SCIENTIFIC OASIS

Decision Making: Applications in Management and Engineering

Journal homepage: www.dmame-journal.org
ISSN: 2560-6018, eISSN: 2620-0104

Optimizing Vehicle-to-Vehicle (V2V) Charging in Electric Vehicles by Adaptive Q-learning Implication for Smart Tourism

Pitchaya Jamjuntr[1], Chanchai Techawatcharapaikul[2,*] and Pannee Suanpang[3,*]

1,2 Department of Electrical Engineering, Faculty of Engineering, King Mongkut's University of Technology Thonburi, Bangkok, Thailand
3 Department of Information Technology, Faculty of Science & Technology, Suan Dusit University, Bangkok, Thailand.

ARTICLE INFO

Article history:
Received 5 April 2024
Received in revised form 19 October 2024
Accepted 7 December 2024
Available online 29 December 2024

Keywords:
Electric vehicles; Adaptive Q-learning; Vehicle-to-Vehicle charging; Optimization; Smart tourism

ABSTRACT

This study investigates the utilisation of adaptive Q-learning to optimise Vehicle-to-Vehicle (V2V) charging among electric vehicles (EVs) in dynamic smart tourism destinations within developing nations. V2V charging presents a viable solution to extend the range of EVs and improve operational efficiency by enabling direct energy transfer between vehicles. However, refining this process in volatile and high-demand sectors requires complex decision-making to ensure both energy efficiency and system integrity. To address these issues, this research introduces an advanced adaptive Q-learning approach that evaluates the current state and adjusts learning parameters accordingly. A bespoke simulation environment was developed to model a fleet of EVs capable of charging one another, incorporating factors such as energy demand, state of charge, and geographical location. The simulation environment also considers real-world variables, such as the vehicles' state of charge, their spatial positioning, and variable energy demands. The reward function favours an even and efficient energy flow, ensuring compatibility with the specific needs of smart tourism destinations. The simulation results demonstrate that the adaptive Q-learning algorithm significantly outperforms rule-based methods, achieving a 20% increase in energy efficiency, a 25% improvement in the average state of charge (SOC), better transfer efficiency, and enhanced system robustness. These findings underscore the potential of adaptive Q-learning as a scalable and effective solution for intelligent energy management in V2V charging systems. Future research should explore its integration with real-time traffic and vehicle movement patterns to further enhance its applicability in smart tourism ecosystems.

1. Introduction

The transition to EVs is a vital aspect of global efforts to reduce carbon emissions and promote sustainable transportation [1;2]. As traditional internal combustion engine vehicles are major contributors to greenhouse gas emissions, the shift to EVs is a key strategy for mitigating climate change [3]. In the context of smart tourism destinations, where advanced technologies are integrated

* Corresponding author.
E-mail address: pannee_sua@dusit.ac.th, chanchai.tec@kmutt.ac.th;

https://doi.org/10.31181/dmame7220241330

to enhance visitor experiences and promote sustainability, the adoption of EVs becomes particularly crucial. Smart tourism destinations are renowned for their capacity to foster environmentally sustainable practices and employ technological innovations to reduce the ecological footprint of tourism activities [4]. However, the growth of EVs introduces new challenges, particularly in high-tourism areas where energy and infrastructure provision are under pressure. While the EV ecosystem will require traditional charging infrastructure, the existing systems may not meet the immediate energy supply demands, especially in popular or densely visited tourist spots. These challenges have driven research into innovative solutions, such as V2V charging, where vehicles can charge each other through direct power transfer. Given the limited availability of charging stations, V2V charging offers an opportunity to increase charging access by enabling vehicles to connect and share energy. It represents a flexible and scalable solution for enhancing existing charging networks, alleviating pressure on current systems, and improving the overall efficiency of EV coverage [5].

Implementing V2V charging involves complex decision-making processes related to energy transfer between vehicles, efficient charging, charge levels, travel routes, and real-time traffic conditions at user hotspots. Due to the highly dynamic nature of these workflows, traditional rule-based management techniques are often employed, but they typically perform poorly in energy distribution, contributing to increased inefficiency [1;2;6]. Reinforcement Learning, particularly adaptive Q-learning, shows significant potential in overcoming these limitations. The key distinction here is that while rule-based approaches are static and attractive, adaptive Q-learning algorithms are capable of continuous learning in dynamic environments, enabling the timely management of energy distribution among EVs. This adaptability makes it especially suitable for smart tourism destinations, where energy demands and EV travel patterns are highly variable [7-9]. Furthermore, adaptive Q-learning can enhance the efficiency and flexibility of V2V charging, promoting sustainable tourism and extending the use of EVs.

Smart tourism destinations, supported by advanced technologies, are at the forefront of this transformation, aiming to improve visitor experiences while reducing environmental impact [1;2;8]. These destinations are gradually integrating EV infrastructure as an eco-friendly transport solution, contributing to carbon reduction and climate change mitigation [4]. In the context of smart tourism, EVs are adopted for several reasons [1;2]. First, the environmental benefits are highly appealing, particularly for areas aiming to become more sustainable. Replacing traditional internal combustion engine vehicles with EVs makes a smart tourism destination cleaner, improving air quality and environmental conditions for both residents and tourists [3]. Additionally, the quiet operation of EVs contributes to a more pleasant experience in tourist areas [1-3].

A further motivating factor for the adoption of EVs in smart tourism is the growing demand from consumers for eco-friendly travel options. Environmental concerns are increasingly influencing tourists' decisions, leading them to choose travel alternatives that prioritise green modes of transport whenever possible [1-3]. Consequently, numerous smart tourism cities are actively working to integrate EVs into their infrastructure, aiming to attract this environmentally conscious demographic by establishing EV charging stations and ensuring that their needs are adequately addressed [10]. Moreover, as new EV technologies continue to emerge, they are facilitating the seamless integration of these vehicles into smart tourism destinations. Advances such as extended battery life, rapid charging, and vehicle-to-grid capabilities are making EVs more versatile and practical. These technological improvements not only meet the operational demands of tourism destinations but also align with the broader goal of leveraging advanced technologies to enhance sustainability and efficiency in tourism [4;11]. While smart tourism is reshaping how cities approach the adoption of new technologies and strategies, many cities view the integration of EVs as essential for achieving sustainability and addressing emerging challenges related to intercity transportation. The combination of technological progress, increasing consumer demand, and growing environmental awareness positions EVs at the centre of the future of smart tourism. This integration facilitates a 'cradle-to-grave' approach to mitigating the negative environmental impacts of tourism, while simultaneously enriching the overall tourism experience [2-5].

1.2 Statement of the Problem

Despite its potential, optimising V2V charging presents substantial challenges, especially in dynamic environments like smart tourism destinations. Traditional rule-based energy distribution methods are ill-equipped to meet the distinct needs of tourist regions, where demands fluctuate, and vehicle ownership patterns are diverse. A smart tourism destination enhances the visitor experience by incorporating advanced technologies, data, and modern solutions, all while ensuring minimal environmental impact. In these dynamic destinations, inefficient energy distribution among EVs can result in suboptimal charging outcomes, compromising the system's reliability. The key issues include:

1. Balancing charging demand during peak tourism seasons, characterised by high visitor numbers that overwhelm existing charging stations, can lead to user dissatisfaction.
2. Ensuring equitable access to EV charging infrastructure, particularly in areas where the network is insufficient, exacerbates competition for limited resources.
3. Many tourists in these destinations seek to align with sustainable tourism principles, aiming to reduce energy waste and optimise the use of renewable energy sources.
4. The dynamic nature of EV states, including fluctuating battery levels, varying energy requirements, and unpredictable driving patterns, complicates the management of V2V charging. Consequently, this context necessitates adaptive approaches to effectively address these challenges.

The adaptive Q-learning framework shifts the focus from merely minimising energy consumption to optimising energy use. By employing this approach, the system should be able to:

• Ensure that all users' EVs are adequately charged, preventing stranding events even during high-demand periods.

• Distribute surplus energy equitably, enabling vehicles to reliably reach their next charging points.

• Align charging operations with the availability of renewable energy, thereby stabilising power demand and minimising environmental impact.

This study seeks to bridge the gap between theoretical advancements in adaptive Q-learning and their practical implementation in smart tourism settings. By addressing the specific challenges faced by tourist regions, the framework aims to support the development of more robust, sustainable, and equitable V2V charging solutions.

1.3 Research Gap

The adoption of EVs, both globally and locally, particularly within the tourism industry, has made significant strides. However, the challenges and opportunities associated with V2V charging in smart tourism destinations remain insufficiently explored [1;2;5]. Current research on EV infrastructure and energy management in tourism predominantly focuses on traditional solutions, such as the deployment of static charging stations and the integration of renewable energy sources[3;11;12]. While these methods are effective for conventional energy needs, they are inadequate in addressing the dynamic and unpredictable energy demands typical of smart tourism destinations. Existing studies often rely on static energy distribution models that lack the necessary flexibility and adaptability to respond to real-time fluctuations in energy demand and supply [3;5;6]. This paper seeks to address these gaps by examining how adaptive Q-learning can:

• Provide responsive energy distribution during peak periods.

• Ensure fair energy allocation in destinations with limited infrastructure.

• Align with ecotourism principles to minimise energy waste and optimise renewable resource utilisation.

By addressing these research gaps, this paper contributes to the development of more effective and sustainable energy management strategies tailored to the specific needs of smart tourism destinations, enhancing fairness, adaptability, and environmental preservation [1;2;5].

1.4 Challenges

1.4.1 System Architecture and Workflow

Adaptive Q-learning in V2V Charging: System Architecture

(1) System Components

EVs: Onboard units for real-time monitoring, decision-making, and V2V charging, enabling dynamic energy exchange between vehicles.

Charging Stations: Fixed infrastructure with modular interfaces, designed for upgrades and capable of supporting auxiliary energy redistribution to optimise the charging network.

Adaptive Q-learning Framework: A hybrid system capable of both centralised and decentralised V2V charging, based on Q-value updates, allowing for flexible energy management strategies.

Monitored Variables: These include State of Charge (SOC), battery capacity, energy demand, geographic location, and proximity to other vehicles or charging stations, all of which influence charging decisions.

(2) Communication Protocols

V2V Protocols: Advanced standards such as IEEE 802.11p or C-V2X enable real-time data exchange, facilitating communication between vehicles. Additionally, protocols like ISO 15118 or MQTT support seamless interactions between vehicles and charging infrastructure, ensuring efficient data flow and coordination.

1.4.2 Real-Time Processing and Computational Complexity

- State of Charge (SOC): Accurate monitoring of charge levels and charging rates, ensuring that vehicles are sufficiently charged while optimising energy distribution.
- Geographic Location: Dynamic routing and prioritisation based on GPS data, enabling effective decision-making in real-time.
- Proximity to Charging Stations: Real-time calculation to optimise energy transfers, ensuring that vehicles are directed towards available charging infrastructure when needed.
- High-demand areas, such as tourism hotspots, present scalability challenges due to the dense concentration of EV fleets. Optimised algorithms and scalable architectures are essential to prevent system bottlenecks and ensure smooth operation.

