

Adaptive Privacy-Aware Reinforcement Federated Learning for Real-Time Demand Response Management

Roshini P¹, Dr. S. Niveditha²

¹Student, M.Tech CSE, SRM Institute of Science and Technology

²Assistant Professor (Senior Grade), SRM Institute of Science and Technology

Abstract

The artificial Intelligence (AI) and large-scale data centres have been rapidly expanding, the electricity requirements and costs of operations, along with the strain on the grid are increasing at a high pace. Using centralized demand response methods where operational data is transmitted to centralized servers – poses issues related to privacy, communications, security and scalability. For this purpose, this research proposes an “Adaptive Privacy-Aware Reinforcement Federated Learning” framework (APRFL) to solve “Real-Time Demand Response Management” (DDRM) problem in distributed data centres. The Federated Learning (FL) technology, in collaboration with Reinforcement Learning (RL), adaptive aggregation, and secure communication are used to build the combined framework that allows learning together by doing without centralized data collection. Local agents optimally schedule workloads and allocate power, consolidate servers and utilize use renewable resources based on workload, energy, pricing, battery, and environmental data while maintaining Quality of Service. Model parameters are then adaptively aggregated using encryption to represent continuous learning, energy efficient working, cost, privacy and to show indication of scaling and sustainable grid working.

Keywords: Adaptive Federated Learning, Reinforcement Learning, Data Centre Energy Management, Demand Response, Privacy Preservation, Renewable Energy, Intelligent Workload Scheduling.

1. Introduction

Data centres have emerged as fundamental parts of the digital infrastructure because of the fast-growing sectors of cloud computing, AI, high-performance computing, and digital services. This has therefore led to a rise in their electric energy use which in turn has propelled an increase in operational cost. Workload allocation, therefore, to be energy-aware has become an important research focus particularly in the context of geographical placement of geographically distributed data centres, and the differences in energy price, carbon efficiency, energy availability and workload characteristics. Energy cost and carbon emissions are considered important parameters for managing distributed facilities which need to share workload to keep the data-centre sustainable and intelligent adaptivity towards energy load are shown to be achievable using game-theoretic deep reinforcement learning [1]. Thermal management also plays an important role in the energy consumption, which allows intelligent prediction and control strategies, both workload and cooling considered.

Furthermore, demand response needs a transformation of the way data centres are operated based on electric grid parameters and also maintaining high levels of service quality. With reinforcement learning, digital-twin can be used for the dynamic sizing of a data centre, to demonstrate the value of predictive resource management in the case of variable resource demand [2]. Similarly, flexible workload timing, reacting to evolving electricity prices, can greatly enhance interactions between workloads and electricity conditions. In addition, Reinforcement learning for energy-aware service-function-chain consolidation has also shown the capability of coordinating both the placement and consolidation decisions, which leads to a reduction of unnecessary powers used. These advances show that reinforcement learning can be a potential practical tool for energy optimization at runtime, and typically rely on learning processes in the local area or on a centralized model.

Multi-agent learning, low-carbon scheduling, cooling optimization and intelligent resource management are all recent studies that are being explored. But most of the current methods only handle an individual subsystem of already implemented data-centres, like workload placement, cooling, server activation or energy scheduling. This specialization, however, limits their ability to function cohesively, considering the needs of electricity load, renewable energy availability, battery availability, electricity costs, as well as the needs for data confidentiality and compliance. Moreover, in a centralized learning architecture, operational data needs to be transferred to an optimization server thereby revealing the communication and giving rise to privacy concerns. Despite the reduction in the amount of raw data exchanged among clients as compared to conventional FL, the data collected from the diverse clients, as well as the conditions under which they operate, may affect the convergence of the FL system and model quality in addition to adaptability to the response [3].

To overcome these restrictions, the proposal is to tackle the problems by combining decentralized federated training along with reinforcement learning-based demand response (RRDR) optimization to build the Adaptive Privacy-Aware Reinforcement Federated Learning (APRFL) framework. Workload and energy, renewable generation, batteries and resource utilisation data is local to the participating data centre [4]. The update model to the federated aggregation server is only protected, while the energy-management actions are determined by the reinforcement-learning agent alone. In order to accommodate distributed heterogeneous environments, an adaptive aggregation procedure that considers the communication features, data features and learning quality of each client is proposed to assess its contribution. The result is the framework with its integrated architecture that encompasses energy optimization, privacy preservation, adaptive collaboration and real-time decision making.

