Renewable Energy, Environment, and Power Quality Journal (REE&PQJ)
ISSN:3020-531X
Volume 25, February 2026

Assessing deep reinforcement learning control of permanent magnet-assisted synchronous reluctance machines

Cristina Martín, M. A. González-Cagigal, Álvaro Rodríguez del Nozal and Juan M. Mauricio

Department of Electrical Engineering
E.T.S.I., Seville University
41092 Seville (Spain)
e-mail: cmartin15@us.es, mgcagigal@us.es, arnozal@us.es, jmmauricio@us.es

Abstract. **This paper presents an application of deep reinforcement learning (DRL) for controlling permanent magnet-assisted synchronous reluctance machines (PMA-SynRMs). A model-free DRL agent is trained to control the power converter switching states, aiming to accurately track current references. The DRL-based control scheme is compared against a traditional finite control set model predictive control (FCS-MPC) strategy employing a simplified linear model of the PMA-SynRM. Simulation results demonstrate that the DRL controller achieves superior performance in terms of tracking accuracy and harmonic distortion reduction, effectively handling the machine's inherent nonlinearities. Furthermore, the DRL agent exhibits robustness against measurement errors. The findings highlight the potential of DRL as a viable alternative to conventional model-based control methods for high-performance PMA-SynRM drives, offering improved adaptability, robustness, and operational flexibility.**

Key words. Artificial neural networks, deep reinforcement learning, electrical drives, model predictive control, permanent magnet-assisted synchronous reluctance machines.

1. Introduction

Electrical machines play an important role in numerous industrial and domestic applications, such as electric mobility, renewable energy production, industrial machinery, and HV-AC systems, among others. They constitute an important component of the electric power system that eases the integration of renewable energy sources and contributes to reducing the environmental impact. Among the various types of electrical machines, induction machines (IMs) and permanent magnet synchronous machines (PMSMs) have traditionally been the preferred choices due to their well-established control techniques and robust performance. However, synchronous reluctance machines (SynRMs), particularly in their permanent magnet-assisted version (PMA-SynRM), have recently gained attention as a viable alternative. These machines offer high efficiency, reliability, and a broad operational speed range while maintaining a lower cost due to the reduced dependence on expensive rare-earth magnets. As a result, PMA-SynRMs present an attractive solution for variable-speed applications [1].

To regulate the operation of these modern electrical drives, advanced control strategies are required. Finite control set model predictive control (FCS-MPC) has emerged as a strong competitor to conventional field-oriented control (FOC) techniques [2]. Unlike FOC, which relies on modulation strategies, FCS-MPC directly determines the optimal power converter state at each control interval by minimizing a predefined cost function that encapsulates the control objectives. To this end, the future state of the plant is predicted using a mathematical model of the system. This approach enables fast dynamic response and flexible control implementation. However, the effectiveness of FCS-MPC heavily depends on the accuracy of the system model employed for the predictions. In the case of PMA-SynRMs, the nonlinear flux-linkage to current characteristic introduce significant modeling challenges. Traditional control approaches often consider constant inductance models to reduce computational burden [3], which compromises control performance. Other proposals use high-dimensional current-flux maps derived from finite element analysis (FEA), where saturation and cross-magnetization effects are considered, or complex analytical models that partially reflect these nonlinearities [4], [5]. Those solutions improve the accuracy of the model at the expense of increased computational cost, making real-time implementation challenging.

In this context, deep reinforcement learning (DRL) offers a paradigm shift by leveraging artificial neural networks to approximate the motor control policy. By mapping system states, such as motor speed and currents, to optimal control actions, DRL-based controllers can learn complex patterns and relationships from training data without requiring an explicit mathematical model of the machine. Although the training process demands extensive data and computational resources, the online deployment of a trained DRL agent is significantly more efficient in terms of computational cost compared to the previously mentioned FEA-based methods.

