

## RESEARCH ARTICLE

# Reinforcement learning algorithm for improving speed response of a five-phase permanent magnet synchronous motor based model predictive control

Ahmed M. Hassan<sup>1,2\*</sup>, Jafar Ababneh<sup>3</sup>, Hani Attar<sup>4,5</sup>, Tamer Shamseldin<sup>6</sup>, Ahmed Abdelbaset<sup>2</sup>, Mohamed Eladly Metwally<sup>2</sup>

**1** Department of Electrical Power and Machines Engineering, Faculty of Engineering, Benha University, Shoubra, Cairo, Egypt, **2** Department of Electrical Power and Machines Engineering, Higher Institute of Engineering (HIE), El-Shorouk Academy, El-Shorouk City, Egypt, **3** Cyber Security Department, Faculty of Information Technology, Zarqa University, Zarqa, Jordan, **4** Faculty of Engineering, Zarqa University, Zarqa, Jordan, **5** College of Engineering, University of Business and Technology, Jeddah, Saudi Arabia, **6** Technical Research Center, Cairo, Egypt

\* [ahmed.hassanin@feng.bu.edu.eg](mailto:ahmed.hassanin@feng.bu.edu.eg)



## OPEN ACCESS

**Citation:** Hassan AM, Ababneh J, Attar H, Shamseldin T, Abdelbaset A, Metwally ME (2025) Reinforcement learning algorithm for improving speed response of a five-phase permanent magnet synchronous motor based model predictive control. PLOS ONE 20(1): e0316326. <https://doi.org/10.1371/journal.pone.0316326>

**Editor:** Omer Saleem, National University of Computer and Emerging Sciences - Lahore Campus, PAKISTAN

**Received:** August 9, 2024

**Accepted:** December 10, 2024

**Published:** January 3, 2025

**Copyright:** © 2025 Hassan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript.

**Funding:** The author(s) received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist.

## Abstract

Enhancing the performance of 5ph-IPMSM control plays a crucial role in advancing various innovative applications such as electric vehicles. This paper proposes a new reinforcement learning (RL) control algorithm based twin-delayed deep deterministic policy gradient (TD3) algorithm to tune two cascaded PI controllers in a five-phase interior permanent magnet synchronous motor (5ph-IPMSM) drive system based model predictive control (MPC). The main purpose of the control methodology is to optimize the 5ph-IPMSM speed response either in constant torque region or constant power region. The speed responses obtained using RL control algorithm are compared with those obtained using four of the most recent metaheuristic optimization techniques (MHOT) which are Transit Search (TS), Honey Badger Algorithm (HBA), Dwarf Mongoose (DM), and Dandelion-Optimizer (DO) optimization techniques. The speed response are compared in terms of the settling time, rise time, maximum time and maximum overshoot percentage. It is found that the suggested RL based TD3 give minimum settling time and relatively low values for the rise time, max time and overshoot percentage which makes the RL provide superior speed responses compared with those obtained from the four MHOT. The drive system speed responses are obtained in the constant torque region and constant power region using MATLAB SIMULINK package.

## 1. Introduction

IPMSMs are highly suitable for high-performance drive systems due to their various advantages. These include exceptional performance, high efficiency, compact size, low noise, and reliability [1–3]. IPMSMs are widely used in various applications, with robots and electric vehicles being among the most common [2, 3]. Due to their inveterate advantages, multiphase

motor drives are gaining prominence as a viable alternative to conventional three-phase motor drives. These advantages encompass minimized torque pulsations, heightened power density, and enhanced capability for fault-tolerant [4, 5]. MPC is an increasingly prominent control method for drive systems, demonstrating superior performance and optimization [6]. Enhancing the performance of 5ph-IPMSM control plays a crucial role in advancing various innovative applications. In the drive systems, the tuning of PI controllers affects the speed response of the motor. The tuning is usually achieved by metaheuristic techniques. Recently, RL are used instead to tune the PI controller [7] in DC drive system because it is an excellent multi-objective optimization technique due to its strong search capabilities and rapid convergence rate. To the best knowledge of the authors, no existing literature has investigated the optimization of the speed response of five-phase IPMSM based MPC using RL based TD3 algorithm across a wide speed range. Several authors presented several methods of control. These publications are explored as follows:

In [8], a harmonic elimination method utilizing vector control was introduced for the five-phase IPMSM. This method, known as harmonic elimination SVM, operates on the principles of space vector theory. In [9], a study was conducted on a 5ph-PMSM VC. A SVM algorithm was proposed to achieve high-performance VC in the drive system. Similarly, [10] introduced a VC strategy for a five-phase VSI utilizing SVM that powers a 5ph-PMSM. This approach optimized the performance of the 5ph-PMSM drive system. Furthermore, in [11], a mathematical model for the five-phase PMSM and corresponding control algorithms were presented. The modeling was achieved using coordinate transformation, and both the two-VC algorithm and four-VC algorithm SVPWMs were investigated. Additionally, a modified four-VC algorithm was introduced to modify the 3rd harmonic current, thereby improving the motor's torque performance. In [12], a speed sensorless DTC method was proposed for five-phase IPMSM. This method relies on measuring the per-phase currents and the voltage of the DC bus. Additionally, [13] introduced a model reference adaptive control, incorporating a neural network, for a five-phase IPMSM. Combined with a hysteresis current controller, this controller enables motor speed control across a wide range of speeds. Furthermore, [14] presented a sensorless control methodology based on 3rd harmonic space for a five-phase PMSM, eliminating the need for motor parameters throughout the entire speed range. In [15], a super-twisting SMC was proposed to control a five-phase PMSM. This control technique exhibits superior response and performance compared to vector control in various conditions. DTMPC of a 5ph-PMSM drive system was presented in [16]. This control strategy effectively reduces the ripples of torque associated with conventional DTC, leading to reduced harmonics with high-orders and losses of the system. A PCC strategy based on a model of finite control set was introduced for a 5-ph PMSM in [17]. This approach reduces the THD to 9.47% and eliminates third harmonic currents. In [18], a MPDTC method was proposed based on the QEM method and a HVE method. The THD was reduced to 11.54% by employing this technique while disregarding third harmonic currents. [19] introduced a DTMPC technique for a five-phase PMSM. This approach aims to optimize torque development, improve speed regulation, reduce the ripples in torque, reduce current harmonics having higher orders, decrease losses in the drive system, and increase the power train's efficiency. In [20], a MPTC technique based on double virtual voltage vectors utilizing geometric principles was presented for controlling a five-phase PMSM. This method achieves reduced processing time without the need for weighting factors. Additionally, [21] proposed an MPTC technique with additional weighting factors, reducing current harmonics and torque ripple. The THD was decreased to 7.11% using this method while disregarding third harmonic currents. Reference number [22] introduced a model-free PCC methodology. This methodology was based on a model called an ultra-local and the outputs of the motor for five-phase PMSM drives. This strategy mitigates

the impact of motor parameter variations on current predictions. Moreover, [23] presented a MPCC methodology for a five-phase PMSM, utilizing pre-selection for the voltage vector. The purpose of this methodology is to reduce computation time compared to conventional MPC by selecting the optimal voltage vector based on deviations in stator current and changes in the position of the stator current vector.

In [24], an ANN-based MPC strategy was suggested to control the speed of a 3-phase IPMSM. This approach utilized a predictive algorithm of back propagation (BP) network and MPC. The optimal selection of the gains of a PI controller in a VC three-phase PMSM drive system was addressed in [25]. Optimization algorithms called RGA and BBO were employed for this purpose. It was observed that using the BBO algorithm exhibited superior transient as well as steady-state performance in the PMSM drive system. In [26], a three-phase surface-mounted PMSM was controlled using an adaptive ANN internal model utilizing the PSO algorithm. Various optimization techniques were compared in [27] to obtain the optimal selection of PI controller gains for a three-phase PMSM drive system. A neural network-based MPC technique to regulate the speed of a 3-phase IPMSM was presented in [28]. Similarly, in [29], a control technique utilizing a BP ANN algorithm was given for a 3-phase IPMSM. Sensorless control of surface-mounted PMSM was achieved in [30] using an adaptive speed observer and a PID controller. In [31], a control technique for a three-phase IPMSM utilizing a MPC based on ANN was introduced. [32, 33] presented ANN-based MPC techniques for a three-phase IPMSM to overcome the effects of parameter mismatches. A dual-vector-based Particle Swarm Optimization MPC technique for a three-phase IPMSM was introduced in [34].

