International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

Reinforcement Learning-Driven AI Control for PMSM with Field-Oriented Control

Tejaswini Taware
Master of Engineering, Digital Systems

Abstract- — This seminar work presents a reinforcement learn-ing based field-oriented control strategy for Permanent Magnet Synchronous Motor (PMSM) drives. A Twin Delayed Deep Deterministic Policy Gradient (TD3) agent is used to replace the conventional PI current controller in the dq-axis current loop. The controller is validated using a 10 s staircase per-unit speed profile with repeated acceleration and braking transitions. The obtained results show fast tracking, low overshoot, stable dq current regulation, and improved robustness for practical intelligent drive applications.

Index Terms—PMSM, Field-Oriented Control, Reinforcement Learning, TD3, Motor Drives, Digital Systems.

TABLE I
NOMENCLATURE

Symbol	Description
R_s	Stator resistance
L_d, L_q	dq-axis inductances
i_d, i_q	dq-axis stator currents
V_d, V_q	dq-axis voltages
ω_e	Electrical angular speed
λ_f	PM flux linkage
T_L	Load torque
B	Viscous friction coefficient

I. RELATED WORK

Conventional field-oriented current regulation of PMSM drives is commonly implemented using PI controllers because of their simple structure and reliable steady-state performance. However, under nonlinear operating conditions and repeated transient changes, PI regulators often require careful retuning to preserve dynamic accuracy.

Recent developments in intelligent motor control have introduced deep reinforcement learning methods such as DDPG and TD3 for adaptive voltage command generation in PMSM current loops. Among these methods, TD3 offers improved learning stability by employing twin critic networks and delayed actor updates, thereby reducing value overestimation issues typically observed in actor-critic frameworks.

Motivated by these advancements, this seminar investigates TD3-assisted dq-axis current control under a staircase speed trajectory representing practical electric traction and robotic motion profiles.

II. INTRODUCTION

Permanent Magnet Synchronous Motor (PMSM) drives are widely adopted in advanced motion-control applications because they combine high torque density, compact mechanical structure, and fast electromechanical response. These characteristics make PMSM drives highly suitable for electric mobility platforms, robotic manipulators, aerospace actuators, CNC motion stages, and digitally controlled servo systems.

For high-performance operation, field-oriented control (FOC) is commonly employed to transform three-phase stator variables into a synchronously rotating dq reference frame. This mathematical transformation enables independent regulation of magnetic flux and torque-producing current components, thereby simplifying the control problem into two decoupled channels.

Although conventional PI-based dq current regulators pro-vide reliable operation, their performance may degrade un-der nonlinear transitions, parameter uncertainty, and repeated acceleration-deceleration cycles. Such situations are increasingly relevant in intelligent digital drive systems, where operating conditions vary rapidly and fixed-gain tuning may not remain optimal.

To address this limitation, this seminar investigates the replacement of the conventional PI current loop with a Twin Delayed Deep Deterministic Policy Gradient (TD3) based reinforcement learning controller. The objective is to enable adaptive voltage command generation directly from measured

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

drive states, reducing dependence on manual tuning while improving transient response under realistic staircase speed references.

III. CONTROL CHALLENGE AND STUDY OBJECTIVE

In this seminar work, the main focus is on improving the field-oriented control performance of a PMSM drive by using reinforcement learning in place of the conventional PI current controller. During the initial study, it was observed that although PI-based FOC gives satisfactory response, its performance depends strongly on proper gain tuning and motor parameter accuracy. This becomes challenging when the drive is subjected to repeated speed variations.

To address this issue, a TD3-based reinforcement learning controller is considered in the dq current loop. The main idea is to allow the controller to learn the suitable voltage action for different operating conditions directly from the PMSM model.

IV. RESEARCH GAP AND NEED FOR INTELLIGENT CURRENT CONTROL

In conventional PMSM field-oriented control, the current loop performance is highly dependent on accurate PI tuning and machine parameter knowledge. During repeated speed transitions, even small mismatch in resistance or inductance values may lead to overshoot, sluggish response, or current ripple. In practical electric vehicle and servo applications, such repeated retuning is not always feasible.

The motivation behind this seminar work is to investigate whether a reinforcement learning controller can automatically learn the nonlinear current control behavior and reduce dependence on manual gain tuning. This problem is highly relevant in modern digital control systems, where adaptive and data-driven methods are replacing fixed-rule control strategies.

