International Journal of Electrical Engineering and Sustainability (IJEES)
ISSN (online): 2959-9229
https://ijees.org/index.php/ijees/index
Arab Impact Factor: 1.51 Dolj 2026 = Q1
Volume 4 | Number 2 | July-December 2026 | Pages 48-62

Reducing Renewable Energy Curtailment in Solar–Wind Integrated Power Grids Using Uncertainty-Aware Multi-Agent Reinforcement Learning

Abdussalam Ali Ahmed[1*], Omar A. M. Edbeib[2]
[1,2]Mechanical and Industrial Engineering Department, Bani Waleed University, Bani Waleed, Libya

Email: [1]abdussalam.a.ahmed@gmail.com, [2]aldbib220@yahoo.com

Received: April 16, 2026 Accepted: June 28, 2026 Published: July 23, 2026
DOI: 10.65998/ijees.v4i2.173

Abstract: High levels of solar and wind generation may exceed local electricity demand and available export capacity, thereby necessitating renewable-energy curtailment. This study evaluates an uncertainty-aware Multi-Agent Proximal Policy Optimization framework for coordinating battery storage and flexible demand. The analysis uses hourly German load, solar generation, wind generation, and electricity-price data obtained from Open Power System Data for the period from January 2015 to June 2020. Data from 2015 to 2018 are used for model training, 2019 is reserved for forecast calibration, and the first half of 2020 is retained as a fully held-out test period. The proposed system comprises three agents controlling two battery-energy-storage units and one flexible-demand resource. Forecast uncertainty is represented through prediction intervals generated using quantile boosting and split-conformal calibration, while forecast errors are randomized during training to improve policy robustness. Learned policies are evaluated across three random seeds using weekly test episodes and compared with no-control, rule-based, and deterministic MAPPO benchmarks. The results show that UA-MAPPO reduced mean renewable-energy curtailment by 1.12% relative to deterministic MAPPO, with the paired Wilcoxon test indicating statistical significance at ($p = 0.0043$). However, UA-MAPPO did not significantly outperform the no-control benchmark, yielding a mean difference of −2,560 MWh per week with ($p = 0.4678$). The rule-based controller achieved the lowest curtailment, operating cost, imports, and emissions, although it increased peak imports by 33.17%. In addition, prediction-interval coverage during testing fell below the 90% target, indicating the presence of temporal distribution shift. Overall, the study therefore provides a transparent and reproducible benchmark, together with practical design insights for future research on intelligent renewable-energy management.

Keywords: Renewable Curtailment; Multi-Agent Reinforcement Learning; MAPPO; Conformal Prediction; Battery Storage; Demand Flexibility; Power-System Operation.

1. Introduction

Realising the full potential of rapidly expanding solar photovoltaic and wind energy requires the implementation of proactive and coordinated integration strategies. Between 2018 and 2023, the installed capacity of these technologies more than doubled, while their contribution to global electricity generation increased to nearly twice its previous level. Governments are increasingly positioning solar PV and wind as central pillars of energy-sector decarbonisation. Their deployment is expected to accelerate further toward 2030, supported by favourable policy frameworks, continuing technological advances, and substantial reductions in generation costs [1-3].

Delays in implementing measures to facilitate the integration of variable renewable energy could place up to 15% of projected solar photovoltaic and wind generation at risk by 2030 and reduce the anticipated decline in power-sector carbon dioxide emissions by as much as 20%. Under a pathway aligned with national climate targets, failure to deploy the necessary integration measures could jeopardise approximately 2,000 TWh of global variable renewable electricity generation by 2030, thereby undermining progress toward national energy and climate commitments. This volume is equivalent to the combined solar and wind output of China and the United States in 2023 [4-7].

The potential loss would arise primarily from increased technical and economic curtailment, grid-connection constraints, and delays in commissioning new renewable-energy projects. Consequently, the combined share of solar PV and wind in the global electricity mix could reach only 30% by 2030, compared with approximately 35% under a scenario in which integration measures are implemented in a timely manner. Should the resulting generation shortfall be compensated through greater reliance on fossil-fuel-based power generation, the expected reduction in power-sector CO_2 emissions could be diminished by up to one-fifth [8-10].

The rapid expansion of solar photovoltaic and wind power generation is transforming conventional electricity networks into low-carbon, decentralized, and increasingly dynamic energy systems. Although the growing penetration of variable renewable energy supports decarbonization and reduces dependence on fossil fuels, it also introduces substantial operational challenges associated with intermittency, limited predictability, network congestion, and mismatches between electricity generation and demand. Under periods of high renewable production and insufficient system flexibility, grid operators may be required to curtail available solar or wind generation to maintain network security and power balance. Renewable energy curtailment consequently represents a loss of clean electricity, reduces the economic returns of renewable projects, and limits the effective utilization of installed generation capacity [11-13].

Curtailment in solar–wind integrated power grids is influenced by several interacting factors, including transmission constraints, insufficient energy-storage capacity, limited ramping capability of conventional generators, forecasting errors, voltage and frequency restrictions, and inadequate coordination among distributed energy resources. The complementary characteristics of solar and wind generation can partially reduce aggregate variability; however, their combined operation remains highly uncertain because of rapidly changing weather conditions and spatially distributed generation patterns. Traditional optimization and rule-based energy-management methods generally rely on deterministic forecasts or simplified uncertainty models [14,15]. These approaches may become computationally demanding in large-scale systems and may not adapt effectively to real-time operating conditions, particularly when multiple autonomous entities must coordinate their actions across a constrained network.

Reinforcement learning has emerged as a promising data-driven framework for sequential decision-making in complex energy systems. By continuously interacting with the environment, a reinforcement-learning agent can learn operational policies that improve long-term system performance without requiring an exact mathematical representation of every system component. Nevertheless, centralized reinforcement-learning methods may encounter scalability, communication, and privacy limitations when applied to geographically distributed power grids. Multi-agent reinforcement learning addresses these limitations by assigning decision-making responsibilities to multiple cooperative agents representing renewable generators, battery energy-storage systems, flexible loads, substations, or regional grid controllers. Through coordinated learning, these agents can jointly manage power dispatch, storage charging and discharging, demand flexibility, and network power flows [16-18].

