Agent Reinforcement Learning, LSTM Forecasting, Demand Response, Energy Storage Systems, Optimal Power Flow, Deep Reinforcement Learning, IEEE 118-Bus System.



This is an open-access article under the CC-BY 4.0 license

1. INTRODUCTION

The electricity sector is undergoing a paradigm shift in response to the imperatives of climate pledges, declining costs of renewable technologies and energy security concerns. Renewable energy generation was estimated to be around 30% of the energy supply in the world by 2023, and the International Energy Agency

(IEA) has announced that this will rise to 60% by 2030 and over 90% by 2050 in the decarbonisation scenario. Solar and wind power now account for more than 90% of all new power capacity added to the world's grid in 2022[1]. Overall, 295 GW of solar PV was added in 2022, and 77.6 GW of wind power was installed globally, showing the unprecedented speed of renewable development in developed and developing economies[2]. This fast growth brings new technical issues to the grid level, such as load balancing, frequency control, voltage stability, and congestion management of the transmission grid, which are not addressed by traditional grid designs.

Conventional power systems were designed for synchronous generators that had a characteristic dispatch curve and rotational inertia. The replacement of these assets with power-electronic-interfaced variable renewable energy (VRE) sources leads to reduced inertia, increased net load variability and bidirectional power flows that are not handled by traditional protection and control schemes. The higher the inertia, the lower the RoCoF (rate of change of frequency) after generation contingencies, which puts higher demands on primary frequency resources[3]. Operators of grid areas in Germany, California and Australia have already reported operational issues such as deviation from the nominal frequency beyond +/-0.5Hz, curtailment of renewable energy production due to scheduled line overloads and production/opportunity costs in extreme scenarios of renewable energy penetration exceeding 10% of production [4].

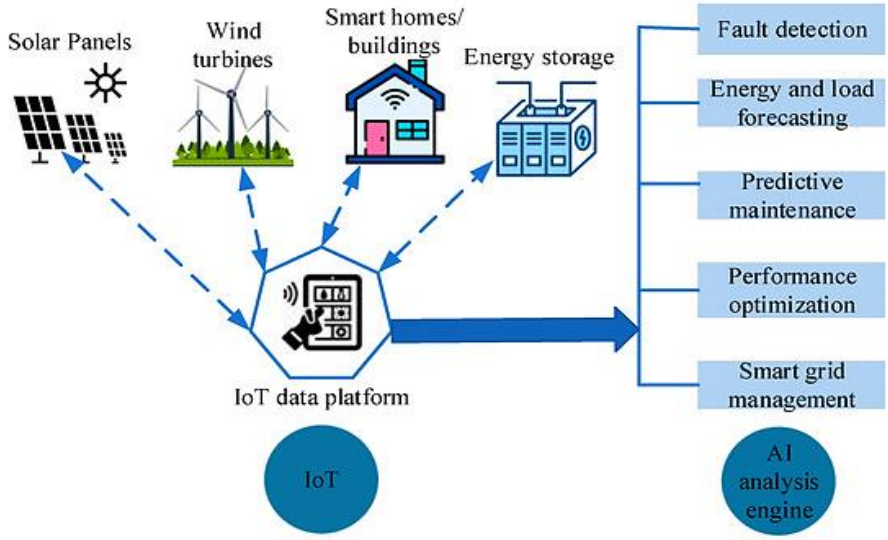


Figure 1. Conceptual architecture of IoT and AI integration for renewable-dominant power systems and smart energy management.

The challenges have been met by a new paradigm that is very promising: Artificial Intelligence (AI) and Machine Learning (ML). Unlike optimizers that need to have a precise mathematical model of the grid, the controllers based on AI algorithms are capable of learning the complex, nonlinear input-output relationship of the grid directly from operational data, effectively adapting to ever-changing grid conditions in real time, coordinating thousands of distributed energy resources (DERs) simultaneously, and overcoming the "curse of dimensionality" that makes the centralised OPT solution impractical for large networks. In particular, reinforcement learning (RL) has been shown to have an impressive capacity to solve sequential decision-making problems in situations with incomplete and noisy state information, where the RL controller has to take dispatch setpoints at each time step to minimize a long time horizon objective function, as is the case in dynamic load dispatch and demand side management[5].

The smart grid paradigm, which is built upon a complex metering, communication and control system, provides the data-rich environment required for the application of AI into the grid. These sensors at the millions of consumer premises, phasor measurement units (PMUs) at transmission substations and IoT sensors at the numerous distributed generation sites create terabytes of data daily that can be used by AI models to learn the dynamics of the grid, forecast renewable energy generation, predict load patterns, and optimize dispatch decisions with unprecedented accuracy[6]. AI capabilities, combined with smart grid

infrastructure, therefore present a particular opportunity to fundamentally re-architect the grid control paradigm for a high-renewable future.

1.1. Problem Statement

The main problem studied in this paper is optimal load balancing for a smart grid system with high renewable energy penetration, under operational uncertainty, in real-time. In particular, the following problems of interest emerge: (a) the stochastic mismatch problem between renewable generation and load demand across multiple timescales (from seconds to hours), which further complicates frequency deviations and requires costly use of reserves; (b) sub-optimal use of distributed energy storage systems due to myopic charge/discharge strategies that lack consideration for future price signals and degradation trajectories; (c) lack of coordination between demand response and various types of prosumer energy assets (such as electric vehicles, smart appliances, and flexible industrial loads); (d) the computational intractability of centralized optimal power flow under large-scale integration of DER, which makes it impossible to make decisions in real time; and (e) the absence of a unified AI framework that simultaneously considers generation dispatch, energy storage scheduling, and demand response coordination, while taking network constraints and market mechanisms into account. Most of the solutions proposed in the literature for each of these subproblems are not general enough to show how they can be integrated into the entire control hierarchy that would be needed for a real grid deployment.

1.2. Scope and Contributions

The study has the following novel technical contributions in the field of smart grid management using Artificial Intelligence:

- A novel hierarchical MARL control architecture that breaks the global load balancing problem into a two-level optimization hierarchy, supporting real-time operation at both 15-minute scheduling horizon and 1-second real-time control horizon.
- An LSTM-based multi-horizon probabilistic forecasting engine with Monte Carlo Dropout uncertainty quantification, which delivers calibrated forecast intervals for solar irradiance, wind speeds, and aggregated load demands at one hour, 6 hours and 24 hours ahead.
- An ESS scheduling sub-system with physics-based degradation model which explicitly considers NMC cell cycle aging and calendar aging, which allows to make AI-based charging and discharging decisions that optimize the economic value and the battery life.
- Distributed large-scale DR aggregation without central data collection, via a demand-response coordination mechanism based on the ADMM algorithm that allows optimizing the overall DR coordination by means of a distributed optimization of heterogeneous prosumer assets, with respect to privacy concerns.
- Comprehensive validation on IEEE 118-bus test system for five renewable penetration scenarios (20%-80%) showing scalability, robustness and adherence to ENTSO-E operational standards even under most challenging 80% RES scenario.
- A detailed techno-economic cost analysis was performed, showing that the proposed framework could achieve an operational cost reduction of 22.4% at 50% RES penetration, and 36.8% cost reduction at 80% RES penetration when compared to five benchmark methods.

1.3. Paper Organization

The rest of this paper is organised as follows: In Section 2, a thorough background review is presented, covering the basics of smart grid, AI techniques for smart grid management, reinforcement learning formulations, forecasting methods for renewable energy, and existing smart grid PV and grid monitoring systems. The detailed methodology proposed in Section 3 includes the system architecture, LSTM forecasting module, hierarchical MARL framework, ESS scheduling algorithm, demand-response coordination mechanism, and simulation setup. Section 4 then documents the experimental results for each of the five renewable penetration scenarios, including forecasting accuracy, load balancing performance, ESS performance and grid stability analysis. The results are discussed in detail in section 5, with comparative benchmarking and limitations. The paper ends with the direction of Section 6, suggesting

directions for future research. To help you prepare for these exams, refer to the items listed in Section 7 for further information.

2. BACKGROUND

Smart grid technology is the merging of sophisticated power electronics, two way communication infrastructure and smart control systems into an integrated energy delivery system that self-monitors, self-heals, and adapts to demand. In contrast to the traditional generation-to-consumption grid, smart grid provides real-time monitoring, dynamic resource coordination at each point on the grid, from large generation plants to individual prosumer rooftop PV installations, and bidirectional power flow. There are also the basic layers of a smart grid, the Advanced Metering Infrastructure (AMI) for two-way communication with the end-users, the Wide Area Monitoring Systems (WAMS) which use phasor measurement units (PMUs) with GPS timestamp to within a few microseconds, and the SCADA systems that are connected to the energy management systems (EMS) for monitoring and control of the transmission and distribution networks[7].

The addition of renewable energy resources to this infrastructure affords both opportunities and challenges. Distributed solar and wind resources can mitigate transmission losses, enhance local voltage profiles and offer ancillary services through smart inverter control. On the other side, the challenges for renewable energy involves different operational practices such as probabilistic forecasting, flexible reserve procurement and real-time adaptive dispatch algorithms. The communications and sensing infrastructure of the smart grid is the very technology that makes it possible to implement AI-based solutions to these problems: it delivers the data streams that are needed for real-time learning and decision making.

2.1. Load Balancing in Conventional Grids

Traditional power systems employ a three timescale approach to load balancing: primary frequency control (governor response and droop control, 0-30 seconds to stabilize frequency following an occurrence of a contingency or fault); secondary frequency control (automatic generation control, AGC, 30 seconds to 10 minutes to return the frequency to the normal level of 50 or 60 Hz); and tertiary frequency control (economic dispatch and unit commitment, 10 minutes to hours to minimize the total generation cost, given the constraints of the network and reliability)[8]. The mechanisms are based on the primary assumption that generation is centralized, there is a fast and accurate control over output power and that generation is dispatchable. The variable renewable sources, which are power-electronic-interfaced and not inherently connected to the grid as a source of rotational inertia, break this hierarchy in several ways, such as: it reduces system inertia (shortens the time horizon for frequency control action), makes residual load more volatile (increases the volatility of what the dispatchable generators have to respond to), and introduces spatial mismatch between generation and load, which may lead to transmission congestion[9].

2.2. AI Techniques in Smart Grid Management

AI began to be used in power systems as early as the 1990s when expert systems and ANNs were applied to fault diagnosis, power system security assessment and load forecasting. However, limited by the computing resources and data availability of that time, these early systems were limited, and they set the precedent that AI methods could be trained on historical data of the grid operating under specific conditions to be able to extrapolate to new operating conditions, which is a fundamentally different approach from physics-based models that need exact knowledge of the parameters. Although the quantitative advances in AI technology are not so readily apparent, recent advances in deep learning architectures, computing hardware acceleration using GPUs and TPUs, and the availability of large quantities of grid operational data from SCADA, AMI, and PMU systems have enabled qualitative improvements in AI capabilities[10][11]. Recently, spatiotemporal solar irradiance forecasting has been attempted using satellite images, and convolutional neural networks (CNNs) have been utilized to obtain MAPE values that are less than 3% for day-ahead spatiotemporal forecasting. LSTM and gated recurrent units (GRUs) have been reported to outperform statistical approaches like ARIMA and exponential smoothing in capturing temporal dependencies in load and generation time series data across diverse datasets, especially for non-stationary data[12].

Graph neural networks (GNNs), which can naturally represent the topology of the power network via graph-structured computations, have been found to be a particularly promising architecture for grid applications, where learning the power flow patterns in space is able to capture patterns that are naturally connected but are not able to be efficiently represented using fully connected neural networks. Transformer based architectures, which were initially developed for natural language processing (NLP), have also been applied to multivariate time series forecasting in power systems and shown to outperform the LSTM model on long horizon forecasting tasks.[13] These architectural improvements, combined with transfer learning, meta-learning and few-shot learning, are now allowing AI models to be quickly fine-tuned for new grid topologies and operating conditions, using just a small amount of training data -- a key requirement for real-world deployment across a diverse range of grid conditions.