This data-driven approach has been recently applied to the control of electrical drives, where most research focus on the PMSM case [6]-[9]. The main advantages that this approach provides include improved adaptability to nonlinearities and uncertainties, and greater tolerance to external disturbances. Additionally, DRL allows greater operational flexibility, as the objective function (reward function) in the training phase can be tailored to meet specific performance criteria, such as energy efficiency, torque ripple minimization, or thermal constraints [10].

This study aims to provide insights into the potential of DRL-based controllers as an alternative to traditional FCS-MPC approaches for high-performance PMA-SynRM applications. Thus, a model-free DRL-based control algorithm for PMA-SynRM constitutes the main contribution of the paper.

Additional contributions of the paper are summarized in the following points:

- Analysis of the influence of measurement errors in the overall performance of the implemented agent.
- Comparative analysis of the proposed DRL-based control scheme against FCS-MPC, in terms of accuracy in the desired output, assuming nonlinearities in the machine model.

The findings will contribute to the ongoing development of intelligent control strategies that enhance efficiency, robustness, and computational feasibility in next-generation electrical drive systems.

The remainder of this paper is organized as follows: Section 2 details the PMA-SynRM drive system model, highlighting the challenges associated with accurate modeling of magnetic nonlinearities. Section 3 describes the proposed DRL-based control scheme and the training process employed. Section 4 presents the simulation results, comparing the performance of the DRL controller against the benchmark FCS-MPC approach. A discussion on the impact of measurement errors is also included. Finally, Section 5 summarizes the key findings and concludes the paper, outlining potential future research directions.

2. PMA-SynRM based drive model

The system under study consists of a three-phase PMA-SynRM powered by a standard two-level three-phase voltage source converter (VSC). Key parameters and specifications for the machine are detailed in Table I.

The dynamic behavior of the PMA-SynRM is commonly modeled in the rotor *dq* reference frame. The stator voltage equations in this frame are:

$$v_d = R_s i_d + \frac{d\lambda_d}{dt} - p\Omega\lambda_q \tag{1}$$

$$v_q = R_s i_q + \frac{d\lambda_q}{dt} + p\Omega\lambda_d \tag{2}$$

where v_d, v_q, i_d, i_q, λ_d, and λ_q are the stator voltages, currents, and flux linkages in the *d*- and *q*-axes, respectively. R_s is the stator resistance, p denotes the number of pole pairs, and Ω is the rotor mechanical speed. The electromagnetic torque (T_{em}) generated is expressed as:

$$T_{em} = \frac{3}{2} p(\lambda_d i_q - \lambda_q i_d) \tag{3}$$

As highlighted in the introduction, accurately modeling the relationship between flux linkages (λ_d, λ_q) and currents (i_d, i_q) is a significant challenge for PMA-SynRMs. These machines exhibit pronounced nonlinear magnetic characteristics, including saturation and cross-saturation effects, which are inherent to their design and operation.

Despite this known complexity, conventional control approaches like FCS-MPC often rely on simplified models to remain computationally tractable for real-time implementation. For the FCS-MPC controller implemented and evaluated in this comparative study, a simplified linear magnetic model is deliberately employed. This model assumes constant values for the d-axis inductance (L_d) and q-axis inductance (L_q), relating flux linkages to currents as follows:

$$\lambda_d = L_d i_d + \lambda_{pm} \tag{4}$$

$$\lambda_q = L_q i_q \tag{5}$$

where L_d and L_q are treated as fixed parameters, and λ_{pm} represents the constant flux linkage from the permanent magnets.

While this simplification drastically reduces the modeling effort and computational requirements for the FCS-MPC predictions, it inherently neglects the significant magnetic nonlinearities present in the actual PMA-SynRM (see the flux-linkage maps in Fig. 1 obtained via FEA). This discrepancy between the simplified model used by the FCS-MPC and the real machine behavior is a primary source of suboptimal performance and potential inaccuracies in the conventional control scheme. The dependency on this simplified, and known-to-be-inaccurate, model is a key limitation motivating the investigation of model-free DRL, which does not require prior knowledge or simplification of the machine's magnetic characteristics. The comparison presented in this paper aims to assess how well DRL can perform relative to an FCS-MPC operating with these common model simplifications.

