Artificial Intelligence-Based Adaptive Control in AC/DC Microgrid Energy Management

by

Zhalgas Bekeyev

Submitted to the Department of Electrical and Computer Engineering
in partial fulfillment of the requirements for the degree of

Master of Science in Electrical and Computer Engineering

at the

NAZARBAYEV UNIVERSITY

June 2025

Author .....
Department of Electrical and Computer Engineering
20.04.2025

Certified by .....
Prashant Kumar Jamwal
Associate Professor
Thesis Supervisor

Certified by .....
Gulsim Kulsharova
Assistant Professor
Thesis Co-Supervisor

Accepted by .....
Yelizaveta Arkhangelsky
Dean, School of Engineering and Digital Sciences

Artificial Intelligence-Based Adaptive Control in AC/DC Microgrid Energy Management

by

Zhalgas Bekeyev

Submitted to the Department of Electrical and Computer Engineering
on 20.04.2025, in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical and Computer Engineering

Abstract

The integration of artificial intelligence into adaptive control systems is a novel strategy for enhancing energy management in AC/DC microgrids. The research constructs an extensive simulation model in MATLAB Simulink that replicates a hybrid microgrid system linked to the utility grid, focusing on renewable energy sources such as solar panels and wind turbines. Instead of relying on traditional fixed-parameter Proportional-Integral (PI) controllers, reinforcement learning (RL) agent have been used which constantly adjusts control settings based on changes in load and generation conditions. The Twin-Delayed Deep Deterministic Policy Gradient (TD3) method proved to be a much better fit for the job. The RL-based controller helped stabilize the DC link voltage and reduce those annoying overshoots during transients, making the system more efficient and adaptable. Experimental results show that, unlike traditional methods, this approach does a better job of handling system nonlinearity and the unpredictable nature of renewable energy sources. Overall, this adaptive control system highlights how AI-driven solutions can improve the stability, reliability, and overall performance of complex microgrid systems.

Thesis Supervisor: Prashant Kumar Jamwal
Title: Associate Professor

Thesis Co-Supervisor: Gulsim Kulsharova
Title: Assistant Professor

Acknowledgements

I would like to express my deepest gratitude to all those who have supported and guided me during my research and thesis preparation.

Firstly, I would like to thank my thesis supervisor, Dr. Prashant Kumar Jamwal and thesis co-supervisor Dr. Gulsim Kulsharova, for their continuous guidance, valuable feedback, and encouragement throughout the course of this work. Their expertise and support have been invaluable. I would also like to extend my sincere appreciation to all the faculty members and staff in the Department of Electrical and Computer Engineering for their help and encouragement during my time at Nazarbayev University.

Special thanks to my friends and classmates for their moral support and for providing a motivating and collaborative environment. Writing the thesis together and motivating each other was unforgettable experience.

I sincerely appreciate everyone's contribution to the completion of this work.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Today, electricity is an integral part of the world and every technology depends on it. Currently, the main energy production comes from the combustion of coal, natural gas, and oil. Although it is the main energy source, it has a significant negative impact on the environment [1] and these resources may run out [2]. Therefore, renewable energy has become a focal point for research as a result of the emergency of depletion of mineral resources in the future. One of the main solutions is a microgrid. It is a system that includes different types of renewable energy system, such as solar and wind energy. Microgrids can provide reliable, pure and efficient energy, and the productivity and stability of the main electric power increases [2]. There are two types of microgrids: alternating-current microgrids and direct-current microgrids. In Figure 1a, the detailed structure of the AC-based microgrid is given. An AC-based microgrid system can easily supply the AC-distributed systems, therefore it is cost efficient [3]. In Figure 1b, the structure of the DC-based microgrid is shown. Compared to AC-based microgrids, it works faster, because it does not consider the frequency and reactive power, and the power efficiency is high. However, since houses and factories are mainly based on AC, they need an additional DC to AC converter which will consider reactive power and frequency. However, there exists an issue related to renewable energies. Since solar and wind energy depend on the weather conditions, energy forecasting is ineffective in producing enough energy. Consequently, missing energy will be taken from the grid. Adaptive control in microgrids is necessary to determine the state of the load and renewable energies [4]. There are mainly three types of modes: enough energy from renew-

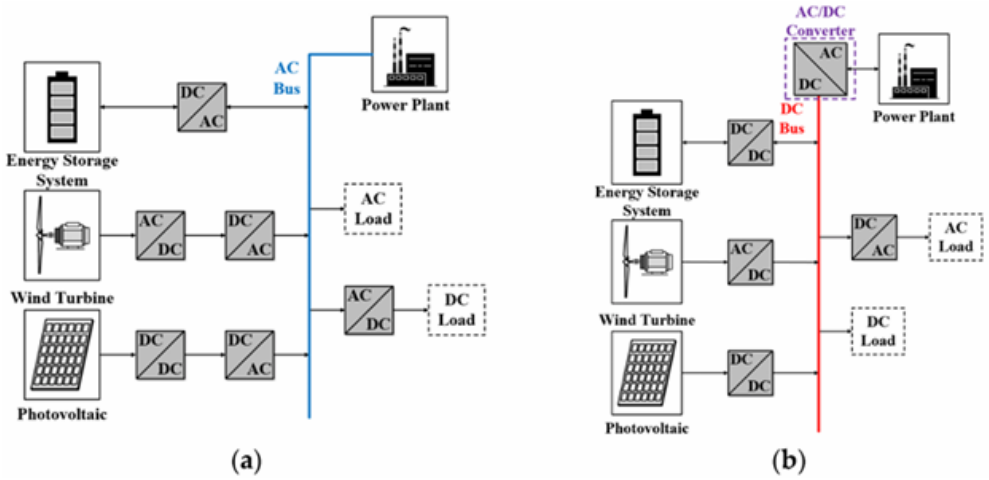

Figure 1-1: (a) AC-based microgrid, (b) DC-based microgrid.

able energy side, insufficient amount of energy provided from renewable energies, therefore grid will supply the load with energy, excess amount of renewable energy, which will directly go to the grid. After the selection of the mode, it helps to maintain the system, by calibrating supplied power, estimating the power loss in converters, and taking into account frequency change and power factor of the system. However, energy optimization and estimation of missing power with mode selection should be more precise and accurate.

As shown in Figure 1(b), the DC-to-AC converter, which is an inverter, employs its own control technique. This technique includes phase locked loop (PLL), abc-to-dq0 transformation, and a proportional-integral (PI) controller or Model Predictive Control (MPC). In the case of the PI controller, as system complexity increases, it becomes difficult to tune the PI controller coefficients accurately using a transfer function. Furthermore, the PI controller parameters are fixed, making it less effective for nonlinear systems since it is primarily designed to handle linear systems [5]. However, MPC faces computational challenges, as it requires optimization at each step to update the system, making it unsuitable for real-time applications [6]. In addition, MPC relies on accurate system models, which can become difficult to adjust when dealing with multiple renewable energy systems, the utility grid, and converters, leading to complex parameter tuning. Therefore, artificial intelligence will help in energy management and distribution of the power supply between wind, solar and grid. With the integration of artificial intelligence techniques such as fuzzy logic controllers, artificial

neural networks, and reinforcement learning, the energy management of the system can be stabilized, system resilience will be improved, and the overall performance of the microgrid will be enhanced. This research's main goal is to show that artificial intelligence technique can demonstrate better results than traditional methods such as PI controller and MPC with increasing the energy efficiency and adaptability of the system. Mainly, research will be simulation-based study which focuses MATLAB Simulink using blocks of the microgrid and using MATLAB functions to support the Simulink based model. MATLAB functions will include artificial intelligence technique algorithms.

This project will be organized into several chapters: literature review, methodology, results, discussion, and conclusion. The literature review will cover traditional methods of microgrid energy management, explaining each part of the microgrid and its working principles. Additionally, it will provide information on PI controllers and Model Predictive Control (MPC), as well as discuss existing artificial intelligence-based approaches related to the project. This section will highlight the gaps in the field and suggest potential improvements to enhance the developed systems. The methodology chapter will focus on the artificial intelligence techniques employed in the simulation. It will provide a detailed explanation of the system, including calculations for the variables of the renewable energy system. The proposed model, the training method, and the hyperparameters used in the AI techniques will also be discussed to clarify the working principles of the system. The results chapter will present the outcomes obtained from the proposed model, accompanied by figures and numerical values. In the discussion chapter, a comparison will be made between the performance of the proposed AI technique and the PI controller. This section will analyze the advantages and disadvantages of each approach, including comparison graphs and an evaluation of the different hyperparameters and their impact on the results. The limitations of the system will also be addressed. Finally, the conclusion chapter will summarize the entire project, emphasizing its contributions to the field and providing suggestions for future work.

Chapter 2

Literature Review

2.1 Microgrid

The microgrid is the system that obtains energy from renewable energy and supplies with it the load and the grid.

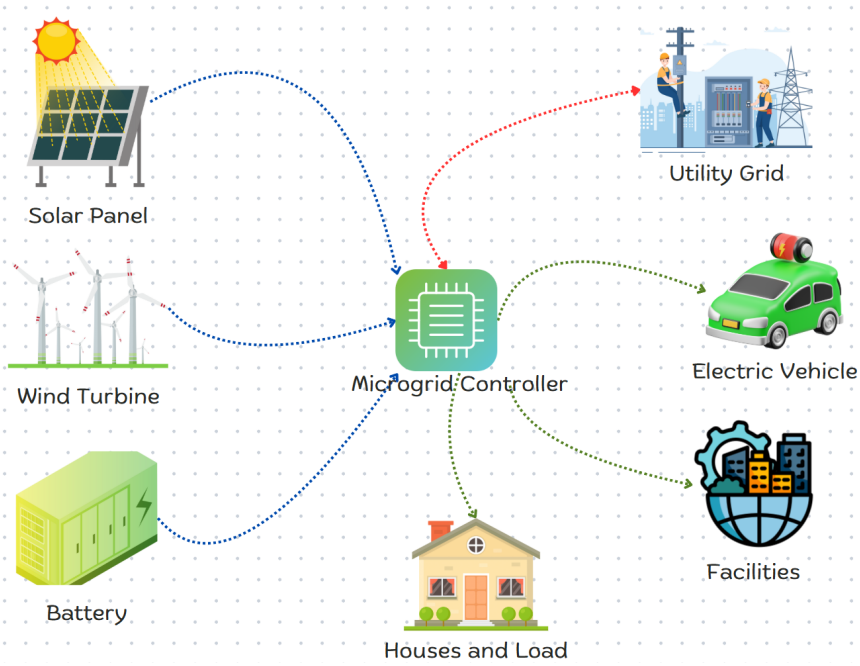

Figure 2-1: Microgrid system.

Figure 2.1 illustrates the general structure of the microgrid system. It consists primarily of solar photovoltaic panels (PV), wind turbines, and battery storage as renewable energy

sources. Loads such as facilities, electric vehicles, and other electrical appliances consume the power generated by these renewable energy sources. In addition, a microgrid controller manages and regulates the energy flow, ensuring the stable and efficient operation of the entire system.

2.1.1 Working Principle

The main difference between a microgrid and a conventional grid is that a microgrid can operate independently or in parallel with the central utility grid. When operating independently, it can supply power autonomously without relying on the central grid. Alternatively, it can connect to the main grid to either draw additional power when demand exceeds local generation or feed surplus power back into the utility grid.

In grid-connected mode, the utility grid and the microgrid can exchange power between each other. Specifically, in this mode, energy flows bidirectionally depending on demand and generation [7]. When the microgrid produces excess energy, the surplus power can flow back into the utility grid. Conversely, when the microgrid's generation is insufficient to meet the load demand, the utility grid provides the necessary additional power. Furthermore, in the connected mode of the grid, electricity cost management is crucial to optimize energy pricing and improve overall energy efficiency [8].

In islanded mode, the microgrid operates independently, without a connection to the utility grid. During this autonomous operation, the microgrid is entirely based on renewable energy sources and battery storage systems [9]. The availability of renewable energy depends mainly on the weather conditions, while battery storage ensures stable power supply during periods of low generation. Effective operation in islanded mode requires careful management to maintain voltage stability, frequency control, and overall system reliability [10]. In addition, since renewable sources are primarily dependent on weather conditions, proper battery management is crucial to maintain consistent power availability under variable conditions [11].

Furthermore, there exists a microgrid control hierarchy [12]. There are primary, secondary and tertiary control. Primary control hierarchy is responsible for immediate changes. Basically, it controls system disturbances and changes the power consumption of the load. This

layer ensures stable voltage and frequency in the microgrid through fast response mechanisms [13]. At this level, stable voltage and frequency are maintained in the microgrid by fast response mechanisms that represent droop control, in which distributed energy resources adapt their output power, in real-time, according to frequency and voltage deviations from setpoints [14].

Based upon the primary control actions, the secondary control restaurants have their voltages and frequencies displaced from nominal [15]. This degree of control considers long-term changes since it corrects errors through communication-based control methods. The response of different resources in the microgrid is coordinated by a centralized or distributed architecture of the secondary control [16].

Tertiary control supervises the function of the microgrid, such as optimizing the economic performance and interacting with the main grid. It ensures that the grid operates long-run economically and sustainably by influencing aspects such as operational planning, energy management strategies, market participation, and demand management [12].

To run a microgrid properly there should be a strong control and communication architectures. These consist of centralized, decentralized, and distributed control schemes [12]:

In the centralized control architecture, a centralized controller collects information from all the components in the microgrid and makes control decisions. Simple as it is, centralized control exposes a single point of failure risk and demands an extensive communication infrastructure [17].