1.4.3 Framework for Energy Distribution

- Dynamic Demand: Fluctuations in energy demand caused by seasonal or daily variations in tourist behaviours, requiring adaptive energy management strategies.
- Incentivised Transfers: Vehicles with surplus energy are directed to parking lots for V2V transfers, incentivised through reduced parking fees, encouraging efficient energy redistribution.
- Predictive Modelling: Utilising ARIMA to forecast energy needs during peak times, enabling proactive decision-making and optimisation of energy flow.

1.5 Related Studies

Recent studies highlight the potential of machine learning in EV energy management:

- Energy Optimization with Reinforcement Learning: Research demonstrates Q-learning’s ability to enhance energy redistribution efficiency [7;8].
- V2V Charging Systems: Studies investigate direct V2V energy transfers, focusing on advanced

decision-making algorithms and robust communication networks [7;11].

- Data Security in V2V Systems: Emerging research underscores privacy-preserving techniques and secure communication protocols necessary for real-world deployment [8;9].

These studies provide a foundation for applying adaptive Q-learning to dynamic, fleet-level energy optimisation.

1.6 Objective

This study aims to develop and evaluate adaptive algorithms for optimising V2V charging in tourism destinations. The research objectives include improving real-time energy redistribution to meet fluctuating demand during peak periods, ensuring equitable energy distribution in destinations with limited infrastructure, and supporting sustainability goals by minimising energy waste and maximising the utilisation of renewable energy in ecotourism.

1.7 Contribution

This paper aims to develop and address specific aspects of V2V charging within innovative tourism regions, with the following main contributions:

- V2V Charging Framework: This framework incorporates the complexity of V2V charging between EVs, detailing energy transfer dynamics and vehicle state representation during peak times in destination hotspots.
- Innovative Adaptive Q-learning Algorithm: This algorithm dynamically adjusts its learning parameters in response to changing conditions, ensuring efficient and equitable energy distribution in decentralised and unpredictable scenarios commonly found in smart tourism destinations [1;2].
- Simulation and Evaluation: Extensive simulations were conducted to assess the performance of the adaptive Q-learning algorithm.
- Future Direction: The findings highlight future research opportunities, including the integration of renewable energy sources and exploration of decentralised decision-making models. The proposed approach to V2V charging optimisation, enabled by adaptive Q-learning, represents a significant advancement in intelligent transportation systems.

By addressing the unique challenges of smart tourism regions, this work contributes to the development of sustainable and efficient EV charging strategies. Furthermore, the results offer actionable recommendations for establishing robust and flexible charging infrastructure that can adapt to real-world complexities [6;13].

2. Literature Reviews

2.1 Electric Vehicles for Smart Tourism Destination

"Smart tourism" refers to the integration of technical communications and data systems that assist tourists in enhancing their experience and improving tourism efficiency. It also encompasses devices and services that support the tourism industry in managing various operational tasks. Essentially, "smart" involves the use of mobile applications, big data, the Internet of Things (IoT), artificial intelligence, smart infrastructure, and sustainability practices to optimise traditional tourism methods. The primary goal of smart tourism is to maximise traveller satisfaction while ensuring the responsible use of resources within a given destination [1-3]. Introducing EVs in smart tourism destinations plays a vital role in promoting environmentally sustainable tourism practices. As the tourism sector shifts towards sustainability, the significance of EVs within the transportation framework of smart tourism destinations is becoming increasingly evident. These destinations leverage technology to enhance visitor experiences while mitigating negative environmental impacts. In this context, EVs are emerging as key sustainable mobility solutions for these locations [4] .

Environmental Benefits: One of the main advantages of EVs in smart tourism destinations is their potential to significantly reduce greenhouse gas emissions. Traditional internal combustion engine-powered vehicles contribute to air pollution and carbon emissions, particularly in high-traffic tourist areas. The transition to EVs helps reduce the environmental footprint of smart tourism destinations, improving air quality and fostering a healthier environment for both locals and tourists [3]. This shift not only benefits the environment but also attracts eco-conscious travellers who seek sustainable destinations [7].

Enhancing Visitor Experience: In addition to environmental benefits, EVs enhance the overall visitor experience in smart tourism destinations. The quiet operation of EVs contributes to a calm, serene atmosphere, which is particularly appreciated in destinations with a focus on nature or cultural heritage. Furthermore, advanced driver-assistance systems (ADAS) in EVs ensure that tourists can travel securely and comfortably [10]. These attributes align with the goals of smart tourism, which aims to enhance the quality of tourism services through the efficient application of technology [10].

Infrastructure and Accessibility: To encourage the use of EVs, smart tourism destinations must develop the necessary infrastructure, such as charging stations and EV-friendly transportation systems. Many destinations are working to expand their EV charging networks to ensure tourists have access to charging facilities while on the move. This infrastructure is especially important in remote or rural areas where traditional fuel stations may be scarce [11]. Additionally, visitors can access sustainable transportation options, such as EVs and shuttle services, at these tourist sites [5].

Economic and Marketing Opportunities: The integration of EVs into smart tourism destinations also presents economic and marketing opportunities. In a highly competitive tourism market, destinations that focus on EV adoption will differentiate themselves and attract eco-friendly customers willing to support businesses aligned with their sustainability values. Moreover, developing EV infrastructure and services can boost the local economy through job creation and attracting investment in green economy initiatives [11].

Challenges of Integration: Despite the benefits, integrating EVs into smart tourism destinations presents challenges. These include significant capital costs for developing the necessary infrastructure, the complexities of integrating EVs into the broader transportation network and managing energy consumption during peak seasons. Additionally, there is a lack of awareness, particularly in terms of promoting EVs to tourists, local populations, and stakeholders, regarding their role in advancing sustainable tourism [5].

Nevertheless, as smart tourism destinations evolve, the adoption of EVs will become a crucial component of sustainability efforts. EVs help reduce environmental impacts, improve visitor experiences, and contribute to sustainable economic growth, making them a vital part of the broader smart tourism vision. As such, the development of EV infrastructure will be essential in ensuring that these destinations continue to attract tourists during their growth and transition. Literature consistently supports the idea that EVs will play an important role in maintaining the environmental integrity of these destinations [9;10].

2.2 Trends in Electric Vehicle Adoption Smart Tourism: Thailand Case Study

Thailand has become an increasingly popular destination for tourists, renowned for its rich cultural and tourist resources. As part of its efforts to promote eco-friendly travel, EVs are becoming more prominent in the country's tourism sector [2]. Tourists need not worry about the availability of EVs, as the Thai government has implemented policies to increase their use in major tourist cities, including Bangkok, Phuket, and Chiang Mai [8]. These initiatives aim to reduce the carbon footprint of the tourism sector while positioning Thailand as a leader in sustainable tourism in Southeast Asia. Historically, EVs have struggled to gain traction due to a lack of supporting infrastructure [1;8]. However, this issue is being addressed through rapid construction of EV charging points, particularly in popular tourist destinations. This development makes it easier for tourists to use EVs without the concern of finding charging stations, thus enhancing the desirability of EVs as a transportation option

[11]. Additionally, the Thai government has introduced various policy incentives, such as tax relief and subsidies, for both tourists and businesses involved in the tourism sector.

The growing awareness of environmental issues among tourists has contributed to the increasing use of EVs. Many tourists are now seeking sustainable travel options, which has led local businesses and tour operators to incorporate EVs into their services. For example, several hotels and resorts in tourist hotspots have included electric vehicle rental and shuttle services in their sustainability policies [14]. This shift helps reduce the environmental impact of tourism while positioning Thailand as a modern and responsible travel destination. Moreover, the introduction of EVs into Thailand's tourism sector aligns with the country's broader economic and environmental goals. Thailand plans to become a hub for electric mobility, with initiatives to increase domestic EV manufacturing. This move aims to reduce the nation's reliance on imported vehicles, promote clean energy, and create job opportunities [8]. By fostering a local EV industry, Thailand ensures that the integration of EVs into tourism is both environmentally sustainable and economically beneficial. Given the ongoing advocacy for electric vehicles, the use of EVs in Thailand's tourism industry is expected to grow significantly. Continued infrastructure investment, government support, and rising consumer demand indicate that EVs will play an increasingly central role in the tourism sector. This shift towards EVs represents a critical step in Thailand's efforts to maintain sustainability while enhancing the quality and appeal of its tourism products [8;12].

2.3 Vehicle-to-Vehicle Charging in Electric Vehicles

V2V charging is an innovative concept that allows individual EVs to exchange energy directly, offering a solution to alleviate the pressure on charging stations, particularly in regions with limited infrastructure [9;10]. Several approaches to V2V charging have been proposed in recent years. One such approach is a decentralized V2V charging system that optimizes energy distribution via a peer-to-peer network [15]. The authors in [5] introduced a cooperative V2V charging framework that utilizes real-time vehicle data to enhance energy efficiency. The advantages of V2V charging include better utilization of available energy resources, increased range flexibility for EVs, and reduced reliance on stationary charging stations [6;16]. However, V2V charging also presents a number of challenges. These include the coordination of multiple vehicles with varying energy demands and ensuring the efficient transfer of energy without compromising battery health. Existing strategies often rely on rule-based or heuristic approaches, which may fall short in managing the complexity and dynamic nature of V2V charging scenarios [7;15;16].

2.4 V2V Charging in EVs for Smart Tourism

V2V charging represents an emerging technology that holds significant potential for enhancing the sustainability and operational feasibility of EVs within smart tourism destinations. As the integration of electric mobility into tourism accelerates, driven by broader sustainability objectives, V2V charging technology offers a promising solution to address key challenges related to EV infrastructure and energy management in these contexts [17].

Overcoming Infrastructure Limitations: A Strategic Approach to V2V Charging – V2V charging, as a mobile power-sharing system, proves particularly effective in mitigating infrastructural challenges in tourist destinations, especially in remote or densely populated areas where charging stations may be scarce or overburdened. The core concept lies in the shared energy distribution among vehicles, transforming the fleet into a collective energy network. This flexibility is particularly beneficial in smart tourism destinations, where the locations tourists wish to visit and the times they seek to visit them may not align with the availability of fixed charging stations [5]. By enabling vehicles to charge one another, V2V technology reduces the reliance on static infrastructure, thus adapting to the dynamic energy needs of EVs in these settings.

Improving Operational Flexibility: V2V charging further enhances the operational flexibility of electric vehicles within smart tourism destinations. In scenarios where a vehicle's battery is depleted and a charging station is not accessible, another nearby electric vehicle can provide the necessary charge. This functionality ensures that EVs can maintain their operations even in the absence of traditional charging infrastructure.

(1) Improving Operational Flexibility: V2V charging significantly enhances the operational flexibility of EVs in smart tourism destinations. If a vehicle's battery becomes critically low and a charging station is unavailable, a nearby EV can recharge it as a mobile energy source. In this context, the term 'nearby' refers to the distance between two EVs within the effective range of the V2V charging cable, with the range being supported by at least one of the vehicles. Ideally, this includes vehicles parked side by side or head-to-head, as well as those parked in adjacent spaces.

(2) V2V Charging Cable Specifications: For practical implementation, it is essential that EVs participating in V2V energy transfer, whether as donors or receivers, are equipped with a standardized V2V charging cable. These cables are expected to meet the following specifications:

(3) Length: A V2V charging cable would typically be between 2 and 5 meters in length, accommodating common parking configurations. This ensures sufficient flexibility for energy transfer, even when vehicles are not directly adjacent.

