IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal
IEEE POWER & ENERGY SOCIETY SECTION

Received 24 April 2024, accepted 1 May 2024, date of publication 8 May 2024, date of current version 15 May 2024.
Digital Object Identifier 10.1109/ACCESS.2024.3398059

RESEARCH ARTICLE

Data-Driven Energy Management of an Electric Vehicle Charging Station Using Deep Reinforcement Learning

G. S. ASHA RANI[1], P. S. LAL PRIYA[1], (Member, IEEE), JINO JAYAN[2], RAHUL SATHEESH[3], (Member, IEEE), AND MOHAN LAL KOLHE[4]

[1]Department of Electrical Engineering, College of Engineering Trivandrum, APJ Abdul Kalam Technological University, Thiruvananthapuram, Kerala 695016, India
[2]Department of Electronics and Communication Engineering, College of Engineering Trivandrum, APJ Abdul Kalam Technological University, Thiruvananthapuram, Kerala 695016, India
[3]Amrita School of Artificial Intelligence, Amrita Vishwa Vidyapeetham, Coimbatore 641112, India
[4]Faculty of Engineering and Science, University of Agder, 4604 Kristiansand, Norway

Corresponding authors: G. S. Asha Rani (asha.eecet@gmail.com) and Mohan Lal Kolhe (mohan.l.kolhe@uia.no)

This work was supported by the University of Agder, Norway.

ABSTRACT A charging station that integrates renewable energy sources is a promising solution to address the increasing demand for electric vehicle (EV) charging without expanding the distribution network. An efficient and flexible energy management strategy is essential for effectively integrating various energy sources and EVs. This research work aims to develop an Energy Management System (EMS) for an EV charging station (EVCS) that minimizes the operating cost of the EVCS operator while meeting the energy demands of connected EVs. The proposed approach employs a model-free method leveraging Deep Reinforcement Learning (DRL) to identify optimal schedules of connected EVs in real time. A Markov Decision Process (MDP) model is constructed from the perspective of the EVCS operator. The real-world scenarios are formulated by considering the stochastic nature of renewable energy and the commuting behavior of EVs. Various DRL algorithms for addressing MDPs are examined, and their performances are empirically compared. Notably, the Truncated Quantile Critics (TQC) algorithm emerges as the superior choice, yielding enhanced model performance. The simulation findings show that the proposed EMS can offer an enhanced control strategy, reducing the charging cost for EVCS operators compared to other benchmark methods.

INDEX TERMS Deep reinforcement learning, electric vehicle, energy management strategy, Markov decision process, renewable energy, truncated quantile critics.

I. INTRODUCTION

In recent years there have been significant advancements in battery technology along with increasing awareness of climate change and global warming. All these resulted in new opportunities for the widespread deployment of EVs [1]. The EVs directly replace fossil fuels with electricity, positively affecting the environment and the economy [2]. The increasing number of EVs on the electric distribution grid is beyond the capacity of the existing EVCS. Since the number of charging stations is not increasing at the same rate as the number of EVs being produced, there will inevitably be an infrastructure capacity shortfall issue. Consequently, more EVCSs are needed in public areas like highways, office buildings, and residential neighborhoods. Profit mechanisms should be suitably developed to sustain the continuing operation of EVCSs, as they are typically affected by issues related to high investment and costly operations [3].

Incorporating renewables into EV charging infrastructure offers a global solution to the dual challenges of fossil

The associate editor coordinating the review of this manuscript and approving it for publication was Yifan Zhou.

fuel dependence and air pollution [4]. It has several advantages besides reducing harmful emissions, which creates a healthier earth and cleaner air. EVCS could become energy independent through on-site renewable energy generation, eventually making EVCS operation more profitable inspite of the enormous initial capital required to install renewable energy sources (RES) [5]. This will also ease the burden on the grid. Renewable energy surplus can be returned to the grid, increasing resilience and overall efficiency [6]. However, the energy management of EVCS is a challenging problem due to the variability in the power from RES.

The increasing number of EVs will significantly change the electric load profiles of power systems by bringing unpredictable and dynamic electricity needs [7]. This presents new challenges and limitations for the grid. An optimal charging strategy for EVs is crucial to overcome these challenges. Numerous algorithms, encompassing optimization and rule-based methodologies, have been proposed in the literature. To ascertain how energy is distributed across several power sources, rule-based techniques are developed utilizing predetermined regulations, human skills or heuristics [8]. These techniques are quite strong and dependable, but they lack the adaptability and flexibility required for charging cycles [9]. Conversely, optimization methods entail the development of a mathematical model of the environment to compute the optimal behavior of the system. To improve the optimization of EVCS infrastructure, Zeng et al. [10] carried out a study using actual EV travel data and a multi-objective genetic algorithm. The basic concept is that an optimally designed charging infrastructure can mitigate load fluctuation and the high energy consumption costs on the grid. However, in the context of an EV charging management application, the system is highly dynamic, making it challenging for optimization techniques to assess the optimal behavior for EV charging management systems.

To efficiently supervise and monitor the available energy resources at EVCS, the implementation of an intelligent EMS is essential.The International Electrotechnical Commission defines EMS as a computer system comprising a software platform offering essential support services and a set of applications that deliver the required functionality for the effective operation of electricity generation and transmission facilities. The primary objective is to ensure a secure and ample energy supply while minimizing costs [11]. The EVCS has to manage how energy is distributed among available renewable energy sources, ESS and the main grid in addition to dispatching EV charging power. The design of such systems is intricate due to the intermittent electricity output of PV systems and the irregular traffic at EVCS [12]. The uncertainty associated with the arrival and departure times of EVs significantly affects the energy management and scheduling techniques of EVCS. It is imperative to guarantee the stability and dependability of the EMS, particularly when dealing with uncertainties related to renewable energy and EV charging demand.

In [13] the Model predictive control(MPC) strategy is utilized to develop an EMS for a hybrid charging station for EVs. MPC is widely used in literature for the effective implementation of EMS. The MPC strategies use deterministic and stochastic control; the latter can withstand unforeseen interruptions without compromising performance quality [14]. MPC is a model-based approach with limited generalization capabilities and necessitates an accurate and high-quality system model. It can become computationally difficult when employed in real-time EMS, which makes it inappropriate for effective online applications [15].

Data-driven methods for modeling and controlling EVCS infrastructure have drawn more attention in recent times. As a result, researchers are trying to use machine intelligence. Algorithms with an artificial intelligence (AI) foundation are effective methods that offer remarkable benefits in challenging decision-making situations [16]. They can be used in systems that undergo constant modification and uncertainties [17]. The reinforcement learning (RL) method is a generalized machine learning (ML) method in which an agent learns from past actions in the environment without requiring an environment model. It is a powerful tool for decision-making problems [18]. An MDP model of the EV charging problem was addressed in most publications in order to construct an RL model capable of accomplishing a certain goal for the EVCS [19].

In [20] the authors employed the fitted Q-iteration batch RL algorithm in to determine the optimal charging policy. It is possible to scale their proposed MDP formulation to any number of charging stations. Wang et al. in [21] suggest an RL method based on SARSA (State-Action-Reward State-Action) algorithm for charging schedule and pricing optimization that maximizes an EVCS's system objective. The proposed approach is described as model-free, indicating that it does not depend on assumed probabilistic models for uncertain events. The drawback of employing RL lies in the curse of dimensionality, which hinders effective utilization of numerous input features.

