GAZİOSMANPAŞA BİLİMSEL ARAŞTIRMA DERGİSİ (GBAD)
Gaziosmanpasa Journal of Scientific Research
ISSN: 2146-8168
http://dergipark.gov.tr/gbad
Araştırma Makalesi (Research Article)

Cilt/Volume : 13
Sayı/Number: 3
Yıl/Year: 2024
Sayfa/Pages: 30-46

***Alınış tarihi (Received)*:** 11.09.2024
***Kabul tarihi (Accepted)*:** 20.11.2024

Controller Design for DC-DC Boost Converter using PI, State Feedback and Q Learning

Murat Erhan ÇiMEN[1,*]

[1]*Department of Electrical and Electronic Engineering, Sakarya University of Applied Sciences, Serdivan, Sakarya, Türkiye.*
**Corresponding author: muratcimen@subu.edu.tr*

ABSTRACT: In recent years, the global decline in fossil fuel reserves and the alarming rise in greenhouse gas emissions have significantly heightened the need for renewable energy sources. This urgent shift towards sustainability has made the development and optimization of efficient energy systems a top priority for countries and communities worldwide. This study focuses on the modeling and control of a DC-DC boost converter, a critical component widely utilized in renewable energy applications such as solar panels, battery systems, and fuel cells. The research explores three control strategies: conventional Proportional-Integral (PI) controller, State Feedback Controller with integral action, and Q Learning-based controller, which employs reinforcement learning principles. Comparative experiments were conducted in the Matlab/Simulink environment to evaluate the performance of each controller. Results have demonstrated that the Q Learning controller outperformed the traditional methods in terms of performance metrics, including Integral Squared Error (ISE), Integral Absolute Error (IAE), and settling time, showcasing its potential for enhancing the efficiency and stability of renewable energy systems

Keywords *– Boost Converter, Q learning, PI Control, State feedback*

1. Introduction

In recent decades, the escalating severity of climate change and environmental pollution has highlighted the urgent need for enhanced attention and decisive action. The progressive depletion of fossil fuel reserves has further intensified the global emphasis on the imperative of harnessing renewable and clean energy sources. This depletion, coupled with its significant contribution to climate change, has accelerated efforts toward the development and deployment of alternative energy technologies (Abdalla et al., 2022). Consequently, there has been a marked increase in interest and research focused on renewable energy sources as a pathway to a sustainable future (Abdalla and Önbilgin, 2024; Garip et al., 2021). Renewable energy systems, such as those based on solar, wind, wave, and tidal sources, often require efficient energy storage solutions, given the intermittent nature of energy production. To address this need, power electronic circuits, particularly DC-DC converters, play a critical role. These converters are fundamental components in power electronics, enabling the conversion of electrical voltage levels through their switching capabilities (Çimen et al., 2021). Due to their functions, DC-DC converters are widely utilized across various applications, including power supplies, portable industrial devices, and electric vehicles (Farajdadian et al., 2024). A Boost converter, also known as a step-up converter, is a type of switched-mode DC-DC converter designed to produce a constant output voltage that exceeds the input voltage. To accurately model the behavior of DC-DC converters, averaging techniques, such as the small-signal model, are commonly employed (Alkrunz and Yazıcı, 2016).

Proportional-Integral (PI) and Proportional-Integral-Derivative (PID) controllers are traditional linear control methods commonly used across various applications due to their simplicity, reliability, and strong performance (Borase et al., 2021). PID controllers have historically been tuned through a combination of experience, trial-and-error, and methods such as Cohen–Coon and Ziegler–Nichols (Ibrahim et al., 2016). For example, Güngör and Yüksek (2020) modeled Boost and small DC-DC converters for photovoltaic (PV) panels, applying PID control for maximum power tracking. (Alkrunz and Yazıcı, 2016) used PI control with discrete-time, state feedback, and Linear Quadratic Regulator (LQR) methods to model a Boost converter. (Sezen and Keskin, 2021) compared PI and fuzzy logic controllers for DC-DC Boost converter control, while (Bououden et al., 2014) employed Model Predictive Control (MPC) for Boost converter regulation and compared it with fuzzy logic control. (Uçmaz and Yakut, 2024) controlled a Boost converter in Proton Exchange Membrane Fuel Cells (PEMFC) using PI and fractional PI controllers, optimizing parameters with Particle Swarm Optimization (PSO). In addition to linear controllers, nonlinear control strategies have been explored for DC-DC converters, such as Sliding Mode Control (SMC) (Güldemir, 2005) and backstepping control (El Fadil and Giri, 2007). Recently, (Palpandian et al., 2024) applied machine learning techniques to control DC-DC converters. Among these, Reinforcement Learning (RL) has gained significant attention due to its ability to learn and adapt to its environment, making it suitable for controlling nonlinear systems and industrial processes (Harmon and Harmon, 1996). RL operates by having an agent interact with its environment, learning to select actions that maximize cumulative rewards, which is particularly useful in dynamic environments (Angiuli et al., 2022). RL is especially promising for controlling DC-DC converters in cases of model uncertainties or nonlinearities, where traditional control methods struggle. However, RL methods like Q Learning face scalability challenges, particularly with the size of the Q-table (You et al., 2023).

Advancements in computational power due to evolving technology have made reinforcement learning particularly applicable in power systems with low time constants, that is, systems operating at high speeds. In their study (Meng et al., 2024), they used reinforcement learning in the control of a microgrid to which wind, photovoltaic panels, battery systems, electric vehicles and residences are connected. They state that it provides adaptation to changing situations especially by performing online pre-programming. In their study (Rajamallaiah et al., 2024), they controlled the DC-DC converters used in the microgrid with the reinforcement method. In their study (Sun and Lue, 2024), they carried out a study on the distribution of distributed energy resources (DERs) with reinforcement learning. They evaluated their applications on the IEEE 33-node system. In contrast to this study, (Lui et al., 2024) implemented Volt-Var control using reinforcement learning in distributed networks. (Kang et al., 2024) achieved successful results by using reinforcement learning in the storage of battery systems fed by photovoltaic panels. Working on a similar topic, (Li et al., 2024) modified the reinforcement learning method and carried out studies on a different case study and achieved successful results. In terms of power converters (Alfred et al., 2024), a Linear Quadratic Regulator (LQR) controller for a buck converter, using RL to determine its parameters, achieved better results than traditional PID controllers. (Nishanthi and Kanakaraj, 2024) demonstrated the effectiveness of a Deep Q-Network (DQN) controller for a buck-boost converter, showing high accuracy and improved transient response. (Muktiadji et al., 2024) compared Genetic Algorithms (GA), Particle Swarm Optimization (PSO), and Twin Delayed Deep Deterministic Policy Gradient (TD3) for determining PI controller parameters in a Boost converter, with TD3 yielding the best results. Similarly, (Ahmed and Ahmad, 2024) explored wireless energy

transfer in electric vehicles, using fuzzy logic and DQN to optimize power transfer, offering reduced communication overhead and improved scalability.