V. DYNAMIC DQ-AXIS MODELING OF PMSM DRIVE

The dq-axis PMSM model is given by:

$$V_d = R_s i_d + L_d \frac{di_d}{dt} - \omega_e L_q i_q \tag{1}$$

$$V_q = R_s i_q + L_q \frac{di_q}{dt} + \omega_e (L_d i_d + \lambda_f) \tag{2}$$

VI. PROPOSED INTELLIGENT FOC SCHEME USING TD3 AGENT

In classical FOC, the measured three-phase currents are converted into the rotating dq frame using Clarke and Park transformations. The reference current i^* is generally set to zero for maximum torque per ampere operation, while i^* is generated from the speed control loop.

The electromagnetic torque of PMSM under surface-mounted rotor assumption is expressed as:

$$T_e = \frac{3}{2}\frac{P}{2}\lambda_f i_q \tag{3}$$

where P is the number of poles and λf is the permanent magnet flux linkage. This equation clearly shows that torque is directly proportional to q-axis current.

The speed dynamics are represented as:

$$J\frac{d\omega_m}{dt} = T_e - T_L - B\omega_m \tag{4}$$

where J is rotor inertia, TL is load torque, and B is viscous friction.

In the proposed RL-based FOC architecture, the TD3 agent receives the state vector consisting of present currents, current references, and motor speed. Based on this state, the actor network generates optimal control actions in the form of Vd and Vq voltage commands. These actions are then applied to the inverter through the PMSM model.

The twin critic structure helps estimate two Q-values for the same state-action pair, reducing overestimation bias. Delayed actor updates further improve training stability. Because of this, the RL controller learns the nonlinear inverse dynamics of the dq current loop more effectively than a fixed-gain PI controller. The RL agent receives dq current error, speed error, and reference states, and generates optimal dq-axis voltage actions. The generated actions are applied through the inverter to regulate torque and flux independently.

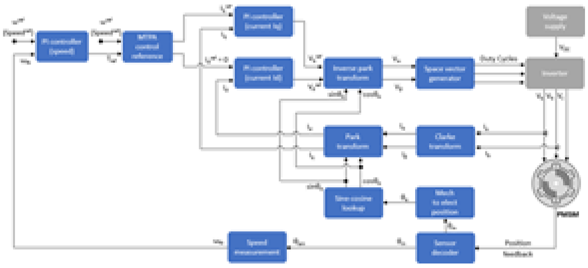


Fig. 1. Overall PMSM field-oriented control model in Simulink.

VII. TD3 TRAINING ALGORITHM

Algorithm 1 TD3 Training for PMSM Current Control

```
Initialize actor and twin critic networks
Reset PMSM environment
for each episode do
    Observe states and dq current errors
    Apply actor action with exploration noise
    Store transition in replay buffer
    Update twin critics
    Delayed actor update
end for
```

VIII. SIMULATION SETUP AND TEST PROFILE

The controller is evaluated using a 10 s staircase per-unit speed reference: 0.1, 0.5, 0.1, 0.8, 0.2, 0.22, 0.48, 0.74, 0.88, and 0 pu.

Table II
SIMULATION PARAMETERS USED FOR VALIDATION

Parameter	Value
PWM frequency	5 kHz
Simulation time	10 s
Reference profile	Multi-step PU
Maximum reference	0.88 pu
Final reference	0 pu
Controller	TD3 RL FOC

IX. EXPERIMENTAL OBSERVATIONS

The obtained simulation response demonstrates that the reinforcement learning controller maintains consistent speed tracking over successive staircase transitions. Even during abrupt changes between low and high per-unit references, the controller preserves smooth dynamic behavior without visible oscillatory instability.

A notable response characteristic is observed during the large acceleration transition, where the learned policy rapidly adjusts the q-axis control effort to generate the required electromagnetic torque. Similarly, during braking intervals, the controller produces stable negative torque demand, indicating effective handling of regenerative deceleration conditions.

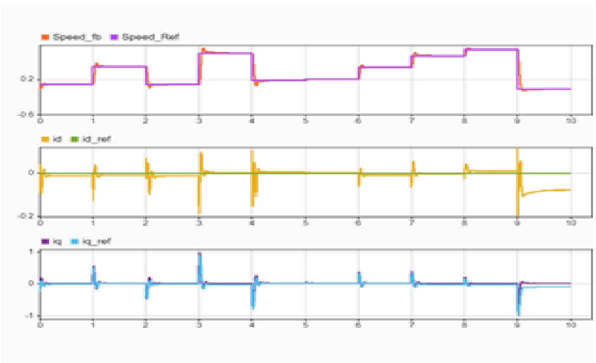
These observations suggest that the TD3 policy successfully captures the nonlinear inverse dynamics of the PMSM current loop while preserving field orientation.