A standard solution to the curse of dimensionality in RL is the adoption of DRL approaches. Deep neural networks are used in DRL for identifying patterns in high-dimensional state spaces and to approximate complex functions. In [22], Wan et al. implemented a model-free methodology utilizing DRL to establish the most effective strategy for EV charging. They employed a Q network to approximate the optimal action-value function. To minimize charging costs and maximize customer satisfaction with charging requirements, the authors in [23] employed a deep deterministic policy gradient (DDPG) algorithm for EV charging management. In [24], the DDPG algorithm is utilized to maximize the profit of a distribution system operator trying to compute an optimal EV charging strategy in a distribution network.

The focus of this study is optimizing the energy management of EVCS by incorporating multiple objectives and considering relevant constraints. The approach employs DRL techniques to assess the impact of uncertainties on charging strategies, conducting thorough simulations based on real-world data. DRL combines the decision-making process of RL with the nonlinear perceptual capabilities of deep neural networks [25]. It makes the technique well-suited for situations involving continuous action domains, showcasing superior performance compared to RL algorithms and deterministic optimization methods, particularly in uncertain environments.

The model-free technique in DRL is essential for the adaptive decision-making process employed by agents to maximize energy-saving measures for EVCS. The model-free method, in contrast to model-based approaches, allows agents to learn optimal policies directly from their interactions with the environment, without prior knowledge of its underlying dynamics [22]. The agent essentially uses trial and error to determine the optimal course of action as it investigates its surroundings by taking actions and receiving feedback in the form of rewards or penalties. This strategy is especially beneficial in the context of EVCS, where the environment is characterized by dynamic and uncertain factors such as fluctuating energy prices, varying grid power and EV charging demands.

This work aims to develop an EMS for EVCS using DRL. The problem is framed as an MDP from the perspective of the EVCS operator. The main objective is to discover cost-effective charging schedules that minimize the EVCS operator charging cost while meeting the energy demands of EV users. The proposed approach employs a model-free method leveraging DRL. Unlike traditional model-based techniques that require a forecast model to handle uncertainties and optimize energy management, this approach doesn't rely on knowledge of the system model.

The main contributions of this study are as follows.

1) A MDP is constructed from the perspective of the EVCS operator. The real-world scenarios are formulated by considering the stochastic nature of renewable energy and the commuting behavior of EVs.
2) A DRL-based model-free technique that doesn't require any information about system model is proposed to determine an optimal strategy for the effective utilization of EVCS resources.
3) Various DRL algorithms for solving the MDP are studied and the performances are empirically compared. The DRL algorithms considered are Soft Actor-Critic (SAC), Proximal Policy Optimization (PPO), Twin Delayed DDPG (TD3), Deep Deterministic Policy Gradients (DDPG) and TQC. A comparative analysis of average rewards among the DRL algorithms shows the superior performance of TQC algorithm for solving the problem.
4) The simulation results demonstrate that a well-trained agent can offer an improved control policy, efficiently shifting EV charging load to time slots with lower electricity prices. Maximum utilization of PV power is achieved and the cost of the EVCS operator is minimized. This accomplishment is realized without relying on any energy storage or backup power generation systems.

This paper is structured as follows: Section II provides an overview of the fundamental structure of the EVCS, section III discusses DRL, Section IV briefly explains the simulation results and analysis and Section V concludes the work.

II. SYSTEM DESCRIPTION

In this section the structure of EVCS considered for this work is demonstrated. An overview of the configuration of the proposed EVCS is shown in Fig.1.

A. SYSTEM CONFIGURATION

The EVCS is equipped with photovoltaic (PV) panels for renewable energy generation and is also connected to the main electrical grid power. However, the power drawn from the main grid is limited due to transformer constraints. By combining the EV chargers, solar panels and the main power grid into a single direct current (DC) system, the EVCS can reduce cost on infrastructure and convert energy more efficiently. Bidirectional AC/DC power converters guarantee a steady DC power connection, which transmits power between the EVCS and the main grid. To meet the energy requirements of EV charging, grid power and PV power are used. The PV generation system comprises several PV modules, each with a boost converter directly connected to the shared DC power line. It is expected that the PV generation system continually runs at the maximum power point to achieve optimal utilisation. When solar power from the PV system is available, it is primarily used to fulfill the charging requirements of EVs. Any surplus PV power is sent back to the main grid. However, if the power generated by the PV system is insufficient to meet the EV charging demand, the connected EVs are charged using power from the main grid [26].

The operation of an EVCS is examined across a time frame divided into T time slots and the time index $t \in T = \{1, 2, \ldots, T\}$. When there is low or insufficient PV power, the total available charging power for EVs needs adjustment. At the start of each time slot, the EV charging schedule based on past and present factors are performed. These factors include the charging demand, departure times of already arrived EVs and electricity prices. The decisions made by the EMS impact the remaining charging demands for future time slots. The number of charging ports available in the charging station is given by N_p. It is assumed that the arrival of N_{ev} EVs are in sequential order, denoted from 1 to N_{ev} with an index $j \in N_{ev} = 1, 2, \ldots, N_{ev}$. The arrival time and departure time of EV_j are designated as t_{aj} and t_{dj}, respectively. The charging demand of EV_j is represented by D_j. In offline models, the charging station possesses advance knowledge of EV profiles, including D_j, t_{aj}, and t_{dj}, which is

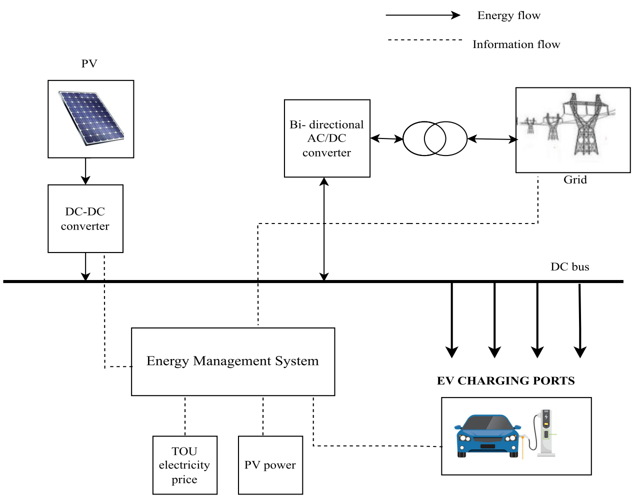

FIGURE 1. System structure of EVCS.

impractical. But here the focus is on an online model where the charging station lacks foresight into future information regarding EV profiles.

