Speed Control of IM Using RL-Based TD3 Agent

Ugur Ufuk Korpe
Dept. of Electrical-Electronics Engineering
Kirsehir Ahi Evran University
Kirsehir, Turkey
ugur.korpe@ahievran.edu.tr

Mustafa Gokdag
Dept. of Electrical-Electronics Eng.
Industrial Electronics Lab. (KBU-IEL)
Karabuk University
Karabuk, Turkey
mgokdag@karabuk.edu.tr

Ozan Gulbudak
Dept. of Electrical-Electronics Eng.
Industrial Electronics Lab. (KBU-IEL)
Karabuk University
Karabuk, Turkey
ozangulbudak@karabuk.edu.tr

2024 6th Global Power, Energy and Communication Conference (GPECOM) | 979-8-3503-5108-8/24/$31.00 ©2024 IEEE | DOI: 10.1109/GPECOM61896.2024.10582561

***Abstract*— Induction machines (IM) have been used daily since the 19th century. The induction machine should be effectively controlled to achieve high performance. Since the late 20th century and 21st century, field-oriented control (FOC), direct torque control (DTC), and model predictive control (MPC) techniques have been used in high-performance control applications. However, these techniques are dependent on the motor parameters and inverter model. These parameters change non-linearly depending on magnetic saturation, temperature, and operating point which negatively affects the performance of the system. In order to eliminate the negative effects of parameter changes, reinforcement learning (RL)-based methods have become increasingly popular in the literature in recent years. In this study, for the first time, the speed control of IM is performed using TD3 agent, which is one of the RL-based methods. The dynamic and steady-state performance of the control system designed with the TD3 agent is compared with the traditional FOC technique. Extensive simulation results have shown the robustness of the proposed drive system.**

Keywords—Induction motor, model predictive control, reinforcement learning, TD3 agent

I. INTRODUCTION

IMs are very popular in electric vehicles like Tesla Model S and Audi E-tron S, and industrial and household applications because of their high reliability due to the absence of magnets in their rotors, low maintenance costs, and cheaper compared to the permanent magnet synchronous machines (PMSMs) [1], [2]. The machine must have high efficiency and a wide operating range in previously mentioned applications. Therefore, the control of the machine must be performed with modern control techniques such as FOC, DTC, and MPC [3].

FOC and DTC were discovered by F. Blaschke and Takahashi&Noguchi in 1972 and 1986, respectively [4], [5]. In FOC, the torque and flux components of the machine are controlled indirectly using *dq*- axis currents. However, in DTC, the torque and flux components of the machine are controlled directly by estimating the torque and flux produced by the machine using an estimator [6], [7]. The advantages of the FOC are that the system has a constant switching frequency and good steady-state response due to modulators such as sine pulse width modulation (SPWM) and space-vector PWM (SVPWM). The advantages of DTC however, this technique is easier to apply than FOC and the dynamic response of the system is faster due to direct control of the torque and the flux of the machine [8], [9]. However, the disadvantages of the FOC are that the system has a slow dynamic response due to the indirect control of the torque and the flux of the machine, the inner and outer controller coefficients are complex to adjust, and these coefficients vary when the machine parameters are changed due to magnetic saturation and temperature. Disadvantages of the other technique, DTC are that the system has a variable switching frequency due to the absence of a modulator block, and the current or voltage observer used in the system is dependent on the machine parameters [10], [11]. The discovery of FOC and DTC coincided with the period when digital signal processor technology was not as widespread as it is today. However, in recent years, the application of control techniques used in different fields for motor controllers has been opened due to developments in digital signal processor technology. One of these control techniques is finite-control set model predictive control (FCS-MPC), and this control technique has been successfully applied to power electronics and motor controllers in recent years [12]–[14]. FCS-MPC uses the discrete-time model of the system to predict the state variables for all system inputs. These estimated state variables are evaluated with their references in the cost function to find the optimum input to be applied to the system [15]. Among the advantages of this method are that it is easy to use, multi-variable control can be performed via cost function, the dynamic response of the system is fast due to the absence of the modulator block, non-linearities and system constraints can be included via the discrete-time mathematical model of the system [16]. However, the disadvantage of this technique is that the system has no modulator block. Thus, the same voltage vector is applied to the inverter at consecutive sampling times. This reduces the steady-state performance of the system [17]. This drawback can be eliminated by decreasing the sampling time of the system. However, this increases the mathematical burden on the controller [18]. Another method to eliminate this drawback is to include a modulator block. The technique that combines the FCS-MPC and modulator block is called the modulated model predictive control (M^2PC) [19]. M^2PC has improved the steady-state performance of the system, and it has been shown extensively in studies in the literature [20]–[22]. However, the machine models are used in the previously mentioned control techniques. The machine model contains non-linear parameters depending on the magnetic saturation, temperature, and operating region [23]. If the machine model and parameters are not defined correctly in these control techniques, this may cause the system not to work properly [24], [25]. As an alternative solution to the previously mentioned problems, the current control of PMSM was performed for the first time in 2020 with actor-critic based deep machine learning, one of the reinforcement learning methods [26].