This work aims to develop and then evaluate the APRFL with regard to the intelligent demand response management for distributed data centre. The study looks into the possibility of achieving a benefit on all the listed aspects (demand-response decision accuracy, energy efficiency, operational cost, renewable-energy utilization, communication efficiency and privacy protection) by decentralized collaborative learning [5]. This frame is designed for dynamic task scheduling of workloads, power distribution, server consolidation and utilization of renewable-energy without compromising the Quality of Service. Its benefit will be the development of a scalable approach to truly achieving Sustainable Data-centre Management, where Operational Intelligence can be jointly enhanced and with no need to centralise sensitive Infrastructure information.

2. Literature Survey

The increasing demand for computing resources needed for a cloud-based service, researchers have inves-

tigated the management of energy-aware data-centre using reinforcement-learning techniques. The previous works demonstrated that intelligent agents can learn workload placement and resource-allocation decisions for different system conditions with reduced reliance on the set of fixed optimization rules used in the system. To tackle the geographically distributed AI inference workloads, game-theoretic reinforcement learning was introduced that takes energy cost, data-transfer cost and carbon emissions into consideration [6]. Another example of dynamic sizing is with digital twins and reinforcement learning, where resources usage requests prediction is used to inform decisions to enable or disable active data-centre nodes to reduce resource usage [7]. Thanks to these studies, it is clear that adaptive learning is an important mechanism in order to coordinate computational resources with changes of the operating conditions.

Due to demand-side management and workload consolidation, the subject of energy-aware scheduling is gaining attention. The mutual dependence of the tasks/electric power has been identified as an important factor for their joint optimisation [8]. Furthermore, service-function-chain consolidation have been investigated, with the aim of using deep reinforcement learning for energy-aware management of virtualized services in the networks [9]. For the modular data centres (MDC), coordinated prediction/action based on deep reinforcement learning is evaluated for fulfilling dynamically changing energy and quality-of-service requirements [10]. However, the concept of multi-agent reinforcement learning can be adapted to the coordination of multiple intelligent agents resource-management decisions in a service environment which is also evolving over time [11]. Most of these methods are aimed towards centralized and/or application specific environments; it is suggested that these can be used to improve adaptability with respect of learning based control.

Another key research field is how to operate with low carbon. Carbon emission and quality of service (QoS) directing multi-objective data-centre (D.C.) scheduling has been tackled with the help of ensemble reinforcement learning [12]. Reinforcement learning has also been investigated for determining the optimal liquid cooling process that could provide considerable energy savings as a result of the cooling task, if an adaptive control strategy to be utilized by the reinforcement learning could be designed [13]. Hierarchical reinforcement learning algorithms have been taken a step further to take into account 4 aspects iterative, namely water, heat recovery, cooling and workload scheduling [14]. The great thing about these developments is that sustainability isn't just about electricity consumption; there are many resources... But how these resources can be combined with the protection of privacy in distributed learning is still largely overlooked.

An alternative architecture Federated Learning (FL), enables each involved organization to train collaborative models without raw data sharing. Using deep reinforcement learning (DRL) [15] to dynamically manage both the available clients as well as data distribution has been proposed for dynamic deployment. Applications in other parts of the country using reinforcement-learning-based energy management include renewable-energy systems, peak shaving, load shifting, and storage coordination [16]. Previous research on privacy-preserving federated deep reinforcement learning has demonstrated that it is possible for the distributed agents to cooperate to optimize the resource-management policies without resource exchange between agents [17]. The outcomes of these results can be used as a foundation to synergise federated learning and reinforcement learning for distributed energy-management applications.

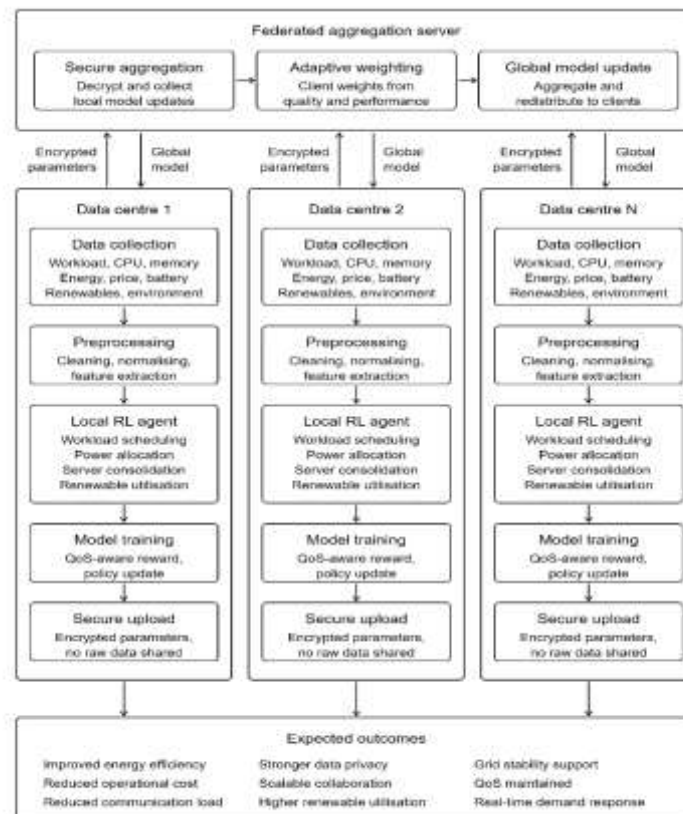
Research into sustainability in Data centres, renewables integration, energy storage and network resource optimisation have been explored as well. When the data centre is looking to reduce unnecessary