In [35] the reinforcement Q-learning algorithm was presented to tune fuzzy PD and PI controllers for SISO and TITO systems. The deep RL (DRL) was used to improve the tuning process for classical PID controller was presented in [36]. The algorithm that was used in RL is the DDPG. The concept of RL based FOC of a three-phase induction motor was presented in [37]. In [38] the design of adaptive RL PID controller under the structure of Actor-Critic which was based on RBF network for nonlinear systems was presented. Online training for a RL used to control real motor drive system was presented in [39]. The adopted drive system was composed of three-phase PMSM fed by VSI. In [40] DRL method for speed control of a three-phase PMSM servo drive system was presented. The presented DRL control improved the system performance especially in case of load variations. An adaptive PID controller using the algorithm of asynchronous advantage actor-critic was proposed in [41]. In [42] speed control of a PMSM is achieved by applying the PI based PSO and DDPG algorithms. In [43] a combined approach that leverages DRL and MPC was introduced to enhance the efficiency of electric vehicles. In [44] an open source toolbox called GEM was developed for training of RL-agents for the controlling of electric motors. The improved control performance for a three-phase PMSM was proposed in [45]. This improvement was achieved using four optimization techniques and RL. The used optimization techniques were PSO, SA, GA, and GWO. An RL control algorithm to control a three-phase PMSM based on the TD3 was suggested in [46]. In [47], an adaptive PID controller was designed for controlling the speed of a DC motor using RL based TD3 algorithm. In [48], a comparison was made between the conventional PID methodology and a TD3 RL algorithm to control three-phase PMSM based on the strategy of vector control (VC). In [7], researchers employed the TD3 method to learn PI controller for optimal dynamics in the simulation environment of controlling the speed of a DC motor. Energy management strategies for hybrid electric vehicles (HEVs) was dealt with in [49]. A hierarchical architecture that combines RL algorithms was proposed. The DDPG algorithm demonstrates superior performance in the energy management of HEVs. In [50], an intelligent system for energy management for a conventional autonomous vehicle using RL was proposed. A novel exploration strategy called self-adaptive Q-learning was introduced.

In this study, we suggest a novel RL control algorithm utilizing the TD3 approach. The algorithm is designed to fine-tune two cascaded PI controllers in a 5-phase voltage source inverter (VSI) / 5ph-IPMSM based MPC drive system. The primary goal of this control methodology is the optimization of the speed response of the 5ph-IPMSM under various operating conditions. We compare the speed responses obtained using the RL control algorithm with those achieved using four recent MHOT: TS, HBA, DM, and DO. The speed response are compared in terms of the settling time, rise time, maximum time and maximum overshoot percentage.

The following points summarize the main contributions of this paper:

- A new RL control algorithm utilizing the TD3 algorithm for tuning the two cascaded PI controllers in the drive system under consideration is proposed.
- The tuning is achieved using the most recent optimization techniques: TS, HBA, DM, and DO.
- A comparative study is achieved among the suggested RL control methodology and the MHOT adopted in this research for the drive system under consideration.
- A MATLAB SIMULINK is accomplished for the drive system under consideration to obtain the results of simulation and verify the validity of the proposed control methodology.

The subsequent sections of this study are arranged as follows. In Section 2 the modeling of the drive system is introduced. Explanation of the MPC is given in Section 3. MHOT are explained briefly in Section 4. In Section 5, the suggested RL utilizing TD3 algorithm is explained. In Section 6, the proposed control methodology of the drive system is explored. In Section 7 simulation results are given. In Section 8, key conclusions drawn from our study are summarized.

## 2. Modeling of the drive system

### 2.1 Modeling of the five-phase VSI

The 5-phase VSI is expressed using the per-phase voltages ( $v_a$  to  $v_e$ ), which depend on the inverter switching functions [19]

$$\begin{bmatrix} v_a \\ v_b \\ v_c \\ v_d \\ v_e \end{bmatrix} = \frac{V_{dc}}{5} \begin{bmatrix} 4 & -1 & -1 & -1 & -1 \\ -1 & 4 & -1 & -1 & -1 \\ -1 & -1 & 4 & -1 & -1 \\ -1 & -1 & -1 & 4 & -1 \\ -1 & -1 & -1 & -1 & 4 \end{bmatrix} \cdot \begin{bmatrix} s_a \\ s_b \\ s_c \\ s_d \\ s_e \end{bmatrix} \tag{1}$$

In Eq (1),  $V_{dc}$  represents the DC voltage supplied to the inverter. The switching functions, denoted as  $S_a$  to  $S_e$ , correspond to the various states of the inverter. Specifically, a switching function equals one when the upper semiconductor switch in a leg is active, and it equals zero when the lower switch is inactive in that leg.

In a five-phase VSI, there exist 32 possible switching states. The MPC determines the most suitable switching state to obtain minimum cost function, and based on these states, gate pulses are generated for the ten switches.

### 2.2 Model of the 5ph-IPMSM

The 5ph-IPMSM is described using the DQ model in the synchronous frame of reference. The voltage ABCDE to DQ transformations are provided by references [19, 51]:

$$\begin{bmatrix} v_{d1} \\ v_{q1} \\ v_{d3} \\ v_{q3} \\ v_0 \end{bmatrix} = \frac{2}{5} \begin{bmatrix} \cos\theta & \cos(\theta - \alpha) & \cos(\theta - 2\alpha) & \cos(\theta + 2\alpha) & \cos(\theta + \alpha) \\ -\sin\theta & -\sin(\theta - \alpha) & -\sin(\theta - \alpha) & -\sin(\theta + 2\alpha) & -\sin(\theta + \alpha) \\ \cos 3\theta & \cos 3(\theta - \alpha) & \cos 3(\theta - 2\alpha) & \cos 3(\theta + 2\alpha) & \cos 3(\theta + \alpha) \\ -\sin 3\theta & -\sin 3(\theta - \alpha) & -\sin 3(\theta - 2\alpha) & -\sin 3(\theta + 2\alpha) & -\sin 3(\theta + \alpha) \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \cdot \begin{bmatrix} v_a \\ v_b \\ v_c \\ v_d \\ v_e \end{bmatrix} \quad (2)$$

where  $\alpha = 2\pi/5$ ,  $\theta$  is the rotor position angle,  $v_{d1}, v_{q1}$  are the fundamental stator voltages in the DQ frame of reference,  $v_{d3}, v_{q3}$  are the third-harmonic DQ components of stator voltage.