RL is a subcategory of machine learning that aims to solve various decision-making and control problems in a data-driven manner. In order to perform the learning process, a reward function similar to the cost function in FCS-MPC must be defined. Then, the learning process is completed, unlike FCS-MPC, when the reinforcement learning agent interacts with the environment around it, learns the control actions that maximize the defined reward function, and creates optimal policy according to these actions [27]. In RL, an agent is referred to by the selected RL method, and the policy relates to the control technique created by the agent after the training.

RL methods can be divided into two categories depending on the control action. The first category is the continuous set, in which the control actions consist of continuous signals, and the second is the discrete set, in which the control actions consist of discrete signals. Deep deterministic policy gradient (DDPG) and twin delay deep deterministic policy gradient (TD3) are the example agents for the first category, and deep-Q network (DQN) is the example agent for the second category [28]. For motor control applications, the environment in which the agent interacts is the inverter and electric motor, and the agent interacts with the environment using the signals received from the inverter and the electric motor. The agent interacts with the signals coming from the environment during training and learns the control process, which consists of optimal control actions with the help of the reward function and creates optimal policy [29]. The control technique based on the DDPG agent has been compared to MPC and FOC techniques [30]. For the first time in 2021, torque control of PMSM is performed using a DQN agent in simulation, and the agent performance has been compared with the MPC-DTC technique in the constant torque and constant speed regions. It has been shown that the dynamic and steady-state performance of the designed agent is better than FCS-MPC [31]. In 2022, the speed control of PMSM is performed using a DDPG agent. The speed and current controllers replace the agent in the FOC, and the performances are compared. Although both controllers show similar results in steady-state, the dynamic response of the controller designed with the DDPG agent is better [32]. The speed control of PMSM is performed in a simulation environment using an FCS-MPC based DQN agent. The mathematical burden on the microcontroller of the designed controller and the FCS-MPC technique is compared with each other. It was shown that the designed controller reduced the mathematical burden on the microcontroller by %24.8 [33]. In , dq-axis inductances, stator resistance and permanent magnet flux linkage of PMSM are estimated using offline deep neural network method. The performance of the created network is compared with different online parameter estimation techniques.

According to the aforementioned informations, the agent incorporates the non-linearities of the inverter and the electric motor into the control by interacting with the signals it receives from the environment during the learning process. Hence, the most important advantage of RL-based control is that it is model-independent, as the agent and the environment interact directly in a data-driven manner [30]. In this study, the speed control of IM with the TD3 agent is performed for the first time in a simulation environment. Inner PI controllers are replaced with TD3 agent. The dynamic and steady-state performances of the designed agent are compared with the FOC-based controller. Extensive simulation results have shown that the dynamic and steady-state performance of the designed agent is better than the FOC. In the introduction, which is the first part of this study, a literature review of the control methods of IM is given. In the second part, mathematical models of the electric motor and inverter and the environment the agent will interact with are given. In the third part, mathematical expressions of the TD3 agent used as a control technique in this study are given. The fourth section provides the control with results with the TD3 agent and compares this control with FOC.

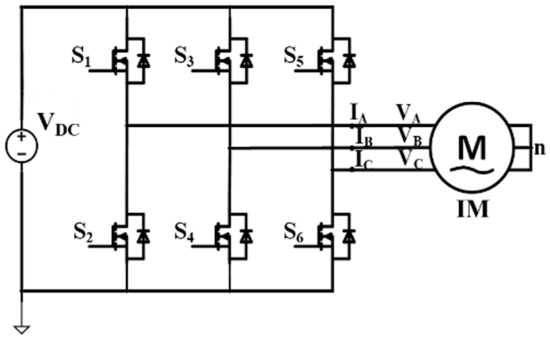

Fig. 1. A voltage-source inverter connected to IM

II. ENVIRONMENT

A. Voltage-source inverter model

Fig. 1. shows the three-phase two-level voltage source inverter (VSI) fed IM. This study uses an ideal inverter, and the VSI is assumed to be in a balanced load state. The phase-to-neutral voltages of phases *A,B,C*, respectively of the inverter are given in (1) [34].

$$\begin{bmatrix} V_{An} \\ V_{Bn} \\ V_{Cn} \end{bmatrix} = \frac{V_{DC}}{3} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} S_1 \\ S_3 \\ S_5 \end{bmatrix} \tag{1}$$

where switch position $S_j \in \{0,1\}$.

B. IM Model

This study gives a dynamic model of the IM in *dq*-axis reference frame where Park transforms are applied to three-phase stator currents. The stator currents of the machine in *dq*-axis are given in (2)-(3) [35].

$$\begin{aligned} \frac{di_{sd}}{dt} = &-\frac{L_r R_s + L_m^2/\tau_r}{\sigma L_s L_r} i_{sd} + \omega_g i_{sq} \\ &+ \frac{L_m}{\tau_r \sigma L_s L_r} \psi_{rd} + \frac{v_{sd}}{\sigma L_s} \\ &+ \frac{L_m}{\sigma L_s L_r} \omega_e \psi_{rq} \end{aligned} \tag{2}$$

$$\begin{aligned} \frac{di_{sq}}{dt} = &-\frac{L_r R_s + L_m^2/\tau_r}{\sigma L_s L_r} i_{sq} - \omega_g i_{sd} \\ &+ \frac{L_m}{\tau_r \sigma L_s L_r} \psi_{rq} + \frac{v_{sq}}{\sigma L_s} \\ &- \frac{L_m}{\sigma L_s L_r} \omega_e \psi_{rd} \end{aligned} \tag{3}$$