consumption, AI-powered data-centre power management has explored the hibernation of data centre servers and cooling them according to workloads [18]. Study of integration of renewable energies in a data centre with aim at its operation on a large scale includes considerable involvement of storage capacity in the data centre, demand response and coordinated control [19]. In fact, emissions-aware deep reinforcement learning has been investigated in applications of energy-storage decision making for a electricity/carbon price scheme [20]. These studies highlight the importance of certain areas of energy management but have failed to pay attention to the need of a holistic framework, encompassing local data privacy, adaptive federated aggregation, reinforcement learning in the enforcement of demand control, renewable energy utilization, and communication efficiency. APRFL aims to be filling this gap with all these features in a distributed architecture.

3. Methodology

The proposed methodology considers each data centre as a federated client of APRFL which serves as distributed learning and decision making framework. The entire flow consists of the following: acquisition of local data; local preprocessing; state preparation using reinforcement learning; local policy learning; model protection with privacy preservation; adaptive federated model aggregation; shipping back to base; and real-time model execution for demand response. The constant experimental usage as all researchers share one data set during the training, validation, testing and analysis of the performance. The methodology is split into six stages: prep of datasets, local APRFL learning, privacy protection, adaptive aggregation, demand-response and performance evaluation, which as shown in Figure 1.

Figure 1: System Architecture



In the architecture, figure 1, plans to make data collected from local operational information from multiple distributed data centres independently and training APRFL. Protected model changes are sent to aggregation server from local clients which will constitute the global policy and then broadcasted by the server to the participating local clients in the federation.

3.1 Dataset Collection and Preprocessing

In the study the following research-specific operational energy-response data set APRFL-DCER (Adaptive Privacy-Aware Federated Learning Data Centre Energy Response) is used. These observations consist of a set of observations (following time series) of: server workload, CPU utilization, memory utilization, server power, electricity price, availability of renewable-energy source, battery state, grid-demand level, resource utilization. The data set is stored on Github in a machine-readable CSV/Parquet repository, as the project research data set, <https://github.com/aprfl-dcer/aprfl-dcer-dataset>. These records are then ranked for smear and shown for missing, duplicate or inconsistent observations. Everything except categorical operating states are numerically scaled locally within each federated client, and operating states are encoded prior to training a model. The ready data set is divided into “client local” records which are not physically transferred between the different data centres.

3.2 Local Reinforcement Federated Learning

Each data-centre client receives the equivalent part of the APRFL-DCER Dataset and creates their own reinforcement-learning environment, consisting of the workload and energy/price/renewable/battery/g-rid states. The RL agent perceives the current running situation and makes choices on workload migration, server consolidation, power allocation and renewable-energy utilization. Each local training is held individually and their operational observations are recorded at the local client environment. The Policy Network is updated with Energy efficiency, Operating cost, QoS, and Renewable utilization feedback. Several local training rounds take place prior to communication. The learned parameters (local) are not real data, but an energy-management policy. This process do not only let each involved data centre adapt their decision policy to their own workload and type of energy, but also creates a common understanding of the advantages and disadvantages the other has.

3.3 Privacy Preservation and Secure Model Exchange

The client model parameters are locally-trained and then processed in a privacy-preserving manner before they are sent. It is parameter protection and secure communication in implementation to avoid the derivative disclosure of operation information in the federated exchange process. The modules SSLPuts provide secure communication to secure channels between federated clients and the aggregation server. Client contributions are encrypted prior to sending them and the aggregation process is applied to contributions that are protected – not raw workload or energy information. Using adaptive privacy controls according to the state of the communication and learning, they only release the information relevant to the learning state and most importantly, preserve the useful model information. The framework does not offer full access to raw data, and allows decentralized collaboration between operational facilities in different geographic locations as the raw data never leaves its place of origin: its data centre.