B. PROBLEM STATEMENT

The main objective of the EMS is to minimise the charging cost of EVCS operator by minimizing power absorbed from the grid, while still meeting the charging demand of all connected EVs before their departure. The grid balance equation is given by (1). Let $P_{\text{pv}}(t)$ as the PV power available at time t, $P_{\text{g}}(t)$ total electricity drawn from the grid and $P_{\text{ev}}(t)$ the sum of all connected EV's charging requirements.

$$P_{\text{pv}}(t) + P_{\text{g}}(t) - P_{\text{ev}}(t) = 0 \tag{1}$$

A penalty fee, denoted as C_p, is granted to EV owners if the charging requirements are not fulfilled before their scheduled departure time. This fee applies to the portion of the charging demand that isn't satisfied. The EVCS operator's total cost of operation is given in (2). Excess PV power sold to the grid and electricity purchased from the grid are denoted by $P_{g,s}(t)$ and $P_g(t)$ respectively. The prices for electricity purchased from and sold to the grid are $C_g(t)$ and C_s respectively. P_{euj} is the unsatisfied charging demand of j^{th} EV.

$$\mathrm{C_T} = \sum_{t}^{T} \left(P_g(t)C_g(t) - P_{g,s}(t)C_s \right) + \sum_{j}^{N_{ev}} P_{euj}C_p \tag{2}$$

III. DEEP REINFORCEMENT LEARNING

Traditional model-based optimization algorithms face difficulties in effectively managing the challenges posed by the uncertainty in the energy demand of EVs and the variability in PV power. At its core, optimizing the operation of an EVCS is essentially a stochastic sequential decision problem. RL emerges as a class of efficient algorithms designed to address such sequential decision problems by leveraging MDP [27]. MDP provides a robust framework for modeling situations where an agent interacts sequentially with an environment to maximize a long-term reward. In reinforcement learning, the main goal is to discover an optimal policy that maximizes the expected cumulative reward over time, enabling efficient decision-making.

A. MARKOV DECISION PROCESS FORMULATION

MDP has five-tuple structures, denoted as (S, A, P, R, γ) which capture the essence of sequential decision-making under uncertainty. S denotes a finite set of states the agent can occupy. The set of available actions is given by A, the state transition probabilities P under each action, a reward function R defining immediate rewards for state transitions and a discount factor γ weighing future rewards. During each time step, the agent which is situated in a particular state, selects an action based on a policy and the environment responds by transitioning to a new state according to a probabilistic transition model. The agent then receives a reward associated with the chosen state-action pair. A pivotal assumption in MDPs is the Markov property, stipulating that the future state is solely dependent on the current state and action, rendering the entire past history irrelevant.

Environment: The environment denotes the system employed in the training algorithm. In this study the environment refers to the EVCS which is shown in Fig.1.

Agent: The agent serves as a substitute for the system operator. Following training, a well trained agent can offer an optimal or close-to-optimal energy management strategy for the EVC, based on the real-time information.

State space: The state space should reflect the total electricity demand of the charging piles, available PV power at the current time slot and predicted PV power during further time slots, Time of Use (TOU) price of electricity from the grid. The EVCS state at each time is defined as

$$\begin{aligned} S_t = \{t, TOU, P_{PV}, P_{PV}(t+1), P_{PV}(t+2), \\ P_{PV}(t+3), E_d(t), EV^k_{SOC}, t^k_d\} \end{aligned} \tag{3}$$

where t is the information regarding the current time slot, TOU is the time of use price of electricity, P_{PV} is the PV power available at the time slot t, $P_{PV}(t+1), P_{PV}(t+2), P_{PV}(t+3)$ are the predicted PV power during the next 3 time slots, $E_d(t)$ is the total EV charging demand at the current time slot for all N_p charging ports available at the charging station, EV^k_{SOC} and t^k_d are the State Of Charge(SOC) and departure time of the EV connected to the k^{th} charging port.

Action space: It represents all valid actions for a given environment in each time slot. It includes a_g and a_k. where a_g represent the grid power purchased from the grid and a_k represent the charging rate of k^{th} charging port. The charging action of EV batteries is constrained by the maximum power that can be transmitted through the EV battery socket.

Reward: In an MDP, the agent executes an action that transitions from the current state to the next state and calculates the reward associated with that transition. The

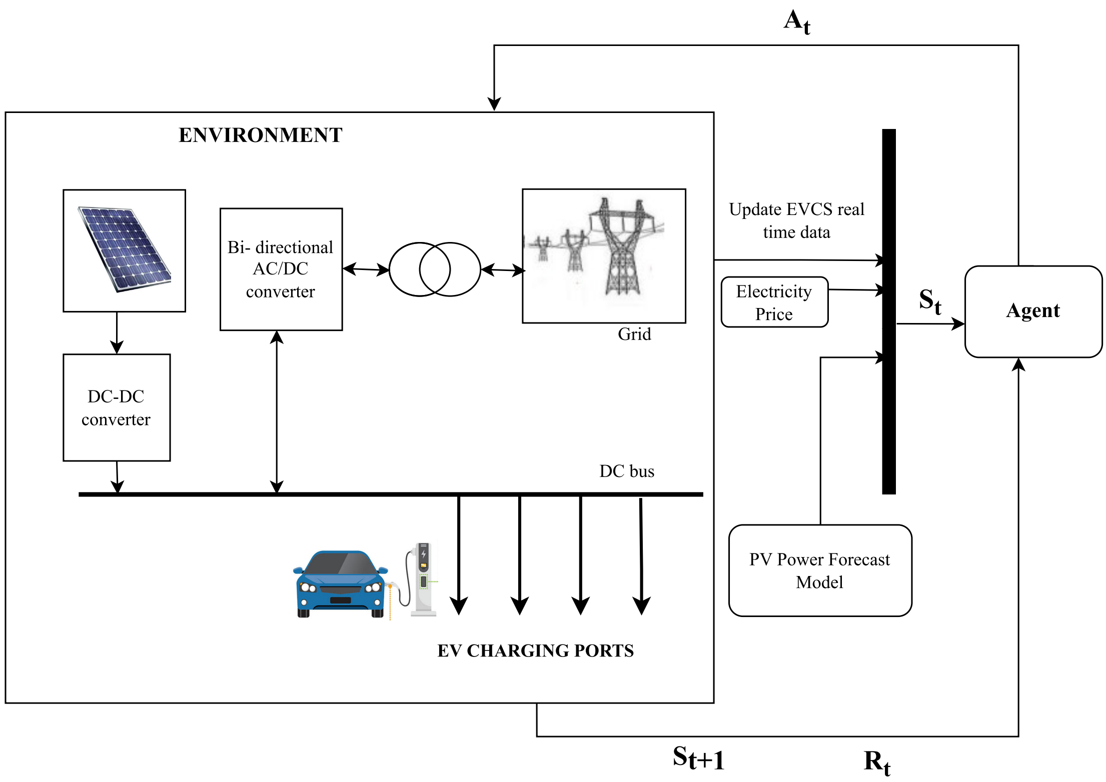

FIGURE 2. EMS using DRL.

goal of the DRL algorithm is to optimize the cumulative reward, in alignment with the objective function given in (2). As a result, the algorithm incorporates the current decision period's objective function value as an immediate reward for the agent. The aim is to minimise the operating cost of the EVCS operator which is given by (2). The reward function can be defined as (4)

$$r_t = -C_T \tag{4}$$

In the t^{th} time step, considering the state S_t obtained through interaction with the environment, the agent issues an action to the environment. Subsequently, the agent receives an immediate reward. The environment then transitions to the next state S_{t+1}. However, due to the environment's inherent randomness, repeating the same action may not yield the same reward. Therefore, the introduction of a discount factor(γ) accounts for the environment's unpredictability. This approach converts the energy management of the EVCS into the task of training an agent to acquire the optimal control policy aiming to maximize the discounted cumulative reward.