In renewable energy, especially PV systems, reinforcement learning methods like Deep Q-Network (DQN), Deep Deterministic Policy Gradient (DDPG), and Temporal Difference effectively enhance power transfer and address efficiency limitations (Panggabean et al., 2023). (Xu et al., 2024) applied a Sliding Mode Observer to control DC voltage from PEMFC, outperforming traditional PI controllers. (Gheisarnejad and Khooban, 2023) integrated a full-bridge DC-DC converter with classical PI, Sliding Mode, and Quantum-based DQN methods, finding DQN to be most effective in both simulations and real-time testing on the OPAL-RT platform.

The main contribution of this study are as follows;

- Adaptation of the machine learning method Q Learning Algorithm to the control of the power electronic circuit -DC-DC boost converter- which is widely used in renewable energy sources,

- Obtaining the small signal model by linearizing the nonlinear power electronic circuit DC-DC boost converter and then designing PI and State Feedback Controller,

- Controlling the DC-DC boost converter in the Matlab Simulink environment after the designs are realized and comparing the obtained results according to performance criteria such as Integral Squared Error (ISE) and Integral Absolute Error (IAE).

In this study, PI and State Feedback Control designs with integral action were implemented for a boost converter. Subsequently, the Q Learning Algorithm, a method within reinforcement learning, was directly applied to the boost converter's control, effectively governing the system. Analysis of the results indicates that Q Learning Algorithm outperformed traditional controllers in terms of control performance. When it comes to the layout of the article, the first section consists of introduction that provides a review of the literature relevant to the study. The second section presents material and methods that includes the modeling of the boost converter and an introduction to the Q Learning Algorithm. Also, this section also includes the linearization of the boost converter, along with the design of PI and State Feedback Controller. The third section, results and discussion, provides and analyzes the simulation results for the PI, state feedback, and Q Learning controllers, as well as the overall performance of the boost converter. Finally, the fourth section offers a summary of the study, assesses the effectiveness of the proposed method, and discusses directions for future research.

2. Material and Methods

The materials and methods used in the study are divided into titles and explained.

2.1 Boost Converter

A boost converter is an electrical circuit designed to produce an output voltage higher than its input voltage. The voltage level adjustments are achieved through a control signal, which utilizes a fast-switching element, such as a transistor, MOSFET or IGBT, to regulate the charge and discharge cycles between the inductor and capacitor within the circuit. Boost converter is depicted at Figure 1 (Güngör and Yüksek, 2020).

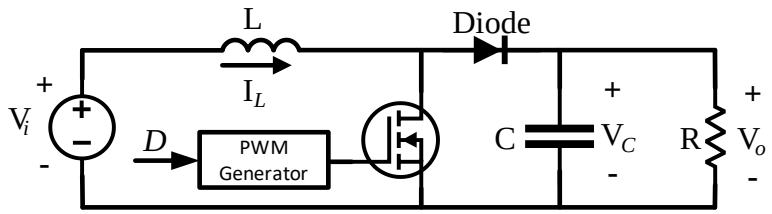

Figure 1. Circuit of DC-DC boost converter

Initially, state-space models are derived for both the "on" and "off" modes of operation. These two modes are then combined using the averaging method to produce an average state-space model. When the fast-switching element is in the "on" state, the current from the voltage source flows through the inductor. The switching element determines the operation between on mode and off mode through the pulse width modulation (PWM) method. The generated signal has a specific duty cycle (*D*) at the switching frequency. Based on this duty cycle, on mode and off mode models are activated (Saha et al., 2023). Consequently, an average model is derived by combining these two modes. The average system matrix (*A*) and the average control matrix (*B*) are obtained in Equation 1. After rearranging the system states, Equation 1 is derived (Güngör and Yüksek, 2020; Saha et al., 2023).

$$\dot{x} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \underbrace{\begin{bmatrix} I_L \\ \dot{V}_C \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 0 & -\frac{1-D}{L} \\ \frac{1-D}{C} & -\frac{1}{RC} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} I_L \\ V_C \end{bmatrix}}_{x} + \underbrace{\begin{bmatrix} \frac{1}{L} \\ 0 \end{bmatrix}}_{B} \underbrace{V_{in}}_{u} \qquad 1$$

$$\dot{x} = Ax + Bu$$

The parameters of the boost converter used in this study are given in Table 1.

Table 1. DC-DC Boost Converter Parameters and Values

Component	Definition	Value
V_{in}	Supply Voltage	12 V
V_{out}	Output Voltage	24 V
L	Inductor	46mH
R	Resistor	15 Ω
C	Capacitor	1.360 mF
f	Switching Frequency	50 kHz

2.2. Design of Controller

The modeled system is required to operate at a certain performance. For this purpose, as mentioned before, many controllers have been proposed in the literature. In this study, the control of the system is realized with PI, state feedback and Q learning.

2.2.1 Q learning

RL is a technique used to solve sequential decision-making problems across various environment, including natural and social sciences, as well as engineering, by enabling an agent to interact with its environment and learn an optimal policy through trial and error (Smart and Kaelbling, 2000; Wang, H., Emmerich, M., and Plaat, 2018). In RL methods, learning is typically conducted via a Q-table (Wang, H., Emmerich, M., and Plaat, 2018). Various approaches exist for updating this table, such as dynamic programming, Monte Carlo methods, Q Learning, and State-Action-Reward-State-Action (SARSA) (Akyurek

and Bucak, 2012). As illustrated in Figure 2, the system consists of a boost converter environment and an agent. The agent takes an action (α_t) within the environment based on its current state (s_t) and action. This action results in a state transition (s_{t+1}) and generates a reward (r_{t+1}) . Through continuous interaction with the environment, the agent learns by utilizing information such as (s_t,α_t,r_t,s_{t+1}). In this study, the Q Learning Algorithm will be applied to update the Q-table for learning.

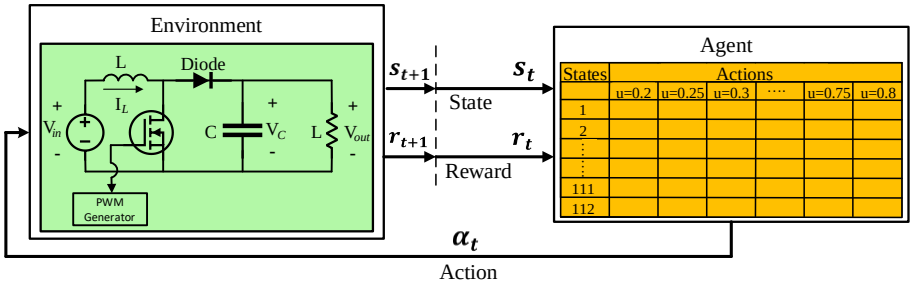

Figure 2. Agent and environent for boost converter control

Bellman equation used in updating the Q table used in Figure 2 is the formula expressed by Equation 2. In Equation 2, state at time *t* obtained from s_t environment, the action that α_t agent *lr* learning factor, γ is the reduction factor. The expression $\max_{\alpha}(Q_t(s_{t+1}, \alpha_t))$ provides the highest value for any action in s_{t+1} state. This approach, called on-policy, constantly updates the Q table in interaction with the environment. The psoudecode of Q Learning is given in Algorithm 1.

$$Q_{t+1}(s_t, a_t) = Q_t(s_t, a_t) + lr\left(r_t + \gamma\left(\max_{\alpha}\left(Q_t(s_{t+1}, \alpha)\right)\right) - Q_t(s_t, \alpha_t)\right) \qquad 2$$

As with every method, the Q learning method has its advantages and disadvantages. The advantages of Q learning method;

• Its simplicity is that it can produce the control signal only according to the states, independent of the model,

• It can be adapted to continuous-time systems,

• It can adapt to the changing states of the system,

• It has the potential to produce control signals that will maximize the reward function by continuously learning,

• Q Learning controller demonstrates superior performance in the dynamic and uncertain environments characteristic of renewable energy systems.

