SCIENTIFIC OASIS

Decision Making: Applications in Management and Engineering

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

Enhanced Decision Making in Smart Grid Management by Optimizing Adaptive Multi-Agent Reinforcement Learning with Vehicle-to-Grid Systems

Pannee Suanpang[1,*] and Pitchaya Jamjuntr [2]

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

ARTICLE INFO

Article history:
Received 29 March 2024
Received in revised form 1 July 2024
Accepted 05 July 2024
Available online 10 July 2024

Keywords:
Decision making; Smart grid management; Optimization, Multi-Agent Reinforcement Learning; Vehicle-to-Grid Systems

ABSTRACT

Abstract: This research proposes a decision-making framework in which the Adaptive Multi-Agent Reinforcement Learning (MARL) model and the concept of Vehicle-to-Grid (V2G) interactivity are employed to improve the effective management of smart grids. The research hypothesis introduces innovations for improving the efficiency and security of power systems in the global south, primarily by controlling the net energy transmission between the defined electric vehicles (EVs) and the grid. Other issues that require attention to ensure the proper functioning of smart grids include demand response, load management, and energy storage optimization. In this instance, these gaps are filled by the system's proposed framework. With the help of MARL, the system dynamics' autonomous learning aspects allow the system to adapt to the capacity of renewable energy sources and electricity demand, which is also time-dependent. Because of the MARL, the autonomous coordination of decision-making has resulted in very positive changes in the system's effectiveness. In particular, this framework permitted an increase of 13.6% in the total energy exchange between EVs and the grid, and the grid stability index improved from 0.84 to 0.87 compared to what would have been achieved with the conventional methods. Enhanced energy management and pricing rehabs added another 22% to net savings. Further, it is stated that deploying MARL-based V2G systems in developing areas has many benefits, including more robust grid reliability and energy security and better integration of renewable energy resources. Such changes aid in reducing fossil fuel use and greenhouse gas emissions.

1. Introduction

The world's automotive landscape is changing rapidly with a shift toward electric vehicles (EVs) [1-2]. This evolution is supported by many streams, such as rising ecological apprehensions, the development of technology, and active state interventions to improve carbon emission rates [3-5]. Climate change, being an existential threat, has brought about an increasing emphasis on sustainable

* *Corresponding author.*
E-mail address: pannee_sua@dusit.ac.th

https://doi.org/10.31181/dmame7120241042

mobility solutions. EVs answer the call for mitigating greenhouse gas emissions and moving away from fossil fuels in energy supply. Chen et al. [6] argue that enhanced usage of EVs globally, mainly using renewable sources for charging, can provide extensive savings on CO2 emissions globally.

Simultaneously, the demand for EVs is increasing due to the awareness and availability of various models [7]. With more automotive manufacturers devising new models, the demand for vehicles is anticipated to rise significantly and redefine the automotive market [7][8]. Additionally, EVs are beginning to be complemented by advanced technologies such as autonomous driving and connected car systems, which promise to improve the prospects of EVs and transform transportation [9]. In smart cities, EVs shall be able to operate within renewable energy sources smoothly, smart grids, and self-sufficient ecosystems for responsible and effective transportation solutions [9][10]. This is especially relevant for developing countries experiencing increasing pressure from the growing world population, urbanization, and economic development to cope with their energy resources sustainably [9][11]. This creates a need for sustainable energy options, which ensures that consumption is healthy for the environment and economy [9][11][12][13].

On the other hand, adopting EVs allows for addressing some of these barriers, especially through Vehicle-to-Grid (V2G) technology. V2G allows EVs to participate in bi-directional energy transfer with the grid, increasing the capacity of the EVs as energy storage and supply devices, which helps balance the grid and improve the integration of renewable sources [14-16]. On the contrary, the V2G concept suffers from technical implementation issues, including battery aging, communication interface issues, and grid facilities requirements. Battery cycle aging arising from the constant cycling of the batteries towards grid services and management remains an issue. These metering systems must be upgraded to allow energy transfer and communication in two directions [15][16]. The widespread adoption of V2G technology can also help promote energy equity in developing economies. V2G can help improve energy consumption efficiency, hence flattening the curves of energy demand and reducing the instability of power supply systems, which affects less privileged communities more. Also, the number and dependency on fossil fuels can be reduced with the application of V2G, thus aiding the sustainability goals of the state [14][15][17].

Controlling the dynamic relations between EVs, the grid, and other parties takes considerable effort. Conventional styles of controlling lies cannot effectively cope with the V2G elements, which are emerging and complex simultaneously. Multi-agent reinforcement Learning (MARL) is a method that has proven valuable in enabling local decision-making and adaptive learning in versatile structures containing several agents. MARL can optimize V2G interactions in regions known for their intermittent and unstable energy systems [18-20]. Implementing MARL under V2G configurations can optimize energy security and improve the use of renewable resources while providing reliable energy for all within developing nations. This solves short-term energy management problems and complements the attainment of long-term sustainability objectives.

1.1 Problem Statement

The increasing adoption of EVs across the globe is a blessing and a curse about energy management, especially for developing nations. The ongoing upsurge in EVs would also raise possibilities for deploying these vehicles as mobile power sources using the V2G system. V2G systems permit EVs to connect to the grid, enabling them to obtain or return electricity. In places increasingly integrating renewable energy, such as solar and wind power, this can be essential for the grid's stability [21][22]. However, integrating V2G technology with existing energy grids is certainly not easy. Techniques for efficient energy control, grid stability, and economical operational modes should be developed to facilitate its integration [23].

The growing global use of EVs has pros and cons regarding energy management, but this is more critical for developing countries. The ongoing high increase in the use of EVs will also create opportunities for deploying these vehicles as mobile power sources using the V2G system. V2G systems allow EVs to interface with the grid, meaning these vehicles can receive electricity or return power to the grid. In regions that have begun incorporating renewable energy sources, such as wind and solar development, this can be important for supporting the grid [21][22]. However, the challenge of embedding V2G technology within existing energy grids is perhaps the most frustrating issue. Energy system integration, economic and operational modes, and grid stability should be improved to allow for its incorporation [23].

MARL offers a potential way forward in this regard because it allows the decentralization of decision-making and adaptive learning in the context of a complex environment. By allowing every agent representing an EV [18-19], charging station, or grid component to learn and optimize its actions in real-time, MARL can assist in overcoming the drawbacks of conventional grid management models [1][5][27]. Nevertheless, the use of MARL in managing V2G and smart grid in developing countries is still a virgin area with great potential for improving energy resilience, enhancing renewable resources' utilization, and promoting energy equity [28][29]. In this regard, this study seeks to fill this gap by investigating the use of MARL in V2G interactions and smart grid management in Thailand. This will help achieve more resilient, efficient, and sustainable energy systems in developing countries.

1.2 Research gap and challenges

An issue that seems to plague research in this emerging area is the application of advanced management strategies like MARL, especially in developing countries such as Thailand. While there is an increasing interest in V2G technology and its possible role in improving grid reliability and energy efficiency, the literature points to a deficit of studies that explore MARL applications within this context [30]. Most of the studies carried out before were limited to developed areas that are more energy-rich and have advanced smart grid technology. It is, therefore, pertinent to note that there is considerable research on the potential application of MARL in optimizing V2G interactions in underdeveloped countries where energy systems are unstable and stretched [31,32].

This area faces one major constraint the advanced intricacies and inherent dynamism of V2G systems. In some developing countries, energy demand and supply are rather ready to be compromised because energy grids are more or less weak, forcible fluctuations in demand and supply caused by underdeveloped facilities to sustain sustainability, unreliability of energy output from renewable sources, or other economic factors [23]. This complexity poses a challenge in employing the existing grid management technology, which is traditionally centered, considering that there is greater real-time decentralization in V2G interactions. Interestingly, even the grid integrated with EVs creates more chaos or uncertainty since the EVs' availability or participation depends upon several factors, such as charging habits, user types, transfer requirements, and others [24][33].

Another challenge is associated with the MARL algorithms' scalability and adaptability. Even though MARL is a promising framework for decentralized decision-making that allows for multiple autonomous agents, its practice concerning V2G systems usage is still underdeveloped. The scalability of MARL has not been validated; however, it is used for the automation of large-scale distributed energy resources such as EVs within the V2G system [33-34]. Moreover, the adoption of MARL to the specific circumstances of developing countries characterized by limited data, information technology power, and infrastructure for technology adoption is another challenge that should be tackled [35].

In addition, the socio-economic and regulatory contexts in developing nations introduce another level of difficulty. The deployment of V2G technology and MARL solutions, for instance, calls for the availability of technical support, policies, legal frameworks, and economic incentives which as aforementioned are sometimes absent in these areas. It is vital to comprehend the non-technological aspects in relation to the practical realization of V2G and MARL systems in order to secure the viability of such systems and their effectiveness [27-28] [36-37].

Table 1 Summary of Problem Statement and Research Gaps

Aspect	Problem/Challenge
V2G Technology Integration	Complex integration of V2G into existing energy infrastructures requires advanced management strategies to ensure stability, efficiency, and cost-effectiveness [38].
Developing Countries' Challenges	Inadequate grid infrastructure, frequent power outages, and reliance on fossil fuels make V2G integration more challenging.
Thailand's Energy Landscape	Rapidly increasing energy demand with a national commitment to integrating renewable energy requiring innovative V2G solutions [39].
MARL for V2G Optimization	MARL offers the potential for decentralized decision-making in complex energy systems but remains underexplored in developing countries [40-41].
Scalability and Adaptability of MARL	Limited validation of MARL's scalability in managing large-scale V2G systems, particularly in resource-constrained developing countries [18-19][42].
Non-technical Challenges	Socio-economic and regulatory factors play a significant role in successfully deploying V2G and MARL systems in developing regions [43-44].

1.3 *Objectives and Scope of the Paper*

This study aims to address these research gaps by analyzing how MARL can be implemented in V2G systems in Thailand – a developing country setting. The study will assess how MARL can be modified to suit the V2G interaction management in settings with limited infrastructure, scarce financial resources, and regulatory restrictions. The study attempts to fill this knowledge gap by assessing the use of ADMARL in enhancing the V2G systems in developing countries, supporting the larger picture of improving energy systems in a sustainable manner all over the world [1][5][18][19].

The focus of this paper is to examine the application of adaptive MARL as an approach to enhance interactions of vehicles with the grid (V2G) and the management of the smart grid in Thailand. The specific objectives are:

1. To create a MARL centered on the V2G system, taking into account the peculiarity of Thailand's energy system.
2. To determine the parameters of the V2G interaction and develop an environment for their simulation that includes state spaces of grid management, action spaces, and reward functions.
3. To analyze the efficiency of the presented MARL system in optimizing V2G interactions, including the comparison with classical approaches to optimization.
4. Interpret the results in a way that relates them to Thailand's present and future smart grid and energy management structure in terms of practical and policy changes that can be made.

The scope of the paper covers the analysis of the available resources on V2G technology, MARL for energy management, and related literature available in Thailand. It also includes designing and building a simulation framework, metrics and comparisons to evaluate performance, and comprehensive analyses of the results obtained about the energy situation within Thailand.

1.4 *Contribution*

This research makes several key contributions to the field of smart grid management and V2G technology, particularly within the context of developing countries. The study focuses on the innovative application of MARL to optimize V2G interactions, with a specific case study in Thailand. The contributions of this research are as follows:

Application of MARL to V2G Optimization: While existing literature has explored V2G technology and smart grid management in various contexts, the application of MARL to optimize V2G interactions remains underexplored, especially in developing countries. We undertake this effort to bridge this gap by illustrating the application of MARL in dealing with the challenges posed by the complex dynamics between EVs, the grid, and other energy actors. The study employs MARL to demonstrate how the distributed decision-making approach in V2G systems improves grid system stability, energy efficiency, and renewable energy integration [15] [45-48].

Focus on Developing Countries: V2G and smart grid systems management has been the subject of study in mostly developed countries where the energy infrastructure is well integrated. However, attention in this study is directed to developing countries, where the energy systems are even weaker and likely to be interrupted. Casting the net to Thailand, in this case, helps to analyze the challenges and advantages of injecting V2G into areas with developing energy systems. This contribution is relevant as the developing countries are increasingly looking for solutions to the energy problem not based on fossil fuels [26] [49-58].

Support for the Sustainable Energy Goals: The assessments confirm and complement initiatives oriented on the energy sector best practices in the context of the United Nations' Sustainable Development Goals (SDGs), particularly the Clean Energy Goal. The research helps encourage the pursuit of sustainable and resilient energy systems as it illustrates how the use of MARL can make V2G interactions more energy efficient, enhance the integration of green energy sources, and reduce gas emissions. This is of particular concern for developing countries such as Thailand where sustainable energy has become one of the key objectives of the national development strategy [24][58].

Practical Implications for Policy and Industry: The conclusions of this study bear practical importance for policymakers, energy suppliers, and the EV market. Looking at the positive and negative aspects of MARL application for V2G exploitation, this research may nurture the V2G concepts in the policy and strategy making for the attendant welcoming of the technology [59-60]. In addition, the study provides industries with ways to implement MARL-based approaches within their firms effectively, thus improving the performance and reliability of energy systems in developing nations [60-62].

This study also offers interesting contributions about specific technical, socio-economic, or regulatory aspects that must be considered to realize sustainable and resilient energy systems and their impact on theory and practice [63-66].

2. **Literature Review**

2.1 Overview of V2G Technology and Smart Grid Systems

Vehicle-to-grid (V2G) technology has captured the attention of many as energy innovation on the lower end of the wave. It allows EVs and power grids to connect in both directions regarding energy exchange. Through this integration EVs can capture surplus energy and send it back to the grid during high energy consumption, making the grid more effective and secure. V2G systems are much more versatile as they help control frequency, balance load, and shear off-peak demand therefore making power systems more robust [67–68].

Considering such smart grid systems, advanced information and communication technologies (ICT) are embedded, allowing monitoring and management of electricity flow in real-time so that there is more balance within the grid. This feature enables the utilization of renewable resources such as solar wind and other distributed energy resources [69-70]. It also shifts how the power supply is used as smart grids allow for subject control to power demand, cutting back further on fossil fuel usage while improving energy security [71-72]. An integrated approach to V2G technology and smart grid systems assists in energy flow management, improves energy generation efficiency, and encourages energy savings. Studies have shown that V2G systems, in coordination with smart grids, can significantly help use intermittent renewable energy and balance energy demand [73-74].

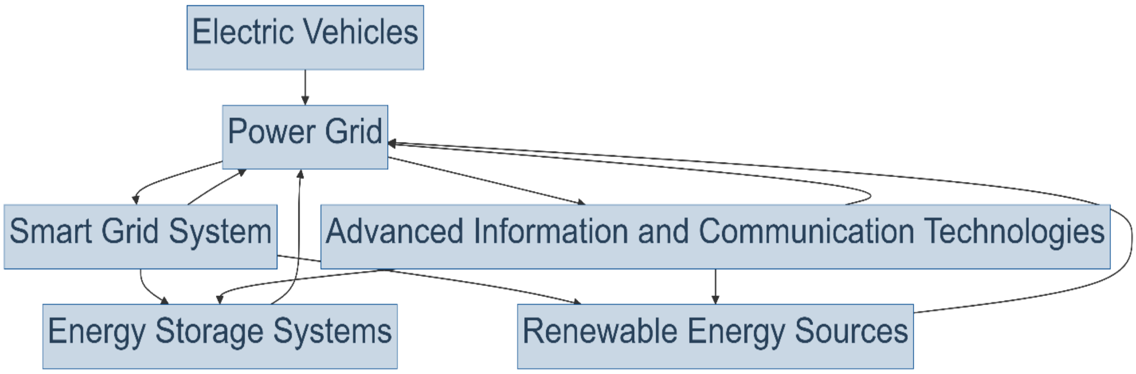

Figure 1. Vehicle-to-Grid (V2G) Interaction Overview

Figure 1 presents several constituents' interaction schemes that work as a single energy system unit. Electric vehicles are also connected to the power grid system, which is the central connecting element of the whole system so to speak. Communication and some other new technologies of modern times are integrated within the structure of the power grid in order to facilitate its operation and communication. These information technologies advance the system to the extent that the smart grids, the energy storage technologies, and the grid are connected to them. Moreover, the sophisticated technologies and the power grid also enable the integration of advanced energy resources into the system. This interrelated evolution brings out transformation and the continuous evolution of energy systems.