B. TRUNCATED QUANTILE CRITICS(TQC)

Policy iteration and value iteration are classical algorithms designed for solving MDP. These methods necessitate an exhaustive enumeration of the entire state space, making them impractical for realistically sized problems due to computational constraints. To address this limitation, this study introduces a DRL algorithm, specifically TQC. Unlike traditional methods, TQC does not require explicit enumeration of the complete state space. Instead, it utilizes neural networks to generalize and approximate the optimal policy based on experiences. This approach proves advantageous in handling large-scale problems where explicit enumeration becomes computationally. This enhances the suitability of DRL methods for real-world applications [28].

The design of TQC incorporates three concepts: distributional representation of a critic, truncation of approximated distribution and ensembling [29]. The foundational idea in TQC is to use a distributional representation for the critic, as opposed to a single scalar value as in traditional Q-learning. This means that instead of estimating just the expected return, TQC aims to model the entire distribution of returns. This is achieved by training multiple Q-networks, each representing different quantiles of the return distribution. The ensemble of these networks provides a richer understanding of the uncertainty and variability associated with different actions in a given state.

The concept of truncation involves limiting the number of quantiles considered during the training process. Not all quantiles need to be explicitly represented. By truncating the distribution, the algorithm focuses on learning the most relevant and informative parts of the distribution. This can help reduce computational complexity and improve the stability of the learning process. Truncation is a key mechanism in TQC for managing the complexity of the distributional representation. Ensembling involves training and combining multiple neural networks, each with different initializations or architectures. In TQC, ensembling is used to capture diverse perspectives on the value distribution. Each network in the ensemble represents a different quantile of the distribution and by combining their outputs. TQC gains the benefit of both exploring various aspects of uncertainty and mitigating the impact of neural network approximations. Ensembling contributes to the robustness and generalization capabilities of the TQC algorithm [30]. The pseudo-code outlining the methodology is presented in Algorithm 1.

Algorithm 1 TQC Algorithm

Initialize Q-networks Q_{θ_i} for $i = 1, \ldots, N$ and policy network π_ϕ with parameters θ_i, ϕ
Initialize target networks $Q'_{\theta'_i} \leftarrow Q_{\theta_i}$ for $i = 1, \ldots, N$
Initialize replay buffer $\mathcal{D}$ with capacity B
Initialize temperature parameter α, entropy target H, and other hyperparameters
for each episode **do**
    Observe initial state s
    for each time step **do**
        Sample action a_t from the policy $\pi_\phi(a_t|s_t)$ with exploration noise
        Execute a_t, observe next state s_{t+1} and reward r_t
        Store transition (s_t, a_t, r_t, s_{t+1}) in replay buffer $\mathcal{D}$
        if training step and buffer size is sufficient **then**
            for $i = 1, \ldots, N$ **do**
                Sample batch of transitions from $\mathcal{D}$
                Compute target values using truncated mixture of quantiles
                Update Q_{θ_i} by minimizing the quantile regression loss
            end for
            Sample batch of transitions from $\mathcal{D}$ for actor update
            Update policy π_ϕ by maximizing the entropy-regularized objective
            Update target networks $Q'_{\theta'_i} \leftarrow \tau Q_{\theta_i} + (1-\tau) Q'_{\theta'_i}$ for $i = 1, \ldots, N$
        end if
    end for
end for

IV. RESULTS AND DISCUSSION

Several case studies based on the real world data are conducted to demonstrate the feasibility and effectiveness of the proposed algorithm.

A. INPUT DATA AND PARAMETER SETTINGS

The power grid has a maximum capacity restriction of 60 kW due to transformer limitations. The assigned upper limit for the PV system's output power is 50 kW. All EVs are assumed to have identical specifications, with battery ratings set at 40 kWh. It is assumed that the charging station is equipped with 10 charging ports. PV power prediction is very important for the effective implementation of the algorithm. The utilization of machine learning models for forecasting solar power based on weather features is crucial for addressing uncertainties arising from the time-dependent and nonlinear parameters of PV systems [31]. To ensure the reliable integration of PV power plants in EVCS, accurate solar power prediction measures are essential. An effective approach involves employing the eXtreme Gradient Boosting (XGBoost) algorithm, a gradient boosting machine introduced by Chen and Guestrin [32]. This algorithm, known for its dynamic selection properties, achieves high prediction accuracy with an efficient trade-off between accuracy and complexity. The dataset utilized for the PV power forecast model is sourced from reference [33].

TABLE 1. TOU price of electricity.

Time of use	Price of electricity
9 pm to 8 am	\$0.12/kWh
8 am to 4 pm	\$0.10/kWh
4 pm to 9 pm	\$0.23/kWh

The gradient boosting algorithm, implemented in an iterative manner, combines weak "learners" to create a robust single learner. The process is easily explained in the context of least-squares regression, where the objective is to train a model F to predict values (5) by minimizing the mean squared error.

$$\hat{y} = F(x) \tag{5}$$

$$\min \frac{1}{n} \sum_i (\hat{y_i} - y_i)^2 \tag{6}$$

where $\hat{y}_i$ is the predicted value of $F(x)$ which is compared to the observed value y_i and i indexes over training set of size n. Here,y_i corresponds to the observed PV power and x represents the available weather data in the short-term PV power prediction model. Further information can be found in reference [34]. The ACN dataset, which can be accessible from [35] via a web portal or a python application programming interface, contains the data on EV charging. We have used a scaled version of TOU electricity pricing scheme in [36] which is given in Table 1.

B. TRAINING PROCESS

The DRL framework implementation comprises two primary components: offline training and online application processes. The training phase is crucial for the neural network to acquire sequential decision making skills from data generated through interactions with the environment. Each training episode represents a simulated day, featuring varying solar energy production, TOU pricing profiles and EV schedules and demands. The DRL algorithms employed undergo training using diverse sets of data. The online

TABLE 2. Hyperparameter configuration.

Parameters	Value
learning rate	0.0003
number of epochs	50
number of iterations per epoch	20000
buffer size	1000000
batch size	256
Discount factor	0.99

TABLE 3. Actor and critic network architectures.

Component	Architecture
Actor	
	Input: State vector
	$\rightarrow$ MLP with ReLU activation
	$\rightarrow$ Hidden Layer 1: 256 units
	$\rightarrow$ Hidden Layer 2: 256 units
	$\rightarrow$ Output Layer: $2 \times$ action_dim (mean and log std)
	Output: Action vector (tanh-squashed)
Critic	
	Input: Concatenation of State and Action vectors
	$\rightarrow$ MLP with ReLU activation
	$\rightarrow$ Hidden Layer 1: 512 units
	$\rightarrow$ Hidden Layer 2: 512 units
	$\rightarrow$ Hidden Layer 3: 512 units
	$\rightarrow$ Output Layer: *n_quantiles*
	Output: Quantiles of the value distribution

application flow of DRL algorithm is shown in Fig.2. The hyperparameter configurations for the TQC algorithm are outlined in Table 2.