The disadvantages of Q learning method;

• It may not perform well when there are undiscovered states while controlling the DC-DC boost converter,

• Learning may take a long time when the number of states in the Q table increases,

• It may not discover new states in case of excessive learning,

• It may take time to find the best path if there is more than one result that maximizes the reward function.

Algorithm 1. Q learning Algorithm

1: Input: Learning Rate (*lr*) Discount factor (γ), $Q(s,a)$ table

2: Output: Updated Q(s,a) Table

3: Determine initial Q table

4: For $i = 1: iter\_max$

5: Initialize state s_t

6: $done == False$

7: $t = 1$

8: While $done == False$

9: Choose α_t with ε greedy probability

10: Apply action α_t to environment and observe s_{t+1}, reward r_{t+1} and $done$

11: Update table $Q_{t+1}(s_t, \alpha_t)$

12: $t = t + 1$

In addition, states are presented as in Equation 3. Also, reward function obtained from the environment is specified as in Equation 4. When the denominators of the fractional expressions in the reward function are 0, the values will take the value of infinity. For this reason, the parameters μ_1 and μ_2 in the fractional expressions are chosen to be non-zero and are determined as μ_1=0.025 and μ_1=0.2. Nevertheless, it is useful to know that these states and reward functions could be changed as to designer.

$$s_t = (V_{ref}(t), V_o(t)) \quad (3)$$

$$reward = -\frac{1}{V_{ref}(t) - V_o(t) + \mu_1} + \frac{1}{V_o(t+1) - V_o(t) + \mu_2} - (V_{ref}(t) - V_o(t))^2 \quad (4)$$

The Q learning parameters used to apply the Q algorithm to the boost converter are given in Table 2. The done function for the boost converter environment takes the value true when the simulation time is 1.5 seconds. Otherwise, it takes the value false. The learning parameters of the algorithm, iteration number and discounted factor, were selected similarly in the studies in the literature. However, when it comes to the learning rate, epsilon greedy parameters, the sampling time of the boost converter is very small, so it was selected specific to the problem compared to the parameters in the literature.

Table 2. Simulation parameters of Q learning for boost converter

Paramater	Value	Paramater	
Iteration number ($iter\_max$)	1,000	Epsilon Greedy threshold value	$1x10^{-3}$
Learning Rate (lr)	0.01	Epsilon Greedy decreasing value	$2x10^{-4}$
Discount factor (γ),	0.95	Boost Converter Sampling Time	$1x10^{-6}$ second
Reference Values	20V, 24V, 30V	Q learning controller Sampling Time	$1x10^{-4}$ second
Epsilon Greedy initial value	$5x10^{-1}$	Simulation Duration	1.5 second

Due to the controller running on the Q table, the Q table of the system was created according to the actions and states. $x(t)$ state is composed of $V_{ref}(t)$ and $V_o(t)$. The $V_{ref}(t)$ value only takes the discrete values 20V, 24V and 30V. But $V_o(t)$ selected between 19 and 37 with a certain (0.5) step length. By using these discrete values, Q table is formed. A small part of the created Q table is presented in Table 3.

Table 3. Q Table for Boost Converter

State (s_t)		Action (α_t)						
$V_{ref}(t), V_o(t)$		d						
Real	Discrete	0.00	0.05	0.1	0.15	…	0.7	0.02
(20, 19)	1	3.951x102	2.527 x102	3.897x102	2.395x102	…	-3.655	4.454x102
(20, 19.5)	2	5.535x102	5.532x102	5.594 x102	5.585x102	…	6.656x10-2	5.285x102
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
(20, 36.5)	36	-3.681x1014	-3.675x102	-3.638x102	-3.645 x102	…	-3.603x1014	-3.602x1014
(20, 37)	37	-4.447x1014	-4.400x1014	-4.421x1014	-4.411x1014	…	-4.451x1014	-4.449x1014
(24, 19)	38	-1.635x1014	-1.085x1014	-1.284x1014	-1.075x1014	…	-1.075x1014	-1.083x1014
(24, 19.5)	39	-8.887x10-1	-9.126x10-3	-1.030x1014	3.086x1014	…	1.063x1014	5.019x10-2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
(24, 36.5)	73	-1.625x102	-1.628x102	-1.641x102	-1.625x102	…	-1.625x102	-1.627x102
(24, 37)	74	-2.107x102	-2.107x102	-2.089x102	-2.101x102	…	-2.106x102	-2.105x102
(30, 19)	75	-5.211x102	-3.700x102	-4.259x102	-3.703x102	…	-3.713x102	-3.716x102
(30, 19.5)	76	-3.696x102	-3.989x102	-3.7212	-4.001x102	…	-3.636x102	-3.689x102
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
(30, 36.5)	110	-1.637x102	-1.675x102	-1.670x102	-1.664x102	…	-1.624x102	-1.610x102
(30, 37)	111	-3.769x102	-3.775x102	-3.790x102	-3.543x102	…	-3.766x102	-3.771x102

2.2.2 Feedback Controller Design

A system is characterized by its inputs and outputs, and controllers modify its performance and response dynamics. For effective controller design, the system model should ideally be specified. This section addresses the linearization of the nonlinear Boost Converter, subsequently design PI and State Feedback Controller with integral action.

2.2.2.1 Linearization of Boost Converter

The boost converter model is presented in Equation 1. Due to the multiplication of the state variables (x_1, x_2) and the control signal (u), linearization is necessary. This is achieved by expressing the state variables as functions, as shown in Equation 5, and differentiating them

at the equilibrium point. Assuming negligible changes over a small time interval, the linearized dynamic model is represented in Equation 6. By calculating the derivatives of the state variables and the control signal at equilibrium, the system and control matrices are obtained. The final linearized model is presented in Equation 7.

$$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} \dfrac{-(1-D)}{L} x_2 + \dfrac{1}{L} u \\ \dfrac{(1-D)}{C} x_1 - \dfrac{1}{RC} x_2 \end{bmatrix} \tag{5}$$

$$\begin{bmatrix} \Delta\dot{x}_1 \\ \Delta\dot{x}_2 \end{bmatrix} = \begin{bmatrix} \left.\dfrac{df_1(x_1,x_2,u)}{dx_1}\right|_{\substack{x_1^*=x_1 \\ x_2^*=x_2 \\ u^*=u}} & \left.\dfrac{df_1(x_1,x_2,u)}{dx_2}\right|_{\substack{x_1^*=x_1 \\ x_2^*=x_2 \\ d^*=u}} \\ \left.\dfrac{df_2(x_1,x_2,u)}{dx_1}\right|_{\substack{x_1^*=x_1 \\ x_2^*=x_2 \\ d^*=d}} & \left.\dfrac{df_2(x_1,x_2,u)}{dx_2}\right|_{\substack{x_1^*=x_1 \\ x_2^*=x_2 \\ d^*=d}} \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix} + \begin{bmatrix} \left.\dfrac{df_1(x_1,x_2,u)}{du}\right|_{\substack{x_1^*=x_1 \\ x_2^*=x_2 \\ u^*=u}} \\ \left.\dfrac{df_2(x_1,x_2,u)}{du}\right|_{\substack{x_1^*=x_1 \\ x_2^*=x_2 \\ u^*=u}} \end{bmatrix} \Delta u \tag{6}$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = C \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + C \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix}$$

$$\begin{bmatrix} \Delta\dot{x}_1 \\ \Delta\dot{x}_2 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & -\dfrac{(1-D)}{L} \\ \dfrac{(1-D)}{C} & -\dfrac{1}{RC} \end{bmatrix}}_{A} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix} + \underbrace{\begin{bmatrix} \dfrac{1}{L} \\ 0 \end{bmatrix}}_{B} \Delta u \tag{7}$$

