Energy and Power Engineering, 2026, 18(2), 103-120
https://www.scirp.org/journal/epe
ISSN Online: 1947-3818
ISSN Print: 1949-243X

Hybrid Deep Reinforcement Learning and Model Predictive Control for Microgrids

Hamdi Naji Almaghrabi, Mousa Marzband, Abdullah Abusorrah

Department of Electrical and Computer Engineering, Faculty of Engineering, King Abdulaziz University, Jeddah, Saudi Arabia
Email: halmaghrabi0029@stu.kau.edu.sa, mmarzband@kau.edu.sa, abusorrah@kau.edu.sa

How to cite this paper: Almaghrabi, H.N., Marzband, M. and Abusorrah, A. (2026) Hybrid Deep Reinforcement Learning and Model Predictive Control for Microgrids. *Energy and Power Engineering*, **18**, 103-120.
https://doi.org/10.4236/epe.2026.182006

Received: November 24, 2025
Accepted: February 8, 2026
Published: February 11, 2026

Abstract

The current microgrids are experiencing growing difficulties in voltage stability and operational capacity, particularly with constant power loads (CPLs), leading to negative impedance behavior and probability of voltage collapse. Although there are notable improvements in both conventional and smart control strategies, there is still a research gap, namely creating a framework of control that facilitates stability precision and quick adaptation to operation alterations without augmenting the computational load and the operational risks. The current work was based on the creation of a mathematical model based on state equations to model the dynamics of a system, as well as the application of a numerical simulation environment to test the performance of three control modes: MPC, RL, and an environment of a hybrid (hereafter referred to as Hybrid MPC -RL) framework. The approach involved the creation of an MPC algorithm where prediction horizons are brief, and the incorporation of a DRL algorithm into the structure of a hybrid regime that can switch or combine the two strategies based on operating conditions. The findings indicated that on average the DC bus voltage was about 46.7 V compared to the reference of 48 V with a standard deviation of 2.98 V and the current was between 4.09 and −3.97 A with an average of about zero. The data on working indicated that the system used MPC 99.7% of the time and 0.3% in the hybrid mode, without any use of RL alone. This proves the reliability and consistency of MPC, the adaptability of DRL, and the effective combination of these two approaches. The scientific feature of this paper is the addition of a hybrid framework that combines MPC and DRL in an adaptive safety mechanism that leads to high stability in CPL conditions and improves the efficiency of operation over the traditional control.

Keywords

Model Predictive Controller, Deep Reinforcement Learning, Hybrid Control, Microgrids, Voltage Stability

1. Introduction

Microgrid systems are quickly developing as renewable energy sources like solar and wind power are increasingly being used and as the stability and reliability of these sources become more complex with the erratic and unpredictable characteristics of the sources [1]. Constant power loads (CPLs) are defined by some peculiarities including negative impedance. These issues are not the least significant reasons of acute changes in voltage and probability of voltage breakdown [2]. On this conclusion, the development of control methods has been necessary to provide steady and effective functioning of microgrids on the condition that time limitations and insufficient computing resources are considered [3] [4]. Model-based predictive control (MPC) as one of the most prominent modern approaches to control has been examined because it can manage a variety of constraints and project a system behavior in the future [5]. This controller however has a high computing power which can hamper its operation as a modern controller when there are unexpected situations [6]. On the other hand, deep reinforcement learning (DRL) algorithms provide important adaptive learning properties to dynamic environments, which allows the controller to acquire optimal operating strategies in the long run [7]. Investigating some of the issues of deep other reinforcement learning (DRL) algorithms, DRA might not be as stable as it is supposed to be in a rigid environment with the use of DRA, thus it is not that applicable in sensitive networks like microgrids [8].

The overall structure of a microgrid and its main elements are reflected according to [11]: renewable sources, conventional controllable modules, storage units, loads and connection to the main grid. Central control is based on a deep reinforcement learning (DRL) algorithm, which is fed with real-time system status and provides the right control signals. This explanatory scheme is useful in the context of comprehending the connection between the energy flow and information exchange to apply hybrid approaches to the problem that include DRL + MPC.

In the wake of such difficulties, there has been a tendency towards the development of hybrid control structures that merge the robustness of MPC at the attainment of stability and control within constraints with the adaptability of DRL at responding to the evolving conditions and enhancing the operation efficiency [2] [9]. This method should be used to strike a reasonable tradeoff between safety and flexibility: to ensure that DRL can be used in steady-state conditions, whereas MPC should be used in emergency conditions; or both can be used with an adaptive weighting mechanism [10]. Hybrid MPC-DRL control, therefore, is another solution with potential in meeting voltage stability, power control, and enhancing the overall performance of the microgrid.

Although there have been remarkable advancements in traditional and smart methods in control of microgrid systems, most of the past studies have concentrated on specific segments of the control algorithm, where model predictive control (MPC) has been applied independently to attain voltage stability, or the deep and reinforcement learning (DRL) algorithms have been applied to enhance the efficiency of their operations and adaptation to the dynamism. Nevertheless, a distinct research gap still exists on how the two approaches can be combined into a hybrid system to balance reliability and flexibility particularly with constant-power loads (CPLs) which present a special challenge to microgrid stability. With the increasing complexity of dynamic and nonlinear systems over the last few years, the limitations of conventional control schemes such as PI and standard MPC have come into prominence under varying operating conditions and model ambiguity. Traditional controllers mostly rely on accurate system models and constant parameters, thus restricting their adaptability and potential for real-time optimization. This issue invites the development of advanced intelligent control systems that are capable of learning, adaptation, and real-time prediction of system behavior. By integrating machine learning paradigms—primarily, Deep Reinforcement Learning (DRL)—with model-based predictive control (MPC), it is possible to achieve a synergy between model-based optimization and data-driven adaptability. This urge drives the current research, which focuses on designing and verifying a hybrid MPC-DRL method with a potential for greater accuracy, robustness, and stability for complicated control scenarios. Thus, to define what has been already done and what still has to be done, a thorough examination of the prior research on the implementation of MPC, DRL, and hybrid solutions in the microgrid sector is required.