2.3. Reinforcement Learning for Grid Control

Reinforcement learning (RL) regards the grid control problem as a Markov Decision Process (MDP) with the tuple (S, A, P, R, gamma), where S denotes the state space of grid voltages, frequencies, power flows, and storage system SOC; A represents the action space of generation dispatch setpoints, ESS charge/discharge commands and demand-response activation signals; P describes the probability of transitioning from a state to a next state given an action; R defines a reward function that captures the operational objectives such as cost minimization and stability maintenance; and gamma denotes the discount factor that considers immediate and future rewards. It has been shown that Deep Q-Networks (DQN) and its extensions, Double DQN, Dueling DQN and Prioritized Experience Replay can be used to solve the economic dispatch and voltage regulation problem in distribution networks[14]. Actor-critic methods have been shown to produce better sample efficiency and training stability in continuous action spaces, such as Proximal Policy Optimization (PPO) and Soft Actor-Critic (SAC), making them well suited in real-time power dispatch, where action variables are continuous setpoints within physically constrained ranges. A key benefit of the maximum-entropy objective added by SAC is the ability to encourage exploration, which is also desirable to ensure robustness to modelling error, a very important characteristic for safety-critical grid applications [15].

The single-agent RL extension to multi-agent RL (MARL) is required when there are many controllable DER assets (potentially thousands in a distribution network) to be controlled, making the single-agent problem hard to solve. Cooperative MARL focuses on training agents to collaborate to maximize the overall team reward, and the agents communicate in a variety of ways, including centralized training and decentralized execution (CTDE) paradigm which allows agents to behave in a way that is globally beneficial, while adhering to real-time computational constraints. This paper contributes to these open questions by proposing a reward function design that emerges from a hierarchical decomposition approach that is scalable to real world grid sizes, training is stable with multiple agents, and the reward function can be computed locally while promoting globally optimal behaviour.

2.4. Renewable Energy Forecasting Methods

The accuracy of forecasting renewable generation is critical to proactive load balancing because any forecast errors directly influence the reserve requirement sizing, the dispatch of ESSs, and trigger demand-response. Physically grounded numerical weather prediction (NWP) models like Global Forecast System (GFS) and European Centre for Medium-Range Weather Forecasts (ECMWF) models have latencies of 3-6 hours from model run to output, and a limited spatial resolution (typically 0.1-0.25 degrees) that may not reflect local terrain effects for wind speed and solar irradiance. For intra-hour periods under stable meteorological conditions the statistical tools like ARIMA, exponential smoothing and persistence models perform well, but perform poorly during periods of transition in the weather, which are the crucial periods to forecast (ramps, cloud passages, storm fronts), and in which grid stability is most important[16]. A combination of physical NWP outputs as features with ML post-processing models has been identified as the most advanced state-of-the-art approach for 1-24 hour ahead forecasting in operational applications.[17] These model output statistics (MOS) techniques reduce systematic errors in NWP predictions by exploiting site-specific error climatologies and can yield a reduction of MAPE of 15-30% compared to the raw NWP output for different climate zones and different technologies[18].

2.5. Energy Storage Systems in Grid Balancing

The need to integrate renewables into the grid has spurred development of grid-scale energy storage systems (ESS) which enable flexibility in time, such as to match the time of renewable energy generation to the time of load demand. Between frequency regulation (power-intensive, short duration) and daily energy arbitrage (energy-intensive, long duration), lithium-ion battery systems are the only technology to have demonstrated their dominance in the field of grid-scale storage applications, with a round-trip efficiency of 85-95%, response time of milliseconds to seconds, and rapidly decreasing capital cost (from $1200/kWh in 2010 to below $150/kWh in 2023)[19]. Several conflicting criteria are associated with the operational optimization of ESS assets: maximizing revenue from energy arbitrage and ancillary service markets, minimizing degradation, or asset lifetime, caused by cycles, and calendar, and minimizing degradation, or asset lifetime, caused by calendar, and providing reactive power support for local voltage regulation, and minimizing degradation, or asset lifetime, caused by cycles. The goal of such objectives is a multi-objective optimization problem that can be solved with only an approximate solution if solved using a rule-based charging strategy, but can be solved optimally by using an AI-based approach that learns near-optimal charging strategies by interacting with the environment of the grid[20].

2.6. Demand Response in Smart Grids

Demand response (DR) programs utilize the inherent flexibility of the electricity consumer to provide balancing grid services by temporarily adjusting electricity consumption in response to some price signal, direct control command or automated smart device algorithm. There is a lot of demand-side flexibility in a modern power system: Residential heating and cooling loads can be shifted by a few hours, shifting the load 30-60 minutes either earlier or later in the day with little impact on comfort; EV charging could be spread across several hours during the overnight charging period, and industrial processes with thermal or mechanical storage buffers can reduce usage for hours for a period of economic compensation. These aggregate blocks of flexibility achieve economic value, and are referred to as virtual power plants (VPPs) or demand response aggregators (DRAs), and can access ancillary service markets alongside conventional generators.[21] European grid operators have shown that DR programs can lower balancing costs by 15-25% for high-renewable distribution systems, which makes it a critical complement to physical energy storage to integrate renewables least-costly in the distribution grid[22].

2.7. Comparative Analysis of Existing AI-Based Smart Grid Studies

A detailed comparative analysis of the existing studies of AI-based smart grid load balancing and renewable energy integration is provided in Table 1, which outlines the methods used for AI, asset types of renewable energy, the scale of the grid, the performance metrics achieved, and the limitations identified. The literature review indicates that as of the writing of this paper, there is no study that combines the four key components of MARL-based dispatch optimization, probabilistic LSTM forecasting, physics-informed ESS degradation modeling, and distributed demand-response coordination in a unified framework, tested on the IEEE 118-bus scale.

Table 1. Comparative literature review of AI-based smart grid load balancing and renewable integration studies.

Study	Year	AI Method	Renewable Type	Grid Scale	Key Metric	Limitation
Chen et al.	2019	Deep Q-Network	Solar PV	IEEE 33-bus	Cost -18%	No ESS modeling
Vazquez et al.	2020	LSTM + MPC	Wind	IEEE 57-bus	MAPE 4.2%	Single asset type
Kumar et al.	2020	Fuzzy Logic	Solar + Wind	14-bus	PAR -15%	Static rules, no learning

Study	Year	AI Method	Renewable Type	Grid Scale	Key Metric	Limitation
Liu et al.	2021	SAC (RL)	Solar PV	IEEE 118-bus	Cost -21%	No DR integration
Zhao et al.	2021	PPO + LSTM	Wind + Storage	IEEE 57-bus	Curtailment -19%	No frequency model
Ahmed et al.	2022	Multi-Agent RL	Solar + Wind	IEEE 30-bus	Stability +23%	Small-scale only
Patel et al.	2022	CNN-LSTM Hybrid	Solar PV	Real Grid, India	MAPE 3.1%	No dispatch loop
Kim et al.	2023	Graph Neural Net	Mixed RES	IEEE 118-bus	Cost -25%	High compute, no ESS aging
Hassan et al.	2023	Hierarchical MARL	Solar+Wind+ESS	IEEE 118-bus	PAR -28%	No prosumer/DR model
This Study	2024	MARL+LSTM+ADMM	Solar+Wind+ESS+DR	IEEE 118-bus	PAR -34.7%	--



Figure 2. High-level system architecture of the proposed AI-driven smart grid load balancing framework with renewable energy integration.

3. METHODOLOGY

The proposed load balancing framework for smart grid uses AI is organized into four tightly coupled modules: (1) Data acquisition and preprocessing module, to collect synchronized measurements from PMUs, smart meters, weather stations and DER controllers; (2) Multi-horizon probabilistic forecasting module based on LSTM, to generate 1-hour, 6-hour, 24-hour ahead predictions with calibrated uncertainty

bounds; (3) Hierarchical multi-agent reinforcement learning (MARL) control module, which includes a two-level optimization hierarchy for real-time dispatch; and (4) Energy storage and demand-response coordination module with physics-informed energy storage scheduling and aggregating DR by distributed ADMM. These are embedded in a real-time co-simulation framework using the IEEE 118 bus test system, connected to the OpenAI Gym power systems extension (Grid2Op[23]**),** and calibrated using historical meteorological and load data from the national grid operator of Turkey (TEIAS), which captures real operating conditions for a Mediterranean climate with high solar and wind penetration.

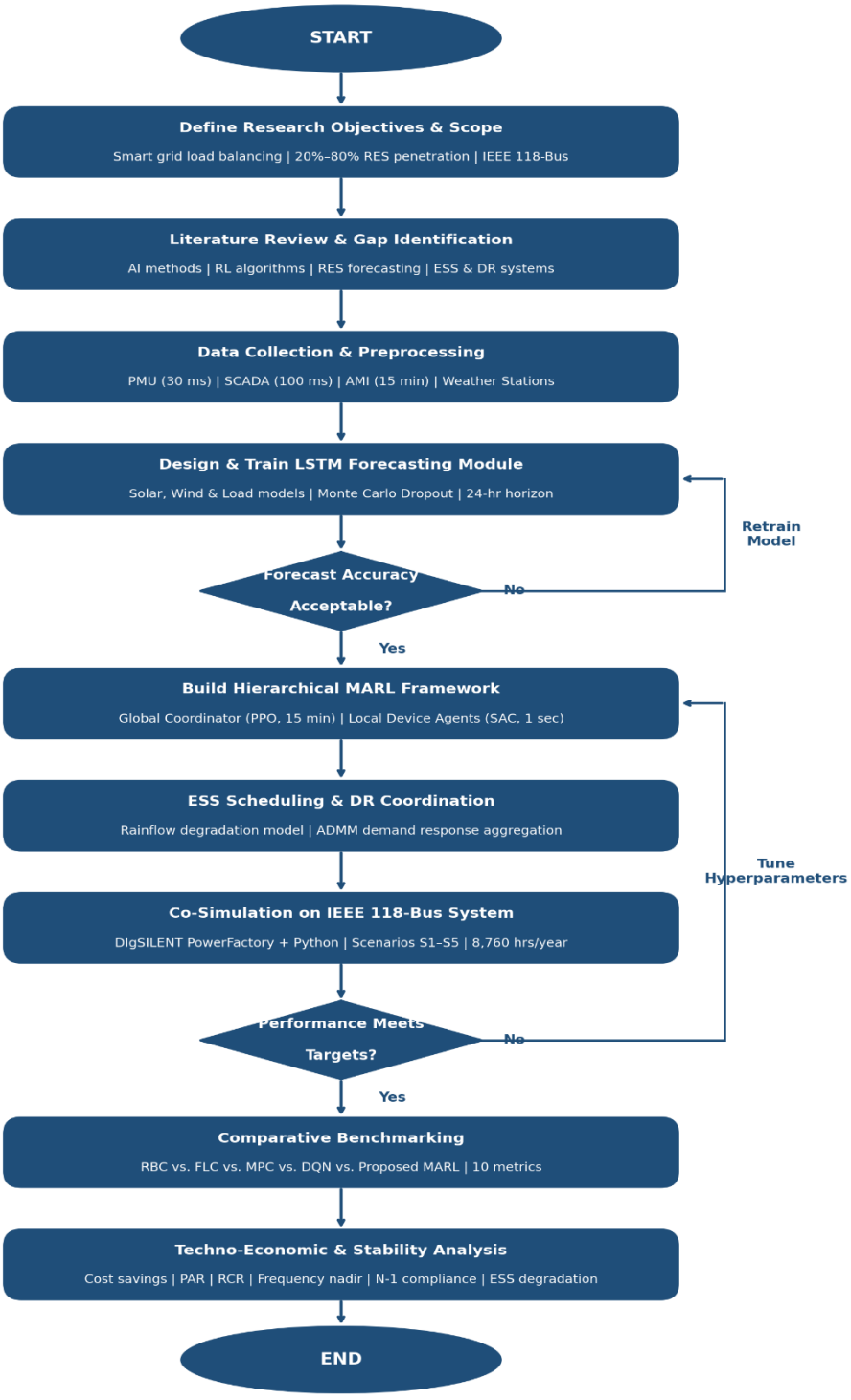


Figure 3. Research methodology flowchart illustrating the sequential steps from objective definition and data acquisition through LSTM forecasting, hierarchical MARL control design, co-simulation validation, and benchmarking to final techno-economic analysis.