The mechanical dynamics of the PMA-SynRM are described by the equation:

$$J\frac{d\Omega}{dt} = T_{em} - T_L - B\Omega \tag{6}$$

where J is the total inertia, B is the viscous friction coefficient, and T_L is the applied load torque.

Finally, the VCS equation completes the system model. The stator *dq* voltages can be computed from the dc-link voltage (V_{dc}), the VSC switching states (S_i), and the transformation matrix (M) as follows:

$$\begin{bmatrix} v_d \\ v_q \end{bmatrix} = \frac{V_{dc}}{3} M \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} S_a \\ S_b \\ S_c \end{bmatrix} \tag{7}$$

$$M = \frac{2}{3}\begin{bmatrix} \cos\theta & \cos(\theta-\epsilon) & \cos(\theta-2\epsilon) \\ -\sin\theta & -\sin(\theta-\epsilon) & -\sin(\theta-2\epsilon) \end{bmatrix} \tag{8}$$

Table I. – Parameters of the PMA-SynRM drive.

Parameter		Value
Pole pair	p	2
Stator resistance	R_s (Ω)	0.197
Nominal *d* inductance	$L_{d,n}$ (mH)	22.27
Nominal *q* inductance	$L_{q,n}$ (mH)	3.24
dc-link voltage	V_{dc} (V)	310
Nominal speed	$\omega_{m,n}$ (rpm)	2500
Nominal current	I_n (A)	22
Maximum current	I_{max} (A)	44
Maximum Torque	T_{max} (N·m)	43

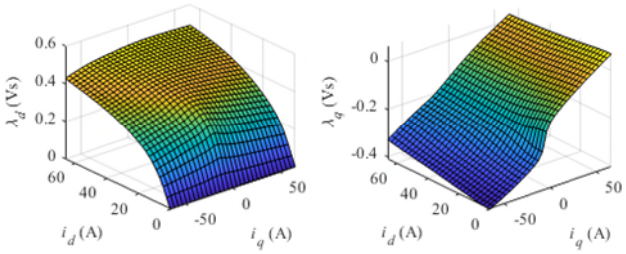

Fig. 1. Flux-linkage maps of the PMA-SynRM.

where $\boldsymbol{\theta}$ is the rotor angle and $\boldsymbol{\epsilon} = 2\pi/3$. Each component of the switching vector is a binary value that identifies the state of the VSC legs: S_i = 1 if the upper leg is ON, and S_i = 0 if it is OFF. Consequently, a set of 2^3 = 8 possible switching states and voltage vectors appear.

3. DRL-based control and training

Deep reinforcement learning is a subset of machine learning that combines deep learning and reinforcement learning to enable an agent to make sequential decisions in complex environments. In DRL, an agent learns by interacting with its environment, receiving feedback in the form of rewards, and optimizing its actions to maximize long-term performance. Thus, the agent learns a policy π, which maps the best action given a state of the environment [11]. The optimal policy π^* should provide actions that result in the highest expected cumulative reward $Q(s,a)$:

$$\pi^*(s) = \arg\max_{a} Q(s,a), \tag{9}$$

where s represents the state of the environment, a the action taken by the agent and $Q(s,a)$ represents the action-value function that estimates the total reward the agent can expect to collect over time by following a specific policy.