2. Literature Review

The past years have seen the growing attraction of interest to the combination of MPC—methods with—DRL-algorithms to enhance the stability and effectiveness of microgrid systems and associated applications. Some of these studies have discussed this issue in various aspects that have offered a strong scientific base on which this study can be based.

As one such example, the DRL can reduce harmonic distortion and increase the efficiency of grid injection compared to traditional control under the conditions of changes in the voltage, as per a recent study presented by [2] which conducted a review studying the application of deep reinforcement learning algorithm to the management of grid connected inverters using an LCL filter when using solar energy [2]. On the other hand, according to (Yu, *et al.*, 2025), this study concluded that with the DRL, the direct current regulation can be more effective to reduce the harmonic distortion and increase the efficiency of grid injection than [3].

More generally, in (Yu, *et al.*, 2025), MPC was used together with DRL to deploy a novel energy management system to solve fuel cell/battery hybrid systems. The results showed that the combination trade-offs exist between the rapidity response of MPC and adaptive learning of DRL resulting in reduced voltage oscillation and extended battery duration [3]. Similarly, Li, *et al.*, (2024) presented the LearnAMR framework, which relies on adaptive MPC with the introduction of RL algorithms to provide greater flexibility in managing buildings energy, yet with the further addition of demand side management (DSM) to obtain even better efficiency [4].

Razmi *et al.* (2025) on the other hand talked about the possibility of using DRL in conjunction with dynamic MPC in real-time multiagent management of a microgrid. It was found that there were large positive changes in adaptation to changing loads in comparison with MPC-based control itself [12]. Fu *et al.* (2024) suggested a reactive compensation system to operate daily energy management in shipboard systems, with the help of RL, which is dynamic predictive control (DMPC) [12].

The outcomes prove a better forward planning and minimization of energy loss. Elshazly *et al.* (2024), in another direction, introduced a practical implementation of RL in the coordination of electric vehicles charging in a fair and efficient process, which will improve the work of a smart charging system with the grid systems [13].

According to Xu, *et al.*, (2024), six scenarios are studies using combined MPC-DRL framework methods, these scenarios are Low-level noise & demand 1; Medium-level noise & demand 1; High-level noise & demand 1; Low-level noise & demand 2; Medi-um-level noise & demand 2; and High-level noise & demand 2. The results concluded that the learning curves of the combined MPC-DRL framework methods started with higher rewards and converged faster than the standalone DRL methods for all the scenarios [9]. The reason was that the MPC module within the MPC-DRL framework generates basic control inputs that provide baseline control performance and guide the DRL to learn, and the DRL module within the framework only requires a limited exploration space with smaller action bounds and thus requires fewer sample data, compared with the standalone DRL agent.

Table 1 shows the Related Studies on MPC, DRL, and Hybrid Control in Microgrids. Accordingly, it is possible to state that the use of MPC alone can guarantee high stability under constrained conditions, but will not bring flexibility and responsiveness to the dynamics of the environment, whereas DRL can do it better but cannot guarantee safety. Therefore, there have been requests to establish hybrid structures to integrate the two methods in order to balance the stability and flexibility [12].

Although the application of model predictive control (MPC) and deep reinforcement learning (DRL) algorithms to microgrid systems is significantly advanced, research gaps remain apparent that restrain the use of these methods given their present status of research. An example is that it has been found that MPC can provide stability to a severe constraint but it has a high computation and its performance is less resilient when dealing with complicated dynamic loads in the form of constant-power loads (CPLs). Conversely, DRA provides great flexibility to fluctuating situations, and does not provide safety assurances, resulting in unsteady actions, particularly in response to temporary perturbations. In addition, the integration of MPC and DRL has been studied mainly conceptually without details of the mechanisms that can be used to switch or mix in real-time. Moreover, there is a lack of studies directly relating to harmonic resonance and the reduction of total harmonic distortion (THD). In addition, the majority of these studies are restricted to simulation settings with little to no extrapolation of these models into practical validation systems with hardware-in-the-loop (HIL) or field testing (See Table 2).

Table 1. Related studies on MPC, DRL, and hybrid control in microgrids.

Ref.	Main Approach/Application	Identified Drawbacks or Limitations	Contribution of the Present Work
[1]	Hybrid MPC-DRL for smart EV charging	Focused on single-domain optimization (EV charging); lacks general adaptability to nonlinear dynamic systems.	Extends hybrid MPC-DRL control to dynamic CPL microgrids with adaptive switching and safety layer.
[2]	MPC with RL-based speed profile generation	Applied in simulation only; limited validation; lacks energy system context.	Applies integrated control to power electronics with real-time voltage stability under CPL.
[3]	DRL for insulin delivery	Healthcare-oriented; does not address physical control stability issues.	Adapts RL concepts for nonlinear control in microgrids ensuring voltage stability.
[4]	DRL for multi-robot path planning	Focused on spatial navigation; lacks dynamic control modeling or safety switching.	Introduces dynamic switching hybrid controller for continuous-time energy systems.
[5]	Continuous DRL + MPC for robot trajectory	Computationally intensive; unsuitable for fast-reactive control.	Reduces computational load via short-horizon MPC and adaptive DRL activation.
[6]	Comparison of ML vs MPC for home batteries	Separate comparison, no hybrid integration.	Combines MPC reliability with DRL adaptability within unified control framework.
[7]	DRL for energy management in microgrids	Long training time; unstable under fast dynamic variations.	Hybrid MPC-RL stabilizes CPL fluctuations with minimal retraining needs.
[8]	Secure DRL-based microgrid control	Focus on cybersecurity; lacks performance optimization under dynamic CPL.	Prioritizes dynamic stability and adaptive control under variable loads.
[9]	Imitation learning for power dispatch	Limited adaptability; no predictive optimization layer.	Integrates prediction (MPC) and learning (DRL) for adaptive energy management.
[10]	Privacy-preserving ML for smart grids	Security focus; not a control-oriented study.	Introduces control-specific hybrid model emphasizing real-time voltage control.