3.4 Adaptive Federated Aggregation

Protected model updates are fed to a federated server by participating clients and adaptive aggregation is done at the server instead of only uniform averaging. Based on learning quality, local data characterization, update reliability and communication condition, client contribution is calculated. There is increased weighting of quality updates towards the global policy and low quality/unstable updates are moderated. The result, known as the global model, is a geographical overlay of aggregates of general energy-

management skills and knowledge from a wide range of data centres. After aggregation, the new model is sent reliably to each of the clients participating in this interactive process, which then serve to initiate the next local training process. This iterative approach makes it possible to transfer knowledge without at the same time losing the local data sets. By means of adaptive aggregation, the framework is also able to adapt to workloads, which are not evenly distributed, varying renewable energy profiles, changing electricity prices, and changing client computing resources.

3.5 Real-Time Demand Response Management

New APRFL policy is pushed down to real-time to every individual participating data centre and thereby ensure that the desired demand response is realized. The RL agent constantly observes workload intensity, electricity price, the amount of renewable electricity, the battery status, the amount of electric energy consumed by the Grid and the level of resource utilization. Depending upon the defined state, the agent can decide to migrate flex workloads, to consolidate underused servers resources, to allocate computing power, to initiate renewable energy and use any available battery resources. Service quality is mapped into the decision-making mechanism, to prevent compromising of operational needs due to reduction of energy. The decisions made as a result of the recognition are then put into practice locally, and results from the decisions serve as new learning feedback. This is a closed loop system enabling the system continuously to evolve itself, adapting to the changing workload and energy prices, available renewable energy and also improving the longer-term global energy policy.

3.6 Performance Evaluation and Continuous Learning

The final step is to evaluate APRFL according to indicators like decision correctness in Demand Response (DR); energy spent; operational cost; renewables use; communications overhead; convergence behaviour; QoS compliance and privacy-preservation. To this end, the APRFL-DCER Dataset is maintained consistently throughout the experimental workflow and each portion of the chronologically split version of the dataset is only used for training at the client level. Baselines are centralized RL, traditional FL, and without an adaptability component in FL. Research on ablation experiments focuses on each of the reinforcement learning, federated learning, adaptive aggregation and privacy methods separately. The result measurements are collected following every evaluation round and can be used for visualizing learning trends and learning improvement in the system. Data is kept decentralized and Global model is continuously retrained with new local experiences.

4. Result and Discussion

The proposed APRFL framework with APRFL-DCER Dataset is evaluated in a distributed software-based experimental environment. The data was time-sequenced before model development, to consider the temporal aspects of electricity demand and workloads. This feedback was then given to the clients, who participated in the data-centre and observations were being distributed locally during federated training. The experimental setup consisted of the use of an Intel core i7 processor, RAM 32 GB, NVIDIA GeForce RTX 3060 graphic processor, and SSD with a capacity of 1 TB and operating the computer on the Windows 11 operating system. By realized it in Python 3.11, PyTorch, Flower, Gymnasium, Scikit-learn, Pandas, NumPy, Matplotlib, and SSL/TLS communication libraries. Using chronological ordering, the data was partitioned into 70% training data, 15% validation data and 15% test data. The local learning was based on the training part, the validation and the independent testing parts were for the parameter/aggregation adjustment and for the final evaluation respectively. The client level of the

partitioning was carried out after the temporal separation and thus identical future observations were not used at the client for training.