Despite its potential, the application of multi-agent reinforcement learning to renewable-energy curtailment remains challenging because many existing approaches do not explicitly account for uncertainty in renewable generation, load demand, electricity prices, equipment availability, and network operating conditions. Policies trained using deterministic or narrowly defined scenarios may perform poorly when exposed to unforeseen fluctuations or forecast errors. An uncertainty-aware learning framework is therefore required to ensure that agent decisions remain effective and reliable under diverse operating conditions. Such a framework may incorporate probabilistic forecasts, scenario-based training, risk-sensitive reward functions, uncertainty estimation, or robust policy-learning mechanisms to balance renewable-energy utilization against operational security. Moreover, Forecast uncertainty constitutes a fundamental challenge in the real-time operation of renewable-energy-integrated power systems [19-21]. Conventional point forecasts provide only a single expected value and therefore conceal the range, likelihood, and asymmetry of potential outcomes. Controllers relying exclusively on such estimates may adopt overly confident decisions, including excessive commitment of energy-storage resources or unnecessarily aggressive deferral of flexible demand. By contrast, prediction intervals provide explicit information regarding forecast dispersion and enable the controller to assess prospective operational risk before selecting an action.

Zheng et al. [22] proposed a deep learning–assisted distributionally robust optimization framework to manage uncertainties in renewable-integrated power systems. The approach combines probabilistic forecasting, adaptive ambiguity sets, and hierarchical multi-agent dispatch to improve cost efficiency, reliability, and renewable utilization. Simulation results showed an 11.0–13.5% reduction in operating costs, an increase in renewable utilization from 85.6% to 89.7%, improved reliability, and a 28.6% reduction in carbon emissions compared with the baseline.

According to [23], a Deep Q-Network-based optimization framework was developed for adaptive decision-making in dynamic power-system environments. Unlike conventional Mixed-Integer Linear Programming methods, the proposed approach continuously learns optimal control actions under renewable-generation uncertainty, load variability, and grid disturbances. Case-study results demonstrated improved real-time adaptability, reduced computational burden, and enhanced resilience, highlighting the potential of reinforcement learning for intelligent and cost-effective power-system operation.

In [24], a reinforcement learning–enhanced decision support system was proposed for the joint sizing and real-time operation of grid-connected photovoltaic–energy storage systems. The framework combines multi-objective optimization with a policy-gradient agent that dynamically adjusts the relative importance of cost, emissions, curtailment, and battery degradation according to changing system conditions. Two-year simulations showed reductions in energy costs of up to 35%, PV self-consumption above 70%, and curtailed energy below 8%, outperforming static and rule-based strategies.

In [25], a comprehensive review examined reinforcement learning–based control strategies for renewable-energy-integrated power grids, microgrids, and building energy systems. The study classified recent literature according to energy domain, RL algorithm, agent architecture, reward design, benchmark controller, integrated technology, and control objective. Its findings identified prevailing methodological trends and design patterns while highlighting the need for more scalable, reliable, and practically deployable RL-based energy-management solutions for next-generation smart grids.

This study evaluates an uncertainty-aware Multi-Agent Proximal Policy Optimization framework using publicly available German energy-system data. The proposed controller incorporates calibrated forecast-interval widths into the observation space of each agent, thereby allowing operational decisions to reflect both the expected trajectory and the associated level of uncertainty. In addition, forecast errors are systematically randomized during training to expose the agents to a broad range of plausible operating conditions and improve policy robustness. The resulting framework is referred to as uncertainty-aware MAPPO, or UA-MAPPO. The investigation addresses three principal research questions. First, does the explicit incorporation of forecast uncertainty improve the performance of conventional deterministic MAPPO? Second, can UA-MAPPO outperform transparent rule-based and heuristic benchmark strategies? Third, to what extent does the learned policy maintain stable performance under variations in renewable generation, demand, forecast accuracy, and system operating conditions? Addressing these questions requires rigorous and balanced comparisons rather than evaluation against a single, potentially weak baseline. The primary contribution of this work is a reproducible evaluation pipeline that integrates public datasets, chronological train–validation–test partitioning, probabilistic forecasting, uncertainty-informed multi-agent control, and multi-seed reinforcement-learning experiments. The analysis also reports unfavorable and statistically inconclusive outcomes rather than selectively emphasizing positive results. Such transparent evidence is essential for preventing overstated performance claims, identifying the practical limitations of uncertainty-aware reinforcement learning, and guiding the development of more reliable control architectures for renewable-energy-integrated power grids.

3. Materials and Methods

3.1 Public data and study period

The experiment used the Open Power System Data time-series package. The package compiles records from European system sources. These include ENTSO-E records (ENTSO-E, n.d.). The package supports transparent electricity modeling. The exact dataset version was released on October 6, 2020. The selected German series contain hourly load, solar generation, wind generation, and day-ahead prices. The period starts on January 1, 2015. It ends on June 30, 2020. All timestamps use Coordinated Universal Time. Measured load and generation anchor the counterfactual grid experiment. The study is not a physical field trial. Curtailment is calculated inside a constrained simulation. This distinction prevents an inflated claim about real grid deployment.

Table 1. Data sources, chronology, and preprocessing.

Item	Study setting	Rows or value
Measured series	German load, solar, wind, and day-ahead price	Hourly
Forecast and RL fit	1 Jan 2015–31 Dec 2018	35,015
Conformal calibration	1 Jan–31 Dec 2019	8,760
Held-out evaluation	1 Jan–30 Jun 2020	4,368
Missing before cleaning	Load / solar / wind / price	0.00% / 0.21% / 0.15% / 68.18%
Missing after cleaning	All four selected series	0.00%
Renewable scaling	Mean target penetration	65%; factor 2.485956
Base power	Maximum measured demand	70,923.6 MW

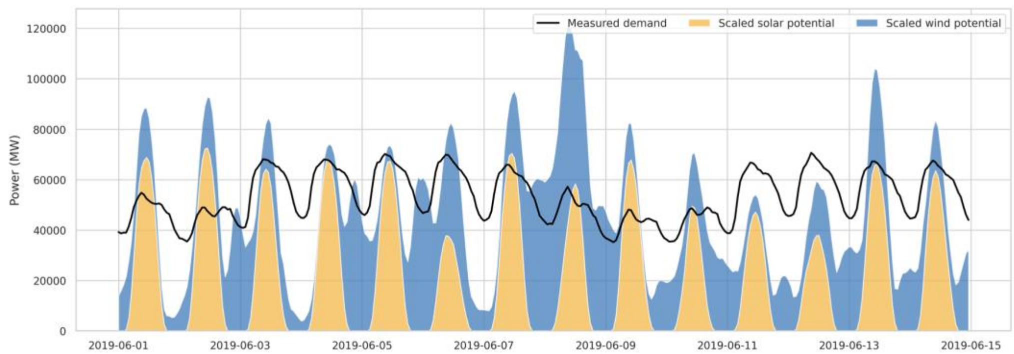

Figure 1. Measured German demand and capacity-scaled renewable potential.