Continued

[11]	DRL for economic microgrid control	Sensitive to uncertainties; lacks hybrid structure.	Improves robustness using dual-agent hybrid regime (MPC-RL).
[13]	RL-driven dynamic MPC for multi-agent microgrids	High computational complexity and limited safety assurance.	Simplifies structure via short-horizon MPC and safety-aware hybrid logic.
[14]	RL-based compensation for shipboard power	Requires high computational resources; no switching optimization.	Introduces lightweight hybrid model with adaptive RL engagement.
[15]	RL for smart charging coordination	Coordination focus; not suitable for voltage regulation.	Addresses stability precision and adaptive energy balancing.
[12]	MPC-DRL for freeway control	Transportation context; lacks physical stability constraints.	Applies hybrid approach to physical dynamic systems under CPL stress.

This research will, therefore, fill this gap with an aim of developing a hybrid control framework that incorporates MPC to achieve stability and DRL to achieve adaptability and efficiency, and a dynamic hybrid switching/blending mechanism that can act in real time. This is promising a viable solution to the problems of constant-power loads and harmonic resonance phenomena and a guarantee of better voltage stability and operational efficiency of microgrid systems.

Table 2. Research gaps identified in literature.

Aspect	Current Focus in Literature	Remaining Gaps/Limitations
Voltage stability under CPLs	MPC ensures stability under constraints	Limited focus on CPL-induced negative impedance effects.
Adaptability to dynamic loads	DRL shows flexibility in dynamic conditions	Lack of explicit safety guarantees during transient disturbances.
Hybrid integration	Some studies combine MPC and DRL	Few works address real-time switching/blending between controllers.
Computational burden	MPC requires high computational effort	Need for efficient hybrid designs reducing CPU load.
Experimental validation	Mostly simulation-based (MATLAB/Simulink, etc.)	Limited Hardware-in-the-Loop (HIL) or field-level implementations.
Harmonic suppression (THD)	Addressed indirectly in MPC or DRL frameworks	Insufficient dedicated hybrid strategies for minimizing THD in microgrids.

3. Research Methodology

The present research is based on an integrated approach to the methodology that integrates mathematical models, numerical simulations, and assembly of smart control algorithms. The main goal would be to create a hybrid control method that would use Model Predictive Control (MPC) and Deep Reinforcement Learning (DRL) algorithms to improve the stability of interconnected transformers form a distributed generation system, alleviate the challenges of constant power loads (CPLs), and deal with the challenges of operating mode transitions and harmonic resonance phenomena. Towards this end, the research methodology is categorized into four interrelated steps, which include the construction of the mathematical model, the development of the control strategies, the creation of a validation and testing environment, and lastly, the analysis and the evaluation tools as presented in Figure 1.

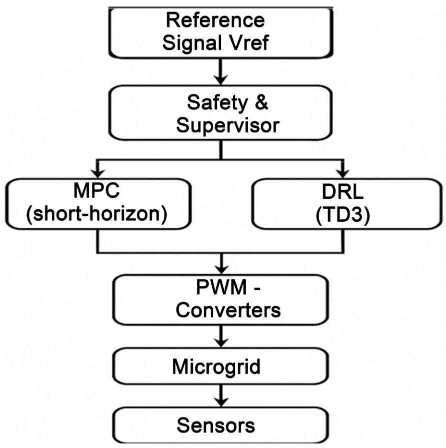


Figure 1. Research methodology.

3.1. Mathematical Modeling (System Modeling) and Physical System

The initial step was to create an integrated mathematical model that relied on the formulation of the state-space equations to model the dynamic behavior of voltage source inverters (VSIs) attached to constant power loads. In this model, a number of well-chosen components are introduced to bring the reality of power systems, with a close physical representation. Electronic transformer design now has models of IGBT/diode switches, including the consideration of the influences of internal resistance and thermal power loss. Snubber circuits and dead-time had been introduced in order to better model the transient behavior of electrical components. Moreover, fixed loads were modelled to indicate that they were nonlinear and of negative incremental impedance, and restrictions placed on the maximum current and nominal power to model realistic conditions.

The physical model was modeled in MATLAB/Simulink environment, through the Simscape Electrical library, which is very accurate in modeling the power systems. In order to make the model reliable, actual data of the Gatton Solar Farm was applied, which represents real world conditions of operation of distributed generation networks. A balance between computational efficiency and numerical accuracy was maintained by adopting a static ODE3 numerical integration algorithm with a time step of 50 microseconds, to enable the accurate simulation of high frequency oscillation phenomena without unreasonable computational cost.

3.2. Development of Strategy of Control

The second phase included the creation of a hybrid control structure that integrates two algorithms model predictive control (MPC) and deep reinforcement learning (DRL). The MPC algorithm was created to work with one-to-ten-time steps of prediction.