Decentralized Control: Instead, decentralized control provides local controllers to the components, which diminishes the extent to which a wide-scale communication network is needed. This limits coordination though and can result in lower system performance, but it does improve reliability [18].

Distributed control: Distributed control is a hybrid of local and global controllers, combining the benefits of centralized and decentralized methodologies. Friction from implementing complex coordination between control units, local control units can carry out tasks autonomously, but they communicate with a central hub for a discreet time to improve the functionality of the entire system by allowing them to communicate what to do next within the limit of available communication resources [19].

Particularly in the case of microgrid systems with high renewable penetration and dynamic load conditions, the effective integration and interaction of these control architectures play a critical role to achieve the operation of microgrid systems in an efficient, stable, and robust manner.

2.1.2 Solar Panel

A solar photovoltaic (PV) panel converts sunlight directly into electricity. It operates using silicon-based cells that produce electric power through the photovoltaic effect. When sunlight strikes these silicon cells, photons transfer their energy to electrons, generating electricity. By configuring these cells in series or parallel arrangements, the PV panel's output voltage and current can be tailored to meet specific requirements. Specifically, connecting cells in series increases voltage, while connecting them in parallel increases the current.
Moreover, the system incorporates Maximum Power Point Tracking (MPPT), a technique designed to continuously optimize the solar panel's output by adjusting current and voltage [20]. This ensures that the solar panel consistently operates at its maximum power output, despite fluctuations in sunlight intensity and other external conditions. Furthermore, a DC-to-DC boost converter is utilized to step up and regulate the voltage to a stable, desired level while simultaneously decreasing current to minimize power losses. This significantly enhances energy efficiency and improves overall system stability.

2.1.3 Wind Turbines

While photovoltaic cells convert photons from sunlight directly into electricity, wind turbines generate electrical energy by converting kinetic energy from wind. A typical wind turbine system consists of several key components. The rotor blades convert wind energy into mechanical energy. This mechanical energy can then be regulated and converted into electrical energy. Specifically, the rotor blades' rotation can vary based on their design and orientation. Wind turbine systems are classified mainly into two types: horizontal-axis wind turbines (HAWT) and vertical-axis wind turbines (VAWT). As their names imply, their primary difference is in their axis of rotation is horizontal for HAWTs and vertical for VAWTs

[21].

The mechanical energy produced by the rotor blades drives a permanent magnet synchronous generator (PMSG), producing alternating current (AC) [22]. Since AC output from the generator needs to be stabilized, it's converted into direct current (DC) using an AC-to-DC rectifier. A DC-to-DC boost converter is then employed to step up and regulate the voltage level, ensuring consistent and stable power output. The rectification from AC to DC, coupled with voltage regulation, helps minimize fluctuations and improves the system's efficiency and stability.

Furthermore, wind energy systems are particularly suited to windy regions, such as Astana, due to their reliance on consistent and sufficient wind speeds for optimal performance.

2.1.4 Battery Energy Storage System (BESS)

The battery storage system facilitates effective energy management within the microgrid by operating in different modes, depending on the availability of renewable energy and load demand. There are five primary operational modes [23]:

1. **Mode 1:** Renewable energy generation is sufficient to meet the demand of the load.

2. **Mode 2:** Renewable energy generation is insufficient; however, the battery storage provides the deficit energy to meet load demands.

3. **Mode 3:** Renewable energy generation meets the load requirements, with excess energy available. In this case, the battery is charged using surplus power.

4. **Mode 4:** Renewable energy generation supplies the load and charges the battery fully. Excess energy, beyond battery capacity, is transferred back to the utility grid.

5. **Mode 5:** Renewable energy generation is insufficient to supply the load, and the battery storage is unable to cover the energy shortfall. Thus, additional required energy is drawn from the utility grid.

2.2 Adaptive Control Techniques

Adaptive control is the main technique used in the energy management of microgrids. Mainly two of them are in use, however there exist at least 7-8 adaptive control techniques.

2.2.1 Proportional Integral Derivative

Proportional Integral Derivative is one of the main adaptive control techniques used in microgrids. Microgrids use them to increase the stability of the system by accurate control of voltage and frequency of the system. For this purpose, metaheuristic algorithms can be applied to optimize the PID controller [24]. It has been proven by simulations which say that the proposed methodology is working. By controlling PID coefficients such as voltage and frequency using metaheuristic algorithms, the stability and efficiency of the system will be increased. Good voltage regulation can provide stability to the system and the power of the system will not vary; however, regulation of the voltage can be achieved through a PID controller [25]. The voltage regulation through the PID controller will give more efficiency than manual or heuristic algorithms. However, computation complexity for PID is missing. Including this type of adaptive control in microgrids will be more efficient than manual control of the voltage and other hyperparameters.

2.2.2 Model Predictive Control

Model Predictive Control is another type of adaptive control. It is not based on voltage and frequency regulation. Mainly it is based on controlling the full system with predictive models and cost functions [26]. It can decide a given small amount of time which increases the response from the system [27]. Depending on the level of the microgrid, it can be varied. Converter level and grid level have different prediction models because they are not the same and do not give similar voltage. Energy efficiency with a reliable and stable system without any voltage variation can be achieved through MPC. Based on the microgrid, the specific MPC should be chosen such as linear, nonlinear, stochastic, and distributed MPCs [27]. It depends on their response time and formulation of the prediction model. Moreover, it is more cost-effective than PID. Even though PID and MPC achieved energy efficiency in

simulations, the real-world applications are missing and for PID computational complexity is not defined. However, every adaptive control method can be optimized by integration of artificial intelligence techniques into it.

2.2.3 Sliding Mode Control

Sliding mode control is a nonlinear adaptive control technique. It employs a predetermined manifold, referred to as the sliding surface, to force the system state to converge onto this surface. One of the key advantages of SMC is its resilience to uncertainties and disturbances, as well as its ability to achieve finite-time convergence, which enables rapid stabilization in complex nonlinear systems. Additionally, its design process is relatively straightforward compared to other nonlinear methods. However, several drawbacks exist. The most notable is chattering, a phenomenon characterized by high-frequency oscillations around the sliding surface. These oscillations can result in undesirable high-frequency dynamics, making practical implementation challenging. Furthermore, in real-world applications, the finite sampling time of digital controllers necessitates smoothing or other modifications to reduce chattering. Finally, external noise can interfere with the ideal switching condition required for SMC, exacerbating chattering and diminishing performance [28, 29].

2.2.4 Adaptive Backstepping

By breaking down the overall control objective into more manageable stages, adaptive backstepping controllers use Lyapunov stability theory to regulate system outputs. At each iteration, a control Lyapunov function (CLF) is generated to incorporate both state and parameter estimation errors, ensuring that the derivative of the Lyapunov function is negative semi-definite. This property guarantees asymptotic or finite-time convergence of the tracking error. A key advantage of adaptive backstepping is its robustness: controller parameters are treated as unknowns and are continuously estimated in real time, thereby maintaining stability without requiring precise system models [30, 31]. Moreover, as shown in [31], an adaptive reference signal can reduce steady-state errors and total harmonic distortion (THD), thus improving power quality in inverter-based microgrids. Because of their nonlin-

ear design, these controllers operate effectively under a wide range of conditions, including demand fluctuations and intermittent renewable energy. Despite these benefits, there are some drawbacks. In contrast to simpler approaches (e.g., PI controllers), adaptive backstepping requires careful tuning of multiple observer settings and gains, adding to design complexity. Additionally, measurement noise may degrade control performance and undermine the accuracy of parameter estimates. Real-time parameter estimation and computation also incur a higher computational cost compared to more traditional methods.

2.2.5 Artificial Neural Network

Neural networks are mathematical models that have been created to solve certain problems without the help of the human brain. Hence, they are capable of learning complex patterns and relationships from data, making them effective in modeling nonlinear and dynamic behaviors in microgrids. An energy management system was proposed for the AC/DC hybrid microgrid by using an artificial neural network (ANN) which includes 6 different modes of controlling [32]. The results demonstrated were promising because it was cost-effective, stable and system efficiency was increased. Another proposed system was similar to the previous one, the only difference being that the new system included a two-stage ANN [4]. It includes 2 main stages where the first stage is given to identify the mode, and the other is to find the power of the battery that will be used to supply the system. However, unlike the previous study, it does not include the training algorithms and optimization techniques. All things considered, this research presents a new approach to energy flow optimization in hybrid AC/DC microgrid systems, which may enhance their dependability, efficacy, and affordability. However, proposed energy management strategies (EMS) can be applied only to small-scale microgrids [4, 32].

2.2.6 Fuzzy Logic Controller

Another type of artificial intelligence technique is fuzzy logic controllers. Fuzzy logic systems offer a framework for expressing and simulating uncertainty in decision-making processes. They are capable of handling nonlinearity and imprecise inputs to microgrid systems. Hence,

by integrating expert knowledge into the fuzzy rule sets, controllers can develop adaptive control strategies based on the current state of the microgrid. An energy allocation solution based on adaptive optimal fuzzy logic (AOFL) has been proposed as a solution to the real-time optimal energy management problem in a microgrid [33]. The main innovation of the system was the integration of fuzzy logic controllers with meta-heuristic optimization techniques. Adaptability and robustness of their solution by comparing it with benchmark strategies, such as online Rule-Based Control (RBC) and offline scheduling based on meta-heuristic optimization. Another research created a model to facilitate efficient regional EMS planning under various uncertainties [34]. The model is a fuzzy-random interval programming (FRIP). The system included probabilistic modeling techniques to estimate uncertainties associated with various factors. Also, the authors deploy mathematical optimization models, such as stochastic programming and robust optimization, to explore decision alternatives and associated risks. The developed solutions offer energy resource/service optimality allocation that minimize system costs, maximize system dependability, and maximize energy security. Nevertheless, unlike the above-mentioned research, there is no validation that includes the real-world application of the proposed system.

2.2.7 Reinforcement Learning

The third type of artificial intelligence technique for energy forecasting is reinforcement learning. Reinforcement learning (RL) is defined as a technique that trains software to make decisions through the trial-and-error learning process to achieve the most optimal results. In microgrid energy management systems it is used to adaptively adjust control actions based on performance feedback. The system which uses model-free RL and deep neural network (DNN) was proposed for multi-microgrid (MMG) [35]. Mainly, one of the RL methods Monte Carlo is used. The distribution system operator (DSO) seeks cost-effective energy selling and minimizing the demand peak-to-average ratio (PAR) simultaneously. The simulation findings reported in the research demonstrate that when faced with issues that lack precise mathematical explanations or have a high computational complexity Monte Carlo approach exhibits strong exploration ability when compared to an intuitive selection method. However, a new RL algorithm is suggested for scheduling a battery for a user [36]. The goal is to create

effective energy forecasting in ever-changing settings to increase profit. The system uses a method called Q-learning for optimization. This method looks at rewards or punishments for actions based on the environment's changes. From the simulation results, the proposed RL algorithm in comparison with other RL techniques demonstrates promising results in energy forecasting by battery scheduling. However, both RL algorithms did not have real-world applications, but the Monte-Carlo approach can deal with high computational complexity and another algorithm concentrates mainly only on a single consumer [35, 36]. According to the literature review, there are mainly three types of artificial techniques that can enhance energy management in AC/DC microgrids. For this purpose, the main drawbacks and advantages of the different AI techniques will be analyzed. Every technique does not include real-world applications, however, simulations proved that the energy management of the microgrid can be boosted depending on the technique. Neural networks can propose a method that chooses the battery by itself, but it applies to only small-scale microgrids. Fuzzy logic controllers maximize the stability and security of the microgrids while dealing with nonlinearity and imprecise inputs of the system. Reinforcement learning is a process that works in a trial-and-error strategy. It works to optimize the battery scheduling, and it can work where high computations are required.

Chapter 3

Methodology

This chapter provides a comprehensive overview of the methods and approaches employed throughout the research for the thesis. Moreover, this study is going to contribute to the power estimation and PI controller tuning using reinforcement learning. This chapter will include two main sections: microgrid and reinforcement learning. It includes all essential graphical representations, mathematical model and system model.

Basically, the main objective of the project is to tune PI controller using the reinforcement learning approach to avoid manual tuning. Currently, tuning the PI controller in the Simulink can be done using 2 types of techniques: Transfer Function Based and Frequency Response Based. Both technique results do not always show perfect results. Therefore, reinforcement learning based model is going to be developed, which will predict proportional coefficient (Kp) and integral coefficient (Ki) and adapt to dynamic conditions.

Additionally, this chapter outlines the specific research methods and approaches utilized throughout the thesis. It provides a clear explanation of the procedures and justifications for selecting the chosen techniques. Moreover, the main goal of this research is to introduce reinforcement learning into the model, to increase the PI controller value optimization and search for optimal values Kp and Ki.

This section will include two main subsections:

1. Microgrid in Simulink.

2. Reinforcement Learning Based PI controller.

Specifically, they will discuss the microgrid model which will be used, component discussions and their calculations via mathematical formulas, control technique and reinforcement learning implementation through Simulink.

3.1 Microgrid

This section will consist of three different subtopics. There are photovoltaic panels, wind turbines and grid connected system. Overall, the main model is described in figure 3.1.