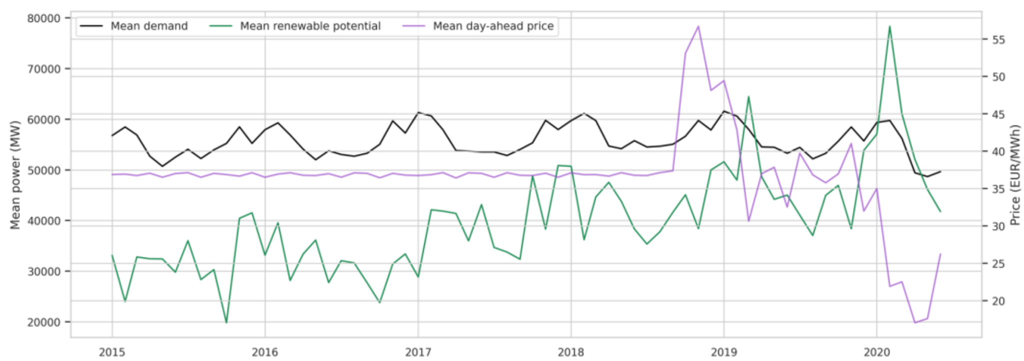

Figure 2. Monthly demand, renewable potential, and day-ahead price conditions.

3.2 Cleaning, scaling, and chronological separation

Negative load and generation values were clipped at zero. Small solar and wind gaps were interpolated. Price gaps received interpolation and weekday-hour median values. No selected variable remained missing afterward. Price completeness presents an important weakness. The raw price column was missing for 68.18% of all timestamps. Therefore, monetary findings receive less weight than energy findings. This limitation is discussed again later. The data were separated by time. Models used records through December 2018. Forecast intervals used 2019 for calibration. The first six months of 2020 remained untouched until testing. This order reduces information leakage. Renewable output was scaled to a 65% mean penetration target. The required factor was 2.485956. This step creates a future high-renewable case from measured profiles. It does not claim that Germany operated at that exact level.

$$renewable\ penetration\ = \frac{\Sigma(solar\ +\ wind)}{\Sigma(load)}$$

3.3 Probabilistic forecasting

Separate gradient-boosting models predicted load, solar, and wind. Each target used the 10th, 50th, and 90th quantiles. The models were fitted only on the training period.Features included one, two, 24, 48, and 168-hour lags. They also included rolling means and standard deviations. Cyclic hour, weekday, and yearly features represented calendar patterns. Lagged price was included as another predictor.
Each boosting model used a 0.06 learning rate and 180 iterations. The tree limit was 31 leaves. Minimum leaf size was 30. L2 regularization was 0.1. Predicted quantiles were sorted to remove crossing. Negative values were clipped at zero. Calibration then widened each interval using 2019 nonconformity scores. The 90th percentile score supplied the adjustment.

$$\hat{q} = Quantile0.90[\max(L\ -\ y, y\ -\ U, 0)]$$

The final interval was [max(0, L − q̂), U + q̂]. Load and wind needed positive adjustments. Solar needed no extra widening. Coverage, width, mean absolute error, and root mean squared error were recorded.

Table 2. Forecast accuracy and interval coverage.

Split	Target	MAE (p.u.)	RMSE (p.u.)	Coverage	Mean width
2019 validation	Load	0.0089	0.0121	90.01%	0.0429
2019 validation	Solar	0.0092	0.0184	91.27%	0.1811
2019 validation	Wind	0.0193	0.0284	90.01%	0.1032
2020 test	Load	0.0127	0.0171	82.69%	0.0494
2020 test	Solar	0.0148	0.0318	87.23%	0.2134
2020 test	Wind	0.0256	0.0408	85.65%	0.1289

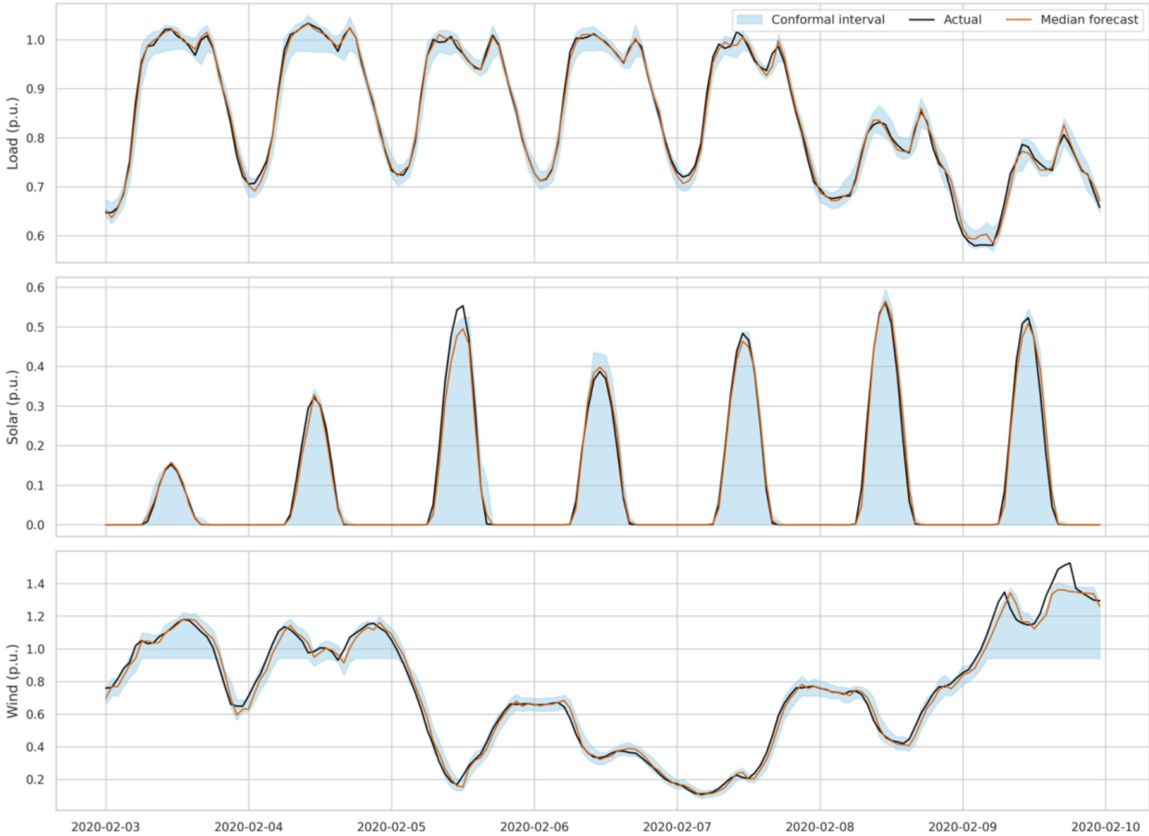

Figure 3. Held-out forecasts and conformal uncertainty intervals during one test week.

