Reinforcement Learning Approaches for Load Forecasting in Microgrids: A Comprehensive Review

Misoon Rashid Ali Almarzouqi[1], Reem Ali Mohammed Al-Rushdi[2]
[1,2]Department of Information Technology, University of Technology and Applied Sciences
Shinas, Oman

Article Info

Article History:

Received Oct 11, 2025
Revised Nov 07, 2025
Accepted Dec 06, 2025

Keywords:

Reinforcement Learning (RL)
Load Forecasting
Microgrids
Smart GridDeep Q-Network (DQN)
Proximal Policy Optimization (PPO)
Time-Series Prediction

ABSTRACT

Increased use of renewables makes good load forecasting crucial to the more efficient running of microgrids, and to the proper management of their energy. The conventional methods of prediction are not typically capable of dealing with the highly non-linear, stochastic and time-varying dynamics commonly observed in modern microgrid systems. During the last several years, reinforcement learning (RL) has started to be preferred due to its capabilities of assisting systems to progress independently and follow the best learning patterns depending on their real-world experiences. In this review, a large number of RL-based load forecasting approaches applied in microgrid environments are considered. The method is used to arrange past research according to forecasting horizon, the kind of algorithms employed, character of the information and judgment basis of outcome. Their efficiency and drawbacks in terms of real-time forecasting assignments are compared in terms of Q-learning, Deep Q-Networks (DQN), Proximal Policy Optimization (PPO) and Actor-Critic. Hybrid models, computation problems and challenges of merging IoT and edge computing layouts are also examined by the authors. It talks about the fields in which the recent research has been lacking and outlines how to proceed, naming federated learning, multi-agent reinforcement learning and the standardization of datasets as requirements. This work aims at demonstrating to the research and developer communities how they can deploy solid RL techniques to achieve smart, scalable and reliable microgrid load forecasting.

Corresponding Author:

Misoon Rashid Ali Almarzouqi,
Department of Information Technology,
University of Technology and Applied Sciences, Shinas, Oman.

1. INTRODUCTION

The integration of additional renewables, electric vehicles and DER is transforming the way power systems are operated, primarily within the microgrid environments. Accurate and timely demand forecasting is the main ingredient in stable supply, optimal utilization of system resources and demand management cost-efficiency. Over the last several years, statistical methods, such as ARIMA, exponential smoothing and multiple linear regression have been relied upon to perform power system load forecasting [1]. Such models can perform adequately under simple, regular conditions, but they tend to fail at capturing the complex, uncertain and stochastic behaviour of energy demand in microgrids when weather-dependent renewable generation, fluctuating demand and user behaviour are in play.

With the implementation of machine learning tools beginning in the last ten years, specifically the use of models like SVR [2], ANNs and LSTM networks, predictions are now more accurate due to the data collected through information. These methods, nonetheless, require an unchanging environment or load to perform best and will have to be retrained in the case of changes. Hence, scientists have been experimenting with techniques that are adaptive and self-learning and reinforcement learning (RL) has been shining in this territory.

Other systems are highly dynamic and reinforcement learning lets agents adapt to such dynamic systems through feedback in their course of action. In addition to the fact that labeled data is not needed, RL is a perfect choice to predict the loads in microgrids as it is capable of assuming fast, challenging decisions when data is uncertain [3]. Recent studies demonstrated that Q-learning, Deep Q-Networks (DQN) and Proximal Policy Optimization (PPO) could assist in making the forecasts more flexible, minimize prediction error and enable energy-efficient scheduling in smart grids.

Nevertheless, certain issues exist. A number of RL-based models are incapable of dealing with different microgrid configurations and struggle to do so because they lack data, are unstable to converge and have high computational requirements. Consequently, researchers need to develop architectures that cluster RL with edge computing, federated learning and multi-agent systems to offer secure and scalable prediction options.

In this paper, types of reinforcement learning have been discussed in anticipation of microgrid loads to bridge gaps in knowledge. We meticulously study accessible algorithms, their system configurations, the mode of operation and the ease of deployment. It also identifies the major research gaps and recommends how to develop robust, intelligent and short-term prediction techniques of future microgrids. Load forecasting evolution in microgrids is illustrated in Figure 1.