The transformation from the DQ frame of reference to the ABCDE frame of reference can be represented as follows [19]:

$$\begin{bmatrix} i_{as} \\ i_{bs} \\ i_{cs} \\ i_{ds} \\ i_{es} \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & \cos 3\theta & -\sin 3\theta & \frac{1}{\sqrt{2}} \\ \cos(\theta - \alpha) & -\sin(\theta - \alpha) & \cos 3(\theta - \alpha) & -\sin 3(\theta - \alpha) & \frac{1}{\sqrt{2}} \\ \cos(\theta - 2\alpha) & -\sin(\theta - 2\alpha) & \cos 3(\theta - 2\alpha) & -\sin 3(\theta - 2\alpha) & \frac{1}{\sqrt{2}} \\ \cos(\theta + 2\alpha) & -\sin(\theta + 2\alpha) & \cos 3(\theta + 2\alpha) & -\sin 3(\theta + 2\alpha) & \frac{1}{\sqrt{2}} \\ \cos(\theta + \alpha) & -\sin(\theta + \alpha) & \cos 3(\theta + \alpha) & -\sin 3(\theta + \alpha) & \frac{1}{\sqrt{2}} \end{bmatrix} \cdot \begin{bmatrix} i_{d1} \\ i_{q1} \\ i_{d3} \\ i_{q3} \\ i_0 \end{bmatrix} \quad (3)$$

where  $i_{d1}, i_{q1}$  are the fundamental stator currents DQ components,  $i_{d3}, i_{q3}$  are the third-harmonic DQ components of stator currents. The DQ voltage equations for the 5ph-IPMSM, after removing the zero-sequence component, can be expressed as follows [19, 51]:

$$\begin{bmatrix} v_{d1} \\ v_{q1} \\ v_{d3} \\ v_{q3} \end{bmatrix} = \begin{bmatrix} R_s & 0 & 0 & 0 \\ 0 & R_s & 0 & 0 \\ 0 & 0 & R_s & 0 \\ 0 & 0 & 0 & R_s \end{bmatrix} \cdot \begin{bmatrix} i_{d1} \\ i_{q1} \\ i_{d3} \\ i_{q3} \end{bmatrix} + \begin{bmatrix} \frac{d\lambda_{d1}}{dt} \\ \frac{d\lambda_{q1}}{dt} \\ \frac{d\lambda_{d3}}{dt} \\ \frac{d\lambda_{q3}}{dt} \end{bmatrix} - \omega \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & -3 & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda_{d1} \\ \lambda_{q1} \\ \lambda_{d3} \\ \lambda_{q3} \end{bmatrix} \quad (4)$$

where  $R_s$  is the stator resistance,  $\omega$  is the motor speed in electrical rad/sec,  $\lambda_{d1}, \lambda_{q1}$  are the fundamental DQ stator flux linkages and  $\lambda_{d3}, \lambda_{q3}$  are the 3<sup>rd</sup> harmonic DQ stator flux linkages. The DQ stator flux linkages take the following form:

$$\begin{aligned} \lambda_{d1} &= L_{d1}i_{d1} + L_{m13}i_{d3} + \lambda_{1m} \\ \lambda_{q1} &= L_{q1}i_{q1} + L_{m13}i_{q3} \\ \lambda_{d3} &= L_{d3}i_{d3} + L_{m13}i_{d1} + \lambda_{3m} \\ \lambda_{q3} &= L_{q3}i_{q3} + L_{m13}i_{q1} \end{aligned} \quad (5)$$

where  $L_{d1}$ ,  $L_{q1}$  are the fundamental direct and quadrature self-inductances,  $L_{m13}$  is the mutual inductance and  $\lambda_{1m}$  and  $\lambda_{3m}$  are the fundamental and 3<sup>rd</sup> harmonic components of the rotor PM flux linkages respectively.

The 5ph-IPMSM differential equation, Eq (4) can be rearranged to take the following formula:

$$D[I] = [LL][V] - [LR].[I] - \omega[L\lambda].[I] + \omega[LG].[I] \tag{6}$$

where D is the operator d/dt,  $[I] = [i_{d1} \ i_{q1} \ i_{d3} \ i_{q3}]^T$ ,  $[V] = [v_{d1} \ v_{q1} \ v_{d3} \ v_{q3}]^T$ ,

$$[LL] = \begin{bmatrix} \frac{-L_{d3}}{L_{m13}^2 - L_{d1}L_{d3}} & 0 & \frac{L_{m13}}{L_{m13}^2 - L_{d1}L_{d3}} & 0 \\ 0 & \frac{-L_{q3}}{L_{m13}^2 - L_{q1}L_{q3}} & 0 & \frac{L_{m13}}{L_{m13}^2 - L_{q1}L_{q3}} \\ \frac{L_{m13}}{L_{m13}^2 - L_{d1}L_{d3}} & 0 & \frac{-L_{d1}}{L_{m13}^2 - L_{d1}L_{d3}} & 0 \\ 0 & \frac{L_{m13}}{L_{m13}^2 - L_{q1}L_{q3}} & 0 & \frac{-L_{q1}}{L_{m13}^2 - L_{q1}L_{q3}} \end{bmatrix}$$

$$[LR] = \begin{bmatrix} \frac{-L_{d3}R_s}{L_{m13}^2 - L_{d1}L_{d3}} & 0 & \frac{L_{m13}R_s}{L_{m13}^2 - L_{d1}L_{d3}} & 0 \\ 0 & \frac{-L_{q3}R_s}{L_{m13}^2 - L_{q1}L_{q3}} & 0 & \frac{L_{m13}R_s}{L_{m13}^2 - L_{q1}L_{q3}} \\ \frac{L_{m13}R_s}{L_{m13}^2 - L_{d1}L_{d3}} & 0 & \frac{-L_{d1}R_s}{L_{m13}^2 - L_{d1}L_{d3}} & 0 \\ 0 & \frac{L_{m13}R_s}{L_{m13}^2 - L_{q1}L_{q3}} & 0 & \frac{-L_{q1}R_s}{L_{m13}^2 - L_{q1}L_{q3}} \end{bmatrix}$$

$$[LG] = \begin{bmatrix} 0 & \frac{3L_{m13}^2 - L_{d3}L_{q1}}{L_{m13}^2 - L_{d1}L_{d3}} & 0 & \frac{3L_{m13}L_{q3} - L_{d3}L_{m13}}{L_{m13}^2 - L_{d1}L_{d3}} \\ \frac{L_{d1}L_{q3} - 3L_{m13}^2}{L_{m13}^2 - L_{q1}L_{q3}} & 0 & \frac{L_{m13}L_{q3} - 3L_{d3}L_{m13}}{L_{m13}^2 - L_{q1}L_{q3}} & 0 \\ 0 & \frac{L_{m13}L_{q1} - 3L_{d1}L_{m13}}{L_{m13}^2 - L_{d1}L_{d3}} & 0 & \frac{L_{m13}^2 - 3L_{d1}L_{q3}}{L_{m13}^2 - L_{d1}L_{d3}} \\ \frac{3L_{m13}L_{q1} - L_{d1}L_{m13}}{L_{m13}^2 - L_{q1}L_{q3}} & 0 & \frac{3L_{d3}L_{q1} - L_{m13}^2}{L_{m13}^2 - L_{q1}L_{q3}} & 0 \end{bmatrix}$$

$$[L\lambda] = \begin{bmatrix} 0 \\ \frac{3\lambda_{3m}L_{m13} - \lambda_{1m}L_{q3}}{L_{m13}^2 - L_{q1}L_{q3}} \\ 0 \\ \frac{\lambda_{1m}L_{m13} - 3\lambda_{3m}L_{q1}}{L_{m13}^2 - L_{q1}L_{q3}} \end{bmatrix}$$

The motor torque equation can be represented as [19, 51]:

$$T_e = \frac{5p}{2^2} [\lambda_{d1} i_{q1} - \lambda_{q1} i_{d1} + 3\lambda_{d3} i_{q3} - 3\lambda_{q3} i_{d3}] \tag{7}$$

Substitution of Eq (5) into Eq (7) gives [19]:

$$T_e = \frac{5p}{2^2} [(L_{d1} - L_{q1}) i_{d1} i_{q1} + 2L_{m13} (i_{d1} i_{q3} - i_{q1} i_{d3}) + 3(L_{d3} - L_{q3}) i_{d3} i_{q3} + (\lambda_{1m} i_{q1} + 3\lambda_{3m} i_{q3})] \tag{8}$$

where p is the poles total number. When considering motor speed variations during the transient time interval, Eq (6) becomes nonlinear. As a result, the 5ph-IPMSM currents need to be numerically solved. To achieve this, we employ the 5ph-IPMSM mechanical equation, which is expressed as:

$$D\omega_m = \frac{T_e - T_l(\omega_m)}{J} \tag{9}$$

In the given equation, the motor speed is denoted by  $\omega_m$  in mechanical radians per second, J denotes the inertia, and  $T_l(\omega_m)$  is defined as follows:

$$T_l(\omega_m) = T_L + T_{fw} \tag{10}$$

Where  $T_L$  and  $T_{fw}$  are the load and friction and windage torques respectively.