3.4 Grid environment and agents

The simulation used 168-hour episodes. Each episode therefore represented one week. Three agents shared the system reward. Two agents controlled separate solar and wind batteries. The third agent controlled flexible demand. Each battery stored 0.20 per-unit hours. Its power limit was 0.08 per unit. Charging and discharging efficiencies were both 95%. The two devices began each week at half charge. The export link was limited to 0.12 per unit. Flexible demand equaled 8% of current load. Deferred energy could reach 0.80 per-unit hours. All remaining debt was forced into the final hour. Each agent selected one of five actions. The normalized levels were −1, −0.5, 0, 0.5, and 1. Battery signs represented discharge and charge. Demand signs represented deferral and recovery.

UA-MAPPO observed median forecasts and three interval widths. Deterministic MAPPO received the same median forecasts. Its interval widths were set to zero. Both policies observed price, time, recent measurements, storage states, and demand debt. The common reward penalized curtailment, net cost, imports, throughput, and wasted discharge. Curtailment received weight 8. Import weight was 0.35. Battery throughput received 0.08. Wasted discharge received 12.

$$r_t = -(8C_t + K_t + 0.35I_t + 0.08B_t + 12W_t)$$

Net cost was normalized by the price scale. Imported power produced emissions at 0.35 tonnes per MWh. Deferred energy was settled exactly. This rule preserved energy neutrality across each week.

Table 3. Control and learning configuration

Component	Setting	Value
Episode	Weekly horizon	168 hours
Agents	Solar battery / wind battery / flexible demand	3
Actions	Discrete normalized levels	5 per agent
Observation	Common variables plus agent identity	20 values
Actor	Two hidden tanh layers	128 units each
Critic	Centralized two-layer tanh network	128 units each
Learning rates	Actor / critic	3×10^{-4} / 8×10^{-4}
PPO settings	γ / GAE λ / clip / entropy	0.99 / 0.95 / 0.20 / 0.015
Optimization	Epochs / batch / gradient clip	6 / 64 / 0.5
Training	Episodes per policy and seed	250
Seeds	Independent runs	7, 21, 42

3.5 MAPPO training and uncertainty treatment

Each agent used its own actor network. A centralized critic received all three observations. This structure allowed decentralized actions with shared value learning. Actor and critic networks used two 128-unit tanh layers. Moreover, PPO clipped each probability ratio between 0.8 and 1.2. The discount factor was 0.99. Generalized advantage estimation used 0.95. Entropy regularization encouraged early action exploration. UA-MAPPO faced randomized forecast errors during training. The stress multiplier was sampled uniformly between zero and one. The median forecast shifted using interval width. The deterministic policy received no such stress. Both policies trained for 250 weekly episodes per seed. Six PPO epochs followed each episode. Mini-batches contained 64 transitions. Gradient norms were clipped at 0.5. Training used PyTorch on a CUDA device.

Figure 4. MAPPO training rewards across three random seeds

3.6 Baselines, metrics, and statistical tests

No control supplied the first baseline. It used no batteries and no demand shifting. The rule-based controller formed a stronger baseline. It charged both batteries when forecast surplus exceeded 0.03 per unit. The rule discharged both batteries when deficit exceeded 0.03 per unit. It also deferred flexible demand. Near balance, it recovered deferred demand when possible. This rule used clear forecast thresholds. It required no learning. Primary performance was weekly curtailed energy. Secondary measures included curtailment rate, utilization, cost, imports, emissions, throughput, and peak import. Reported means gave each held-out week equal weight. Learned policies were also averaged across seeds.

$$curtailment\ rate\ = \frac{\Sigma(curtailed\ energy)}{\Sigma(available\ solar\ +\ wind)}$$

Bootstrap intervals used 5,000 resamples of weekly means. Paired Wilcoxon tests compared UA-MAPPO with each baseline. The paired unit was a held-out week. Two-sided p-values tested any difference.

4. Results

4.1 Forecast performance and temporal shift

Validation coverage closely matched the 90% target. Load and wind reached 90.01%. Solar reached 91.27%. This result confirms correct calibration within 2019. Coverage declined during the 2020 test. Load reached 82.69%, solar reached 87.23%, and wind reached 85.65%. Test errors also increased for every target. The pattern indicates a temporal distribution shift. Using the 70,923.6 MW base, test MAE was about 897 MW for load. Solar MAE was about 1,048 MW. Wind MAE was about 1,813 MW. Wind therefore remained the hardest target.

4.2 Training behavior

Training rewards improved unevenly and remained noisy. The last 50 episodes did not show uniform convergence. Deterministic seed means ranged from about −217 to −209. UA-MAPPO means ranged from about −216 to −202. Training curtailment rates were much lower than held-out rates. They generally ranged between 3.6% and 4.4%. Test rates were near 15%. This gap confirms that random training weeks did not represent 2020 fully.

4.3 Main held-out comparison

The rule-based controller achieved the lowest weekly curtailment. Its mean was 1.370 million MWh. No control curtailed 1.519 million MWh. Deterministic MAPPO curtailed 1.533 million MWh. UA-MAPPO curtailed 1.516 million MWh each week. This value was 1.12% below deterministic MAPPO. It was only 0.17% below no control. It remained 10.65% above the rule-based result. The rule also produced the lowest mean cost, imports, and emissions. However, its peak import reached 60,018 MW. That level was 33.17% above no control. The rule reduced energy totals by creating sharper recovery periods.

Table 4. Held-out weekly performance.

Policy	Curtailment (MWh)	Rate	Net cost (€)	Imports (MWh)	Emissions (t)	Peak import (MW)
No control	1,518,848	14.84%	51,632,366	1,807,762	632,717	45,067
Rule based	1,370,381	13.11%	46,538,284	1,630,438	570,653	60,018
Deterministic MAPPO	1,533,482	14.99%	51,468,348	1,826,105	639,137	65,888
UA-MAPPO	1,516,288	14.78%	51,400,296	1,808,043	632,815	62,016

Values are equal-weight weekly means. Learned policies are also averaged across seeds 7, 21, and 42.