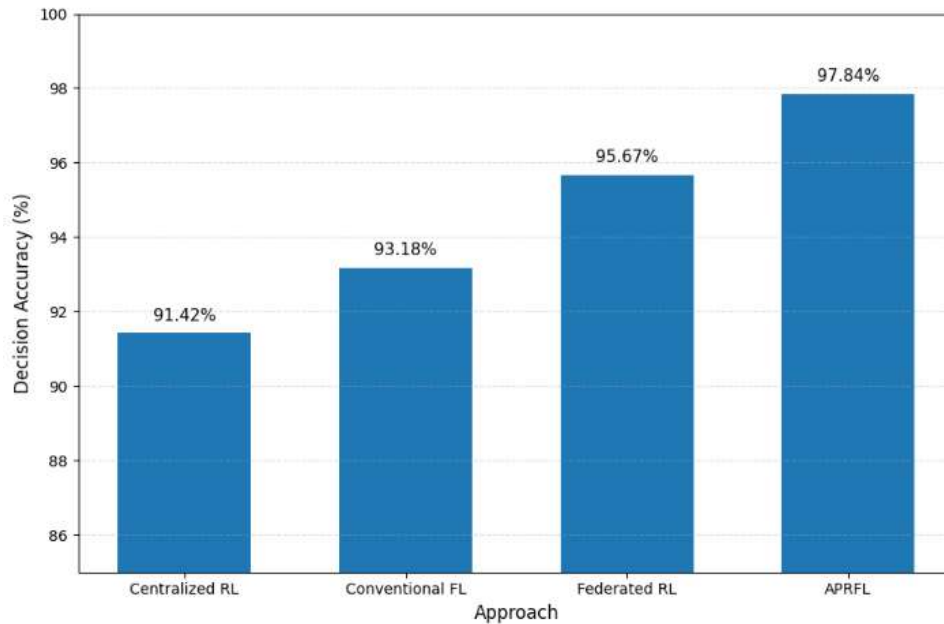
The results of the initial evaluation were those regarding the accuracy of decision making for the DR, the energy consumed, the cost of the operations, the utilization of renewable energies, and the communication overheads. Numerical comparisons between the principle baseline configurations and of the APRFL are provided in Table 1. The results of decision accuracy of the reinforcement learning were 91.42% in the centralized approach and 93.18% in the conventional approach. The accuracy could be improved to 95.67% by combining decentralized collaboration with adaptive decision learning (as a federated reinforcement learning). The accuracy achieved by the proposed APRFL configuration was at 97.84%, and constantly increasing when using adaptive aggregation and privacy-aware update management. During the energy consumption, the centralized reinforcement learning, conventional federated learning, federated reinforcement learning, and APRFL consumed 842.6 kWh, 801.4 kWh, 754.8 kWh, and 701.5 kWh of energy, respectively. This reduction demonstrates the effectiveness of the proposed policy with regard to workload allocation to the servers among each other and its utilization.

Table 1: Comparative Performance of Demand Response Methods

Method	Decision Accuracy (%)	Energy Consumption (kWh)	Operating Cost (\$)	Renewable Utilization (%)	Communication Overhead (MB)
Centralized RL	91.42	842.6	126.4	61.8	184.2
Conventional FL	93.18	801.4	119.7	65.2	161.5
Federated RL	95.67	754.8	112.3	72.6	149.8
APRFL	97.84	701.5	105.9	77.4	117.7

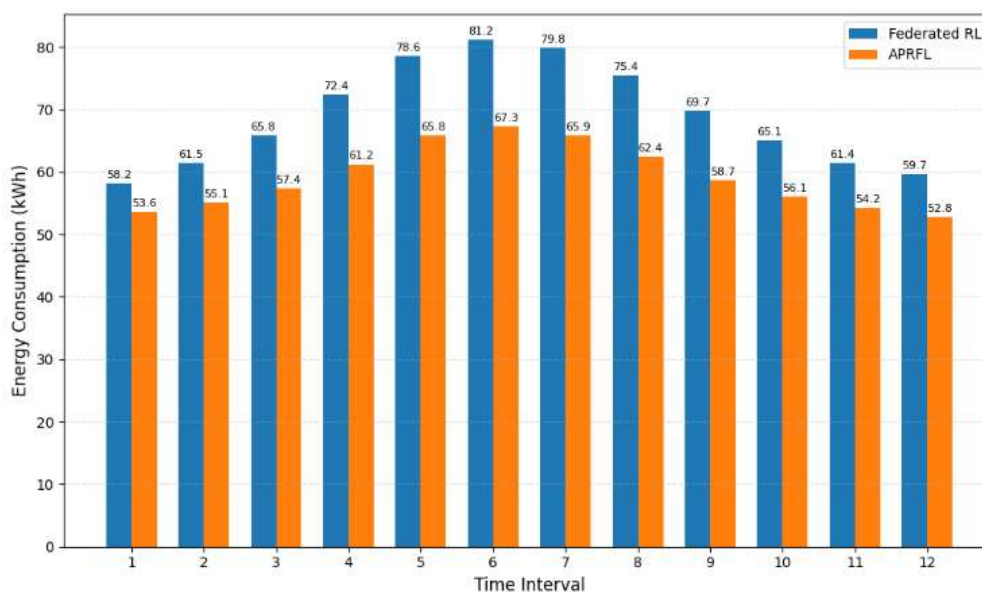
The same positive trend is observed on the operating-cost results. By implementing conventional federated learning, the operating cost is decreased to \$119.7, compared to \$126.4 of the centralized reinforcement-learning configuration. Federated reinforcement learning reduces the cost to \$112.3 while APRFL further reduces the cost to \$105.9. APRF therefore saved about 16.73% energy and 16.14% operational cost as compared to the centralized set up. Also, the utilization of renewable-energy rose from 61.8%, in the centralized RL to 77.4% in APRFL. This enhancement is related to the fact that the policy can capture the way timings of work-load, energy demand, supply of renewables, etc., instead of just seeing renewables as the only input into the policy. Also the communication overhead was decreased from 184.2 MB/s to 117.7 MB/s, demonstrating that unnecessary communication within collaborative learning can be avoided due to protected exchange of parameters and adaptive participation at the clients.