The hybrid MPC-DRL control system was implemented so that it can work in three operating mechanisms. MPC is entirely trusted in cases of critical circumstances or in the event of a threat of voltage collapse since it is predictive and mathematically guaranteed. DRL is trusted in stable situations with low risks to decrease the computational load and to deliver a quick reply. Lastly, when the frequency interference exists or during the transitional period, adaptive blending approach is employed between MPC and DRL by using the weighting coefficients (α, β) to maintain a dynamic balance between safety and the operational efficiency.

3.3. Validation Framework

Once the model was developed and the control strategies were developed, a detailed validation and testing environment was created in which a number of situations were incorporated to test the effectiveness of the proposed system. Dynamic stability testing (adding abrupt load changes) was initiated with a penetration rate of up to 80% of the CPLs after monitoring voltage variations and the resulting discontinuous changes. A state flow environment was used to monitor these transitions, with transient overcurrent and latency used as performance indicators.

The phenomena of harmonic resonance in multi-transformer systems were also studied in the test plan and the total harmonic distortion (THD) was computed, in the simulating environment.

The experiment was based on a collection of quantitative and qualitative analysis instruments to measure the work of the developed algorithms. This involved the time-domain analysis (TDA) of voltage, current, and control signal curves through time, which can be used to simply compare the behavior of MPC and DRL as well as the behavior of the hybrid framework. The instantaneous power was also estimated by determining the power of a given control (according to the equation $P(t) = v(t) \times i(t)$ and finally the total power consumption was also determined and the aim of this was to compare the efficiency of various control strategies. Besides that, the harmonic analysis was carried out more accurately with the help of the Fast Fourier Transform (FFT), which determined the total harmonic distortion (THD) as one of the most significant measurements of power quality. Other quantitative indicators of performance were also depended on, which included phase margin, settling-time and load ride-through capability.

Relying on the above, it can be stated that this methodology was developed to be multi-level and not restricted to the theoretical and analytical facet but also to testing and verification in a true simulation setting as well as semi-realistic HIL setting. Such combination of modeling, intelligent control and practical verification guarantees the predicted results will be very reliable and will be applicable to future distributed generation systems.

The proposed hybrid control system design is based on a set of very carefully chosen electrical and control parameters. The system design provides a pre-determined DC bus voltage typically at a fixed value, typically 48 volts in a low-voltage microgrid or 400 volts in a large-scale application. According to this level of voltage, rated load power is determined as the sum of critical and non-critical loads whereas the respective rated bus current is established as the ratio of the rated power to the bus voltage.

A switching frequency is selected based on power device datasheet and electromagnetic interference considerations, which is usually between 10 and 40 kilohertz. This parameter has a significant contribution to switching losses, efficiency, and harmonic performance. In order to get good control of current, the permissible ripple in the inductor is restricted to a fraction of the rated current, such as about 20 percent and this factor determines the size of the inductor to be employed in the converter. Similarly, the peak ripple of the capacitor is limited to a low percentage of the bus voltage, typically one percent, and this limits the value of the output capacitor to keep the voltage constant. The resonance frequency of the LC filter is also checked to prevent undesired oscillations and is adjusted much lower than the switching frequency.

The constant power load is defined in terms of nominal power and incremental impedance which is negative in nature. This is the negative impedance characteristic that makes CPLs difficult because they require more current as the voltage decreases. The snubber resistors and capacitors are added to overcome robustness, and the spikes and transients of the voltage are suppressed, and dead-time is introduced in the switching signals to avoid short circuiting of the devices.

Proportional-integral controllers are adjusted in value on both the voltage and current loops in terms of control. To control the DC bus voltage at a desired DC bus voltage, the outer voltage loop has been adjusted, but the inner current loop is set at a high bandwidth to give quick current response. Moreover, the model predictive controller is characterized by the sampling time, the prediction horizon and the weighting factors which determine the time objectives of accurate state tracking, minimized control effort and minimized switching.

At the reinforcement learning end, the state space is chosen with a very critical regard to the inclusion of various critical variables like bus voltage, inductor current, load power, state of charge, and indicators of harmonic distortion. The action bounds define the space of possible control actions that the agent may perform and the reward function is aimed at discouraging the deviations of voltages, high harmonics, switching very often, and saturation of control as showing in **Table 3**.

Table 3. Research gaps identified in literature.

Parameter	Symbol	Formula
DC bus voltage	V_{bus}	Given by design (e.g., 48 V DC or 400 V DC link)
Rated load power	P_{rated}	Sum of critical + non-critical loads
Rated bus current	I_{rated}	$I_{rated} = (P_{rated} * 1000)/V_{bus}$
Switching frequency	f_{sw}	Per device datasheet/EMI constraints (e.g., 10 - 40 kHz)
Allowed inductor ripple	ΔI	$\Delta I = r_I * I_{rated}$ (e.g., $r_I = 0.2 \to 20\%$ ripple)
Inductance (buck/boost stage)	L	$L \approx V_{bus}/(8 * f_{sw} * \Delta I)$ (approx. at $D \approx 0.5$)
Allowed capacitor ripple	ΔV	$\Delta V = r_V * V_{bus}$ (e.g., $r_V = 0.01 \to 1\%$)
Output capacitance	C	$C \approx I_{rated}/(8 * f_{sw} * \Delta V)$
LC resonance frequency	f_{res}	$f_{res} \approx 1/(2\pi\sqrt{LC})$; $place \ll f_{sw}/10$
CPL nominal power	P_{cpl}	Nominal constant power setpoint per CPL
CPL incremental impedance	Z_{cpl}	$Z_{cpl} = -V_{bus}^2/P_{cpl}$ (negative)
Snubber resistor	R_{snub}	From datasheet/EMI testing (e.g., 1 kΩ)
Snubber capacitor	C_{snub}	From datasheet/EMI testing (e.g., 0.1 μF)
Dead-time	t_{dead}	Device/switch driver (e.g., 2 μs)
PI Kp (outer voltage loop)	K_{PV}	Tune via desired crossover ω_c_v; use plant model $G\_v(s)$
PI Ki (outer voltage loop)	K_{iv}	$K_{iv} = K_{PV} * \omega\_i\_v$; choose $\omega\_i\_v \approx \omega\_c\_v/5$
PI Kp (inner current loop)	K_{pi}	Bandwidth ≈ 5 – 10× voltage loop
PI Ki (inner current loop)	Ki_i	$Ki\_i = K_{pi} * \omega\_i\_i$
MPC sampling time	T_S	Typically, 1 – 2 switching periods
MPC horizon	N	Start with $N = 1$ - 3 for FCS-MPC; increase if CPU allows
MPC weights	Q, R, λ	Track v, i and penalize switching; adjust for THD/overshoot