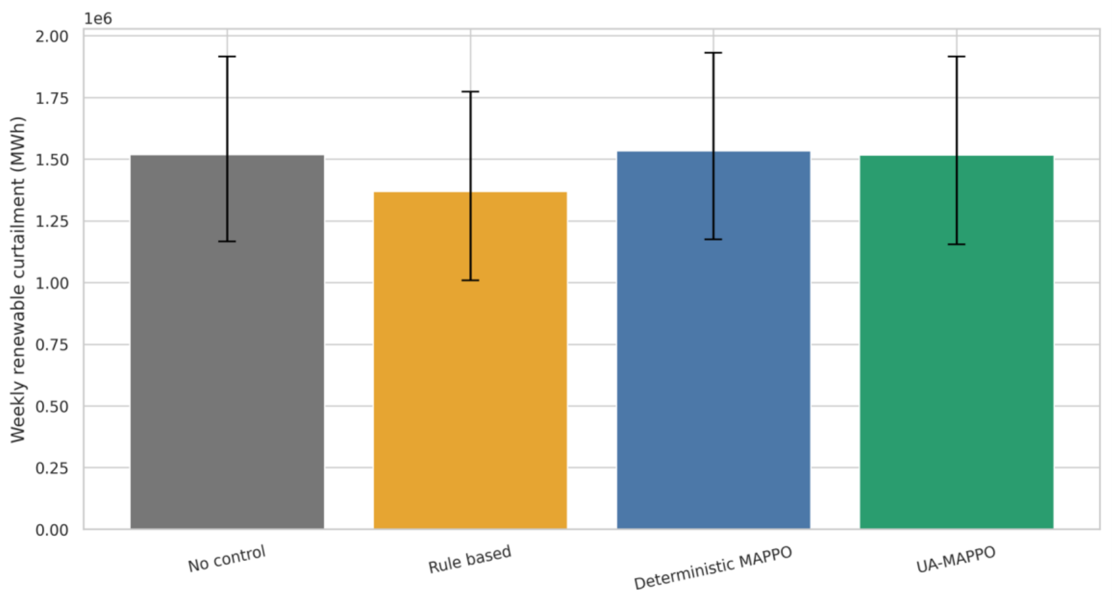

Figure 2. Held-out curtailment with bootstrap confidence intervals

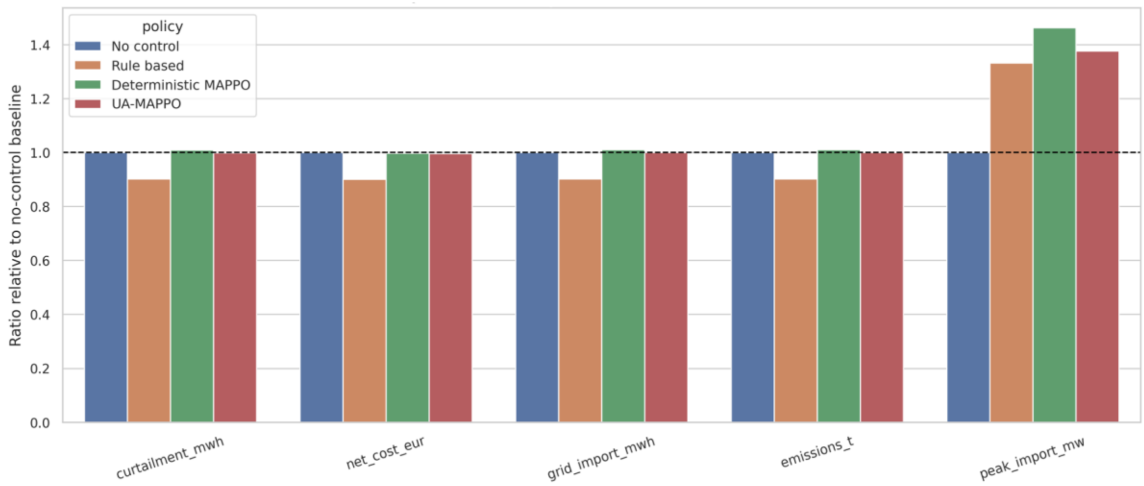

Figure 6. Normalized performance across curtailment, cost, imports, emissions, and peak import.

4.4 Statistical evidence

UA-MAPPO significantly improved deterministic MAPPO. The mean paired difference was −17,194 MWh per week. Its 95% bootstrap interval excluded zero. The Wilcoxon p-value was .0043. The difference against no control was not significant. The mean was −2,560 MWh per week. Its interval ranged from −13,181 to 9,248 MWh. The p-value was .4678. UA-MAPPO performed significantly worse than the rule. Its mean disadvantage was 145,907 MWh per week. The interval remained fully positive. The two-sided p-value was below .000001.

Table 2. Paired weekly curtailment tests for UA-MAPPO.

Comparator	Mean difference (MWh)	95% bootstrap interval	Wilcoxon p	Interpretation
No control	−2,560	−13,181 to 9,248	.4678	No significant difference
Rule based	+145,907	119,334 to 173,299	< .000001	UA-MAPPO is worse
Deterministic MAPPO	−17,194	−26,480 to −7,627	.0043	UA-MAPPO is better

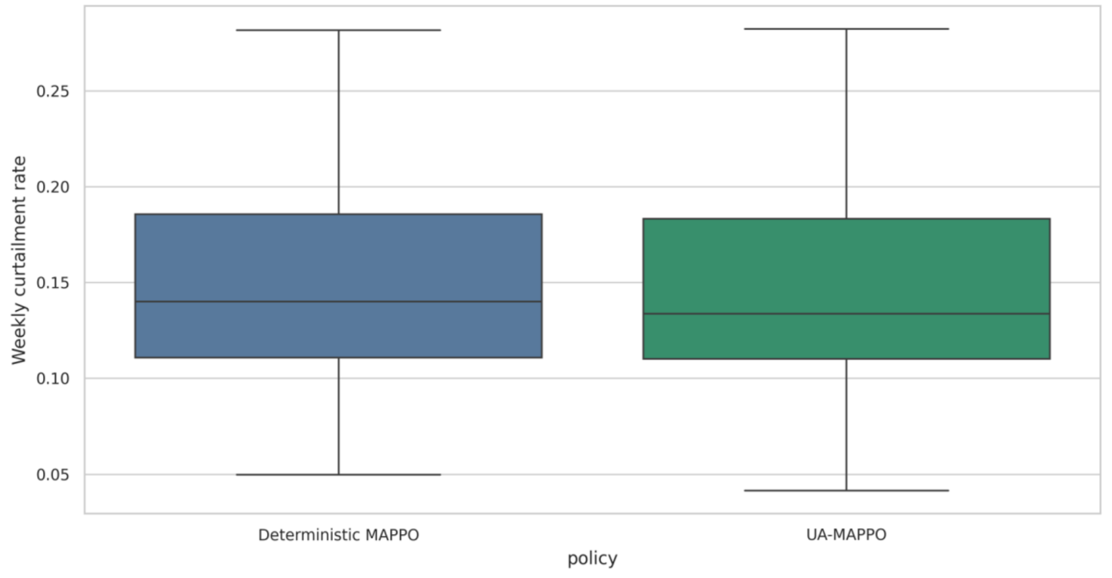

Figure 3. Weekly uncertainty ablation across held-out episodes.