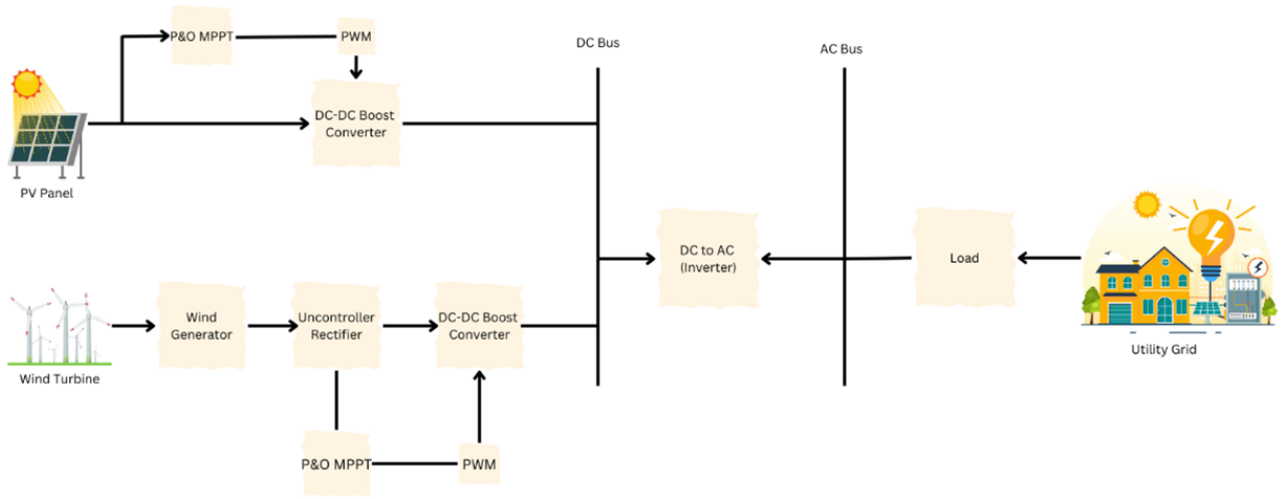

Figure 3-1: Proposed model.

In the figure above, the proposed microgrid model is illustrated. Photovoltaic (PV) panels are connected to a DC-to-DC boost converter to minimize power dissipation and increase voltage levels. Additionally, a Perturbation and Observation (P&O) and Maximum Power Point Tracking (MPPT) algorithm is used to get a Pulse Width Modulation (PWM) signal which directly connected to the DC-to-DC boost converter. The output current from the

PV system is subsequently fed into the DC bus.

The wind turbine subsystem is comparatively more complex. Initially, the wind turbine is connected to a wind generator, producing alternating current (AC). This AC is then converted to direct current (DC) through a rectifier. Like the PV panels, the wind turbine system is also integrated with a DC-to-DC boost converter and employs a parallel P&O MPPT controller. The resulting DC current is then delivered to the DC bus.

On the grid side, the main grid connects directly to the load and the AC bus. An intermediate DC-to-AC bidirectional inverter is placed between the DC bus and the AC bus, facilitating bidirectional energy flow. Detailed explanations of each subsystem and their interactions will be discussed comprehensively in the following subchapters.

3.1.1 Photovoltaic Panel

In this subchapter, everything that was included in the photovoltaic system will be discussed. In Figure 3.2, the proposed photovoltaic panel model is presented. It consists of four main

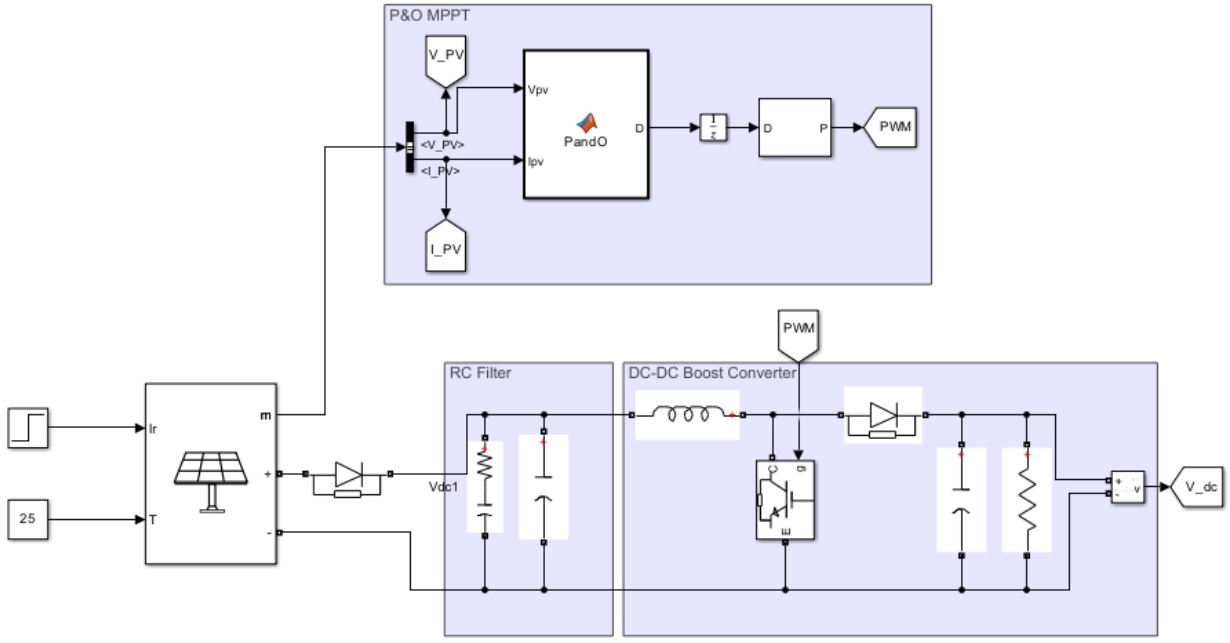

Figure 3-2: Photovoltaic Panel Model.

parts: a perturbation and observation-based maximum power point tracking (MPPT) module, a DC-DC converter, a photovoltaic panel, and the load.

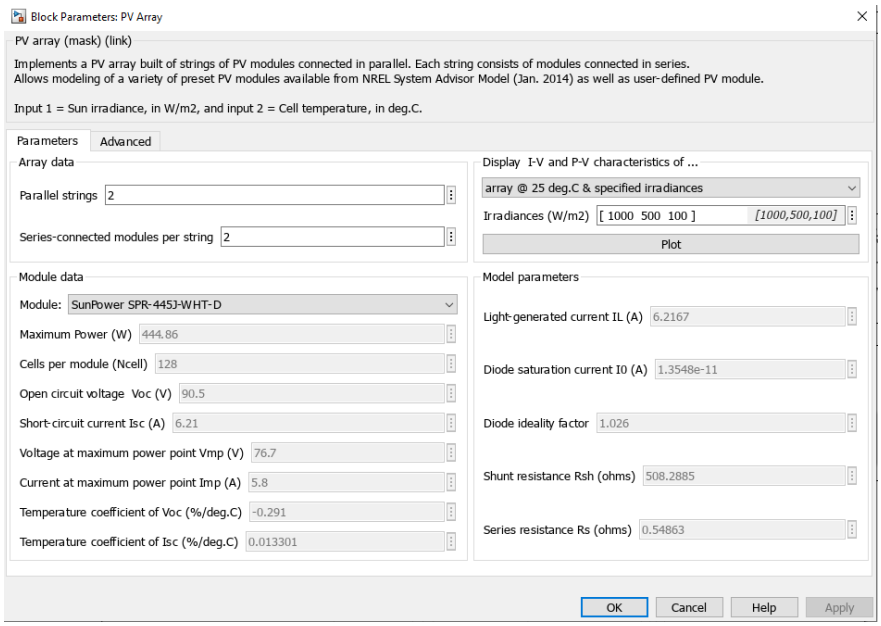

Figure 3-3: Photovoltaic Panel Description.

The photovoltaic panel model selected for this study was the "SunPower SPR-445J-WHT-D." The primary inputs for the PV panel are solar irradiation and temperature. In this analysis, a temperature of 25℃ and irradiation of 1000 W/m^2 were used, as these represent standard test conditions (STC) for photovoltaic panels. Figure 3.2 shows the proposed photovoltaic panel configuration, which includes four main components: a Perturbation and Observation (P&O) Maximum Power Point Tracking (MPPT) system, a DC-DC converter, the photovoltaic panel, and the load.

As depicted in Figure 3.2, the system configuration clearly outlines the connections among these parts. Additionally, Figure 3.2 demonstrates how the components interact to optimize performance. Figure 3.3 presents the voltage and current values at maximum power($V_{\text{mp}} = 153.4\,\text{V}$ and $I_{\text{mp}} = 5.8\,\text{A}$) will be utilized in the maximum power calculations. The maximum power can be obtained as:

$$P_{\text{max}} = I_{\text{max}} \cdot V_{\text{max}} \tag{3.1}$$

Where every value changes according to the number of connections in parallel and in series:

$$V_{\max} = N_{\text{series}} \cdot V_{\text{mp}} = 2 \cdot 76.7\,\text{V} = 153.4\,\text{V},$$

$$I_{\max} = N_{\text{parallel}} \cdot I_{\text{mp}} = 2 \cdot 5.8\,\text{A} = 11.6\,\text{A},$$

$$P_{\max} = 11.6 \cdot 153.4 = 1779.44\,\text{W}.$$

All calculations can be checked using figure 3.4. As shown in Figure 3.2, the system includes

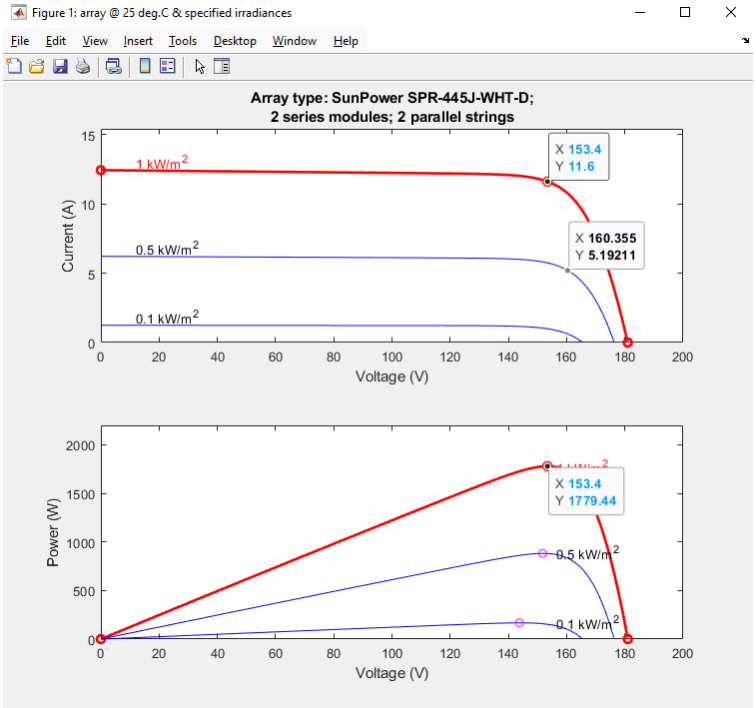

Figure 3-4: Power and Current of the PV Panel.

an RC filter and a diode. The diode prevents current from flowing back into the photovoltaic panel, protecting it from potential reverse currents. Additionally, the RC filter reduces voltage ripples originating from the photovoltaic panel, resulting in a smoother voltage output. This contributes significantly to improved stability and overall system efficiency. The cutoff frequency of an RC filter is given by

$$f_c = \frac{1}{2\pi RC} \tag{3.2}$$

where f_c is the cutoff frequency. Low ripple from DC voltage can be obtained if the cutoff

frequency is small. Here, $f_c = 1.5\,\text{Hz}$ with $RC = 0.1$ (where $R = 20\,\Omega$ and $C = 5300\,\mu\text{F}$). The DC-DC boost converter includes an inductor (L), a diode, and a capacitor (C). The diode prevents current from flowing back toward the photovoltaic panel, while the capacitor reduces voltage ripples, resulting in a smoother output. The inductor stores and releases energy to boost voltage levels. Together, these components help improve the stability and efficiency of the system. Typically, the design aims for minimal ripple to ensure stable operation, with standard ripple values selected based on around 5% variation from the target output voltage.

$$\Delta I_L = 0.05 \cdot \frac{P_{\max}}{I_{\max}} = 0.05 \cdot \frac{1779.44}{153.4} \approx 0.58\,\text{A}$$

The switching frequency (f_s) is set to 10 kHz. Additionally, the desired output voltage (V_o) for the model is 700 V. Using these parameters, the duty cycle (d) can be calculated. This duty cycle determines the switching operation of the DC-DC boost converter and is essential for achieving the targeted output voltage.

$$d = 1 - \frac{V_{\max}}{V_o} = 1 - \frac{153.4}{700} \approx 0.78$$

After obtaining duty cycle, inductor value can be obtained using the formula given below: The inductor value is calculated as:

$$L = \frac{V_0(1-d)}{\Delta I_L f_s} = \frac{700(1-0.78)}{0.58 \times 10000} \approx 26.55\,\text{mH}$$

Calculation of the Resistive Load (R_{load}):

$$R_{\text{load}} = \frac{V_o^2}{P_{\max}} = \frac{700^2}{1779.44} \approx 275.4\,\Omega$$

Now, the value of the capacitor can be obtained using the following formula, where ΔV_o is 1% of V_o (voltage ripple):

$$C = \frac{V_0 d}{2\Delta V_o R_{\text{load}} f_s} = \frac{700 \times 0.78}{2 \times 0.01 \times 700 \times 275.4 \times 10000} \approx 14.16\,\mu\text{F}$$

The IGBT switch requires a Pulse Width Modulation (PWM) input. This PWM signal is

determined by the duty cycle calculated using the Perturbation and Observation (P&O) Maximum Power Point Tracking (MPPT) method. At each step, the main parameters—voltage (V_{pv}), current, and power (P_{pv})—are updated based on logical conditions.
Initially, the previous voltage (V_{prev}) and previous power (P_{prev}) are both set to zero, while the previous duty cycle (D_{prev}) is initialized to 0.78. A fixed step size is defined for the perturbation. In each step, the differences in voltage (dV) and power (dP) are calculated and evaluated against logical conditions to update the duty cycle.