When sytem is executed for the system parameters and equilibrium values with D=0.5, the equilibrium values of the system are obtained as x_1*=1.804, x_2*=22.890. These values correspond to the circuit configuration illustrated in Figure 3. The time-dependent graphs of the current and voltage values are shown in Figure 4 and Figure 5, repectivelty.

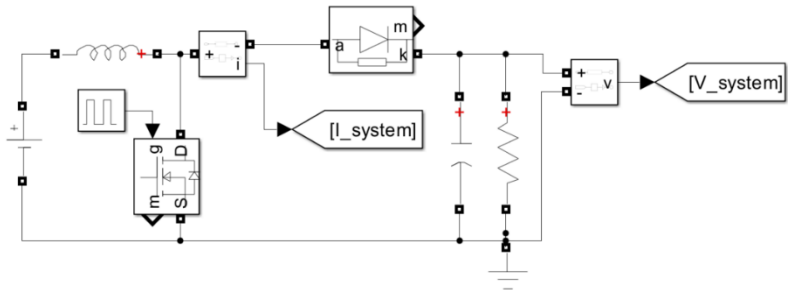

Figure 3. Boost converter circuit diagram

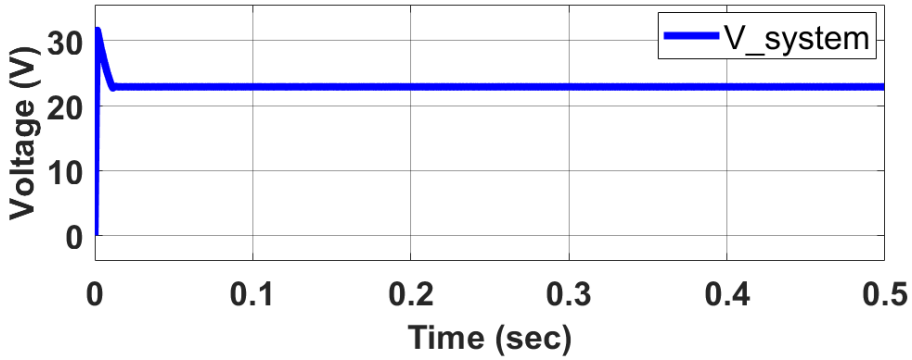

Figure 4. Voltage response (V_system) of Boost converter for D=0.5

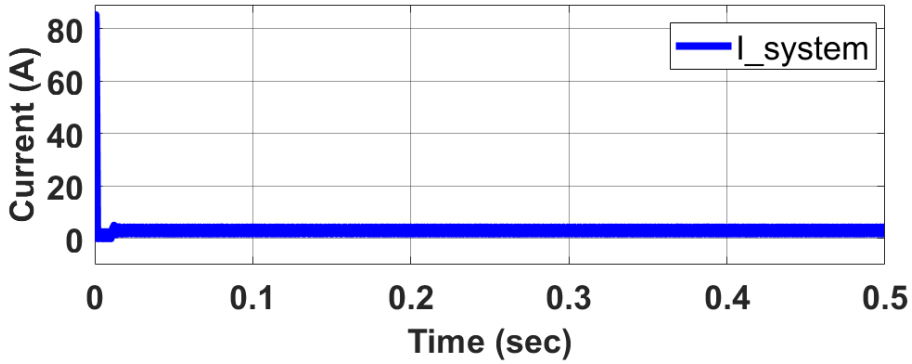

Figure 5. Current response (I_system) of Boost converter for D=0.5

When the obtained values are substituted in matrix *A* and matrix *B*, equation 8 is obtained.

$$\begin{bmatrix}\Delta\dot{x}_1\\ \Delta\dot{x}_2\end{bmatrix} = \underbrace{\begin{bmatrix}0 & -1.087\times10^4\\ 367.647 & -49.0196\end{bmatrix}}_{A}\begin{bmatrix}\Delta x_1\\ \Delta x_2\end{bmatrix} + \underbrace{\begin{bmatrix}4.9761\times10^5\\ -1.325\times10^3\end{bmatrix}}_{B}\Delta u \quad (8)$$

In order to derive the system's transfer function, the transformation from state space to the transfer function, as outlined in Equation 9, is applied. Substituting the values into this equation yields Equation 10. After performing the necessary operations, the system's transfer function is obtained, as shown in Equation 11.

$$\frac{\Delta x_2}{\Delta u} = C[sI - A]^{-1}[B] \quad (9)$$

$$\frac{\Delta x_2}{\Delta u} = [0 \quad 1]\left[\begin{bmatrix}s & 0\\ 0 & s\end{bmatrix} - \begin{bmatrix}0 & -1.087\times10^4\\ 367.647 & -49.0196\end{bmatrix}\right]^{-1}\begin{bmatrix}4.9761\times10^5\\ -1.325\times10^3\end{bmatrix} \quad (10)$$

$$G_p = \frac{\Delta x_2}{\Delta u} = \frac{-1.3265\times10^3 s + 1.8294\times10^8}{s^2 + 49.0196s + 3.996\times10^6} \quad (11)$$

2.2.2.2 PI Control Design

In many studies, PI controllers can be designed using classical methods such as Zeigler-Nichols, Cohen, or pole assignment. The goal is to position the poles of the closed-loop system to achieve the desired performance. This desired pole location depends on the

designer's experience and the root locus curve. Figure 3a shows that while the settling time is low, the overshoot is high. The system's low time constant indicates a fast response, as seen in its step response. For $t_{settling}$=0.1sec, the desired pole location is determined using Equations 12.

$$ts = 0.1 = \frac{4}{\xi wn} \rightarrow \xi wn = 40.00 \rightarrow s_1 = -40 \quad (12)$$

The general form of the PI controller is given in Equation 13.

$$G_C = K_p + \frac{K_i}{s} \quad (13)$$

The closed-loop structure of the system is illustrated in Figure 6. The closed-loop response is derived by substituting the values of the transfer function and the controller, as detailed in Equation 14. The characteristic equation ($F(s)$) of the system is given by Equation 15. After arranging this equation, the result is shown in Equation 16.