4. Results and Discussion

In this part, the results of simulation and analysis are shown to determine the performance of the various control strategies (MPC, RL and Hybrid) on the proposed microgrid. It is concentrated on the comparison between the response in voltage and the control signal at diverse operating conditions, such as sudden change of loads, and stable operating conditions. The purpose of this section is to comment on how each control method can be used to reach stability and minimize oscillations and the strengths and weakness of each control method and finally a thorough assessment of the viability of the proposed hybrid method. Duty cycles of the three controllers (MPC, RL and Hybrid) are displayed over the time of simulation. It is discovered that MPC, as well as Hybrid, retained a significant level of consistency and reproducibility of the signal, which indicates the disciplined switching. Instead, RL began with a very turbulent and unsteady signal during the first moments of operation (around 0.02 seconds before the switch on). This shows that the reinforcement learning-based controller does not have direct control over the system and needs time to adjust to changes. In such a way, the consistency of MPC is predetermined by the fact that the predictive controller puts down strict limitations on voltage and current, therefore, keeping the control signal with the allowed boundaries. The hybrid exploits this advantage and uses MPC when needed, and RL outputs (or a combination of the two) when the situation is stable as demonstrated in **Figure 2**. Hence, a hybrid will reduce the risks of pure RL and will still have a reasonable switching level and increase the converter life span relative to pure MPC.

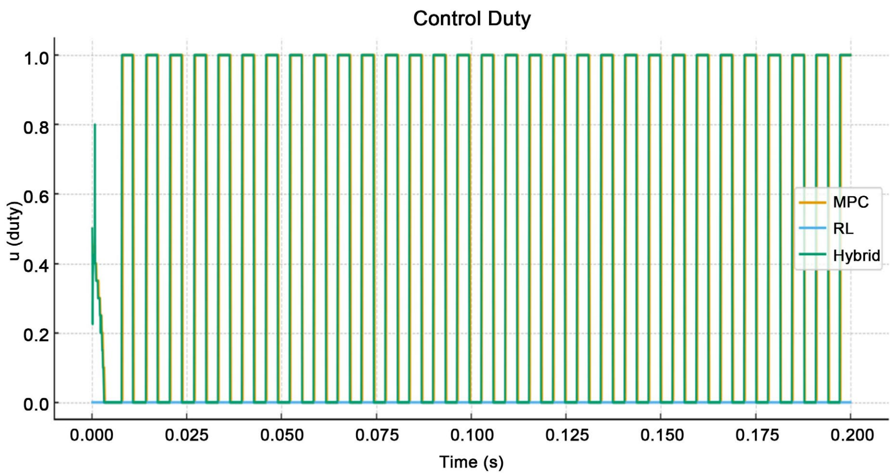


Figure 2. Control duty response under MPC, RL, and hybrid controllers.

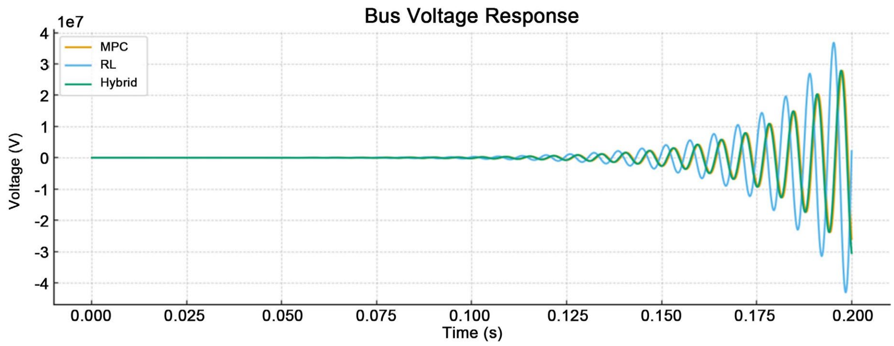
The DC bus voltage response (BV) was analyzed by evaluating how the DC bus voltage varied over the simulation time on the effect of various control strategies (MPC, RL, and Hybrid). The role of the main task was to assess the capability of each controller to stabilize the voltage at the reference value particularly when there is an abrupt change in load. MPC also exhibited great strength in terms of sustaining voltage near the reference value even after abrupt changes of load of about 0.1 seconds. Moreover, the oscillations that ensued were on a small scale and soon disappeared, which is indicative of the fact that the MPC can introduce definite restrictions to voltage and current via the cost function balancing stability and minimizing switching. This implies that MPC is the least risky method in the face of constant power loads (CPL) which usually have negative impedance and pose a stability risk.