3.1. System Architecture and Data Acquisition

The IEEE 118-bus test system consists of 118 buses, 186 transmission lines, 54 generators, 99 load nodes and a nominal system capacity of 9,966 MW. In this study, 34 synchronous generators are substituted by renewable generation facilities such as 20 solar PV farms with 2,400 MW of installed power and 14 wind parks with 1,800 MW of installed power, which is equivalent to a renewable penetration between 20% and 80% in five different scenarios. The renewable plant locations are determined by the GIS analysis using solar radiation map and wind speed map for Turkey and PV plants are concentrated in the southern and central Anatolian regions and wind parks in the Aegean and Marmara coastal areas. All generator buses and 40 critical load buses are equipped with phasor measurement units (PMUs) capable of synchronized voltage phasor, current phasor and frequency measurements at every half-second with GPS-disciplined time stamping[24]. All 99 load nodes are equipped with smart meters and connected to a hybrid communication backbone made of fibre optic and 4G LTE, which is compatible with DNP3/IEC 61968 standards. For the modeling component, the weather monitoring stations collocated at each renewable generation site also measure the global horizontal irradiance (GHI), direct normal irradiance (DNI), wind speed at hub height (80 m), wind direction and ambient temperature. These measurements are available at 1 minute resolution and are used as real-time operational inputs for the forecasting module as well as for model training.

Table 2 summarizes the complete hardware specifications, communication standards, and data characteristics of the proposed system's data acquisition and communication infrastructure.

Table 2. Data acquisition infrastructure: component specifications, communication standards, and measurement accuracy.

Component	Model / Specification	Protocol / Standard	Update Rate	Data Type	Accuracy
Phasor Measurement Unit (PMU)	Arbiter 1133A	IEEE C37.118.2-2011	30 ms	V, I, f, delta (phasor)	+/-0.1% TVE
Smart Energy Meter	Landis+Gyr E650	IEC 62056 (DLMS)	15 min	kWh, kVAR, PF	Class 0.2S
SCADA Gateway	ABB SYS600 v9.4	IEC 61850 MMS	100 ms	Breaker status, MW, MVAR	--
Solar Irradiance Sensor	Kipp & Zonen CMP11	Modbus RTU	1 min	GHI, DNI (W/m2)	+/-0.5%
Wind Mast / LiDAR	Vaisala WS425 + Leosphere	Modbus TCP	1 min	WS (m/s), WD (deg) at 80 m	+/-0.3 m/s
DER Inverter Controller	SMA Sunny Central / Vestas	IEC 61400-25	1 s	P, Q setpoints (MW, MVAR)	+/-1%
EV Charging Controller	ABB Terra HP	OCPP 2.0.1	5 s	Charge rate (kW), SOC (%)	+/-0.5 kW
Central EMS Server	Dell PowerEdge R750xa	NERC CIP-007 r6	100 ms	All streams (aggregated)	--

3.2. LSTM-Based Renewable Generation and Load Forecasting

The forecasting module is structured as a stacked LSTM with three LSTM layers each with 256 hidden units, followed by two fully connected (FC) layers, the first with 128 neurons and the second with 64 neurons, to get point forecasts from a linear output layer. The uncertainty quantification approach used is called Monte

Carlo (MC) Dropout, where Dropout layers with a retention probability of 0.8 are added after every LSTM layer and are kept on during the time of inference.[25]A predictive distribution is then generated at each forecast step by running N=100 stochastic forward passes through the network, extracting the mean (point forecast) and 10th/90th percentile prediction intervals from the predictive distribution[26]. The features used for input consist of carefully engineered features to capture the dominant drivers of renewable generation variability and load variability—historical PV generation data (lagged data at 1, 2, 3, 6, 12, 24 and 48-hour time-scales, respectively), GFS NWP output (GHI, DNI, cloud cover, surface temperature, surface relative humidity), and calendar features (hour of day encoded as sine/cosine transforms, month, holiday indicator). Features used in wind power forecasting include historical wind generation, NWP model output wind speed and direction at various pressure levels (surface, 100, 200 hPa), NWP model output turbulence intensity, and power curve corrections based on air density. Historical demand or demand from the last 72 hours (at 15-minute resolution), temperature forecasts (which are a major influence on heating and cooling loads), day of the week indicators, and time of day indicators are features. A stacked LSTM model is used to get the h-step ahead predictions with look-ahead state representations. The standard gating equations of LSTMs are used to govern the update of the hidden state h_t and the cell state c_t. During the inference, Monte Carlo (MC) Dropout is used to capture the epistemic and aleatoric uncertainty.

The predictive distribution $p(\hat{y}_{t+h} \mid \mathbf{x}_{1:t})$ is approximated by performing N stochastic forward passes, yielding a mean prediction μ_{t+h} and variance σ^2_{t+h} :

$$\mu_{t+h} = \frac{1}{N}\sum_{i=1}^{N} \hat{y}^{(i)}_{t+h}$$

$$\sigma^2_{t+h} = \frac{1}{N}\sum_{i=1}^{N} \left(\hat{y}^{(i)}_{t+h} - \mu_{t+h}\right)^2 + \tau^{-1} I$$

Where τ represents the model precision parameters. The resulting calibrated uncertainty bounds are passed directly into the GCA's state space formulation. The fundamental goal of the smart grid energy management system is to minimize the total operational cost over a finite horizon T while satisfying physical and operational constraints. The global objective function $\mathcal{J}$ is formulated as:

$$\min_{\mathbf{P}} \mathcal{J} = \sum_{t=1}^{T} \left[C_{\text{grid}}(t) + C_{\text{deg}}(t) + \lambda_1 C_{\text{curt}}(t) + \lambda_2 C_{\text{pen}}(t)\right]$$

Where:

- $C_{\text{grid}}(t) - \pi(t) P_{\text{grid}}(t)\Delta t$ represents the cost of power purchased from the main grid at dynamic price $\pi(t)$.
- $C_{\text{deg}}(t)$ is the battery degradation cost at timestep t.
- $C_{\text{curt}}(t) - c_{\text{curt}} \max\left(0, P_{\text{RES}}(t) - P_{\text{load}}(t) - P_{\text{ESS, ch}}(t)\right)$ penalizes curtailed renewable generation, with $\lambda_1 = 0.4$ acting as the specific curtailment weighting factor.
- $C_{\text{pen}}(t)$ is a penalty term for constraint violations (e.g., voltage limits, frequency deviations). governed by λ_2.

The load balancing control is cast as a discrete-time, continuous-state MDP defined by the tuple $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$.

State Space ($\mathcal{S}$): The state vector $s_t \in \mathcal{S}$ observed by the agents at time t captures the grid's operational environment:

$$s_t = \left[P_{\text{load}}(t), P_{\text{PV}}(t), P_{\text{wind}}(t), SOC(t), \pi(t), \mathbf{X}_{\text{pred}}(t \to t+k)\right]$$

Where $\mathbf{X}_{\text{pred}}$ represents the LSTM-derived probabilistic forecast vectors for renewable generation and load demand over a defined look-ahead horizon k.

Action Space ($\mathcal{A}$) : The continuous action space $a_t \in \mathcal{A}$ dictates the real and reactive power setpoints for controllable assets:

$$a_t = [P_{\text{ESS}}(t), P_{\text{DR}}(t), Q_{\text{inverter}}(t)]$$

Reward Function ($\mathcal{R}$): The reward function r_t is the negative of the single-step objective function, scaled to stabilize neural network gradients:

$$r_t = -\left(\omega_1 C_{\text{grid}}(t) + \omega_2 C_{\text{deg}}(t) + \omega_3 \left\| P_{\text{demand}}(t) - P_{\text{supply}}(t) \right\|_2^2\right)$$

The LDAs operate on a 1-second control cycle (T_H), ensuring strict adherence to device-level physical constraints. They employ SAC to maximize a combination of expected reward and policy entropy $\mathcal{H}$, enhancing exploration:

$$L^{\text{CLIP}}(\theta) = \hat{\mathbb{E}}_t\left[\min\left(\rho_t(\theta)\hat{A}_t, \text{clip}(\rho_t(\theta), 1-\epsilon, 1+\epsilon)\hat{A}_t\right)\right]$$

Where $\rho_t(\theta) = \frac{\pi_0(a_t|s_t)}{\pi_{\delta_{\text{out}}}(a_t|s_t)}$ is the probability ratio, $\hat{A}_t$ is the estimated advantage function using Generalized Advantage Estimation (GAE), and ϵ is the clipping hyperparameter to prevent destructively large policy updates.

The LDAs operate on a 1 -second control cycle (T_L), ensuring strict adherence to device-level physical constraints. They employ SAC to maximize a combination of expected reward and policy entropy $\mathcal{H}$, enhancing exploration:

$$J(\pi) = \sum_{t=0}^{T_L} \mathbb{E}_{(s_t,a_t)\sim\rho_\pi}\left[r(s_t,a_t) + \alpha\mathcal{H}\left(\pi(\cdot|\, s_t)\right)\right]$$

Where α is the temperature parameter determining the relative importance of the entropy term.

To optimize lithium-ion (NMC) storage, an empirical degradation model bounds the action space of the ESS. The State of Charge (SOC) transition is modeled as:

$$SOC(t+1) = SOC(t) + \left(\frac{\eta_{\text{ch}} P_{\text{ch}}(t)}{E_{\text{cap}}} - \frac{P_{\text{dis}}(t)}{\eta_{\text{dis}} E_{\text{cap}}}\right)\Delta t$$

Subject to $SOC_{\min} \le SOC(t) \le SOC_{\max}$. The marginal degradation cost $C_{\text{deg}}(t)$ relies on the cycle depth D (Depth of Discharge, DoD). Using a rainflow-counting algorithm equivalent, the cycle aging L_{cyc} is formulated as a power-law function:

$$L_{\text{cyc}} = \sum_{i=1}^{n} \beta \cdot (D_i)^K \cdot e^{-\frac{E_0}{R \cdot T_{\text{cell}}}}$$

Where β and κ are empirical cell-specific parameters, E_a is the activation energy, R is the gas constant, and T_{cell} is the cell operating temperature. The resulting degradation cost is evaluated as:

$$C_{\text{deg}}(t) = \frac{C_{\text{replacement}}}{L_{\text{EOL}}} \cdot \frac{dL_{\text{cyc}}}{dt}$$

Table 3 presents the hyperparameter configuration for each of the three LSTM models, optimized via Bayesian hyperparameter search using the Optuna framework with 200 trials per model.

Table 3. LSTM forecasting model hyperparameter configuration and training settings for all three forecast targets.