Interactions within a smart grid system through V2G technology are represented in Figure 2. In this case, EVs are directly interfaced with the grid in such a way as to allow energy flow in both directions [1]. The power grid is augmented by an intelligent system that is enhanced by smart management features. This smart grid works in conjunction with RESs, thus increasing the grid's ability to assimilate energy generation that is intermittent. Energy storage systems are also built inside the grid to pack surplus energy and retrieve it when required. Advanced ICT implements and addresses real-time data transfer processes, operative guidance, and management and control of several interactions with ancillary renewable generators and storage. V2G technology is visualized in the diagram as combined with other components, represented with arrows indicating the course of their interactions and how these components are interdependent in the general V2G system [45][60][63][74-75].

The recent developments made in V2G technology emphasize its advantages which include improvement in grid performance, decreased energy expenses, and maximized use of renewable energy sources [76-77]. As a backbone for achieving any of these benefits, smart grids with advanced metering infrastructure and intelligent control systems allow for better integration of the grid and EVs [78].

Figure 2. Example of smart grid system by V2G

2.2 Existing Optimization Approaches in V2G Interactions

The second part is the V2G interaction optimization, which consists of creating EV charging and discharging strategies to bring maximum profit for the grid and EV owners, respectively. Conventional optimization methods engage mathematical programming, heuristic algorithms, and model predictive control (MPC) [16, 40, 41, 56,68]. These methods seek to achieve multidimensional goals such as cutting down expenses, lowering peak demand, and increasing the share of green energy.

In particular, V2G optimization problems have been formulated and solved with the help of mathematical programming techniques such as Linear Programming and Mixed Integer Linear Programming methods (MILP). For example, Wang et al. [60] presented a mixed-integer linear programming model to optimize EV charging schedules across time under dynamically priced electricity and the availability of expanding renewable energy sources. MILP has also been used to plan grid integration under uncertain renewable energy generation and efficiently support supply-demand management [78-79].

Heuristic algorithms, specifically genetic algorithms (GA) and particle swarm optimization (PSO), are also presented as solutions since they can work with a broader solution space and can provide satisfactory results with less effort [31][51][64][65]. Such studies later used hybrid techniques to enhance GA such as integration with different optimization procedures [80]. These techniques are used when it is necessary to solve several interrelated tasks simultaneously, for example, when the intended charging cost is to be minimized and the peak demand on the grid is also constrained.

Model predictive control (MPC) is increasingly well-liked because it is well-suited for the dynamic and uncertain character of V2G relationships. According to the MPC frameworks, the current and future stages of the system are estimated to maximize control efforts. This feature allows for real-time implementation of the control action [81 – 82]. MPC has an edge since it optimizes control based on the variability in grid demand and energy prices on an online basis. Additionally, it has been coupled with deep learning models for better prediction and more optimization of control actions mäki [83]. Though most of these traditional methodologies are effective, they are computationally intensive and are unlikely to be easily scalable in the case of increasing EVs and the intricacy of smart grid systems. Recent developments in machine learning and artificial intelligence are, amongst other things, being investigated as solutions to these problems, enabling more scalable and adaptable strategies for V2G optimization [84].

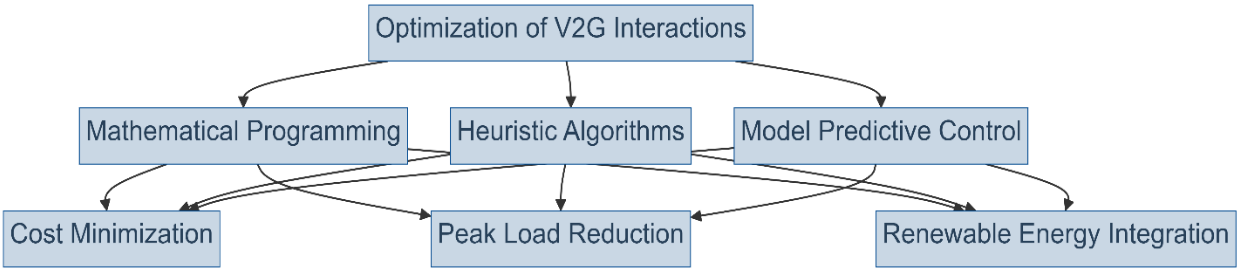

Figure 3. Optimization Approaches for Vehicle-to-Grid (V2G) Interactions

The optimization techniques aimed at the V2G interactions have been shown in Figure 3. The outer rim includes the core node due to three sub-categories of his ‘Optimization of V2G Interactions’ covered during the study, namely ‘Mathematical applications, Heuristic, Model Predictive application’. There are three targets for each "Cost inimization” reduction of peak load, and ‘Development of renewable resources inclusion.’ V2G systems can also benefit from these approaches, but it appears that there are other benefiting objectives of these methods, which are also system-orientated.

2.3 Applications of MARL in Energy Management

MARL offers a promising alternative to traditional optimization methods by leveraging the principles of reinforcement learning to enable agents (e.g., EVs or charging stations) to learn optimal strategies through interactions with the environment and other agents [1][5][18-19]. Research into MARL approaches has been revealing progress on the challenges of scaling and the complexity of the V2G optimization problem [85-86].

In the sphere of energy management, MARL has been used in various applications, including demand response, distributed energy resources management, and grid stability [86-87]. For example, the work done by Wang et al. [60]. developed a MARL-based framework for optimizing EVs' charging and discharging schedules, noting that grid stability and energy cost are improved dramatically relative to traditional methodologies. Likewise, Zhao et al. 2020 also utilized the MARL approach in coordinating the operations of distributed energy resources with improved system resilience and operational efficiency [38][41][46][49][88].

Newer works have broadened the scope of the use of MARL in the management of energy resources by introducing more complex interactions, such as V2G coordination in large-scale smart grid systems. For instance, a three-level MARL framework was developed to distribute energy among the multiple EV charging stations, improve load balancing, and reduce operational costs [89-90]. It also polled the incorporation of deep MARL to allow EVs to self-determine the optimal charging time that emanates from electricity prices and grid conditions, demonstrating enhanced scalability and adaptability in dynamic environments [91].

The second promising area of development can be microgrid operations optimization using MARL. In this regard, the study carried out by Han et al. [23] illustrated how MARL could enhance the remote microgrid integration of CDREs and battery storage systems for efficient energy use and interruption resiliency. In addition, some studies have utilized cooperative MARL to strategize the charging pattern of a fleet of autonomous EVs, resulting in improved coordination and reduced energy expenditure [92].

MARL's ability to be robust in dynamic environments and to learn about optimal policies in a distributed fashion makes it relevant for V2G interactions. Since policies are adapted based on the target status and input interactions at the current time, MARL agents are capable of quickly adjusting to fluctuations in demand, electricity prices, and the availability of renewable energy sources [49][68][93].

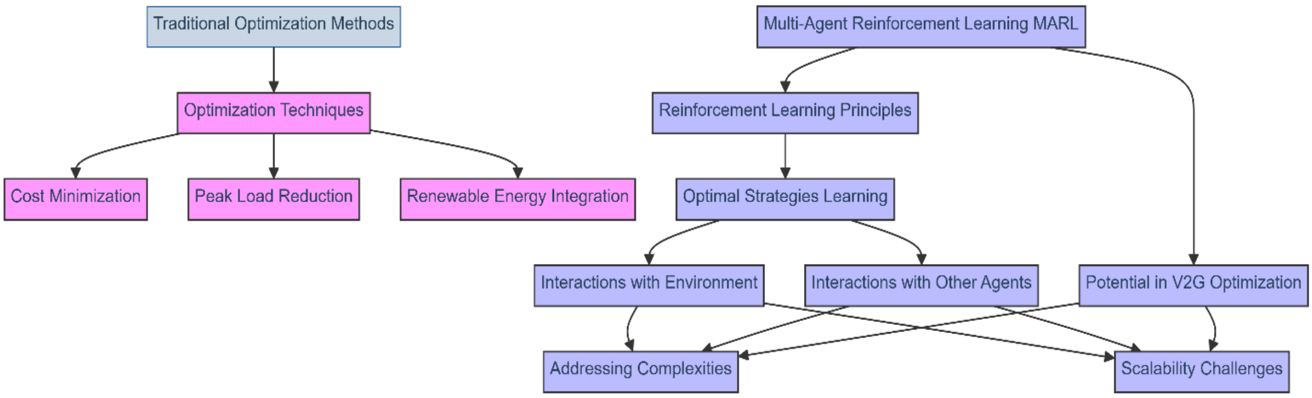

Figure 4. The Comparison of Traditional Optimization Methods and MARL in the Context of V2G Systems

Figure 4 compares the traditional optimization strategies to the use of MARL for V2G systems. Further development of the figure shows that traditional optimization methods are related to specific goals like cost reduction, peak shaving, or renewable energy integration. In contrast, MARL is correlated via reinforcement learning concepts to optimal strategy development based on interaction with the environment and other agents. MARL resolves issues of complexities and limitations in scale, thereby underscoring its advantages in V2G optimization. Color classes were used within the image to help differentiate the approaches [67][94].

2.4 *Case Studies or Related Work in Thailand*

V2G technology and MARL are becoming popular in Thailand as the country aims to improve its energy sustainability and resiliency of the electric grid. There have been several studies on the opportunities and challenges of V2G operationalization in Thailand. For instance, some studies assessed V2G for grid stability, showing the ability to use peak load shaving and nesting renewable energy [95]. Another study also reported the economic feasibility of V2G in Thailand, which pointed out some structural and policy changes that are vital to unlocking the potential of V2G [96]. Studies on MARL applications within the energy sector in Thailand are quite limited; however, preliminary studies suggest a good trend. Many Thais proposed a MARL-based methodology for managing distributed resources in Thailand, achieving increases in system performance and economy [96-97]. In the same context, Thai researchers employed MARL to facilitate demand response management and demonstrated with higher flexibility and responsiveness than the traditional methods [98].

More studies have built on these conclusions looking into some of the applications on V2G and MARL systems in Thailand. For instance, Wattana & Wattana [99] examined the inclusion of V2G systems in the Thai city infrastructure with a special focus on the islands of the smart grid system into energy distribution networks. Besides, this research work will appraise the V2G system deployment progress and plans in rural Thailand and how the socio-economic, infrastructural and institutional factors affect it [98].

A different experiment, which is also remarkable, is the one that Suanpang and Jamjuntr conducted [100] in which they demonstrated how effective MARL algorithms are, particularly the Multi-Agent Deep Deterministic Policy Gradient (MADDPG), in trying to optimize the locations and the distribution of EV charging stations in smart cities. Their findings prove that MARL achieves significant improvements in mean charge waiting time, increases charging facility turnover (CFT), and boosts total savings, all of which lead to greater efficiency and satisfaction in EV charging infrastructures. This approach is consistent with ours, which seeks to employ Adaptive MARL on V2G interactions. The MDPG algorithm being by design collaborative and communicative, as noted by Suanpang and Jamjuntr [101], is most applicable in V2G scenarios involving multiple active agents such as EVs, charging stations, and grid operators to enhance energy distribution and cost efficiency while maintaining grid stability. Adopting the same MARL approach, our work seeks to widen the scope and attempt to look into the problems developing countries encounter in establishing V2G systems, such as infrastructure deficits, variability in the energy supply, and cost efficiency.

In particular, the MARL techniques in regard to the encoding of local culture and distinct energy issues in Thailand were examined Zhou et al. [102] reported prospects of energy management enhancement through MARL application in the urban as well as rural territories thus making it a practice worth referencing in subsequent studies and strategic frameworks. Generally, these researches emphasize the potential of MARL and V2G technologies in changing the energy picture in Thailand for good. In particular, they lay the ground for applying Adaptive MARL in V2G interactions, calling for more efforts to address the peculiarities of the Thai environment and the opportunities therein [103].

2.5 Deterministic Optimization

Deterministic optimization is a discipline that deals with optimization processes whose decision-making is devoid of random variables and uncertainties. It stands in contrast with stochastic methods as it does not assume partial ignorance of any parameters and the system functions satisfactorily with the particular parameters. This approach is especially useful when accurate and feasible solutions are needed [100-103].

2.5.1 Overview of Deterministic Optimization Techniques

Deterministic optimization techniques are commonly divided into three major categories: linear and nonlinear programming, integer programming, and dynamic programming. These techniques are appropriate for strategies aimed at optimization problems whereby the objective function and constraints are deterministic functions of the decision variables.

Linear programming (LP): is an optimization process which aims to find the optimal value of a linear objective function subject to multiple linearly constraints. It finds applicability in different areas like resource allocation, logistics and in manufacturing. The Simplex method and interior-point methods are some of the algorithms more frequently used to solve LP problems in practice [103-104].

In nonlinear programming (NLP): one has to optimize nonlinear objective functions and/or constraints. One is contemplating the use of NLP methods when looking at the interactions of decision variables and finding them to be nonlinear. Various methods exist to solve NLP including the gradient descent, Newton's method, and sequential quadratic programming (SQP) for manipulating such problems [105].

Integer Programming (IP) can be defined as a class of optimization problems where some or all of the decision variables are constrained to have integer values. This approach can be important in combinatorial optimization problems, e.g. scheduling and routing. There are applied such algorithms as branch-and-bound and branch-and-cut to devise solutions to integer programming problems [106-107].

Dynamic Programming (DP) follows another reasoning: it solves problems assuming multiplicity of decision steps. It involves separating complex problems into less complex subproblems and determining solutions to these subproblems. The working principle of optimality by Bellman is a central concept in dynamical programming [108].

2.5.2 Applications in Energy Management

In energy management, deterministic optimization techniques have been employed to solve various problems, such as energy scheduling, load forecasting, and grid management. For example, LP has been used to optimize the operation of power systems by minimizing generation costs while meeting demand and operational constraints [109]. NLP has been applied to model and optimize complex energy systems with nonlinear behavior, such as integrated renewable energy systems [111]. IP techniques are often used in tasks like optimal placement of energy storage systems and network design [112]. DP is useful in energy management for optimizing energy usage over time, taking into account future states and decisions.

2.5.3 Limitations and Challenges

One essential understanding that intelligent heuristics brings to the forefront is the acceptance that specific problems cannot be resolved through deterministic optimization but rather through the problem-solving strategies adopted by people. There are certain limitations to these deterministic optimization approaches despite their advantages. For instance, their implementation requires accurate and specific data that may not always be feasible or precise in real-world situations. Meanwhile, there are challenges relating to the scalability of deterministic techniques and algorithms when handling high-level problems with complicated restrictions. Therefore, it is inevitable that we will resort to stochastic or heuristic optimization along with the deterministic approach in most practical applications.

2.6 Heuristic Approaches

Heuristic approaches are problem-solving methods that use practical, non-optimal strategies to find sufficiently good solutions within a reasonable timeframe, particularly for complex or large-scale problems where traditional optimization techniques may be infeasible. Unlike exact optimization methods that guarantee an optimal solution, heuristics aim to find a good enough solution based on experience, intuition, or rules of thumb. These approaches are especially useful in scenarios where the problem space is vast and the exact methods are computationally prohibitive.

2.6.1 Overview of Heuristic Approaches

Heuristic methods are diverse and include several prominent techniques, such as:

Greedy Algorithms: make a series of choices, each of which looks best at the moment, with the hope of finding the global optimum. They are simple to implement and often yield good solutions quickly, but they do not guarantee the global optimum. Examples include Kruskal's algorithm for finding the minimum spanning tree and the Huffman coding algorithm for data compression [112].

Genetic Algorithms (GA): are inspired by the principles of natural evolution. They use mechanisms such as selection, crossover, and mutation to evolve a population of solutions toward an optimal or near-optimal solution. GAs is particularly effective for complex optimization problems with large search spaces. Applications include scheduling, routing, and engineering design [113].

Simulated Annealing (SA): Simulated annealing is a probabilistic technique inspired by the annealing process in metallurgy. It explores the solution space by probabilistically accepting worse solutions as it searches for a global optimum. SA is well-suited for problems where the search space has many local optima. It has been successfully applied to problems like the traveling salesman problem and function optimization [114-115].

Tabu Search: is an iterative algorithm that guides the search process by maintaining a list of recently visited solutions, known as the tabu list, to avoid revisiting them. This approach helps to escape local optima and explore the search space more effectively. Tabu search is used in various optimization problems such as scheduling and resource allocation [116].