In Table 3, we present the architectural details of the actor and critic components in our model. The actor network takes the state vector as input and consists of a multi-layer perceptron (MLP) with Rectified Linear Unit (ReLU) activation. The architecture includes two hidden layers with 256 units each, followed by an output layer with $2 \times$ action_dim units representing the mean and log standard deviation of the action vector. The output is transformed using the tanh function to ensure values are within the range $[-1, 1]$.

On the other hand, the critic network takes the concatenation of state and action vectors as input. It also employs an MLP with ReLU activation, featuring three hidden layers with 512 units each. The output layer produces a distribution of values with *n_quantiles* quantiles, providing a comprehensive representation of the value function.

C. TRAINING RESULTS

The training phase plays a crucial role in determining the effectiveness and robustness of the learned policy in achieving optimal charging strategies for EVCS. Various training metrics are analyzed to provide insights into the learning process and the quality of the learned policy [30]. These metrics have been carefully selected to evaluate key aspects of the method's performance, including reward maximization, policy adaptability, value estimation accuracy, and exploration-exploitation trade-off. By examining these training results, a comprehensive understanding of the TQC method's capability to learn and execute effective energy management strategies is sought, ultimately contributing to the advancement of sustainable and efficient electric vehicle charging infrastructure. The training results for the TQC algorithm are presented in Table 4. The table summarizes the performance metrics of the algorithm across different episodes during the training process.

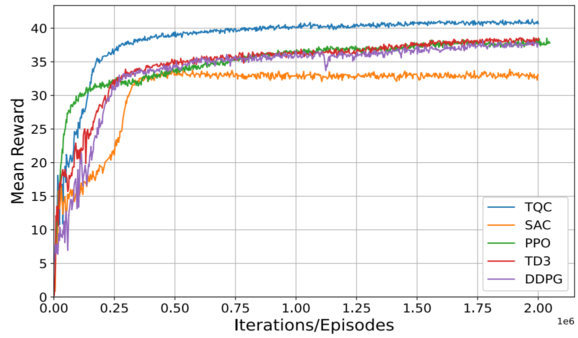

FIGURE 3. Comparison of Mean Episodic rewards.

Throughout the training, the average reward steadily increases, indicating that the algorithm is effectively learning and improving its performance over time. The actor loss, which measures the adaptability of the policy network, decreases consistently, suggesting that the policy network is becoming more proficient in selecting actions that lead to higher rewards. Similarly, the critic loss, representing the quality of learned state-action values, demonstrates a decreasing trend, indicating improvements in value estimation.

The entropy coefficient and entropy loss also exhibit consistent behavior during training. The entropy coefficient, which reflects the level of exploration in the policy, decreases gradually, suggesting a shift towards exploiting known strategies. Conversely, the entropy loss, which measures the adjustment of the entropy regularization term, fluctuates slightly but remains within acceptable bounds. Notably, the algorithm achieves a peak average reward of 40.93 at the 50th episode, indicating its effectiveness in maximizing rewards in the given environment. The actor loss and critic loss reach their respective minima at this point, indicating optimal performance of the policy and value networks. Overall, the training results demonstrate the capability of the TQC algorithm to effectively learn and adapt to the task environment, ultimately achieving high rewards and stable performance.

D. COMPARISON WITH BENCHMARK ALGORITHMS

In this section a comparative analysis of the performance of various DRL algorithms and TQC is presented. Fig. 3 provides a comparative analysis of average rewards among other DRL algorithms like SAC [37], PPO [38], TD3 [39] and DDPG [40]. A higher average reward for TQC compared to other algorithms would indicate superior performance.

Fig. 4 shows a comparison of mean episodic rewards during the last 50,000 steps of training. This figure provides

TABLE 4. Training results for TQC algorithm.

Episode	Step	Average Reward	Actor Loss	Critic Loss	Entropy	Entropy Loss
1	1.82e+03	2.782761	-17.414577	1.365271	0.584816	-8.422409
6	1.64e+05	33.835642	16.337158	0.332693	0.006751	0.664340
11	3.52e+05	38.382265	12.435920	0.074012	0.003440	-0.044572
16	5.23e+05	39.128622	11.628864	0.100487	0.003365	1.617796
21	7.18e+05	39.706541	11.326030	0.053088	0.003311	1.289257
26	8.85e+05	39.860409	10.799242	0.039020	0.003196	-4.312737
31	1.06e+06	40.561131	11.112980	0.040024	0.003155	0.647313
36	1.24e+06	40.606504	11.093746	0.036517	0.003001	-1.364750
41	1.53e+06	40.695110	10.300406	0.049312	0.002940	3.916893
46	1.77e+06	40.789433	10.625885	0.034926	0.002918	2.535536
50	2.00e+06	40.929733	10.125869	0.036132	0.002743	-2.728273

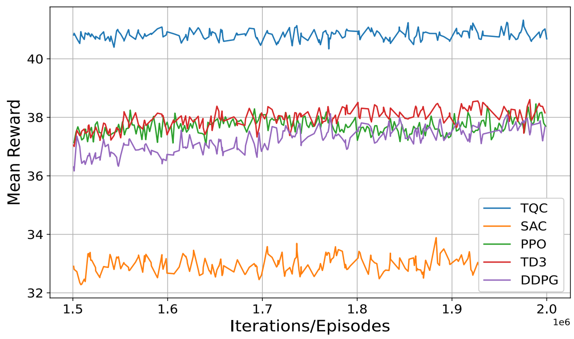

FIGURE 4. Comparison of Mean Episodic rewards (last 50k steps).

TABLE 5. Training results for different algorithms.

Algorithm	Mean Value	Median Value	Standard Deviation
TQC	38.30	40.14	5.54
SAC	30.59	32.87	5.81
PPO	35.06	36.45	3.90
TD3	34.58	36.21	5.37
DDPG	33.63	35.76	6.43

insights into the performance of algorithms in the later stages of training, capturing any convergence or divergence trends.

Table 5 presents the training results for different DRL algorithms. In this set of training results, the TQC algorithm exhibits the highest mean and median values, indicating superior overall performance compared to other algorithms considered. TQC also has a relatively low standard deviation, suggesting consistent performance. Conversely, SAC performs the poorest among the algorithms, with the lowest mean and median values and a higher standard deviation, reflecting less stable learning.

The percentage increases highlight the relative performance differences between the algorithms. TQC outperforms SAC, PPO, TD3, and DDPG, with percentage increases ranging from 8.45% to 20.12%. Among the other algorithms, PPO demonstrates a balanced performance, showing percentage increases against SAC, TD3, and DDPG.

The TQC algorithm demonstrates superior performance compared to other algorithms, as evidenced by its higher mean and median values of 38.30 and 40.14, respectively, with a relatively lower standard deviation of 5.54. This indicates that TQC consistently achieves higher rewards across training episodes while exhibiting less variability in performance. The effectiveness of TQC can be attributed to its innovative approach of utilizing quantile regression to predict a distribution for the value function, rather than solely focusing on mean values. By capturing the entire distribution of returns, TQC gains a richer understanding of the uncertainty and variability associated with different actions, enabling more robust and stable learning. Additionally, the truncation mechanism in TQC focuses on learning the most relevant parts of the distribution, enhancing computational efficiency and improving the stability of the learning process. Overall, TQC's combination of distributional representation, truncation, and ensemble learning contributes to its superior performance in optimizing the operation of EVCS compared to other state-of-the-art algorithms such as SAC, PPO, TD3, and DDPG.