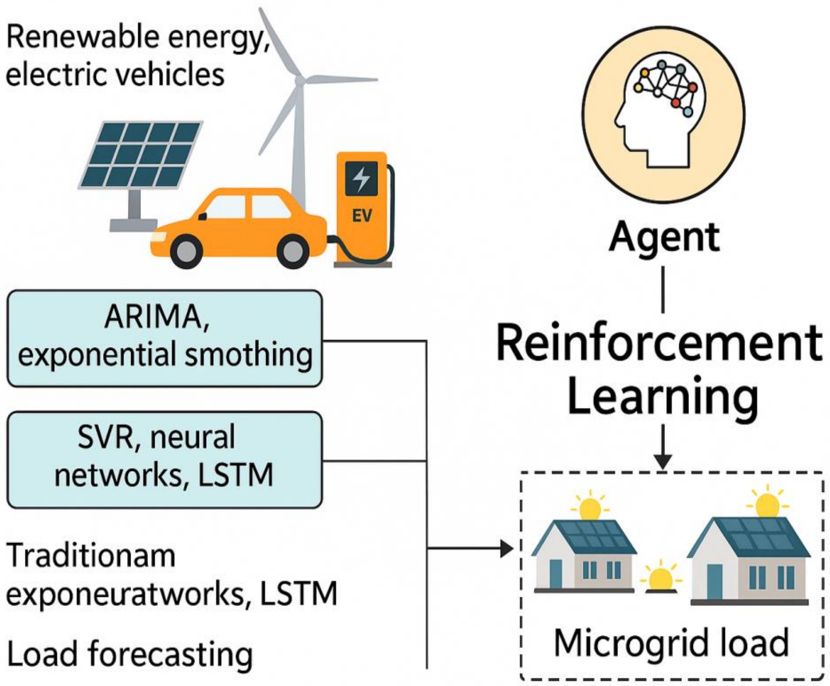

Figure 1. Load Forecasting Evolution in Microgrids: Traditional to RL Approaches

2. BACKGROUND

2.1 Microgrid Structure and Load Forecasting Needs

They are small systems that integrate such things as DERs, energy storage and controllable equipment to serve a local area [4]. They can operate grid-tied or off-grid that provides them with flexibility, good performance and capability of producing their own energy. Major components of a microgrid include a solar or wind power machine, a diesel or gas-powered generator, battery pack, smart meter and an EMS to direct the interaction between them.

One primary reason why a microgrid is important is that it can maintain constant balance between the produced energy and the consumed energy. This is why power demand forecasting is required since it determines how power is distributed [5], how batteries are charged and discharged, and how customers react to the energy market change. Due to the potential volatility of renewable energy and unpredictable needs of the consumers in microgrids, the demand of the power becomes more significant and challenging to forecast.

2.2 Classification of Forecasting Horizons

Short-term, intermediate-term and long-term forecasting are referred to as load forecasting and each type has its purpose in terms of value to operations and planning.

STLF is a technique of extrapolating electrical load over several minutes to hours. Primarily it is used to make run decisions such as which generators to turn on, how to instruct energy usage and when to charge or discharge batteries. In MTLF, the requirement of swift and precise data is quite essential. In this level, its primary applications are in the scheduling of maintenance, bidding of electricity and fuel purchase timing. In the case of LTLF, we consider the timeframe of months to years because it is applied in the development of new infrastructure, establishment of laws and expansion of supply. So trends and scenarios are significant to STLF as it is highly involved in microgrid systems that direct daily operations and are a main factor in grid stability [6].

2.3 Key Performance Metrics in Load Forecasting

The usefulness of the forecasting models depends on checking the accuracy of the predictions they give. The literature contains many articles that employ various measures of performance to determine the accuracy of forecasts.

Mean Absolute Percentage Error (MAPE):

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \times 100 \qquad (1)$$

It expresses the average error as a percentage of actual values and is widely used due to its interpretability. However, it can be biased when actual values are near zero.

Root Mean Square Error (RMSE):

RMSE penalizes large errors more than smaller ones and is suitable when larger forecasting errors are particularly undesirable.

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(A_t - F_t)^2} \qquad (2)$$

Mean Absolute Error (MAE):

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|A_t - F_t| \qquad (3)$$

MAE provides an absolute measure of average forecast error and is less sensitive to outliers than RMSE.