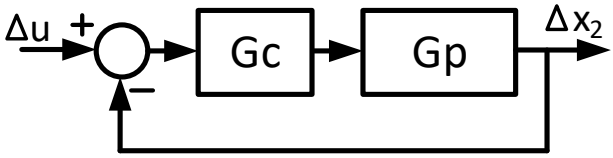

Figure 6. Closed loop system

$$G = \frac{G_c G_p}{1 + G_c G_p} \quad (14)$$

$$F(s) = 1 + G_c G_p$$

$$F(s) = 1 + \left(K_p + \frac{K_i}{s}\right)\left(\frac{-1.3265 \times 10^3 s + 1.8294 \times 10^8}{s^2 + 49.0196s + 3.996 \times 10^6}\right) \quad (15)$$

$$F(s) = s(s^2 + 49.0196s + 3.996 \times 10^6)$$
$$+K_p(s + K_i/K_p)(-1.3265 \times 10^3 s + 1.8294 \times 10^8) = 0 \quad (16)$$

When K_i/K_p=1,000 is selected for the design, the rlocus curve that gives the root locus curve of the system is obtained as in Figure 7a and Figure 7b. Figure 7b is obtained, when Figure 7a is zoomed in to see it in more detail. When the pole value seen in this graph reaches -40, the gain value K_p in the Rclous curve is found to be 0.00091. In this case, K_i=0.91 is calculated.

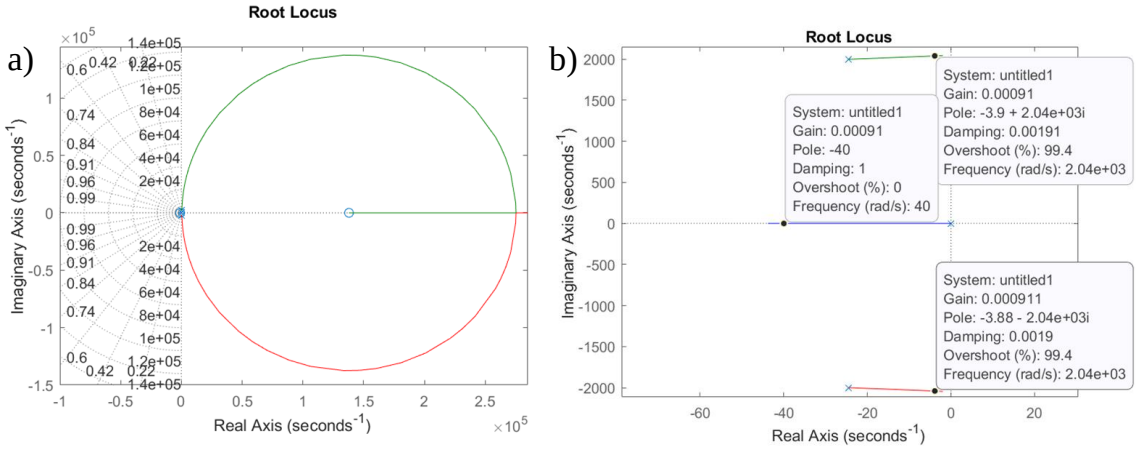

Figure 7. Root locus of closed loop system

In addition, the controller design for the boost converter was carried out with the Zeigler Nichols method, which is widely used in the literature, and K_p=0.00011, K_i=1.00 was obtained.

2.2.2.3 State Feedback Design

The development of modern control theory was significantly advanced through the application of matrix operations to control systems. State Feedback Control is a powerful method that facilitates controller design by representing the system in state-space form. A key requirement for this approach is the ability to measure all system states. If direct measurement of states is not possible, observers such as the Luenberger or Kalman observers must be designed to estimate the states. The closed-loop block diagram used for controlling the Boost Converter is shown in Figure 8.

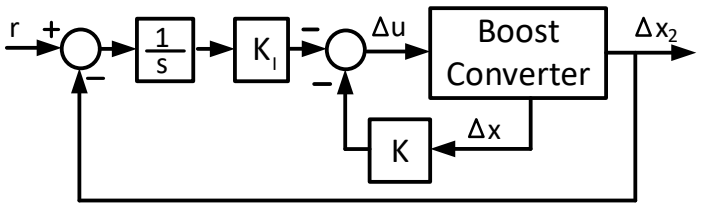

Figure 8. State feedback control of boost converter

In Figure 8, control system uses both an integrator and state feedback for control. Feedback is calculated using Equation 17, where x_I represents the integral of the output-input error. The integral of x_I is expressed in Equaiton 18. System equations are provided in Equation 19. Substituting Δu into the state-space equations leads to Equation 20. Subseqently Equation 20 is expressed in matrix form as Equation 21. Simplifying results in Equation 22, where the closed-loop matrix A_{cl} defines the system's behavior.

$$\Delta u = -K\Delta x - K_I x_I \quad (17)$$

$$\dot{x}_I = r - y = r - C\Delta x \quad (18)$$

$$\Delta\dot{x} = A\Delta x + B\Delta u \quad (19)$$
$$\dot{x}_I = r - C\Delta x$$

$$\Delta\dot{x} = (A - BK)\Delta x - BK_I x_I \quad (20)$$
$$\dot{x}_I = r - C\Delta x$$

$$\begin{bmatrix}\Delta\dot{x}\\ \dot{x}_I\end{bmatrix} = \begin{bmatrix}A - BK & -BK_I\\ -C & 0\end{bmatrix}\begin{bmatrix}\Delta x\\ x_I\end{bmatrix} + \begin{bmatrix}0\\ 1\end{bmatrix} r \quad (21)$$

$$\begin{bmatrix}\Delta\dot{x}_1\\ \Delta\dot{x}_2\\ \dot{x}_I\end{bmatrix} = \underbrace{\begin{bmatrix}A - B[K_1 & K_2] & BK_I\\ 0 & -1 & 0\end{bmatrix}}_{A_{cl}}\begin{bmatrix}\Delta x_1\\ \Delta x_2\\ x_I\end{bmatrix} + \begin{bmatrix}0\\ 0\\ 1\end{bmatrix} r \quad (22)$$

In Equation 22, unknown K_1, K_2 and K_I values in A_{cl} will be determined according to the performance criteria of the system. This criteria, namely the settling time, were determined as 0.2 sec. Therefore, the poles to be assigned are desired to be $s_{1,2}$=-3.9±2040i and s_3=-40 as in Figure 5b. Ackherman formula was used to determine the parameter values controlling the system by assigning these poles. The parameters obtained using this formula are obtained as K_1=-2.3014x10^{-8}, K_2=9.1181x10^{-4} and K_I=-0.9099.

3. Results and Discussion

Simulation studies were conducted using a computer with an Intel(R) Core(TM) i5-9400 CPU @ 2.90GHz, 64-bit, and 8GB RAM, with Matlab as the primary tool. Q Learning Algorithm was integrated into Simulink to control the boost converter, alongside PI and State Feedback Controllers. This section presents and compares the results and system responses of the boost converter controlled by PI, state feedback, and Q Learning. Different reference signals (24V, 20V, and 36V) were applied to the boost converter over a 0-1.5 second range. Due to the system's non-linearity and the discrete nature of Q Learning, no reference or control signals were applied during the first 100 ms, after which the signals were introduced.

Q Learning Algorithm operated iteratively within Simulink over 1.5 seconds, learning with each of the 1,000 iterations (iter_max). In contrast, PI and State Feedback Controller require only a single run post-design. The system responses and calculated ISE and IAE values were presented graphically. These results were compared and analyzed, with a comparison table provided for clarity.