### 2.3 Five-phase IPMSM maximum torque per ampere operating mode of operation model

To maximize efficiency, the 5ph-IPMSM operates at maximum torque per ampere, particularly for speeds up to the motor’s rated speed. This section derives the reference fundamental and 3<sup>rd</sup> harmonic DQ currents components.

When  $L_{m13}$  is disregarded, the torque equation for the fundamental component can be formulated using Eq (8) as follows:

$$T_{e1} = \frac{5p}{2^2} [(L_{d1} - L_{q1}) i_{d1} i_{q1} + \lambda_{1m} i_{q1}] \tag{11}$$

The reference fundamental direct current component is derived by differentiating the fundamental torque equation, Eq (11), w.r.t. the fundamental direct current and setting the result to zero, i.e.  $\frac{dT_{e1}}{di_{d1}} = 0$ . Consequently, the reference fundamental direct current component is determined as shown in the following equation:

$$i_{d1} = \frac{\lambda_{1m}}{2(L_{q1} - L_{d1})} - \sqrt{\frac{\lambda_{1m}^2}{4(L_{q1} - L_{d1})^2} + i_{q1}^2} \tag{12}$$

The reference fundamental quadrature current component can be derived from Eq (11) as follows:

$$i_{q1} = \frac{T_{e1}}{\frac{5p}{2^2} [(L_{d1} - L_{q1}) i_{d1} + \lambda_{1m}]} \tag{13}$$

The reference 3<sup>rd</sup> harmonic direct current component can be obtained from [52]:

$$i_{d3} = k\sqrt{i_{d1}^2 + i_{q1}^2} \sin\{3[\tan^{-1}\left(\frac{i_{d1}}{i_{q1}}\right)]\} \tag{14}$$

Also, the reference 3<sup>rd</sup> harmonic quadrature current component can be obtained from [52]:

$$i_{q3} = k\sqrt{i_{d1}^2 + i_{q1}^2} \cos\{3[\tan^{-1}\left(\frac{i_{d1}}{i_{q1}}\right)]\} \tag{15}$$

### 2.4 Five-phase IPMSM field weakening operating mode model

To extend the operating speed range beyond the rated speed, the 5ph-IPMSM will be operated in field weakening mode. The equations for the reference fundamental and 3<sup>rd</sup> harmonic DQ currents components are derived as follows.

The fundamental direct and quadrature steady-state voltages can be obtained from Eq (4) and neglecting the stator resistance voltage drops, we have:

The fundamental direct and quadrature steady-state voltages can be derived from Eq (4). By neglecting the stator resistance voltage drops, we obtain the following equations:

$$\begin{aligned} v_{d1} &= -\omega\lambda_{q1} \\ v_{q1} &= \omega\lambda_{d1} \end{aligned} \tag{16}$$

By substituting  $\lambda_{d1}$  and  $\lambda_{q1}$  from Eq (5) into Eq (16) and ignoring  $L_{m13}$ , we obtain:

$$\begin{aligned} v_{d1} &= -\omega L_q i_{q1} \\ v_{q1} &= \omega L_d i_{d1} \end{aligned} \tag{17}$$

The reference fundamental direct and quadrature components should satisfy the following equation to guarantee the maximum fundamental voltage,  $V_{m1}$ , of the IPMSM:

The reference fundamental direct and quadrature components must satisfy the following equation to ensure the maximum fundamental voltage,  $V_{m1}$ , of the 5ph-IPMSM:

$$V_{m1} = \sqrt{v_{d1}^2 + v_{q1}^2} \tag{18}$$

By substituting Eq (17) into Eq (18) and solving for the fundamental direct current component, we obtain:

$$i_{d1} = \frac{-\lambda_{1m}}{L_{d1}} + \frac{1}{L_{d1}} \sqrt{\frac{V_{1m}^2}{\omega^2} - L_{q1}^2 i_{q1}^2} \tag{19}$$

In this mode of operation, equations analogous to Eqs. (13), (14), and (15), expressed in terms of the fundamental direct current component provided in Eq (19), are utilized.

### 3. Model predictive control technique

The MPC consists of two primary components: the model of the plant and the optimizer. The core concept of MPC is to choose the optimal sequence of inputs for the plant utilizing predictions of its future behavior. These predictions are achieved using the model of the plant, which employs previous states to predict future states. At each discrete interval, the optimizer leverages the predicted states and the desired trajectory to address the optimization problem over the prediction horizon, thereby identifying the optimal set of inputs for future operations.



To successfully implement MPC, it is crucial to discretize the motor model. Thus, Eq (6) is converted into its discrete form using the Forward Euler approximation method. Consequently, the discrete model for the 5-phase IPMSM can be represented as follows:

$$[I(k+1)] = [I(k)] + T_s \{ [LL][V] - [LR].[I(k)] - \omega[L\lambda].[I(k)] + \omega[LG].[I(k)] \} \quad (20)$$

In this context,  $T_s$  denotes the sampling interval of the discretized system.

The MPC is specifically designed to minimize torque error as its primary objective. Since the electromagnetic torque equation for a five-phase IPMSM involves contributions from both the d-q axes currents, it is necessary to control these currents to reduce the error in torque. Consequently, the problem can be simplified by minimizing the error in current error instead of torque error. Therefore, the basic CF aimed at reducing current error can be formulated as follows:

The primary objective of the MPC is to reduce torque error. Given that the electromagnetic torque equation for a 5ph-IPMSM includes contributions from both the d-q axes currents, controlling these currents is essential to decrease torque error. Thus, the problem can be made simpler by focusing on reducing current error instead of torque error. Therefore, the basic cost function (CF) aimed at reducing current error can be formulated as follows:

$$C.F = [i_{d1r} - i_{d1}(k+1)]^2 + [i_{q1r} - i_{q1}(k+1)]^2 + [i_{d3r} - i_{d3}(k+1)]^2 + [i_{q3r} - i_{q3}(k+1)]^2 \quad (21)$$

Where  $i_{d1r}$ ,  $i_{q1r}$ ,  $i_{d3r}$ ,  $i_{q3r}$  represent the reference direct and quadrature fundamental and third harmonic currents. These currents are determined using either Eqs. (12), (13), (14), and (15) for motor operation in the region of constant torque, or Eqs. (19), (13), (14), and (15) motor operation in the region of constant power. The optimal inverter switching functions are chosen based on the minimum cost function.

## 4. Metahuristic optimization techniques

### 4.1 PI controller

The PI (proportional-integral) controller is commonly utilized in various industrial applications because of its simplicity, ease of implementation, and robust performance. Achieving optimal performance with a PI controller involves fine-tuning two key parameters: the proportional gain ( $k_p$ ) and the integral gain ( $k_i$ ). To optimize the performance of PI controllers, researchers have employed various optimization algorithms. In this study, four of the most recent optimization techniques are used for obtaining the optimum PI gains. These techniques are Transit Search (TS) [53], Honey Badger Algorithm (HBA) [54], Dwarf Mongoose (DM) [55], and Dandelion-Optimizer (DO) [56]. These techniques are compared with the RL utilizing TD3 algorithm. The goal is to minimize the error associated with the control of speed of a 5-phase IPMSM, to ensure the best possible performance.

The optimization process involves reducing the error  $e(t)$ , which is generated by comparing the reference speed with the actual speed, as well as the reference torque and actual torque, using four standard performance indicators: IAE, ISE, ITAE, and ITSE. This optimization is achieved through the utilization of the following equation, Eq (22), that accurately represents

the superior results obtained with the proposed control technique.

$$\left\{ \begin{array}{l} IAE = \int_0^{tss} |e(t)| \cdot dt \\ ISE = \int_0^{tss} e^2(t) \cdot dt \\ ITAE = \int_0^{tss} t \cdot |e(t)| \cdot dt \\ ITSE = \int_0^{tss} t \cdot e^2(t) \cdot dt \end{array} \right. \quad (22)$$

The optimization objective involves minimizing the steady-state time response (*tss*) and the error function *e*(*t*).