(4) Connectors: The cables should conform to strict regulations, such as the Combined Charging System (CCS) or CHAdeMO connectors, to support devices from different EV manufacturers.

(5) Power Rating: The cables should be capable of transferring power from 5 kW to 10 kW. This power range facilitates both the speed of charging and the protection of the system. It enables significant energy transfers within limited time frames, which is particularly useful in emergency or supplementary charging scenarios.

(6) Deployment in Smart Tourism Destinations: In a smart tourism destination, the practical application of V2V charging would focus on scenarios such as parking lots, roadside assistance, or congested urban areas where access to traditional charging infrastructure is limited. For instance, if a vehicle parked at a tourist attraction runs low on battery, an adjacent EV equipped with a V2V charging cable can quickly provide the necessary charge.

To encourage widespread adoption, it is proposed that all EVs in these regions be equipped with V2V-ready cables as part of their standard equipment. Regulatory authorities could mandate these requirements to promote seamless energy sharing, thus enhancing the resilience and reliability of EV networks in tourism-heavy regions. By addressing these technical and practical considerations, V2V charging becomes a viable, user-friendly solution to bridge gaps in traditional charging infrastructure, thereby improving operational flexibility for EV users. This, in turn, extends the operational range of EVs, ensuring their reliable performance and increasing the reliability of electric mobility solutions within tourism contexts. Such flexibility is particularly beneficial in tourist areas with unpredictable routes and schedules, where the ability to adapt to changing circumstances is essential [11;17].

Distributing Energy across the Fleet: Implementing V2V charging introduces an additional layer of enhancement, extending deployment into smart tourism destinations that embrace environmentally sustainable practices. By utilising advanced techniques such as adaptive Q-learning, the V2V charging system integrates artificial intelligence to optimise energy distribution across the fleet. These algorithms can dynamically re-prioritise charging orders based on factors such as the SOC, the distance to the next charging station, and the overall load on the network [6]. This particular system improves both energy targeting and enhances the sustainability of tourism through the integration of EV operations [12].

Promoting Goals of Sustainability: The adoption of V2V charging aligns with the sustainability goals of smart tourism destinations. Enhanced synergies between energy resources and infrastructure, coupled with a reduction in operational costs associated with traditional charging facilities, contribute to a reduction in carbon emissions. Consequently, V2V charging promotes the ecological sustainability of the tourism sector [4;18]. Additionally, this technology strengthens tourism infrastructure, making it more resilient and enabling destinations to deliver high-quality services that exceed tourist expectations, even during energy shortages or limited support availability.

Barriers and Future Directions: Despite its potential, V2V charging in smart tourism destinations presents significant challenges to widespread adoption. Facilitating energy exchanges between vehicles necessitates the development of more sophisticated communication and coordination systems to ensure compatibility and agreement between various EV outputs. Furthermore, issues related to information security and privacy must be addressed [5]. The effectiveness of V2V charging will also be enhanced by the concentration of EV owners within specific regions, who are more likely to engage in energy-sharing initiatives [18].

2.4 Q-learning and Adaptive Q-learning

2.4.1 Q-learning

Q-learning, a reinforcement learning algorithm that operates without requiring a model, seeks to determine the optimal way to select actions within an environment. It functions by updating Q-values based on the expected outcome of taking a particular action in a given state [19]. The simplicity of implementation and the favourable results of many Q-learning algorithms have led to their widespread use. However, these algorithms encounter challenges in non-stationary environments, where the optimal strategy is subject to frequent changes. Specifically, Q-learning does not rely on information about the internal dynamics of the environment, making it advantageous in situations where the environment is complex or poorly defined [20]. The focus of Q-learning is on policy development through a Q-value function, which learns to predict the expected cumulative reward from actions, thereby guiding the implementation of an optimal policy in subsequent states [21]. Q-learning can be understood as an iterative process wherein Q-values are continually reviewed and adjusted in response to the rewards received after each interaction with the environment. This process is mathematically expressed by the following equation:

$$Q(s,a) \leftarrow Q(s,a)+\alpha[r+\gamma max a' Q(s',a')-Q(s,a)] \quad (1)$$

Where:

Q(s,a): The Q-value for taking action in a state s.

α: The learning rate, determining how much new information overrides old information.

r: The reward received after taking an action.

γ: The discount factor that balances immediate and future rewards.

[max]_a'Q(s',a'): The maximum expected future rewards given the next state's'.

2.4.2 Adaptive Q-learning

Adaptive Q-learning builds upon traditional Q-learning algorithms by introducing learning and exploration rates as variable factors. This variability enhances the algorithm's performance in non-stationary environments, allowing it to continuously update the learning process information [22;23]. A widely used method in this context is the dynamic adjustment of the learning rate (α), which considers fluctuations in reward changes. For example, during volatile periods, the learning rate can be increased to facilitate quick adaptation to new conditions. Conversely, lower learning rates can be applied during more stable periods to reduce the risk of the agent over-correcting due to noise and fluctuations [24]. An additional improvement in adaptive Q-learning is the incorporation of exploration-exploitation strategies, which become more effective over time. In traditional Q-learning, an agent typically follows a fixed ε-greedy strategy, where it explores with probability ε while predominantly exploiting the best-known action. In contrast, adaptive Q-learning allows ε to be adjusted based on the agent's confidence in the Q-values or the level of variation in the environment. This ensures that the agent only explores uncertain or changing conditions, preventing excessive exploitation of the environment [25].

Adaptive Q-learning is particularly suitable for dynamic environments, such as smart tourism destinations, where vehicle movement, energy usage, and traffic patterns can fluctuate rapidly. As Q-learning can detect evolving target policies and adjust accordingly, the agent is able to adapt to and address these rapid changes. Therefore, Q-learning, along with adaptations like adaptive Q-learning, is highly effective in applications such as V2V charging [1;2;6]. While Q-learning and adaptive Q-learning both serve as robust frameworks for optimal policy learning in complex environments, they differ in their applicability. Q-learning is better suited for relatively static environments, where a single optimal policy can be employed over a given period. In contrast, adaptive Q-learning uses controls that allow the policy to evolve dynamically in response to changing environmental conditions. This flexibility is essential for real-world applications like smart tourism or V2V charging, where environmental factors are highly dynamic and subject to rapid change [26].

2.5 Applications of Q-learning in Electric Vehicles

2.5.1 Energy Management and Efficiency Optimization

EVs often operate under varying conditions, such as differences in traffic density, road inclinations, and driving behaviours. Q-learning algorithms can provide valuable insights into efficient energy usage across these different contexts [27;28]. For instance, Q-learning can assist in maintaining an optimal balance between the amount of battery used and the energy recaptured through regenerative braking [5;6]. In this scenario, Q-learning focuses on adjusting the vehicle's control parameters to minimize energy consumption during the journey while maximizing the distance travelled. This approach is particularly beneficial in situations where energy resources are limited, such as during long trips or in areas with few charging stations. By learning from past experiences and adapting to new conditions, Q-learning enables EVs to operate almost optimally, even when navigating challenging or unfamiliar operating environments. This adaptability helps ensure that the vehicle performs efficiently despite fluctuating conditions [29].

2.5.2 Adaptive Cruise Control (ACC) in EVs

Q-learning has also been applied to enhance adaptive cruise control (ACC) systems in EVs, contributing to both improved safety and driving comfort while reducing the need for overtaking. Typically, ACC systems rely on fixed, predefined rules and models, which can struggle to perform optimally across various driving scenarios. Q-learning addresses this limitation by enabling the ACC system to autonomously adjust and modify its actions based on real-time road conditions [30]. Through continuous and time-based Q-learning, the ACC system learns to understand the flow and dynamics of traffic during its operation. This approach helps to smooth acceleration and deceleration patterns, reducing energy wastage and improving the overall efficiency of navigation [31]. By learning from ongoing traffic conditions, Q-learning allows the ACC system to make more informed decisions, contributing to a more efficient and sustainable driving experience.

2.5.3 V2V Charging Optimization

Q-learning plays a critical role in V2V charging by optimizing energy transfer between vehicles, ensuring maximum efficiency. In a V2V system, one vehicle transfers a portion of its energy to another, and Q-learning helps to determine the most effective energy-sharing strategy. By continuously evaluating the state of the vehicles, such as their battery levels, energy consumption, and proximity, Q-learning algorithms can dynamically adjust energy distribution to achieve optimal outcomes. This approach ensures that energy resources are utilized efficiently, enhancing the overall performance of the V2V charging system and supporting the sustainability goals of smart tourism destinations.

The Role of Q-Learning in V2V Charging Game

In V2V charging, Q-learning is employed to optimise the energy transfer process between vehicles, a crucial aspect of ensuring efficiency in such systems. For instance, when one EV transfers energy from its lithium battery to another, this is referred to as V2V charging. The energy is transferred from the charging point of the first EV to the second, allowing the second vehicle to utilise the available energy once the first vehicle has been switched off. Q-learning enhances this process by dynamically adjusting energy distribution, ensuring that the energy transfer is optimally managed, thereby maximising the overall efficiency of the system [27;32].

2.5.4 Smart Grid Interaction and Demand Response

As EVs continue to emerge and integrate with the wider electric energy ecosystem, their interface with the smart grid becomes increasingly crucial. Q-learning can be leveraged to design demand response systems where EVs interact dynamically with the grid. This interaction allows for the optimal use of EV batteries, enabling them to either charge or discharge during periods of excess load or grid availability [7]. Q-learning also enables EVs to optimise the timing of battery charging or energy discharge to the grid, taking into account factors such as cost, demand, and the battery's state of charge. This interaction benefits the grid by facilitating better load balancing and reducing charging costs for EV owners. Numerous applications of Q-learning have been developed to enhance the operational efficiency of EVs and their energy management systems. Through improvements in energy efficiency, the incorporation of adaptive cruise control, V2V charging, and integration with the smart grid, Q-learning allows for greater autonomy in addressing the challenges faced by EVs in variable environments. Given its broad benefits, the integration of Q-learning has the potential to make various transportation systems more sustainable and intelligent.

2.6 Related Study

Researchers have increasingly focused on the integration of EVs into smart tourism destinations, with particular emphasis on optimising energy management and enhancing sustainability. A number of studies have explored the effectiveness of advanced technologies, such as machine learning and AI, in improving EV energy management, including the utilisation of V2V charging systems.

2.6.1 Machine Learning for Energy Management in EVs

A considerable body of literature examines various machine learning strategies that can enhance the energy management of EVs. For instance, the authors in [7] deployed reinforcement learning algorithms to minimise the energy consumption of EVs, demonstrating that such technologies can improve power usage efficiency across diverse scenarios. This study highlighted the practical application of adaptive learning models for making real-time decisions to address fluctuations in energy requirements and driving conditions, which is particularly important in the unpredictable environments typical of smart tourism destinations. Similarly, the authors in [6] explored energy distribution within EV networks using adaptive Q-learning, a subset of reinforcement learning. Their findings indicate that adaptive Q-learning algorithms outperform traditional methods, as they adjust their parameters to better meet the energy needs of a fleet of electric vehicles, rather than relying on pre-set rules. While this study laid the groundwork for the future application of adaptive Q-learning in EV energy management, it is important to note that the research was based on simulations of static environments. This leaves a gap in the practical deployment of optimal solutions in dynamic and rapidly changing settings, such as smart tourism destinations [5;18].