In this application, the Double Deep Q-Network (DDQN) algorithm is used to improve learning stability and decision-making accuracy [12]. To achieve this, DDQN employs two neural networks: an online network, which selects actions, and a target network, which evaluates them. This separation improves the accuracy of value estimates and avoids instability during training. The training process involves repeated experiences and periodic updates of the target network, which ensures more reliable learning. In this context, an agent interacts with a model of the VSC-PMA-SynRM system to achieve a specific goal: tracking a current reference. The agent follows a trial-and-error strategy, in which it selects control actions based on observations from the model. The effectiveness of each action is quantified by a reward signal, which provides information on whether the chosen action contributes to the tracking of the current reference. From this information, the neural network that defines the policy, that is, the one that estimates $Q(s,a)$, is updated, which improves future decision making and optimizes motor control. To stabilize the training, a duplicate $\tilde{Q}(s,a)$ is employed, which estimates the reward obtained by performing the action selected by the other neural network. Thus, during the training of both networks, the Q-learning algorithm follows an iterative approach to refine Q-values by incorporating observed rewards and estimated future discounted rewards. Its objective is to minimize the discrepancy between the predicted values and the actual rewards received, ultimately converging to an optimal policy. This process involves updating Q-values in a way that progressively reduces the Bellman error:

$$\begin{aligned} Q(s,a) &\leftarrow Q(s,a) \\ &+\alpha\left[r + \gamma\, \tilde{Q}\left(s', \arg\max_{a} Q(s',a')\right) - Q(s,a)\right], \end{aligned} \tag{10}$$

where α is the learning rate, r is the immediate reward obtained for taking action a in state s, and γ is the discount factor that weighs future rewards.

In this case, the system control involves selecting the optimal switching state of the VSC driving the PMA-SynRM every 50 microseconds. At each step, after the agent takes an action, it receives observations about the state of the motor, including rotor angle (θ), currents in dq coordinates (i_{dq}), and reference currents (i^*_{dq}). Finally, the reward function penalizes deviations between actual and reference currents, with logarithmic scaling to emphasize smaller errors and encourage precise tracking. During training, the agent decides between exploring new actions or using what it has learned. With a certain probability, it picks a random action to explore. Otherwise, it chooses the best action based on its current knowledge. In order to balance this exploration-exploitation of the network, a ε-greedy policy is employed. The agent has a probability ε to perform random actions. With a probability of $1-\varepsilon$, the agent selects the action that the Q-function suggests as optimal, prioritizing exploitation of learned knowledge. At the beginning of training, ε is set to a high value to encourage exploration and collect a diverse range of experiences. As training advances, ε is gradually decreased to favor the exploitation of learned knowledge from previous experiences.

To approximate both Q-functions, i.e. $Q(s,a)$ and $\tilde{Q}(s,a)$, a dense neural network with four layers is employed. The first three layers consist of 128, 64, and 64 neurons, respectively, each using a ReLU activation function to extract and refine features from the input state s. The final layer has 8 neurons with a linear activation function, corresponding to the 8 possible actions in the motor control problem. The network outputs 8 Q-values, each representing the expected cumulative reward for a specific action in the given state. The agent selects the optimal action by choosing the one with the highest Q-value.

Fig. 2 shows a schematic representation of the proposed controller. A PI-based external control loop regulates the rotor speed (Ω) and generates the reference electromechanical torque (T^*_{em}). Then, an optimal reference generator procures the dq-current references that are used as observations for the agent, together with the measured currents and rotor angle. This reference generator looks for the optimal dq-currents that provide the desired torque and minimizes the copper losses while respecting the maximum reachable peak values of currents and voltages of the system. The formulation of this optimization problem is:

$$\begin{aligned} &\min R_s\left(i_d^2 + i_q^2\right) \\ \text{s.t.}\quad &\max(i_{abc}) \le I_{max} \\ &\max(v_{abc}) \le \frac{V_{dc}}{2} \\ &T_{em} = T^*_{em} \\ &\text{Eqs. (1)-(3)} \end{aligned} \tag{11}$$