where v_{sd}, v_{sq} are *dq*- axis stator voltages, i_{sd}, i_{sq} are *dq*- axis currets, ψ_{rd}, ψ_{rq} are *dq*- axis rotor fluxes, L_s, L_r, L_m are stator, rotor and mutual inductances respectively, $\tau_r = L_r/R_r$ is rotor time constant, $\sigma = (L_r L_s - L_m^2)/L_r L_s$ is leakage constant, ω_g is synchronous angular speed, $\omega_e = P\omega_m$ is electrical angular speed where P is machine pole pair number and ω_m is mechanical angular speed. The dynamic expressions of rotor fluxes with respect to *dq*- axis are given in (4)-(5), and electromagnetic torque expression is given in (6).

$$\frac{d\psi_{rd}}{dt} = \frac{L_m}{\tau_r} i_{sd} - \frac{\psi_{rd}}{\tau_r} + (\omega_g - \omega_e)\psi_{rq} \tag{4}$$

$$\frac{d\psi_{rq}}{dt} = \frac{L_m}{\tau_r} i_{sq} - \frac{\psi_{rq}}{\tau_r} - (\omega_g - \omega_e)\psi_{rd} \tag{5}$$

$$T_e = \frac{3P}{2}\frac{L_m}{L_r}(\psi_{rd} i_{sq} - \psi_{rq} i_{sd}) \tag{6}$$

Rotor flux orientation is one of the field-orientation techniques. In this technique, the rotor flux vector is aligned with the *d*- axis rotor flux, and this ensures $\psi_{rq} = d\psi_{rq}/dt = 0$ in steady-state. If this term is written into (5), the slip angular speed is obtained as given in (7) [15].

$$\omega_{sl} = \omega_g - \omega_e = \frac{L_m}{\tau_r \psi_{rd}} i_{sq} \tag{7}$$

Field-orientation is achieved by writing the expression $\psi_{rq} = d\psi_{rq}/dt = 0$ into (4). Thus, the flux values desired to be produced in the rotor can be obtained with the *d*- axis current as given in (8).

$$i_{sd} = \frac{\psi_{rd}}{L_m} \tag{8}$$

With the completion of rotor flux orientation, the electromagnetic torque expression given in (6) simplifies as given in (9). Electromagnetic torque becomes proportional to *q*- axis current, which is the torque-producing current component.

$$T_e = K_t i_{sq} \tag{9}$$

$K_t = \frac{3P}{2}\frac{L_m \psi_{rd}}{L_r}$ is the torque constant. Thus, using (9), the reference electromagnetic torque value is controlled by i_{sq}. After expressing the mathematical model of the environment in which the agent will interact, the mathematical model of the TD3 agent will be given in the next section.

III. TD3 AGENT

TD3 is one of the RL methods, and RL can be defined as the problem of learning how to best behave in the environment using the trial-error approach. In RL, the agent interacts with the environment and discovers the action that will receive the highest cumulative reward instead of telling the agent which action to perform [36]. In the steps of an RL algorithm, the agent examines the current state of the environment, s_k, based on the observations the agent receives from the environment, o_k, in time step *k*, and performs an action, a_k, based on this interaction. The agent determines the state of the environment based on the observations it receives from the environment. The agent receives a reward, R_{k+1}, based on its action. If the action performed by the agent is an action that brings the environment closer to the desired value, the agent receives a positive reward; otherwise, the agent receives a negative reward. After receiving the reward based on the action performed, the agent moves to the next state, s_{k+1}. The agent creates an optimum policy, π^*, according to these steps, and this π^* acts as a controller. The first expression that the agent uses to find the optimum policy is the Bellman Equation given in (10) [31].

$$g_k = q(o_k, a_k) = E\{R_{k+1} + \gamma q(O_{k+1}, A_{k+1}) \mid O_k = o_k, A_k = a_k\} \tag{10}$$

where g_k is cumulative reward, $E[\cdot]$ is the expected value, $q(o_k, a_k)$ is Q value, and γ is the discount factor. Using (10), Q value is calculated iteratively using the immediate reward and discounted future Q values. The resulting Q value represents the state-action pair. The optimum action can be defined as the action that will give the maximum Q value as given in (11) [23].

$$a^* = \underset{a_{k+1}}{\arg max}\, q(o_k, a_{k+1}) \tag{11}$$

Where a^* is the optimum action. The agent creates the optimum policy by mapping the optimum actions according to the state of the environment. This is called Q-learning. However, due to the non-linearities of the system and reward behavior, it is challenging to map every state action corresponding to each state. To overcome this issue, the neural network, which $\widehat{q_\theta}$ represents the neural network and θ represents the weights of the neural network, can be combined with Q-learning. This is called DQN agent [37]. In DQN, the action that gives the maximum Q value is applied to the environment. In order to update the weights of the neural network, the loss function given in (12) is used. This expression is a mean square error function and $\widehat{q_*}(o_k, a_k)$ is the optimum Q value [31].