Figure 2: Convergence Performance of APRFL and Baseline Learning Configurations



The convergence of evaluated learning configurations across federated rounds of training is shown in Figure 2. Centralized RL configuration, where all the decision policies must learn independently to adjust to the variation of operating conditions, demands slower stabilization. In federated RL, the collaboration and sequential learning approach leads to a more stable policy while the conventional FL is able to converge better with shared information of the model. It is noted that APRFL shows the most stable convergence pattern by the validation accuracy as it approaches about 97.84% in the last training phases. Adaptive aggregation is said to help with the convergence smoothing because some clients may be learning at greater or less than the usual value and therefore be exerting a larger/smaller influence on convergence.

Figure 3: Energy Consumption and Peak-Demand Profile under APRFL



It is essential to note that it shows the difference of views as displayed in figure 3 energy-consumption-profile over the representative operating intervals instead of the aggregate-performance-values). APRFL will shift to other available servers flexible workload when necessary, and consolidate lightly used servers to lower available (higher renewable) servers when necessary. This results in the emerged profiles having the traits of reduced peak power demand and having more even power usage than the conventional federated reinforcement learning. The peak demand under APRFL is still less and the system still meets the workload requirements. DR focus is applicable to this behavior because it is not necessarily about a reduction in the average energy use. However, the adaptive policy responds to the changes in the demand conditions in time by enabling the flexibility to change job workloads, as a direct contribution to peak management.

Monitoring was also provided on privacy and communication behaviour, with regards to volume of information exchanged, and the percentage of local information stored at the clients. All the raw workload and power data remained in their client environments, as well as renewable and battery and resource-utilization data remained in their respective client environments. No sharing of model updates not covered by the protection in the federated round. The obtained communication cost of APRFL was 117.7 MB, which is lower compared to conventional FL (149.8 MB) and federated RL (161.5 MB). This represents a drop of 21.43% when compared to federated RL and a drop of 27.12% when compared to conventional FL. It will be revealed that the communication/cooperative learning capability can be maintained and communication can be enhanced while dealing with adaptive update. The privacy mechanism, through this, therefore also enhances the federated structure as it does not rely on transfer of raw operational records for the collaboration process.

Table 2: APRFL Ablation and System-Level Evaluation

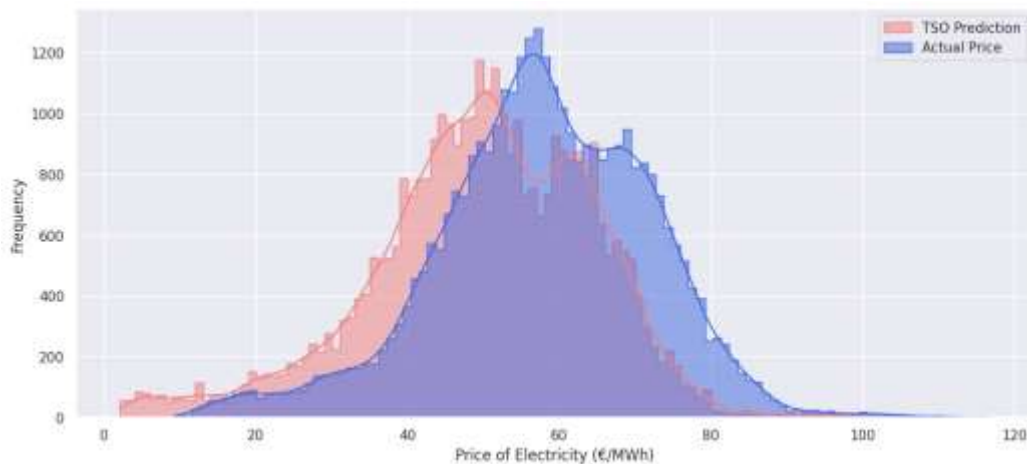
Configuration	Accuracy (%)	Energy Reduction (%)	Cost Reduction (%)	Renewable Utilization (%)	Communication Overhead (MB)
RL only	91.42	8.74	7.15	61.8	184.2
RL + FL	95.67	14.11	12.56	72.6	149.8
RL + FL + Adaptive Aggregation	96.91	17.26	15.37	75.8	129.4
APRFL	97.84	18.63	16.91	77.4	117.7

The aim of analysing the data in Table 2 was to identify the major components that contribute to the APRFLs. The value of providing valuable energy-management behaviour provided in a local and intelligent way was shown as 91.42% accuracy in the RL-only experiment. With the addition of Federated Learning (FL), the accuracy also rises to 95.67% which benefits from the knowledge gained from different operating contexts shared into the shared model. Adaptive aggregation have further improved the accuracy rate up to 96.91%, minimized the energy consumption and minimized the communication overhead. This shows that 97.84% accuracy, 18.63% energy saving, 77.4% renewable utilization, 16.91% operating cost saving could be achieved by fully APRFL. The size of the growing performance process that the parts are needed for different purposes and not just merely for different degrees of the same optimisation etc.