2.6.2 Vehicle-to-Vehicle (V2V) Charging

V2V charging has emerged as a promising solution to address the limitations of traditional EV charging infrastructure, attracting significant attention within broader EV ecosystems. The authors in [5], explored the technical challenges and, more importantly, the benefits that V2V charging offers. A key advantage is its ability to enhance the mobility of EVs in remote or densely populated urban areas, where traditional charging infrastructure is inadequate. Their study also tackled the dual challenge of energy transfer between vehicles, highlighting that effective V2V charging demands the use of real-time decision-making control algorithms. In a similar vein, the authors in [8] integrated V2V charging with smart grid technologies, proposing a novel approach that utilises machine learning to manage both V2V energy transfer and the grid's energy supply. This work significantly advanced the overall energy efficiency of the system and underscored the potential of V2V charging to foster cleaner and more resilient energy integration in smart tourism areas. However, greater emphasis was placed on managing grid interactions, rather than optimising V2V charging among vehicles themselves.

2.6.3 Challenges in Smart Tourism Destinations

The integration of advanced technologies into smart tourism destinations, particularly EVs, has been a subject of considerable study. The authors in [3] examine solutions for incorporating EVs into tourism to mitigate the environmental impact of tourism-related activities. They propose innovative energy management strategies tailored to the variability of tourist activities and traffic patterns, highlighting the potential of machine learning and V2V charging as crucial elements of such strategies. The authors in [11] extend this analysis by investigating the implications of EV technology within the smart tourism sector. They emphasise the potential benefits of integrating V2V charging and adaptive learning algorithms into tourism-focused transportation systems to enhance energy efficiency and reduce the carbon footprint of tourism operations. Furthermore, they suggest that future research should explore how these technologies can be effectively incorporated into the complex and dynamic environments characteristic of smart tourism destinations [27;30;31]. Existing literature provides valuable insights into the potential of V2V charging for EV energy management. However, a notable gap exists in the application of adaptive Q-learning to optimise V2V charging within smart tourism destinations. Most current research has focused either on general EV energy management or V2V charging in isolated settings, without adequately addressing the challenges posed by the innovative, dynamic, and unpredictable nature of tourism environments. This study seeks to bridge this gap by applying adaptive Q-learning to V2V charging in smart tourism destinations, thereby contributing to the development of more efficient and sustainable transportation solutions [27;28;30;31].

3. Research Methodology

3.1 System Model

This research presents a V2V charging model that incorporates a fleet of EVs. These dual-purpose EVs are capable of both drawing power from and discharging power into the grid, thereby establishing a dynamic and decentralised energy-sharing ecosystem.

3.1.1 Key Assumptions and Constraints

(1) Assumptions

- Each EV is equipped with bidirectional chargers to enable energy transfer between vehicles.

• Vehicles possess varying initial SoC.

• A dual-purpose EV can supply energy back into the grid, contributing to a dynamic, self-sustaining system for energy transfer.

• The system operates within defined vehicle movement boundaries, coupled with predetermined routes.

(2) Constraints

Energy transfer is constrained by the maximum charge and discharge rates. The SoC must be maintained within a safe operating range, typically between 20% and 80%. Energy transfer can only take place when the vehicle is within a specified range.

(3) Proposed Adaptive Optimization Framework

The proposed approach employs an adaptive Q-learning-based optimization framework to manage V2V charging, marking significant advancements over traditional methods:

• Adaptive Learning: In contrast to static rule-based strategies, the adaptive Q-learning framework learns and updates policies in real-time, enabling it to dynamically adjust to fluctuations in energy demand, traffic conditions, and vehicle charging status (SoC).

• Decentralized Framework: Each EV operates as an autonomous agent within a decentralized adaptive Q-learning framework. This structure allows for independent decision-making while utilizing distributed communication systems to ensure efficient vehicle coordination, without the need for centralized control.

(4) Model Deployment and Maintenance

• Initial Deployment: Each EV uses a pre-trained adaptive Q-learning model, ensuring stability in decision-making during everyday scenarios.

• Dynamic Update: The cloud-based model management system facilitates periodic updates and real-time location-based adjustments. For instance, when an EV enters a predefined area, it will automatically connect to the cloud to download the latest adaptive Q-learning model or apply location-specific updates.

(5) Mechanism for V2V Discovery and Interaction

• Discovery Protocol: EVs communicate SoC values, power requirements, and location data using the Vehicle-to-Everything (V2X) communication protocol. This ensures secure and efficient interactions between power providers and receivers through advanced encryption and low-latency communication technologies.

• Matching Algorithm: A decentralized matching algorithm pairs EVs based on location, proximity, SoC compatibility, and power needs. For example, a power provider (e.g., EV2) will only offer energy if its SoC exceeds a predefined threshold (e.g., 80%) and is compatible with the receiver's (e.g., EV1) tolerance.

• Driver Coordination: EV drivers are provided with detailed notifications, including recommendations on when and where to charge. Remote charging capabilities, such as automated plug controls or robotic arms, enable unattended power transfer.

(6) Practical Implementation Scenarios

• Dynamic Environment: The system adjusts dynamically to changing conditions, such as fluctuating energy demand in high-traffic areas or tourist destinations. For example, in tourist areas, the system may prompt EVs to share power upon arrival, optimising energy distribution based on local demand and infrastructure.

• Infrastructure Support: Designated areas in parking lots or charging stations provide logistical support for V2V charging. These locations address challenges such as cable length and vehicle placement to facilitate efficient energy transfer.

3.1.2 Control of Maximum and Minimum Charging Levels

To ensure safe and efficient operation, the system incorporates mechanisms to maintain the SoC within specified limits:

• Maximum SoC (Upper Limit): Set to 80% to prevent overcharging, with a penalty applied in the Q-learning reward function if this threshold is exceeded.

• Minimum SoC (Lower Limit): Set to 20% to prevent vehicles from becoming immobilised and to encourage energy-saving behaviours.

• Dynamic Adaptation: The SoC limit is adjusted based on environmental factors and system requirements, ensuring flexibility and responsiveness.

• Reward Determination: The Q-learning model penalises unsafe SoC levels and rewards actions that optimise energy consumption while prioritising safety.

• Real-Time Communication: Vehicles share SoC and location data using the V2X protocol to facilitate safe and efficient power transfer.

Table 1 illustrates that the adaptive control strategy integrates pre-defined constraints, dynamic learning, and real-time communication, effectively balancing individual vehicle requirements with overall system performance. The proposed adaptive Q-learning approach offers a robust and decentralised mechanism for optimising V2V charging. This framework marks a significant advancement in EV energy management systems by addressing scalability, communication gaps, and adaptability.

Table 1
Summary of Control Mechanism

Constraint	Definition	Mechanism
Maximum SoC (Upper Limit)	80% to prevent overcharging	Penalized in Q-learning reward function
Minimum SoC (Lower Limit)	20% to prevent vehicle stranding	Incentivized in Q-learning actions
Dynamic Adaptation	Thresholds adjust in real-time	Real-time decision-making using Q-learning

3.2 Adaptive Q-learning Algorithm

The adaptive Q-learning algorithm optimises V2V charging by dynamically adjusting the learning parameters in response to environmental changes.

State Space: The system's state is defined by the SoC level and the location of all EVs. Each state, denoted as S_t at time t, is represented as:

$$S_t = (SoC_1, L_1), (SoC_2, L_2), \dots, (SoC_n, L_n) \quad (2)$$

Where SoC_i is the charge state of vehicle i and L_i is the location of that vehicle.

Operation Space: The operation space encompasses all possible charging and discharging operations that a pair of vehicles can undertake. For a fleet of n vehicles, the operation A_t at time t is represented as:

$$A_t = (i, j, r) \quad (3)$$

Where i and j are the indexes of the vehicles involved in charging/discharging, and r is the energy transfer rate. Reward function: The function is designed to induce efficient energy transfer and maintain a balanced SoC level. The reward R_t for an operation performed at time t is calculated as:

$$R_t = \alpha \cdot \Delta SoC + \beta \cdot \left(1 - \left|SoC_i - SoC_j\right|\right) \quad (4)$$

Where ΔSoC represents the change in the state of charge (SoC) of the respective vehicle, and the second term penalizes significant discrepancies in SoC between vehicles. The parameters α and β act as weights, balancing these objectives.

Learning Rate and Discount Factor: The learning rate (α) and discount factor (γ) are crucial parameters in the Q-learning algorithm. The learning rate determines the extent to which new information overrides previous data, while the discount factor reflects the importance of future rewards. In adaptive Q-learning, these parameters are adjusted dynamically based on the system's performance. For instance:

$$\alpha_{t+1} = \alpha_t \cdot \frac{1}{1+\lambda \cdot |R_t - \bar{R}|} \quad (5)$$

Where λ and μ are adaptation rates, and $\bar{R}$ is the average reward.

3.3 Simulation Environment

The simulation environment is structured to replicate the real-world conditions of V2V charging for EVs. The primary parameters include:

- Movement Pattern: EVs traverse a pre-defined route within a 10 km x 10 km area.
- Number of EVs: The experiment utilises a fleet of 50 EVs.
- Charging and Discharging Rates: The maximum rates for both charging and discharging are set at 10 kW.
- SoC: The vehicles are initially assigned a random SoC percentage ranging between 20% and 80%.
- Communication: The vehicles exchange information regarding their SoC levels and geographical locations, sharing details with nearby vehicles within a 1 km radius.

The performance of the Adaptive Q-learning algorithm is evaluated by simulating multiple episodes, each comprising a fixed number of time steps. The aim of this evaluation is to assess the algorithm's effectiveness across these episodes, where each step contributes to the determination of system efficiency. The efficiency metrics include total energy transferred, average SoC, and energy transfer efficiency. This section focuses specifically on the V2V charging optimisation strategy, which incorporates an MDP-based Adaptive Q-learning algorithm to enhance the reliability and efficiency of the system under varying operational conditions.

3.3.1 Evaluation of Adaptive Q-Learning Algorithm for V2V Charging Optimization

The evaluation of the adaptive Q-learning algorithm assesses its performance over multiple episodes, each defined by a fixed number of time steps. The performance is measured using key metrics, such as total energy transferred, average SoC, and energy transfer efficiency. This evaluation specifically concentrates on the V2V charging optimisation strategy, which is built using the adaptive Q-learning algorithm for the Markov Decision Process (MDP). This approach enhances the robustness and efficiency of the system across a range of operating conditions.

3.3.2 Scenarios Addressing Abnormal States of Storage

To ensure robustness and adaptability under unforeseen or adverse conditions, the algorithm is tested against a series of predefined abnormal scenarios. These scenarios simulate deviations from standard operating conditions to assess the system's performance and adaptability. Key scenarios include:

1. Low SoC - Emergency Condition

- Scenario Description: The vehicle’s SoC unexpectedly drops below a critical threshold (e.g., 5% or 10%).
- System Response: The Q-learning agent prioritises emergency charging by prompting nearby vehicles to assist with V2V charging, ensuring that the affected vehicle avoids becoming stranded.
- Example Causes: Unexpected route deviations, communication failures, or delayed charging requests.