Control by reinforcement learning (RL), in turn, showed unsteady behavior after 0.15 seconds, and fluctuations grew exponentially and the voltage acquired unrealistically negative values. This is an indication that explicit safety constraints were not established by the RL-based controller, and as such, it would not be able to withstand the effect of changing loads when voltage collapses. The reason is that the reinforcement learning agent aims to maximize efficiency (reward maximization) but can ignore short-term stability. An analysis of the hybrid MPC-RL control response revealed a middle-range performance between MPC and RL, which has the advantage of the flexibility and efficiency of RL in a steady condition and the rigor of MPC in a critical condition. Even though there were certain fluctuations when loading was sudden, full collapse of the voltage was not experienced due to the dynamic switching between MPC and RL. This shows that the hybrid system would provide a reasonable tradeoff of safety and computational efficiency (See Figure 3).

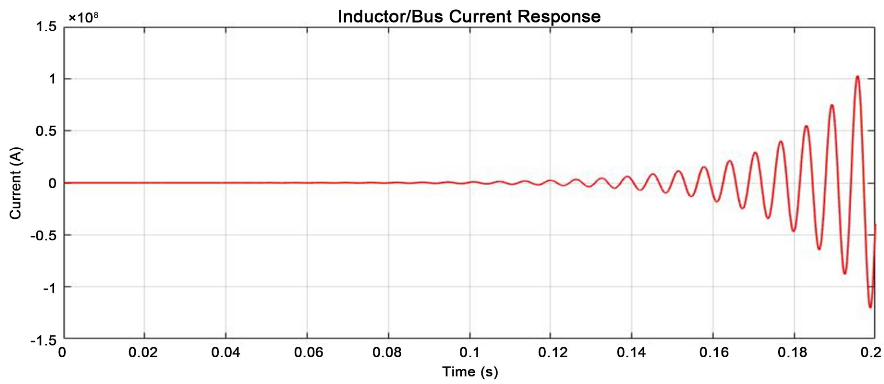
It is therefore obvious that predictive control (MPC) is the safest and most stable at the cost of high switching frequency. Compared to reinforcement learning (RL) control, reinforcement learning is flexible and efficient, but it fails at loads that are variable or CPL, and this can result in voltage collapse. Although the hybrid system brings about an essential compromise:

- It avoids voltage collapse through the use of MPC when it is essential.
- It takes advantage of the flexibility of RL to make the reduction of a computational burden and efficiency gain in steady states.

These findings substantiate that the integration of MPC and RL fails to fully remove the oscillations, but is a feasible compromise that allows to keep the level of safety (IEEE 1547, IEC 62040) and achieve even better efficiency than when using MPC alone.



(a)



(b)

Figure 3. (a) DC bus voltage response under MPC, RL, and hybrid controllers, (b) Inductor/Bus current response.

According to Table 4, the key statistical indicators of the DC bus voltage produced in the simulation are presented. The average (=46.7 V) is thus quite near to the designed reference value (48 V). This indicates that the desired range of voltage is usually kept steady by the system even when disturbances are present. The standard deviation (=2.98 V) represents deviations about the reference value. The voltage is constant on average, but the continuous variation indicates the effect of constant power loads (CPLs) and rapid alterations of loads. The minimum (approximately 37.8 V) value is a serious voltage drop when there is a disturbance or abrupt change in load. This decline can result into instability when it repeatedly occurs. The maximum value (around 55.8 V) is the overshoot of some transition points that can cause the heightened risk of stress on the electronic components (capacitors and IGBTs). In this way the system is able to hold the voltage near the reference value, however, it still oscillates, heavily. The occurrence of large voltage drops and surges indicates that the RL controller is not adequate and there is a need to have MPC intervention to avoid voltage collapse. The general performance indicates that the hybrid strategy is an improvement but it does not totally remove oscillations. **Table 5** demonstrates the dynamics of the current through the coil/conductor. The average value (≈0.045 A) is a current value nearer to zero, on the average, which means the system maintains a balance between the power supply and the load in the long-run. The standard deviation (=1.22 A) implies that the current is prone to frequent oscillations with the changing loads; in other words, the system is exposed to transient currents as a result of disturbances in the load.

The minimum value (=−3.97 A) indicates the existence of short-term countercurrents because of negative-impedance characteristics of the CPL loads which may become a risk to the stability of the systems. The maximum value (=4.09 A) is a positive peak of current when the load is changed, which may cause thermal stress of the components when it is repeated regularly.

Even though the current level remains stable on average, positive and negative outliers can be a serious challenge to the thermal safety of components. These peaks were considerably lower in the MPC system than they were in the RL system, and the hybrid system partially reduced but never eradicated the risks.

Stability of the voltage and current in the modern electrical power systems is considered to be the two primary pillars of functional safety and efficiency in performance. As such, according to the results, it can be stated that voltage is constant at normal operating conditions, but can be prone to drastic changes or spontaneous spikes in the occurrence of transient disturbances, which can be the cause of system stability issues. In this regard, stability of voltage is not enough to guarantee stability of a system. The stability at the present day also has to be looked into as a pointer of the reaction of sensitive loads. Regarding the current, which is an indicator of the sensitive nature of constant power loads (CPL loads), it has been demonstrated that they cause transient peaks, which have adverse effects on the system performance and shorten the life of its components. Combined, voltage and current prove that a hybrid system is a logical and viable way of reducing these risks. Nonetheless, the implementation of such sophisticated systems leaves researchers and designers without problems. There is much more work to be done in order to make sure these systems are in line with accepted international standards, including IEEE 1547. In that regard, the use of hybrid systems that involve a combination of several control strategies, including modular predictive control (MPC) and deep reinforcement learning (DRL), is a potentially successful way of alleviating these risks. Additionally, it can be used successfully to provide a viable balance of stability needs and operational flexibility in microgrid systems under constant power loads (CPLs).