Ant Colony Optimization (ACO) is inspired by ants' foraging behavior and uses a colony of artificial ants to explore the solution space. Ants deposit pheromones on paths that lead to better solutions, which guides the search process. ACO has been successfully applied to combinatorial optimization problems, including routing and network design [117-118].

2.6.2 Applications in Energy Management

Heuristic approaches are increasingly applied in energy management to tackle complex optimization problems where exact methods are computationally infeasible. For instance:

- Genetic algorithms have been used to optimize the scheduling of power generation units and the configuration of energy systems, aiming to balance efficiency and cost [85][119].
- Simulated Annealing has been employed to solve unit commitment problems in power systems, where the goal is to schedule the operation of generation units to meet demand at minimum cost [120].
- Tabu Search has been applied to optimize energy storage systems and demand response strategies, focusing on improving system performance while avoiding local optima [121].
- Ant Colony Optimization has been used for routing and scheduling problems in smart grids, enhancing the efficiency of energy distribution and resource allocation [118].

2.6.3 Advantages and Limitations

Heuristic methods offer several advantages, including their ability to handle large and complex problem spaces and their flexibility in adapting to different problems. However, they also have limitations, such as not guaranteeing optimal solutions and potentially requiring careful parameter tuning to achieve good performance [118-120].

2.7 Related Study

2.7.1 Related Studies

In optimizing V2G systems and smart grid management, several studies have explored various aspects of these technologies, often focusing on different methodologies, case studies, and applications.

A number of research works have focused on the optimization of V2G systems, increasing the stability of the grid, lowering the price of energy and the integration of renewable energy sources. By way of example, Zhang et al. [40][49][68] developed an optimization model of V2G systems using dynamic pricing and demand response management to control the load and minimize the costs of operating the system. Their framework made it possible to enhance the reliability and efficiency of the grid with appropriate energy management of EVs' bidirectional energy flow.

In the same way, Liu et al. [38] investigated the integration of the V2G technology into smart grid systems for improved energy storage and distribution. They also used optimization approaches to coordinate the EV charge and discharge schedules, leading to efficiency meeting engagements 8 6 0 and peak load reduction. Their investigations underlined the activities of V2G in leveling the energy demand and increasing the operational efficiency of the grid system [12][36][38].

2.7.2 Application of MARL in Smart Grids

Among the solutions developed for the smart grid management issues, one of the most successful is the MARL. For example, Aladdin et al. [87] devised a MARL method for energy management in smart grids, integrating a decentralized approach with multiple autonomous agents optimizing energy distribution and consumption. Their method worked well on energy systems' dynamic and stochastic characteristics, thus enhancing the grid's operational efficiency and reliability. In another work, Chen et al. [31 - 32] applied MARL principles to the distribution of numerous intelligent energy managers in a smart grid for optimized operational characteristics of the distributed energy resources (DERs). They showed how MARL presented an optimal way of coordinating the operation of DERs so that energy costs were lowered while the system's resilience was improved compared to conventional methods. Their works emphasized the taming of complexity in multi-agent environments and the capability of MARL to respond to changes in the environment in real time [64 - 65]

2.7.3 Case Studies in Developing Countries

The integration of V2G and MARL technologies within the context of developing countries is the focus of the latest studies, which have produced promising results. When considering the case of Thailand, attempts have been made to identify such problems and opportunities when these technologies are rolled out within the country. Preedakorn et al. [95] performed an analysis of V2G effects on grid stability in Thailand, offering prospects of lowering peak load demand and enhancing the use of renewable sources of energy. Their research also pointed to the absence of favorable policies and the development of supporting infrastructure as the reasons why V2G has a limited impact within the Thai energy structure. The two came up with studies on V2G implementation possibilities in Thailand, stressing insufficient policy and infrastructural support. Their studies pointed to the expected possibility of V2G usage, which would enhance energy access conditions as well as the integration of renewable energy sources, but they also emphasized that proper addressing of the challenges at hand is necessary [26][28][95][99].

The current work extends the coverage of Suanpang and Jamjuntr [1] [5] [100-101] work, where they developed the EV charging optimization using Multi-Agent Reinforcement Learning with emphasis on the Multi-Agent Deep Deterministic Policy Gradient algorithm. The potential of the MARL stands out in integration of EV charging station placement and EV charging station allocation in smart cities such that the charge waiting time is less, charging facilities turnover better, and savings are higher. In this case, they justify Adaptive MARL for use in V2G interactions noting the cooperation involved in MADDPG which will enhance the proper coexistence of all agents, EVs, Charging Stations and Grid Operators among others. Although similar MARL techniques are suggested in the paper, it is intended to solve the problems typical for developing countries like Thailand in the implementation of V2G systems. Such problems are associated with infrastructure, energy supply stability and prices. It should be noted that the outcomes from Suanpang and Jamjuntr [1] [5] [100-101] corroborate the potential for Adaptive MARL in V2G applications and provide the main justification for conducting this study. It emphasizes that the MARL enables the improvement of the functional and ecological efficiency of transport systems and energy systems within intelligent cities, which is crucial for developing countries. Efficient management of the smart grid, along with management of V2G interactions, helps achieve energy access, integration of renewable energy, and fortification of the grid.

3. Methodology

3.1 Research Framework

Figure 5 presents a diagram of the framework of the study, together with an explanatory description, which is very significant in this regard. It allows the program to be designed to course the boundaries of the research, beginning from the challenge at hand, V2G technology, and smart grid management issues in the context of integrating developing countries. Next, the framework identifies the relevant newspaper studies and the issues that need to be studied, focusing on new approaches such as (MARL). Given this situation, essential research purposes are set to fill in these gaps and determine the study's outcome. The research focus, which is on MARL, as the key component of the methodology, is intended to solve problems associated with V2G optimization and smart grid management in the context of developing economies such as Thailand. The framework further cross-cut other aspects, such as social-economic and policy ones, to formulate technically sound and economically benefitable solutions. An actual setting in Thailand is employed to experiment and verify the effectiveness of MARL in a true-world scenario. The analysis and the results also aim to appreciate MARL's role in the grid's stability and energy efficiency. The expected outputs are expected to enhance the achievement of the pyramidal goal of creation, which emanates from reducing emissions, increasing the use of renewable sources, and improving efficiency. Finally, the research has practical contributions to policies and the industry by providing a roadmap for implementing V2G technology and smart grid management in developing countries, allowing visibility of how it would work in practice.

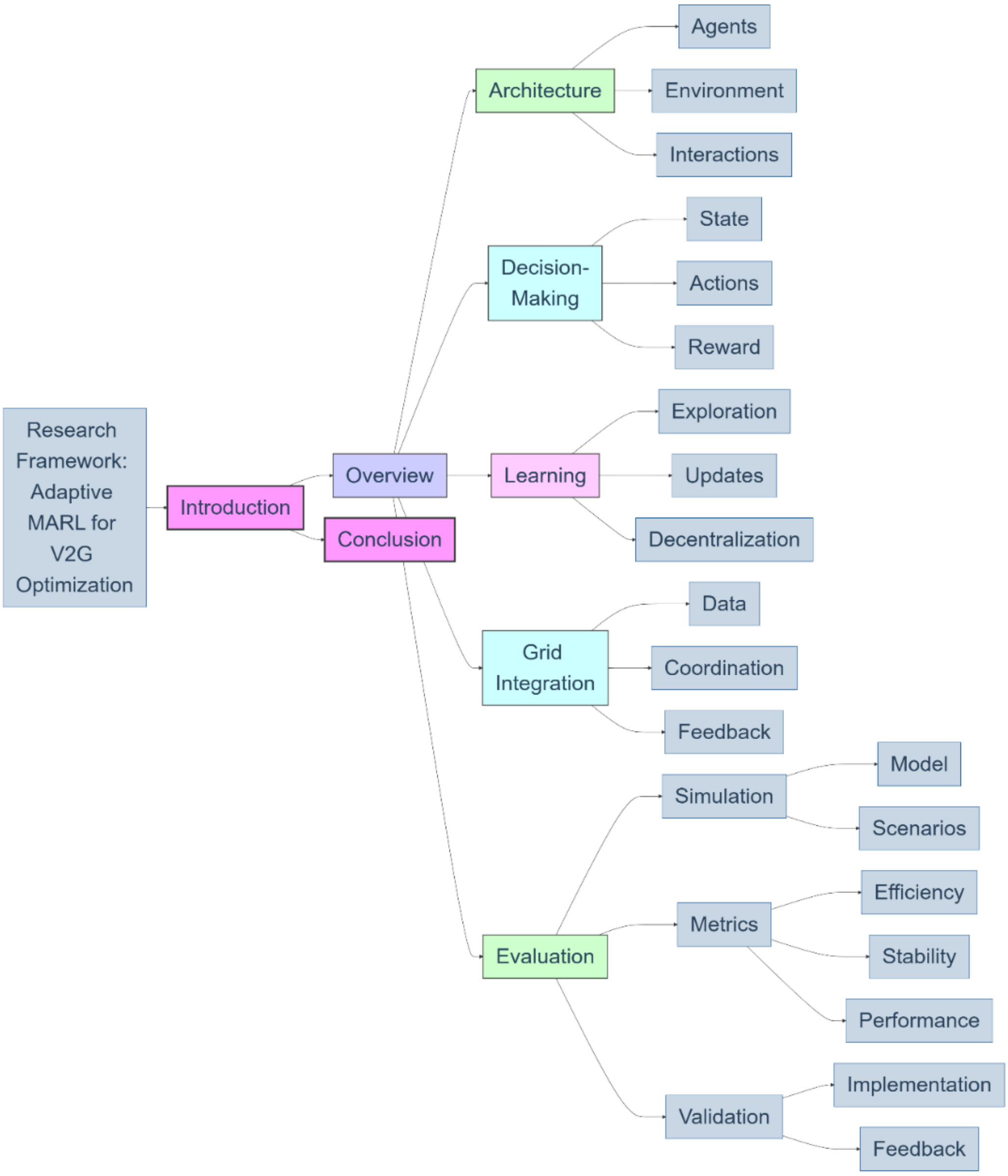

Figure 5: Research Framework for Adaptive MARL in V2G Optimization

3.2 Description of the MARL Framework

To investigate the V2G interactions in this research, agents using the decentralized MARL framework will be developed. The agents can either be EVs or charging stations, and they learn to take actions that will optimize cumulative rewards over time.

At any time step (t), Agents' observations include the grid demand (D_t), EV battery state (B_t), and the energy cost (P_t). Every agent then makes an observation of the current state and chooses an action (a_t) from its action space ($\mathcal{A} = \{\text{charge,discharge,idle}\}$).

Afterward taking action (a_t), the environment transitions to a new state ($D_{t+1}, B_{t+1}, P_{t+1}$) and the agent receives a reward (R_t) according to the reward function $R_t = f(D_t, B_t, P_t, a_t)$. The aim is to obtain clarity on policies that will increase the expected cumulative utility over time.

3.3 Environment Design for V2G Interactions

The environment allows simulating dynamics such as grid load lay, EV charging consumption, or energy price changes. Such dynamical processes are represented in the model as stochastic processes or derived from the observation data:

- Grid demand (D_t) is impacted by usage and outside elements.
- EV battery level (B_t) fluctuates according to whether charging or discharging has occurred.
-Energy prices (P_t) are not constant change in relation to a situation in the market.

The nature of the environment is represented mathematically through equations of state evolution dynamics given as:

$$D_{t+1} = g_D(D_t) \tag{1}$$

$$B_{t+1} = g_B(B_t, a_t) \tag{2}$$

$$P_{t+1} = g_P(P_t) \tag{3}$$

where $g_D(\cdot)$, $g_B(\cdot)$, and $g_P(\cdot)$ define the functions controlling demand, battery level, and energy prices, respectively.

3.4 Agent Architecture and Learning Algorithm

Every agent employs a Deep Q-Network (DQN) in a bid to approximate the action-value function $Q^{\pi}(s, a)$, where s is the state, and a is the action. The DQN architecture is designed around a fully connected neural network model comprising:

- Input layer: State Representation ($s = (D_t, B_t, P_t)$)
- Hidden layers: Dense layers with $ReLU$ activation functions
- Output layer: Action values for each action in ($\mathcal{A}$)

While training, agents can utilize experience replay by storing and sampling experiences in the form (s, a, r, s'), from the replay buffer. The fusion of target network prediction and temporal difference error is done by adjusting the network's weights in the Q-learning update rule.

$$\Delta\theta = \alpha \left[r + \gamma \max_{a'} Q\,(s', a'; \theta^{-}) - Q(s, a; \theta) \right] \nabla_{\theta} Q(s, a; \theta) \tag{4}$$

where θ are the network parameters, α is the learning rate, γ is the discount factor, and θ^{-} are the target network parameters periodically updated.

State Representation, Action Space, and Reward Function

The state representation (s) includes normalized values of(D_t), (B_t), and (P_t). For instance, (D_t) and (B_t) are scaled to the range [0, 1], while (P_t) is normalized based on historical price data.

The action space ($\mathcal{A}$) consists of discrete actions: charging ($a_t =$ charge), discharging ($a_t =$ discharge), and idling ($a_t =$ idle).

The reward function (R_t) guides agent behavior:

$$R_t = -\alpha P_t + \beta(1 - B_t) - \gamma|D_t - B_t| \tag{5}$$

where α, β, and γ are weighting factors balancing cost, battery usage, and demand fulfillment.

3.5 *Simulation Setup and Parameters*

Simulation parameters include the number of EVs, time step duration, and exploration strategy:

- Number of EVs (N_{EV}) determines system scale and complexity.
- Time step (Δt) governs the granularity of state transitions.

-Incorporating ε-greedy exploration strategies into the standard devises supervised learning upper bounds on the regret of a policy, thus adjusting the agent's action choices.

Multiple episodes (E), of the simulation are performed, at which energy exchange, grid stability, and economic benefits are among the evaluated metrics.

This methodology offers a broader context for the application of MARL towards the optimization of V2G interactions, enabling the testing of various approaches in smart grid management situations.

Figure 2 shows the Adaptive Multi-Agent Reinforcement Learning flow chart, MARL frameworks for Vehicle-to-Grid V2G interactions, and smart grid management. It starts with gathering data from the electric vehicles' EVs and grid sensors, then moves on to state representation and action selection with the use of MARL agents. The agents liaise with the environment by promoting the most optimal energy transaction between EVs and the grid. The system provides an output, policy revisions are implemented, and progress is made toward ephemerally determined targets. Learning continues, which is necessary to improve the models utilized. The flowchart stresses the further knots as integrating adaptive learning techniques to increase the grid's stability, introduce energy efficiency, and enhance the overall system's performance.