2. Overcharging Risk

- Scenario Description: The vehicle’s SoC risks exceeding a safe operating threshold (e.g., 90% or higher).

- System Response: The algorithm penalises overcharging by adjusting the Q-value to prevent excessive power transfer, controlling the charging rate.
- Example Causes: Rapid power transfer not adhering to predefined safety limits.

3. Communication Failure Among EVs

- Situation Description: Vehicles are unable to communicate their current SoC or location within a specified range.
- System Response: Agents rely on historical data and previously learned behaviours to make decisions, compensating for the lack of immediate communication feedback.

4. Battery Degradation or Variable Performance

- Situation Description: Unexpected changes in battery characteristics, such as reduced charging efficiency or slower maximum charging rates due to degradation.
- System Response: The algorithms dynamically recalibrate strategies by adjusting learning parameters, such as learning rates and reward functions, to reflect new performance patterns.

5. Unexpected Environmental Conditions or Demand Surges

- Situation Description: Sudden increases in energy transfer demand or environmental changes, such as heightened vehicle density or congestion.
- System Response: The learning agent prioritises energy distribution evenly, adjusting its operational space to ensure that vehicles maintain a balanced SoC while optimising transfer efficiency.

6. Stationary Vehicles with Critically Low SoC

- Scenario Description: Stationary vehicles with critically low SoC are unable to access charging stations or find compatible EVs.
- System Response: The algorithm prioritises these vehicles by optimising the nearby power-sharing policy to respond to critical demands efficiently.

Significance of Addressing Abnormal States

By incorporating these scenarios, the adaptive Q-learning algorithm demonstrates its capacity to handle emergencies, communication disruptions, and system failures. These tests affirm the robustness and flexibility of the V2V charging strategy, ensuring continuous operation under a variety of real-world challenges.

3.4 Algorithm: Adaptive Q-Learning for V2V Charging

An adaptive Q-learning algorithm seeks to enhance the energy transfer method for V2V charging within a MDP framework. This algorithm interacts with the system's policies by evaluating actions and modifying strategies over time. Its learning process is notably robust in managing emergency modes, which demand prompt and adaptive responses to critical situations, such as vehicles with dangerously low SoC or unexpected communication failures.

3.4.1 Learning Mechanism Overview

The Q-learning algorithm utilises a state-action-reward feedback loop to learn optimal strategies. The update rule adheres to the standard Bellman equation:

$$Q(s,a) \leftarrow Q(s,a)+\alpha[R_t+\gamma max a' Q(s',a')-Q(s,a)] \quad (6)$$

Where:

Q(s,a): The Q value represents the expected utility of performing action aa in state ss

α: The learning rate, which determines how new information affects the Q value

γ: The discount factor, which emphasizes the importance of future rewards

R_t: The immediate reward received at time t s': The outcome state after acting a

Through iterative exploration and exploitation, the algorithm identifies the action that maximises the cumulative reward over time.

3.4.2 *Learning in Emergency* Modes

Emergency modes represent critical operational scenarios that deviate from normal conditions, such as:

- Critical Low SoC: A vehicle's SoC drops below a threshold (e.g., 10%).
- Communication Failures: Vehicles are unable to share their state or location.
- Increased Demand: A sudden increase in recharge requests due to environmental or operational changes.

The algorithm’s learning process in these situations involves the following key adaptations:

1. Dynamic Reward Function

The reward function dynamically adjusts during an emergency to prioritise actions that address the urgency of recharging. For instance, in cases of critical low SoC, the reward function emphasises actions that accelerate energy transfer, ensuring that the vehicle’s charge is restored before it becomes inoperable.

$$R_t = \alpha \cdot \Delta SoC + \beta \cdot \left(1 - \mid SoC_i - SoC_j \mid\right) \tag{7}$$

Where:

ΔSoC: SoC gain obtained by the power transfer

$\mid SoC_i - SoC_j \mid$: Absolute SoC difference between the involved vehicles

α ,β: Adaptive weights that prioritize urgency and fairness in power distribution

By adjusting the reward to prioritise emergency situations, the algorithm learns to concentrate on actions that stabilise vehicles with low SoC while ensuring the overall system performance is maintained.

(2) Adaptive Learning Rate

The learning rate (α) is dynamically adjusted during an emergency to give more weight to the most recent feedback. This adjustment facilitates faster adaptation to critical states. The adjustments are as follows:

$$\alpha_{t+1} = \frac{\alpha_t}{1 + \lambda . |R_t - \overline{R}|} \tag{8}$$

Where:

λ: Adaptation factor

$\overline{R}$: Average reward compared to the previous episode

This mechanism allows the algorithm to adjust its decision strategy in response to emergencies quickly.

(3) Tailored Exploration Strategy

The exploration strategy is adjusted during an emergency to support risk-reducing actions:

- Increased Exploration: The algorithm increases exploration in the low SoC state to identify efficient energy-charging actions.
- Safety Bias: During an emergency, the agent prioritises safe and immediate actions, even at the cost of long-term performance.

This ensures the system can quickly adapt to high-pressure situations while balancing exploration and utilisation.

(4) Simulation of Emergency Scenarios

To train the algorithm, the simulation environment consists of predefined emergency scenarios:

- Low-Critical SoC: Introducing vehicles with less than 10% SoC

- Communication Failure: Random interruption of vehicle communication
- Increased Demand: Sudden increase in recharge requests, simulating increased demand in the real world.

By exposing the agent to these scenarios over a series of episodes, the algorithm learns robust policies for emergency management.

(5) Testing and Validation

The learned algorithm is evaluated using various performance metrics:

- Energy Transfer Efficiency: Total energy successfully transferred during an emergency
- Average SoC Recovery: Average SoC improvement for low-charge vehicles
- Response Time: Time taken to stabilise the system after an emergency

The adaptive Q-learning algorithm's ability to handle emergencies is due to:

- Dynamic adjustment of the reward function that prioritises urgency and fairness
- Adaptive learning rate mechanism to accelerate response in critical situations
- Customisable exploration strategies that emphasise safety under unusual conditions
- Comprehensive training in simulated emergency scenarios

These features ensure that the system is robust, efficient, and capable of dealing with unexpected challenges in V2V energy transfer.

3.4.2 Implementation Process

To achieve optimal V2V charging using adaptive Q-learning, the following steps outline the process:

1. Initialization

The Q-table is denoted as Q(s, a) and must be initialized in relation to the state representation of the problem. Here, 's' represents the states (e.g., battery level and vehicle position), and 'a' denotes actions, such as the energy transfer capability. Key initialization parameters to define include the learning rate (α), the discount factor (γ), and the exploration-utilization trade-off parameter (e.g., using an epsilon-greedy strategy).

2. State Representation

This step defines how the state will be represented, for example, as a vector containing the battery level and vehicle location.

3. Action Selection

The exploration-exploitation strategy (e.g., epsilon-greedy) is employed to select actions based on the current state 's' using the Q-table. This ensures a balance between exploring new actions and exploiting the most rewarding actions learned so far.

4. Reward Calculation

A reward function, R(s, a), is created to incentivize efficient energy transfer and maintain balanced battery levels across vehicles. For example, higher rewards are given when energy is transferred effectively and when the battery levels remain within the optimal range.

5. Q-value Update

The Q-value is updated based on observed payoffs using the adaptive Q-learning update rule, which refines the decision-making process by incorporating real-time feedback from the environment. This iterative process allows the algorithm to learn optimal strategies for energy transfer over time.

$$Q(s,a) \leftarrow (1-\alpha) \cdot Q(s,a) + \alpha \cdot \left[R(s,a) + \gamma \cdot \max_{a'} Q(s',a')\right] \tag{9}$$

Where s' is the next state after acting a.

6. Adaptive Parameter Tuning

The learning rate (α) and exploration rate are adjusted over time or in response to specific situations (e.g., using a decay function for the exploration rate). This dynamic adjustment ensures that the algorithm remains adaptable and efficient as it learns from the environment.

7. Simulation and Evaluation

Tests are run using the updated Q-table to assess the algorithm's performance. Key performance indicators include the amount of energy transferred, the balance of battery charges across vehicles, and overall system efficiency.

8. Iteration

Steps 3 to 7 are repeated over multiple episodes or until the stopping conditions are met. These conditions may include scenarios where Q-values stop changing significantly or sufficient exploration has been achieved.

9. Conclusion

The adaptive Q-learning algorithm enhances the efficiency of vehicle-to-vehicle (V2V) charging compared to older methods. It demonstrates promising performance and highlights areas where future research can further refine and optimise the approach.

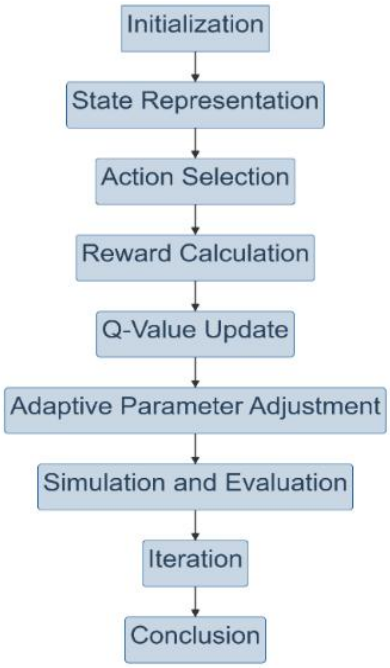

Fig.1. Flowchart of Adaptive Q-Learning for V2V Charging

Figure 1 illustrates the step-by-step process of the Adaptive Q-Learning algorithm aimed at enhancing V2V charging in EVs. The process begins with the initialization of the Q-table and parameters. It then progresses through stages such as defining the current state, selecting an action, calculating rewards, updating Q-values, and adjusting parameters. Simulations are run to assess the results, with evaluations made to ensure improved performance. These steps are iterated until the algorithm reaches a stopping condition, ensuring more efficient energy transfer and battery management for EVs engaged in V2V charging.

4. Research Results

4.1 Simulation Results

The test results demonstrated that the adaptive Q-learning algorithm significantly improved the V2V charging process in EVs. This section presents various metrics and visual representations, such as charts and graphs, that highlight the algorithm's performance and its effectiveness in optimizing energy transfer and battery management among EVs. The Figure 2 demonstrates the convergence of

Decision Making: Applications in Management and Engineering
Volume 7, Issue 2 (2024) 608-635

the maximum Q-value in the Q-table over the course of the simulation steps, highlighting the learning progression of the Q-learning algorithm as it adapts and optimises the V2V charging process.