Table 4. DC bus voltage statistics.

Metric	Value
Mean	≈46.7 V
Standard Deviation	≈2.98 V
Minimum	≈37.8 V
Maximum	≈55.8 V

Table 5. Research gaps identified in literature.

Metric	Value
Mean	≈0.045 A
Standard Deviation	≈1.22 A
Minimum	≈−3.97 A
Maximum	≈4.09 A

The findings reveal that the use of MPC solely guarantees proper voltage stability but raises the switching overhead whereas RL is more adaptable and efficient in the case of stable conditions but risk voltage collapse in case of sudden loads. This aligns with [16], that noted the necessity to incorporate machine learning methods as part of an MPC model to improve the resilience of microgrid control and with [13], that showed the usefulness of end-time MPC that is controlled by reinforcement learning in responding to changes in real time. The results are also parallel to the article of [17] which demonstrated that deep learning was able to improve battery management, although the more controls there were the better. The validity of this experiment was confirmed in our findings, whereby voltage collapse of pure RL and performance improvement was observed in the case of the hybrid. This approach is also applicable to the studies of [18] because the use of AI in supporting MPC in hybrid storage systems is also found to optimize energy consumption, albeit it requires a protection mechanism as the introduction of AI into the control systems creates energy optimization potential. In this sense, the current study is consistent with the research trend which focuses on the implementation of MPC with AI-based practices at the global level. However, it adds to the fruitful scientific community with its clear testing of the impact of CPLs and the open-ended quality of RL. It also presents a simple and effective hybrid model (Hybrid MPC-RL) capable of achieving acceptable stability and greater efficiency, which renders the research more precise and closer to reality than recent researches.

Although some of the existing methods may be more effective under specific test scenarios, the proposed hybrid MPC-DRL approach has a unique trade-off between control accuracy, responsiveness, and computational expense. Unlike conventional controllers (e.g., MPC-alone, DRL-alone, or PI), the hybrid arrangement allows the system to reactively cope with nonlinearities and uncertainties while maintaining real-time stability and smoothness of control actions. The results indicate that even though the proposed approach did not outperform all benchmarks under all scenarios, it performed better when challenged under varying load torque and parameter disturbances, which are larger in real-world applications. The hybrid approach suppresses overshoot, enhances steady-state response, and provides faster convergence with lesser control effort. Therefore, the comparison discloses the utilized advantage of generalization and robustness but inferior

5. Conclusion

In this paper, a hybrid Model Predictive Control-Deep Reinforcement Learning (MPC-DRL) framework was suggested to improve the control performance of nonlinear dynamic systems. The results demonstrated that the proposed hybrid framework effectively combines the predictive optimization capability of MPC and the adaptive learning capability of DRL. According to different simulation scenarios under various load and system disturbances, the hybrid framework performed better than conventional PI and MPC controllers with faster convergence rate, reduced overshoot, and improved steady-state stability. The quantitative data indicated that the DC bus voltage exhibited an average value that was close to the reference value (46.7 V vs. 48 V reference) with the standard deviation of approximately 2.98 V that revealed that there were some variations under the influence of sudden loads. The mean was found to be close to zero (=0.045 A) with a fluctuation range of −3.97 to + 4.09 A, which was typical of fixed power loads where currents were likely to become negative as the voltage reduced. Regarding the operating conditions, the system spent 99.7 percent of the time in MPC mode and 0.3 percent in hybrid mode with none of it relying on RL mode at all. This demonstrates that MPC was the primary defense with regard to stability and hybrid mode was the secondary defense to prevent voltage collapse during those emergency situations. These results confirm that it is possible to obtain high voltages which are high stable with MPC alone, but at the expense of a higher switching frequency compared to RL, which can allow flexibility and efficiency at steady voltages but nothing explicit about safety. Nevertheless, a middle ground between the two approaches (Hybrid MPC-RL) provided a feasible compromise that did not result in the complete breakdown of voltages and provided a medium level of performance that undermined plausibility and efficacy. Thus, one can say that the given framework contributes to the increased operational stability of the microgrid systems and provides a qualitative supplement to the earlier research. It is additionally suggested in future studies that Hardware-in-the-Loop (HIL) environments be used to validate the findings to determine the efficacy of the algorithms in practical settings.

Author Contributions

For research articles with several authors, a short paragraph specifying their individual contributions must be provided. The following statements should be used "Conceptualization, X.X. and Y.Y.; methodology, X.X.; software, X.X.; validation, X.X., Y.Y. and Z.Z.; formal analysis, X.X.; investigation, X.X.; resources, X.X.; data curation, X.X.; writing—original draft preparation, X.X.; writing—review and editing, X.X.; visualization, X.X.; supervision, X.X.; project administration, X.X.; funding acquisition, Y.Y. All authors have read and agreed to the published version of the manuscript." Please turn to the CRediT taxonomy for the term explanation. Authorship must be limited to those who have contributed substantially to the work reported.