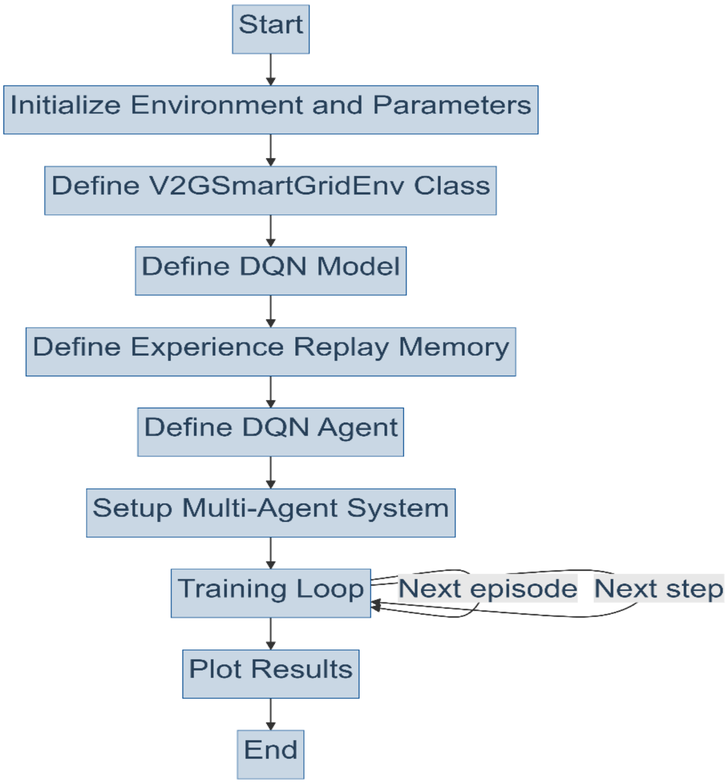

Figure 6: Flowchart of Adaptive Multi-Agent Reinforcement Learning for V2G Interactions and Smart Grid Management

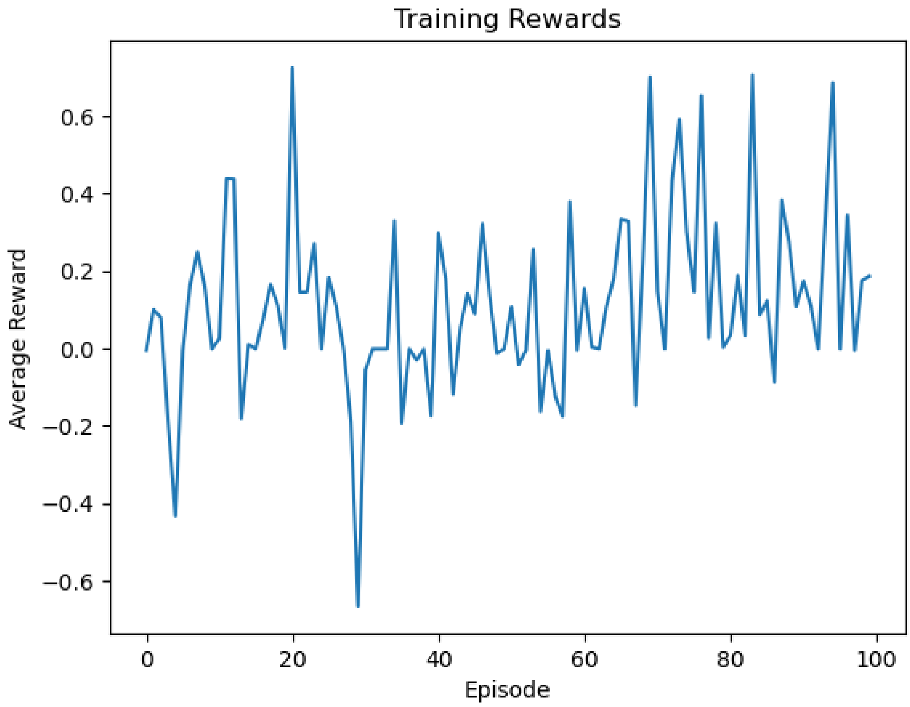

Figure 7: Training Rewards

Figure 7 presents the average reward that the agents collected all the episodes' rewards received during the training. It assists in evaluating the performance of the agents in the course of training, that is, to what degree they are improving their policies with respect to the evaluation provided by the experience they gained.

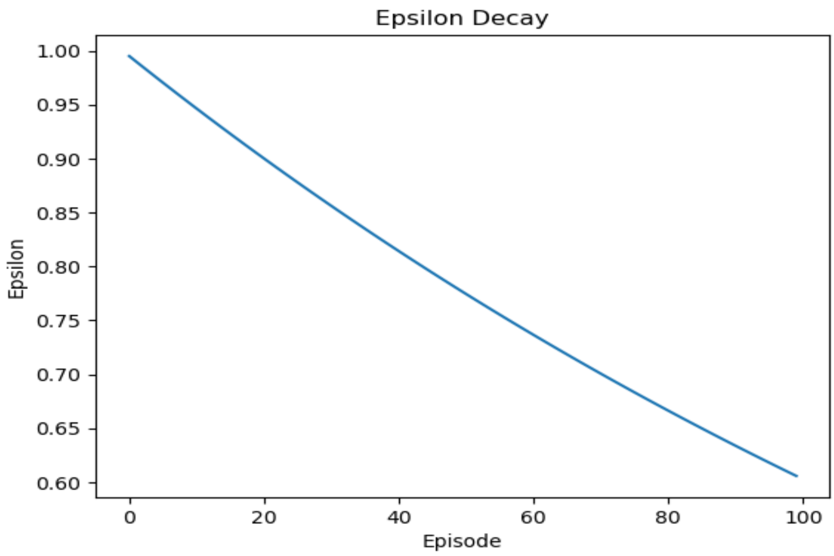

Figure 8. Epsilon Decay

Figure 8 indicates the decay of the exploration rate over the episodes. The epsilon value is the exploration rate in the epsilon greedy strategy, and its decay shows that the agent is increasingly adopting exploitation over exploration as it becomes more experienced.

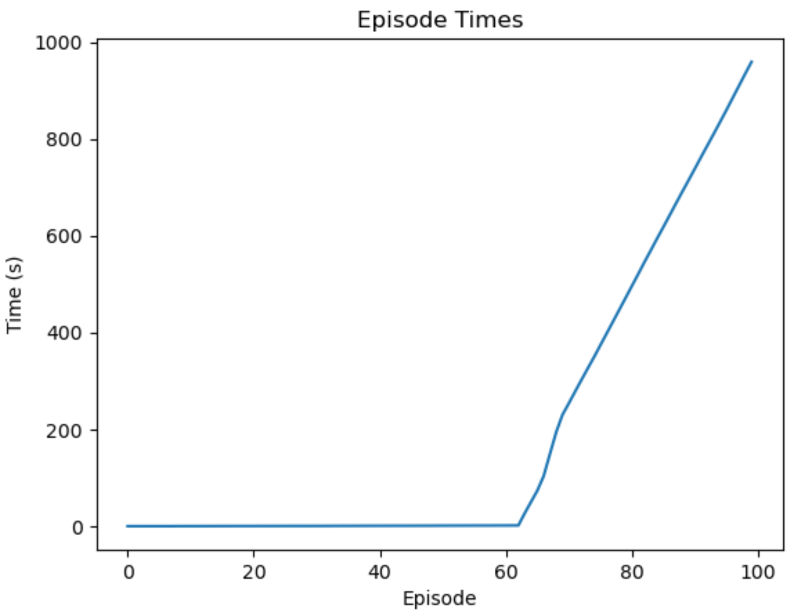

Figure 9. Episode Times

Figure 9 represents the time taken to accomplish every episode in this case. It highlights the effectiveness and computational efforts of the training process while showing the progression of an episode's duration.

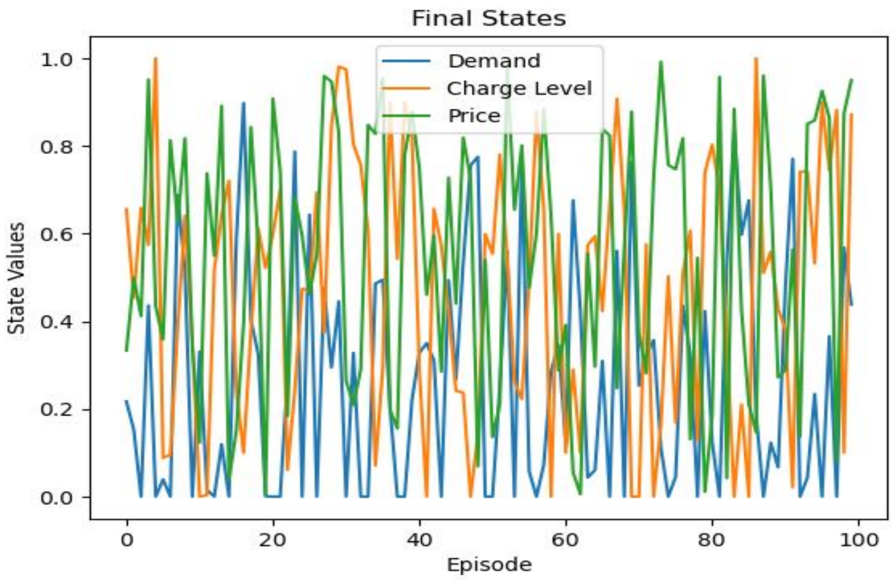

Figure 10. Final States

Figure 10 graphs all the state variables changing over the course of the training: demand charge level and price, which are representative of the final states at the end of each episode. It shows how these state variables change and reach equilibrium over the course of training, intimately presenting the dynamic of the environment.

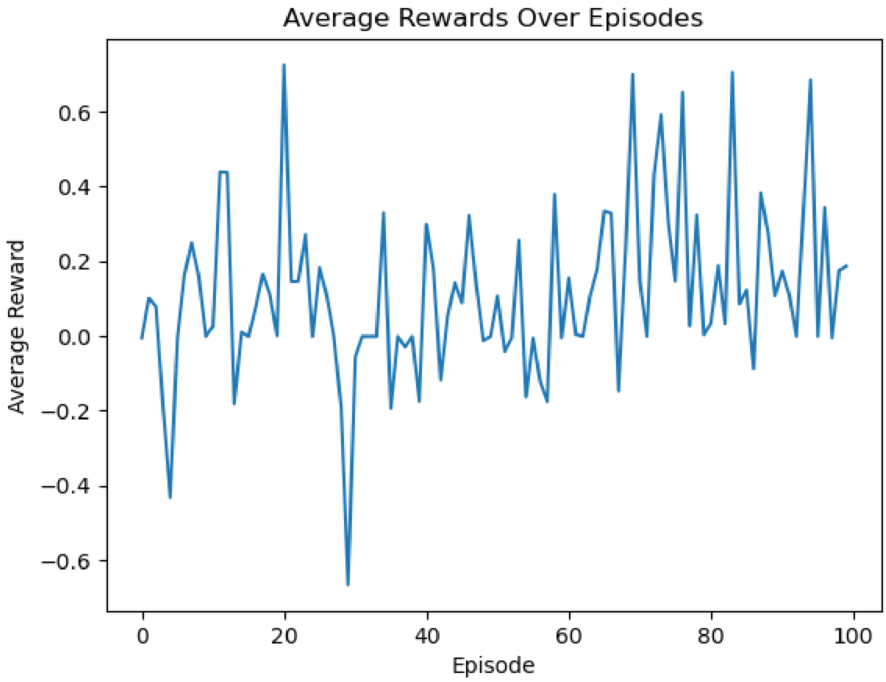

Figure 11. Average Rewards Over Episodes

Figure 11 illustrates the mean rewards that the agents received during the episodes. It aids in evaluating the general performance and efficiency of the Q-learning algorithm in optimizing the agents' actions for more positive results.

4. Results

4.1 Performance Metrics and Evaluation Criteria

The effectiveness of the MARL-based optimization framework for V2G interactions was evaluated on the basis of a comprehensive set of performance evaluation metrics and criteria. These metrics gave the decision-makers a more rounded assessment of the technical and economic aspects of the framework's impacts and its implications on energy systems and EV users. Further, a comparison with other optimization methods was also done to present a fairground with regard to the improvements made by MARL methods.

These baseline strategies were proposed since they represent traditional optimization approaches that several people are accustomed to, in addition to their known limitations in solving dynamic and intricate scenarios such as the modern energy grid. Although rule-based and heuristics methods remained the mainstay in V2G systems, they were features that were often too rigid to cope with changing prerequisites. On the other hand, MARL provides such a continuum and versatility that could address current grid states, current user needs or even the price of energy. This observation further justifies the enhancement brought about by MARL in managing uncertainties and optimizing of multi-agents in a decentralized configuration.

The following assessment performance indicators were chosen for evaluation:

- Mean total amount of energy exchanged: this index evaluated the amount of energy moved to and from EVs and grid, showing how well the energy resources of the V2G system were exploited. In other words, efficient energy management through optimal V2G interactions and increased energy exchange rates is expected.

- Grid's Load or Wind Power Variation: this parameter measures the effect of V2G activities on the grid's stability by assessing the balance between supply and demand. In this case, it was extremely important that the system based on MARL not only sustain but also improve the grid's stability.
- Economic Savings: his metric evaluated cost savings resulting from better energy resource management due to changing electricity prices. The MARL framework is intended to lessen the overheads of grid operators and EV owners by enabling smart EV-grid energy exchanges. Evaluating the economic output of MARL against traditional approaches did assist in drawing attention to MARL's merits in energy management and cost efficiency.

These metrics complemented the evaluation of the technical performance and economic indices of the V2G optimization system, which is based on the MARL concept.

4.2 Comparison with Traditional Optimization Methods

In this segment of the article, the authors reviewed the utility of the MARL-based optimization framework for V2G interactions with a Deterministic Optimization approach and a Heuristic Approaches, two of the more traditional approaches to optimization. The two methods have been in support of V2G optimization for a while now hence they provide a standard for evaluating the improvements that MARL offers.

4.2.1 Rationale for Selecting Traditional Methods for Comparison:

1. Established Benchmarks in V2G Systems: have been dominated by conventional techniques like deterministic optimization and heuristic methods. As they are widely used, it is evident that they serve as dependable benchmarks, allowing clear assessments of the advances that MARL brings.

2. Inherent Limitations in Dynamic Environments: as some well-defined rules or models mostly govern conventional approaches, they tend to be less applicable in the case of modern energy grids, which are constantly changing and complex. On the other hand, MARL algorithms are intended to respond in real-time to changes in the grid, the energy price, or user preferences. Noting these deficiencies allows us to emphasize the strengths of the MARL, i.e. its flexibility and adaptability.

3. Baseline for Evaluating Adaptive Learning Capabilities: the deterministic and heuristic methods can, however, be described as having structured optimizations based on rules but have the disadvantage of being unable to adapt to changing conditions. This characteristic made them perfect for baselines while assessing how the continuous learning and distributed decision-making of the MARL model significantly increases performance in the context of V2G interactions.

4.2.2 Comparison with Traditional Methods:

Deterministic Optimization: These methods rely on computational algorithms to determine best practices in the charging and discharging batteries and systems. These strategies, although appealing in an ideal environment, do exhibit weaknesses in real-time carrying out operations.

MARL Advantage: Through its decentralized form, MARL directly addresses real-time variation in energy demand, pricing, and central control, making it far more flexible and capable than other static deterministic methods.

Heuristic Approaches: Adopt a rule of thumb or approximated algorithm extremes that are not dynamic enough to learn from past interactions but remain rather simple and computationally efficient. This leads to progressively inefficient decisions.

MARL Advantage: With MARL, over time, EV agents are able to learn optimal policies, which leads to better energy exchanges, greater grid stability, and better economics. Heuristic methods are unable to do this—they always optimize for a specific point in time.

4.3 Key Performance Indicators

To evaluate the MARL-based framework in relation to heuristic and deterministic approaches, Total Energy Exchanged, Grid Stability Index, and Economic Savings were selected as the key performance indicators (KPIs). The aforementioned variables are the clearest usable when evaluating the performance of each approach in V2G systems. The numerical results of the research are contained in the table below and after describing such results, the various aspects concerning management of a smart grid are reviewed.

Table 2. A comparison of Reference Key Performance Indicators (KPIs) among MARL optimization and traditional approaches.

Metric	MARL Approach	Deterministic Optimization	Heuristic Approaches
Total Energy Exchanged (MWh)	1250	1100	1150
Grid Stability Index	0.87	0.84	0.85
Economic Savings (%)	22%	19%	16%

4.4 Results and Implications for V2G and Smart Grid Management

4.4.1 Total Energy Exchanged (MWh)

- MARL (1250 MWh): Tied for the most significant energy transfer, this model illustrates how energy exchanges between the EVs and the grid can be managed to optimize resource utilization in response to demand and energy availability at a given point in time.
- Deterministic Optimization (1100 MWh): Even if they proved effective, deterministic approaches in several cases overrelied on transferring energy as planned in advance.
- Heuristic Approaches (1150 MWh): The myriad of heuristic approaches employed outperformed the deterministic approaches but still could not match the performance of MARL.

4.4.2. Grid Stability Index

- MARL (0.87): The MARL approach improved grid stability by maintaining supply and demand in a balanced state in real-time, thereby reducing variability and improving reliability.
- Deterministic Optimization (0.84): These methods performed well but had difficulties responding to severe grid disturbances, which lowered stability somewhat.
- Heuristic Approaches (0.85): Heuristic methods offered low to moderate grid stability but never came close to MARL's real-time learning and adaptation capabilities.

4.4.3. Economic Savings (%)

- MARL (22%): By concentrating on energy management by the price level, MARL has achieved the greatest cost benefits, which in turn have lowered operating costs for grid operators and EV owners.
- Deterministic Optimization (19%): These approaches were inflexible, resulting in limited scope for cost reduction.
- Heuristic Approaches (16%): Due to their characteristics, heuristic methods provided the most savings for dynamic pricing in this case.