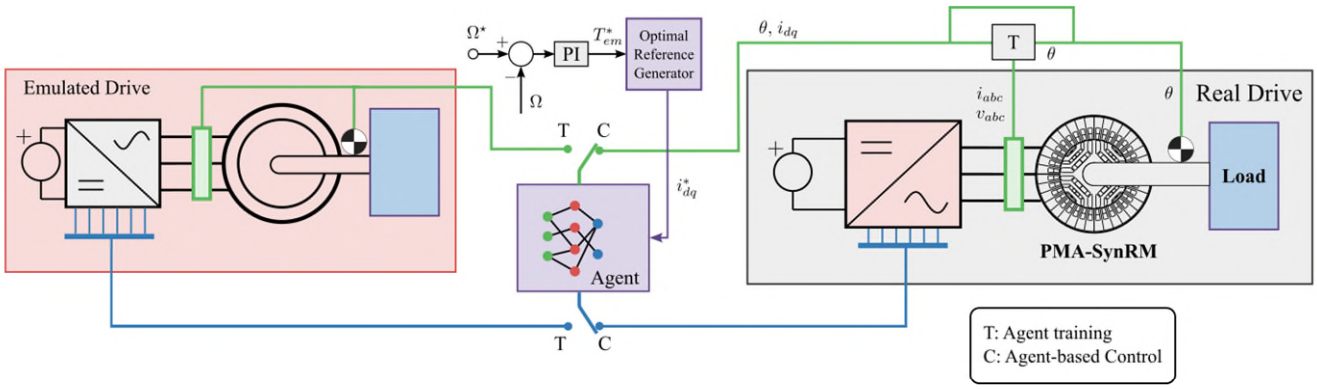

Fig. 2. Proposed DLR-based control scheme.

This problem is solved off-line providing a look-up table (LUT) with the optimal *dq*-current references for each pair of speed and torque references values. That LUT is employed in the control process. Finally, the DRL agent computes the control actions to be released to the converter based on references and measured system variables. As can be seen, the DRL agent replaces the model-based inner current control loop in the conventional FCS-MPC algorithm. To train the agent, the switch of Fig. 2 is fixed in position "T" where it interacts with an emulated machine with a wide range of speed references.

4. Results

The effectiveness of the proposed approach has been validated through simulations in a Python-based testing platform. In order to simulate the differences between the electrical drive used in the training process and the real one, artificial error has been added to the current measurements. In the base case, 1 % Gaussian error is considered.

As mentioned in the previous section, the training process encompasses a wide range of operating points, varying the load torque and the reference speed for the electrical drive. The results of a sample test scenario are included in Fig. 3, where dq currents and rotor speed are represented together with the corresponding reference values. The start-up period has not been included in the graph in order to exclusively evaluate the steady-state performance. Additionally, Fig. 4 depicts a 50-ms sample in the abc domain. It can be noted the good performance of the DRL-based control scheme, being the machine currents effectively regulated to their reference values. Similar conclusions are obtained when the operating point is changed.

A comparative assessment between the proposed DRL implementation and a FCS-MPC approach, where a linear model is considered assuming constant values for the dq inductances of the machine, has been also conducted. The following representative metrics are considered for this analysis:

- Root mean square error for the dq currents, $RMSe_{dq}$, in order to assess the tracking performance.
- Total harmonic disorder, THD, for the abc currents. These results are summarized in Table 1 for different operating conditions, together with the previously mentioned errors.

In light of the results presented in Table 2, it can be concluded that the proposed implementation of DRL-based control schemes outperforms the FCS-MPC approach in all the operating points considered. Regarding the tracking accuracy and the THD, these results can be explained given that the agent in the DRL algorithm has been trained using a nonlinear model of the electrical drive, whereas the MPC formulation assumes an average linear model for the machine.