$$L_Q = \frac{1}{|B|}\sum_{k=1}^{B}\left(\left(R_{k+1} + \gamma \underset{a_{k+1}}{max} \widehat{q_{\theta,target}}(o_{k+1}, a_{k+1})\right) - \widehat{q_\theta}(o_k, a_k)\right)^2 \tag{12}$$

where B is the mini-batch size. $\widehat{q_{\theta,target}}$ represents the target neural network, which is the same with $\widehat{q_\theta}$. DQN agent is a critic-based agent since only the Q values of the actions are estimated. TD3 is an actor-critic-based agent. Therefore, in the TD3 agent, in addition to the critic network, there is another neural network called actor. Thanks to being an actor-critic, actions can be continuously taken at the output of the agent. The actor creates a policy with continuous actions using the policy gradient approach that is given in (13)-(14) [38].

$$\nabla_{\theta^\pi} J(\pi) = E\left[\nabla_a Q(o_k, a_k|\theta^Q)|_{a=\pi(o_k)} \cdot \nabla_{\theta^\pi}\pi(o_k|\theta^\pi)\right] \tag{13}$$

$$J_{k+1} = J_k - \eta \nabla J_k \tag{14}$$

where η is the learning rate. The critic calculates the policy value created by the actor with the temporal difference approach. There are two critical networks in TD3, also called wthe double Q network. These two critic networks reduce the overestimation problem of other actor-critic-based agents, and the critic evaluates the actor's policy with a delay [39]. There are two target Q networks since there are two Q networks in TD3 agent. There is also one target network for the policy network. Weight updates for these target networks can be done using (15)-(16) [40].

$$\theta^{\pi_{target}} \leftarrow (1-\rho)\theta^{\pi_{target}} + \rho\theta^\pi \tag{15}$$

$$\theta^{Q_{target}} \leftarrow (1-\rho)\theta^{Q_{target}} + \rho\theta^Q \tag{16}$$

where ρ is the target smoothing factor. The same mathematical expression given in DQN is used to update the critic network in TD3. The actor updates the policy according to the value calculated by the critic. Then, the optimal action that maximizes cumulative reward is selected and applied to the environment. The developed drive system with TD3 agent and block diagram of TD3 agent are given in Figure 2 and Figure 3, respectively.

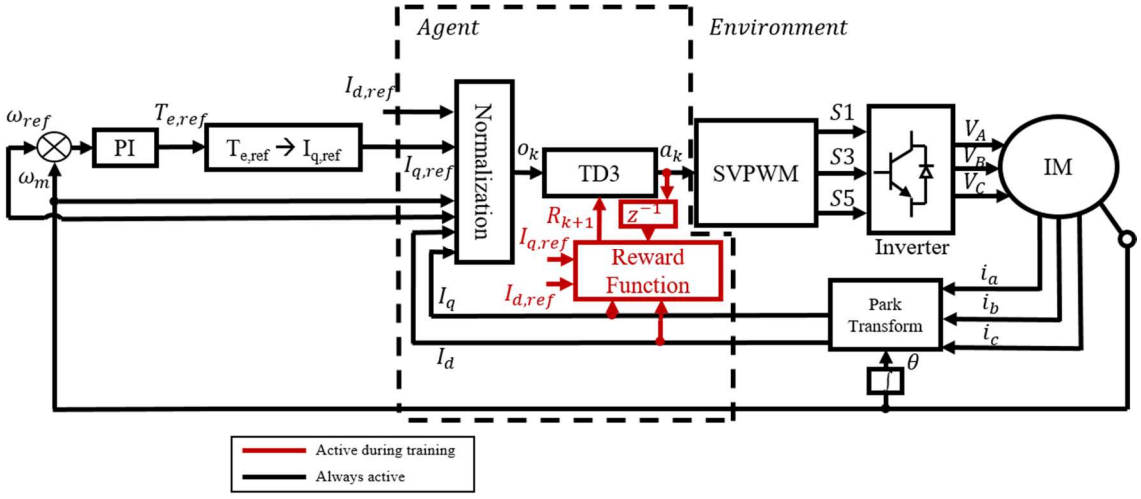

Fig. 2. Block diagram of the drive system with TD3 agent

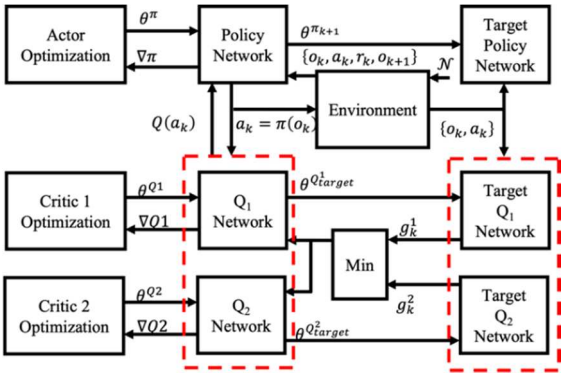

Fig. 3. Block diagram of TD3 agent [40]

The block diagram of the drive system with the TD3 agent is given in Figure 3. As can be seen from Figure 3, the inner PI controllers in the FOC technique have been replaced by TD3 agents. The normalized observation vector of the control system is given in (17).

$$o_k = \left[I_{d_{ref}}, I_{q_{ref}}, I_d, I_q, I_{d_{error}}, I_{q_{error}}, \omega_{ref}, \omega_m \right] \quad (17)$$

Using the observation vector given in (17), the agent creates the policy that will maximize the reward function given in (18). The reward function is similar to the cost function in FCS-MPC.

$$R_k = -\left(C_1 \left| I_{d,ref} - I_d \right|^2 + C_2 \left| I_{q,ref} - I_q \right|^2 + C_3 a_{k-1} \right) \quad (18)$$

The action at the output of the TD3 agent is set as reference *dq*- axis voltages as in FOC. The optimal switching vectors are applied to the inverter using space vector pulse width modulation (SVPWM) so that the reference voltages generated in the output of the TD3 agent can be produced in IM. Hence, the system has a constant switching frequency.