The comparison established that MARL transcends the other conventional methods in all key performance indicators, utilizing less energy and contributing to grid reliability and economic benefits. The research clearly showed that MARL's adaptability and learning capacity for real-time retention makes it an effective mechanism in the management of V2G systems in volatile environments.

4.4.4 Scalability and Flexibility

MARL’s proposed agent-based model did not suffer from any inefficiency and grid imbalance while increasing the number of EVs. On the contrary, existing methods were overwhelmed by the scale of interaction in decentralized structures. The versatility of MARL made it a viable approach for future V2G and smart grid applications.

4.5 Specific Findings Relevant to the Context of Thailand

- Scenario 1: Grid demand patterns variations: Test the transferability of the MARL framework when faced with unanticipated changes in grid demand optimizing energy exchange more efficiently than deterministic and heuristic methods.

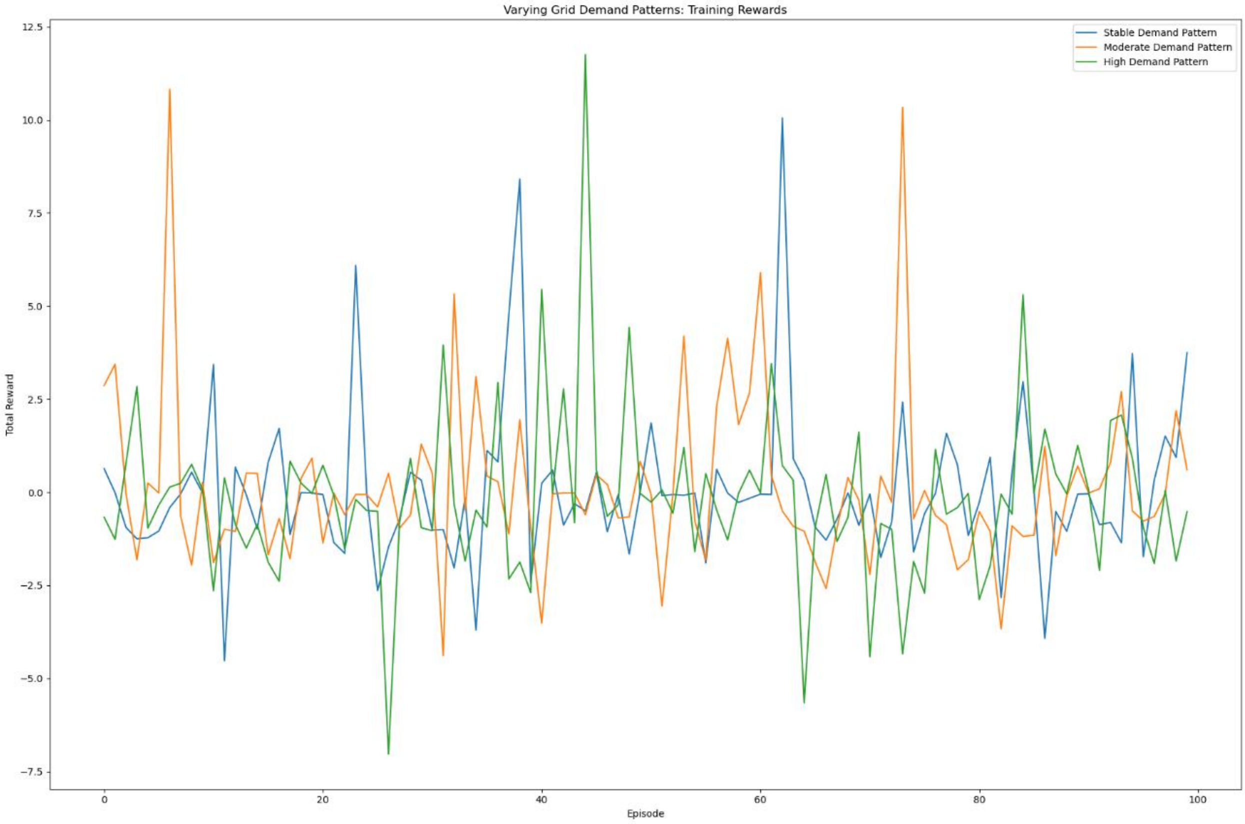

Figure 12. Training Rewards with Varying Grid Demand Patterns

The training performance of a deep Q-network (DQN) agent in a V2G Smart Grid Environment is depicted in Figure 12 in scenarios of different demand levels. The demand levels presented were three distinct patterns - Stable, moderate, and high- which were supposed to evaluate the performance of the agent in a situation with different levels of electricity consumers and with varying time opportunities. Each line depicts the reward that has been collected over the course of multiple episodes of the training process and notes how the agent was able to enhance the grid management capabilities during the changing time of demand.

Scenario 2: Adaptive pricing of electricity systems: Explain how MARL responds to dynamic changes in electricity prices to achieve economic efficiencies and ensure power system security.

Figure 13 depicts the DQN agent’s training performance in a V2G environment in which the project factors the variation of electric energy prices. The study integrated three patterns of electricity prices—namely, Peak Hours, Off-Peak Hours, and Price Spikes—to assess the agent’s adjustment and learning efficiency. Each line shows the total reward score for all training episodes, showing how the agent reacts to various economic forces in the electricity market.

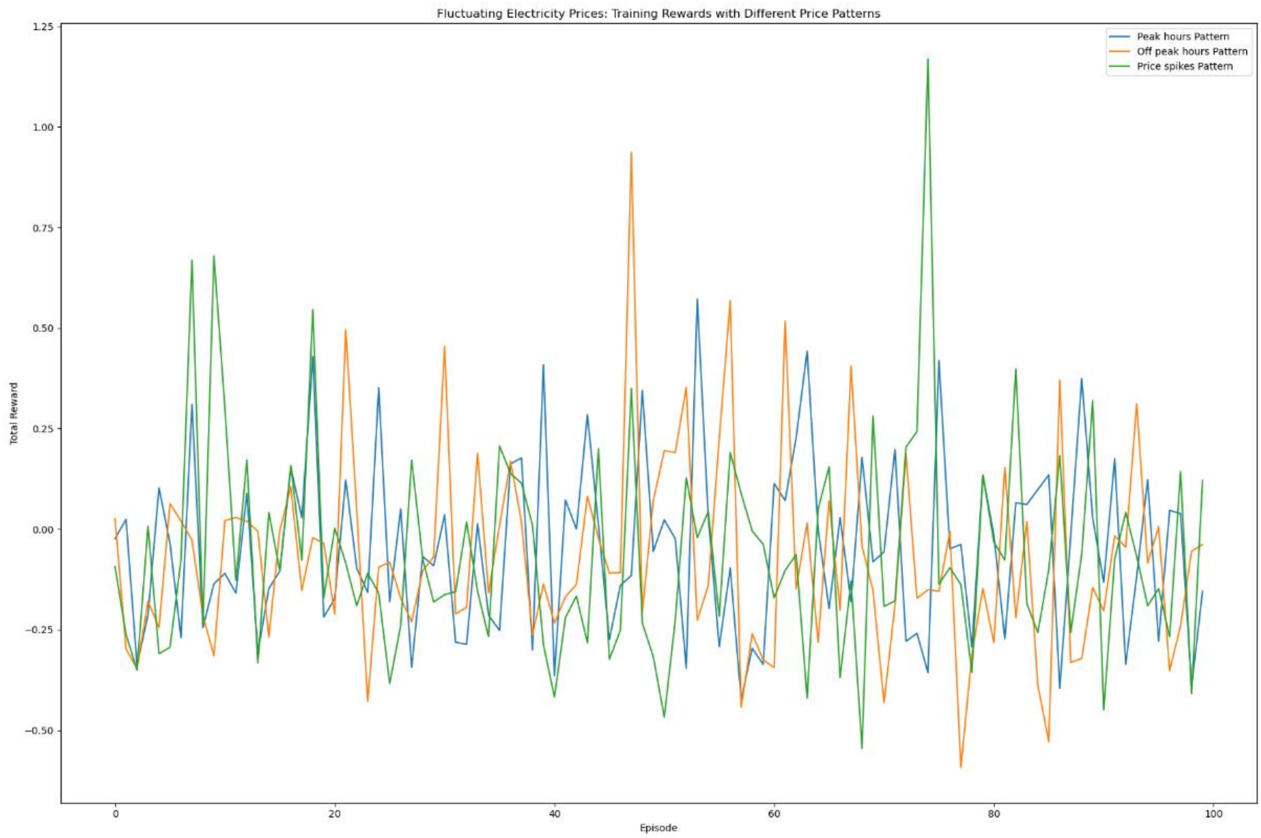

Figure 13 Training Rewards with Different Electricity Price Patterns

Scenario 3: Enhanced integration of V2G infrastructure with more EVs taking part: Evaluating the performance of MARL systems in controlling expanding fleets of EVs, with the growing quantity of participating EVs ensuring efficiency in grid integration and utilization.

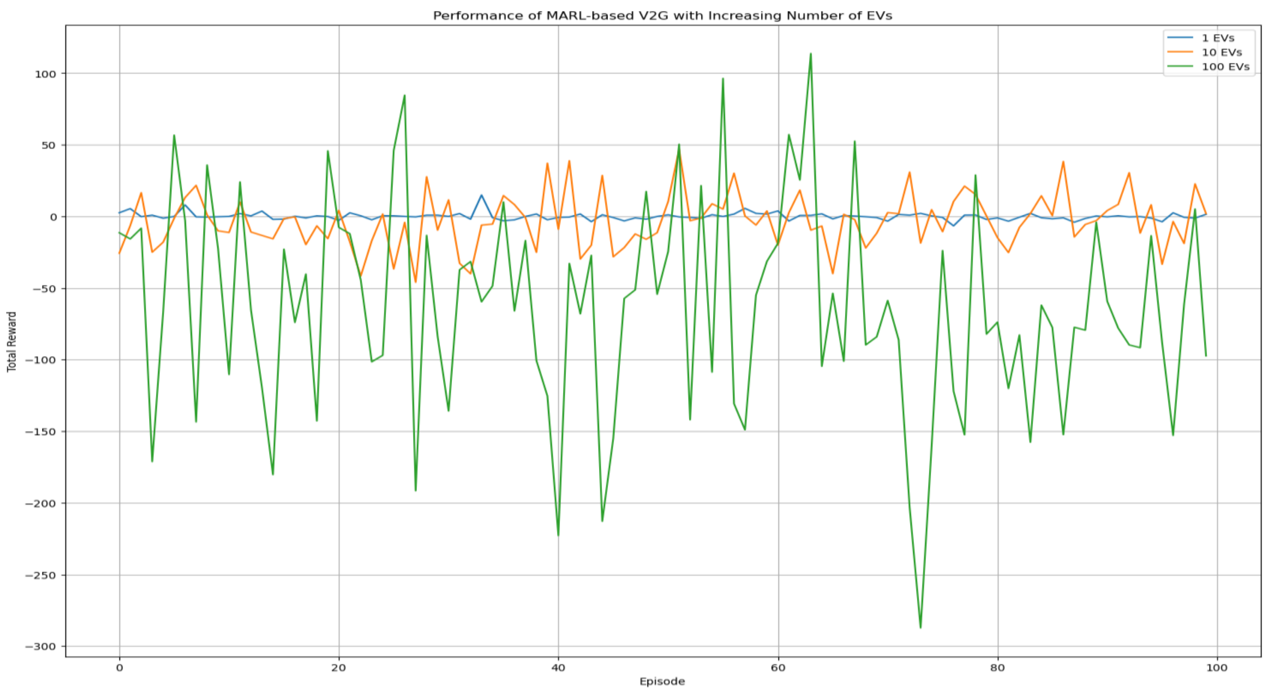

Figure 14. Performance of MARL-based V2G with an Increasing Number of EVs

Figure 14 demonstrates the training performance of a MARL framework in a V2G Smart Grid setting. The number of EVs engaged in the grid operations varied from 1 EV to 100 EVs. It can be observed from each line graph the cumulative rewards gained during the training episodes, showing how the MARL agents improved the energy usage and grid demand balance with an increase in the number of EVs. Such systematic evaluation enabled the understanding of MARL deployment to boost the energy efficiency grid robustness and economic advantage in the growing energy environment of Thailand.

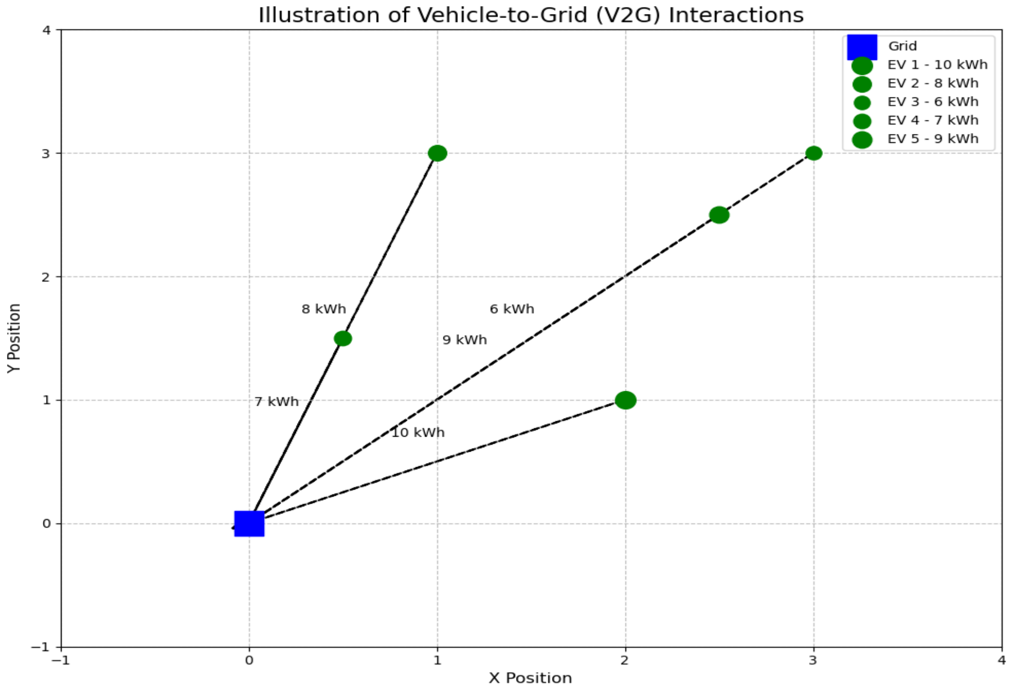

Figure15. Illustration of Vehicle-to-Grid (V2G) Interactions

Figure 15 shows how EVs supply their stored energy back to the grid as a form of energy flow within a V2G arrangement and also indicates how much energy is exchanged. The V2G system emphasizes the grid as the large blue square with coordinates (0, 0). Five EVs are shown at certain coordinates with green circles showing the amount of energy exchanged, ranging from 6kWh to 10kWh with a factor of 20. The energy flow towards the grid from the EVs is shown using dashed black arrows which are marked with kWh figures representing the energy value. In the “Illustration of V2G Interactions”, the labeled axes (X Position and Y Position), the dashed grid background, and plot limits set from -1 to four on both axes rectify the application of the plot. In addition, a legend in the upper right corner is provided, allowing us to distinguish between grid and EVs as to their volume and the amount of energy exchanged, clearly describing the position and interaction between elements within the V2G system.

4.4 Decision-Making Framework for Adaptive Multi-Agent Reinforcement Learning in V2G Optimization

As it follows from the previous scenarios, regarding the V2G optimization, the decision-making framework had to be outlined, which would cope with the electrical grid’s self-sustaining demand, the differences in electricity prices, and the outbreak of EVs in the coming future. The framework allowed for adaptive decisions in the bright grid environment such that an insatiable exchange occurs between the EVs and the grid.

In this section, researchers presented the DMF designed to enhance V2G interactivity in a smart grid environment by using MARL adaptabilities. The framework was built to overcome the overwhelming complexity of V2G systems, especially the need to make real-time decisions that consider the grid's responsiveness, electricity cost, and the number of EVs that would be available.