R-squared (R^2): quantifies the ratio of the variance in the observed load that is predicted by the model. It is commonly employed together with measures of error in assessing the overall goodness-of-fit.

In the case of reinforcement learning based load forecasting these are the metrics either used to train objectively or to measure following training [7]. The optimal forecasting model to be utilized together with microgrids must not alone be precise in its forecasts but also respond rapidly to rapid changes and support critical operational decisions. Figure 2 represents the reinforcement learning-enabled load forecasting and control framework in a microgrid.

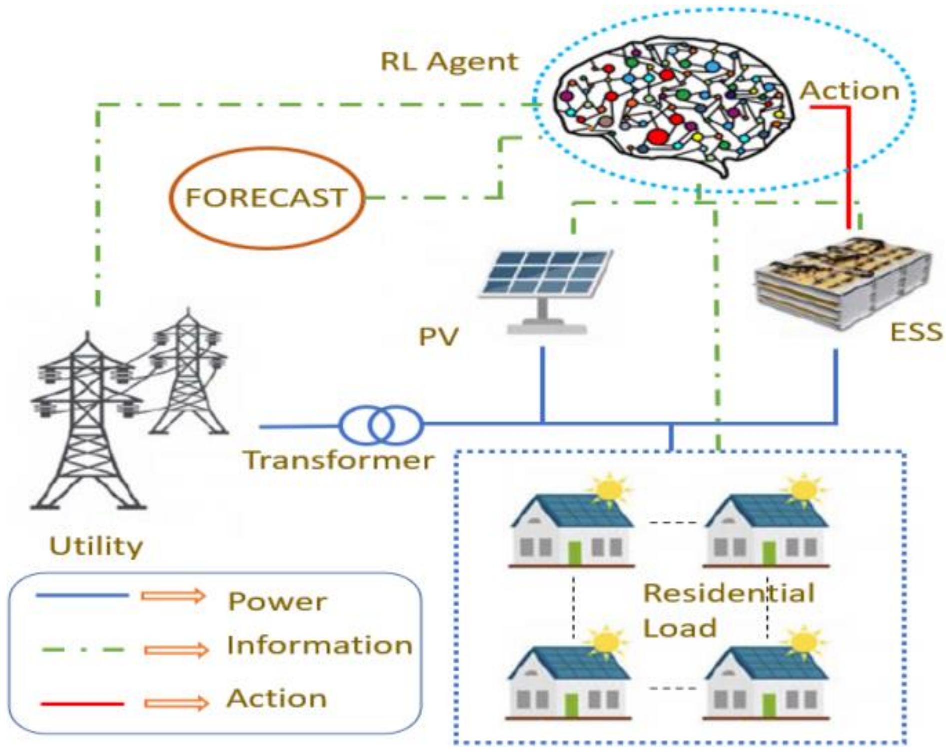

Figure 2. Reinforcement Learning-Enabled Load Forecasting and Control Framework in a Microgrid

2.4 Reinforcement Learning: Concepts and Algorithms

Such a learning approach is called Reinforcement Learning (RL) [8], which assists an independent system to make recurrent choices in a setting based on the rewards it receives. Reward Learning consists of four components: the agent, the environment, a reward signal and a policy. The agent constructs strategies such that the agent achieves the most benefit in the long run, most often by considering an MDP where the future state is only dependent on the present state and the action of the agent. RL can manage energy systems and foretell loads, unlike supervised learning that utilizes fixed data and needs to be re-trained each time the environment adjusts; RL can use its findings about the world.

Algorithms Reinforcement learning algorithms are currently finding application in energy demand prediction and the smart grid domain [9]. The selection of Q-Learning is very popular since the best actions are easy to learn. DQN Deep neural networks are utilized in DQN to assist the Q-learning to dealing with high-dimensional spaces, which is common in multivariate load prediction. Addressing the issue of overestimation in Q-learning, Double DQN furnishes the process of action selection and evaluation as separate.