Figure 9 is exhibitted the system responses when controlled using Zeigler-Nichols and Figure 10 is indicated system responses for pole assignment methods. The graph shows that while the system output exhibited minimal oscillation or chattering, the model displayed noticeable chattering and zigzags. This is attributed to the system having a zero in the right half plane in the closed loop. Despite this, the system generally achieved the desired reference signal within the expected settling time with no steady-state error. Furthernore, Figure 11 shows the response of the boost converter when controlled with state feedback including an integrator. The system have reached the reference signal within the designed settling time without overshoot and exhibited less chattering compared to the PI controller. Additionally, Figure 12 displays the system response of Q Learning control. Although the Q Learning Algorithm allowed the system to quickly reach the desired reference, it results in overshoot.

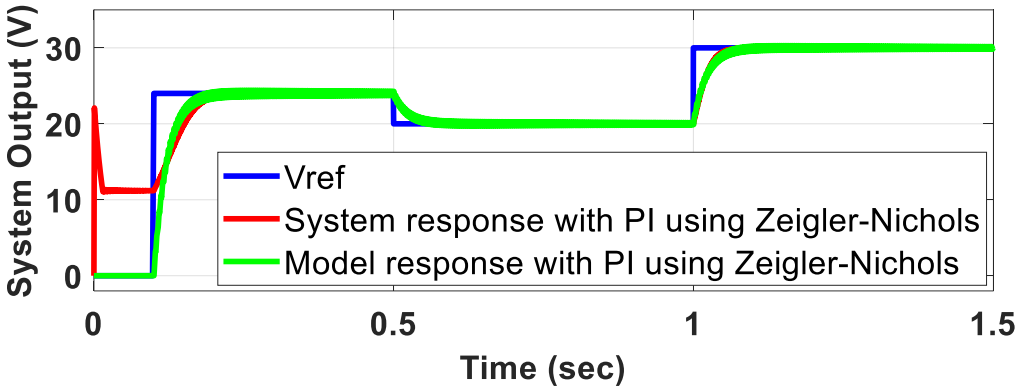

Figure 9. Model and system responses of boost converter controlled by PI using Zeigler-Nichols

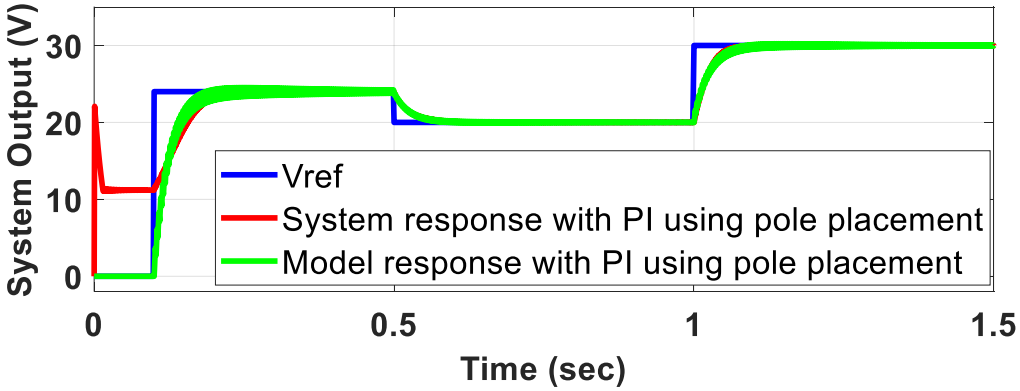

Figure 10. Model and system responses of boost converter controlled by PI using pole placement

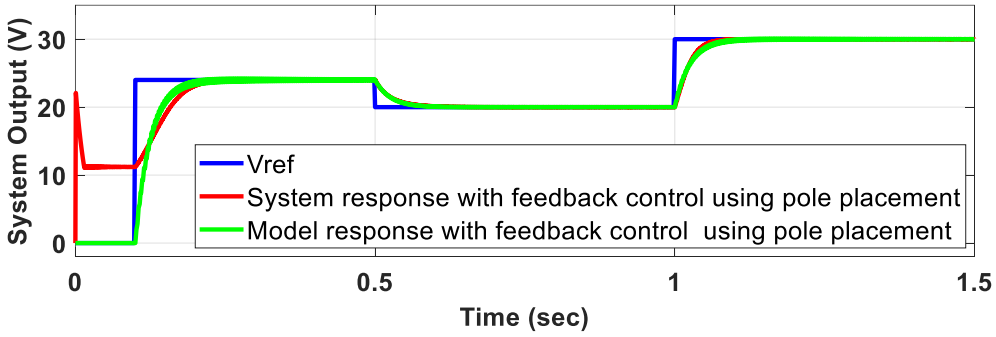

Figure 11. Model and system responses of boost converter controlled by feedback control

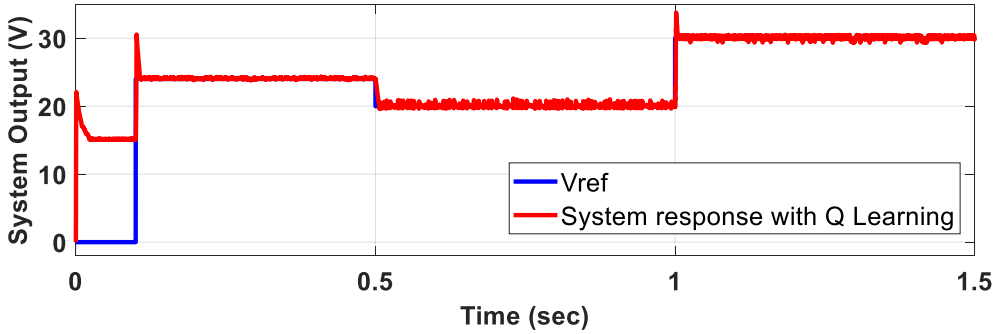

Figure 12. Model and system responses of boost converter controlled by Q learning controler

The display of all controller method results together was given in Figure 13. In particular, separate graphs were drawn for the 0-1.5 seconds and 0-0.5 seconds intervals to enable the results to be displayed in more detail. When Figure 13 is examined, it is seen that the Q learning algorithm reached the desired reference signal faster but it has overshoot.

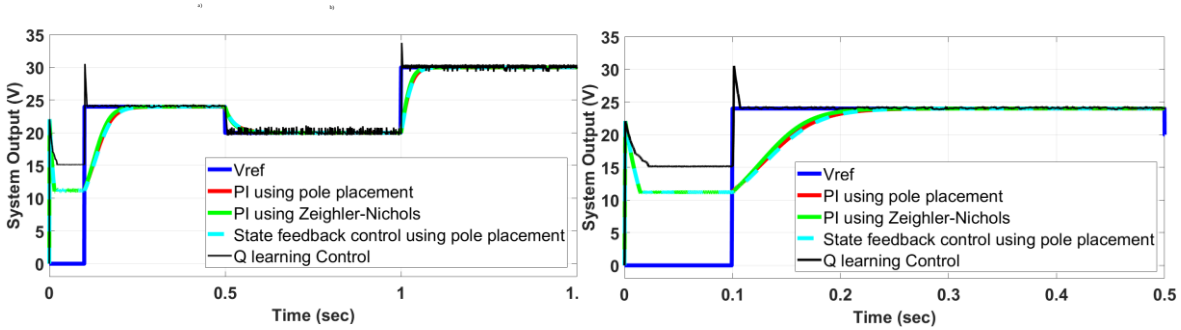

Figure 13. PI control using pole placement, Zeigler Nichols, state feedback control with integral action and Q learning responses for (a) 0-1.5 sec and (b) 0-0.5 sec Boost Converter

IAE and ISE values calculated according to the system responses are important and generally accepted performance indices in the design of controllers. The lower these values indicats that the performance of the controller is better. The performance of the designed controllers with respect to IAE and ISE is given in Figure 14. As can be seen, these values obtained with PI and State Feedback Control produced higher values than the Q learning algorithm.