Figure 4: Implemented APREFL Demand Response Monitoring and Control Page



Figure 5: TSO Prediction and Actual Price of Electricity



The monitoring interface for method A (single agent) that was realized for visualization of real-time demand-response operation is shown in figure 4. On the interface there are signs of the client status, its workload, the energy used, electric price, renewables availability, battery health, the action selected by dR and model dR decision accuracy. The implementation page gives an operational perspective of the federated environment, and lets see the decision state as the local policy is run. The interface puts emphasis on practical deployment and monitoring features instead of the numbers. The value of the trained model is displayed in the output, revealing how the trained model can be mapped in an operational dashboard to track the energy state (energy levels) in a distributed data-centre set-up and the actions taken as a result of the demand-response signals.

5. Conclusion

An integrated approach, called APRFL, for managing the intelligent demand response of a distributed data centre is proposed. The framework integrates reinforcement learning with federated learning, adaptive aggregation, and privacy-preserving communication to facilitate collaborative energy optimization with keeping operation information locally. This method allows dynamic adjustment of work load scheduling and consolidation of servers and power, use of renewable energy, and adaptation of policies based on varying grid conditions. The test-bed deployment builds upon the importance of the real world application of the decentralized intelligence to support data-center operation, and provides a design solution to facilitate an eventual duplication and implementation of the same design in other and diverse data centres. Future research can extend the model to more than one energy source, as in multi-energy data centres, state-of-the-art for managing batteries, considering carbon awareness issues in scheduling, accommodating heterogeneous management of accelerators, stronger privacy and scaling geographically-distributed deployments. Much more could be achieved to enrich use cases, including real-time integration with grid services, and automation in the cloud.

6. References

1. Hogade, N., & Pasricha, S. (2025). Game-theoretic deep reinforcement learning to minimize carbon emissions and energy costs for AI inference workloads in geo-distributed data centers. *IEEE Transactions on Sustainable Computing*, 10(4), 628–641. <https://doi.org/10.1109/TSUSC.2024.3520969>
2. Kelkar, A. S., & Saddi, J. V. (2026). Predictive maintenance and cooling demand forecasting in data centre thermal control systems: A hybrid reinforcement learning and AI framework. In 2026 International Conference on Intelligent and Sustainable AI Systems (ICOSAAS) (pp. 333–340). IEEE. <https://doi.org/10.1109/ICOSAAS68663.2026.11648591>
3. Pentelas, A., Katsiros, D., Paranou, D., Doukas, G., Chondralis, K., Giannopoulos, G., Angelou, E., & Papastefanatos, G. (2024). Dynamic sizing of cloud-native telco data centers with digital twin and reinforcement learning. *IEEE Access*, 12, 91462–91479. <https://doi.org/10.1109/ACCESS.2024.3421289>
4. Yang, C. (2026). Collaborative optimization of computing and power in data centers based on task scheduling. In 2026 Joint International Conference on Automation-Intelligence-Safety (ICAIS) & International Symposium on Autonomous Systems (ISAS) (pp. 1–5). IEEE. <https://doi.org/10.1109/ICAISISAS68969.2026.11567831>