DDPG is a favorite choice since deterministic policy learns in continuous action spaces, which are prevalent in reinforcement learning with actor-critic architecture. PPO aims at encouragement of stability as well as great performance, achieved through clipping of the probability ratios during leaning. Actor-Critic models with A2C and A3C maintain the policy and value function in sync with each other and are able to adaptation in real-time scenarios [10]. Whereas other methods, such as SVMs or LSTMs, are fixed, RL algorithms can enhance themselves during interaction, which is why they are well suited to be used in microgrids which continuously vary. Figure 3 shows the independent illustrations of supervised learning and reinforcement learning paradigms.

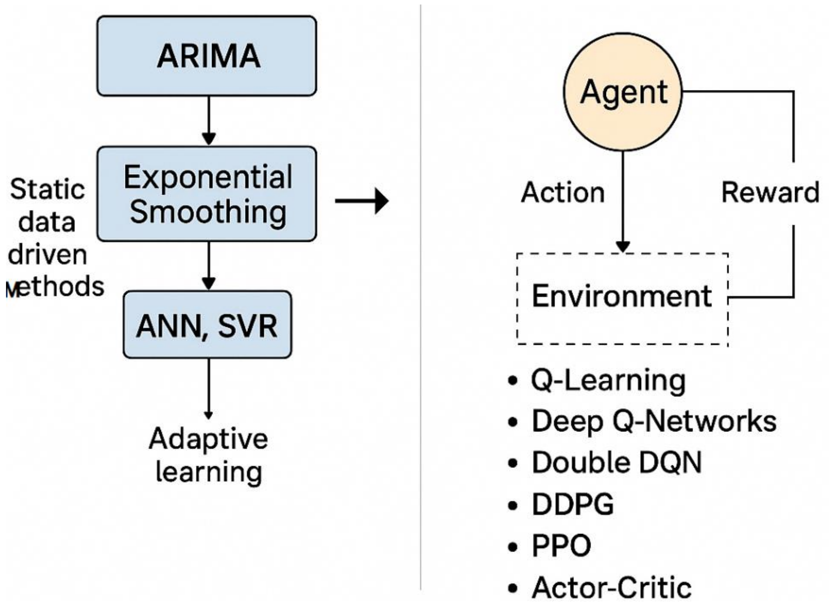

Figure 3. Independent Illustrations of Supervised Learning and Reinforcement Learning Paradigms

Benefits of Applications and Algorithms for Reinforcement Learning

- **Capacity to Manage Nonlinear and Complex Systems**

Reinforcement learning algorithms are well suited for environments characterized by nonlinear, stochastic, and time-varying dynamics. Unlike traditional model-based or supervised learning approaches, RL does not require explicit system modeling, making it effective for complex real-world applications such as microgrids, robotics, and autonomous systems.

- **Model-Free Learning Capability**

A major advantage of RL is its model-free nature, where optimal policies are learned directly through interaction with the environment. This is particularly beneficial in systems where accurate mathematical models are difficult [11] or expensive to obtain, such as energy systems with renewable integration or dynamic consumer behavior.

- **Flexibility in Changing and Unpredictable Situations**

RL algorithms continuously learn and adapt based on real-time feedback [12]. This enables them to respond effectively to changes in system conditions, uncertainties, and disturbances, making RL suitable for applications involving fluctuating loads, renewable energy variability, and evolving operational constraints.

- **Sequential Decision-Making and Long-Term Optimization**

Unlike traditional optimization methods that focus on immediate outcomes, RL optimizes long-term cumulative rewards. This characteristic makes RL ideal for sequential decision-making problems, such as energy management, load forecasting, traffic control, and resource allocation.

- **Deep Reinforcement Learning's Scalability**

The integration of deep neural networks with reinforcement learning (Deep RL) enables scalability to high-dimensional state and action spaces. Techniques such as Deep Q-Networks (DQN), Proximal Policy Optimization (PPO) [13], and Actor–Critic methods allow RL to handle complex sensory inputs and large-scale systems.

- **Decreased Reliance on Labeled Information**

RL does not require labeled datasets, unlike supervised learning methods. Instead, learning is guided by reward signals, reducing data labeling costs and enabling continuous online learning in real-world environments.

- **Capability to Make Decisions in Real Time**

Many RL algorithms can operate in near real time once trained, enabling fast and autonomous decision-making. This is critical for applications such as smart grids, autonomous vehicles, robotics, and industrial automation.