The experiments were carried out on a system running Ubuntu 22.04.3 LTS with Python version 3.11.5. The hardware included an AMD Ryzen 5 5600H processor with 12 cores and 6 threads per core, along with an NVIDIA GeForce RTX 3050 Mobile GPU and 16.0 GB of memory.

E. CASE STUDIES AND PERFORMANCE ANALYSIS

To demonstrate the effectiveness of the proposed EMS, various case studies and numerical simulations are performed on the EVCS. This subsection uses EVCS shown in Fig.1 as an example for simulation study. The simulations were performed over a period of one day, with each day being divided into 48 intervals, each lasting 30 minutes. The charging service fee is $\alpha_C = 0.07\$/kWh$ and the PV power selling price is taken as constant and the value is given by $C_s = 0.09\$/kWh$. The penalty price for unsatisfied EV charging demand is taken as $C_p = 0.04\$/kWh$.

A widely acknowledged and commonly employed charging methodology known as First-Come First-Served (FCFS) is considered as a benchmark method. FCFS charging, also referred to as immediate charging, is a prevalent approach in EVCSs. This strategy entails initiating the charging process for an EV immediately upon its arrival, and the process persists until the battery attains full charge. In the FCFS charging model, the EV is charged promptly upon arrival at

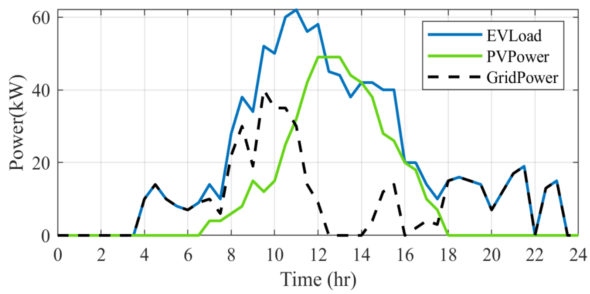

FIGURE 5. Power flow in one day.

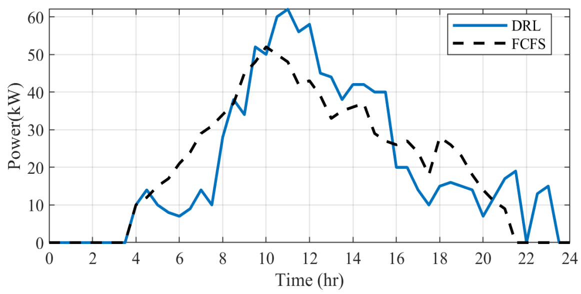

FIGURE 6. Charging load profile of EVCS.

the EVCS, disregarding the availability of PV power or the prevailing electricity price.

Various test cases are simulated to study the effectiveness of the proposed EMS.

Case 1: In this case study EV charging demand is met through a combination of the grid power and PV power. Dynamic scheduling outcomes from a randomly chosen day in the test set are illustrated in Fig.5. The EV charging profile, PV power and power received from the grid for the EV charging purpose are depicted. The proposed EMS aims to maximize the utilization of available PV power. If the grid power is utilized, the EMS tries to shift the EV load to time periods where the energy price is low. From hours 16:00 to 21:00, the energy price reaches its peak. Upon examining the results, it is evident that the EV load is minimal during these periods of high energy prices. Eventhough after hour 18.00 when PV power is unavailable and the grid power becomes the sole available source.

The comparison of EV charging profiles for the proposed EMS employing DRL with TQC algorithm and FCFS strategy is shown in Fig.6. When the TOU price of electricity is higher, that is from hours 16.00 to 21.00 FCFS consumes more power compared to the proposed EMS. In contrast, in the proposed EMS the EV charging is systematically scheduled within specific time slots. This scheduling takes into account periods when there is a surplus of PV power or when electricity prices are lower. Unlike the FCFS charging approach, the scheduling in EMS with DRL ensures that EV charging aligns with optimal conditions, optimizing the utilization of available PV power and minimizing electricity costs.

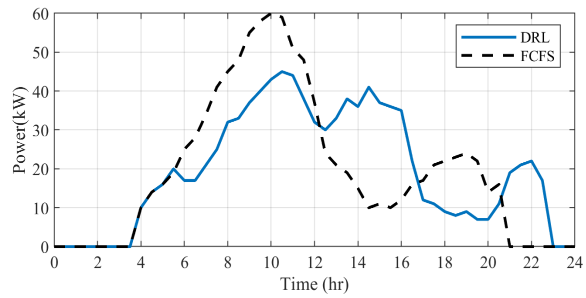

FIGURE 7. Charging load profile of EVCS when there is no PV power.

TABLE 6. Comparison of cost of EVCS based on the control strategy.

Control strategy	Cost in case 1	Cost in case 2
DRL	88 $	106 $
FCFS	115 $	130 $

Case 2: The EV power demand is met by power from the grid alone, that is during cloudy or winter days, when sufficient PV power is not available. Both FCFS and DRL charging schemes has initially same charging power. The proposed EMS employing DRL with TQC algorithm can effectively transfer EV charging load to those periods were TOU electricity price is lower. It could maximizes charging power at each time step without sacrificing profit, allowing for greater flexibility to reduce charging power during peak hours. From the Fig.7 we can infer that in the EMS using DRL, the grid power is not exceeding its maximum limit of 60kW at any of the time slots. But in FCFS strategy, the grid power is reaching its maximum and more power is consumed during peak demand periods from 16.00 to 21.00. This considerably increases the charging cost of EVCS operator.

The benefits of the suggested scheme are also demonstrated by cost comparisons of various charging schemes. The daily cost comparison of the two methods for both cases is given in Table 6. It can be seen that the EMS using DRL has almost 23% cost reduction compared to FCFS in case 1 and almost 18% reduction in case 2.

V. CONCLUSION

In this work, an EMS for PV powered EVCS is developed. The EMS minimizes the operating cost of the EVCS operator while meeting the energy demands of connected EVs. A MDP model is constructed from the perspective of the EVCS operator. The real-world scenarios are formulated by taking into account the stochastic nature of renewable energy and the commuting behavior of EVs. Unlike traditional model-based techniques that require a forecast model to handle uncertainties, this approach doesn't rely on the system model. Five different DRL algorithms are implemented for addressing MDP and their performances are empirically evaluated and compared. Out of the selected algorithms, TQC shows better performance. Various case studies are simulated to show the effectiveness of the proposed EMS. The simulation results demonstrate that a well-trained agent can provide a better control policy and reduce charging cost of EVCS operator in comparison with other benchmark method. EMS with DRL ensures that EV charging aligns with optimal conditions, maximizing the utilization of available PV power and minimizing electricity costs. This is accomplished without relying on any energy storage or backup power generation systems.

REFERENCES

[1] S. I. Jeong and D.-H. Choi, "Electric vehicle user data-induced cyber attack on electric vehicle charging station," *IEEE Access*, vol. 10, pp. 55856–55867, 2022, doi: 10.1109/ACCESS.2022.3177842.