### 4.2 Objective

The optimization problem involves the simultaneous pursuit of two distinct objectives, allowing for a comprehensive optimization strategy tailored to address the specific aims of each target. The optimization focuses on minimizing the speed and torque errors of the individual controllers by fine-tuning their respective gain parameters.

There are two PI controllers—one for the regulation of speed and another for controlling the torque. Each controller has two adjustable gains, the proportional and integrals gains (*k<sub>p</sub>* and *k<sub>i</sub>*). Optimization must determine the optimal values for these four gain parameters, which are constrained within the range of [0, 300] for each gain. By optimizing these controller gains, the goal is to achieve high precision in both speed and torque regulation, thereby enhancing the overall system performance. The optimization strategy must navigate this multi-objective landscape to identify the set of gain values that best satisfies the conflicting targets of minimizing both speed and torque errors simultaneously.

$$Target = \min (error)$$

$$error = \int (T_{act} - T_{ref})^2 dt + \int (w_{act} - w_{ref})^2 dt \quad (23)$$

## 5. Reinforcement learning based TD3 approach

The TD3 approach is an advancement over the DDPG algorithm. TD3 addresses the function approximation error that can occur in DDPG. The TD3 algorithm significantly enhances both the learning speed and performance of DDPG across various challenging continuous control tasks. The algorithm of TD3 outperforms many state-of-the-art methods. Due to the simplicity of TD3 modifications, they can be easily integrated into any other actor-critic algorithm [57, 58]. In addition to this, the TD3 algorithm learns two Q-value functions and uses the minimum estimate during policy updates.

It does this by using twin critic networks (CN), delayed target network updates, and added exploration noise, which together help to stabilize the process of training and improve the efficiency of the learned policy. The goal of RL is to obtain the optimum policy  $\pi$  that makes the predicted reward to have maximum value which is achieved by tuning the parameter. This is typically achieved by renewing the parameter by employing the gradient  $\nabla_{\phi} J(\phi)$  RL employs

an actor-critic structure, where the policy (actor) is renewed according to the DPG algorithm. In case of large value of state space, the Q-value function  $Q(s, a)$  is nearly determined using a function approximator  $Q_{\theta}(s, a)$  with the aiding of a tuning parameter  $\theta$ . To maintain a stable learning objective, a frozen target network  $Q_{\theta'}(s, a)$  is used to maintain a stable learning goal  $y$  across many updates [57].

$$y = r + \gamma Q_{\theta'}(s', a'), \quad a' \leftarrow \pi_{\phi'}(s') \tag{24}$$

Actions are selected by the algorithm from a desired actor-network (AN)  $\pi_{\phi}(s)$  that is separated from the main AN. The weights of the target network are periodically renewed to accurately approximate the weights of the current network using a soft update rule [57]. This helps maintain a stable learning objective during the training process.

In actor-critic (A-C) methods, the current and desired networks may be very similar, leading to inaccurate value prediction. In order to address this, the algorithm uses a set of two actors  $(\pi_{\phi_1}, \pi_{\phi_2})$  and critics  $(Q_{\theta_1}, Q_{\theta_2})$  [57]. The actors are optimized according to their respective critics, but this can cause overestimation. To mitigate this, a trimmed double Q-learning approach is used, that selects the minimum of the two critic estimates as the target update [58]. The AN and CN are then updated according to this formulation.

$$y = r + \min_{i=1,2} Q_{\theta_i}(s', \pi_{\phi_1}(s')) \tag{25}$$

The algorithm uses a reward value  $r$  and a discount factor  $\gamma$  which determines the influence of previous reward values on next decisions. The value of discount factor  $\gamma$  ranges from 0 to 1 [57, 59]. To address the issue of deterministic policies overfitting sharp peaks in the value approximation, the algorithm adds a tiny amount of random noise to the desired policy, which is then trimmed to keep the objective within a limited range [57].

$$y = r + \min_{i=1,2} Q_{\theta_i}(s', \pi_{\phi_1}(s') + \epsilon) \quad \epsilon \approx \text{clip}(N(0, \sigma), -c, c) \tag{26}$$

The TD3 approach is presented as flowchart in Fig 1, as proposed by [57, 60]. Table 1 lists some of the key parameters used in the TD3 algorithm.

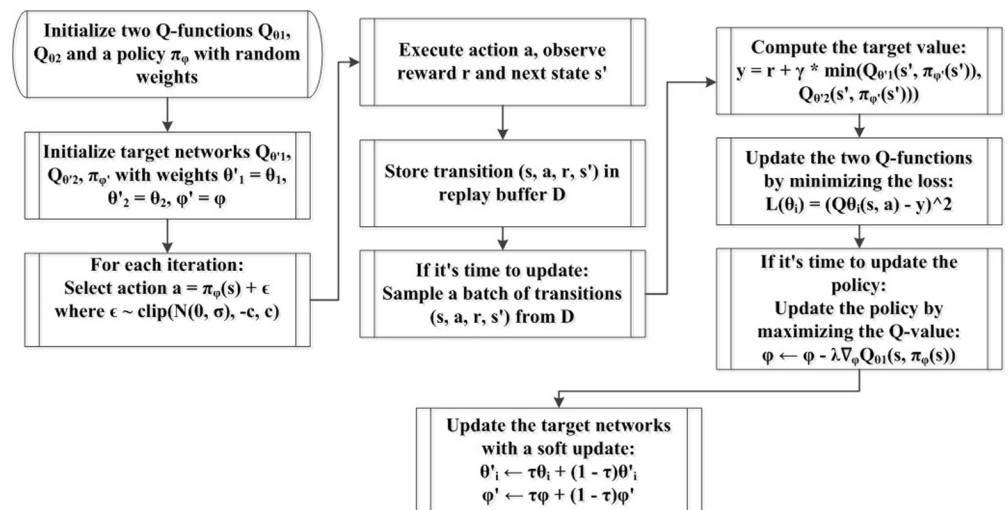


Fig 1. TD3 algorithm.

<https://doi.org/10.1371/journal.pone.0316326.g001>

**Table 1. Parameters for TD3 approach [60].**

Parameters	Explanations
$Q_{\theta_1}, Q_{\theta_2}$	two CNs, where $\theta_1$ and $\theta_2$ represent the respective weights of the CNs.
$Q_{\theta'_1}, Q_{\theta'_2}$	two target CNs, where $\theta'_1$ and $\theta'_2$ represent the respective weights of the target CNs.
$\pi_\varphi, \pi_{\varphi'}$	AN, $\pi_\varphi$ , and a target AN, $\pi_{\varphi'}$ , where $\varphi$ represents the weights of the AN and $\varphi'$ represents the weights of the target AN.
tuple (s,a,r,s')	The system transition tuples, where $s$ represents the current state, $a$ is the action taken, $r$ is the reward received, and $s'$ is the next state.
$\gamma$	A CN discount factor
$\theta'_1, \theta'_2$	The weight parameters are updated according to the next value calculation.
$\varphi'$	represent the next value for the AN.
$\tau$	represent the weights used for Polyak averaging.

<https://doi.org/10.1371/journal.pone.0316326.t001>

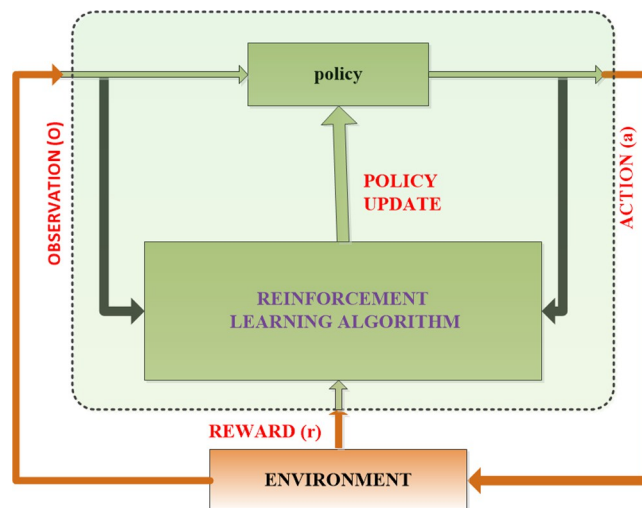
The block diagram depicted in Fig 2 illustrates the RL agent utilizing the TD3 algorithm for process control. In this context, the goal of obtaining the optimum policy for the RL-TD3 agent may be interpreted as determining the appropriate command signals for the 5ph-PMSM in order to enhance a given effectiveness metric.