- If power (dP) increases, the algorithm checks the voltage difference (dV):
  - If voltage also increases, the duty cycle continues increasing by adding the step size.
  - If voltage decreases, the duty cycle is decreased by the step size.
- Conversely, if the power decreases:
  - When the voltage increases, the perturbation direction is reversed, and the duty cycle decreases.
  - If voltage decreases as well, the duty cycle increases.

After each iteration, these logical conditions update the duty cycle, voltage, and power values ($d, V_{\text{pv}}, P_{\text{pv}}$), optimizing power extraction from the PV panel.
The PWM signal generated from the calculated duty cycle at a switching frequency of 10 kHz is then fed into the IGBT switch. Finally, the voltage output from the photovoltaic panel is directed toward the DC bus.

3.1.2 Wind Turbine

In this subchapter, the wind turbine system is discussed. Figure 3.5 illustrates the proposed model for the wind turbine system, which consists of a wind turbine, a permanent magnet synchronous generator (PMSG), an uncontrolled rectifier, and an RC filter.
Figure 3.5 shows that the main inputs affecting the wind turbine system are wind speed and the pitch angle of the blades. Maximum power capture occurs under standard conditions,

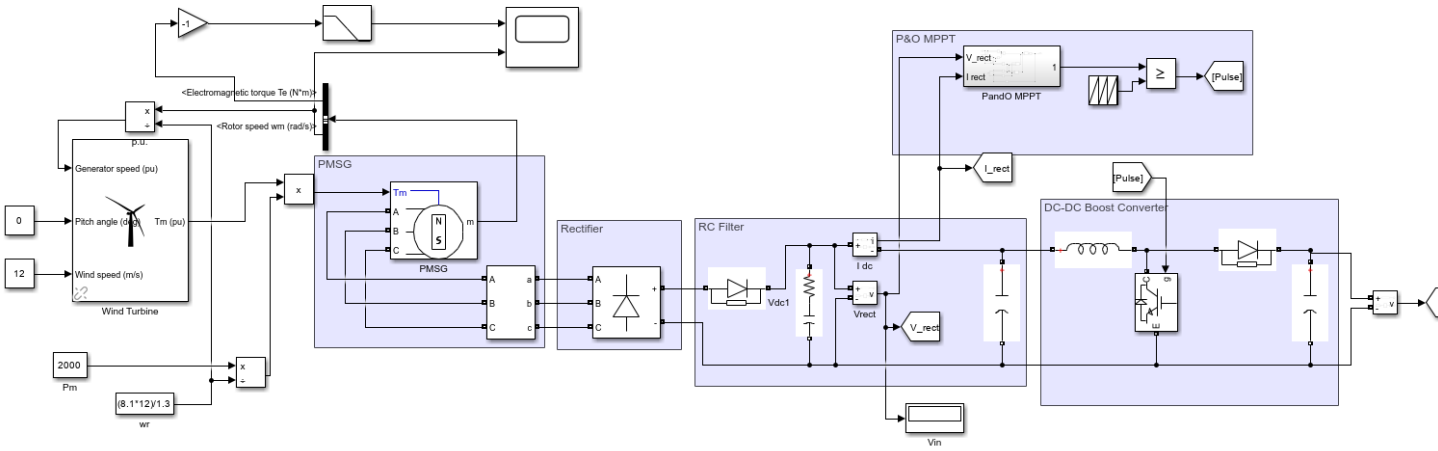

Figure 3-5: Wind Turbine Model.

specifically when the pitch angle is set to 0° and the wind speed is relatively low. Adjusting the pitch angle significantly affects the aerodynamic efficiency and, consequently, the maximum power that can be captured. At higher wind speeds, increasing the pitch angle reduces the aerodynamic effect and limits mechanical power, helping maintain stable and controlled power output.

In figure 3.6, the ideal wind speed can be found related to the power. As it shows, the

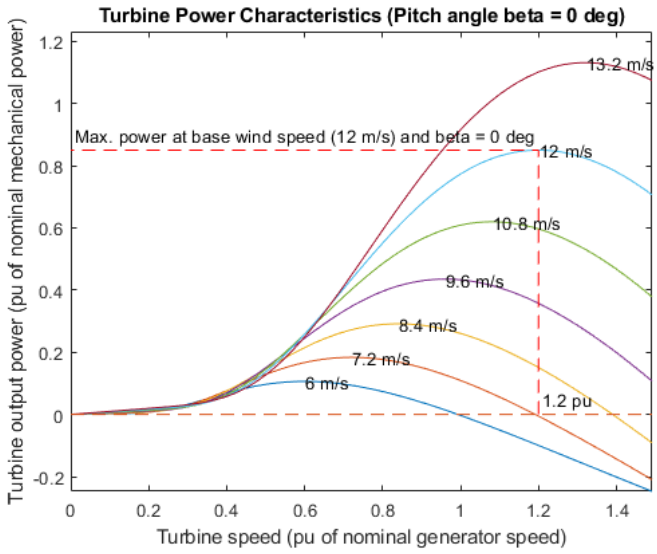

Figure 3-6: Wind Turbine Power Characteristics.

maximum power can be obtained at pitch angle equal to 0, the wind speed is 12 meters per second. Figure 3.7 shows the main description of the wind turbine. The nominal mechanical

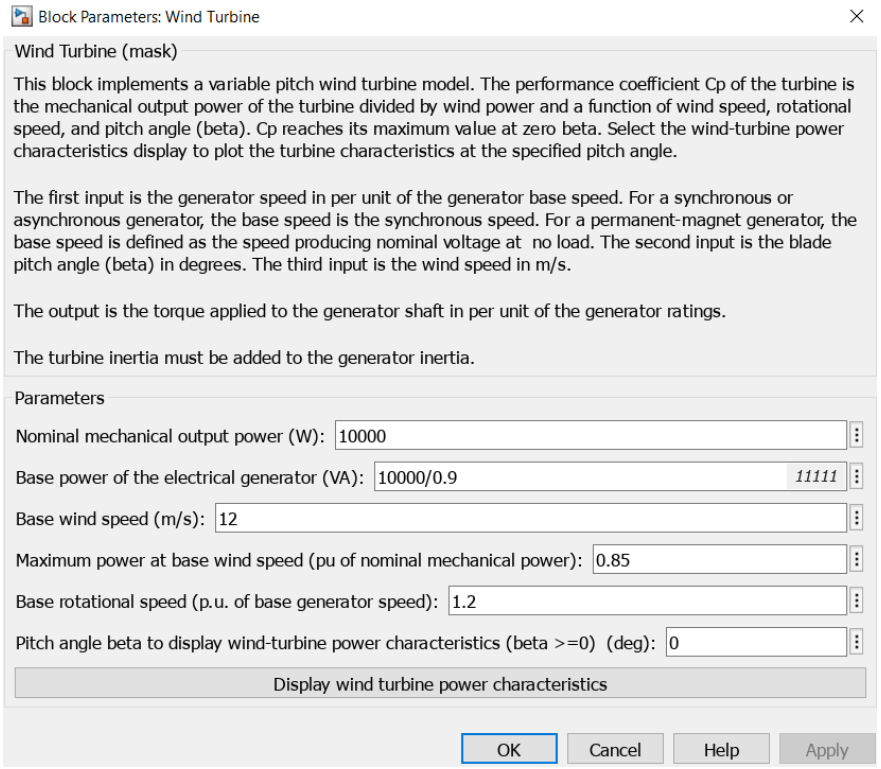

Figure 3-7: Wind Turbine Description.

power of the wind turbine is 10 kW. The base power of the electrical generator is 11111 VA. To calculate the generator speed in the per-unit system, the rotor speed (ω_r) should be determined using the following formula:

$$\omega_r = \frac{V_{\text{wind}} \cdot \lambda}{R} = \frac{12 \times 8.1}{1.3} \approx 74.77\,\text{rad}\,\text{s}^{-1}$$

where:

- V_{wind} – wind speed,
- λ – tip speed ratio,
- R – effective gear factor.

The tip-speed ratio and the power coefficient are typically given under standard test conditions. The output of the wind turbine is expressed as mechanical torque in a per-unit system. Additionally, the actual mechanical torque can be calculated using:

$$T_m = \frac{P_m}{\omega_r} = \frac{10000}{74.77} \approx 133.75\,\mathrm{N\,m}$$

By multiplying the mechanical torque obtained in per unit by the rated torque, the actual mechanical torque input to the generator can be determined. The outputs of the Permanent Magnet Synchronous Generator (PMSG) include mechanical torque, rotor speed, and three-phase electrical power.

The rotor speed also feeds back into the system. The ratio of the actual rotor speed to the rotor speed obtained from the PMSG gives the generator speed in the per-unit system. The three-phase electrical power produced by the generator enters an uncontrolled rectifier, which converts alternating current (AC) into direct current (DC). This rectification is necessary to integrate the generated power into a DC bus. However, a major disadvantage of the rectifier is the significant ripple in the direct current. The DC voltage can be calculated as:

$$V_{\mathrm{dc}} = \frac{3\sqrt{2}}{\pi} V_{\mathrm{ll,rms}}$$

As discussed earlier, the RC filter prevents current from flowing in the reverse direction and reduces voltage ripple, resulting in smoother current flow and enhanced system performance. The Perturbation and Observation (P&O) MPPT and DC-DC boost converter operate based on the same principles described previously.

To design the boost converter, required values for the inductor (L_{wind}), duty cycle (d), capacitor (C_{wind}), and resistive load (R_{wind}) must be calculated. The switching frequency is set to 10 kHz, the desired output voltage is 700 V, and the input voltage from the generator is 78.5 V. Additionally, the electrical power converted from mechanical power is calculated to be 6.045 kW. Using these parameters, appropriate values for each component of the boost converter can be determined. The duty cycle (d) is calculated as:

$$d = 1 - \frac{V_{\mathrm{wind}}}{V_{o,\mathrm{wind}}} = 1 - \frac{78.5}{700} \approx 0.888$$

The rectifier current (I_{rect}) is obtained using:

$$I_{\text{rect}} = \frac{P_{\text{rect}}}{V_{\text{wind}}} = \frac{6045}{78.5} \approx 77\,\text{A}$$

Using the formula above, the current from the rectifier can be calculated as:

$$\Delta I_L = 0.2 \cdot I_o \cdot \frac{V_{o,\text{wind}}}{V_{\text{wind}}} = 0.2 \cdot 77 \cdot \frac{700}{78.5} \approx 137.3\,\text{A}$$

Inductor Calculation: After obtaining the duty cycle, the inductor value (L) can be determined using:

$$L = \frac{V_{\text{wind}}(V_{o,\text{wind}} - V_{\text{wind}})}{\Delta I_L \cdot f_s \cdot V_{o,\text{wind}}}$$

$$= \frac{78.5(700 - 78.5)}{137.3 \times 10000 \times 700} \approx 50.76\,\mu\text{H}$$

Resistive Load Calculation: The resistive load (R_{load}) is given by:

$$R_{\text{load}} = \frac{V_o^2}{P_{\text{max}}} = \frac{700^2}{1800} \approx 272\,\Omega$$

Capacitor Calculation: The capacitor value (C) can be determined using:

$$C = \frac{I_{\text{rect}} \cdot d}{\Delta V_o \cdot f_s}$$

$$= \frac{22.9 \times 0.888}{0.01 \times 700 \times 10000} \approx 976.8\,\mu\text{F}$$

The Perturbation and Observation (P&O) MPPT algorithm provides the input signal to the IGBT switch. The MPPT scheme used here is identical to the one previously described for the solar panel system.

3.1.3 Utility Grid & Inverter

In this subchapter, the grid and inverter descriptions will be discussed.

The grid connection is illustrated in the figure above. It is a three-phase AC system con-

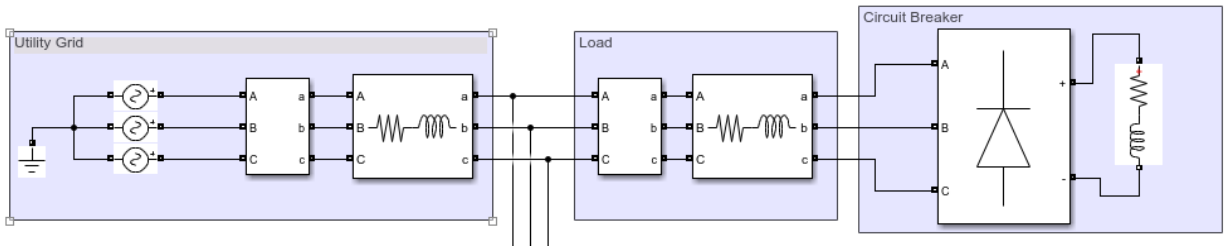

Figure 3-8: Utility grid in Simulink.

nected to the output of the inverter, which provides alternating current. The grid's peak voltage is 328 V. As shown, the utility grid and inverter output from renewable energy sources are connected in parallel. If the load requires more power, the deficit is provided by the utility grid. Conversely, if there is excess power generated, it flows back into the utility grid.

The configuration also includes an RC filter, which reduces voltage ripple and improves the stability and efficiency of the overall system. The RC filter smooths current and prevents reverse current flow. Additionally, if the load requires more power than the inverter can supply, the utility grid provides the deficit. Conversely, excess power generated by the renewable system is fed back into the grid.