IV. SIMULATION RESULTS

The drive system given in Figure 3 is developed in simulation, and speed control of IM is performed with a TD3 agent. First, the simulation results with the TD3 agent are given. Afterwards, the dynamic and steady-state results of the agent are compared with the FOC technique. Parameters of the IM and the inverter used in the simulations and hyperparameters of the TD3 agent are given in Table 1.

Table I. SIMULATION PARAMETERS

Parameter	Description	Values
V_{DC}	DC bus voltage	150 V
P	Number of pole pairs	2
T_e	Nominal torque	7.5 Nm
L_m	Mutual inductance	442.3 mH
L_{ls}	Stator leakage inductance	10.145 mH
L_{lr}	Rotor leakage inductance	10.061 mH
R_s	Stator resistance	3.919 Ω
R_r	Rotor resistance	4.9618 Ω
T_s	Sampling time	100 µs
$f_{a,c}$	Critic-actor activation function	ReluLayer
B	Mini-batch size	512
D	Buffer length	2×10^6
γ	Discount factor	0.995
ρ	Target smoothing factor	0.005
η	Learn rate	1×10^{-3}
K	Number of episodes	1500

ω_{ref}=500 rpm, and the demanded electromagnetic torque from the machine is set to T_L=1.5 Nm between t=0-2 seconds. At t=2 s, the reference speed is increased to ω_{ref}=750 rpm, and the demanded electromagnetic torque is kept the same until t=3 s. At t=3 s, the demanded electromagnetic torque from the machine is increased by %100 to T_L=3 Nm. The dynamic and steady-state performances of the TD3 agent are evaluated by applying step changes for speed and electromagnetic torque. TD3 agent is trained for K=1500 training steps by creating the environment according to the information. The agent with the highest average reward is the optimum agent and, therefore, the controller. The simulation results are presented in Fig. 4. When the system operates under the transient conditions at the time interval t=0-0.5 s, there is no overshoot in rotor speed. As shown in Figure 4, the drive system controlled by the TD3 agent reaches the reference speed of ω_{ref}=500 rpm and an electromagnetic torque of T_e=1.5 Nm in ~0.5 s, and then the system becomes the steady-state. At ω_{ref}=500 rpm and T_e=1.5 Nm torque, total harmonic distortion (THD) of the phase A current is %1.36 which is quite good. When the speed step is applied to the system at t=2 s, and the speed command is increased to ω_{ref}=750 rpm, TD3 agent reacts to this increase by increasing the frequency of phase currents. As seen in Figure 4, the system's settling time is ~0.5 s, and there is no overshoot in the system's dynamic response. When the demanded torque value of the system is increased from T_e=1.5 Nm to T_e=3 Nm in t=3 s, the TD3 agent responds to this step change by adjusting the phase currents magnitudes. There is a slight decrease in the speed of the motor in response to this increase in torque. However, the outer PI controller changed

the reference torque value, and the agent adjust the magnitude of the currents. Then, the speed rapidly reached the reference value ω_{ref}=750 rpm. When the system operates at these reference values, the THD of phase A current is %0.63, which is significantly low. In summary, the dynamic response of the drive system using the TD3 agent at different speeds and torque steps is fast, and there is no overshoot. In the steady-state response of the system, the THD of the phase currents and the ripple in electromagnetic torque are very low, as can be seen in Figure 4. These responses increase the power quality of the system and extend the life of the motor.

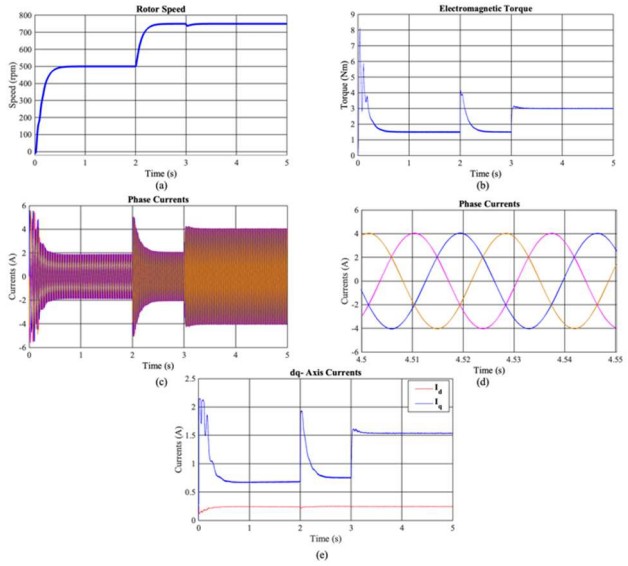

Fig. 4. Results of the drive system with TD3 agent; a) Rotor speed, b) Electromagnetic torque, c) Phase currents, d) Zoom-in phase currents, e) dq- axis currents