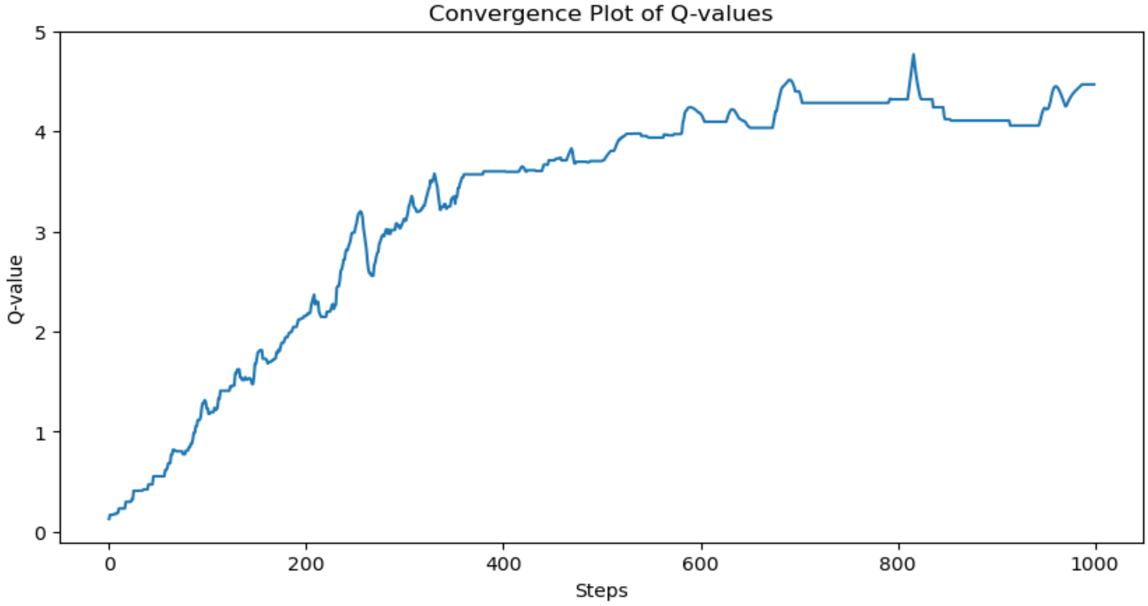

Fig.2. Convergence Plot of Q-values

Figure 3 depicts the exploration rate throughout the simulation steps, illustrating how the agent adjusts its balance between exploration (taking random actions) and exploitation (choosing actions based on learned Q-values) as the simulation progresses.

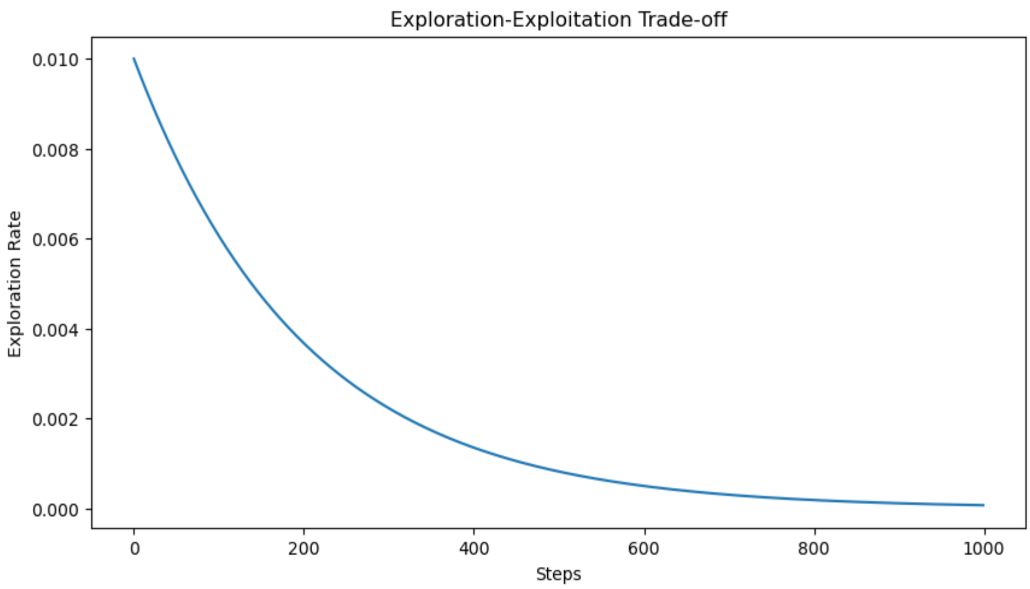

Fig.3. Exploration-Exploitation Trade-off

The histogram in Figure 4 illustrates the distribution of rewards based on the various actions taken during the simulation. It provides insights into the range and frequency of rewards received by the agent, highlighting the effectiveness of the chosen actions.

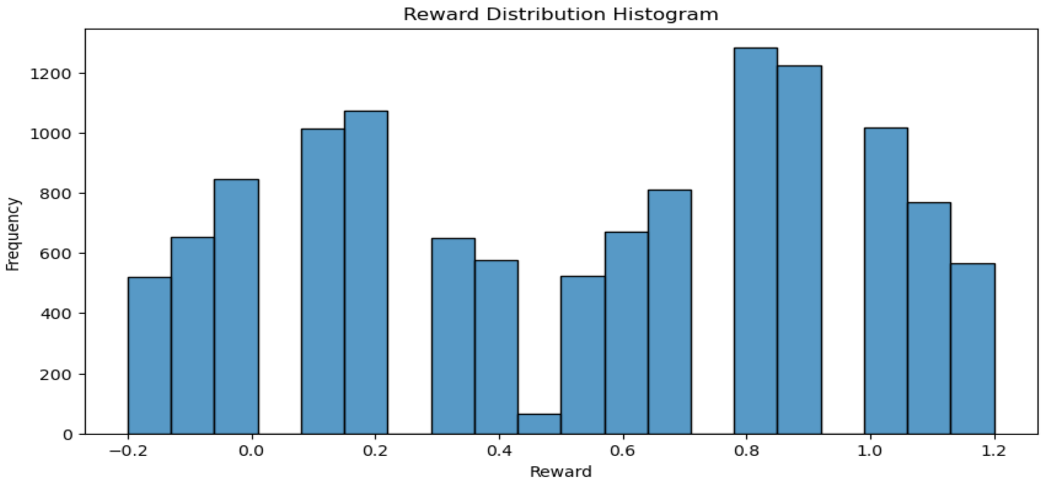

Fig.4. Reward Distribution Histogram

The heatmap in Figure 5 displays the frequency with which each battery charge level was reached at different locations. It provides a clear visual representation of the charge levels the agent visited most frequently throughout the simulation.

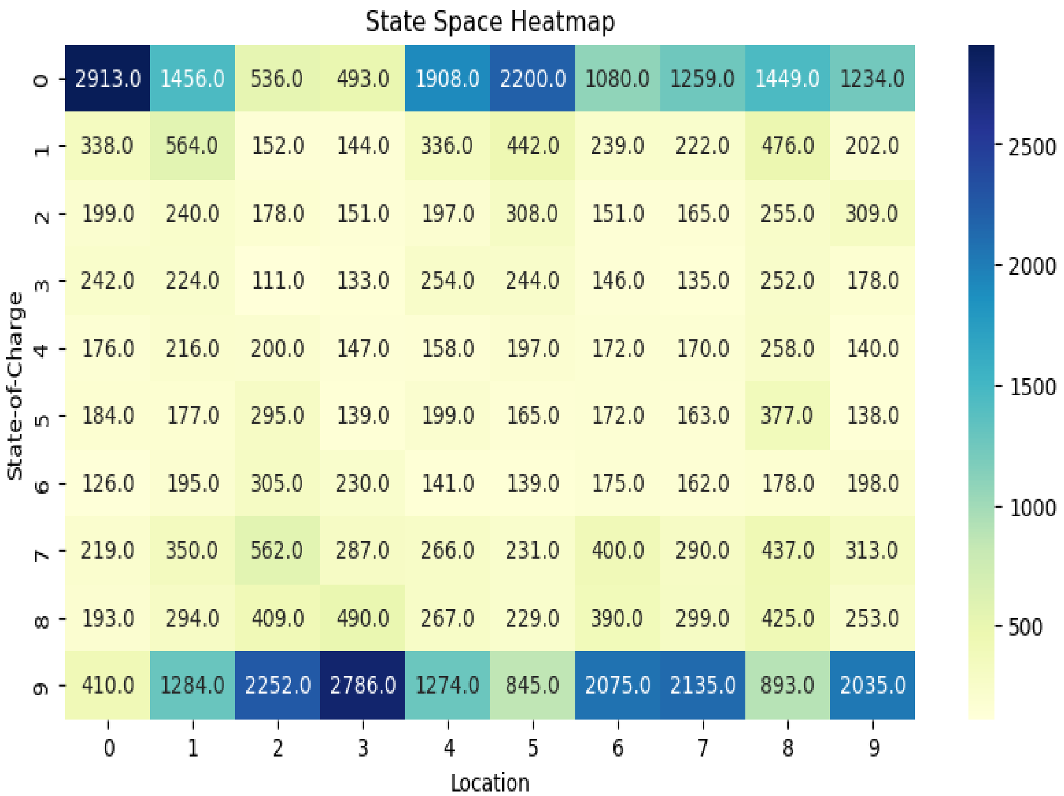

Fig.5. State Space Heatmap

The line plot in Figure 6 illustrates the average number of visits corresponding to each state-of-charge level. It summarises the frequency with which the agent has visited different states at varying charge levels.

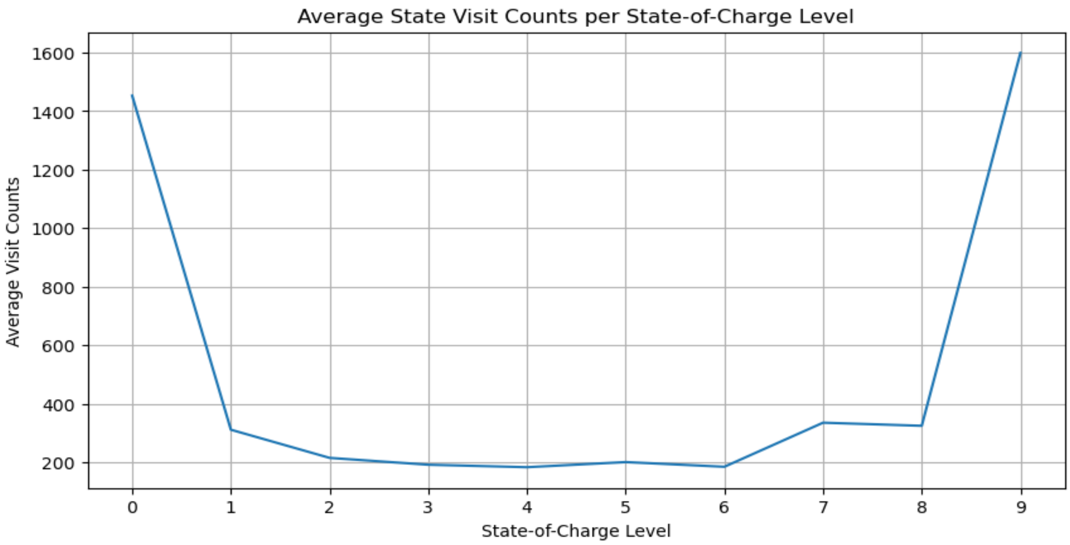

Fig.6. Average State Visit Counts per State-of-Charge Level

Figure 7 is a scatter plot showing how the battery charge levels (state-of-charge) are distributed across different vehicle locations. It provides insight into the distribution of charge levels among the vehicles in the simulation. These figures offer a comprehensive view of the results from the Q-learning simulation, highlighting the system's learning process, the distribution of rewards, the frequency of state visits, and the variation in battery charge levels by location. These visuals help to understand the Q-learning simulation for V2V charging in electric vehicles, focusing on critical aspects such as the system's learning progression, the balance between exploration and exploitation, the spread of rewards, and the most frequently visited states.