In addition, the simulation time of the Q learning algorithm is given in Figure 15, calculated using Equation 4. Especially in the 100ms when there was no control, the reward value was very low, but it began to decrease when the control process started. It also peaked when the reference signals changed, but it still started to decrease on average.

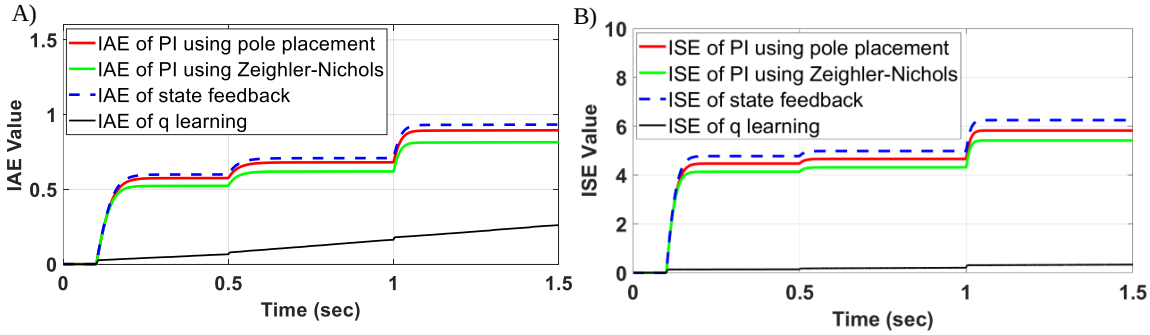

Figure 14. IAE and ISE values of Boost Converter Responses for PI, State feedback Control and Q Learning Control

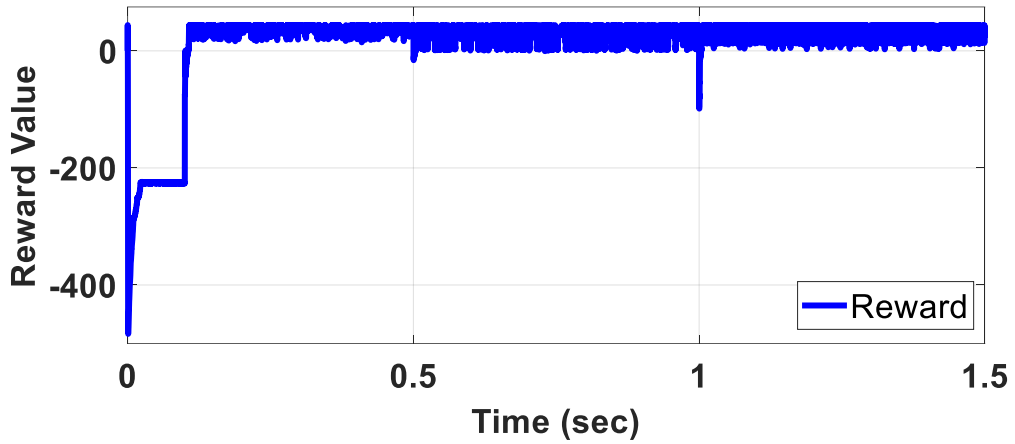

Figure 15. Reward Values for Boost Converter for Q Learning Control

ISE, IAE, $t_{settling}$ and Overshoot values of the boost converter controlled by different controllers are given in Table 4. When the results are examined in terms of ISE, IAE and $t_{settling}$, the Q learning method showed the best performance. However, when examined in terms of overshoot, the PI and State Feedback Controllers gave results without overshoot. Note that, $t_{settling}$ and Overshoot values were values calculated for the period between 0-0.5 sec. ISE and IAE were values calculated for the period between 0-1.5 sec. The reason why the overshoot was especially high in the Q learning algorithm is that the Q learning algorithm run depending on the state. The control signal was generated in a way that would maximize the recommended reward function for the state. The overshoot can be reduced by

Table 4. Performance results of Control Methods

Control Method	ISE	IAE	$T_{settling}$	$Overshoot\%$
PI using pole placement	5.825	0.895	0.08989 sec	0%
PI using Zeigler Nichols	5.413	0.815	0.08061 sec	0%
State Feedback Control	6.246	0.933	0.09195sec	0%
Q learning Control	0.331	0.260	0.0062sec	27.08%

revising reward function that using both adding different terms and constraints.

Generally speaking, control methods such as PI, PID, and state feedback have been widely used in the industry for the control of real-time systems. However, in the coming years, machine learning-based experimenters that can learn and make decisions in the control of real-time systems will become more common, especially as the processing capabilities of computers improve, their computational speeds increase, and their accessibility decreases and production costs increase. This will enable the use of machine learning-based control methods for many equipment used in renewable energy sources. This means that in the coming years, the use of flexible controllers that can quickly adapt to changing dynamics will become widespread. Especially in renewable energy sources, the high ability of

controllers to adapt to changing dynamics will have an effect of reducing costs as well as increasing system performance.

4. Conclusion

This study investigates the design of controllers for a boost converter using PI control, State Feedback Control, and Q Learning Algorithm. The boost converter was modelled, linearized and simulated in MATLAB/Simulink, with the converter represented in state-space form and an average model derived. For linear controller design, the average model was used in order to design PI controller using Ziegler-Nichols and pole placement methods. State Feedback Controller with integral action was then designed using pole placement. Q Learning Algorithm has been applied by configuring the states, reward function, and actions. Although Q Learning demonstrated superior performance in terms of ISE, IAE, and settling time, it has exhibited less favorable overshoot performance. In future studies, this deficiency will be addressed by making changes to the reward function for overshooting. Overall, Q Learning is advantageous for its simplicity, model-free implementation, and capacity for continuous learning and adaptation, making it suitable for applications in electrical energy systems, power systems, and electric vehicles, where boost converters play a crucial role in enhancing energy efficiency in the context of renewable energy.

5. References

Abdalla, A. O. M., Ibrahim, A. A. Z., Fadul, S. M. E. 2022. Modeling single diode PV using particle swarm optimization (PSO) techniques. *Gaziosmanpaşa Bilimsel Araştırma Dergisi*, *11*(1), 44–56.

Abdalla, A. O. M., Önbilgin, G. 2024. An Approach to Optimized the Output Power of Photovoltaic System Using Artificial Neural Networks. *Gaziosmanpaşa Bilimsel Araştırma Dergisi*, *13*(1), 18–30.

Ahmed, S. H., Ahmad, I. 2024. Optimal wireless power transfer to hybrid energy storage system for electric vehicles: A comparative analysis of machine learning-based model-free controllers. *Journal of Energy Storage*, *75*, 109534.