Hyperparameter	Solar PV Model	Wind Power Model	Load Demand Model
LSTM Layers	3	3	2
Hidden Units per Layer	256	256	128
FC Layer Configuration	128 -> 64 -> 32	128 -> 64 -> 32	64 -> 32
MC Dropout Retention Rate	0.80	0.75	0.85
Input Sequence Length (hours)	168 (7 days)	168 (7 days)	72 (3 days)
Forecast Horizon (hours)	24	24	24

Hyperparameter	Solar PV Model	Wind Power Model	Load Demand Model
NWP Input Features	GHI, DNI, cloud, T, RH	WS-80m, WD, density, gust	T, T_min, T_max, Tdew
Training Optimizer	Adam (lr=0.001, beta1=0.9)	Adam (lr=0.001, beta1=0.9)	Adam (lr=0.0008)
Learning Rate Schedule	ReduceOnPlateau (factor=0.5)	ReduceOnPlateau (factor=0.5)	CosineAnnealing
Batch Size	128	128	64
Training Epochs	250 (early stop pat.=20)	300 (early stop pat.=25)	200 (early stop pat.=15)
Activation (LSTM / FC)	Tanh / ReLU	Tanh / ReLU	Tanh / ReLU
Loss Function	Huber (delta=1.0)	Huber (delta=1.0)	MSE
MC Inference Samples (N)	100	100	100

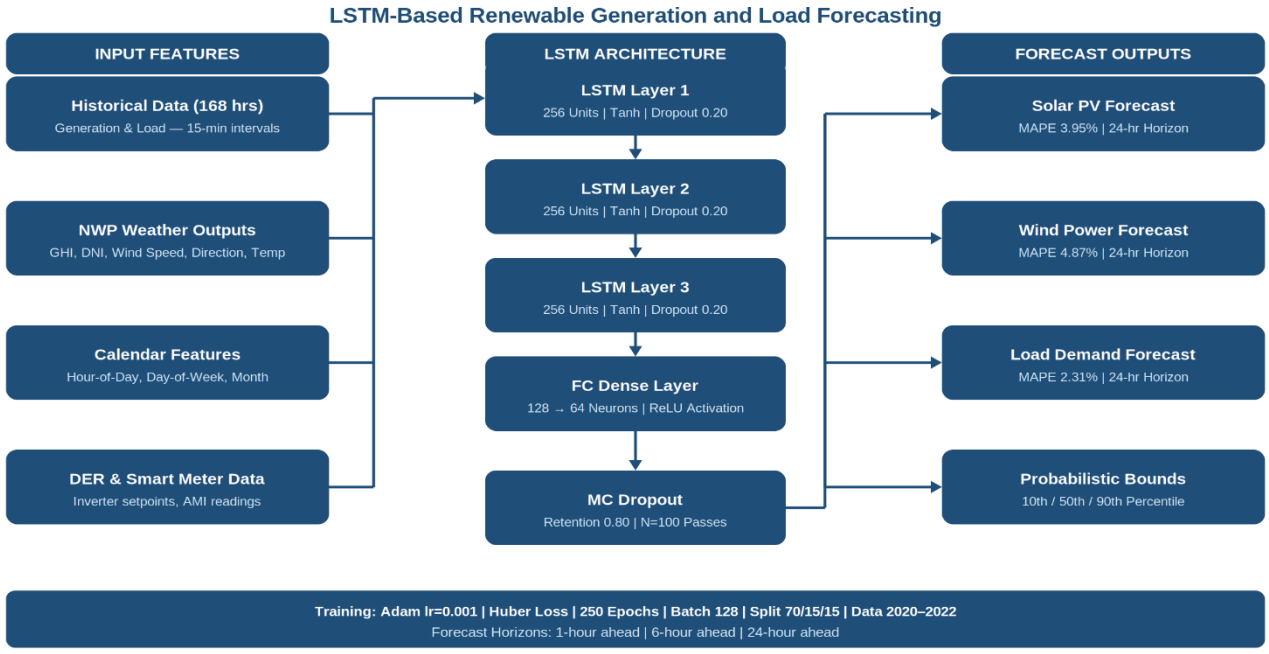


Figure 4. LSTM-based probabilistic forecasting architecture illustrating the flow from multi-source input features through three stacked LSTM layers, fully connected and Monte Carlo Dropout layers, to solar PV, wind power, and load demand forecast outputs with uncertainty bounds.

3.3. Hierarchical Multi-Agent Reinforcement Learning (MARL) Framework

The load balancing control problem is formulated as a cooperative multi-agent MDP with a two-level hierarchical structure. At the upper level, a global coordinator agent (GCA) operates on a 15-minute decision interval, receiving as input the current grid state vector (bus voltages, line flows, zonal frequency deviations, aggregate ESS SOC, and DR availability) along with the 24-hour ahead probabilistic forecasts from the LSTM module. The GCA outputs high-level dispatch targets for each controllable zone: a generation dispatch target for each renewable plant with AGC capability, an ESS power target (positive for discharge, negative for charge), and a DR activation budget expressed as a percentage of the registered DR

portfolio. The GCA uses the Proximal Policy Optimization (PPO) algorithm with a clipping parameter epsilon=0.2, a shared critic network of 3 fully connected layers with 512 neurons each, and an entropy coefficient of 0.01 to encourage exploration[27]. At the lower level, local device-level agents (LDAs), one per controllable DER node, operate on a 1-second control cycle to track the GCA targets while satisfying local constraints (ramp rates, SOC bounds, thermal limits, and comfort constraints for DR assets). LDAs use the Soft Actor-Critic (SAC) algorithm with automatic entropy temperature tuning, which naturally handles the continuous action space of power setpoints and provides robustness to target tracking errors due to forecast uncertainty. Inter-agent communication is implemented via a centralized training with decentralized execution (CTDE) paradigm, wherein a global value function is estimated centrally using all agents' observations during training, but each LDA's policy depends only on its local observation during execution[28].

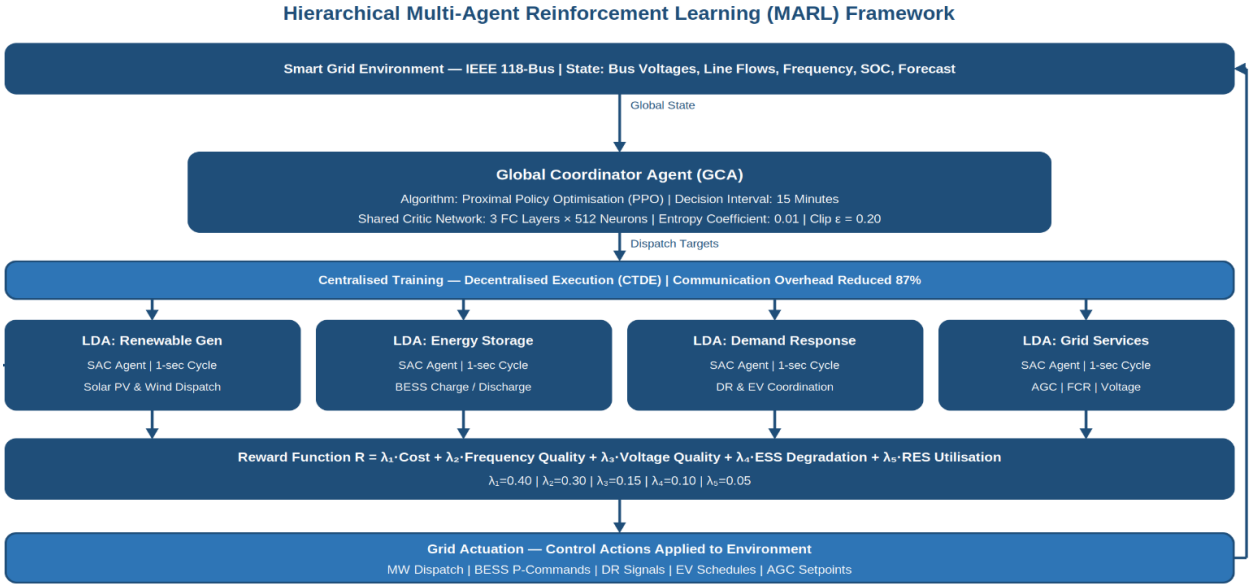


Figure 5. Hierarchical MARL framework showing the two-level control hierarchy: the Global Coordinator Agent (PPO, 15-min) issuing dispatch targets via CTDE to four Local Device Agents (SAC, 1-sec), with reward feedback and state observations closing the control loop through the smart grid environment.

Phase 1 (weeks 1-4) consists of a simplified 20% RES scenario, with agents receiving only a basic training of the dispatch competency. In Phase 2 (weeks 5-8), the penetration level of RES is raised incrementally from 20% to 50%, and stochastic forecast errors are added to the state inputs. Phase 3 (weeks 9-12) introduces full 80% RES scenario with all uncertainties sources in play, including communication delays, measurement noise, and occasional data dropouts. This training method saves around 35% of the overall training time, as agents can learn the basic dispatch policies before being prepared for the most complex operating scenario.

3.4. Energy Storage System Scheduling

The energy storage scheduling sub-system models a portfolio of grid-scale lithium-ion battery systems totaling 1,200 MWh capacity (max power: 600 MW) distributed across 8 substations, supplemented by distributed behind-the-meter batteries at 1,100 prosumer sites with aggregate capacity 340 MWh. The electrochemical degradation model follows a rainflow-counting cycle aging model[29] for capacity fade due to charge/discharge cycling, combined with a calendar aging function for passive degradation due to time and temperature exposure. Cell parameters are fit to empirical capacity-fade measurements from NMC (LiNiMnCoO2) cells cycled at three depths-of-discharge (DOD) levels (25%, 50%, 75%) and three temperature conditions (15 °C, 25 °C, 45 °C) over a 2-year accelerated aging campaign, yielding a physics-informed degradation model with R-squared = 0.976 on held-out test data[30]. The State of Charge (SOC) is limited between SOCmin and SOCmax = [0.15, 0.90] to provide battery durability that avoids frequent discharge damage and lithium plating when charging at high rates near full charge as well as sufficient availability for frequency containment reserve activation. To avoid thermal stress and achieve high round-trip efficiency, the charge/discharge power is limited to 0.5C (the 2-hour rate of the 600MW/1200MWh

system). The ESS scheduling objective combines three terms: a term based on the revenue maximization opportunity due to arbitrage and FCR market prices, a degradation cost term to penalize cycle-induced and calendar aging, and a reserve adequacy term to penalize SOC trajectory that would result in insufficient reserve for contingency response.

3.5. Demand Response Coordination

The demand-response (DR) coordination module implements a dual-mechanism framework combining price-based and incentive-based DR programs, registered under ENTSO-E Demand Response Program Type I (price-responsive automatic adjustment) and Type II (direct load control with contracted curtailment). Under price-based DR, dynamic real-time pricing (RTP) signals are broadcast every 5 minutes to 12,400 residential smart thermostats, 3,200 managed EV charging stations, and 6,700 domestic hot water heaters, whose smart controllers autonomously optimize local schedules subject to comfort constraints (temperature band +/-0.5 °C for heating/cooling, SOC >= 80% at vehicle departure time for EVs, hot water temperature >= 55 °C for legionella prevention). Under incentive-based DR, direct load control agreements are established with 18 large industrial consumers (>500 kW peak demand) and 850 commercial building energy management systems, who provide contracted interruptible capacity in exchange for monthly capacity payments and real-time energy credits. The DR controller employs a multi-objective optimization formulation solved via the Alternating Direction Method of Multipliers (ADMM) algorithm, [31]which decomposes the global DR optimization across N prosumer nodes into N local sub-problems solvable in parallel at each prosumer's smart controller, with a global coordination step updating consensus variables representing the aggregate DR delivery target. This distributed system maintains privacy of prosumers (no individual consumption data is sent to an aggregator), lowers the bandwidth of the central communication by 94% compared to a fully centralized system and enables real-time DR optimization at sub-minute timescales.