4.5 Storage and flexible-demand behavior

The wind battery often moved between empty and full states. The solar battery stayed near empty during the displayed week. This contrast reflects resource timing and learned action choices. It may also indicate weak coordination. Deferred demand often reached its maximum level. Recovery occurred only during selected hours. Any remaining debt was imposed in the final hour. This settlement created a sharp end-of-week demand increase. The settlement rule ensures energy neutrality. It can still distort peak imports. A smoother rolling debt constraint may be better for future studies. Peak measures should therefore be read with care.

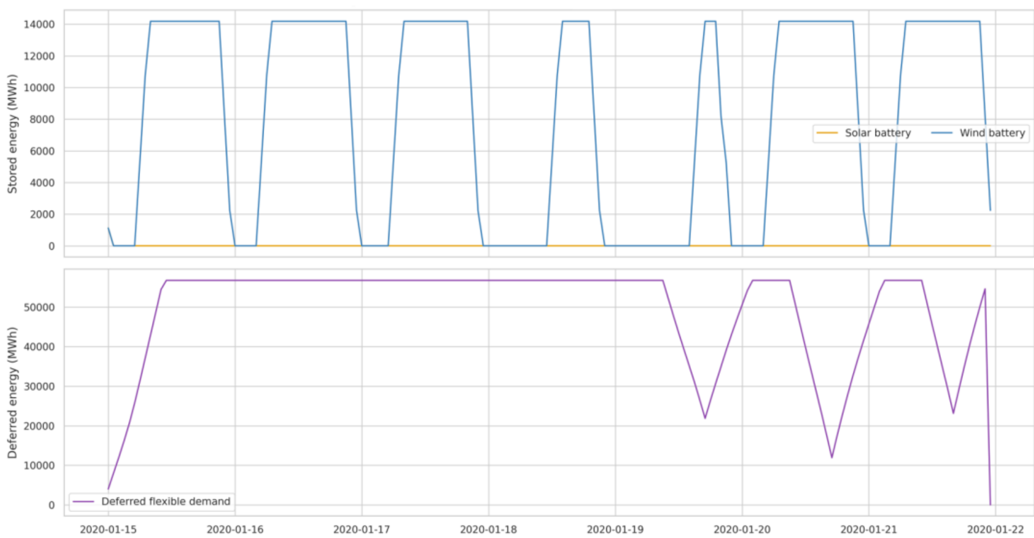

Figure 4. Storage states and energy-neutral flexible demand in a held-out week.

4.6 Sensitivity to penetration and storage

Curtailment increased strongly with renewable penetration. UA-MAPPO averaged 5.03% at 50% penetration. It reached 14.60% at 65%. The rate rose to 24.17% at 80% and 32.36% at 95%. Greater storage did not improve the learned policy in this test. Its curtailment rate rose slightly at each penetration level. The policy was not retrained after capacity changed. Therefore, this result shows out-of-distribution deployment. The finding does not prove that added storage increases curtailment. It shows that a fixed policy may misuse changed hardware. Capacity planning should include policy retraining. Joint optimization may also reduce this mismatch.

Table 3. Sensitivity at the original storage capacity

Mean renewable penetration	No control rate	Rule-based rate	UA-MAPPO rate
50%	5.13%	4.08%	5.03%
65%	14.84%	13.11%	14.60%
80%	24.40%	22.77%	24.17%
95%	32.58%	31.20%	32.36%

The learned controller used the seed-7 model in this sensitivity experiment.

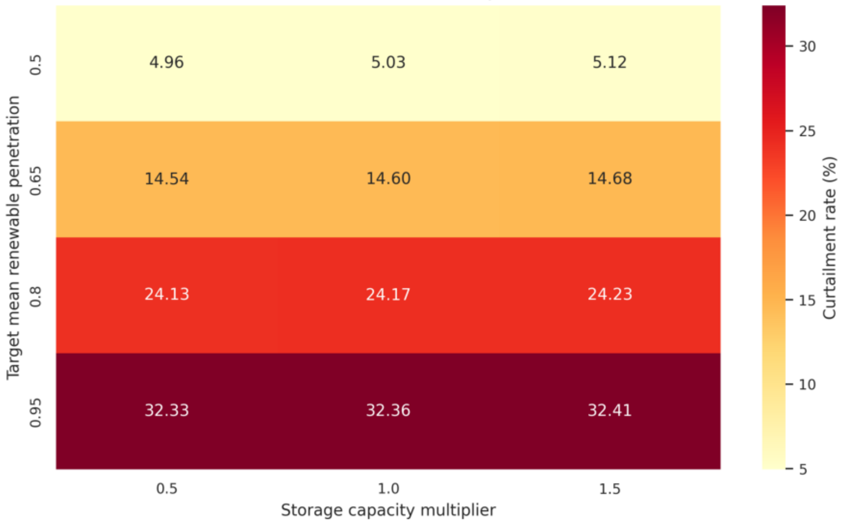

Figure 5. UA-MAPPO sensitivity to renewable penetration and storage capacity

4.7 Robustness under forecast stress

The robustness test increased random median-forecast shifts. UA-MAPPO remained slightly below deterministic MAPPO across all stress levels. Both curves were nearly flat. Confidence intervals were broad and strongly overlapping.This test used only seed 7. It therefore provides supportive, not decisive, evidence. The main three-seed paired comparison remains stronger. Future robustness testing should repeat every seed and stress condition.