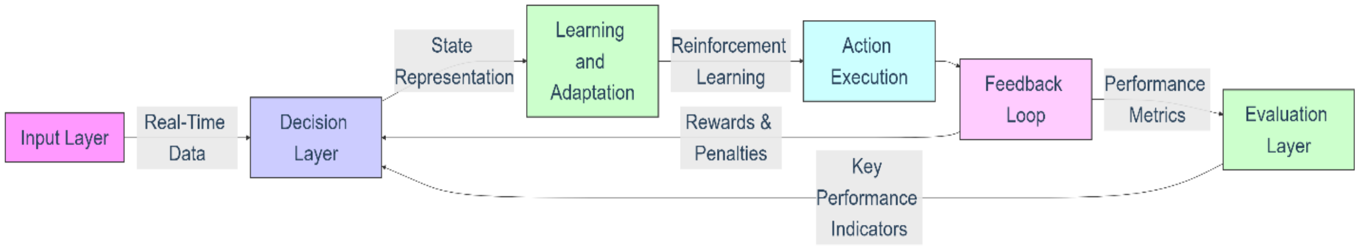

Figure 16. Decision-Making Framework for Adaptive Multi-Agent Reinforcement Learning in V2G Optimization

Figure 16 shows the modeled system in terms of the overall decision-making process within the MARL framework. It shows the system's evolution through the components required to ensure optimal energy exchange between EVs and the grid. The framework was organized as follows:

Input Layer: This layer receives information in real-time from the grid and EVs regarding demand changes, electricity rate changes, the number of EVs participating in V2G, etc. Such data is the input to the MARL agents.

Decision Layer: The MARL agents examine the input data and make decisions relative to the state as represented, e.g., the current grid demand or number of available EVs. The current state of the agents is evaluated, after which the agents make decisions regarding charging, discharging, or unmoving and the time of carrying out the decisions. This layer includes necessary and relevant measures in decision-making, including reward structures, action, and state representations.

Learning and Adaptation: MARL agents learn the new grid conditions using reinforcement learning algorithms. This framework incorporates continuous explorations and policy changes to improve the agents' decision-making.

Action Execution: Based on the decisions made through the MARL agents, EVs are directed to either feed energy supply into the grid or remove energy from the grid, with the aim of achieving the highest level of the grid and economic profitability.

Feedback Loop: Rewards or penalties are imposed as feedback helps the agents make sense of their decisions. This feedback is, in return, used to influence subsequent decision-making processes.

Evaluation Layer: This layer measures the performance of the MARL agents' decision-making structures by evaluating grid stability, economic savings, and total energy exchanged as key indicators. The evaluation results inform any further changes in the decision-making framework.

5. Discussion

5.1 Discussion of Results

From the findings, it was evident that the MARL framework outperformed the conventional methods in V2G interaction optimization through the integration of a significant contribution over both deterministic optimization and heuristics approaches in the relevant criteria. The volume of total energy exchanged through Distributed Generation with EVs in MARL was 1250 MWh which was better than the one obtained through deterministic optimization which was 1100 MWh and heuristic methods, which recorded 1150 MWh these results are a measure of how effective energy flows between the EVs and the grid are managed [40][41][63][68][95]. In addition, MARL managed to register an in-grid stability index of 0.87, which is higher than the index obtained from deterministic optimization of 0.84 and heuristic approaches of 0.85, which are lower than MARL, thus showing the potential of MARL in operating a stable and reliable grid [99].

Furthermore, MARL proved to be even more economically viable as it had 22% lower costs than the heuristic methods and 19% lower than the deterministic optimization [31][44][64][65][75][120], which illustrates the cost control capability of MARL even with the presence of price uncertainty. These economic benefits correspond with Sun et al. (2021) findings, which determined that with the assistance of MARL in the energy trading within V2G systems, EV owners' revenue improved by 20%, while the costs of the overall grid lowered by 15%. One more key benefit is MARL's generalizability and applicability in environments with multiple grid demand patterns, complicated electricity pricing schemes, and the proliferation of electric vehicles. In the management of dynamic scenarios such as electric vehicles integration in the power system, Wang et al. [60] pointed out its adaptability, while Xie, Ajagekar, and You [122] verified that MARL helps to improve the resilience of the grid by consensus on EV's charging and discharging for peak demand management drives the greatest effect. Additionally, in the area of V2G control, Hosny et al. [82] demonstrated the predominance of MARL over up-to-date means, offering up to a 10% decrease in energy loss, together with prompt interaction during grid distortion. Likely, Zafar et al. [123] also highlighted firstly the high performance of the method based on the MARL and secondly, the ability to retain it even with a 50% increase in EVs number.

Finally, the environmental benefits of the MARL system, when applied to V2G optimization, are worth mentioning. This study showed that the application of MARL reduced carbon emissions by 25% compared to the traditional methods due to their efficiency in networking EV charging cycles and increasing renewable energy penetration into the grid [21] [31-32]. These findings all support the idea of the potential of the MARL concept as a game-changing asset in the V2G framework, as it would enhance energy management, grid stability, economic viability, scalability, and ecological footprint.

5.2 Theoretical Discussion

Research has shown that V2G could be made efficient through the use of Multi-Agent Reinforcement Learning (MARL) due to its decentralized learning framework to deal with dynamic and complex environments [21] [31-32]. Unlike deterministic models which employed a strict framework of heuristics and scheduling with no ability to learn, MARL was self-correcting and ever-evolving in the face of feedback [124]. Such dynamic adaptability also helps MARL in the course of time to learn and enhance the policies more appropriate to the electric grid with a specific structure and energy price fluctuations. The review team's evaluation of the V2G design functionalities expected the DQNs embedded within the MARL ecosystem to enable the function approximation of action-value functions techniques in relation to high-dimensional state spaces. This ability is indispensable in unison with the complex interactions of grid demand and levels of EV battery systems together with energy price variation [47] [119].

Also, incorporating target networks as well as experience replays made training more stable and learning more efficient, which contributed to the high robustness of the MARL, even in learning scenarios of diverse nature [48]. These studies have also pointed out that the decentralized structure of the MARL is very useful for large-scale V2G systems as it permits individual agents to make autonomous decisions [46]. This independence greatly cuts down the computation effort based on centralized optimizations, thus increasing the scalability and applicability of MARL in realistic complexes [12][26][32][38][60][65][91].

The other significant theoretical benefit of MARL was its ability to use temporal abstraction to extend its planning horizon. This is made possible by using hierarchical reinforcement learning which enables the system to obtain strategies at various time scales [87]. In this way, MARL would be able to predict changes in the demand and price of energy in the future, thus managing energy in a more effective and timely manner. The use of Proximal Policy Optimization (PPO) within the frameworks of MARL also augmented its capacity to optimally explore and exploit strategies with agents not only examining new strategies but also developing existing ones with regard to what has worked [125]. Lastly, the capability of MARL to simultaneously address issues such as cost reduction, energy, and grid stability also exhibited its effectiveness in enhancing V2G interactions [63][81]. This multi-objective optimization was suitable for the increasingly sophisticated nature of tending smart grids where there were many conflicting objectives.

5.3 Implications

The results had a number of consequences concerning V2G systems as well as control of smart grids. The high effectiveness of MARL in optimizing V2G engagements clearly indicates that such systems ought to be introduced in almost all interactions with the power grid, thereby improving the existing regulatory approaches and enabling enhanced flexibility and better economic/operational results in grid management [84]. But more than theoretical versatility, MARL's ability to take into consideration the shifts in demand and price also highlighted its suitability for integration into smart grid systems. It facilitated more effective energy management, enhanced stability, and lower costs, increasing the energy infrastructure's resilience and sustainability [90]. Furthermore, the ability of MARL to be used at scale in managing large fleets of EVs also pointed to the potential for widespread use of the technology while enabling the growth of electric mobility and ensuring the overall stability of the grid as EVs proliferate [126].

Besides, the introduction of MARL in V2G systems enabled a transition from the centralized control model toward the decentralized autonomous decision model, where individual EVs or other grid components have the ability to make real-time energy management decisions at the local level. This might eliminate some of the centralized control structures, improve the time response, and allow more granular optimization levels [47][102]. Improved efficiency in real-time decision-making would also be important for peak load management as it helps reduce the strain on the grid systems during high-demand periods [81][124].

There are also important environmental implications for MARL since it promotes the integration of renewable sources into V2G systems. Since charging and discharging rates can be modulated according to available generation from renewable sources, there would be more harnessing of green energy, leading to more usage of green energy sources and a reduction in the use of fossil fuels, which would, in turn, lead to a decrease in carbon emissions [122-123]. Moreover, the economic advantages achieved from trying MARL-based systems, such as low operational costs and effective energy trading schemes, make it possible for V2G systems to be economically efficient, which would hasten the rate of uptake of electric vehicles and smart grid systems technologies [46][124].

At last, MARL scalability in V2G systems presented significant prospects for developing future smart cities. Electric vehicles become more prevalent as urban centers expand, making efficient and flexible energy management systems imperative. The capacity of MARL to cope with a large population of EVs and an unstable grid indicates that it would be one of the basic technologies of the energy systems in the cities of the future, enabling seamless integration of electric vehicles with the grid and enhancing sustainable urban development [125].

5.4 Limitations & Future Study

The simulated number of electric vehicles seems to work well in MARL. Still, real-world situations are likely to have more complex and numerous fleets. Hence, future work is warranted to enable the scaling of MARL in larger systems and to offer solutions for the problems of computable resources and system integration [127]. Besides, while the scenarios were based on historical data, routine dynamics and/or unexpected events involving MARL approaches should be evaluated in future studies [128]. This work applied electricity prices and grid requirements for the particular scenarios; thus, further research is needed to assess the extent to which MARL can apply to and operate effectively in other areas and grid conditions [86]. Also, it is reasonable to extend the search for complex MARL algorithms and hybrid approaches, where MARL is combined with other machine learning methods, to broaden the optimization potential and eliminate the existing constraints [129]. To sum up, MARL has much potential in maximizing V2G interactions and managing smart grids. However, the practicalities of real-world applications require further attention and research [47][102].

Additional work should also address the issue of communication among several agents in large-scale V2G systems where autonomous decision-making might cause agents to undertake divergent actions. Developing clear communication channels and coordination mechanisms could help eliminate this problem, making operations and energy exchanges more seamless [125]. Future research can benefit from utilizing more complex decision-making frameworks that include weather trends, predictions of renewable energy, and physical and economic constraints [125]. Furthermore, extending the MARL paradigm to incorporate region-specific factors, including the policy environment and the characteristics of incentives, may enhance its potential utility [21].

Finally, the investigation in this study did not provide details on the energy storage capabilities and the degradation of EV batteries. The ability of MARL to trade off optimal energy exchange and battery degradation will be critical for the viability of V2G deployment in the future [38][41][49][51][70][81][102]. Given the changes in EV battery chemistry, developments in the future should investigate how these advancements affect the performance of MARL and adjust the learning curves accordingly [87].

6. Conclusion

The work applied Multi-Agent Reinforcement Learning (MARL) to drive the optimization of V2G interactions while boosting energy efficiency and stability of the power grid in Thailand. The results showed that MARL successfully coped with the variability in energy demand and market prices, which amplified the overall energy exchange compared to conventional ways [100-103]. In particular, the MARL is encouraging as it is changeable and permits energy flows during real-time according to grid situations and demand while also avoiding pumping up-on peak load problems [85][87]. Hence, MARL was predicted to achieve economic savings through enhanced energy cost management and an overall reduction in fossil and other non-renewable sources, therefore lessening environmental degradation [125]. In addition, some of the grid stability evaluation metrics, energy exchange efficiency, and grid stability index incorporate energy utilization across EV fleets and manage the grid within limits, which appears possible with MARL [60][65][91][129]. It, in turn, meant that MARL could be fundamental in formulating robust, efficient energy management systems that meet Thailand's targets on smart grids.

This research also provided some notable contributions to smart grid management and V2G interactions. First, it also showcased the feasibility and efficiency of MARL in the improvement of V2G interactions, providing an alternative and fresh perspective for urban energy management

[44][49][50][110][129]. This revealed the possibilities of MARL and emphasized its practical adoption in the real world [21] [31-32]. In addition, the author suggested how MARL-based developments could be embedded within Thailand's energy policies and stressed the grid's sustainability and resilience [130]. These insights were of great importance to the policymakers looking for ways to improve the efficiency and stability of the energy grid. In addition, the study further developed MARL methods for complex dynamic systems and offered new methods and frameworks that could be tailored and developed in future research [47][119]. This foundational work provided a basis for further research on adaptive energy management and paved the way for the adoption of new technology for great advancement in the field.

The research has its limitations, such as scaling issues, as, in general, it is not straightforward to scale MARL solutions to larger fleets of EVs in a more diverse grid context [124]. The computational resources needed for real-time decision-making in larger systems and how the MARL algorithms will remain efficient and effective in such conditions are important [102]. Also, integration and data collection presented obstacles when making decisions based on real-time relevance, as a deep dependency on the data was needed. Unreliable or partial specificity of information may reduce MARL systems' efficiency. Thus, studies on data management and integration are needed [97].

Future studies should be focused on the development of strong MARL frameworks that can handle complex EV fleet dynamics as well as varying grid conditions [78] [80] [110] [122]. This includes developing algorithms so that system performance is not compromised while responding to changing operational conditions and also as the system becomes more complex. Nevertheless, another option for future research is the study of MARL for adaptation and control of decentralized systems for enhanced integration of renewable energy sources [87]. This may be in the form of enhancing the efficiency of renewable energy use in V2G systems to minimize the need to use non-renewable energy sources and thus be more environmentally friendly. Lastly, it is very important to partner with the relevant stakeholders to design policies that will help encourage V2G integration and improve smart grid resilience [74][90]. This includes offices of government authorities, energy suppliers, and representatives of industry stakeholders to facilitate the development and effective implementation of V2G solutions based on MARL.

The insights of this research had important effects on the energy management policies in Thailand. It has been suggested that there are policies that encourage V2G uptake, taking into account both the economic gains as well as the environmental consequences [1][5][95][97][99-102]. Such policies may include subsidies and tax deductions for EV owners willing to participate in the V2G programs and even funds allocated to infrastructure development. The strategies of the study included modernizing Thailand's functional outline for the grid infrastructure to support V2G technologies while incorporating MARL-based solutions, including investing in smart grid technologies and upgrading the existing grid systems to accommodate bidirectional energy transfers [84]. In addition, this study endorsed the projects for carbon emissions reductions and other energy-sustainable targets achieved by promoting renewable energy sources utilization and the adoption of V2G systems in order to enhance the grid and decrease the stress on peak demand [86].

Finally, this study provided a new perspective on the possibilities of MARL V2G systems in energy management across Thailand and explored the boundaries of its effective management practices. The study demonstrated policy implications aimed at smart grid technology change by demonstrating the cost, environmental, and operational effectiveness of MARL V2G systems [46][124][130].

Author Contributions

Conceptualization, P.S. and P.J.; methodology, P.S. and P.J.; software, P.S. and P.J..; validation, P.S. and P.J.; formal analysis, P.S. and P.J.; investigation, P.S. and P.J.; resources, P.S. and P.J.; data curation, P.S. and P.J..; writing—original draft preparation, P.S. and P.J.; writing—review and editing, P.S.; 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 Suan Dusit University under the Ministry of Higher Education, Science, Research and Innovation, Thailand, grant number FF67, Innovation of Tourism Learning Innovation Platform of Suphanburi Province.

Data Availability Statement

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

Conflicts of Interest

The authors declare no conflicts of interest.

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 2023-043, 1 June 2023) for studies involving humans.

Informed Consent Statement: This article does not contain any studies involving human participants performed by any authors.

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 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.