[2] F. Mohammad, D.-K. Kang, M. A. Ahmed, and Y.-C. Kim, "Energy demand load forecasting for electric vehicle charging stations network based on ConvLSTM and BiConvLSTM architectures," *IEEE Access*, vol. 11, pp. 67350–67369, 2023, doi: 10.1109/ACCESS.2023.3274657.

[3] A. S. Al-Ogaili, T. J. Tengku Hashim, N. A. Rahmat, A. K. Ramasamy, M. B. Marsadek, M. Faisal, and M. A. Hannan, "Review on scheduling, clustering, and forecasting strategies for controlling electric vehicle charging: Challenges and recommendations," *IEEE Access*, vol. 7, pp. 128353–128371, 2019, doi: 10.1109/ACCESS.2019.2939595.

[4] A. K. Karmaker, Md. A. Hossain, H. R. Pota, A. Onen, and J. Jung, "Energy management system for hybrid renewable energy-based electric vehicle charging station," *IEEE Access*, vol. 11, pp. 27793–27805, 2023, doi: 10.1109/ACCESS.2023.3259232.

[5] A. A. S. Mohamed, M. Jun, R. Mahmud, P. Mishra, S. N. Patel, I. Tolbert, S. Santhanagopalan, and A. Meintz, "Hierarchical control of megawatt-scale charging stations for electric trucks with distributed energy resources," *IEEE Trans. Transport. Electrific.*, vol. 9, no. 4, pp. 4951–4963, Dec. 2023, doi: 10.1109/TTE.2022.3167647.

[6] M. L. Kolhe and M. J. M. A. Rasul, "3-phase grid-connected building integrated photovoltaic system with reactive power control capability," *Renew. Energy*, vol. 154, pp. 1065–1075, Jul. 2020.

[7] J. Singh and R. Tiwari, "Impact analysis of different charging models for optimal integration of plug-in electric vehicles in distribution system," *J. Eng.*, vol. 2019, no. 18, pp. 4728–4733, Jul. 2019.

[8] A. R. Bhatti and Z. Salam, "A rule-based energy management scheme for uninterrupted electric vehicles charging at constant price using photovoltaic-grid system," *Renew. Energy*, vol. 125, pp. 384–400, Sep. 2018.

[9] C. Balasundar, C. K. Sundarabalan, N. S. Srinath, J. Sharma, and J. M. Guerrero, "Interval type2 fuzzy logic-based power sharing strategy for hybrid energy storage system in solar powered charging station," *IEEE Trans. Veh. Technol.*, vol. 70, no. 12, pp. 12450–12461, Dec. 2021, doi: 10.1109/TVT.2021.3122251.

[10] L. Zeng, T. Krallmann, A. Fiege, M. Stess, T. Graen, and M. Nolting, "Optimization of future charging infrastructure for commercial electric vehicles using a multi-objective genetic algorithm and real travel data," *Evolving Syst.*, vol. 11, no. 2, pp. 241–254, Jun. 2020.

[11] *Energy Management System Application Program Interface (EMS-API)*, Standard IEC 61970, Int. Electrotechnical Commission (IEC), Geneva, Switzerland, 2004.

[12] D. Yan, H. Yin, T. Li, and C. Ma, "A two-stage scheme for both power allocation and EV charging coordination in a grid-tied PV-battery charging station," *IEEE Trans. Ind. Informat.*, vol. 17, no. 10, pp. 6994–7004, Oct. 2021, doi: 10.1109/TII.2021.3054417.

[13] E. González-Rivera, P. García-Triviño, R. Sarrias-Mena, J. P. Torreglosa, F. Jurado, and L. M. Fernández-Ramírez, "Model predictive control-based optimized operation of a hybrid charging station for electric vehicles," *IEEE Access*, vol. 9, pp. 115766–115776, 2021.

[14] F. Zhou, Y. Li, W. Wang, and C. Pan, "Integrated energy management of a smart community with electric vehicle charging using scenario based stochastic model predictive control," *Energy Buildings*, vol. 260, Apr. 2022, Art. no. 111916.

[15] B. Karg and S. Lucia, "Efficient representation and approximation of model predictive control laws via deep learning," *IEEE Trans. Cybern.*, vol. 50, no. 9, pp. 3866–3878, Sep. 2020.

[16] M. Ganeshkumar, V. Ravi, V. Sowmya, E. A. Gopalakrishnan, and K. P. Soman, "Explainable deep learning-based approach for multilabel classification of electrocardiogram," *IEEE Trans. Eng. Manag.*, vol. 70, no. 8, pp. 2787–2799, Aug. 2023, doi: 10.1109/TEM.2021.3104751.

[17] M. Shin, D.-H. Choi, and J. Kim, "Cooperative management for PV/ESS-enabled electric vehicle charging stations: A multiagent deep reinforcement learning approach," *IEEE Trans. Ind. Informat.*, vol. 16, no. 5, pp. 3493–3503, May 2020, doi: 10.1109/TII.2019.2944183.

[18] A. Hafeez, R. Alammari, and A. Iqbal, "Utilization of EV charging station in demand side management using deep learning method," *IEEE Access*, vol. 11, pp. 8747–8760, 2023, doi: 10.1109/ACCESS.2023.3238667.

[19] S. Mateen, A. Haque, V. S. B. Kurukuru, and M. A. Khan, "Discrete stochastic control for energy management with photovoltaic electric vehicle charging station," *CPSS Trans. Power Electron. Appl.*, vol. 7, no. 2, pp. 216–225, Jun. 2022.

[20] N. Sadeghianpourhamami, J. Deleu, and C. Develder, "Definition and evaluation of model-free coordination of electrical vehicle charging with reinforcement learning," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 203–214, Jan. 2020, doi: 10.1109/TSG.2019.2920320.

[21] S. Wang, S. Bi, and Y. A. Zhang, "Reinforcement learning for real-time pricing and scheduling control in EV charging stations," *IEEE Trans. Ind. Informat.*, vol. 17, no. 2, pp. 849–859, Feb. 2021, doi: 10.1109/TII.2019.2950809.

[22] Z. Wan, H. Li, H. He, and D. Prokhorov, "Model-free real-time EV charging scheduling based on deep reinforcement learning," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 5246–5257, Sep. 2019, doi: 10.1109/TSG.2018.2879572.

[23] F. Zhang, Q. Yang, and D. An, "CDDPG: A deep-reinforcement-learning-based approach for electric vehicle charging control," *IEEE Internet Things J.*, vol. 8, no. 5, pp. 3075–3087, Mar. 2021, doi: 10.1109/JIOT.2020.3015204.

[24] T. Ding, Z. Zeng, J. Bai, B. Qin, Y. Yang, and M. Shahidehpour, "Optimal electric vehicle charging strategy with Markov decision process and reinforcement learning technique," *IEEE Trans. Ind. Appl.*, vol. 56, no. 5, pp. 5811–5823, Sep. 2020, doi: 10.1109/TIA.2020.2990096.

[25] R. Satheesh, N. Chakkungal, S. Rajan, M. Madhavan, and H. H. Alhelou, "Identification of oscillatory modes in power system using deep learning approach," *IEEE Access*, vol. 10, pp. 16556–16565, 2022, doi: 10.1109/ACCESS.2022.3149472.