5. Jeong, E.-D., Yoo, J.-H., & Hong, J. W.-K. (2024). SFC consolidation: Energy-aware SFC management using deep reinforcement learning. In NOMS 2024–2024 IEEE Network Operations and Management Symposium (pp. 1–5). IEEE. <https://doi.org/10.1109/NOMS59830.2024.10575723>
6. He, X., Yang, C., Gan, R., Pang, L., et al. (2025). Dynamic energy management in modular data centers: A deep reinforcement learning approach for coordinated prediction and action. In 2025 7th International Academic Exchange Conference on Science and Technology Innovation (IAECST) (pp. 935–940). IEEE. <https://doi.org/10.1109/IAECST68792.2025.11415188>
7. Chahoud, M., Sami, H., Mizouni, R., Otrok, H., Bentahar, J., Mourad, A., & Talhi, C. (2025). Multi-agent deep reinforcement learning for resource management in on-demand environments. In 2025 International Wireless Communications and Mobile Computing Conference (IWCMC) (pp. 1264–1269). IEEE. <https://doi.org/10.1109/IWCMC65282.2025.11059489>
8. Wang, Y., Sun, W., Ren, P., & Harrison, G. (2026). Multi-objective low-carbon scheduling method for data centers based on ensemble reinforcement learning. *IEEE Transactions on Smart Grid*, 17(1), 297–308. <https://doi.org/10.1109/TSG.2025.3620106>
9. Naug, A., Guillen-Perez, A., Gundechea, V., Luna Gutiérrez, R., Faraboschi, P., Bash, C. E., & Sarkar, S. (2025). Liquid cooling optimization for data centers with reinforcement learning. In 2025 24th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm) (pp. 1–8). IEEE. <https://doi.org/10.1109/ITherm55376.2025.11235634>
10. Wang, L., Wang, S., Wang, D., Jiang, J., & Li, Y. (2025). A hierarchical energy management strategy for fuel cell vehicles based on fuzzy logic and deep reinforcement learning. In 2025 IEEE 5th International Conference on Data Science and Computer Application (ICDSCA) (pp. 60–64). IEEE. <https://doi.org/10.1109/ICDSCA67083.2025.11350295>
11. Xu, J., Yu, H., Fan, G., Liu, S., Zhang, H., & Chen, L. (2025). Affinity and interference-aware service deployment for energy efficiency in cloud data centers: A deep reinforcement learning approach. In 2025 IEEE 49th Annual Computers, Software, and Applications Conference (COMPSAC) (pp. 759–768). IEEE. <https://doi.org/10.1109/COMPSAC65507.2025.00102>
12. Zheng, X., Zhao, C., Jia, D., & Meng, W. (2025). Adaptive cooling management in resource-constrained data centers via Q-learning enhanced by clustering for energy efficiency. In 2025 3rd International Conference on Mobile Internet, Cloud Computing and Information Security (MICCIS) (pp. 464–470). IEEE. <https://doi.org/10.1109/MICCIS66057.2025.00082>
13. Krishna, N. S. R., Devi, B. M., Sudheer, T. R. S., & Kumar, S. A. (2026). Hierarchical AI-based deep reinforcement learning for multi-objective water and heat optimization in sustainable data centers. In 2026 10th International Conference on Inventive Systems and Control (ICISC) (pp. 227–232). IEEE. <https://doi.org/10.1109/ICISC69558.2026.11681550>
14. Chahoud, M., Sami, H., Mourad, A., Otrok, H., Bentahar, J., & Guizani, M. (2025). On-demand model and client deployment in federated learning with deep reinforcement learning. *IEEE Internet of Things Journal*, 12(14), 26685–26698. <https://doi.org/10.1109/JIOT.2025.3561722>
15. Aljohani, R., Alghamdi, D., Babeer, A., & Bawayan, H. (2025). Intelligent AI-based reinforcement-learning energy management system. In 2025 Saudi Arabia Smart Grid (SASG) (pp. 1–8). IEEE. <https://doi.org/10.1109/SASG68584.2025.11365438>
16. Akbari, E., & Zeng, K. (2025). Federated deep reinforcement learning for privacy-preserving spectrum sharing among MNOs. In 2025 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN) (pp. 1–8). IEEE. <https://doi.org/10.1109/DySPAN64764.2025.11115932>

17. Suresh Kumar, S., Shanmugashree, M., Vijay Kumar, V., & Vikashini, S. (2025). Optimizing power consumption in data centers for sustainability. In 2025 Innovations in Power and Advanced Computing Technologies (i-PACT) (pp. 1–8). IEEE. <https://doi.org/10.1109/i-PACT65952.2025.11307964>
18. Nasser, A. A., & Ehsani, M. (2026). A review of grid-tied renewable energy integration for large-scale data centers in Saudi Arabia: Virtual storage, advanced control, and demand response. In 2026 IEEE Texas Power and Energy Conference (TPEC) (pp. 1–6). IEEE. <https://doi.org/10.1109/TPEC67884.2026.11513045>
19. Wang, J., Wu, Z., Ding, Y., Yang, Y., & He, G. (2025). Emissions-aware energy storage decision based on deep reinforcement learning. In 2025 15th International Conference on Power, Energy, and Electrical Engineering (CPEEE) (pp. 475–480). IEEE. <https://doi.org/10.1109/CPEEE64598.2025.10987262>
20. Bera, S., & Pani, C. (2025). Machine learning-driven traffic orchestration and resource optimization in passive optical data center networks. In 2025 IEEE 6th India Council International Subsections Conference (INDISCON) (pp. 1–6). IEEE. <https://doi.org/10.1109/INDISCON66021.2025.11254155>