Akyurek, H. A., Bucak, İ. Ö. 2012. Zamansal-Fark, Uyarlanır Dinamik Programlama ve SARSA Etmenlerinin Tipik Arazi Aracı Problemi için Öğrenme Performansları. *Akıllı Sistemlerde Yenilikler ve Uygulamaları Sempozyumu*.

Alfred, D., Czarkowski, D., Teng, J. 2024. Reinforcement Learning-Based Control of a Power Electronic Converter. *Mathematics*, *12*(5), 671.

Alkrunz, M., Yazıcı, İ. 2016. Design of discrete time controllers for the DC-DC boost converter. *Sakarya University Journal of Science*, *20*(1), 75–82.

Angiuli, A., Fouque, J. P., Laurière, M. 2022. Unified reinforcement Q-learning for mean field game and control problems. *Mathematics of Control, Signals, and Systems*, *34*(2), 217–271.

Borase, R. P., Maghade, D. K., Sondkar, S. Y., & Pawar, S. N. 2021. A Review of PID Control, Tuning Methods and Applications. *International Journal of Dynamics and Control*, *9*, 818–827.

Bououden, S., Hazil, O., Filali, S., Chadli, M. 2014. Modelling and model predictive control of a DC-DC Boost converter. *In 2014 15th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)*, 643–648.

Çimen, M., Garip, Z., Boz, A. 2021. Chaotic flower pollination algorithm based optimal PID controller design for a buck converter. *Analog Integrated Circuits and Signal Processing*.

El Fadil, H., Giri, F. 2007. Backstepping based control of PWM DC-DC boost power converters. *In 2007 IEEE International Symposium on Industrial Electronics*, 395–400.

Farajdadian, S., Hajizadeh, A., & Soltani, M. (2024). Recent developments of multiport DC/DC converter topologies, control strategies, and applications: A comparative review and analysis. *Energy Reports*, *11*, 1019–1052.

Garip, Z., Çimen, M. E., Boz, A. F. 2021. Meta-sezgisel algoritmalar kullanarak güneş pili modellerinin parametre çıkarımında karşılaştırmalı performans analizi. *Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, *2*, 1133–1144. https://doi.org/10.17341/gazimmfd.586269

Gheisarnejad, M., Khooban, M. H. 2023. Quantum Deep Learning for Fast Switching of Full-Bridge Power Converters. *Designs*, *7*(3), 60.

Güldemir, H. 2005. Sliding mode control of DC-DC boost converter. *Journal of Applied Sciences*, *5*(3), 588–592.

Güngör, O., Yüksek, H. İ. 2020. Modeling of Boost and Cuk Converters and Comparison of Their Performance in MPPT. *Sigma Journal of Engineering and Natural Sciences*, *11*(1), 83–101.

Harmon, M. E., Harmon, S. S. 1996. Reinforcement learning: A tutorial. *WL/AAFC, WPAFB Ohio*, *45433*, 237–285.

Kang, H., Jung, S., Kim, H., Jeoung, J., Hong, T. (2024). Reinforcement learning-based optimal scheduling model of battery energy storage system at the building level. *Renewable and Sustainable Energy Reviews*, *190*, 114054.

Ibrahim, O., Yahaya, N. Z., Saad, N. 2016. Comparative studies of PID controller tuning methods on a DC-DC boost converter. *In 2016 6th International Conference on Intelligent and Advanced Systems*, 1–5.

Meng, Q., Hussain, S., Luo, F., Wang, Z., Jin, X. (2024). An online reinforcement learning-based energy management strategy for microgrids with centralized control. *IEEE Transactions on Industry Applications*.

Muktiadji, R. F., Ramli, M. A., & Milyani, A. H. 2024. Twin-Delayed Deep Deterministic Policy Gradient Algorithm to Control a Boost Converter in a DC Microgrid. *Electronics*, *13*(2), 433.

Nishanthi, B., Kanakaraj, J. 2024. Enactment of deep reinforcement learning control for power management and enhancement of voltage regulation in a DC micro-grid system. *Electric Power Components and Systems*, *52*(4), 555–565.

Li, Y., Wu, J., Pan, Y. (2024). Deep reinforcement learning for online scheduling of photovoltaic systems with battery energy storage systems. *Intelligent and Converged Networks*, *5*(1), 28-41.

Liu, Q., Guo, Y., Xu, T. (2024). Robust Deep Reinforcement Learning for Inverter-based Volt-Var Control in Partially Observable Distribution Networks. *arXiv preprint arXiv:2408.06776*.

Palpandian, P., Govindaraj, V., Dinesh, S., Megalan, K., Sivaprasanth, K., Vigneshwaran, G. 2024. Intelligent Voltage Regulation Using Machine Learning-Enhanced Boost Converter. *In 2024 International Conference on Science Technology Engineering and Management (ICSTEM)*, 1–7.

Panggabean, J., Sutisna, N., Syafalni, I., & Adiono, T. 2023. Comparison of MPPT based on Deep Reinforcement Learning by DQN, DDPG and TD3. *In 2023 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)* , 261–266.

Rajamallaiah, A., Karri, S. P. K., & Sankar, Y. R. (2024). Deep Reinforcement Learning Based Control Strategy for Voltage Regulation of DC-DC Buck Converter Feeding CPLs in DC Microgrid. *IEEE Access*.

Saha, U., Shahria, S., & Rashid, A. B. (n.d.). Proximal Policy Optimization-Based Reinforcement Learning Approach for DC-DC Boost Converter Control: A Comparative Evaluation Against Traditional Control Techniques. 2023. *ArXiv Preprint ArXiv:2310.02945*.

Sezen, A., Keskin, K. 2021. Hybrid Control of DC-DC Buck Boost Converter. *Demiryolu Mühendisliği*, *14*, 99–109.

Smart, W. D., Kaelbling, L. P. 2000. Practical Reinforcement Learning in Continuous Spaces. *ICML*.

Su, T., Wu, T., Zhao, J., Scaglione, A., & Xie, L. (n.d.). A Review of Safe Reinforcement Learning Methods for Modern Power Systems. *ArXiv Preprint ArXiv:2407.00304*.

Sun, Z., Lu, T. (2024). Collaborative operation optimization of distribution system and virtual power plants using multi-agent deep reinforcement learning with parameter-sharing mechanism. *IET Generation, Transmission & Distribution*, *18*(1), 39-49.

Uçmaz, B. N., Yakut, Y. B. (2024). PEM yakıt pillerinde PI, PSO ve FOPI kontrollü DC/DC dönüştürücülerine ilişkin performanslarının karşılaştırılması. *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi*, *15*(1), 23–29.

Wang, H., Emmerich, M., Plaat, A. (2018). Monte Carlo Q-learning for General Game Playing. *ArXiv Preprint ArXiv:1802.05944*.

Xu, J. H., Zhang, B. X., Yan, H. Z., Ding, Q., Zhu, K. Q., Yang, Y. R., & Wang, X. D. (2024). Sliding mode–Extended state observer control strategy to improve energy transfer of PEMFC connected DC-DC boost converter system. *Sustainable Energy Technologies and Assessments*, *63*, 103645.

You, W., Yang, G., Chu, J., & Ju, C. (2023). Deep reinforcement learning-based proportional–integral control for dual-active-bridge converter. *Neural Computing and Applications*, *35*(24), 17953–17966.