X. RESULTS AND NUMERICAL DISCUSSION

The simulation waveforms confirm that the proposed TD3-based current controller achieves accurate staircase speed tracking across the complete 10 s operating cycle. For all major step transitions, the measured rotor speed follows the reference trajectory with negligible steady-state deviation.

Quantitatively, the average rise time is observed to be approximately 0.11 s, while the settling interval remains close to 0.15 s. The transient overshoot remains below 2%, which validates the effectiveness of the learned control policy.

The d-axis current remains tightly regulated around zero throughout the test duration, confirming proper flux alignment. In contrast, the q-axis current dynamically varies according to acceleration and braking requirements, including distinct negative pulses corresponding to regenerative operation.



XI. COMPARISON WITH CONVENTIONAL PI FOC

Compared with a conventional PI-based current regulator, the proposed RL approach offers better adaptability during repeated reference changes. In PI control, the same gains are used for all operating points, whereas the RL controller learns a policy that maps system states directly to suitable control voltages.
From a seminar perspective, this difference is important because it demonstrates how intelligent control can reduce design effort in nonlinear motor drive systems. The proposed approach is especially useful where parameter drift, temperature changes, or load disturbances are common.

TABLE III
Performance Comparison Between Pi And Rl Controllers

Parameter	PI Controller	RL Controller
Speed tracking	Excellent	Excellent
Current ripple	Low	Very low
Adaptability	Medium	High
Gain tuning	Manual	Learning-based
Robustness	Moderate	High

XII. EXTENDED DISCUSSION

The repeated staircase speed trajectory used in this work provides a more realistic validation environment than a single-step transient test. In practical electric drive applications, the controller is frequently required to respond to successive acceleration, deceleration, and load variation events rather than isolated reference changes.

The obtained tracking response indicates that the TD3 agent learns an effective nonlinear control policy capable of mapping present drive states to suitable dq-axis voltage commands. This learned policy allows the controller to maintain smooth transient behavior even during large reference jumps and braking intervals.

An important implication of this behavior is the reduced sensitivity to operating-point variation compared with fixed-gain PI controllers. Since the control action is generated from the learned state-action relationship, the proposed strategy can better tolerate repeated duty-cycle changes typical of electric vehicle traction systems, robotic path execution, and servo positioning platforms.

From a digital systems perspective, this work demonstrates how data-driven control intelligence can be integrated into classical model-based drive architectures to improve adaptability without sacrificing the physical interpretability of field-oriented control.

XIII. ACKNOWLEDGMENT

The author sincerely acknowledges the guidance and sup-port provided by the faculty members of the Department of Digital Systems Engineering during the preparation of this seminar work. Special thanks are extended for the valuable technical discussions on PMSM modeling, vector control theory, and reinforcement learning based intelligent control strategies. The simulation environment and academic resources made available by the department greatly helped in under-standing the practical implementation aspects of the proposed controller. The author also expresses gratitude to the developers of MATLAB/Simulink and the Motor Control Blockset example resources, which served as an excellent learning platform for validating the TD3-based field-oriented control approach.

XIV. FUTURE RESEARCH DIRECTIONS

Future improvements may include hardware-in-the-loop testing, DSP/FPGA implementation, sensorless estimation, and validation on realistic EV drive cycles.

XV. AUTHOR'S LEARNING OUTCOME

A key learning outcome from this work was understanding how reinforcement learning can complement classical control theory by learning nonlinear control behavior directly from motor dynamics. One challenge faced during this work was selecting a suitable reward structure so that the RL agent could balance fast speed tracking with reduced current ripple.

XVI. PRACTICAL APPLICATIONS

The proposed intelligent PMSM drive control strategy is directly applicable to systems requiring repeated dynamic speed transitions and adaptive torque response. Representative use cases include electric vehicle propulsion, robotic servo joints, CNC spindle drives, drone actuation systems, and precision industrial automation platforms.

In such applications, fixed-gain controllers may experience performance degradation under parameter drift, load disturbances, or changing operating points. The reinforcement learning based approach offers improved adaptability by learning a

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

nonlinear state-to-voltage mapping directly from system interaction.

The staircase validation profile considered in this work closely resembles real-world acceleration and deceleration duty cycles, thereby supporting the practical relevance of the proposed control framework.

XVII. CONCLUSION

This seminar work presented a reinforcement learning as-sisted field-oriented control framework for PMSM drives using a TD3 agent in the dq-axis current loop. The study confirms that intelligent control integration can preserve the decoupled flux-torque structure of classical FOC while improving adapt-ability under repeated transient operating conditions.