The simulation environment of Matlab Simulink is utilized to implement the TD3 algorithm. The objective is to train an optimum two-stage PI controller for regulating the speed and torque of a 5ph-IPMSM in the control system setup. The process of learning generates the optimum tuning parameters for the two PI controllers, which can effectively address the regulation challenges simulated in the system environment.

This setup can be viewed as analogous to the control of an industrial process, where the observed inputs are the speed error and its integral, as well as the torque, as shown in Fig 3 while the outputs represent the reference signals, and the enhancement objective is framed as a reward function.

The reward function is defined by the subsequent expression:

$$r(t) = -(w_{c1} \int_0^t e_1(t)^2 dt + w_{c2} \int_0^t e_2(t)^2 dt + w_{c3} u(t)^2) \tag{27}$$



**Fig 2. RL block diagram.**

<https://doi.org/10.1371/journal.pone.0316326.g002>

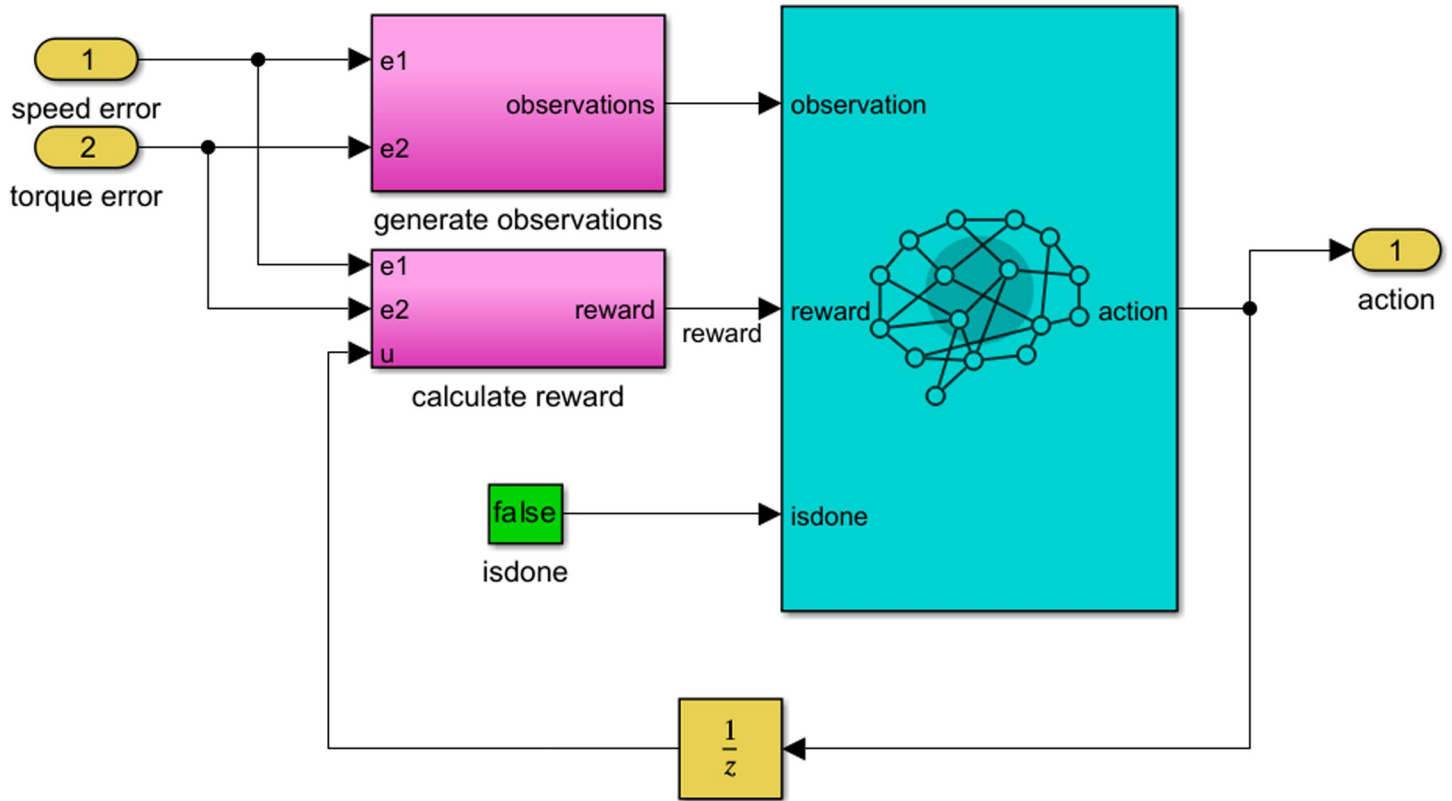


Fig 3. Matlabsimulink block diagram for the TD3 approach.

<https://doi.org/10.1371/journal.pone.0316326.g003>

The weight coefficients  $w_{c1}, w_{c2}$ , and  $w_{c3}$  are used to balance the reduction of the reference error and the control signal. The error signals  $e_1(t), e_2(t)$  represent the observed states from the environment of the control system, and the control signal  $u(t)$  corresponds to the action of the actor.

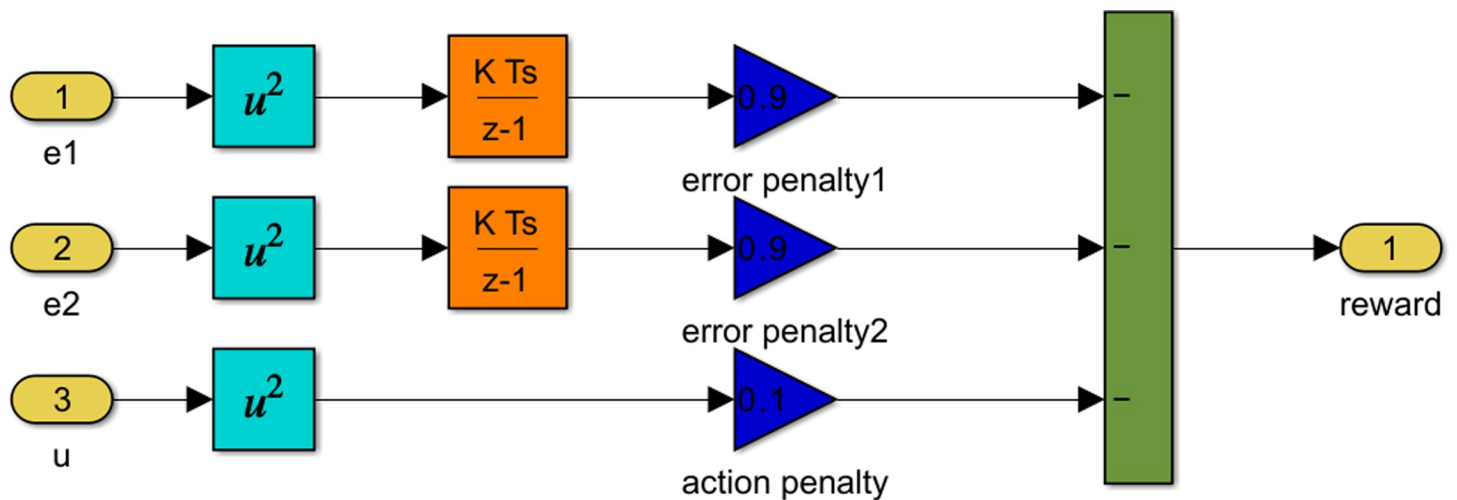


Fig 4. Matlab Simulink block diagram of the reward equation.