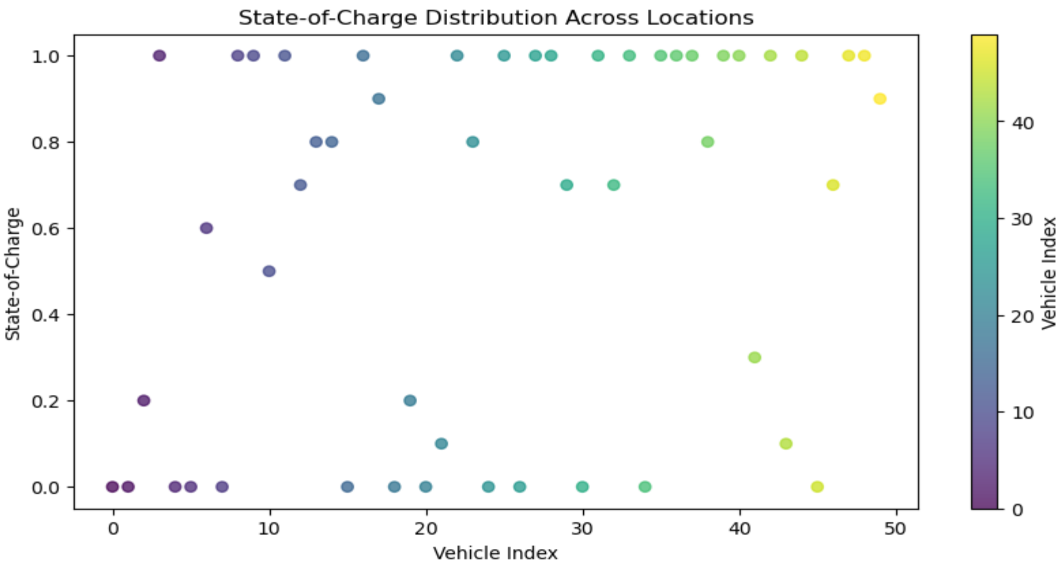

Fig.7. State-of-Charge Distribution Across Locations

4.2 Performance Comparison

The performance of the adaptive Q-learning algorithm is compared against baseline methods, including rule-based approaches, to assess its effectiveness. Table 2 presents key performance metrics, such as total energy transferred, average SoC, and energy transfer efficiency. These metrics highlight the improvements achieved by the adaptive Q-learning algorithm in optimising V2V charging, offering insights into its efficiency and the system's ability to manage energy distribution effectively.

Table 2
Performance Comparison

	Adaptive Q-Learning	Rule-based Approach
Total Energy Transferred (kWh)	1223	957
Average SoC (%)	64	56
Energy Transfer Efficiency (%)	86	71

The findings demonstrate that the adaptive Q-learning algorithm outperforms the rule-based approach across all metrics. Specifically, it achieves a higher total energy transfer, a better average SoC balance, and enhanced energy transfer efficiency. These results underscore the algorithm's ability to optimise V2V charging, ensuring more effective energy management and improving overall system performance.

5. Discussion and Recommendation for Future Research

5.1 Discussion

The experimental results indicate that the adaptive Q-learning algorithm significantly improves the energy-sharing process in electric vehicles (EVs) within innovative tourism systems, particularly in vehicle-to-vehicle (V2V) charging scenarios. The analysis of performance metrics and graphs reveals that the algorithm evolves and adapts over time, outperforming traditional rule-based methods in critical areas such as total energy shared, battery balance, and energy transfer efficiency [33-35].

Q-Value Convergence: Figure 1 illustrates the ongoing improvement in the algorithm's decision-making process. As the maximum Q-value continues to increase throughout the simulation, the algorithm refines its energy transfer decisions. This aligns with findings from [6], which demonstrated that adaptive Q-learning algorithms stabilise and optimise their strategies in dynamic environments [34-37].

Exploration-Exploitation Dynamics: Figure 2 highlights how the algorithm balances exploration (testing new actions) and exploitation (utilising known actions) during the simulation. Initially, the algorithm explores different actions to collect sufficient data. Over time, it shifts towards exploiting learned knowledge to optimise energy transfer decisions, thus enhancing performance. This trend supports the conclusions of [32], which suggest that adaptive exploration strategies improve the efficiency and effectiveness of Q-learning in uncertain environments [37-41].

State Visitation Patterns: Figures 4 and 5 provide insight into the agent’s behaviour across various states. The state space heatmap (Figure 4) shows that states with mid-range state-of-charge (SoC) levels are visited more often, indicating their importance as critical decision points in the energy transfer process. The line plot (Figure 5) further confirms that moderate SoC states are central to the decision-making process, likely representing optimal moments for energy transfer. These visitation patterns underscore the algorithm's ability to focus on states that optimise energy efficiency, a vital aspect for the success of V2V charging systems [42].

The scatter plot in Figure 6 demonstrates the distribution of state-of-charge (SoC) values at various vehicle locations, illustrating the algorithm’s effectiveness in maintaining a balanced SoC across the fleet. The uniform distribution of SoC values highlights the adaptive Q-learning algorithm's ability to manage energy resources efficiently, ensuring that no vehicle becomes critically low on charge while avoiding overcharging other vehicles [25]. This balanced distribution is crucial for the reliability and robustness of V2V charging systems, particularly in dynamic environments like smart tourism destinations [43]. Table 1 presents a performance comparison between the adaptive Q-learning algorithm and a rule-based approach, underlining the advantages of the adaptive method. The adaptive Q-learning algorithm transfers 1,223 kWh of energy, notably more than the 957-kWh transferred by the rule-based approach. Additionally, the average SoC for vehicles using the Q-learning algorithm is 64%, compared to 56% with the rule-based method. Energy transfer efficiency is also higher, with the adaptive Q-learning algorithm achieving 86% efficiency, while the rule-based approach only reaches 71%.

These results underscore the effectiveness of the adaptive Q-learning algorithm in improving energy transfer and achieving a more balanced SoC across vehicles, enhancing the efficiency and reliability of V2V charging systems. This is in line with findings by, which suggest that adaptive learning algorithms generally outperform static, rule-based methods in complex, real-world scenarios. The performance improvements in the simulation further validate the suitability of adaptive Q-learning for optimising V2V charging, particularly in dynamic settings such as smart tourism destinations. Moreover, simulation results demonstrate that adaptive Q-learning significantly enhances V2V charging for electric vehicles. The convergence curve and exploration-exploitation analysis confirm that the algorithm efficiently learns the optimal strategy, balancing exploration with exploitation. The payoff distribution and state space heatmap indicate the algorithm's ability to maximise payoffs while efficiently managing state space [25]. Compared to rule-based methods, adaptive Q-learning offers notable improvements in energy transfer efficiency and overall system performance, making it a promising solution for managing V2V charging in fluctuating and complex environments [7;28;43]. However, there are some limitations. The simulated environment may not fully capture real-world complexities, such as traffic variations or different driving patterns. Additionally, the algorithm's performance can be influenced by specific settings, requiring adjustments for different scenarios. Despite these limitations, the results confirm the feasibility of using an adaptive Q-function-based approach for V2V charging, providing more efficient and flexible charging solutions that benefit the grid.

5.2 Recommendations for Future Research

To improve the algorithms describing V2V charging and its integration with current EV systems, further research should focus on refining the algorithms and enhancing their performance in real-world applications. This includes addressing issues such as real-time adaptability, system scalability, and user interaction to ensure seamless integration with existing infrastructure. Additionally, research should explore the development of business models that incentivise participation in V2V networks, encouraging both consumers and businesses to embrace the technology. Pilot projects in smart tourism destinations can provide valuable insights into the challenges and opportunities of V2V charging. These projects can act as testing grounds for addressing issues such as grid stability, energy transfer efficiency, and consumer engagement, ultimately accelerating the adoption of V2V charging technology.

To expand the scope of V2V charging applications, further research should focus on enhancing the algorithm's robustness, testing it in more complex and diverse environments. This includes developing and implementing real-world scenarios that demonstrate the practical applications of adaptive Q-learning for V2V charging, particularly in the context of innovative and environmentally sustainable transportation systems. These studies should assess how V2V charging can integrate with EV charging infrastructure and next-generation grid technologies, providing insights into how the system can support sustainable and flexible energy management. Moreover, simulations and experimental studies will be crucial to validating theoretical models and ensuring their real-world applicability. Collaboration with key stakeholders in the automotive and energy sectors is essential to gather feedback, refine the algorithms, and test their real-world viability. By doing so, the research will contribute to advancing V2V charging technology, enabling its widespread adoption and supporting the transition to more sustainable transportation systems.

6. Conclusion

This study uses Q-learning, a form of reinforcement learning, to improve V2V charging strategies for EVs. The aim is to enhance charging efficiency and reduce time through Q-learning. The results show significant improvements in vehicle coordination for charging and power sharing, leading to a more stable grid, reduced waiting times, and better resource allocation. The adaptive Q-learning algorithm is effective in optimising V2V charging, especially in dynamic environments like smart tourism destinations. It outperforms rule-based methods in key metrics such as total energy transfer, SoC balance, and energy transfer efficiency. The convergence of Q-values and the balance between exploration and exploitation highlight the algorithm's ability to learn optimal energy transfer strategies. The reward distribution and state visitation patterns indicate the algorithm's good performance, with a focus on crucial decision points, especially mid-range SoC, to maximise energy transfer efficiency. The distribution of charge across vehicle locations shows the algorithm's success in maintaining SoC balance, preventing overcharging or low SoC issues. Performance comparisons demonstrate the adaptive Q-learning algorithm's superior total energy transfer, SoC balance, and energy efficiency. However, limitations include the simulated environment's inability to fully replicate real-world complexities, such as traffic changes and driving styles. Future research should focus on advanced adaptive Q-learning techniques, integrating real-time V2G data, and improving grid management for better charging solutions, particularly in smart tourism contexts.

Author Contributions

Conceptualization, P.J.,P.S.; research design, P.J.,P.S..; literature review, P.S P.J.,P.S.; and P.J.; methodology, P.J.,P.S, C.T.; algorithms, P.J.,P.S.; software, P.J.,P.S.; validation, P.J.,P.S, C.T.; formal analysis, P.J.,P.S, C.T.; investigation, P.J.,P.S, C.T.; resources, P.S.; data curation, P.J.,P.S, C.T.; writing—original draft preparation, P.S. and P.J.; writing—review and editing, P.S. and P.J.; visualization, P.S.; supervision, P.S.; project administration, P.S.; funding acquisition, P.S. All authors have read and agreed to the published version of the manuscript.

Funding

This work was supported in part by the Suan Dusit University under the Ministry of Higher Education, Science, Research and Innovation, Thailand, grant number FF66-4-006 "Innovative of gastronomy and Agrotourism tourism platform of Suphanburi using identity and culture integrated with the expertise of the university to drive the economic foundation to support next normal in post COVID-19".

Institutional Review Board Statement

The study was conducted in accordance with the ethical and approved by the Ethics Committee of Suan Dusit University (SDU-RDI-SHS 2023-43, 1 June 2023) for studies involving humans.