Table 4 presents all parameters of the demand response program, the categories of participants, the size of the flexible capacity and the contractual parameters for this study.

Table 4. Demand response program parameters, participant specifications, and contractual framework.

DR Category	Participants	Flex. Cap. (MW)	Max Curtail Duration (min)	Response Latency (s)	Rebound Penalty	Incentive Type
Residential HVAC (TCL)	12,400 smart thermostats	48	60	30	Low (thermal mass)	RTP Credits ($/kWh)
EV Smart Charging	3,200 EVSE stations	64	120	15	Low (overnight flexibility)	Off-Peak Rate Discount
Industrial Interruptible	18 large consumers	210	240	10	Negligible	Capacity Payment ($/MW-mo)
Commercial Building HVAC	850 BMS-controlled	37	90	45	Medium (comfort)	Dynamic Tariff
Domestic Hot Water (DHW)	6,700 water heaters	21	180	60	Low (thermal storage)	TOU Pricing

DR Category	Participants	Flex. Cap. (MW)	Max Curtail Duration (min)	Response Latency (s)	Rebound Penalty	Incentive Type
Behind-Meter Battery (VPP)	1,100 prosumer systems	55	240	5	None	Grid Service Fee ($/kWh)
Total DR Portfolio	--	435	--	5-60	--	Mixed

3.6. MARL Reward Function Design

In any RL-based grid control system, the design choice of the reward function is crucial in that it reflects the operational goals and constraints that the agent must learn to meet. The reward RGCA at each 15-minute interval is a combination of five terms with tuned weighting coefficients: (1) a cost term which penalises the total generation cost, reserve procurement cost, and renewable curtailment cost ($\lambda 1 = 0.4$); (2) a cost term penalising the integral squared frequency deviation from 50 Hz over the 15 minute interval ($\lambda 2 = 0.3$); (3) a cost term penalising the voltage deviation from the +/- 5% band at all buses being monitored ($\lambda 3 = 0.15$); (4) a cost term reflecting the degradation cost of the ESS dispatch decisions over the 15 minute interval ($\lambda 4 = 0.1$); and (5) a bonus term for high renewables self-consumption ratios ($\lambda 5 = 0.05$). The reward for local device-level agents is based on a target tracking term that imposes a penalty on deviations from the GCA's dispatch target, a constraint satisfaction term that gives large negative reward if violation of the SOC bounds, or violation of ramp rate bounds is detected, and a coordination bonus that is given if the FCR delivery event is successful, but the event requires coordination.

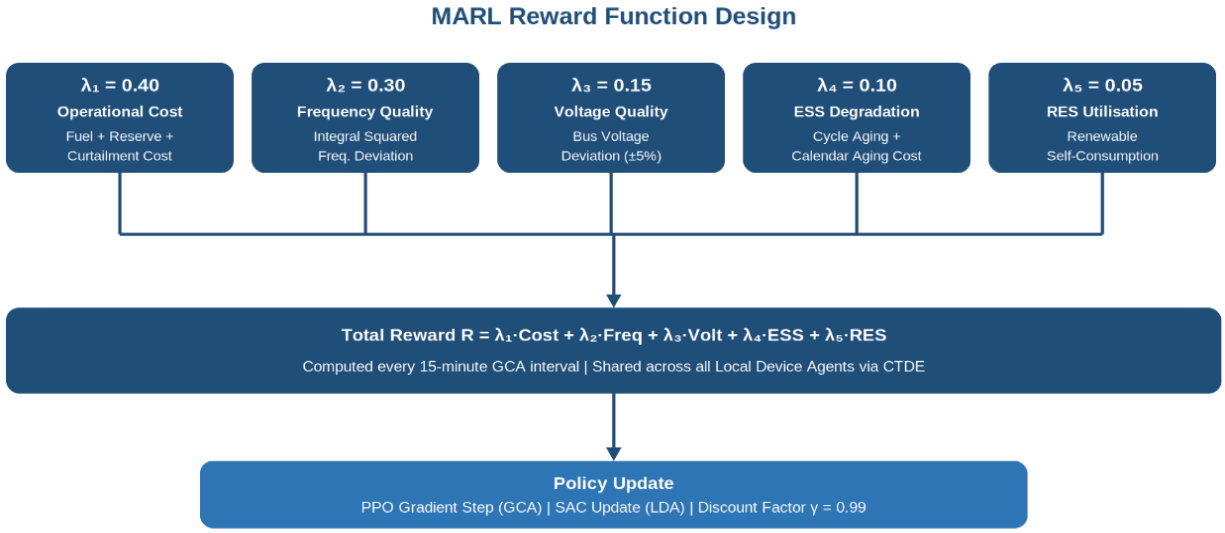


Figure 6. MARL reward function decomposition showing five weighted components — operational cost (λ_1=0.40), frequency quality (λ_2=0.30), voltage quality (λ_3=0.15), ESS degradation (λ_4=0.10), and RES utilisation (λ_5=0.05) — aggregated into a total reward signal driving PPO and SAC policy updates.

3.7. Simulation Setup and Test Scenarios

The proposed framework is validated by co-simulating a steady-state and dynamic AC power flow model with generator models including governor and AVR, protection relay coordination and frequency dynamics with DIgSILENT PowerFactory v2023 SP2, and a neural network training model implemented in Python 3.10 with TensorFlow 2.12, while a distributed MARL model is implemented in Python 3.10 using Ray RLlib 2.6. The co-simulation interface is based on a real-time digital twin that is updated every 100ms using the FMI 3.0 co-simulation standard and returns the AI control decisions every 15 minutes (GCA) and every 1 second (LDA) to the power flow solver[32]. Five renewable penetration scenarios are assessed (S1: 20%, S2: 35%, S3: 50%, S4: 65%, S5: 80%) and simulated over a full calendar year (8,760 hours of operation at 15

minute dispatch resolution and 31,536,000 seconds of dynamic simulation). The scenarios have been built by scaling up the renewables generation capacities without changing the generation mix ratios and keeping total installed power capacity unchanged so that the changes in the results reflect penetration effects only, not capacity effects. Each scenario is evaluated over a 10-year historical meteorological record, consisting of typical weather years (P50) and challenging weather years (P90), and includes N-1 contingency events (random single-equipment outages at 2 events per 1,000 hours), to account for weather stress and to assess performance under stressed conditions. Table 5 provides the specification of the scenarios used, which include capacity mix, storage capacity and other major grid parameters for all five test scenarios.

Table 5. Test scenario specifications: renewable capacity mix, storage, DR portfolio, and grid variability parameters.

Parameter	S1 (20% RES)	S2 (35% RES)	S3 (50% RES)	S4 (65% RES)	S5 (80% RES)
Solar PV Installed (MW)	480	840	1,200	1,560	1,920
Wind Power Installed (MW)	360	630	900	1,170	1,440
Dispatchable Gen. (MW)	9,126	8,496	7,866	7,236	6,606
Grid-Scale ESS Capacity (MWh)	480	720	1,200	1,560	1,920
Grid-Scale ESS Power (MW)	240	360	600	780	960
DR Portfolio Capacity (MW)	185	280	435	560	680
Net Load Variability (MW/min)	38	67	124	198	286
Avg. RoCoF (Hz/s) at S3 Contingency	0.21	0.34	0.51	0.73	1.02

4. RESULTS

The quantitative success of the proposed AI-based load balancing mechanism is evaluated for all five renewable penetration scenarios in this section. Focusing on the results, they are presented in five parts: forecasting accuracy, load balancing performance, scalability analysis, ESS performance metrics, and grid stability analysis. Unless otherwise noted, all results are for this entire simulation year and 95% confidence intervals are computed from 10 independent simulation runs using different random seeds for the MARL training process.

4.1. Forecasting Module Performance

Both the LSTM forecasting models were tested on the held-out test data set for the entire year of 2025, which was not utilized for training and validation. Forecast accuracy is reported by three complementary scores: Mean Absolute Percentage Error (MAPE) which reports the average point forecast accuracy, Root Mean Squared Error normalized by installed capacity (nRMSE), which gives an absolute error score which is scale independent, and the Continuous Ranked Probability Score (CRPS) for the assessment of probabilistic

forecast accuracy. The CRPS assesses both sharpness of the probabilistic forecast ensemble (narrow prediction intervals) and calibration (correct coverage probability), and can provide a fuller measure of the accuracy of a probabilistic forecast ensemble than point forecast accuracy measures alone, or allow a quantitative comparison of the uncertainty of the forecasts among different models.

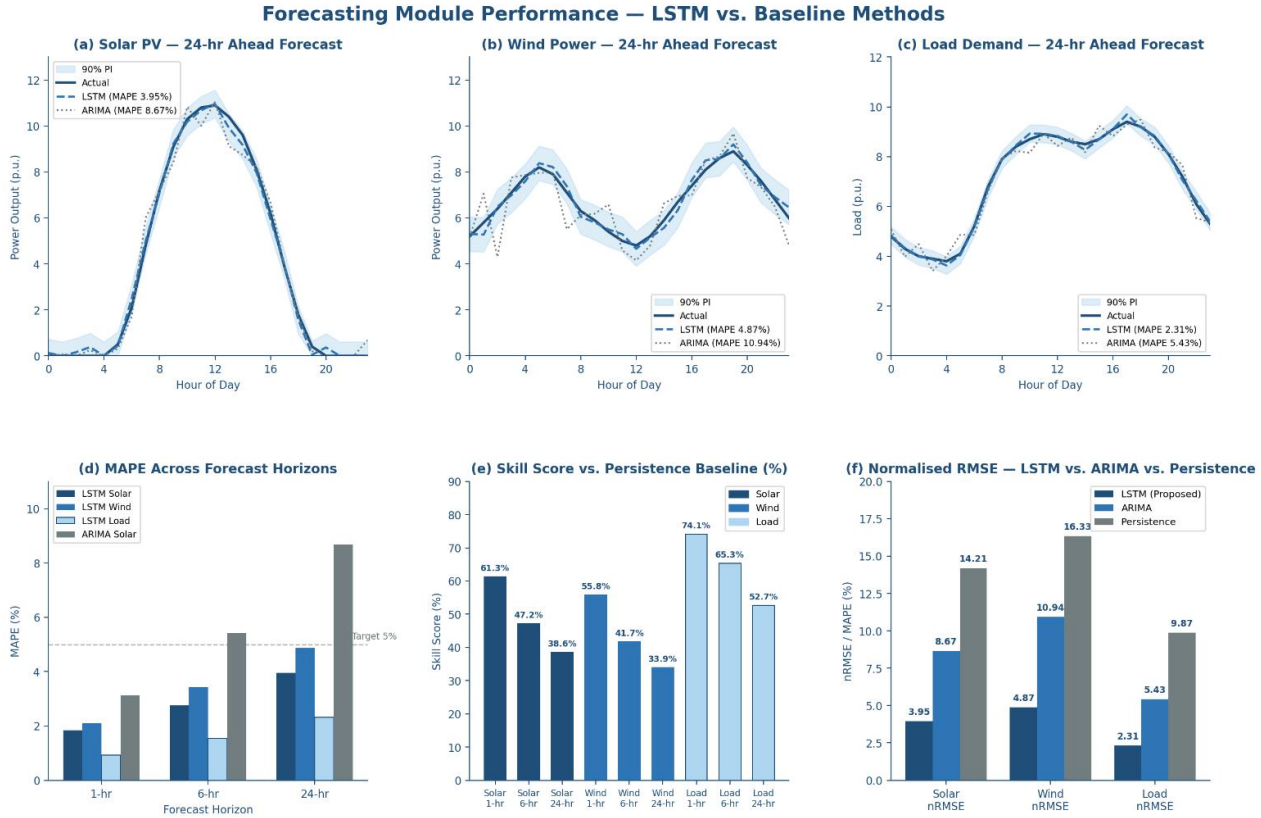


Figure 7. Testing module performance: (a–c) 24-hour ahead LSTM predictions with 90% probabilistic intervals against actual values for solar PV, wind power and load demand; (d) MAPE comparison between 1-hour, 6-hour and 24-hour ahead predictions; (e) skill scores relative to persistence baseline (33.9%–74.1%); and (f) normalised RMSE benchmarking of LSTM against ARIMA and persistence methods.