<https://doi.org/10.1371/journal.pone.0316326.g004>

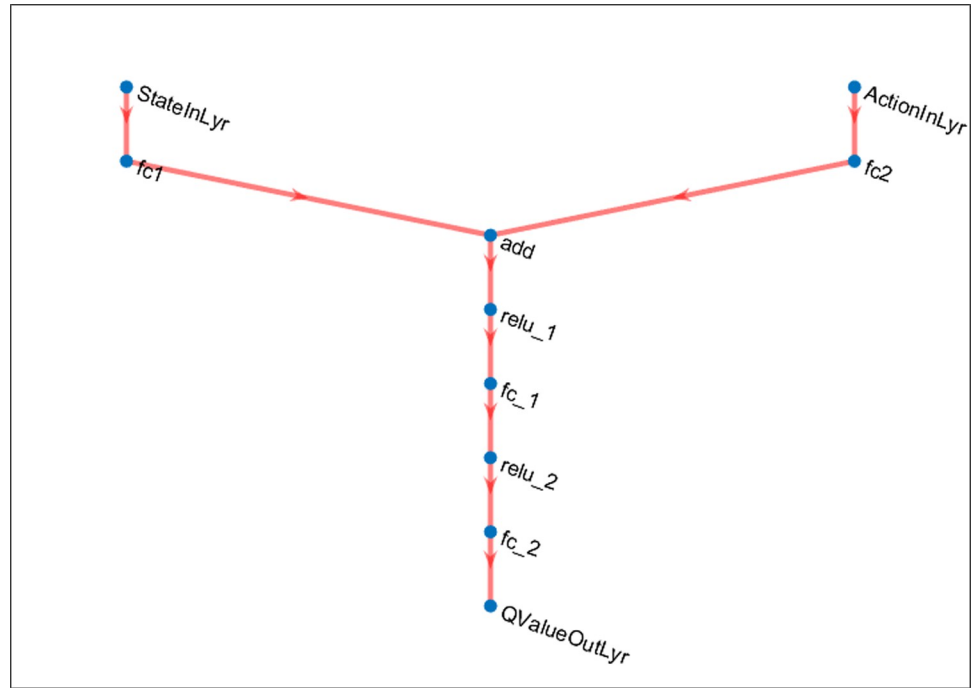


Fig 5. The architecture of CNs utilized in the drive system under consideration.

<https://doi.org/10.1371/journal.pone.0316326.g005>

The optimal control reward equation can be represented by:

$$R(z) = -\mathcal{W}_{c1} \frac{T_s}{z-1} e_1(z)^2 - \mathcal{W}_{c2} \frac{T_s}{z-1} e_2(z)^2 - \mathcal{W}_{c3} u(t)^2 \tag{28}$$

Fig 4 shows the Matlab Simulink block diagram of reward equation, Eq (28). In this figure the weights,  $\mathcal{W}_{c1}$ ,  $\mathcal{W}_{c2}$  and  $\mathcal{W}_{c3}$  are taken to be 0.9, 0.9 and 0.1 respectively [7].

The AN and CN, with their corresponding reference networks, were set up in the MATLAB environment to achieve RL agents using the TD3 approach. The CN utilize the

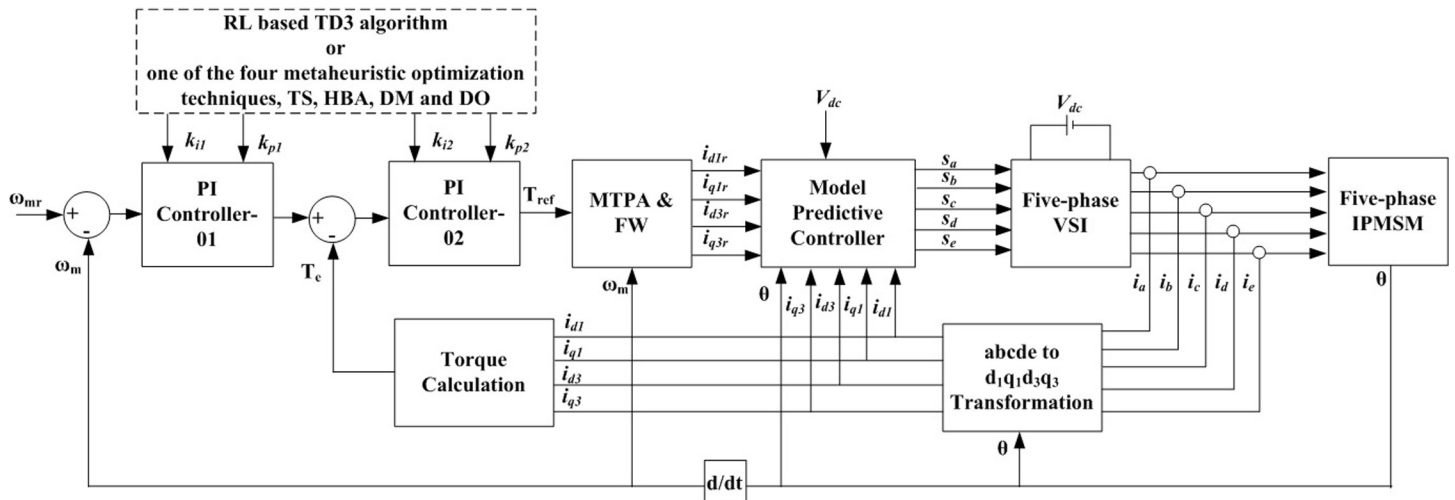


Fig 6. Proposed 5ph-IPMSM control system.

<https://doi.org/10.1371/journal.pone.0316326.g006>

**Table 2. Five-phase IPMSM parameters.**

Parameter	Value
poles (p)	4 poles
Rated power	12 kW
Rated Speed	1800 rpm
Maximum speed	5400 rpm
Rated Torque	63.662 Nm
$R_s$	0.389 $\Omega$
$L_{d1}$	2.7 mH
$L_{q1}$	9.6 mH
$L_{d3}$	1.1 mH
$L_{q3}$	2 mH
$L_{m13}$	0 mH
$\lambda_{1m}$	0.11 WbT
$\lambda_{3m}$	0.0012 WbT
Moment of inertia	0.0036 kg.m <sup>2</sup>
Connection	Star

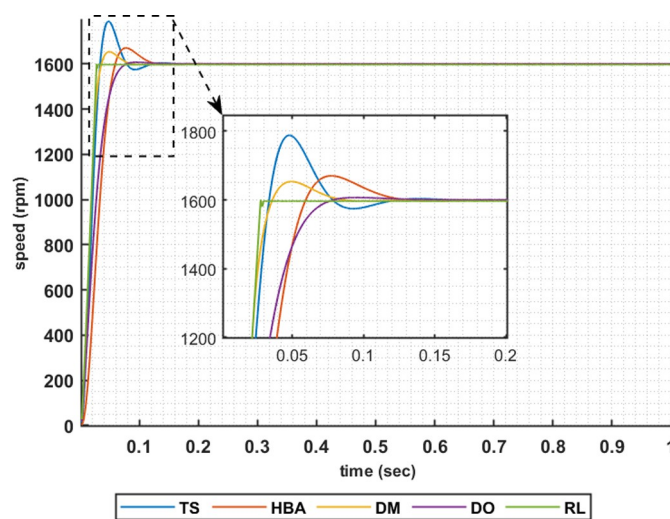
<https://doi.org/10.1371/journal.pone.0316326.t002>

action ( $a$ ) and observed state ( $s$ ) as inputs, and function as an approximation to approximate the quality value  $Q(s, a)$ . The arrangement of CNs utilized in the drive system under consideration is given in Fig 5.

The CNs consist of three layers:

- i. fully connected (fc) input layers for the action and state inputs,
- ii. a fc common path layer,
- iii. a fc output layer with one neuron representing  $Q(s, a)$ .

The input and common path layers use ReLU activation. The actor networks take the observed state ( $s$ ) as input and output is the action ( $a$ ). The TD3 algorithm trains the AN and CN to enhance the reward, which is the optimum control goal [61].