- **Adaptability in a Variety of Applications**

Reinforcement learning has been successfully applied across a wide range of domains, including smart grids, microgrids, robotics, healthcare, finance, transportation, and gaming. Its general-purpose learning framework allows easy adaptation to different problem settings.

- **Combining Emerging Technologies**

RL can be effectively combined with emerging paradigms such as Internet of Things (IoT), edge computing, digital twins, and cyber-physical systems. This integration enhances intelligent decision-making in distributed and resource-constrained environments.

- **Assistance with Distributed and Multi-Agent Systems**

Multi-agent reinforcement learning (MARL) enables coordination and cooperation among multiple agents operating in decentralized systems. This is particularly beneficial for applications such as distributed energy management, swarm robotics, and networked control systems.

- **Sturdiness in the Face of Uncertainty and Noise**

RL algorithms learn optimal behaviors despite noisy measurements, incomplete information, and uncertain environments. This robustness improves reliability in real-world deployments.

- **Constant Improvement with Experience**

RL systems improve performance over time as more interaction data become available. This lifelong learning capability allows systems to adapt to evolving environments and operational objectives.

- **Improved Decision Quality in Relation to Rule-Based Systems**

Compared to fixed rule-based or heuristic approaches, RL can discover more efficient and optimal strategies by exploring the environment and learning from outcomes, often outperforming traditional methods.

- **Enables Intelligence and Automation**

Reinforcement learning enables autonomous systems to operate with minimal human intervention, reducing operational complexity and enabling intelligent automation in complex systems.

3. REVIEW OF RL ALGORITHMS FOR LOAD FORECASTING IN MICROGRIDS

In microgrids, lately, there has been more effort to use reinforcement learning to load forecasting, owing to its proficiency in repeated choice modeling and responding to demand alterations. In this text, I explain the latest RL-based forecasting approaches displaying their algorithms and the way to utilize them, important metrics involved and their practical application in the real world.

3.1 Categorization Based on RL Algorithm

The study indicates that the two categories of RL algorithms are widely applicable in making predictions [14]. Q-learning and DQN are used by many people as they are easy and assist in discrete action problems. Instead, issues with continuous load prediction have led to measures such as adopting policy gradient methods and actor-critic models such as PPO, A2C and DDPG. These algorithms exhibit better behavior in scenarios when both the load and time are affected by other means.

3.2 Model Architectures and Hybrid Frameworks

Most of the researchers have developed DRL models through applying RL and neural networks that enables us to appreciate and utilize information in high-dimensional input data. Moreover, these models succeed in the forecast because they take into account the variation of load with time. One has the thought to combine RL with a supervised technique like SVR or ensemble technique so as to assist the system to use the exploratory policy approach and yet observe the distinction in the representing functions.

3.3 Data Sources and Forecasting Scenarios

The majority of the works reviewed use real-world data available on OpenEI, the UCI Machine Learning Repository or provided by utilities or operate with the data synthesized under the GridLAB-D or Simulink simulation conditions. The majority of the cases contemplate the short-term load forecasting (STLF) with a time step ranging between 15 minutes to 1 hour [15], which is required in real-time control and storage scheduling. There are also a few attempts to use RL in medium-term or multi-purpose forecasting problems, largely in the form of model predictive control.

3.4 Evaluation Metrics and Performance Comparison

MAPE, RMSE and MAE are common performance assessment methods of RL-based forecasting models. In a number of situations, the RL-based techniques perform superior to the conventional methods, primarily where there are alterations in the load. Though they have benefits, the employments of the techniques in the resource constrained microgrid control system may be doubtful because they require more time and are more expensive during the training process. Following Table 1 presents the comparative analysis of RL algorithms.

Table 1. Comparative Analysis of RL Algorithms for Load Forecasting in Microgrids

Study	**RL Algorithm**	**Forecasting Horizon**	**Dataset**	**Performance Metrics**	**Key Findings**
Reinforcement Learning-Based Online Learning Strategy for Real-Time Load Forecasting	Q-Learning	Short-Term	Real-time meter data	MAPE, RMSE	Demonstrated improved adaptability in real-time forecasting scenarios.
A Review on Short-Term Load Forecasting Models for Microgrid	Various (including RL)	Short-Term	Diverse microgrid datasets	MAPE, MAE, RMSE	Provided a comprehensive comparison of models, highlighting the effectiveness of RL in certain contexts.
Energy Forecasting: A Comprehensive Review of Techniques and Applications	Deep Q-Network (DQN)	Short to Medium-Term	Thermal power unit data	MAPE, RMSE, MAE	CNN-LSTM-A model outperformed conventional LSTM, indicating the potential of hybrid deep learning approaches.
Short-Term Load Forecasting of Microgrid via Hybrid Support Vector Regression-Long Short-Term Memory	SVR-LSTM Hybrid	Short-Term	Rural microgrid in Africa	Correlation Coefficient	SVR-LSTM model achieved higher accuracy compared to individual SVR and LSTM models.