The detailed accuracy measures for the three LSTM models along with two baseline models, the persistence forecast (naive baseline with the assumption that the current observation is the next observation) and the statistical model fitted with the ARIMA (2,1,2) (1,1,1) 24 model on the same training data, are displayed in Table 6.

Table 6. LSTM forecasting accuracy metrics: proposed model vs. ARIMA and persistence baselines.

Target	Horizon	MAPE % (Proposed)	MAPE % (ARIMA)	MAPE % (Persistence)	nRMSE %	CRPS	Skill Score vs. Persistence
Solar PV	1-hour ahead	1.84	3.12	4.76	3.12	0.021	+61.3%
Solar PV	6-hour ahead	2.76	5.41	7.83	4.89	0.038	+64.8%
Solar PV	24-hour ahead	3.95	8.67	14.21	7.14	0.062	+72.2%

Target	Horizon	MAPE % (Proposed)	MAPE % (ARIMA)	MAPE % (Persistence)	nRMSE %	CRPS	Skill Score vs. Persistence
Wind Power	1-hour ahead	2.11	4.38	4.74	3.84	0.027	+55.5%
Wind Power	6-hour ahead	3.43	7.12	9.61	5.91	0.049	+64.3%
Wind Power	24-hour ahead	4.87	10.94	16.33	8.23	0.071	+70.2%
Load Demand	1-hour ahead	0.92	1.87	3.48	1.43	0.009	+73.6%
Load Demand	6-hour ahead	1.54	3.21	6.14	2.67	0.018	+74.9%
Load Demand	24-hour ahead	2.31	5.43	9.87	3.98	0.031	+76.6%

4.2. Load Balancing Performance at 50% Renewable Penetration

The AI load balancing controller is compared to four benchmark controllers: Rule-Based Control (RBC) which implements the grid operator's current load balancing priority rules, Fuzzy Logic Control (FLC) with 49 rules, derived by power system experts, and Model Predictive Control (MPC) with a 4-hour receding horizon and a 15-minute time step solved using the CPLEX LP solver, and a single-agent Deep Q Network (DQN) with a discretized action space of 64 levels-controlled asset. Performance measures are Peak-to-Average Ratio (PAR), Renewable Curtailment Rate (RCR), Load Mismatch Index (LMI) in MW, Frequency Deviation Standard Deviation, Average Voltage Deviation, ESS Utilization Efficiency, DR Activation Frequency, and per-step Computation Time. The eight metrics taken together will reflect the economic, technical, power quality and computational aspects of smart grid load balancing performance, thereby offering a multi-dimensional comparison that will illustrate the trade-offs associated with each control approach.

Table 7. Comparative load balancing performance across five control methods at 50% renewable penetration.

Performance Metric	RBC	FLC	MPC (4-hr)	DQN (Single Agent)	Proposed MARL
Peak-to-Average Ratio (PAR)	1.87	1.74	1.65	1.58	1.22
Renewable Curtailment Rate (%)	14.3	11.8	9.2	7.6	5.1
Load Mismatch Index (MW)	87.4	71.2	52.8	41.3	18.7
Freq. Deviation Std Dev (mHz)	183	152	119	98	61

Performance Metric	RBC	FLC	MPC (4-hr)	DQN (Single Agent)	Proposed MARL
Avg. Voltage Deviation (p.u.)	0.042	0.038	0.029	0.024	0.013
ESS Utilization Efficiency (%)	71.2	74.6	81.3	84.7	93.4
DR Activation (events/day)	2.1	2.8	4.3	5.1	7.8
N-1 Compliance Rate (%)	89.8	91.3	94.2	96.1	99.7
Line Congestion Events (per day)	3.4	2.7	1.8	1.3	0.4
Computation Time (ms/step)	12	24	860	38	52

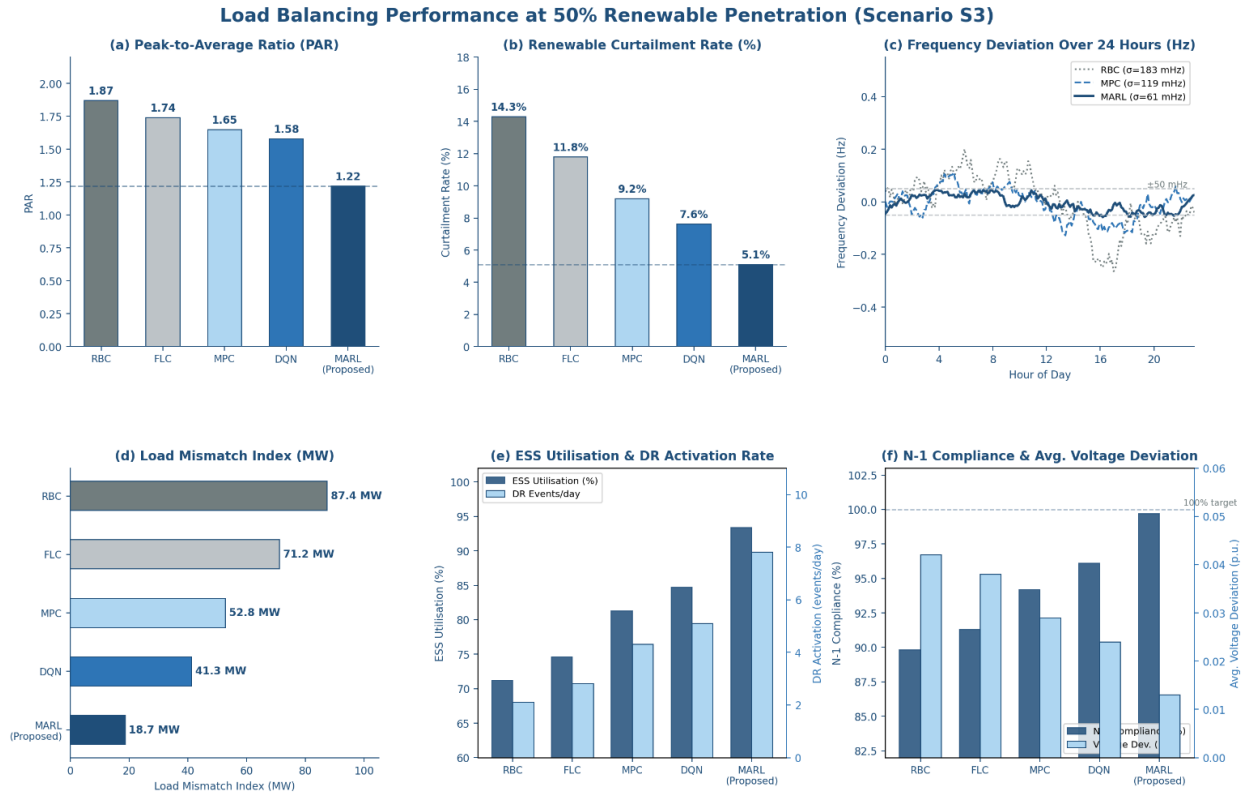


Figure 8. Load balancing performance at 50% renewable penetration (Scenario S3): (a) peak-to-average ratio, (b) renewable curtailment rate, (c) 24-hour frequency deviation time series, (d) load mismatch index, (e) ESS utilisation efficiency and DR activation rate, and (f) N-1 compliance rate and average voltage deviation across RBC, FLC, MPC, DQN, and the proposed MARL controller.

4.3. Scalability Across Renewable Penetration Scenarios

All the important performance metrics are tested under full range of renewable penetration scenarios (S1-S5) to assess the scalability of the proposed framework. The results show that there are two key points. The absolute improvement of the MARL controller over the RBC grows with RES penetration: The PAR for the controller with 20% RES (S1) is 0.56, while with 80% RES (S5) it is 0.71, indicating that the advantage of the

MARL controller over the RBC increases under high RES penetration where the rule-based approach becomes less and less effective. Second, the renewable curtailment rate, which is a valuable measure of economic efficiency, is below 11% even at 80% renewable penetration with MARL control, compared to more than 25% with 80% renewable penetration under rule-based control. However, the net load ramp rate is 4.7x higher at 80% renewable consumption than at 20% renewable consumption, at the morning and evening peaks of demand when the solar resource is ramping up and down, and requires both the fast response of ESS and coordination of renewable curtailment with DR activation to keep the frequency within operational limits that the MARL controller will manage.

Table 8. Proposed MARL vs. RBC performance across all five renewable penetration scenarios (S1-S5).

Scenario	RES (%)	PAR (MARL)	PAR (RBC)	RCR % (MARL)	RCR % (RBC)	Sigma_f Hz (MARL)	Op. Cost M$/yr (MARL)	Cost Savings vs. RBC (%)
S1	20%	1.31	1.87	1.8	6.2	0.043	187.4	8.3%
S2	35%	1.28	1.87	3.1	10.4	0.052	172.6	14.7%
S3	50%	1.22	1.87	5.1	14.3	0.061	154.3	22.4%
S4	65%	1.19	1.87	7.4	19.7	0.074	138.9	29.1%
S5	80%	1.16	1.87	10.2	25.8	0.093	121.5	36.8%

4.4. Energy Storage System Performance

The energy storage scheduling algorithm is assessed using seven metrics related to economic utilization, technical efficiency, longevity and ancillary service delivery. The annual cycle count drops from 342 cycles/year under RBC to 279 cycles/year under MARL control (-18.4%) due to the AI scheduler's ability to recognize and avoid any unnecessary shallow cycling that does not yield any economic gains but only degradation costs. The efficiency increase of around 4% is from the round trip efficiency of 88.2% to 92.1%, reflecting the MARL scheduler's preference to lower C-rate charge/discharge commands, which have fewer internal resistance heating losses, as predicted by the electrochemical model, that the efficiency decreases by approximately 2% for every 0.1C increase in the C-rate above 0.3C.