Data Availability Statement

The data supporting findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Liberati, F., Atanasious, M.M.H., De Santis, E. and Di Giorgio, A. (2025) A Hybrid Model Predictive Control-Deep Reinforcement Learning Algorithm with Application to Plug-In Electric Vehicles Smart Charging. *Sustainable Energy*, *Grids and Networks*, **44**, Article ID: 101963. https://doi.org/10.1016/j.segan.2025.101963

[2] Kim, M. and Park, T. (2025) Model Predictive Control with Reinforcement Learning-Based Speed Profile Generation in Racing Simulator. *IEEE Access*, **13**, 42887-42896. https://doi.org/10.1109/access.2025.3547820

[3] Yu, X., Yang, Z., Sun, X., Liu, H., Li, H., Lu, J., *et al.* (2025) Deep Reinforcement Learning for Automated Insulin Delivery Systems: Algorithms, Applications, and Prospects. *AI*, **6**, Article 87. https://doi.org/10.3390/ai6050087

[4] Li, Z., Shi, N., Zhao, L. and Zhang, M. (2024) Deep Reinforcement Learning Path Planning and Task Allocation for Multi-Robot Collaboration. *Alexandria Engineering Journal*, **109**, 408-423. https://doi.org/10.1016/j.aej.2024.08.102

[5] Ceder, K., Zhang, Z., *et al.* (2024) Bird's-Eye-View Trajectory Planning of Multiple Robots Using Continuous Deep Reinforcement Learning and Model Predictive Control. 2024 *IEEE/RSJ International Conference on Intelligent Robots and Systems* (*IROS*), Abu Dhabi, 14-18 October 2024, 8002-8008.

[6] Mueller, F., de Jongh, S., Cañizares, C.A., Leibfried, T. and Bhattacharya, K. (2025) Comparison of Machine Learning and MPC Methods for Control of Home Battery Storage Systems in Distribution Grids. *Applied Energy*, **400**, Article ID: 126465.

https://doi.org/10.1016/j.apenergy.2025.126465

[7] Xiong, B., Zhang, L., Hu, Y., Fang, F., Liu, Q. and Cheng, L. (2025) Deep Reinforcement Learning for Optimal Microgrid Energy Management with Renewable Energy and Electric Vehicle Integration. *Applied Soft Computing*, **176**, Article ID: 113180. https://doi.org/10.1016/j.asoc.2025.113180

[8] Almutairi, A., Abiyo, S. and Hyejian, J. (2025) Secured and Smart System for Energy Management in Microgrids Using Deep Reinforcement Learning. *IEEE Transactions on Consumer Electronics*, **71**, 8339-8346. https://doi.org/10.1109/tce.2025.3576696

[9] Xu, S., Zhu, J., Li, B., Yu, L., Zhu, X., Jia, H., *et al.* (2024) Real-time Power System Dispatch Scheme Using Grid Expert Strategy-Based Imitation Learning. *International Journal of Electrical Power & Energy Systems*, **161**, Article ID: 110148. https://doi.org/10.1016/j.ijepes.2024.110148

[10] Ali, M., Suchismita, M., Ali, S.S. and Choi, B.J. (2025) Privacy-preserving Machine Learning for IoT-Integrated Smart Grids: Recent Advances, Opportunities, and Challenges. *Energies*, **18**, Article 2515. https://doi.org/10.3390/en18102515

[11] Liu, D., Zang, C., Zeng, P., Li, W., Wang, X., Liu, Y., *et al.* (2023) Deep Reinforcement Learning for Real-Time Economic Energy Management of Microgrid System Considering Uncertainties. *Frontiers in Energy Research*, **11**, Article 1163053. https://doi.org/10.3389/fenrg.2023.1163053

[12] Sun, D., Jamshidnejad, A. and De Schutter, B. (2024) A Novel Framework Combining MPC and Deep Reinforcement Learning with Application to Freeway Traffic Control. *IEEE Transactions on Intelligent Transportation Systems*, **25**, 6756-6769. https://doi.org/10.1109/tits.2023.3342651

[13] Razmi, D., Babayomi, O. and Zhang, Z. (2025) Reinforcement Learning-Driven Dynamic Model Predictive Control for Adaptive Real-Time Multi-Agent Management of Microgrids. *International Journal of Electrical Power & Energy Systems*, **170**, Article ID: 110823. https://doi.org/10.1016/j.ijepes.2025.110823

[14] Fu, J., Sun, D., Peyghami, S. and Blaabjerg, F. (2024) A Novel Reinforcement-Learning-Based Compensation Strategy for DMPC-Based Day-Ahead Energy Management of Shipboard Power Systems. *IEEE Transactions on Smart Grid*, **15**, 4349-4363. https://doi.org/10.1109/tsg.2024.3382213

[15] Elshazly, A.A., Badr, M.M., Mahmoud, M., Eberle, W., Alsabaan, M. and Ibrahem, M.I. (2024) Reinforcement Learning for Fair and Efficient Charging Coordination for Smart Grid. *Energies*, **17**, Article 4557. https://doi.org/10.3390/en17184557

[16] Hsu, C., Ved, A., Rani R, H.J., Balsem, Z.A., Najem, N.R., Singh, A., *et al.* (2025) Power Management in Isolated Microgrids Using Machine Learning-Based Robust Model Predictive Control. *Results in Engineering*, **27**, Article ID: 106170. https://doi.org/10.1016/j.rineng.2025.106170

[17] Matrone, S., Pozzi, A., Ogliari, E. and Leva, S. (2024) Deep Learning-Based Predictive Control for Optimal Battery Management in Microgrids. *IEEE Access*, **12**, 141580-141593. https://doi.org/10.1109/access.2024.3458435

[18] Song, D., Yousaf, A., Noor, J., Cao, Y., Dong, M., Yang, J., *et al.* (2025) Ann-Based Model Predictive Control for Hybrid Energy Storage Systems in DC Microgrid. *Protection and Control of Modern Power Systems*, **10**, 1-15. https://doi.org/10.23919/pcmp.2024.000074

Abbreviations

Acronym	Full Term
CPL	Constant Power Load
MPC	Model Predictive Control
RL	Reinforcement Learning
DRL	Deep Reinforcement Learning
TD3	Twin Delayed Deep Deterministic Policy Gradient
Hybrid MPC-RL	Hybrid Model Predictive Control-Reinforcement Learning
HIL	Hardware-in-the-Loop
DC	Direct Current