[26] M. S. Mathew, M. L. Kolhe, S. T. Kandukuri, and C. W. Omlin, "Data driven approach for the management of wind and solar energy integrated electrical distribution network with high penetration of electric vehicles," *J. Cleaner Prod.*, vol. 421, Oct. 2023, Art. no. 138467.

[27] B. Premjith and K. P. Soman, "Deep learning approach for the morphological synthesis in Malayalam and Tamil at the character level," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 20, no. 6, pp. 1–17, Nov. 2021.

[28] N. Dorka, T. Welschehold, J. Bödecker, and W. Burgard, "Adaptively calibrated critic estimates for deep reinforcement learning," *IEEE Robot. Autom. Lett.*, vol. 8, no. 2, pp. 624–631, Feb. 2023, doi: 10.1109/LRA.2022.3229236.

[29] A. Kuznetsov, P. Shvechikov, A. Grishin, and D. Vetrov, "Controlling overestimation bias with truncated mixture of continuous distributional quantile critics," in *Proc. 37th Int. Conf. Mach. Learn.*, vol. 119, H. D. III and A. Singh, Eds. Jul. 2020, pp. 5556–5566.

[30] K. A. Nugroho, M. Amirul Haq, C.-K. Wang, S.-J. Ruan, M. Polikarpov, D. Wagstyl, and J. Deuse, "Towards smart manufacturing using reinforcement learning in a sparse-rewarded environment for through-hole technology," in *Proc. IEEE 12th Global Conf. Consum. Electron. (GCCE)*, Oct. 2023, pp. 672–673, doi: 10.1109/GCCE59613.2023.10315655.

[31] G. S. A. Rani and P. S. L. Priya, "Robust scheduling of electric vehicle charging for PV integrated parking-lots," in *Proc. Int. Conf. Power, Instrum., Control Comput. (PICC)*, Dec. 2020, pp. 1–6.

[32] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 785–794.

[33] Networks UK Power. *PV Energy Generation Data*. Accessed: Nov. 5, 2023. [Online]. Available: https://data.london.gov.uk/dataset/photovoltaic–pv–solar-panel-energy-generation-data

[34] W. Guo, X. Jiang, and L. Che, "Short-term photovoltaic power-forecasting based on machine learning," in *Proc. IEEE 3rd Int. Electr. Energy Conf. (CIEEC)*, Sep. 2019, pp. 1276–1280.

[35] Z. J. Lee, T. Li, and S. H. Low, "ACN-data: Analysis and applications of an open EV charging dataset," in *Proc. 10th ACM Int. Conf. Future Energy Syst.*, Jun. 2019, pp. 139–149.

[36] Y. Yang, H.-G. Yeh, and R. Nguyen, "A robust model predictive control-based scheduling approach for electric vehicle charging with photovoltaic systems," *IEEE Syst. J.*, vol. 17, no. 1, pp. 111–121, Mar. 2023, doi: 10.1109/JSYST.2022.3183626.

[37] W. Sun, Y. Zou, X. Zhang, N. Guo, B. Zhang, and G. Du, "High robustness energy management strategy of hybrid electric vehicle based on improved soft actor-critic deep reinforcement learning," *Energy*, vol. 258, Nov. 2022, Art. no. 124806, doi: 10.1016/j.energy.2022.124806.

[38] Y. Gu, Y. Cheng, C. L. P. Chen, and X. Wang, "Proximal policy optimization with policy feedback," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 52, no. 7, pp. 4600–4610, Jul. 2022, doi: 10.1109/TSMC.2021.3098451.

[39] Z. E. Liu, Q. Zhou, Y. Li, and S. Shuai, "An intelligent energy management strategy for hybrid vehicle with irrational actions using twin delayed deep deterministic policy gradient," *IFAC-PapersOnLine*, vol. 54, no. 10, pp. 546–551, 2021.

[40] C. Qiu, Y. Hu, Y. Chen, and B. Zeng, "Deep deterministic policy gradient (DDPG)-based energy harvesting wireless communications," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8577–8588, Oct. 2019, doi: 10.1109/JIOT.2019.2921159.

G. S. ASHA RANI received the B.Tech. degree in electrical and electronics engineering from the University of Kerala, India, and the M.Tech. degree from the College of Engineering Trivandrum, Kerala, India, where she is currently pursuing the Ph.D. degree with the Department of Electrical Engineering. She has more than 12 years of teaching and research experience. Her research interests include data-driven control, robust control, machine learning, renewable energy integration, and electric vehicle technology.

P. S. LAL PRIYA (Member, IEEE) received the B.Tech. degree in electrical and electronics engineering and the M.Tech. degree in control systems from the University of Kerala and the Ph.D. degree in control systems from Indian Institute of Technology Bombay (IITB). Since 2001, she has been a dedicated Professor with the EEE Department, College of Engineering Trivandrum (CET). She is currently a Distinguished Professor and the Head of the Department of Electrical Engineering, CET. Her academic journey spans various aspects of control systems. Her research interests include nonlinear dynamical systems, multivariable control, robust control techniques, fractional order systems and control, and machine learning.

JINO JAYAN received the B.Tech. degree in mechanical engineering from the College of Engineering Trivandrum (CET), in 2020, where he is currently pursuing the M.Tech. degree in robotics and automation. His academic journey reflects his commitment to advancing his expertise in cutting-edge technologies. His research interests include robotics, deep reinforcement learning, and other emerging areas at the intersection of technology and automation.

RAHUL SATHEESH (Member, IEEE) received the B.Tech. degree in electrical and electronics engineering from the University of Kerala, India, in 2012, the M.E. degree in power systems from Anna University, Chennai, India, in 2014, and the Ph.D. degree in power systems from the Department of Electrical Engineering, National Institute of Technology Calicut, Kerala, India. He is currently an Assistant Professor with the Amrita School of AI, Amrita Vishwa Vidyapeetham, Coimbatore. He has published several research articles in various journals and attended several conferences. His current research interests include computational and data-driven algorithms, wide-area monitoring systems, power oscillation studies, power grid resilience, smart grids, electric vehicle technology, and AI applications in power systems. He is an Active Member of the IEEE Power and Energy Society. He was honored with the 2019 Young Innovator Award by the Kerala State Government, the 2022 Outstanding Engineer Award from IEEE PES HQ, and the Outstanding Chapter Volunteer Award from IEEE Kerala Section, in 2023.

MOHAN LAL KOLHE is currently a Full Professor of smart grid and renewable energy with the Faculty of Engineering and Science, University of Agder, Norway. He is a leading renewable energy technologist with three decades of academic experience at an international level. He has held various academic positions at prestigious universities, such as University College London, U.K.; University College London, Australia; the University of Dundee, U.K.; the University of Jyvaskyla, Finland; and the Hydrogen Research Institute, Trois-Rivières, QC, Canada. From 2009 to 2011, he was a Board Member of the Government of South Australia's Renewable Energy Board as an Advisor on renewable energy policies. He is an Expert Evaluator of projects for funding at various international research councils, such as European Commission: Erasmus Higher Education International Capacity Building, Royal Society London, U.K., and the Engineering and Physical Sciences Research Council (EPSRC), U.K. He has delivered keynote addresses and expert lectures at many conferences and workshops.