The inverter system comprises six Insulated-Gate Bipolar Transistor (IGBT) switches, la-

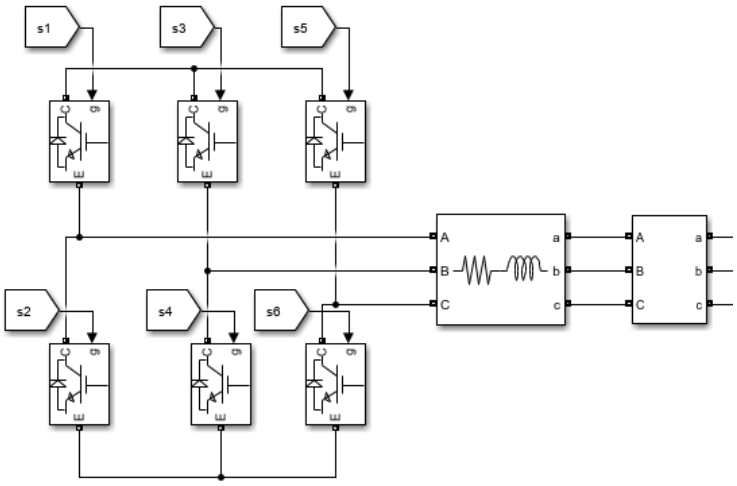

Figure 3-9: Bidirectional inverter.

beled S1 through S6, as shown in the figure. These switches convert DC power, received

from the DC-DC boost converter, into three-phase AC power. Specifically, switches S1 and S2 generate the first phase, switches S3 and S4 handle the second phase, and switches S5 and S6 produce the third phase. Furthermore, the Pulse Width Modulation (PWM) control signals, derived from the inverter control technique, regulate these switches. The inverter output then enters the grid via an RC filter and parallel connection.

This configuration also includes an RL branch, functioning as a load and stabilizing the output. A bi-directional flow of power occurs when the load demand exceeds the generated renewable energy, additional energy flows from the grid. In contrast, when generation exceeds consumption, excess power is sent back into the grid. The overall arrangement enhances efficiency, stability, and ensures effective power management. The inverter requires

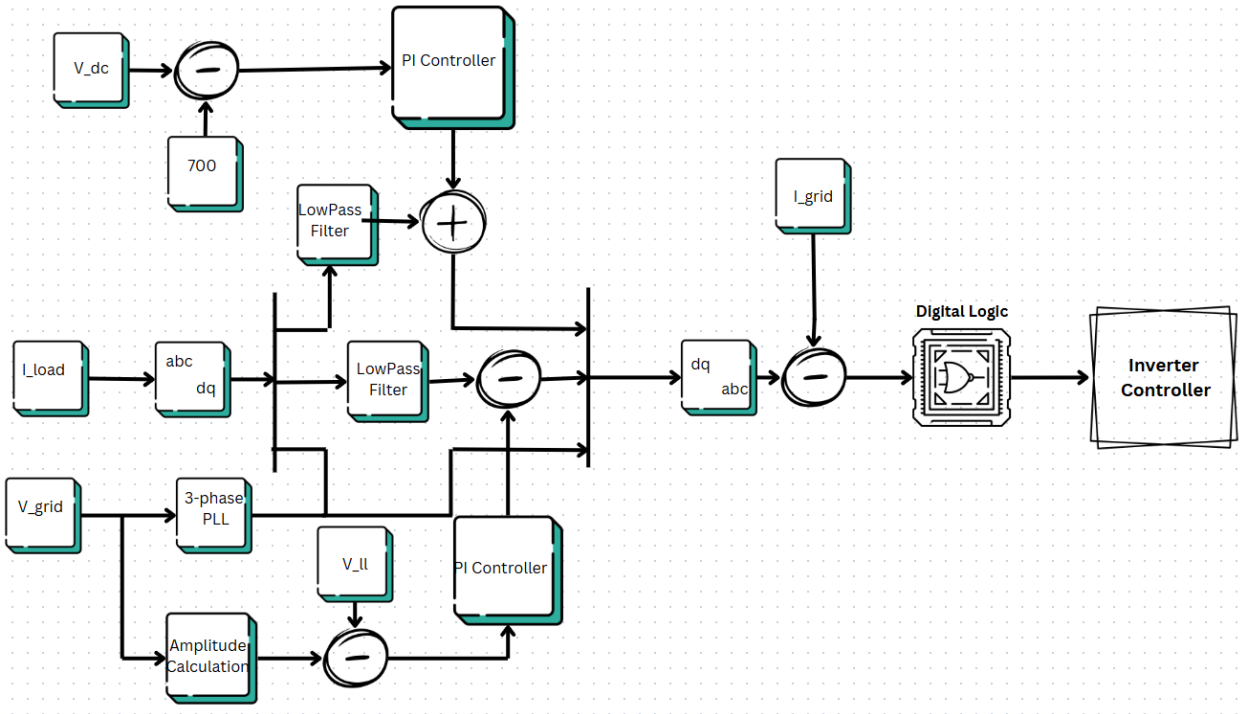

Figure 3-10: Inverter Control Technique.

PWM input, controlled by the Perturbation and Observation (P&O) Maximum Power Point Tracking (MPPT) algorithm, which operates similarly to the solar panel configuration discussed earlier.

The inverter structure includes six Insulated-Gate Bipolar Transistor (IGBT) switches labeled from S1 to S6. These switches convert the DC voltage from the boost converter into

a three-phase AC voltage. Each pair of switches generates one phase of the three-phase output: switches S1 and S2 control the first phase, S3 and S4 handle the second phase, and S5 and S6 manage the third phase.

As depicted in the figure above, the inverter is connected to the grid, forming a three-phase AC system. The peak grid voltage is 328 V. An RC filter is incorporated to reduce voltage ripples and harmonics, enhancing the overall system performance by ensuring a smoother current flow and preventing reverse current from flowing back to the inverter.

The inverter control system uses a feedback loop that involves the measurement of the load current (I_{load}) via three-phase voltage-current (V-I) measurement. This measured load current enters an abc-to-dq0 converter, transforming three-phase currents into direct (d-axis) and quadrature (q-axis) components, representing active and reactive power, respectively. A low-pass filter is utilized to mitigate high-frequency noise from the measured signals. The mathematical formula is given below.

$$\begin{bmatrix} I_d \\ I_q \\ I_0 \end{bmatrix} = \frac{2}{3} \begin{bmatrix} \cos(\theta) & \cos\left(\theta - \frac{2\pi}{3}\right) & \cos\left(\theta + \frac{2\pi}{3}\right) \\ -\sin(\theta) & -\sin\left(\theta - \frac{2\pi}{3}\right) & -\sin\left(\theta + \frac{2\pi}{3}\right) \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} I_{La} \\ I_{Lb} \\ I_{Lc} \end{bmatrix}.$$

As shown, it takes three-phase current load and the $\cos(\theta)$ and $\sin(\theta)$ are calculated using the phase lock loop and 3 phase grid voltages. Moreover, low-pass filter as filter to avoid the AC harmonics and only DC will remain. In addition, the PI controller for active power estimation works to deal with loss and the mathematical representation is given as below:

$$i_w(k) \;=\; i_w\big(k-1\big) \;+\; K_p\left(V_e(k) \;-\; V_e\big(k-1\big)\right) \;+\; K_i\left(V_e(k)\right).$$

where:

- V_{e} – difference between desired and actual voltages,
- k – number of sampling,
- $i_{\text{w(k)}}$ – IGBT switch loss for kth sampling.

Inputs for inverse operation of abc/dq0 transformation are amplitude of reference which is the summation of output (I_{w}) and the lowpass filter output. As quadrature axis serves the difference between line to line voltage which is 328 V and the amplitude of the grid voltage which enters to the another PI controller for reactive power estimation. The mathematical representation given below:

$$i_r(k) \;=\; i_r\big(k-1\big) \;+\; K_p\left(V_{ae}(k) - V_{ae}\big(k-1\big)\right) \;+\; K_i\, V_{ae}(k).$$

where:

- V_{ae} – difference between line to line voltage and actual grid voltages,
- $i_{\mathrm{r(k)}}$ – reactive current kth sampling.

Moreover, inputs for dq0 to abc trnasformation are in mathematical calculation are given below:

$$I_d^{\mathrm{new}} = \mathrm{LPF}(I_d) + I_w,$$

$$I_q^{\mathrm{new}} = \mathrm{LPF}(I_q) + I_r.$$

Additionally, the inverter controller compares the calculated amplitude from the dq0 transformation with the actual line-to-line voltage of 328 V. The resulting signals are converted back to three-phase currents via a dq0-to-abc converter. The difference between the actual grid current (I_{grid}) and the estimated current is then fed into digital logic circuits, including logical gates. This logic ultimately controls the switching operation of the IGBTs. The mathematical representation of the reverse transformation is given below:

$$\begin{bmatrix} I_a^{\mathrm{new}} \\ I_b^{\mathrm{new}} \\ I_c^{\mathrm{new}} \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 1 \\ \cos\!\left(\theta - \frac{2\pi}{3}\right) & -\sin\!\left(\theta - \frac{2\pi}{3}\right) & 1 \\ \cos\!\left(\theta + \frac{2\pi}{3}\right) & -\sin\!\left(\theta + \frac{2\pi}{3}\right) & 1 \end{bmatrix} \begin{bmatrix} I_d^{\mathrm{new}} \\ I_q^{\mathrm{new}} \\ I_0 \end{bmatrix}.$$

Furthermore, if the load demands more power than what the renewable energy sources can supply, the utility grid provides the additional required power. Conversely, excess power generated by renewable energy sources is fed back into the grid, enabling bidirectional energy flow.

3.2 Reinforcement Learning

In this subchapter, the proposed model for PI controller tuner will be discussed. The main objective is energy management by tuning the PI controller using artificial intelligence techniques. In the project reinforcement learning has been used to train and calculate the Kp and Ki coefficients. The overall control technique is the same, but here the PI controller has been replaced with an RL agent which trains on this specific model of the microgrid and effectively estimates coefficients. The main control technique is shown in figure 3.11 and the details are shown in figure 3.12. Moreover, the reinforcement learning based inverter control technique is similar to the PI controller based technique with difference in replacing PI controller with reinforcement learning agent.

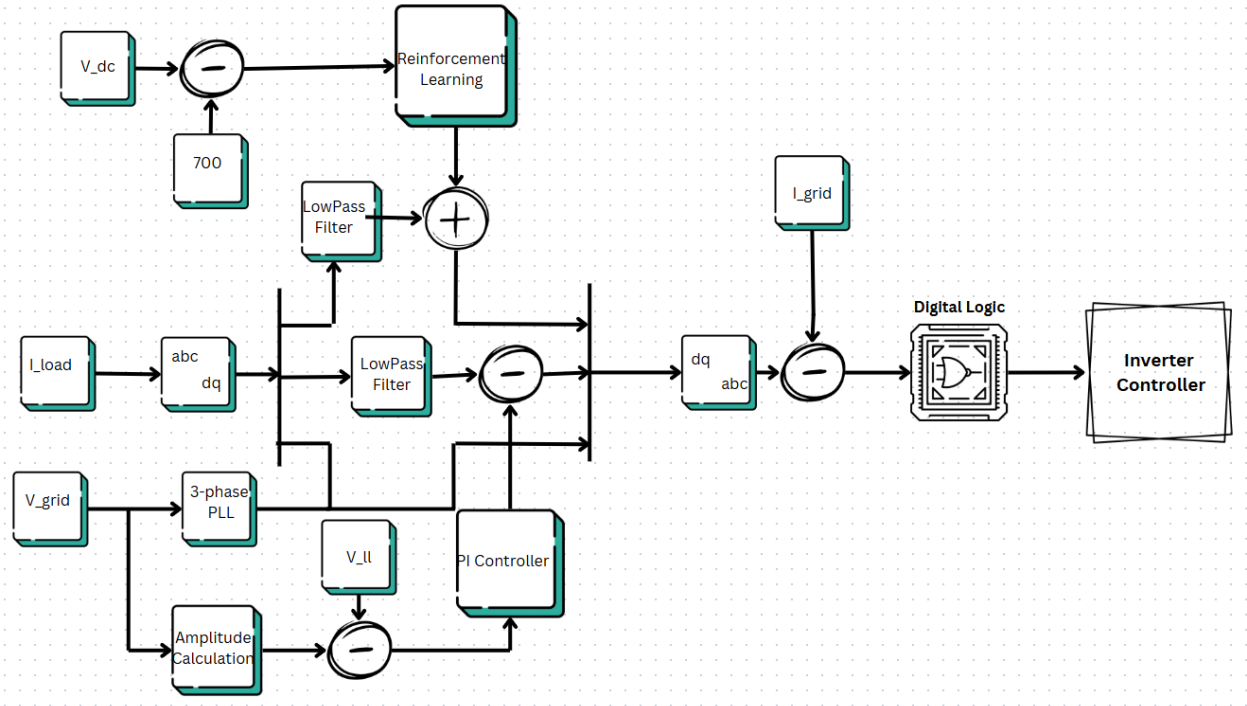

Figure 3-11: Reinforcement learning based inverter control technique.