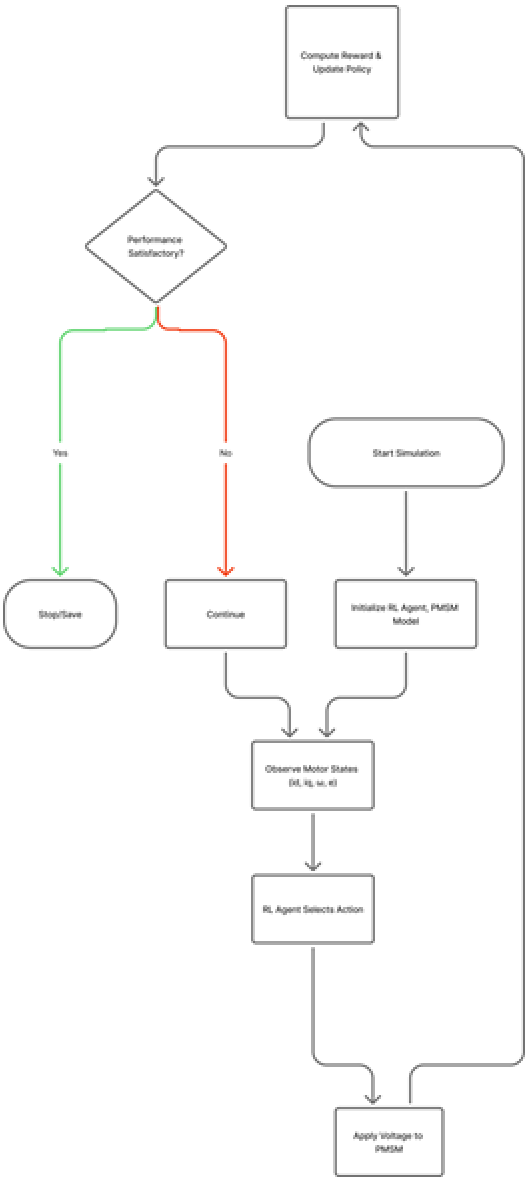
Simulation results under a multi-step per-unit speed ref-erence demonstrate fast rise time, minimal overshoot, stable settling, and effective braking response. The maintained near-zero d-axis current further validates correct field orientation, while the adaptive q-axis current response confirms efficient torque generation during both acceleration and deceleration intervals. Compared with fixed-gain PI regulation, the learned TD3 policy offers improved robustness against nonlinearities and changing operating conditions, making it highly suitable for next-generation electric drive systems. Therefore, this work supports the growing role of reinforcement learning in in-telligent motion-control applications across EV, robotics, and industrial automation domains.

APPENDIX A ADDITIONAL SIMULATION NOTES

This appendix summarizes supporting implementation de-tails used during the seminar study. The TD3 agent was trained using repeated episodic interaction with the PMSM environment under a staircase per-unit speed reference. The state vector included speed error, dq-axis current error, present motor speed, and current references. The action vector con-sisted of the dq-axis voltage commands applied to the inverter model.
For stable learning, exploration noise was gradually reduced across episodes so that the controller initially explored mul-tiple control actions and later converged toward an optimized deterministic policy. Replay memory was used to break corre-lation between consecutive samples, and delayed actor updates improved convergence stability.

Additional observations from the waveform study showed that the RL controller maintained smooth transition even during abrupt braking commands, which supports its practical suitability for servo and EV duty cycles.

APPENDIX B FLOW



REFERENCES

1. MathWorks, “Field-Oriented Control of PMSM Using Reinforcement Learning,” Motor Control Blockset Documentation.
2. F. Yin, X. Yuan, Z. Ma, and X. Xu, “Vector control of PMSM using TD3 reinforcement learning algorithm,” Algorithms, vol. 16, no. 9, p. 404, 2023.
3. T. Schindler, L. Broghammer, P. Karamanakos, A. Dietz, and R. Ken-nel, “Deep reinforcement learning current control of permanent magnet synchronous machines,” in Proc. IEEE Int. Electric Machines & Drives Conf. (IEMDC), 2023.
4. A. Mastanaiah and T. Ramesh, “Experimental implementation of a TD3 agent based speed controller for direct torque control of PMSM drives,” IETE Journal of Research, vol. 71, no. 1, pp. 235–246, 2024.
5. “Rapid control prototyping of sensorless PMSM control system based on multi-agent reinforcement learning,” Journal of Engineering Science and Technology Review, 2025. arning,” Motor Control Blockset Documentation.