Enhancing Microgrid Performance Prediction with Attention-Based Deep Learning Models	Attention-Based GRU	Short-Term	Micro-grid Tariff Assessment Tool dataset	MAE, RMSE, R^2 Score	Achieved MAE of 0.39, RMSE of 0.28, and R^2 of 98.89%, outperforming traditional ML models.
Reinforcement Learning-Based Dynamic Model Selection for Short-Term Load Forecasting	Q-Learning	Short-Term	Two-year load and weather data	Not specified	Implemented a dynamic model selection approach, improving forecasting accuracy by approximately 50% compared to state-of-the-art ML models.
Optimal Scheduling of Isolated [16] Microgrids Using Automated Reinforcement Learning-Based Multi-Period Forecasting	Prioritized Experience Replay AutoRL	Multi-Period	Simulated microgrid data	Not specified	Proposed a forecasting method addressing error accumulation in multi-step forecasting, leading to improved prediction accuracy and reduced operating costs.

3.5 Comparative Summary of Reviewed Studies

To facilitate the comprehension, readers frequently find it helpful when a short comparison of the major studies is provided. The following table contains the description of various representatives of literature that are different in the type of algorithm, dataset, the distance into the future a forecast has to be made, whether hybrid methods are used and what performance is achieved. The systematic methodology enables us to identify the trend in selecting algorithms and highlights which RL methods are compatible with some characteristics of microgrid load forecasting.

4. CHALLENGES AND LIMITATIONS

Although RL may be quite effective in predicting smart loads in microgrids, the absence of decent data restricts its broad usage. In practical microgrids, data to carry out such optimizations is insufficient, fragmented or difficult to access in a uniform way. Deep reinforcement learning algorithms take powerful computers, a lot of memory and a lot of time to run. It poses challenges to microgrids that rely on low-power systems or minor computing locations. When the environment is non-stationary, these models may have difficulties in continuing to learn since the dynamic loads are continuously varying. However, extrapolation of the findings to unmeasured circumstances is yet to be achieved. RL models trained on particular datasets would exhibit an issue of new seasonal, demographic or geographic variations thus become inflexible. Such challenges have to be sorted out to have improved and flexible, reliable and instantaneous predictions in any network.

5. FUTURE RESEARCH DIRECTIONS

To bring the RL-based load forecasting in microgrids more realistic and scalable, several thrilling future investigations are emergent.

With the use of transfer learning and federated reinforcement learning, the adaptation of models and preservation of privacy in data becomes less challenging. Transfer learning, or transfer RL, makes it simple to apply what has been learned in one microgrid setting to another without much retraining. The focus is to ensure that these models are efficient to the point that the forecast can be undertaken on time. This method is also promising as it allows MARL to facilitate intelligent load shifting and peak shaving, with the reminder of real-time updates of forecasts, prices and user preferences. The future of MARL requires the sharing of popular benchmarks and datasets, to ensure fair comparisons and speed up progress. One set of evaluation measures and assessment steps will allow us to appreciate future models in a more reasonable way.

6. CONCLUSION

We have provided the entire literature survey on the position of reinforcement learning (RL) algorithms in microgrid load forecasting. Due to the very complex and dynamic nature of energy systems, the research considers applying Reinforcement Learning instead of the typical statistical or supervised learning. As we have seen, RL-based models can deal with nonlinear loads, real-time prediction and environmental uncertainty, but are not yet used in large scale due to absence of required data, extreme system load, issues of stability and limited adaptability. Meanwhile, new concepts in forecasting are being promoted by transfer learning, federated reinforcement learning, lightweight models and multi-agent systems.