As shown above, the difference between the desired and actual output voltage (700 V) enters the reinforcement learning (RL) block. The RL agent includes three different inputs and one output. The inputs are:

- Observation signal,

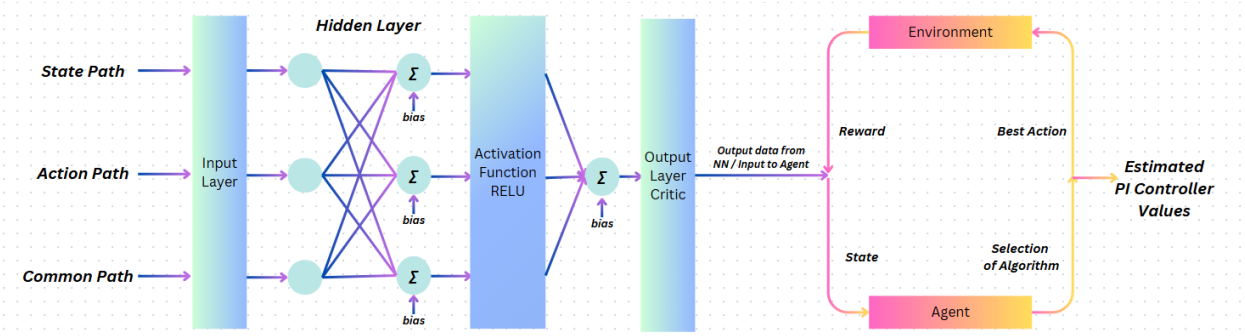

Figure 3-12: Reinforcement learning working principles.

- Reward calculation,
- Termination signal ("done" flag).

The output provides the control action for the system. This output is fed back into the system as an input to update the control decision continuously.
To prevent algebraic loops in Simulink, a unit delay block with a sampling period $T_s = 0.1\,\mathrm{s}$ has been utilized. This delay avoids computational algebraic loops by holding the previous value, allowing sequential computation.

3.2.1 Observation Block

The observation block takes the current error and the historical action output from the RL agent as inputs. Its main purpose is to generate state data, which allows the RL agent to accurately assess the current system state and choose an appropriate action accordingly. Additionally, the observation helps the RL agent monitor changes in the error over time.

3.2.2 Reward Block

The reward block calculates the reward based on two penalty terms: an error penalty and an action penalty. Specifically, the error penalty is calculated as:

$$\mathrm{error}_{\mathrm{penalty}} = \mathrm{error}^2$$

and the action penalty as:

$$\text{action}_{\text{penalty}} = 0.01 \times \text{action}^2.$$

The overall reward is then computed as:

$$\text{reward} = -\text{error}_{\text{penalty}} - \text{action}_{\text{penalty}}.$$

The main purpose of this reward calculation is to minimize both the error and the magnitude of the action taken by the RL agent. Thus, if either the error or the action value is large, the penalties increase, resulting in a lower reward. Conversely, when the error and action are both close to zero, the reward becomes higher, encouraging the system toward stable and efficient operation.

3.2.3 Done Block

The final input is the isdone flag, indicating whether to terminate the training episode. Currently, it remains set to "false," meaning training will only stop upon reaching the predefined number of epochs.

Moreover, several steps are required to build and train the reinforcement learning block with agents. First, it is necessary to set up the experimental neural network and define appropriate blocks, including the agent itself. Essential components such as the actor network, critic network, reward function, and relevant hyperparameters must be initialized. These steps enable the agent to learn effectively through interactions with the environment.

3.2.4 Actor

The actor network is represented by the policy function $\pi_\theta(s)$. This network takes the system's current state as input and determines the appropriate actions.

For the designed system, the inputs to the actor are the integrated error and its historical value. The desired target error is set to zero, meaning the output aims to minimize the error and approach this target value.

The actor network consists of fully connected neural network layers, each having its own

weights and biases, which are adjusted during training to optimize performance.

3.2.5 Q-value

The Q-value is an action-value function that represents the estimated cumulative reward. Mathematically, it is estimated using the following function:

$$Q^{\pi}(s,a) = \mathbb{E}\left\{\sum_{k=0}^{\infty} \gamma^{k} r_{t+k} \mid s_t = s, a_t = a, \pi\right\} \tag{3.3}$$

where:

- r_t – reward at time t,
- γ – discount factor, which varies between 0 and 1,
- $\mathbb{E}$ – expected value,
- s_t – state at time t,
- a_t – action at time t.

The Q-value can be used to select the best action in the current state using the following formula:

$$a^{*} = \arg\max Q^{\pi}(s,a) \tag{3.4}$$

By applying this formula, the maximum Q-value is chosen as the best action.

3.2.6 Critic

The critic is an algorithm used to estimate the Q-value. When the critic's estimation accuracy improves, the overall performance of the reinforcement learning algorithm also improves, enabling a broader and more accurate selection of actions.

In early reinforcement learning methods, only a single Q-value estimation was used; however, this approach often limited performance. Modern approaches utilize a critic network with

two separate estimations, enhancing reliability and performance.
The critic network architecture consists of three paths:

- **State path:** Takes the observation data as input and processes it through fully connected neural network layers, where the number of feature dimensions can be adjusted.
- **Action path:** Processes input corresponding to the number of actions, also through fully connected layers.
- **Common path:** Merges the outputs of the state and action paths by concatenating them and further processing the combined information.

Each neural network layer has its own adjustable weights and biases. The critic network employs the Rectified Linear Unit (ReLU) activation function to prevent the vanishing gradient problem and improve computational efficiency. The ReLU function is defined as:

$$f(x) = \max(0, x) \tag{3.5}$$

The main advantage of ReLU is that it sets negative inputs to zero, ensuring gradient stability and computational efficiency during training. Using the backpropagation algorithm combined with gradient descent optimization, the weights and biases are continuously updated until the critic's estimation becomes accurate and converges to optimal values.

3.2.7 Agent

An agent is a comprehensive object that manages the actor, critic, and memory components, using their interactions to continuously update the control logic. While various types of agents exist, this project specifically utilizes the Twin-Delayed Deep Deterministic Policy Gradient (TD3) agent. The TD3 agent incorporates two critic networks, which provide more accurate and stable estimations of Q-values. Additionally, TD3 employs delayed policy updates and adds small Gaussian noise for target smoothing. Training is performed using minibatches, which helps stabilize gradient descent updates and reduce data correlation issues.

The working principle begins with the agent receiving inputs—the current state and action—from the actor network. These inputs are evaluated by the critic network to estimate the corresponding Q-value. At every time step t, the environment returns two essential outputs: a reward and the subsequent state. The TD3 agent then uses this information to update its predictions. Specifically, the TD3 agent calculates the expected future rewards using the next action and next state according to the following formula:

$$y_t = r_t + \gamma \min \left\{ Q_{\phi_1}^{\text{target}}(s_{t+1}, a_{t+1}), Q_{\phi_2}^{\text{target}}(s_{t+1}, a_{t+1}) \right\} \tag{3.6}$$

where:

- y_t – target Q-value,
- $Q_{\phi_1}^{\text{target}}$ – Q-value of the first critic,
- $Q_{\phi_2}^{\text{target}}$ – Q-value of the second critic.

The minimum function is used to avoid overestimation. The loss function is calculated as:

$$L(\phi_i) = \left(Q_{\phi_i}(s_t, a_t) - y_t\right)^2 \tag{3.7}$$

where:

- Q_{ϕ_i} – estimated Q-value for $i = 1, 2$.

In this approach, the Mean Squared Error (MSE) is used as the loss function to quantify the difference between the predicted and actual Q-values. The MSE guides the updates of the neural network parameters, minimizing prediction errors during training. Specifically, the loss calculation helps the TD3 agent apply the gradient descent to continuously update the weights and biases, optimizing the performance of both actor and critic networks.
Reinforcement learning (RL) with the TD3 agent offers potential advantages over traditional Proportional-Integral (PI) controllers, especially in systems characterized by non-linear dynamics and requiring significant computational resources. Additionally, once trained, an RL-based controller can be effectively adapted to similar system configurations without

extensive redesign. Consequently, the RL controller may outperform the conventional PI controller by efficiently handling system complexity, non-linearity, and reducing error.

Chapter 4

Results & Discussion

In this chapter, the results of the project will be presented and discussed, with a particular focus on the comparison between reinforcement learning (RL)-based adaptive energy management and the traditional PI controller-based approach. The chapter is divided into three main subchapters. The first subchapter, RL Training, will present and analyze the results obtained during the training phase of the RL agent. The second subchapter will focus on the PI Controller-Based Microgrid, including the outcomes of grid current, load current, and the DC link voltage. The final sub-chapter will present the same set of results: grid current, load current, and DC link voltage , but using the RL-Based Control, enabling a direct comparison with the PI controller to evaluate performance differences.

4.1 Reinforcement Learning Training

This chapter presents the results of the RL training. As shown in Figure 4-1, the outcomes are promising. The episode reward shown in the figure represents the cumulative reward obtained over a single episode. In this training process, 160 episodes were used to train the RL agent.

The value Q_0 corresponds to the estimated reward at the starting point of each episode. The average reward refers to the mean of all rewards received by the RL agent during an episode, which includes both the error penalty and the action penalty.

Based on the figure, it can be observed that the RL agent is learning effectively. This is

evident from the decreasing difference between the episode's initial Q_0 value and the final episode reward. As the training progresses, the episode reward approaches zero, indicating that the agent is successfully learning to minimize both the error and the control effort, thus reducing penalties and improving performance.

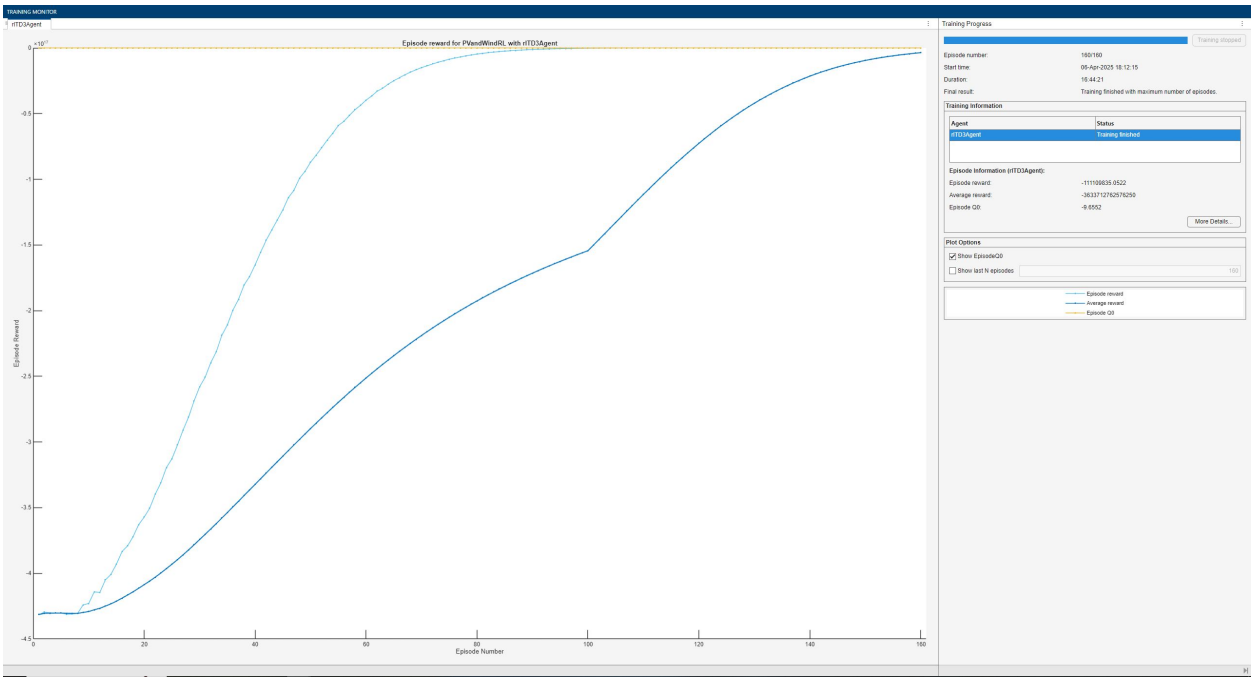

Figure 4-1: Reinforcement learning training (1 hidden layer).

To evaluate the impact of the critic network structure—specifically, the number of hidden layers in the neural network—two different architectures were used to train the reinforcement learning agent. In Figure 4-1, the training was performed using a critic with one hidden layer, while Figure 4-2 shows the results obtained using four hidden layers.

As illustrated, there is no significant difference in the training performance between the two configurations. Therefore, it can be concluded that the number of hidden layers in the critic network has a minimal effect on the overall training process. This impact is small enough to be considered negligible in this context. Additionally, both models were trained using the same learning rate of 0.001.

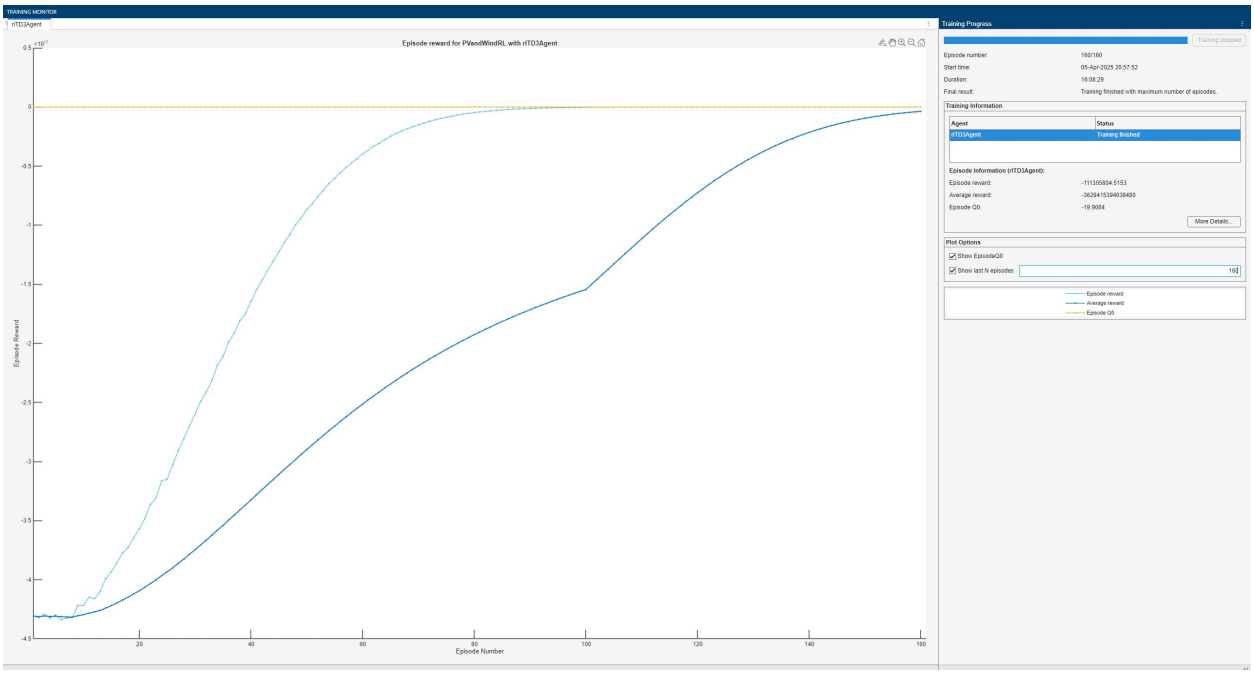

Figure 4-2: Reinforcement learning training (4 hidden layers).