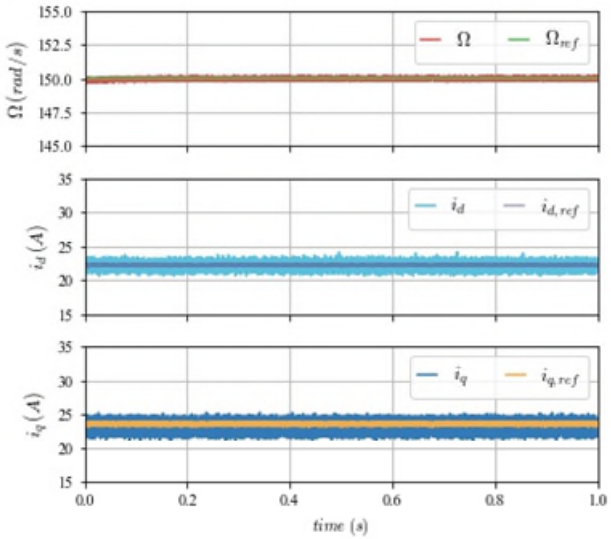

Fig. 3. Response of the DRL-based controller for the operating point $\Omega_{ref} = 150$ rad/s and $T_L = 30$ N·m.

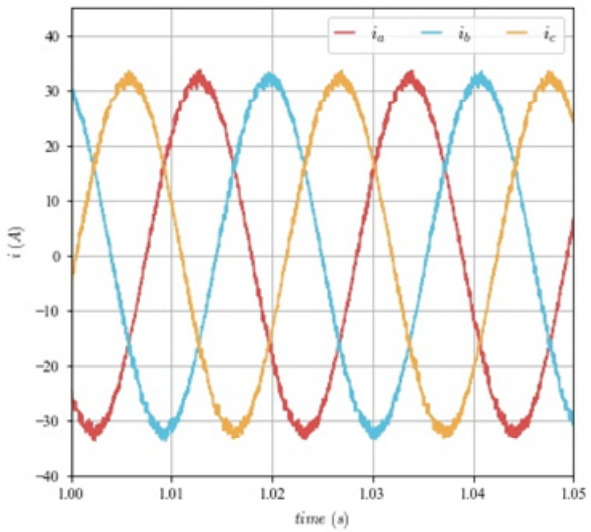

Fig. 4. Sample of the machine currents in the abc domain

Table 2. Performance metrics in the comparison of the DRL-based control scheme and the FCS-MPC approach.

		DRL-based control		FCS-MPC	
Ω_{ref} (rad/s)	T_L (N·m)	RMS_{dq} (A)	THD (%)	RMS_{dq} (A)	THD (%)
100	10	0.394	4.242	0.449	4.781
	30	0.558	2.661	0.685	2.933
150	10	0.364	3.962	0.425	4.490
	30	0.589	2.492	0.697	2.737
200	10	0.377	3.873	0.432	4.155
	30	0.579	2.543	0.729	2.648

Table 3. Variation of the $RMSe_{dq}$, in A, under different levels of measurement errors.

		Measurement error			
Ω_{ref} (rad/s)	T_L (N·m)	0.5 %	1 %	2 %	3 %
100	10	0.386	0.394	0.426	0.804
	30	0.521	0.558	0.659	0.743
150	10	0.360	0.364	0.574	0.637
	30	0.551	0.589	0.689	0.790
200	10	0.375	0.377	0.407	0.672
	30	0.571	0.579	0.687	0.912

Once the effectiveness of the DRL-based technique has been successfully assessed in the base case, Table 3 includes how the performance of the control scheme deteriorates when higher measurement errors are considered. The metric $RMSe_{dq}$ has been considered for this analysis.

It can be noticed that even for higher measurements errors, the $RMSe_{dq}$ does not increase excessively, giving evidence of the robustness of the DRL-based control under noisy observations.

Finally, the dynamic performance of the DRL-based control is assessed under varying operating conditions. For this purpose, Figs. 5 and 6 show the control performance when the rotor speed varies from 100 to 200 rad/s under a constant load torque, and when the load torque changes from 10 to 20 N·m at a constant rotor speed, respectively. The controller effectively regulates both the rotor speed and the stator currents in the dq refence frame (i_d and i_q) in both tests. The shadowed zone in Fig. 5 marks the system start-up period.