The speed control of IM with the TD3 agent has been successfully performed in the simulation environment. In Figure 5, the dynamic and steady-state response of the drive system with the TD3 agent are compared with the FOC technique. The scenario in the simulation performed with the FOC technique is the same as that in the simulation performed with the TD3 agent. The settling time of the motor for both techniques is ~0.5 s at the start-up, and there is no overshoot. THD of the motor phase A current is %1.14 in the FOC technique. The FOC technique has a small overshoot when the speed step is applied. Although the TD3 agent reaches the reference speed value later than the FOC technique, there is no unwanted overshoot. When the torque step is applied, there is a smaller decrease in motor speed for the drive system using the TD3 agent compared to the FOC technique. In addition, the overshoot in the electromagnetic torque is greater, and the system reaches the reference speed, while torque values are slower in the FOC technique. At this torque, the THD of the phase A current is %0.74 and is greater compared to the drive system using the TD3 agent. THD comparison of controllers are summarized in Table 2. The drive system with TD3 agent gives better results than the FOC technique in overshoot and steady-state performance at higher torque values.

Table II. THD Comparison

Control Technique	Torque (Nm)	THD (%)
TD3	1.5 / 3	1.36 / 0.63
FOC	1.5 / 3	1.14 / 0.74

V. CONCLUSION

In this study, speed control of IM has been performed in a simulation environment with a TD3 agent, which is one of the RL methods. The designed TD3 agent was trained in an entirely data-driven manner and independent of system parameters, and the agent with the highest average reward was selected as the controller with the optimum policy. It has been shown through extensive simulation results that the agent can successfully control the system when the speed and the torque steps were applied. Comparisons of the designed drive system with the FOC technique have shown that the designed system gives better results than the FOC technique in terms of overshoot when the speed step is applied. When the torque step is applied, it has been shown that the system gives better overshoot and settling time results than the FOC technique. Thus, it has been shown that motor control can be performed independently of parameters with RL-based methods.

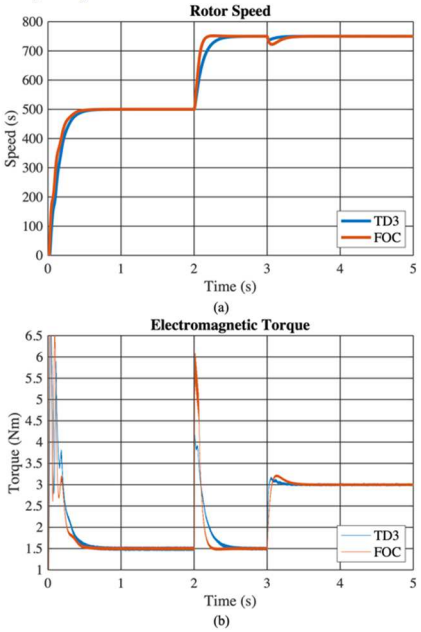

Fig. 5. TD3 agent and FOC comparison; a) Rotor speed, b) Electromagnetic torque

REFERENCES

[1] C. G. Dias and L. C. Da Silva, "Induction Motor Speed Estimation Based on Airgap Flux Measurement Using Hilbert Transform and Fast Fourier Transform," *IEEE Sens. J.*, vol. 22, no. 13, pp. 12690–12699, 2022, doi: 10.1109/JSEN.2022.3176085.

[2] V. Dmitrievskii, V. Prakht, E. Valeev, A. Paramonov, V. Kazakbaev, and A. Anuchin, "Comparative Study of Induction and Wound Rotor Synchronous Motors for the Traction Drive of a Mining Dump Truck Operating in Wide Constant Power Speed Range," *IEEE Access*, vol. 11, no. June, pp. 68395–68409, 2023, doi: 10.1109/ACCESS.2023.3292244.

[3] M. Wu, X. Sun, J. Zhu, G. Lei, and Y. Guo, "Improved Model Predictive Torque Control for PMSM Drives Based on Duty Cycle Optimization," *IEEE Trans. Magn.*, vol. 57, no. 2, 2021, doi: 10.1109/TMAG.2020.3008495.

[4] B. GÖRDÜK, "DESING AND IMPLEMENTATION OF SENSORLESS VECTOR CONTROLLED DRIVE FOR PMSMs," Istanbul Technical University, Electrical Engineering, Istanbul, 2019.

[5] I. Takahashi and T. Noguchi, "A New Quick-Response and High-Efficiency Control Strategy of an Induction Motor," *IEEE Trans. Ind. Appl.*, vol. IA-22, no. 5, pp. 820–827, 1986, doi: 10.1109/TIA.1986.4504799.

[6] I. R. Akhil and C. K. Vijayakumari, "Modified direct torque control scheme for PMSM," in *PEDES 2012 - IEEE International Conference on Power Electronics, Drives and Energy Systems*, 2012, no. 1, pp. 1–6, doi: 10.1109/PEDES.2012.6484328.