The operation of microgrids will be based on RL operating alongside demand response, embedded systems and decentralized systems in order to become resilient, efficient and more intelligent. Any further studies should be based on a solid methodology that will rely on standard data and benchmarking strategies. An important advancement towards constructing intelligent, scalable and self-regulating forecasting systems required to promote sustainable energy in microgrids involves reinforcement learning.

REFERENCES

[1] She, B., Li, F., Cui, H., Zhang, J., & Bo, R. (2022). Fusion of model-free reinforcement learning with microgrid control: Review and vision. arXiv. https://arxiv.org/abs/2206.11398

[2] Yao, F., Zhao, W., Forshaw, M., & Song, Y. (2024). A holistic power optimization approach for microgrid control based on deep reinforcement learning. arXiv. https://arxiv.org/abs/2403.01013

[3] Ghasemi, A., Shojaeighadikolaei, A., & Hashemi, M. (2023). Combating uncertainties in wind and distributed PV energy sources using integrated reinforcement learning and time-series forecasting. arXiv. https://arxiv.org/abs/2302.14094

[4] Yu, L., Qin, S., Zhang, M., Shen, C., Jiang, T., & Guan, X. (2020). A review of deep reinforcement learning for smart building energy management. arXiv. https://arxiv.org/abs/2008.05074

[5] Agupugo, C. P., Tochukwu, M. F. C., &Ogunmoye, K. A. (2025). Review of smart microgrid platform integrating AI and deep reinforcement learning for sustainable energy management. International Journal of Future Engineering Innovations, 2(3), 1–17. https://www.researchgate.net/publication/391394132

[6] Biswal, B., Deb, S., Datta, S., Ustun, T. S., & Cali, U. (2024). Review on smart grid load forecasting for smart energy management using machine learning and deep learning techniques. Energy Reports, 12, 3654–3670. https://doi.org/10.1016/j.egyr.2024.02.039

[7] Ghosh, S., & Saha, H. (2020). Deep reinforcement learning for energy management in a microgrid. Sustainable Energy, Grids and Networks, 21, 100314. https://doi.org/10.1016/j.segan.2020.100314

[8] Sahoo, S., & Mohanty, S. (2023). State-of-the-art review on energy and load forecasting in microgrids using artificial neural networks, machine learning, and deep learning techniques. Renewable and Sustainable Energy Reviews, 165, 112604. https://doi.org/10.1016/j.rser.2022.112604

[9] Zhang, Y., Wang, J., & Li, Y. (2023). Demand-side load forecasting in smart grids using machine learning and deep learning techniques: A review. Frontiers in Energy Research, 11, Article 11065410. https://doi.org/10.3389/fenrg.2023.11065410

[10] Kumar, A., & Singh, R. (2022). A review on short-term load forecasting models for microgrid applications. IET Generation, Transmission & Distribution, 16(5), 1001–1012. https://doi.org/10.1049/tje2.12151

[11] Abdullah, D. (2025). Metamaterial-based antenna for beam steering in 5G mmWave bands. National Journal of RF Circuits and Wireless Systems, 2(2), 8–13.

[12] Madhanraj. (2025). Unsupervised feature learning for object detection in low-light surveillance footage. National Journal of Signal and Image Processing, 1(1), 34–43.

[13] Aslam, S., Herodotou, H., Mohsin, S. M., Javaid, N., Ashraf, N., & Aslam, S. (2021). A survey on deep learning methods for power load and renewable energy forecasting in smart microgrids. *Renewable and Sustainable Energy Reviews*, *144*, 110992.

[14] Arwa, E. O., & Folly, K. A. (2020). Reinforcement learning techniques for optimal power control in grid-connected microgrids: A comprehensive review. *Ieee Access*, *8*, 208992-209007.

[15] Wazirali, R., Yaghoubi, E., Abujazar, M. S. S., Ahmad, R., & Vakili, A. H. (2023). State-of-the-art review on energy and load forecasting in microgrids using artificial neural networks,

machine learning, and deep learning techniques. *Electric power systems research*, *225*, 109792.

[16] Chandrasekaran, K., Kandasamy, P., & Ramanathan, S. (2020). Deep learning and reinforcement learning approach on microgrid. *International transactions on electrical energy systems*, *30*(10), e12531.