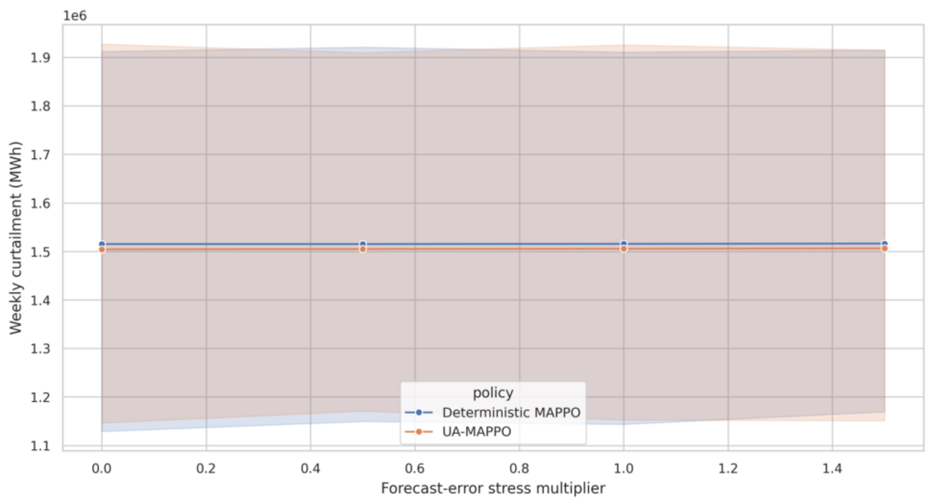

Figure 6. Curtailment under increasing forecast-error stress.

5. Discussion

5.1 Meaning of the main result

Uncertainty awareness improved the learned controller. The improvement over deterministic MAPPO was statistically significant. Forecast widths and randomized errors therefore added useful information. This result supports the proposed contribution. The gain was still small in operational terms. UA-MAPPO did not significantly beat no control. It also lost clearly to the rule-based controller. A publication should state all three findings together. This pattern is scientifically useful. It separates improvement within an algorithm family from practical superiority. Many reinforcement learning studies report only the first comparison. Operators need the second comparison before considering deployment.

5.2 Why the simple rule performed well

The rule directly targeted forecast surplus and deficit. Its threshold matched the main physical imbalance. It also charged both batteries together. That direct structure can be very effective in a compact environment. MAPPO learned from a mixed reward. Curtailment competed with cost, imports, throughput, and wasted discharge. The chosen weights may not reflect the desired priority perfectly. Reward shaping can also create hard learning signals. The action space was coarse. Each agent had only five levels. A battery could not select a finely tuned power value. Continuous actions may reduce wasted charging and forced recovery. They could also raise training difficulty. The rule was not universally better. Its peak import was one third above no control. UA-MAPPO also showed a high peak. These peaks reveal a tradeoff between weekly energy and hourly capacity. A constrained peak limit could improve both policies.

5.3 Forecast shift and policy generalization

- Conformal calibration worked during 2019.

It weakened during 2020. Standard split conformal methods assume stable error behavior. Temporal change can break that condition. The observed coverage loss confirms this risk.

- Training and test curtailment also differed sharply.

The controller saw random weeks through 2018. The held-out 2020 profile contained different renewable and demand patterns. Domain randomization softened this shift but did not remove it.

- Rolling calibration may improve interval reliability.

Weighted conformal scores could emphasize recent errors. Online policy fine-tuning may also help. Such updates must protect the test period during evaluation.

5.4 Practical implications

A grid operator should not select UA-MAPPO from these results alone. The controller needs stronger peak limits and wider scenario tests. It also needs comparison with optimization methods. Safety constraints should remain explicit.The experiment still provides a useful development path. Prediction intervals should enter the control state. Training should expose policies to realistic forecast errors. Evaluation should retain chronological data and simple baselines. Demand flexibility deserves careful modeling. Deferred load is not free energy. Recovery timing affects peaks, costs, and customer comfort. The International Energy Agency also stresses the need for usable flexibility arrangements. Storage capacity and policy design should be studied together. The sensitivity result showed weak transfer across capacities. A controller trained for one device may not suit another. Retraining or conditioning on hardware limits can address this issue.

5.5 Limitations

First, this work uses a simplified single-area grid. It excludes transmission topology, unit commitment, reserves, and voltage constraints. Curtailment is therefore counterfactual. It is not an observed German curtailment record. Second, the price series had extensive missing values. Completed prices supported an internally consistent cost measure. They may not reproduce market history accurately. Cost rankings should remain secondary to energy rankings. Third, only three learning seeds were available. Three seeds are better than one, yet still limited. More seeds would narrow uncertainty. They would also support robust aggregate performance measures. Fourth, all agents shared one reward. This choice promotes cooperation but hides individual credit. Credit-assignment methods may improve learning. Agent-specific constraints could also prevent extreme actions. Fifth, robustness testing used one seed. Sensitivity testing also reused a fixed trained policy. Those analyses show behavior under change, not retrained optimum performance. Their conclusions should remain narrow. Finally, the weekly debt settlement produced an artificial terminal spike. It guaranteed energy neutrality. A rolling deadline would represent flexible loads more realistically. Future work should compare several settlement designs.

6. Conclusion

This study evaluated uncertainty-aware MAPPO for renewable curtailment control. The evidence came from public German hourly data. Forecasting, calibration, training, and testing followed a strict chronological order. UA-MAPPO reduced weekly curtailment by 1.12% against deterministic MAPPO. The paired difference was significant. This result confirms value from interval information and forecast-error randomization. The wider conclusion is more cautious. UA-MAPPO did not significantly beat no control. A transparent rule reduced curtailment by 9.78% against no control. It also increased peak imports. The study therefore rejects a simple superiority claim. Uncertainty-aware learning improved MAPPO, but not the best practical baseline. Future research should add network constraints, adaptive calibration, continuous actions, and safer demand recovery. The reproducible benchmark can support that next step. Its strongest lesson concerns evaluation. Learned control should face real data, temporal shift, simple rules, multiple seeds, and paired statistical tests.

Author Contributions: Authors have contributed significantly to the development and completion of this article.
Funding: This article received no external funding.

Data Availability
The public data are available from Open Power System Data (2020). The versioned dataset DOI is https://doi.org/10.25832/time_series/2020-10-06.

Code and Result Availability
The executed notebook is named Renewable_Curtailment_UA_MAPPO_Colab.ipynb. The results archive is named renewable_curtailment_results_bundle.zip. The archive contains exact tables, figures, trained weights, metadata, and integrity checks.

Acknowledgments: The authors thank the mechanical and industrial engineering department, Bani Waleed University, Libya for providing laboratory facilities and equipment.
Ethics approval: The study used public system-level data and involved no human participants.
Conflicts of Interest: The author(s) declare no conflict of interest.