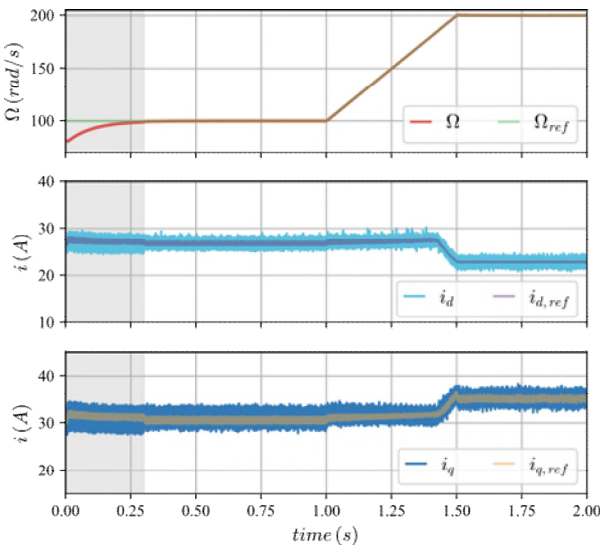

Fig. 5. Dynamic response of the DRL-based controller when rotor speed is varied from 100 to 200 rad/s at a constant load.

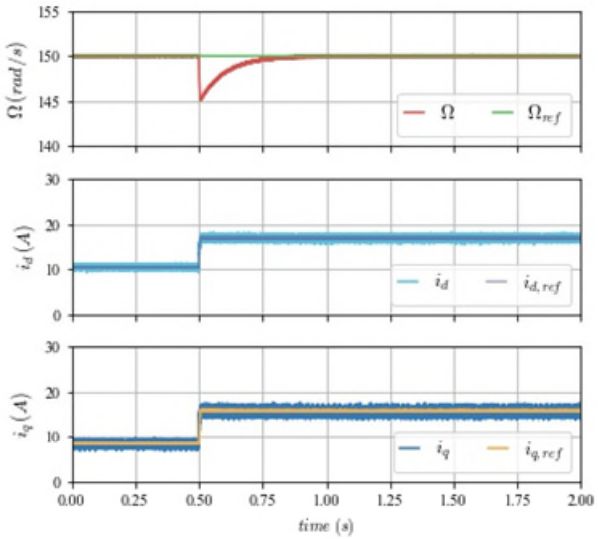

Fig. 6. Dynamic response of the DRL-based controller when load torque is varied from 10 to 20 N·m at a constant speed.

It can be concluded from the presented results that the proposed implementation of the DRL-based technique has adequate performance with time-varying operating conditions.

5. Conclusion

The findings of this study demonstrate the potential of DRL as a high-performance control strategy for PMA-SynRMs. The DRL-based controller consistently outperformed the benchmark FCS-MPC strategy across various operating conditions, achieving superior tracking accuracy and lower harmonic distortion in the motor currents. This enhanced performance can be attributed to the DRL agent's ability to learn and adapt to the complex nonlinearities inherent in the magnetic characteristics of the PMA-SynRM, a feature that conventional model-based approaches like FCS-MPC, which often rely on simplified linear models, struggle to capture effectively.

The robustness of the DRL controller was further validated under the presence of measurement errors. Even with artificially introduced noise in the current measurements, the DRL agent maintained an adequate level of performance, exhibiting only a moderate degradation in tracking accuracy. This robustness is crucial for real-world applications where sensor noise is inevitable. Furthermore, the dynamic performance evaluation demonstrated the controller ability to effectively track varying speed references while maintaining tight control over the stator currents. This adaptability to changing operating conditions highlights the potential of DRL for applications requiring agile and responsive motor control.