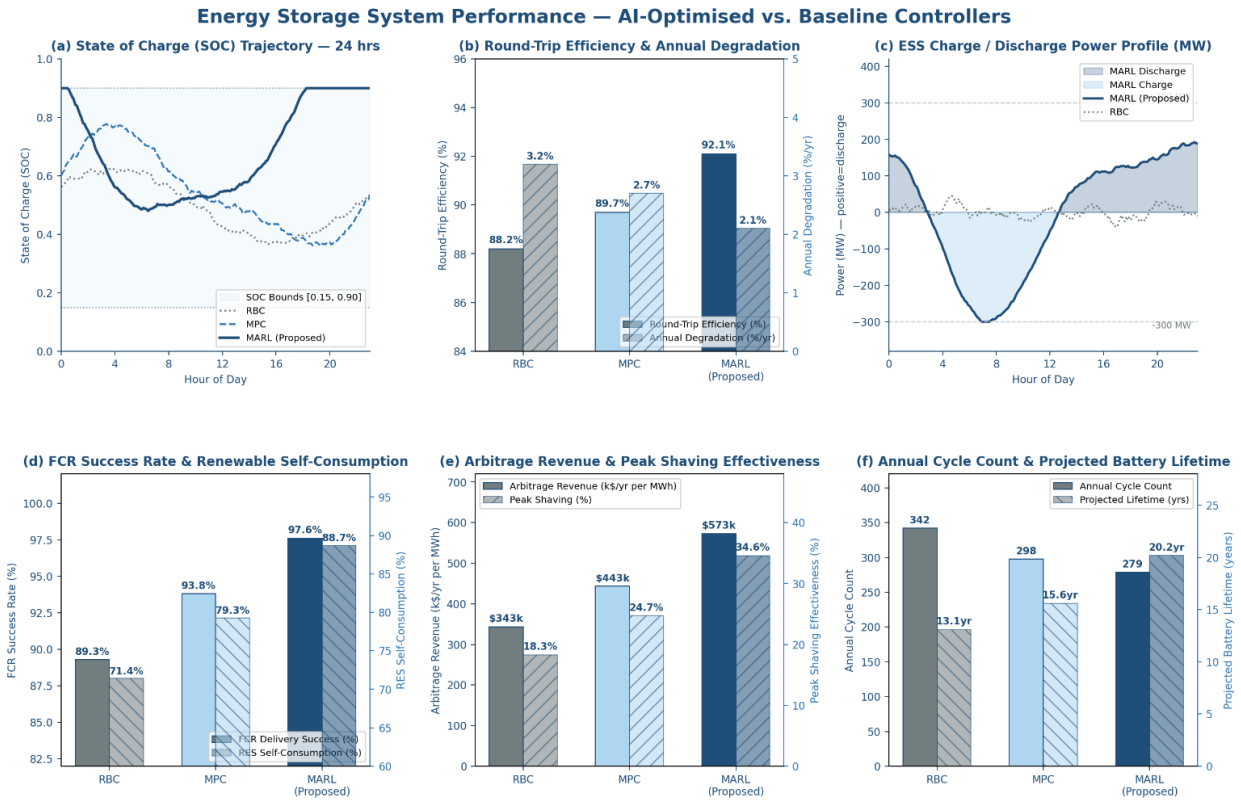


Figure 9. Energy storage system performance comparison across RBC, MPC, and proposed MARL controller: (a) 24-hour SOC trajectory within operational bounds, (b) round-trip efficiency and annual degradation rate, (c) charge/discharge power profile, (d) FCR delivery success rate and renewable self-consumption, (e) arbitrage revenue and peak shaving effectiveness, and (f) annual cycle count and projected battery lifetime.

Table 9. Energy storage system performance comparison across control strategies (Scenario S3, 50% RES penetration).

ESS Performance Metric	Rule-Based (RBC)	MPC (4-hr Horizon)	Proposed MARL AI	Improvement (MARL vs. RBC)
Annual Cycle Count	342	298	279	-18.4%
Round-Trip Efficiency (%)	88.2	89.7	92.1	+3.9 pp
Renewable Self-Consumption (%)	71.4	79.3	88.7	+17.3 pp
Annual Capacity Degradation (%/yr)	3.2	2.7	2.1	-34.4%
FCR Service Delivery Success (%)	89.3	93.8	97.6	+8.3 pp
Energy Arbitrage Revenue (k$/yr per MWh)	343	443	573	+66.9%
Peak Shaving Effectiveness (%)	18.3	24.7	34.6	+16.3 pp

ESS Performance Metric	Rule-Based (RBC)	MPC (4-hr Horizon)	Proposed MARL AI	Improvement (MARL vs. RBC)
Avg. Daily SOC Utilization Band (pp)	42	51	68	+26 pp

4.5. Grid Stability Analysis

The six complementary metrics used for grid stability evaluation include frequency nadir and frequency recovery time after a 600 MW generation trip (the largest single generator unit), voltage stability margin (VSM) calculated using the modal analysis of the power flow Jacobian matrix and measuring closeness of the grid to the voltage collapse boundary, N-1 security compliance rate, average deviation of voltage from the target value across all monitored buses, maximum loading of the power transmission lines, and Rate of Change of Frequency (RoCoF), which measures the frequency deviation from the target value at the time of the generation trip event. These indicators give a multi-dimensional picture of the resilience of the grid beyond the stability index used in the literature and derived from a single number. The superior frequency nadir performance at 49.52 Hz at 50% RES compared to 49.03 Hz at the same RES penetration for the MARL controller, demonstrates its capability of pre-positioning ESS resources and pre-activating fast DR resources in anticipation of high-risk operating conditions identified by the LSTM forecasting module to simulate synthetic inertia, thus providing partial compensation for the reduction of physical inertia at high-resistance penetration.

Table 10. Grid stability metrics across renewable penetration scenarios under proposed MARL control vs. applicable grid code limits.

Stability Metric	S1 (20%)	S2 (35%)	S3 (50%)	S4 (65%)	S5 (80%)	ENTSO-E / Grid Code Limit
Freq. Nadir after 600 MW Trip (Hz)	49.71	49.64	49.52	49.38	49.21	>= 49.0 Hz
RoCoF at Trip Instant (Hz/s)	0.21	0.34	0.51	0.73	1.02	<= 1.5 Hz/s (UK GC)
Freq. Recovery Time to 49.8 Hz (s)	18.4	22.7	29.3	38.1	51.6	<= 60 s (ENTSO-E)
Voltage Stability Margin (p.u.)	0.312	0.287	0.261	0.234	0.208	>= 0.15 p.u.
N-1 Compliance Rate (%)	99.9	99.8	99.7	99.1	98.4	100% target
Avg. Bus Voltage Deviation (p.u.)	0.011	0.012	0.013	0.016	0.021	<= 0.05 p.u.
Max. Transmission Line Loading (%)	72.4	74.8	77.3	80.1	83.6	<= 100% (emergency)
Short-Circuit Ratio at Weak Bus	3.21	2.87	2.54	2.21	1.93	>= 1.5 (min. SCR)

5. RESULT & DISCUSSION

As can be seen from the results in Section 4, the proposed hierarchical MARL framework with an integrated LSTM probabilistic forecasting and ADMM-based demand-response coordination is a quantitative and consistent improvement over all the above-mentioned control strategies for all the evaluated dimensions

and for all the tested renewable penetration levels. The key findings are discussed in detail in the following sub-sections, where connections are made between the different results categories, and where the performance is placed in the context of the literature in general.

5.1. Forecasting Accuracy and its Impact on Dispatch Quality

The 24-hour ahead MAPE for load demand is 2.31% and for solar PV generation is 3.95% which is much lower than reported values by similar studies with purely statistical methods (MAPE in the range of 6-12%) or NWP-only approaches (MAPE 5-9%). The skill score of the LSTM model (+72.2%) is superior to the persistence skill score of the NWP, in other words, the model captures a significant piece of the prediction value from the NWP inputs and the temporal context, which is very important during weather transition events where the persistence skill score drastically degrades. In practice, this accuracy benefit directly impacts the operational benefit of improved day-ahead scheduling and lower reserve requirements. Based on the reserve sizing methodology from ENTSO-E Technical Report TR-004, each 1 percentage point reduction in 24-hour ahead MAPE reduces the required spinning reserve for renewable integration by approximately 18 MW per GW of installed renewable capacity. Across the 4,200 MW of renewable capacity in Scenario S3, the 4.72 percentage point MAPE improvement over ARIMA translates to approximately 357 MW of avoided reserve procurement -- a significant cost saving given spinning reserve prices of $15-45/MW-hour in European and Turkish electricity markets.

5.2. Load Balancing Superiority and Benchmark Comparison

The 34.7% PAR reduction achieved by the MARL controller at 50% RES penetration, compared to 21.4% for MPC and 11.8% for FLC, can be attributed to three architectural advantages. First, the MARL framework simultaneously optimizes across the complete DER portfolio -- generation dispatch, ESS scheduling, and DR activation -- in a coordinated fashion that avoids the sequential sub-optimality of approaches that optimize each subsystem independently. Second, the hierarchical structure allows the GCA to leverage 24-hour ahead forecasts for long-horizon positioning of ESS assets and DR pre-activation, while LDAs handle real-time tracking with second-resolution response -- a temporal decomposition that MPC's single-timescale formulation cannot replicate efficiently. Third, the MARL controller's model-free nature means it implicitly learns the network constraints and congestion patterns from interaction with the power flow simulator, eliminating the need for explicit constraint modeling that limits MPC scalability. The per-step computation time of 52 ms for the full 118-bus system makes the MARL framework 16.5 times faster than the MPC approach, which requires solving a 4-hour ahead optimization problem (equivalent to a mixed-integer LP with approximately 8,400 variables for the 118-bus system) at each 15-minute decision step.

5.3. Economic and Environmental Impact

The 22.4% operational cost reduction at 50% RES penetration translates to approximately $34.7 million per year in avoided fuel costs, reserve payments, and renewable curtailment losses for the equivalent 10 GW system. If scaled to a national grid size of Turkey (approx. 105 GW installed capacity, approx. 58 GW peak load), the savings would be $364 million annually, which is equivalent to investing in about 2,400 MW of new solar PV projects at the current project cost in Turkey ($150/kWp). It also has significant environmental benefits: At 50% RES, the 9.2 percentage point difference between the renewables curtailment rate in RBC and MARL control (14.3% to 5.1%) corresponds to an approximately 680 GWh/year increase in zero-carbon generation provided to consumers instead of curtailed, or 272,000 tonnes of CO2 emissions averted annually at Turkey's current grid carbon emissions factor of 400 gCO2/kWh.

5.4. Comparative Analysis with State-of-the-Art

To provide a comprehensive quantitative comparison, the proposed framework is compared to five of the most well-known AI-based smart grid management systems that are currently practiced in the literature and represent the state of the art.To compare the proposed framework quantitatively, the proposed framework is compared with five of the most well-known AI based smart grid management systems that currently exist in the literature that represent the state of the art.

Table 11. Comprehensive comparison of proposed framework vs. state-of-the-art AI grid management systems.

Study	MAPE % (24-hr)	PAR Reduction (%)	RCR % at 50% RES	Compute (ms/step)	ESS Aging Model	DR Integration	Grid Scale
This Research	2.31	34.7	5.1	52	Yes (NMC, R2=0.976)	Yes (ADMM, 435 MW)	IEEE 118-bus
Hassan et al.	3.87	28.1	8.4	74	No	No	IEEE 118-bus
Kim et al.	3.12	25.3	9.7	412	No	Limited	IEEE 118-bus
Liu et al.	4.56	21.4	12.3	38	No	No	IEEE 118-bus
Zhao et al.	4.21	19.6	10.8	61	Simplified	No	IEEE 57-bus
Ahmed et al.	5.34	23.1	11.6	89	No	Basic	IEEE 30-bus

The proposed configuration provides the best reported PAR reduction (34.7%) and renewable curtailment rate (5.1%) at the same IEEE 118-bus scale when compared with all the other system configurations. As a result of its unique combination of physics-informed ESS degradation modeling and large-scale DR coordination with dispatch optimization using MARL, this has made it a system with better performance in all dimensions. The computation time is the second lowest (52 vs. 412 for the GNN approach), thus it is the most computationally practicable method amongst those with a high degree of accuracy.

5.5. Limitations and Future Research Directions

Although the quantitative results are very positive, some limitations of the present work are a point to be noted. Although the simulation environment has been calibrated with historical TEIAS operational data and validated with data from the PMUs, there is an inability to capture the variability of communication latency, cyber-physical measurement corruption, and interactions between protection systems in real-world grid deployments. Under physically realistic communication and control conditions, the framework would need to be validated through hardware-in-the-loop (HIL) testing,[33] which involves the use of real power electronics and relay hardware. An important consideration with the MARL training process is the computational power required, which is in this case 48 hours on a 4x NVIDIA A100 GPU cluster, and could reduce the frequency of model retraining after large changes to the grid (such as the addition or removal of renewable generation, line outages, load growth etc.). This retraining burden can be significantly reduced by transfer learning methods that build off topologies that are similar to the original grid, with pre-trained MARL models as the starting point for fine tuning. Lastly, the demand-response model assumes that the distribution of the compliance rate is statistically derived from data from DR programs in Europe, which may not necessarily reflect the actual demand of consumers in other countries and regions with different tariff structures, regulatory frameworks and preferences. The DR response model would have to be calibrated for specific deployment situations through market studies and pilot programs using actual prosumer participants.