References ha

[1] M. Swadi, D. Jasim Kadhim, M. Salem, F. Mohammed Tuaimah, A. S. Majeed, and A. Jawad Alrubaie, “Investigating and predicting the role of photovoltaic, wind, and hydrogen energies in sustainable global energy evolution,” *Global Energy Interconnection*, vol. 7, no. 4, pp. 429–445, 2024.

[2] L. Sander, C. Jung, and D. Schindler, “New concept of renewable energy priority zones for efficient onshore wind and solar expansion,” *Energy Convers. Manag.*, vol. 294, no. 117575, p. 117575, 2023.

[3] O. S. Abualoyoun and A. A. Amar, “Advancing the global integration of solar and wind power: Current status and challenges,” *ijees*, vol. 3, no. 2, pp. 73–88, 2025.

[4] P. Tafarte, M. Eichhorn, and D. Thrän, “Capacity expansion pathways for a wind and solar based power supply and the impact of advanced technology—A case study for Germany,” *Energies*, vol. 12, no. 2, p. 324, 2019.

[5] X. Costoya, M. deCastro, D. Carvalho, B. Arguilé-Pérez, and M. Gómez-Gesteira, “Combining offshore wind and solar photovoltaic energy to stabilize energy supply under climate change scenarios: A case study on the western Iberian Peninsula,” *Renew. Sustain. Energy Rev.*, vol. 157, no. 112037, p. 112037, 2022.

[6] M. Khaleel *et al.*, “Emerging issues and challenges in integrating of solar and wind,” *ijees*, vol. 2, no. 4, pp. 1–11, 2024.

[7] H. Jiang *et al.*, “Globally interconnected solar-wind system addresses future electricity demands,” *Nat. Commun.*, vol. 16, no. 1, p. 4523, 2025.

[8] M. Khaleel, Z. Yusupov, O. M. Bshina, M. M. Dali, and W. A. Shah, “Toward renewable energy deployment in North African countries: Potential resources, investments, transition scenario, opportunities, challenges, and policy,” *Unconventional Resources*, vol. 12, no. 100363, p. 100363, 2026.

[9] O. Bamisile *et al.*, “What is the cost and potential of low carbon electricity transition in the MENA region?,” *Energy Strat. Rev.*, vol. 59, no. 101762, p. 101762, 2025.

[10] M. Khaleel, Z. Yusupov, and S. Rekik, “Advancing hydrogen as a key driver for decarbonized power systems,” *Unconventional Resources*, vol. 9, no. 100278, p. 100278, 2026.

[11] Y. Yang *et al.*, “The low-carbon economic scheduling method for regional integrated energy systems considering the joint integration of electric vehicles and concentrated solar power plants for wind power consumption,” *Wind Eng.*, 2024.

[12] U. A. Dodo *et al.*, “Grid-connected hydro–solar hybrid systems for low-carbon energy transition: a techno-economic assessment,” *Environ. Dev. Sustain.*, 2026.

[13] Y. Tian, J. Wu, Z. Xu, and X. Guan, “Integrated modeling and dynamic evaluation indicator system of urban energy systems towards a sustainable, low-carbon pathway,” *Energy (Oxf.)*, vol. 319, no. 134993, p. 134993, 2025.

[14] M. Lafuente-Cacho *et al.*, “State of the art for solar and wind energy-forecasting methods for sustainable grid integration,” *Curr. Sustain./Renew. Energy Rep.*, vol. 12, no. 1, 2025.

[15] A. Hadavi, M. T. Hagh, and S. G. Zadeh, “Critical inertia thresholds for frequency stability in renewable Energy-Integrated power systems,” *Int. J. Electr. Power Energy Syst.*, vol. 169, no. 110733, p. 110733, 2025.

[16] Y. Lin, J. Tang, J. Guo, S. Wu, and Z. Li, “Advancing AI-enabled techniques in energy system modeling: A review of data-driven, mechanism-driven, and hybrid modeling approaches,” *Energies*, vol. 18, no. 4, p. 845, 2025.

[17] L. Zhang, Y. He, H. Wu, and N. D. Hatziargyriou, “An optimal scheduling framework for integrated energy systems using deep reinforcement learning and deep learning prediction models,” *IEEE Trans. Smart Grid*, vol. 16, no. 6, pp. 4620–4634, 2025.

[18] P. Michailidis, I. Michailidis, and E. Kosmatopoulos, “Reinforcement learning for optimizing renewable energy utilization in buildings: A review on applications and innovations,” *Energies*, vol. 18, no. 7, p. 1724, 2025.

[19] R. Shan, X. Jia, X. Su, Q. Xu, H. Ning, and J. Zhang, “AI-driven multi-objective optimization and decision-making for urban building energy retrofit: Advances, challenges, and systematic review,” *Appl. Sci. (Basel)*, vol. 15, no. 16, p. 8944, 2025.

[20] M. Guder, N. Balta-Ozkan, D. Huo, and P. Mirzania, “Data-driven decision making: Harnessing energy data for policy,” in *Simulation and Machine Learning Models for Energy Policy Design*, Elsevier, 2026, pp. 41–73.

[21] G. Dudek and M. Blachnik, "Artificial intelligence and machine learning models for forecasting, optimization, and control in smart energy systems," *Energies*, vol. 19, no. 3, p. 768, 2026.

[22] Y. Zheng *et al.*, "Multi-agent coordination and uncertainty adaptation in deep learning-assisted hierarchical optimization for renewable-dominated distribution networks," *Sci. Rep.*, vol. 16, no. 1, p. 5176, 2026.

[23] D. Li, H. Zheng, and T. Pan, "Resilience and flexibility optimization in solar integrated power systems via deep Q network under extreme weather," *Sci. Rep.*, vol. 15, no. 1, p. 37374, 2025.

[24] A. C. Zournatzis, E. P. Nicolopoulou, V. T. Kontargyri, and C. A. Christodoulou, "An adaptive multi-objective decision system with reinforcement learning for PV-ESS sizing and operation for grid-connected industrial PV-ESS systems in distribution networks," *IEEE Access*, vol. 13, pp. 148547–148558, 2025.

[25] P. Michailidis, F. Minelli, H. H. Coban, I. Michailidis, and E. Kosmatopoulos, "Reinforcement learning for optimizing renewable energy utilization in smart grids: Recent advances in power grids, microgrids, and building energy systems," *Infrastructures*, vol. 11, no. 7, p. 240, 2026.