4.2 PI Controller based Microgrid

The impact of the PI controller and RL on the microgrid. To assess how the PI controller affects the microgrid, it was tested on the DC-DC boost converter and the grid, with simulations of the grid and load currents.

As shown in Figure 4-3, the DC link voltage created in DC bus. The desired output voltage was 700 V; however, there are some fluctuations. After approximately 0.3 seconds, the voltage stabilizes at 700 V, but before that, it rises to 1273 V and drops to 429 V. This indicates that the PI controller, tuned using a transfer function-based tuner, causes these fluctuations. Initially, the system attempts to stabilize and adjust the power output.

Next, let's examine how this affects the grid and load. The grid voltage remains constant, generating three-phase alternating current at 50 Hz, with a peak value of 328 V, as set. However, the grid current fluctuates until it stabilizes after about 0.15 seconds, at which point it maintains a constant value. The distortion is due to the need for regulating the DC link voltage, as well as the presence of harmonics and load dynamics. The peak and trough of the current indicate that the grid is compensating for the power demand or generated power. The peak current reaches 250 A, and within the first 0.01 seconds, it starts to fluctuate. After

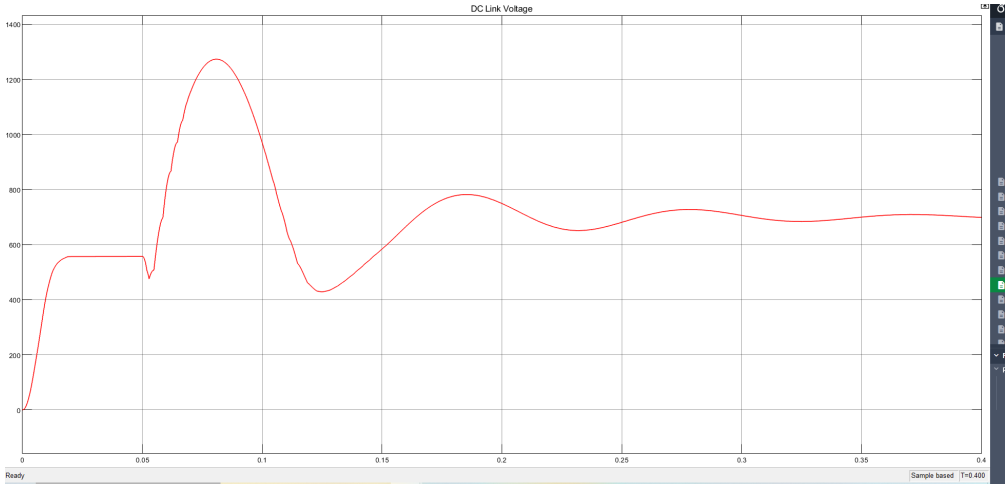

Figure 4-3: DC link voltage.

about 0.05 seconds, it stabilizes at 20 A, but then overshoots again as the DC link voltage fluctuates. The system then begins to stabilize, with more consistent behavior starting from 0.3 seconds.

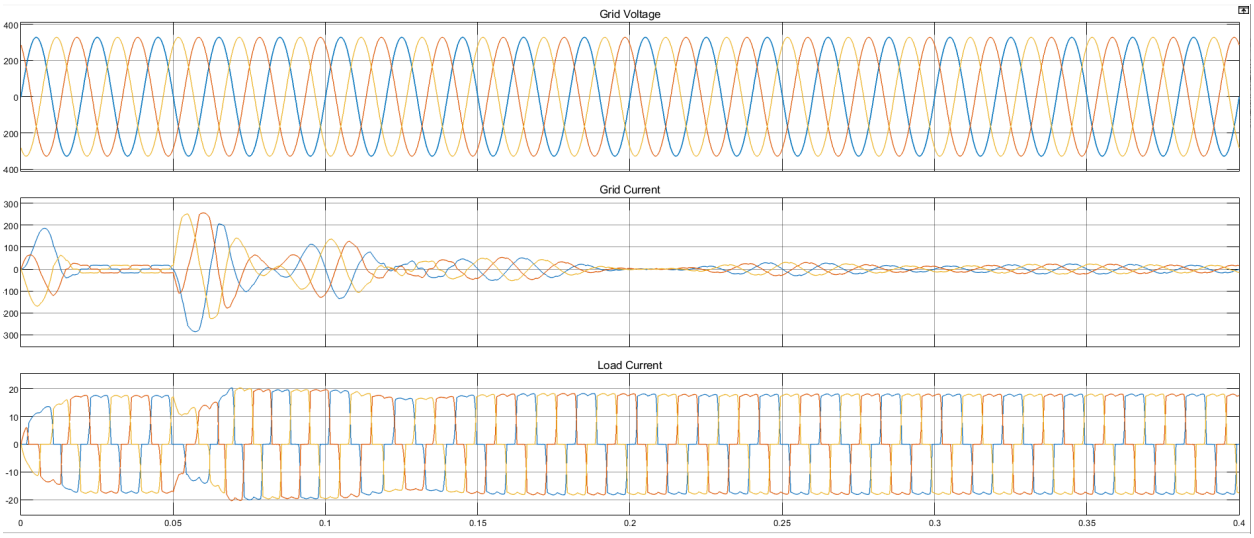

Figure 4-4: Grid and load voltage and currents.

Furthermore, the grid current is related to the grid voltage. The load current shows sudden changes as it responds to varying load demand. Disturbances begin from the PMSG, causing distortions in the current. The system does not always adapt smoothly, leading to instabilities and overshoots in the load. These fluctuations in the load current negatively impact the overall efficiency of the system. Overall, PI controller results is good, but the efficiency of

the system requires improvement in terms of stability.

4.3 Reinforcement Learning based Microgrid

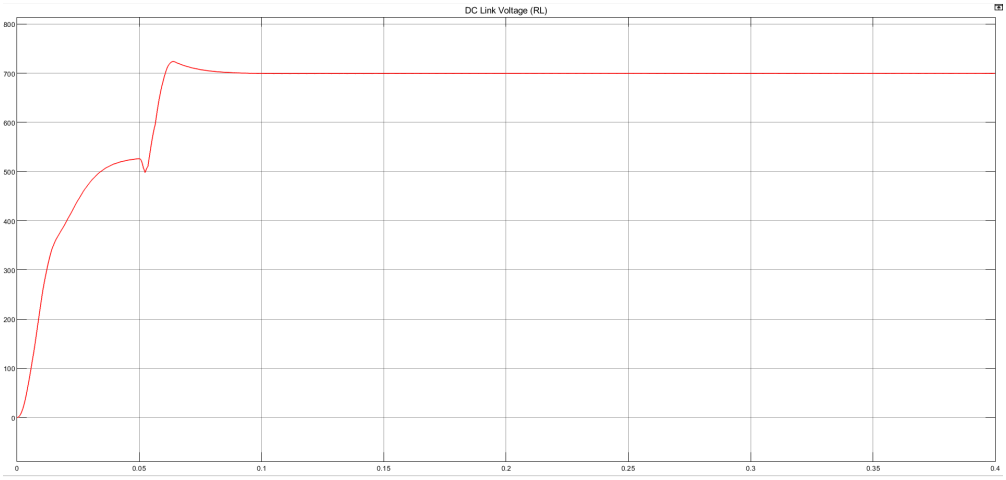

Figure 4-5: DC link voltage (RL).

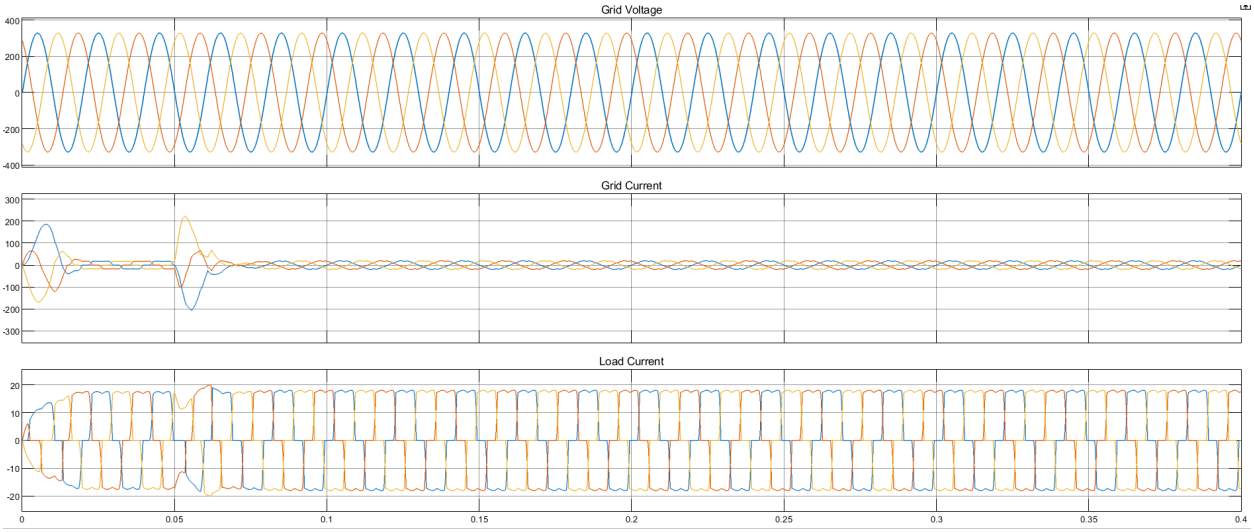

Figure 4-6: Grid and load voltage and currents (RL).

In the case of reinforcement learning (RL), the results are quite different. As shown in Figure 4-5, the DC link voltage is more efficient, and the system is more stable. The desired output voltage is still set to 700 V. The minimum point of the DC link voltage occurs at 500 V around 0.06 seconds, and the maximum point reaches 720 V at 0.07 seconds. After this, the

voltage stabilizes and becomes steady at 700 V, with minimal fluctuations. The RL-based controller adapts well to varying conditions, resulting in reduced fluctuations. Compared to the PI controller's DC link voltage, the RL controller is more optimized and efficient.

As shown in Figure 4-6, the grid voltage remains constant at 328 V. Unlike the PI controller, the RL-based microgrid achieves stability starting from 0.07 seconds, which is significantly better than the PI controller. However, before reaching this stable point, there is an overshoot, with the current peaking at 215 A. Since the inverter has a significant impact on the grid, the RL-based controller offers greater flexibility and adaptability. The load current in the RL system is similar to that of the PI controller, but with smaller deviations, resulting in a smoother and more stable performance.

Table 4.1: Transient Performance Parameters for DC Link Voltage

Parameter	**PI Controller**	**Reinforcement Learning**
Final Value	700 V	700 V
Rise Time	0.02–0.03 s	0.02 s
Overshoot	≈ 500 V $\sim 71\%$ overshoot	≈ 25 V $\sim$ 3–4% overshoot)
Undershoot	≈ 300 V $\sim 40\%$ undershoot	No significant undershoot
Settling Time	≈ 0.12 s	$\approx$ 0.04–0.05 s
Transient Time	≈ 0.12 s	$\approx$ 0.04–0.05 s

Table 4.2: Comparison of Transient Performance Parameters

Parameter	**Reinforcement Learning**	**PI Controller**
Current THD	$\sim$ 3–4%	$\sim 5\%$
Accuracy (overall tracking)	$\sim$ 95–97%	$\sim$ 94–95%
Voltage Ripples (line distortion)	$< 5\%$	$\sim 5\%$
Grid Current (peak amplitude)	Up to ± 60 A	Peaks up to ± 100 A
Load Current (peak amplitude)	Approximately $\pm$20–24 A	Approximately $\pm$20–24 A
Observed Problems (Grid Current)	Moderate overshoot	Larger overshoot
Observed Problems (Load Current)	Small harmonic distortion	Brief distortion

The main results for the DC link voltage are summarized in Table 4.1, which includes the final voltage value, rise time, settling time, and transient response characteristics. Addi-

tionally, the table details both the overshoot and undershoot values. A comparison of these parameters clearly demonstrates that the reinforcement learning approach outperforms the PI controller. In particular, the overshoot and undershoot metrics highlight this improvement: while the PI controller exhibits fluctuations of up to 500 V, reinforcement learning keeps these deviations to around 25 V.