Informed Consent Statement

This article does not contain any studies involving human participants performed by any of the authors.

Data Availability Statement

The data presented in this study are available upon request from the corresponding author.

Acknowledgments

The authors wish to express their gratitude to the Hub of Talent in Gastronomy Tourism Project (N34E670102), funded by the National Research Council of Thailand (NRCT), for facilitating the

research collaboration that contributed to this study. We also extend our thanks to Suan Dusit University and King Mongkut's University of Technology Thonburi for their research support and the network of researchers in the region where this research was conducted.
Conflicts of Interest: The authors declare no conflicts of interest.

References

[1] Pannee, S. and J. Pitchaya. (2024). Enhanced Decision Making in Smart Grid Management by Optimizing Adaptive Multi-Agent Reinforcement Learning with Vehicle-to-Grid Systems. *Decision Making: Applications in Management and Engineering*, *7*(1), 494-530. https://doi.org/10.31181/dmame7120241257

[2] Bozanic, D., et al. (2023). Ranking challenges, risks and threats using Fuzzy Inference System. *Decision Making: Applications in Management and Engineering*, *6*(2), 933-947. https://doi.org/10.31181/dmame622023926

[3] Lim, I., et al. (2023). Design, Simulation, and Prototype of an 18-Wheeler Electric Vehicle with Range Extension using Solar PV and Regenerative Braking. *WSEAS Transactions on Circuits and Systems*, *22*, 251-273. https://doi.org/10.37394/23201.2023.22.27

[4] Santos, D. and V. Santos. (2023). Power Flow Optimization with Energy Storage-Sal Island Case Study. *WSEAS Transactions on Power Systems*, *18*, 216-231. https://doi.org/10.37394/232016.2023.18.23

[5] Mousa, A. (2023). Extended-deep Q-network: A functional reinforcement learning-based energy management strategy for plug-in hybrid electric vehicles. *Engineering Science and Technology, an International Journal*, *43*, 101434. https://doi.org/10.1016/j.jestch.2023.101434

[6] Guo, J., et al. (2023). Model predictive adaptive cruise control of intelligent electric vehicles based on deep reinforcement learning algorithm FWOR driver characteristics. *International Journal of Automotive Technology*, *24*(4), 1175-1187. https://doi.org/10.1007/s12239-023-0096-4

[7] Shafiqurrahman, A., V. Khadkikar, and A.K. Rathore. (2023). Electric vehicle-to-vehicle (V2V) power transfer: Electrical and communication developments. *IEEE Transactions on Transportation Electrification*, *10*(3), 6258-6284. https://doi.org/10.1109/TTE.2023.3345738

[8] Mahdi, A.J. and D. Esztergár-Kiss. (2023). Supporting scheduling decisions by using genetic algorithm based on tourists' preferences. *Applied Soft Computing*, *148*, 110857. https://doi.org/10.1016/j.asoc.2023.110857

[9] Chen, Z., et al. (2021). Environmental and economic impact of electric vehicle adoption in the US. *Environmental Research Letters*, *16*(4), 045011. https://doi.org/10.1088/1748-9326/abe2d0

[10] Hussein, H.M., et al. (2024). A review of battery state of charge estimation and management systems: Models and future prospective. *Wiley Interdisciplinary Reviews: Energy and Environment*, *13*(1), e507. https://doi.org/10.1002/wene.507

[11] Preedakorn, K., D. Butler, and J. Mehnen. (2023). Challenges for the adoption of electric vehicles in Thailand: potential impacts, barriers, and public policy recommendations. *Sustainability*, *15*(12), 9470. https://doi.org/10.3390/su15129470

[12] Zhang, J., et al. (2022). Proximal policy optimization via enhanced exploration efficiency. *Information Sciences*, *609*, 750-765. https://doi.org/10.1016/j.ins.2022.07.111

[13] Fayyazi, M., et al. (2023). Real-time self-adaptive Q-learning controller for energy management of conventional autonomous vehicles. *Expert Systems with Applications*, *222*, 119770. https://doi.org/10.1016/j.eswa.2023.119770

[14] Tsai, J.-F., et al. (2024). Electric vehicle adoption barriers in Thailand. *Sustainability*, *16*(4), 1642. https://doi.org/10.3390/su16041642

[15] Manavaalan, G., et al. (2024). Energy saving and speed control in autonomous electric vehicle using enhanced manta ray foraging algorithm optimized intelligent systems. *Journal of Power Sources*, *619*, 235217. https://doi.org/10.1016/j.jpowsour.2024.235217

[16] Ahmadian, S., M. Tahmasbi, and R. Abedi. (2023). Q-learning based control for energy management of series-parallel hybrid vehicles with balanced fuel consumption and battery life. *Energy and AI*, *11*, 100217. https://doi.org/10.1016/j.egyai.2022.100217

[17] Divya, G. (2023). Design and modeling of hybrid electric vehicle powered by solar and fuel cell energy with quadratic buck/boost converter. *WSEAS Transactions on Circuits and Systems*, *22*, 41-54. https://doi.org/10.37394/23201.2023.22.7

[18] Wiedemann, N., et al. (2024). Vehicle-to-grid for car sharing-A simulation study for 2030. *Applied Energy*, *372*, 123731. https://doi.org/10.1016/j.apenergy.2024.123731

[19] Andrew, B. and S. Richard S. (2018). Reinforcement learning: an introduction. http://dlib.hust.edu.vn/handle/HUST/23971

[20] Watkins, C.J. and P. Dayan. (1992). Q-learning. *Machine learning*, *8*, 279-292. https://doi.org/10.1007/BF00992698

[21] Wu, H., et al. (2024). Adaptive multi-agent reinforcement learning for flexible resource management in a virtual power plant with dynamic participating multi-energy buildings. *Applied Energy*, *374*, 123998. https://doi.org/10.1016/j.apenergy.2024.123998

[22] Haarnoja, T., et al. (2018). Soft actor-critic algorithms and applications. *arXiv preprint arXiv:1812.05905*. https://doi.org/10.48550/arXiv.1812.05905

[23] Mnih, V., et al. (2015). Human-level control through deep reinforcement learning. *nature*, *518*(7540), 529-533. https://doi.org/10.1038/nature14236

[24] Tokic, M. *Adaptive ε-greedy exploration in reinforcement learning based on value differences*. in *Annual conference on artificial intelligence*. 2010. Springer. https://doi.org/10.1007/978-3-642-16111-7_23

[25] Dini, P., S. Saponara, and A. Colicelli. (2023). Overview on battery charging systems for electric vehicles. *Electronics*, *12*(20), 4295. https://doi.org/10.3390/electronics12204295

[26] Bibak, B. and H. Tekiner-Mogulkoc. (2021). Influences of vehicle to grid (V2G) on power grid: An analysis by considering associated stochastic parameters explicitly. *Sustainable Energy, Grids and Networks*, *26*, 100429. https://doi.org/10.1016/j.segan.2020.100429

[27] Cheng, H., et al. (2020). An integrated SRM powertrain topology for plug-in hybrid electric vehicles with multiple driving and onboard charging capabilities. *IEEE Transactions on Transportation Electrification*, *6*(2), 578-591. https://doi.org/10.1109/TTE.2020.2987167

[28] Issa, M., et al., *Grid integrated non-renewable based hybrid systems: Control strategies, optimization, and modeling*, in *Hybrid Technologies for Power Generation*. 2022, Elsevier. p. 101-135. https://doi.org/10.1016/B978-0-12-823793-9.00004-8

[29] Salehpour, M.J. and M. Hossain. (2024). Leveraging machine learning for efficient EV integration as mobile battery energy storage systems: Exploring strategic frameworks and incentives. *Journal of Energy Storage*, *92*, 112151. https://doi.org/10.1016/j.est.2024.112151

[30] Al-Heety, O.S., et al. (2024). Traffic Control Based on Integrated Kalman Filtering and Adaptive Quantized Q-Learning Framework for Internet of Vehicles. *CMES-Computer Modeling in Engineering & Sciences*, *138*(3). https://doi.org/10.32604/cmes.2023.029509

[31] Das, L.C. and M. Won. *Saint-acc: Safety-aware intelligent adaptive cruise control for autonomous vehicles using deep reinforcement learning*. in *International Conference on Machine Learning*. 2021. PMLR. https://proceedings.mlr.press/v139/das21a/das21a.pdf

[32] Lu, G., et al. (2025). Enhanced Dynamic Expansion Planning Model Incorporating Q-Learning and Distributionally Robust Optimization for Resilient and Cost-Efficient Distribution Networks. *Energies*, *18*(5), 1020. https://doi.org/10.3390/en18051020

[33] Alaskar, S. and M. Younis. (2024). Optimized Dynamic Vehicle-to-Vehicle Charging for Increased Profit. *Energies*, *17*(10), 2243. https://doi.org/10.3390/en17102243

[34] Priyan, S.R., Amanesh and J. Niresh. (2022). Vehicle to Vehicle charging (V2V). *International Research Journal of Engineering and Technology (IRJET)*, *9*(10), 51-57. https://www.irjet.net/archives/V9/i10/IRJET-V9I1008.pdf

[35] Regehr, M.T. and A. Ayoub. (2021). An elementary proof that q-learning converges almost surely. *arXiv preprint arXiv:2108.02827*. https://doi.org/10.48550/arXiv.2108.02827

[36] Miyamoto, K., et al. *Convergence of Q-value in case of Gaussian rewards*. in *Progress in Intelligent Decision Science: Proceeding of IDS 2020*. 2021. Springer. https://doi.org/10.1007/978-3-030-66501-2_12

[37] Sun, C. *Fundamental Q-learning algorithm in finding optimal policy*. in *2017 International Conference on Smart Grid and Electrical Automation (ICSGEA)*. 2017. IEEE. https://doi.org/10.1109/ICSGEA.2017.84

[38] Parr, R., et al. *An analysis of linear models, linear value-function approximation, and feature selection for reinforcement learning*. in *Proceedings of the 25th international conference on Machine learning*. 2008. https://doi.org/10.1145/1390156.1390251

[39] De Ath, G., et al. (2021). Greed is good: Exploration and exploitation trade-offs in Bayesian optimisation. *ACM Transactions on Evolutionary Learning and Optimization*, *1*(1), 1-22. https://doi.org/10.1145/3425501

[40] Kamthe, S. and M. Deisenroth. *Data-efficient reinforcement learning with probabilistic model predictive control*. in *International conference on artificial intelligence and statistics*. 2018. PMLR. https://proceedings.mlr.press/v84/kamthe18a/kamthe18a.pdf

[41] Cavinato, V., et al. *Dynamic-aware autonomous exploration in populated environments*. in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. 2021. IEEE. https://doi.org/10.1109/ICRA48506.2021.9560933

[42] El Fallah, S., et al. (2023). State of charge estimation of an electric vehicle's battery using Deep Neural Networks: Simulation and experimental results. *Journal of Energy Storage*, *62*, 106904. https://doi.org/10.1016/j.est.2023.106904

[43] Ghimire, K. (2022). Coil Parameter Analysis in Wireless Electric Vehicle Charging. *International Journal of Electrical Engineering and Computer Science*, *4*, 101-109. https://doi.org/10.37394/232027.2022.4.15