[7] A. Samar, P. Saedin, A. I. Tajudin, and N. Adni, "The implementation of Field Oriented Control for PMSM drive based on TMS320F2808 DSP controller," in *Proceedings - 2012 IEEE International Conference on Control System, Computing and Engineering, ICCSCE 2012*, 2012, pp. 612–616, doi: 10.1109/ICCSCE.2012.6487218.

[8] D. Casadei, F. Profumo, G. Serra, and A. Tani, "FOC and DTC: Two viable schemes for induction motors torque control," *IEEE Trans. Power Electron.*, vol. 17, no. 5, pp. 779–787, 2002, doi: 10.1109/TPEL.2002.802183.

[9] S. E. Rezgui, A. Mehdi, S. Legrioui, H. Meddouce, A. M. Boulahia, and H. Benalla, "IRFOC vs DTC performance comparison analysis," *2013 3rd Int. Conf. Electr. Power Energy Convers. Syst. EPECS 2013*, pp. 1–6, 2013, doi: 10.1109/EPECS.2013.6713066.

[10] S. Dwivedi, S. M. Tripathi, and S. K. Sinha, "Review on Control Strategies of Permanent Magnet-Assisted Synchronous Reluctance Motor Drive," in *2020 International Conference on Power Electronics and IoT Applications in Renewable Energy and its Control, PARC 2020*, 2020, pp. 124–128, doi: 10.1109/PARC49193.2020.236570.

[11] H. Mohan, M. K. Pathak, and S. K. Dwivedi, "Direct power control of induction motor drives," *Proc. - 2019 IEEE 13th Int. Conf. Compat. Power Electron. Power Eng. CPE-POWERENG 2019*, pp. 1–5, 2019, doi: 10.1109/CPE.2019.8862412.

[12] H. Xie, W. Tian, X. Gao, F. Wang, J. Rodriguez, and R. Kennel, "An Ensemble Regulation Principle for Multiobjective Finite-Control-Set Model-Predictive Control of Induction Machine Drives," *IEEE Trans. Power Electron.*, vol. 38, no. 3, pp. 3069–3083, 2023, doi: 10.1109/TPEL.2022.3220289.

[13] S. Alireza Davari, V. Nekoukar, S. Azadi, F. Flores-Bahamonde, C. Garcia, and J. Rodriguez, "Discrete Optimization of Weighting Factor in Model Predictive Control of Induction Motor," *IEEE Open J. Ind. Electron. Soc.*, vol. 4, no. September, pp. 573–582, 2023, doi: 10.1109/OJIES.2023.3333873.

[14] M. Gokdag and O. Gulbudak, "Dual-model predictive control of two independent induction motors driven by a SiC nine-switch inverter," *Int. J. Electron.*, vol. 110, no. 1, pp. 124–142, 2023, doi: 10.1080/00207217.2021.2007545.

[15] M. Gokdag and O. Gulbudak, "Dual-model predictive control of two independent induction motors driven by a SiC nine-switch inverter," *Int. J. Electron.*, vol. 00, no. 00, pp. 1–19, 2021, doi: 10.1080/00207217.2021.2007545.

[16] S. Vazquez, J. Rodriguez, M. Rivera, L. G. Franquelo, and M. Norambuena, "Model Predictive Control for Power Converters and Drives: Advances and Trends," *IEEE Trans. Ind. Electron.*, vol. 64, no. 2, pp. 935–947, 2017, doi: 10.1109/TIE.2016.2625238.

[17] J. P. Zucuni, F. Carnielutti, H. Pinheiro, M. Norambuena, and J. Rodriguez, "Cost Function Design for Stability Assessment of Modulated Model Predictive Control," in *2020 22nd European Conference on Power Electronics and Applications, EPE 2020 ECCE Europe*, 2020, pp. 1–9, doi: 10.23919/EPE20ECCEEurope43536.2020.9215797.

[18] U. U. Korpe, M. Gokdag, M. Koc, and O. Gulbudak, "Modulated Model Predictive Control of Permanent Magnet Synchronous Motors with Improved Steady-State Performance," in *Proceedings - 2021 IEEE 3rd Global Power, Energy and Communication Conference, GPECOM 2021*, 2021, pp. 67–72, doi: 10.1109/GPECOM52585.2021.9587747.

[19] U. U. Korpe, M. Gokdag, M. Koç, and O. Gulbudak, "Modulated Model Predictive Torque Control for Interior Permanent Magnet Synchronous Machines," *El-Cezeri*, vol. 2022, no. 2, pp. 777–787, 2022, doi: 10.31202/ecjse.1008121.

[20] M. Gokdag, "Modulated Predictive Control to Improve the Steady-State Performance of NSI-Based Electrification Systems," *Energies*, vol. 15, no. 6, 2022, doi: 10.3390/en15062043.

[21] T. Jin *et al.*, "Low Complexity Model Predictive Flux Control Based on Discrete Space Vector Modulation and Optimal Switching Sequence for Induction Motors," *IEEE Trans. Ind. Electron.*, vol. 71, no. 1, pp. 305–315, 2024, doi: 10.1109/TIE.2023.3241412.

[22] Z. Zhang, Q. Sun, and Q. Zhang, "A Computationally Efficient Model Predictive Control Method for Dual Three-Phase PMSM of Electric Vehicle With Fixed Switching Frequency," *IEEE Trans. Ind. Appl.*, vol. 60, no. 2, pp. 1105–1116, 2023, doi: 10.1109/TIA.2023.3278255.