Moreover, Table 4.2 presents the load and grid results, including total harmonic distortion (THD), accuracy, voltage and current ripple magnitudes, and overshoot data. In the load current, the reinforcement learning (RL) approach exhibits relatively low harmonic distortion, whereas the PI controller experiences more pronounced distortion. For the grid current, the RL-based system shows moderate overshoot, while the PI controller undergoes a higher degree of distortion. Additionally, the RL method demonstrates favorable accuracy, reflected by a current THD of 3%, in contrast to the PI controller's THD, which is higher by approximately 2%

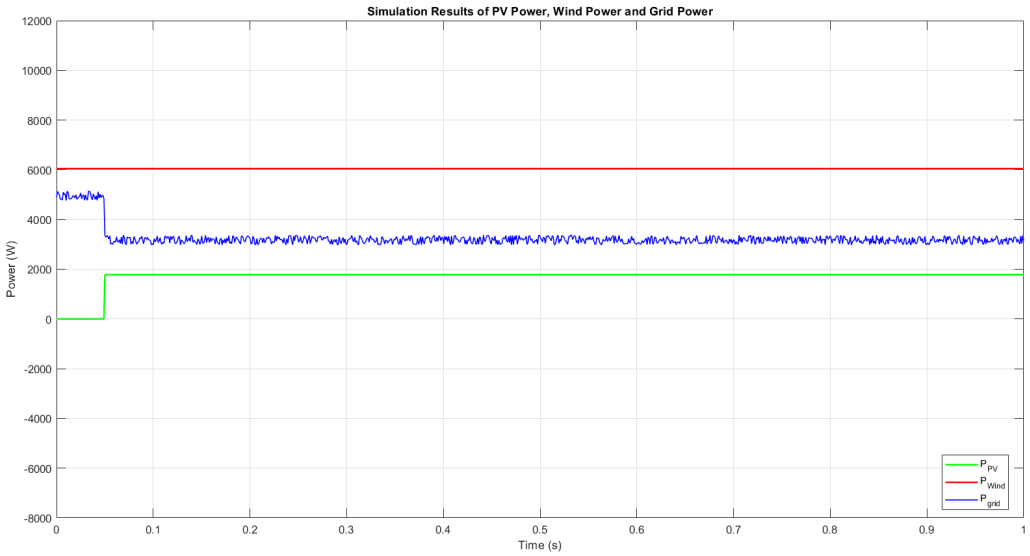

Figure 4-7: Simulation of PV power, grid power and wind power.

Figure 4-7 presents the simulated power contributions from the photovoltaic (PV) array, wind turbine, and utility grid. The sequence of events is as follows:

- **0–0.05 s:** Wind turbine works from starting point, supplying 6.045 kW. Remained

power is supplied by the grid, $11 - 6.045 = 4.955\,\text{kW}$.

- **0.05 s:** PV panel is enabled at 1.780 kW, decreasing grid supply to $11-6.045-1.780 = 3.175\,\text{kW}$.
- **0.05–1 s :** PV remains at 1.780 kW and grid supply is always 3.175 kW.

In steady-state operation ($t > 0.05\,\text{s}$), the power balance is

$$P_{\text{PV}} \approx 1.780\ \text{kW}, \qquad P_{\text{wind}} \approx 6.045\ \text{kW}, \qquad P_{\text{grid}} \approx 3.175\ \text{kW}, \qquad P_{\text{load}} = 11\ \text{kW}.$$

Hence, the coordinated control strategy enables the renewables to supply roughly **86**% of the load, with the grid covering only the remaining $\sim 1.4\,\text{kW}$ and compensating for short-term irradiance dips.

Finally, replacing the PI controller with a reinforcement-learning agent changes the microgrid into a data-driven, self-adaptive platform capable of optimizing energy flows, enhancing power quality, and preserving stability under various and unpredictable operating situations.

Chapter 5

Conclusion

In conclusion, this study successfully demonstrates the efficacy of integrating reinforcement learning into the adaptive control framework for AC/DC microgrid energy management. The proposed RL-based controller not only outperforms conventional PI control by dynamically tuning control parameters in real time, but it also effectively mitigates issues such as transient overshoots and voltage fluctuations. The simulation model created in MATLAB Simulink does a great job of representing a real-life microgrid, incorporating renewable energy sources and interactions with the grid. This helps show how artificial intelligence can handle the complex, unpredictable aspects of energy systems, like sudden changes in energy demand or supply. By using a reinforcement learning approach with a Twin-Delayed Deep Deterministic Policy Gradient (TD3) agent, a solid alternative to traditional control methods could be offered. The TD3 agent turned out to be much more adaptable and efficient, which helped improve the stability and overall performance of the microgrid. This is especially useful for microgrids, which often have to deal with the unpredictability of renewable energy sources and constantly shifting conditions. Although the simulation results are promising, there's still more work to be done to figure out how this can work in the real world. Some issues need to be tackled such as scaling up the RL-based control method and refining it for actual use. Future research could include real-world tests, adding other AI techniques, and making sure the algorithms run efficiently. All in all, this research sets the stage for better energy management solutions, pushing us closer to smarter, more resilient, and sustainable microgrid systems.

Bibliography

[1] P. A. Owusu and S. Asumadu-Sarkodie, "A review of renewable energy sources, sustainability issues and climate change mitigation," *Cogent Engineering*, vol. 3, no. 1, p. 1167990, 2016.

[2] M. R. Rasekh, P. K. Jamwal, V. Gali, and M. J. Ahmadi, "Design and analysis of high gain dc-dc boost converter for grid connected solar photovoltaic system," in *2023 International Conference on Power Electronics and Energy (ICPEE)*, 2023, pp. 1–6.

[3] T. Dragičević, X. Lu, J. C. Vasquez, and J. M. Guerrero, "Dc microgrids—part ii: A review of power architectures, applications, and standardization issues," *IEEE Transactions on Power Electronics*, vol. 31, no. 5, pp. 3528–3549, 2016.

[4] T.-G. Kim, H. Lee, C.-G. An, J. Yi, and C.-Y. Won, "Hybrid ac/dc microgrid energy management strategy based on two-step ann," *Energies*, vol. 16, no. 4, 2023. [Online]. Available: https://www.mdpi.com/1996-1073/16/4/1787

[5] Maneesh, "Frequency control of a microgrid by using pi controller," in *2015 International Conference on Energy, Power and Environment: Towards Sustainable Growth (ICEPE)*, 2015, pp. 1–5.

[6] A. Parisio and L. Glielmo, "Energy efficient microgrid management using model predictive control," in *2011 50th IEEE Conference on Decision and Control and European Control Conference*, 2011, pp. 5449–5454.

[7] D. M. Britz and R. R. Miller, "Mesh free space optical systems: A method to improve broadband neighborhood area network backhaul," in *Proceedings of the IEEE Workshop on Local Metropolitan Area Networks*, 2007, pp. 37–42.

[8] H. E. Brown and S. Suryanarayanan, "A survey seeking a definition of a smart distribution system," in *Proceedings of the North American Power Symposium (NAPS)*, 2016, pp. 1–7.

[9] N. Bressan, L. Bazzaco, N. Bui, P. Casari, L. Vangelista, and M. Zorzi, "The deployment of a smart monitoring system using wireless sensors and actuators networks," in *Proceedings of IEEE SmartGridComm*, 2015, pp. 49–54.

[10] P. Piagi and R. H. Lasseter, "Autonomous Control of Microgrids," in *Proceedings of the IEEE PES Meeting*, Montreal, Canada, Jun. 2006.

[11] F. Katiraei, M. R. Iravani, and P. W. Lehn, "Micro-Grid Autonomous Operation During and Subsequent to Islanding Process," *IEEE Transactions on Power Delivery*, vol. 20, no. 1, pp. —.

[12] Y. Sabri, N. El Kamoun, and F. Lakrami, "A survey: Centralized, decentralized, and distributed control scheme in smart grid systems," in *2019 7th Mediterranean Congress of Telecommunications (CMT)*, 2019, pp. 1–11.

[13] D. M. Britz and R. R. Miller, "Mesh free space optical systems: A method to improve broadband neighborhood area network backhaul," in *IEEE Workshop on Local Metropolitan Area Networks*, 2007, pp. 37–42.

[14] M. Brucoli and T. C. Green, "Fault behaviour in islanded microgrids," in *19th International Conference on Electricity Distribution*, 2017, pp. 1–4.

[15] W. Kempton and J. Tomic, "Vehicle-to-grid power implementation: From stabilizing the grid to supporting large-scale renewable energy," *Journal of Power Sources*, vol. 144, no. 1, pp. 280–294, 2005.

[16] W. Kempton, V. Udo, K. Huber, K. Komara, S. Letendre, S. Baker *et al.*, "A test of vehicle-to-grid (V2G) for energy storage and frequency regulation in the PJM system," Mid-Atlantic Grid Interactive Cars Consortium, Tech. Rep., 2016.

[17] K. Kurohane, T. Senjyu, A. Yona, N. Urasaki, T. Goya, and T. Funabashi, "A hybrid smart AC/DC power system," *IEEE Transactions on Smart Grid*, vol. 1, no. 2, pp. 199–204, 2017.

[18] M. J. Ghadi, A. Rajabi, S. Ghavidel, A. Azizivahed, L. Li, and J. Zhang, "From active distribution systems to decentralized microgrids: A review on regulations and planning approaches based on operational factors," *Applied Energy*, vol. 253, p. 113543, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261919312176

[19] J. E. Rubio, C. Alcaraz, and J. Lopez, "Recommender system for privacy-preserving solutions in smart metering," *Pervasive Mob. Comput.*, vol. 41, no. C, pp. 205–218, Oct. 2017, [Online]. Available: https://doi.org/10.1016/j.pmcj.2017.03.008.

[20] A. H. and, "Maximum power point tracking (mppt) scheme for solar photovoltaic system," *Energy Technology & Policy*, vol. 1, no. 1, pp. 115–122, 2014.

[21] M. K. Johari, M. Jalil, and M. F. M. Shariff, "Comparison of horizontal axis wind turbine (hawt) and vertical axis wind turbine (vawt)," *International Journal of Engineering and Technology*, vol. 7, no. 4.13, pp. 74–80, 2018.

[22] N. A. Orlando, M. Liserre, R. A. Mastromauro, and A. Dell'Aquila, "A survey of control issues in pmsg-based small wind-turbine systems," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 3, pp. 1211–1221, 2013.

[23] M. Eskandari, A. Rajabi, A. V. Savkin, M. H. Moradi, and Z. Y. Dong, "Battery energy storage systems (besss) and the economy-dynamics of microgrids: Review, analysis, and classification for standardization of besss applications," *Journal of Energy Storage*, vol. 55, p. 105627, 2022.

[24] S. Rajamand, "Effective control of voltage and frequency in microgrid using adjustment of pid coefficients by metaheuristic algorithms," *IETE Journal of Research*, vol. 68, pp. 3526 – 3539, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:234108095

[25] I. Alhamrouni, M. Salem, M. Hairullah, N. Omar, A. Jusoh, and T. Sutikno, "Modelling and design of pid controller for voltage control of ac hybrid micro-grid," *International Journal of Power Electronics and Drive Systems*, 03 2019.

[26] K. S. Joshal and N. Gupta, "Microgrids with model predictive control: A critical review," *Energies*, vol. 16, no. 13, 2023. [Online]. Available: https://www.mdpi.com/1996-1073/16/13/4851

[27] J. Hu, Y. Shan, J. M. Guerrero, A. Ioinovici, K. W. Chan, and J. Rodriguez, "Model predictive control of microgrids – an overview," *Renewable and Sustainable Energy Reviews*, vol. 136, p. 110422, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032120307097

[28] M. Cucuzzella, G. P. Incremona, and A. Ferrara, "Design of robust higher order sliding mode control for microgrids," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 5, no. 3, pp. 393–401, 2015.

[29] ——, "Design of robust higher order sliding mode control for microgrids," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 5, no. 3, pp. 393–401, 2015.

[30] T. K. Roy, M. A. Mahmud, A. M. T. Oo, M. E. Haque, K. M. Muttaqi, and N. Mendis, "Nonlinear adaptive backstepping controller design for islanded dc microgrids," *IEEE Transactions on Industry Applications*, vol. 54, no. 3, pp. 2857–2873, 2018.

[31] M. Dehghani, T. Niknam, M. Ghiasi, H. R. Baghaee, F. Blaabjerg, T. Dragicevǐć, and M. Rashidi, "Adaptive backstepping control for master-slave ac microgrid in smart island," *Energy*, vol. 246, p. 123282, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360544222001852

[32] K.-M. Kang, B.-Y. Choi, H. Lee, C.-G. An, T.-G. Kim, Y.-S. Lee, M. Kim, J. Yi, and C.-Y. Won, "Energy management method of hybrid ac/dc microgrid using artificial neural network," *Electronics*, vol. 10, no. 16, 2021. [Online]. Available: https://www.mdpi.com/2079-9292/10/16/1939

[33] W. Dong, Q. Yang, X. Fang, and W. Ruan, "Adaptive optimal fuzzy logic based energy management in multi-energy microgrid considering operational uncertainties," *Applied Soft Computing*, vol. 98, p. 106882, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568494620308206

[34] Y. Cai, G. Huang, Z. Yang, and Q. Tan, "Identification of optimal strategies for energy management systems planning under multiple uncertainties," *Applied Energy*, vol. 86, no. 4, pp. 480–495, 2009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261908002328

[35] Y. Du and F. Li, "Intelligent multi-microgrid energy management based on deep neural network and model-free reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1066–1076, 2020.

[36] E. Kuznetsova, Y.-F. Li, C. Ruiz, E. Zio, G. Ault, and K. Bell, "Reinforcement learning for microgrid energy management," *Energy*, vol. 59, pp. 133–146, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360544213004817