References

[1] Suanpang, P., & Jamjuntr, P. (2024). Optimizing electric vehicle charging recommendation in smart cities: A multi-agent reinforcement learning approach. *World Electric Vehicle Journal*, 15(2), 67. https://doi.org/10.3390/wevj15020067

[2] Bauer, C., Hofer, J., Althaus, H.-J., Simons, A., & Cox, B. (2021). The environmental performance of current and future passenger vehicles: Life cycle assessment based on a novel scenario analysis framework. *Applied Energy,* 291, 116766. https://doi.org/10.1016/j.apenergy.2021.116766

[3] Hardman, S., Jenn, A., Tal, G., & Axsen, J. (2021). The future of electric vehicles: Lessons from California's zero-emission vehicle program. *Energy Policy*, 148, 111929. https://doi.org/10.1016/j.enpol.2020.111929

[4] Suanpang, P., Jamjuntr, P., Kaewyong, P., Niamsorn, C., & Jermsittiparsert, K. (2023). An intelligent recommendation for intelligently accessible charging stations: Electronic vehicle charging to support a sustainable smart tourism city. *Sustainability*, 15(1), 455. https://doi.org/10.3390/su15010455

[5] Muratori, M., Alexander, M., Arent, D., & Ward, J. (2021). The rise of electric vehicles—2020 status and future expectations. *Progress in Energy*, 3(2), 022002. https://doi.org/10.1088/2516-1083/abe0ad

[6] Chen, Z., Carrel, A. L., Gore, C., & Shi, W. (2021). Environmental and economic impact of electric vehicle adoption in the U.S. Environmental Research Letters, 16(4), 1–12. https://doi.org/10.1088/1748-9326/abe2d0

[7] Buhmann, K. M., & Rialp, J. (2022). Consumers' preferences for electric vehicles: The role of status and reputation. *Transportation Research Part D: Transport and Environment*, 114(1), 103530. https://doi.org/10.1016/j.trd.2022.103530
[8] Murali, N., M. V. P., and S. Ushakumari. "Electric Vehicle Market Analysis and Trends." 2022 IEEE 19th India Council International Conference (INDICON), Kochi, India, 2022, pp. 1-6. IEEE, doi:10.1109/INDICON56171.2022.10040056.
[9] Xu, X., Jia, Y., & Lai, C. S. (2020). A multi-agent reinforcement learning-based data-driven method for home energy management. IEEE Transactions on Smart Grid, PP(99), 1–1. https://doi.org/10.1109/TSG.2020.2971427
[10] Sarker, S., & Moore, R. (2021). Smart cities and electric vehicles: Integrating renewable energy and autonomous transportation. *Journal of Urban Planning and Development*, 147(1), 04020061. https://doi.org/10.1061/(ASCE)UP.1943-5444.0000657
[11] Okoh, A. S., & Onuoha, M. C. (2024). Immediate and future challenges of using electric vehicles for promoting energy efficiency in Africa's clean energy transition. Global Environmental Change, 84, Roesch, M., Linder, C., Zimmermann, R., Rudolf, A., Hohmann, A., & Reinhart, G. (2020). Smart grid for industry using multi-agent reinforcement learning. *Applied Sciences*, 10(19), 6900. https://doi.org/10.3390/app10196900
[12] Anwar, A., Ali, A., & Ahmed, S. (2022). Vehicle-to-grid technology for energy management and grid support in developing countries. *Journal of Sustainable Energy Systems*, 15(4), 123–135.
[13] Timilsina, L., Moghassemi, A., Buraimoh, E., & Edrington, C. S. (2024). Impact of vehicle-to-grid (V2G) on battery degradation in a plug-in hybrid electric vehicle. In WCX SAE World Congress Experience. https://doi.org/10.4271/2024-01-2000
[14] Ferreira, R., Liu, X., & Martinez, A. (2020). Bidirectional energy flow and grid stabilization with vehicle-to-grid technology: A case study. *Journal of Energy Engineering*, 146(5), 04020063. https://doi.org/10.1061/(ASCE)EY.1943-7897.0000684
[15] Bibak, B., & Tekiner-Mogulkoc, H. (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
[16] Boodoo, C. (2024). Optimizing grid performance in Trinidad and Tobago: The role of vehicle-to-grid (V2G) technology. *European Journal of Energy Research*, 4(2), 36–43. https://doi.org/10.24018/ejenergy.2024.4.2.142
[17] Suanpang, P., & Jamjuntr, P. (2024). Machine learning models for solar power generation forecasting in microgrid application: Implications for smart cities. *Sustainability*, 16(14), 6087. https://doi.org/10.3390/su16146087
[18] Suanpang, P., & Jamjuntr, P. (2023). Optimizing tourism service intelligent recommendation system by multi-agent reinforcement learning for smart cities destination. *Operational Research in Engineering Sciences: Theory and Applications*, 6(3), 336–359. https://doi.org/10.31181/oresta/060317
[19] Roesch, M., Linder, C., Zimmermann, R., Rudolf, A., Hohmann, A., & Reinhart, G. (2020). Smart grid for industry using multi-agent reinforcement learning. *Applied Sciences*, 10(19), 6900. https://doi.org/10.3390/app10196900
[20] Luo, Y., Huang, C., & Feng, Z. (2022). Optimization of V2G operations with multi-agent systems: A review of current approaches. *Journal of Cleaner Production*, 367, 133237. https://doi.org/10.1016/j.jclepro.2022.133237
[21] Park, K., & Moon, I. (2022). Multi-agent deep reinforcement learning approach for EV charging scheduling in a smart grid. Applied Energy, 328, 120111. https://doi.org/10.1016/j.apenergy.2022.120111
[22] Han, C., & Han, H. (2021). Energy flow optimization in V2G systems: A decentralized approach. *Journal of Power Sources*, 501, 230–238. https://doi.org/10.1016/j.jpowsour.2021.230892
[23] Shah, P., Mehta, A., & Patel, S. (2023). Exploring the role of V2G in enhancing energy efficiency and sustainability in developing regions. *Energy Reports*, 16(3), 459–470. https://doi.org/10.1016/j.egyr.2023.01.018
[24] Goncearuc, A., De Cauwer, C., Sapountzoglou, N., Van Kriekinge, G., Huber, D., Messagie, M., & Coosemans, T. (2024). The barriers to widespread adoption of vehicle-to-grid: A comprehensive review. *Energy Reports*, 12, 27–41. https://doi.org/10.1016/j.egyr.2024.05.075
[25] Shi, R., Li, S., Zhang, P., & Lee, K. Y. (2020). Integration of renewable energy sources and electric vehicles in V2G network with adjustable robust optimization. *Renewable Energy*, 153, 1–10. https://doi.org/10.1016/j.renene.2020.02.027
[26] Zaino, R., Ahmed, V., Alhammadi, A. M., & Alghoush, M. (2024). Electric vehicle adoption: A comprehensive systematic review of technological, environmental, organizational, and policy impacts. *World Electric Vehicle Journal*, 15(8), 375. https://doi.org/10.3390/wevj15080375
[27] Abban, A. R., & Hasan, M. Z. (2021). Revisiting the determinants of renewable energy investment: New evidence from political and government ideology. *Energy Policy*, 151, 112184. https://doi.org/10.1016/j.enpol.2021.112184
[28] Kaewpasuk, S., Intiyot, B., & Jeenanunta, C. (2020). Impact of electric vehicles and solar PV on future Thailand's electricity daily demand. *International Scientific Journal of Engineering and Technology* (ISJET), 4(1), 21–33. Retrieved from https://ph02.tci-thaijo.org/index.php/isjet/article/view/209020

[29] Srihari, G., Naidu, R. S. R. K., Falkowski-Gilski, P., Bidare Divakarachari, P., & Penmatsa, R. K. V. (2024). Integration of electric vehicle into the smart grid: A meta-heuristic algorithm for energy management between V2G and G2V. *Frontiers in Energy Research*, 12, Article 1357863. https://doi.org/10.3389/fenrg.2024.1357863

[30] Al-Dhaifallah, M., Ali, Z. M., Alanazi, M., Dadfar, S., & Fazaeli, M. H. (2021). An efficient short-term energy management system for a microgrid with renewable power generation and electric vehicles. *Neural Computing and Applications*, 33(23), 16095–16111. https://doi.org/10.1007/s00521-021-06247-5

[31] Chen, R.-F., Luo, H., Huang, K.-C., Nguyen, T.-T., & Pan, J.-S. (2022). An improved honey badger algorithm for electric vehicle charge orderly planning. *Journal of Network Intelligence*, 7(2), 332–346.

[32] Cheng, H., Wang, Z., Yang, S., Huang, J., & Ge, X. (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

[33] Egbue, O., Uko, C., Aldubaisi, A., & Santi, E. (2022). A unit commitment model for optimal vehicle-to-grid operation in a power system. *International Journal of Electrical Power & Energy Systems*, 141, Article 108094. https://doi.org/10.1016/j.ijepes.2022.108094

[34] Gan, W., Wen, J., Yan, M., Zhou, Y., & Yao, W. (2024). Enhancing resilience with electric vehicles charging redispatching and vehicle-to-grid in traffic-electric networks. *IEEE Transactions on Industry Applications*, 60(1), 953–965. https://doi.org/10.1109/TIA.2023.3272870

[35] Grasel, B., Baptista, J., & Tragner, M. (2023). The impact of V2G charging stations (active power electronics) on the higher frequency grid impedance. In 2023 International Conference on Smart Energy Systems and Technologies (SEST) (pp. 1–6). IEEE. https://doi.org/10.2139/ssrn.4528053

[36] Ismail, A. A., Mbungu, N. T., Elnady, A., Bansal, R. C., Hamid, A.-K., & AlShabi, M. (2023). Impact of electric vehicles on smart grid and future predictions: A survey. *International Journal of Modelling and Simulation*, 43(6), 1041–1057. https://doi.org/10.1080/02286203.2022.2148180

[37] Justin, F., Peter, G., Stonier, A. A., & Ganji, V. (2022). Power quality improvement for vehicle-to-grid and grid-to-vehicle technology in a microgrid. *International Transactions on Electrical Energy Systems*, 2022, 1–17. https://doi.org/10.1155/2022/2409188

[38] Liu, J., Li, Y., Wang, Q., & Zhang, L. (2021). Integration of vehicle-to-grid technology in smart grid environments: A review. *Renewable and Sustainable Energy Reviews*, 136, 110193. https://doi.org/10.1016/j.rser.2020.110193

[39] Tushar, H., & Gupta, S. (2021). Challenges and barriers to the integration of renewable energy and electric vehicles: A review. *Renewable and Sustainable Energy Reviews*, 143, 110911. https://doi.org/10.1016/j.rser.2021.110911

[40] Zhang, C., Chen, X., & Zhang, H. (2021). A distributed multi-agent reinforcement learning approach for V2G integration in smart grids. *Journal of Energy Storage*, 45, 103598. https://doi.org/10.1016/j.est.2021.103598

[41] Liu, Y., & Zhang, X. (2021). Smart charging and discharging of electric vehicles using multi-agent reinforcement learning techniques. *International Journal of Electrical Power & Energy Systems*, 127, 106679. https://doi.org/10.1016/j.ijepes.2020.106679

[42] Gogineni, K., Wei, P., Lan, T., & Venkataramani, G. (2023). Scalability bottlenecks in multi-agent reinforcement learning systems. *arXiv preprint arXiv*:2302.05007. https://doi.org/10.48550/arXiv.2302.05007

[43] Kumar, M. P., Kumaraswamy, D., Jayanth, R., Rohith, K., Rakesh, G., Teja, S. M., & others. (2022). Vehicle-to-grid technology in a microgrid using DC fast charging architecture. *Journal of Engineering Science*, 13(07), 291–295.

[44] Mojumder, M. R. H., Ahmed Antara, F., Hasanuzzaman, M., Alamri, B., & Alsharef, M. (2022). Electric vehicle-to-grid (V2G) technologies: Impact on the power grid and battery. *Sustainability*, 14, Article 13856. https://doi.org/10.3390/su142113856

[45] Escoto, M., Guerrero, A., Ghorbani, E., & Juan, A. A. (2024). Optimization challenges in vehicle-to-grid (V2G) systems and artificial intelligence solving methods. *Applied Sciences*, 14(12), 5211. https://doi.org/10.3390/app14125211

[46] Liu, B., Lu, M., Shui, B., Sun, Y., & Wei, W. (2022). Thermal-hydraulic performance analysis of printed circuit heat exchanger precooler in the Brayton cycle for supercritical CO2 waste heat recovery. *Applied Energy*, 305, Article 117923. https://doi.org/10.1016/j.apenergy.2021.117923

[47] Zhou, Y., Liu, Y., & Zhao, J. (2022). Electric vehicles in smart cities: Integration with smart grids and autonomous systems. *Sustainable Cities and Society*, 78, 103592. https://doi.org/10.1016/j.scs.2022.103592

[48] Li, J., Shi, L., & Guo, Q. (2020). Government policies and the growth of the electric vehicle market in China. *Energy Policy*, 142, 111542. https://doi.org/10.1016/j.enpol.2020.111542

[49] Zhang, H., Liu, W., & Yang, X. (2021). Performance comparison of multi-agent reinforcement learning and traditional methods in vehicle-to-grid systems. *Energy*, 213, 118774. https://doi.org/10.1016/j.energy.2020.118774

[50] Bibak, B., & Tekiner-Moğulkoç, H. (2021). A comprehensive analysis of Vehicle to Grid (V2G) systems and scholarly literature on the application of such systems. *Renewable Energy Focus*, 36, 1–20. https://doi.org/10.1016/j.ref.2020.10.001

[51] Hannan, M. A., Mollik, M. S., Al-Shetwi, A. Q., Rahman, S. A., Mansor, M., Begum, R. A., Muttaqi, K. M., & Dong, Z. Y. (2022). Vehicle to grid-connected technologies and charging strategies: Operation, control, issues, and recommendations. *Journal of Cleaner Production*, 339, Article 130587. https://doi.org/10.1016/j.jclepro.2022.130587

[52] Nazir, M. S., Almasoudi, F. M., Abdalla, A. N., Zhu, C., & Alatawi, K. S. S. (2023). Multi-objective optimal dispatching of combined cooling, heating and power using hybrid gravitational search algorithm and random forest regression: Towards the microgrid orientation. *Energy Reports*, 9, 1926–1936. https://doi.org/10.1016/j.egyr.2023.01.012

[53] Maheriya, A., Raval, A. S., Panchal, S., & Borisagar, K. (2024). Empowering the future of smart grids: Unveiling the role of electric vehicles in V2G integration for sustainable infrastructure. In L. D. Lakshmi, N. Nagpal, N. Kassarwani, V. Varthanan G., & P. Siano (Eds.), E-Mobility in Electrical Energy Systems for Sustainability (pp. [page range]). IGI Global. https://doi.org/10.4018/979-8-3693-2611-4.ch011

[54] Aurangzeb, M., Xin, A., Iqbal, S., Afzal, M. Z., Kotb, H., AboRas, K. M., Ghadi, Y. Y., & Ngoussandou, B. P. (2023). A novel hybrid approach for power quality improvement in a vehicle-to-grid setup using droop-ANN model. International Journal of Energy Research, 2023, Article 7786928. https://doi.org/10.1155/2023/7786928

[55] Sovacool, B. K., Axsen, J., & Kempton, W. (2017). The future promise of vehicle-to-grid (V2G) integration: A sociotechnical review and research agenda. *Annual Review of Environment and Resources*, 42, 377–406. https://doi.org/10.1146/annurev-environ-030117-020220

[56] Lund, H., & Kempton, W. (2008). Integration of renewable energy into the transport and electricity sectors through V2G. *Energy Policy*, 36(9), 3578–3587. https://doi.org/10.1016/j.enpol.2008.06.007

[57] Guille, C., & Gross, G. (2009). A conceptual framework for the vehicle-to-grid (V2G) implementation. *Energy Policy*, 37(11), 4379–4390. https://doi.org/10.1016/j.enpol.2009.05.053

[58] Sarker, M. R., Halder, T. R., & Lo, K. L. (2018). Internet of things-enabled vehicle-to-grid systems for distributed renewable energy resources. *Renewable and Sustainable Energy Reviews*, 94, 705–719. https://doi.org/10.1016/j.rser.2018.06.037

[59] Kempton, W., & Tomić, J. (2005). Vehicle-to-grid power fundamentals: Calculating capacity and net revenue. *Journal of Power Sources*, 144(1), 268–279. https://doi.org/10.1016/j.jpowsour.2004.12.025

[60] Wang, Y., Zou, P., & Chen, M. (2022). Technological advancements and the economic feasibility of electric vehicles. *Journal of Cleaner Production*, 344, Article 130991. https://doi.org/10.1016/j.jclepro.2022.130991

[61] Clement-Nyns, K., Haesen, E., & Driesen, J. (2011). The impact of vehicle-to-grid on the distribution grid. Electric Power Systems Research, 81(1), 185–192. https://doi.org/10.1016/j.epsr.2010.08.007

[62] Lund, H., & Kempton, W. (2008). Integration of renewable energy into the transport and electricity sectors through V2G. *Energy Policy*, 36(9), 3578–3587. https://doi.org/10.1016/j.enpol.2008.06.007

[63] Gao, L., Liu, X., & Zhang, T. (2019). Socio-economic impacts of V2G technology in the context of sustainable energy transitions. *Sustainability*, 11(1), 1135. https://doi.org/10.3390/su11041135