While the training process for DRL agents is computationally intensive and requires significant data collection, the online deployment of the trained agent is computationally efficient. This characteristic makes DRL attractive for applications where online computational resources are limited. This research contributes to the ongoing exploration of intelligent control strategies for next-generation electrical drive systems, in the way for more efficient, robust, and adaptable motor control

solutions in various applications, including electric vehicles and industrial automation. Future research will focus on exploring different DRL algorithms and network architectures to further optimize performance and reduce training times. Furthermore, experimental validation on a physical PMA-SynRM drive system will be conducted to validate the simulation results and demonstrate the real-world applicability of the proposed DRL-based control strategy.

Acknowledgement

Grant PID2021-127835OB-I00 funded by MICIU/AEI/ 10.13039/501100011033 and by "ERDF/EU".

References

[1] M. D. Nardo et al., "Permanent Magnet Assisted Synchronous Reluctance Machine Design for Light Traction Applications", IEEE Transactions on Industry Applications (2024), Vol. 60, no. 4, pp. 6079-6091.
[2] S. Vazquez, J. Rodriguez, M. Rivera, L. G. Franquelo and M. Norambuena, "Model Predictive Control for Power Converters and Drives: Advances and Trends," IEEE Transactions on Industrial Electronics (2017), Vol. 64, no. 2, pp. 935-947.
[3] Li, Z., Su, J., Gao, H., Zhang, E., Kuang, X., Li, C., Bi, G., and Xu, D. "Sensorless Control of PMaSynRM Based on Hybrid Active Flux Observer", Electronics (2025), Vol. 14, no. 2, 259.
[4] M.A. González-Cagigal, C. Martín, M. Bermúdez, P. Cruz-Romero, "Comparison of nonlinear Kalman filtering schemes for sensorless control of permanent magnet-assisted synchronous reluctance machines", International Journal of Electrical Power & Energy Systems (2025), Vol. 165.
[5] K. Tan, J. Su, B. Zhong and G. Yang, "A Computationally Efficient Full-Speed Domain Control Method for PMaSynRM Considering Magnetic Saturation", IEEE Transactions on Power Electronics (2025), early acess.
[6] M. Schenke, W. Kirchgässner and O. Wallscheid, "Controller Design for Electrical Drives by Deep Reinforcement Learning: A Proof of Concept" IEEE Transactions on Industrial Informatics (2020), Vol. 16, no. 7, pp. 4650-4658.
[7] S. Bhattacharjee, S. Halder, Y. Yan, A. Balamurali, L. V. Iyer and N. C. Kar, "Real-Time SIL Validation of a Novel PMSM Control Based on Deep Deterministic Policy Gradient Scheme for Electrified Vehicles", IEEE Transactions on Power Electronics (2022), Vol. 37, no. 8, pp. 9000-9011.
[8] G. Book *et al*., "Transferring Online Reinforcement Learning for Electric Motor Control From Simulation to Real-World Experiments", IEEE Open Journal of Power Electronics (2021), Vol. 2, pp. 187-201.
[9] M. Schenke, B. Haucke-Korber and O. Wallscheid, "Finite-Set Direct Torque Control via Edge-Computing-Assisted Safe Reinforcement Learning for a Permanent-Magnet Synchronous Motor," IEEE Transactions on Power Electronics (2023), Vol. 38, no. 11, pp. 13741-13756.
[10] Yiming Zhang, Jingxiang Li, Hao Zhou, Chin-Boon Chng, Chee-Kong Chui, Shengdun Zhao, "Comprehensive evaluation of deep reinforcement learning for permanent magnet synchronous motor current tracking and speed control applications", Engineering Applications of Artificial Intelligence (2025), Vol. 149.
[11] Sutton, R. S. & Barto, A. G. (2018). *Reinforcement Learning: An Introduction* (2nd ed.). MIT Press.
[12] Hessel, M., Modayil, J., Van Hasselt, H., Schaul, T., Silver, D., & Sutton, R. (2017). Rainbow: Combining Improvements in Deep Reinforcement Learning. In Advances in Neural Information Processing Systems (NeurIPS), 30.