[23] B. Haucke-Korber, M. Schenke, and O. Wallscheid, "Reinforcement Learning-Based Deep Q Direct Torque Control with Adaptable Switching Frequency Towards Six-Step Operation of Permanent Magnet Synchronous Motors," 2022.

[24] D. Jakobeit, M. Schenke, and O. Wallscheid, "Meta-Reinforcement-Learning-Based Current Control of Permanent Magnet Synchronous Motor Drives for a Wide Range of Power Classes," *IEEE Trans. Power Electron.*, vol. 38, no. 7, pp. 8062–8074, Jul. 2023, doi: 10.1109/TPEL.2023.3256424.

[25] A. Traue, G. Book, W. Kirchgassner, and O. Wallscheid, "Toward a Reinforcement Learning Environment Toolbox for Intelligent Electric Motor Control," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 33, no. 3, pp. 919–928, Mar. 2022, doi: 10.1109/TNNLS.2020.3029573.

[26] M. Schenke, W. Kirchgassner, and O. Wallscheid, "Controller Design for Electrical Drives by Deep Reinforcement Learning: A Proof of Concept," *IEEE Trans. Ind. Informatics*, vol. 16, no. 7, pp. 4650–4658, Jul. 2020, doi: 10.1109/TII.2019.2948387.

[27] S. Zhang, O. Wallscheid, and M. Porrmann, "Machine Learning for the Control and Monitoring of Electric Machine Drives: Advances and Trends," *IEEE Open J. Ind. Appl.*, vol. 4, pp. 188–214, 2023, doi: 10.1109/OJIA.2023.3284717.

[28] T. Schindler, L. Broghammer, P. Karamanakos, A. Dietz, and R. Kennel, "Deep Reinforcement Learning Current Control of Permanent Magnet Synchronous Machines," 2023, doi: 10.1109/IEMDC55163.2023.10238988.

[29] Y. Wang, S. Fang, J. Hu, and D. Huang, "Multiscenarios Parameter Optimization Method for Active Disturbance Rejection Control of PMSM Based on Deep Reinforcement Learning," *IEEE Trans. Ind. Electron.*, vol. 70, no. 11, pp. 10957–10968, Nov. 2023, doi: 10.1109/TIE.2022.3225829.

[30] G. Book *et al.*, "Transferring online reinforcement learning for electric motor control from simulation to real-world experiments," *IEEE Open J. Power Electron.*, vol. 2, pp. 187–201, 2021, doi: 10.1109/OJPEL.2021.3065877.

[31] M. Schenke and O. Wallscheid, "A Deep Q-Learning Direct Torque Controller for Permanent Magnet Synchronous Motors," *IEEE Open J. Ind. Electron. Soc.*, vol. 2, pp. 388–400, 2021, doi: 10.1109/OJIES.2021.3075521.

[32] 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 Trans. Power Electron.*, vol. 37, no. 8, pp. 9000–9011, Aug. 2022, doi: 10.1109/TPEL.2022.3153845.

[33] Z. Tang, C. Ma, J. Rodriguez, C. Garcia, and W. Song, "Data-Driven Finite-Set Predictive Current Control via Deep Q-Learning for Permanent Magnet Synchronous Motor Drives," 2023, doi: 10.1109/PRECEDE57319.2023.10174508.

[34] J. Rodríguez *et al.*, "Predictive Current Control of a Voltage Source Inverter," *IEEE Trans. Ind. Electron.*, vol. 54, no. 1, pp. 495–503, 2007, doi: 10.1109/TIE.2006.888802.

[35] O. Gulbudak and M. Gokdag, "Dual-Hysteresis Band Control of Nine-Switch Inverter to Control Two Induction Motors," *IEEE Trans. Energy Convers.*, vol. 8969, no. c, pp. 1–11, 2021, doi: 10.1109/TEC.2021.3131385.

[36] F. Rossi, G. Gruosso, and G. S. Gajani, "A Reinforcement Learning based controller for optimal speed control of a DC motor using deep Q-network algorithm," in *EUROCON 2023 - 20th International Conference on Smart Technologies, Proceedings*, 2023, pp. 181–186, doi: 10.1109/EUROCON56442.2023.10198954.

[37] V. Mnih *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015, doi: 10.1038/nature14236.

[38] X. Qi, W. Cao, and L. Aarniovuori, "Reinforcement Learning Based Parameter Lookup Table Generating Method for Optimal Torque Control of Induction Motors," *IEEE Trans. Ind. Electron.*, vol. 70, no. 5, pp. 4516–4525, May 2023, doi: 10.1109/TIE.2022.3189103.

[39] F. Yin, X. Yuan, Z. Ma, and X. Xu, "Vector Control of PMSM Using TD3 Reinforcement Learning Algorithm," *Algorithms*, vol. 16, no. 9, Sep. 2023, doi: 10.3390/a16090404.

[40] Y. Wang, S. Fang, and D. Huang, "An Improved Model-Free Active Disturbance Rejection Deadbeat Predictive Current Control Method of PMSM Based on Data-Driven," *IEEE Trans. Power Electron.*, vol. 38, no. 8, pp. 9606–9616, Aug. 2023, doi: 10.1109/TPEL.2023.3280013.