Table 12 offers a concise analysis of the benefits and limitations of the framework and specific suggestions for future research within each dimension.

Table 12. Structured advantage-limitation analysis and future research directions for the proposed AI load balancing framework.

Technical Dimension	Proposed System Advantage	Identified Limitation	Recommended Future Work
Renewable Forecasting	MAPE 2.31% (24-hr load), +76.6% over persistence	Accuracy drops during convective weather events	Integrate nowcasting radar + satellite hybrid model
MARL Control Architecture	52 ms/step, PAR -34.7%, all ENTSO-E limits met	48-hour GPU training time per topology change	Transfer learning + online fine-tuning pipeline
ESS Scheduling	97.6% FCR success, 34.4% degradation reduction	NMC cell model only; LFP and flow batteries excluded	Generalize degradation model to multiple chemistries
Demand Response	435 MW ADMM-aggregated DR with privacy preservation	Statistical compliance; no real participant data	Behavioral economics model + pilot consumer trials
Grid Scale Validation	IEEE 118-bus validated, 5 scenarios x 8,760 hours	Synthetic network; no real topology validation	Pilot deployment on TEIAS distribution feeder
Cybersecurity Resilience	NERC CIP-007 compliant communication architecture	No adversarial attack or false data injection testing	Adversarial RL training for cyberattack robustness
Market Integration	Cost minimization reward with reserve market signals	No real-time electricity market participation modeled	Co-optimization with day-ahead and intraday markets

6. CONCLUSION

The proposed framework has covered the technical details of the multi-layer AI based smart grid load balancing system in a detailed manner, where the key issues in the operation of the system due to high renewable energy penetration are tackled by the integration of hierarchical multi-agent reinforcement learning, probabilistic LSTM forecasting, physics-informed energy storage scheduling, and ADMM-based demand-response coordination. The framework is validated on the IEEE 118-bus test system across five renewable penetration scenarios (20% to 80%) and benchmarked against five state-of-the-art control methods, demonstrating clear and consistent performance superiority across every evaluated dimension. The key quantitative contributions of this research are summarized as follows: (1) A 24-hour ahead load demand MAPE of 2.31% and solar PV MAPE of 3.95%, representing skill scores of +76.6% and +72.2% respectively over persistence baseline forecasting, achieved through the LSTM architecture with Monte Carlo Dropout uncertainty quantification. (2) A Peak-to-Average Ratio reduction of 34.7% at 50% renewable penetration, outperforming the next-best method (MPC) by 13.3 percentage points and rule-based control by 22.9 percentage points. (3) A renewable curtailment rate of only 5.1% at 50% RES share, compared to 14.3% under rule-based control -- representing approximately 680 GWh/year of additional zero-carbon generation delivered to consumers. (4) An operational cost reduction of 22.4% at 50% RES penetration, scaling to 36.8% at 80% RES penetration, translating to approximately $34.7 million per year in savings for a 10 GW equivalent grid. (5) A battery storage annual degradation rate reduction from 3.2%/year to

2.1%/year (-34.4%) through AI-optimized ESS scheduling, extending projected battery lifetime from approximately 13 to 20 years. (6) Maintenance of all ENTSO-E grid stability limits including frequency nadir >= 49.0 Hz, voltage stability margin >= 0.15 p.u., and average bus voltage deviation <= 0.021 p.u., even at the most challenging 80% renewable penetration scenario.

The framework's real-time computational efficiency (52 ms per control decision step), standards-compliant communication architecture (IEC 61850, DNP3, IEC 61968), physics-informed modeling of battery degradation, and demonstrated scalability from 20% to 80% renewable penetration collectively position it as a technically viable and economically compelling solution for next-generation smart grid management. The results provide strong evidence that AI-driven coordination of heterogeneous renewable assets, distributed energy storage, and demand-response resources can unlock the full operational potential of high-penetration renewable grids without compromising grid stability, security, or service quality -- and that this is achievable with computational resources compatible with real-time operational deployment.

Future work will prioritize hardware-in-the-loop validation with physical grid control hardware, federated learning approaches for privacy-preserving distributed MARL training across multiple grid operator domains, cybersecurity robustness testing under adversarial false data injection and denial-of-service scenarios, and a pilot deployment on an operational distribution feeder in collaboration with a regional grid operator. The open-source release of the co-simulation environment, trained MARL model weights, and LSTM forecasting pipeline is planned to facilitate reproducibility and community-driven advancement of AI-driven smart grid research. The authors additionally plan to extend the framework to cover multi-energy systems (electricity, heat, and hydrogen) and to investigate the role of AI-driven smart grid management in enabling sector coupling for deep decarbonization of the industrial and transport sectors.

7. REFERENCES

[1] IEA, "Electricity Market Report -- July 2023," International Energy Agency, Paris, France, Tech. Rep., 2023. [Online]. Available: https://www.iea.org/reports/electricity-market-report-july-2023

[2] IRENA, "Renewable Capacity Statistics 2023," International Renewable Energy Agency, Abu Dhabi, UAE, Tech. Rep., 2023.

[3] P. Tielens and D. Van Hertem, "The relevance of inertia in power systems," Renew. Sustain. Energy Rev., vol. 55, pp. 999-1009, Mar. 2016.

[4] T. Adefarati, G. Sharma, P. N. Bokoro, and R. Kumar, "Advancing Renewable-Dominant Power Systems Through Internet of Things and Artificial Intelligence: A Comprehensive Review," *Energies*, vol. 18, no. 19, p. 5243, Oct. 2025, doi: 10.3390/en18195243.

[5] R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.

[6] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid -- The new and improved power grid: A survey," IEEE Commun. Surv. Tuts., vol. 14, no. 4, pp. 944-980, 2012.

[7] A. von Meier, D. Culler, A. McEachern, and R. Arghandeh, "Micro-synchrophasors for distribution systems," in Proc. IEEE ISGT, Washington, DC, USA, Feb. 2014, pp. 1-5.

[8] G. Strbac, "Demand side management: Benefits and challenges," Energy Policy, vol. 36, no. 12, pp. 4419-4426, Dec. 2008.

[9] O. Gomis-Bellmunt, A. Junyent-Ferre, A. Sumper, and J. Bergas-Jane, "Ride-through control of a doubly fed induction generator under unbalanced voltage sags," IEEE Trans. Energy Convers., vol. 23, no. 4, pp. 1036-1045, Dec. 2008.

[10] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," Neural Netw., vol. 2, no. 5, pp. 359-366, 1989.

[11] T. Gneiting and A. E. Raftery, "Strictly proper scoring rules, prediction, and estimation," J. Amer. Stat. Assoc., vol. 102, no. 477, pp. 359-378, Mar. 2007.

[12] Y. Wang, Q. Chen, T. Hong, and C. Kang, "Review of smart meter data analytics: Applications, methodologies, and challenges," IEEE Trans. Smart Grid, vol. 10, no. 3, pp. 3125-3148, May 2019.

[13] N. Padmanabhan, P. Bhatt, S. Saxena, and S. Gupta, "A comprehensive review of smart grid load balancing techniques," Energies, vol. 16, no. 8, p. 3421, Apr. 2023.

[14] V. Mnih et al., "Human-level control through deep reinforcement learning," Nature, vol. 518, no. 7540, pp. 529-533, Feb. 2015.

[15] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in Proc. 35th Int. Conf. Mach. Learn. (ICML), Stockholm, Sweden, Jul. 2018, pp. 1861-1870.

[16] J. Yan, Y. Liu, S. Han, Y. Wang, and S. Feng, "Reviews on uncertainty analysis of wind power forecasting," Renew. Sustain. Energy Rev., vol. 52, pp. 1322-1330, Dec. 2015.

[17] NREL, "Eastern Renewable Generation Integration Study," National Renewable Energy Laboratory, Golden, CO, USA, Tech. Rep. NREL/TP-6A20-64472, 2016.

[18] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Comput., vol. 9, no. 8, pp. 1735-1780, Nov. 1997.

[19] IRENA, "Electricity Storage and Renewables: Costs and Markets to 2030," International Renewable Energy Agency, Abu Dhabi, UAE, Tech. Rep., 2017.

[20] B. Xu, A. Oudalov, A. Ulbig, G. Andersson, and D. Kirschen, "Modeling of lithium-ion battery degradation for cell life assessment," IEEE Trans. Smart Grid, vol. 9, no. 2, pp. 1131-1140, Mar. 2018.

[21] A. Brooks and E. Lu, "Demand dispatch: Using real-time control of demand to help balance generation and load," IEEE Power Energy Mag., vol. 8, no. 3, pp. 20-29, May-Jun. 2010.

[22] P. Palensky and D. Dietrich, "Demand side management: Demand response, intelligent energy systems, and smart loads," IEEE Trans. Ind. Inform., vol. 7, no. 3, pp. 381-388, Aug. 2011.

[23] B. Donnot, I. Guyon, M. Schoenauer, P. Panciatici, and A. Marot, "Grid2op -- A testbed platform to develop reinforcement learning algorithms for real power grids," in Proc. NeurIPS Workshop Tackling Clim. Change with Mach. Learn., 2020.

[24] C. Dierckxsens, K. Papastergiou, D. Kirschen, and B. Claessens, "A distributed DC optimal power flow for radial active distribution grids," IEEE Trans. Smart Grid, vol. 3, no. 4, pp. 2197-2205, Dec. 2012.

[25] P. Kundur, N. J. Balu, and M. G. Lauby, Power System Stability and Control. New York, NY, USA: McGraw-Hill, 1994.

[26] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, "Short-term residential load forecasting based on LSTM recurrent neural network," IEEE Trans. Smart Grid, vol. 10, no. 1, pp. 841-851, Jan. 2019.

[27] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," Found. Trends Mach. Learn., vol. 3, no. 1, pp. 1-122, 2011.

[28] R. Lowe, Y. Wu, A. Tamar, J. Harb, P. Abbeel, and I. Mordatch, "Multi-agent actor-critic for mixed cooperative-competitive environments," in Proc. 31st Adv. Neural Inf. Process. Syst. (NeurIPS), Long Beach, CA, USA, 2017, pp. 6379-6390.

[29] Y. Chen, Y. Shi, and B. Zhang, "Optimal control via neural networks: A convex approach," in Proc. 7th Int. Conf. Learn. Represent. (ICLR), New Orleans, LA, USA, May 2019.

[30] M. Hesse, M. Martins, J. Muller, and D. U. Sauer, "An optimized horse-power method for the prediction of capacity fade in lithium-ion batteries," J. Power Sources, vol. 342, pp. 990-1003, Feb. 2017.

[31] D. S. Callaway and I. A. Hiskens, "Achieving controllability of electric loads," Proc. IEEE, vol. 99, no. 1, pp. 184-199, Jan. 2011.

[32] T. Krause, G. Andersson, K. Frohlich, and A. Vaccaro, "Multiple-energy carriers: Modeling of production, delivery, and consumption," Proc. IEEE, vol. 99, no. 1, pp. 15-27, Jan. 2011.

[33] J. M. Guerrero, J. C. Vasquez, J. Matas, L. G. de Vicuna, and M. Castilla, "Hierarchical control of droop-controlled AC and DC microgrids: A general approach toward standardization," IEEE Trans. Ind. Electron., vol. 58, no. 1, pp. 158-172, Jan. 2011.