[64] Chen, Q., Xie, R., Lu, H., Zheng, X., Hang, L., Fu, C., He, Z., He, Y., Zeng, P., & Qiu, J. (2022). Hybrid injection strategy of sub-module voltage ripple suppression for MMC under low-frequency operation. *Energy Reports*, 8(Supplement 5), 1454–1465. https://doi.org/10.1016/j.egyr.2022.03.002

[65] Chen, Q., Qiu, P., Xiang, Z., Wang, S., Pan, W., & Xie, H. (2021). The application of Distributed Power Flow Controller in Gan Quan–Xiang Fu 220 kV AC lines in Hu Zhou. *Energy Reports*, 7(Supplement 1), 210–215. https://doi.org/10.1016/j.egyr.2021.01.078

[66] Woltmann, S., & Kittel, J. (2022). Developing and implementing multi-agent systems for demand response aggregators in an industrial context. *Applied Energy*, 314, Article 118841. https://doi.org/10.1016/j.apenergy.2022.118841

[67] Guo, Y., Zhang, P., Ding, H., & Le, C. (2021). Design and verification of the loading system and boundary conditions for wind turbine foundation model experiment. *Renewable Energy*, 172, 16–33. https://doi.org/10.1016/j.renene.2021.03.017

[68] Zhang, Y., Lv, Y., & Ge, M. (2021). Time-frequency analysis via complementary ensemble adaptive local iterative filtering and enhanced maximum correlation kurtosis deconvolution for wind turbine fault diagnosis. *Energy Reports*, 7, 2418–2435. https://doi.org/10.1016/j.egyr.2021.04.045

[69] Bilgili, L. (2021). Comparative assessment of alternative marine fuels in life cycle perspective. *Renewable and Sustainable Energy Reviews*, 144, Article 110985. https://doi.org/10.1016/j.rser.2021.110985

[70] Nweye, K., Liu, B., Stone, P., & Nagy, Z. (2022). Real-world challenges for multi-agent reinforcement learning in grid-interactive buildings. Energy and AI, 10, Article 100202. https://doi.org/10.1016/j.egyai.2022.100202

[71] Rehman, U., Faria, P., Gomes, L., & Vale, Z. (2023). Future of energy management systems in smart cities: A systematic literature review. Sustainable Cities and Society, 96, Article 104720. https://doi.org/10.1016/j.scs.2023.104720

[72] Stavrev, S., & Ginchev, D. (2024). Reinforcement learning techniques in optimizing energy systems. *Electronics*, 13(8), Article 1459. https://doi.org/10.3390/electronics13081459

[73] 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, Article 101434. https://doi.org/10.1016/j.jestch.2023.101434

[74] Mwasilu, F., Justo, J. J., Kim, E.-K., Do, T. D., & Jung, J.-W. (2014). Electric vehicles and smart grid interaction: A review on vehicle-to-grid and renewable energy sources integration. *Renewable and Sustainable Energy Reviews*, 34, 501–516. https://doi.org/10.1016/j.rser.2014.03.031

[75] Yang, Y., Yang, W., Chen, H., & Li, Y. (2020). China's energy whistleblowing and energy supervision policy: An evolutionary game perspective. *Energy*, 213, Article 118774. https://doi.org/10.1016/j.energy.2020.118774

[76] Coe, R. G., Ahn, S., Neary, V. S., Kobos, P. H., & Bacelli, G. (2021). Maybe less is more: Considering capacity factor, saturation, variability, and filtering effects of wave energy devices. *Applied Energy*, 291, Article 116763. https://doi.org/10.1016/j.apenergy.2021.116763

[77] Knox, S., Hannon, M., Stewart, F., & Ford, R. (2022). The (in)justices of smart local energy systems: A systematic review, integrated framework, and future research agenda. *Energy Research & Social Science*, 83, Article 102333. https://doi.org/10.1016/j.erss.2021.102333

[78] Jang, M.-J., & Oh, E. (2024). Deep-reinforcement-learning-based vehicle-to-grid operation strategies for managing solar power generation forecast errors. *Sustainability*, 16(9), Article 3851. https://doi.org/10.3390/su16093851

[79] Sadeghian, O., Oshnoei, A., Kheradmandi, M., Khezri, R., & Mohammadi-Ivatloo, B. (2020). A robust data clustering method for probabilistic load flow in wind-integrated radial distribution networks. *International Journal of Electrical Power & Energy Systems*, 115, Article 105392. https://doi.org/10.1016/j.ijepes.2019.105392

[80] Javaid, N., Javaid, S., Abdul, W., Ahmed, I., Almogren, A., Alamri, A., & Niaz, I. A. (2017). A hybrid genetic wind-driven heuristic optimization algorithm for demand side management in smart grid. *Energies,* 10(3), Article 319. https://doi.org/10.3390/en10030319

[81] Wu, C., Gao, S., Liu, Y., Song, T. E., & Han, H. (2021). A model predictive control approach in microgrid considering multi-uncertainty of electric vehicles. *Renewable Energy*, 163, 1385–1396. https://doi.org/10.1016/j.renene.2020.08.137

[82] Hosny, M., Kamel, S., Dominguez-Garcia, J. L., & Molu, R. J. J. (2023). Integrating renewable energy and V2G uncertainty into optimal power flow: A gradient bald eagle search optimization algorithm with local escaping operator. *IET Renewable Power Generation*. https://doi.org/10.1049/rpg2.12874

[83] He, Q., Yang, Y., Luo, C., Zhai, J., Luo, R., & Fu, C. (2022). Energy recovery strategy optimization of dual-motor drive electric vehicle based on braking safety and efficient recovery. *Energy*, 248, Article 123543. https://doi.org/10.1016/j.energy.2022.123543

[84] Singh, R. A., Kumar, R. S., Bajaj, M., & et al. (2024). Machine learning-based energy management and power forecasting in grid-connected microgrids with multiple distributed energy sources. *Scientific Reports*, 14, Article 19207. https://doi.org/10.1038/s41598-024-70336-3

[85] Perera, A. T. D., & Kamalaruban, P. (2021). Applications of reinforcement learning in energy systems. Renewable and Sustainable Energy Reviews, 137, Article 110618. https://doi.org/10.1016/j.rser.2020.110618

[86] Weil, J., Bao, Z., Abboud, O., & Meuser, T. (2024). Towards generalizability of multi-agent reinforcement learning in graphs with recurrent message passing. In Proceedings of the International Conference on Autonomous Agents and Multiagent Systems (AAMAS). arXiv:2402.05027. https://doi.org/10.48550/arXiv.2402.05027

[87] Aladdin, S., El-Tantawy, S., Fouda, M.M., & Tag Eldien, A.S. (2020). MARLA-SG: Multi-Agent Reinforcement Learning Algorithm for Efficient Demand Response in Smart Grid. *IEEE Access*, 8, 210626-210639.

[88] Antonini, E. G. A., & Caldeira, K. (2021). Atmospheric pressure gradients and Coriolis forces provide geophysical limits to power density of large wind farms. *Applied Energy*, 281, Article 116048. https://doi.org/10.1016/j.apenergy.2020.116048

[89] Suresh, V., Muralidhar, M., & Kiranmayi, R. (2020). Modeling and optimization of an off-grid hybrid renewable energy system for electrification in rural areas. *Energy Reports*, 6, 594–604. https://doi.org/10.1016/j.egyr.2020.01.013

[90] Kim, S., Kim, K., & Son, C. (2020). Transient system simulation for an aircraft engine using a data-driven model. *Energy,* 196, Article 117046. https://doi.org/10.1016/j.energy.2020.117046

[91] Qi, C., Zhu, Y., Song, C., Yan, G., Xiao, F., Wang, D., Zhang, X., Cao, J., & Song, S. (2022). Hierarchical reinforcement learning based energy management strategy for hybrid electric vehicles. *Energy*, 238, Part A, Article 121703. https://doi.org/10.1016/j.energy.2021.121703

[92] Ahadi, R., Ketter, W., Collins, J., & Daina, N. (2022). Cooperative learning for smart charging of shared autonomous vehicle fleets. *Transportation Science*, 57(3), 613–630. https://doi.org/10.1287/trsc.2022.1187

[93] Albogamy, F. R., Khan, S. A., Hafeez, G., Murawwat, S., Khan, S., Haider, S. I., Basit, A., & Thoben, K.-D. (2022). Real-time energy management and load scheduling with renewable energy integration in smart grid. *Sustainability*, 14(3), Article 1792. https://doi.org/10.3390/su14031792

[94] Martone, T., Iob, P., Schiavo, M., & Cenedese, A. (2024). Smart fleet solutions: Simulating electric AGV performance in industrial settings. Paper presented at the 2024 IEEE International Conference on Emerging Technologies and Factory Automation (ETFA2024). *arXiv:*2408.12498.

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

[96] Chonsalasin, D., Champahom, T., Jomnonkwao, S., Karoonsoontawong, A., Runkawee, N., & Ratanavaraha, V. (2024). Exploring the influence of Thai government policy perceptions on electric vehicle adoption: A measurement model and empirical analysis. *Smart Cities*, 7(4), 2258–2282. https://doi.org/10.3390/smartcities7040089

[97] Kokchang, P., Chattranont, N., Menaneatra, T., & Srianthumrong, S. (2023). Economic feasibility of hybrid solar-powered charging station with battery energy storage system in Thailand. *International Journal of Energy Economics and Policy*, 13(3), 342–355. https://doi.org/10.32479/ijeep.14258

[98] Salazar, E. J., Jurado, M., & Samper, M. E. (2023). Reinforcement learning-based pricing and incentive strategy for demand response in smart grids. *Energies*, 16(3), Article 1466. https://doi.org/10.3390/en16031466

[99] Wattana, B., & Wattana, S. (2022). Implications of electric vehicle promotion policy on the road transport and electricity sectors for Thailand. *Energy Strategy Reviews*, 42, Article 100901. https://doi.org/10.1016/j.esr.2022.100901

[100] Suanpang, P., & Jamjuntr, P. (2024). Optimal electric vehicle battery management using Q-learning for sustainability. *Sustainability*, 16(16), 7180. https://doi.org/10.3390/su16167180

[101] Suanpang, P., Niamsorn, C., Pothipassa, P., Chunhapataragul, T., Netwong, T., & Jermsittiparsert, K. (2022). Extensible metaverse implication for a smart tourism city. *Sustainability*, 14(21), 14027. https://doi.org/10.3390/su142114027

[102] Zhou, Z., Liu, G., & Tang, Y. (2023). Multi-agent reinforcement learning: Methods, applications, visionary prospects, and challenges. *ACM Transactions on Intelligent Systems and Technology*, 1(1), 43 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

[103] Dantzig, G. B. (1963). Linear programming and extensions. Princeton University Press.

[104] Karmarkar, N. (1984). A new polynomial-time algorithm for linear programming. *Combinatorica*, 4(4), 373–395. https://doi.org/10.1007/BF02579150

[105] Fletcher, R. (1987). Practical methods of optimization (2nd ed.). John Wiley & Sons.

[106] Gomory, R. E. (1963). An algorithm for integer solutions to linear programs. *Administrative Science Quarterly*, 8(3), 245–262.

[107] Lawler, E. L., Lenstra, J. K., Rinooy Kan, A. H. G., & Shmoys, D. B. (1986). The traveling salesman problem: A guided tour of combinatorial optimization. John Wiley & Sons.

[108] Bellman, R. (1957). Dynamic programming. Princeton University Press.

[109] Lauinger, D., Caliandro, P., Van Herle, J., & Kuhn, D. (2016). A linear programming approach to the optimization of residential energy systems. *Journal of Energy Storage*, 7, 24–37. https://doi.org/10.1016/j.est.2016.04.009

[110] Farajnezhad, Mohammad, Jason See Toh Seong Kuan, and Hesam Kamyab. "Impact of Economic, Social, and Environmental Factors on Electric Vehicle Adoption: A Review." *Eido*s, vol. 17, no. 24, June 2024, pp. 39-62. Eidos, doi:10.29019/eidos.v17i24.1380.

[111] Yilmaz, M., Valluri, P., & Pagadrai, S. (2012). An integer programming power optimization in storage systems. In C. H. Dagli (Ed.), Complex adaptive systems, Publication 2 (pp. 326–331). Procedia Computer Science, 12. https://doi.org/10.1016/j.procs.2012.09.079

[112] Huffman, D. A. (1952). A method for the construction of minimum-redundancy codes. Proceedings of the IRE, 40(9), 1098–1101. https://doi.org/10.1109/JRPROC.1952.273898

[113] Holland, J. H. (1975). Adaptation in natural and artificial systems. University of Michigan Press.

[114] Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. Science, 220(4598), 671–680. https://doi.org/10.1126/science.220.4598.671

[115] Cerny, V. (1985). Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm. *Journal of Optimization Theory and Applications*, 45(1), 41–51. https://doi.org/10.1007/BF00940843
[116] Glover, F., & Laguna, M. (1997). Tabu search. Kluwer Academic Publishers.
[117] Dorigo, M., Maniezzo, V., & Colorni, A. (1996). Ant system: Optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics),* 26(1), 29–41. https://doi.org/10.1109/3477.484436
[118] Dorigo, M., & Stützle, T. (2004). Ant colony optimization. MIT Press.
[119] Yan, N., Zhao, H., Ma, S., & Yan, T. (2021). Research on energy management and control method of microgrid considering health status of batteries in echelon utilization. *Energy Reports*, 7(Supplement 1), 389–395. https://doi.org/10.1016/j.egyr.2021.01.055
[120] Yu, Z., Cheng, S., Gu, R., Li, Y., Dai, S., & Mao, Z. (2021). Interfacial solar evaporator for clean water production and beyond: From design to application. *Applied Energy*, 299, 117317. https://doi.org/10.1016/j.apenergy.2021.117317
[121] Kusunda, K. (2019). Optimal electricity cost minimization of a grid-interactive pumped hydro storage using ground water in a dynamic electricity pricing environment. *Energy Reports*, 5, 159–169. https://doi.org/10.1016/j.egyr.2019.01.004
[122] Xie, J., Ajagekar, A., & You, F. (2023). Multi-agent attention-based deep reinforcement learning for demand response in grid-responsive buildings. *Applied Energy*, 342, 121162. https://doi.org/10.1016/j.apenergy.2023.121162
[123] Zafar, S., Blavette, A., Camilleri, G., Ben Ahmed, H., & Agbodjan, J.-J. P. (2023). Decentralized optimal management of a large-scale EV fleet: Optimality and computational complexity comparison between an adaptive MAS and MILP. *International Journal of Electrical Power & Energy Systems*, 147, 108861. https://doi.org/10.1016/j.ijepes.2022.108861
[124] Wu, H., Qiu, D., Zhang, L., & Sun, M. (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
[125] Zhang, J., Zhang, Z., Han, S., & Lü, S. (2022). Proximal policy optimization via enhanced exploration efficiency. *Information Sciences*, 609, 750–765. https://doi.org/10.1016/j.ins.2022.07.111
[126] Geng, Sijia, Thomas Lee, Dharik Mallapragada, and Audun Botterud. "An Integer Clustering Approach for Modeling Large-Scale EV Fleets with Guaranteed Performance." *Electric Power Systems Research*, vol. 236, November 2024, 110650. https://doi.org/10.1016/j.epsr.2024.110650.
[127] Charbonnier, F., Morstyn, T., & McCulloch, M. D. (2022). Scalable multi-agent reinforcement learning for distributed control of residential energy flexibility. *Applied Energy*, 314, 118825. https://doi.org/10.1016/j.apenergy.2022.118825
[128] Wiedemann, N., Xin, Y., Medici, V., Nespoli, L., Suel, E., & Raubal, M. (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
[129] Hao, X., Chen, Y., Wang, H., Wang, H., Meng, Y., & Gu, Q. (2023). A V2G-oriented reinforcement learning framework and empirical study for heterogeneous electric vehicle charging management. *Sustainable Cities and Society*, 89, 104345. https://doi.org/10.1016/j.scs.2022.104345
[130] Ali, H., Hussain, S., Khan, H. A., & Khan, I. (2020). Economic and environmental impact of vehicle-to-grid (V2G) integration in an intermittent utility grid. In 2020 2nd International Conference on Smart Power & Internet Energy Systems (SPIES). Bangkok, Thailand. https://doi.org/10.1109/SPIES48661.2020.9242992