**Fig 7. The motor speed responses when IAE error criteria is used.**

<https://doi.org/10.1371/journal.pone.0316326.g007>

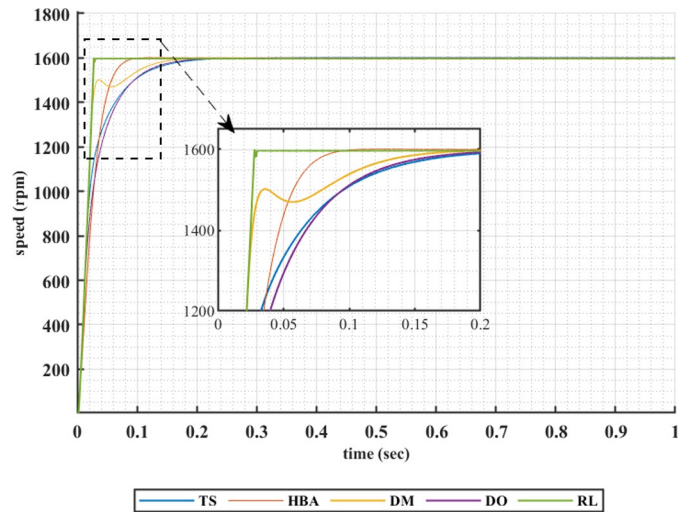


Fig 8. The motor speed responses when ISE error criteria is used.

<https://doi.org/10.1371/journal.pone.0316326.g008>

The gains of the PI controllers  $k_{i1}$ ,  $k_{p1}$ ,  $k_{i2}$  and  $k_{p2}$  can be found using the Matlab functions: *actor = getActor (agent)* and *parameters = getLearnableParameters (actor)*.

### 6. Proposed control methodology

Fig 6 illustrates the proposed control system, which comprises several blocks. The rotor position is measured at a specific load torque and reference speed of the 5ph-IPMSM to determine the actual motor speed, as shown in Fig 6. The speed error is processed by a primary PI controller to generate the reference torque. The motor currents of the 5-ph IPMSM are measured and transformed into the DQ frame of reference using the “abcde to d1q1d3q3 Transformation” block in Fig 6. This block represents Eq (2). The resulting fundamental and 3<sup>rd</sup> direct and quadrature currents are used to calculate the electromagnetic torque via the “Torque Calculation” block, which corresponds to Eq (8).

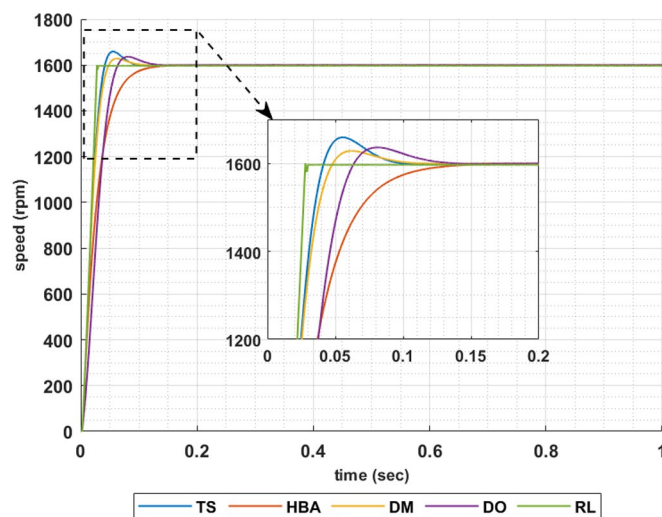


Fig 9. The motor speed responses when ITAE error criteria is used.

<https://doi.org/10.1371/journal.pone.0316326.g009>



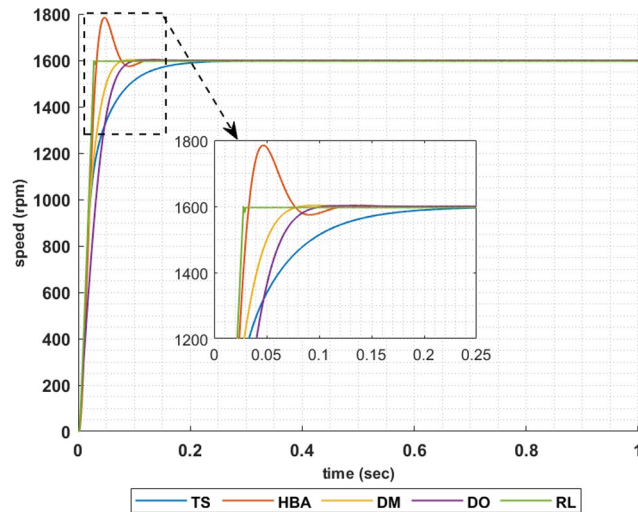


Fig 10. The motor speed responses when ITSE error criteria is used.

<https://doi.org/10.1371/journal.pone.0316326.g010>

The gains of the primary and secondary PI controllers ( $k_{i1}$ ,  $k_{p1}$ ,  $k_{i2}$ , and  $k_{p2}$ ) are determined using one of four metaheuristic optimization techniques (TS, HBA, DM, and DO) or the presented RL-based TD3 algorithm. The torque error is processed by the secondary PI controller, “PI Controller-02,” to obtain the corrected reference torque,  $T_{ref}$ . The motor speed,  $\omega_m$ , and  $T_{ref}$  are used to derive the reference DQ currents from the “MTPA & FW” block in Fig 6. This block represents Eqs (12), (13), (14), and (15) for MTPA operation when the motor speed is not greater than the rated speed, and Eqs (19), (13), (14), and (15) for FW operation when the motor speed exceeds the rated speed.

The reference fundamental and 3<sup>rd</sup> harmonic DQ currents, along with the rotor position angle and DC voltage, are used in the MPC, Eq (20), to generate the inverter gating signals that

Table 3. Summary of the comparison between the RL based TD3 and the metaheuristic optimization techniques at speed 1600 rpm.

Technique	Error criteria	kp1	ki1	kp2	ki2	Settling Time (s)	Rise Time(s)	Max. time (s)	Overshoot Percentage
Transit Search	IAE	0.323	76.0253	149.504	108.5812	0.0723	0.0222	0.048	11.7962
	ISE	0.7304	18.9004	11.908	0.362	0.3476	0.0688	0.5302	0.0888
	ITSE	0.6456	19.8709	31.5824	9.1463	0.1459	0.067	0.5245	<b>0.0121</b>
	ITAE	0.4315	52.7312	23.4413	22.9367	0.0729	0.0266	0.0552	3.7238
Honey Badger Algorithm	IAE	0.309237	44.0829	41.9761	181.3634	0.1022	0.0364	0.0775	4.4114
	ISE	0.441908	34.8417	51.007	0	0.073	0.0443	0.1284	0.0631
	ITSE	0.335703	74.6599	170.868	60.38459	0.0706	<b>0.0217</b>	0.0468	11.4977
	ITAE	0.505046	27.6241	10.3486	25.0843	0.0954	0.0527	0.5336	0.0205
Dwarf Mongoose	IAE	0.40047	61.941	125.741	291.6633	0.0657	0.0226	0.0512	3.476
	ISE	0.641088	26.9701	33.325	57.3394	0.1222	0.0229	<b>0.0273</b>	0.5449
	ITSE	0.441913	41.2805	141.442	193.1191	0.0626	0.0385	0.096	0.2092
	ITAE	0.449873	47.7301	98.6211	8.21241	0.0439	0.029	1.7479	0.0631
Dandelion-Optimizer	ISE	0.549676	21.9602	14.511	37.1645	0.1405	0.0708	0.596	0.0164
	IAE	0.413013	37.3398	127.584	5.874622	0.065	0.0412	0.0941	0.476
	ITSE	0.392725	34.0375	103.825	146.0474	0.0787	0.0507	0.1245	0.1395
	ITAE	0.345308	42.0511	143.562	143.2322	0.0901	0.04	0.0813	2.2936
Reinforcement Learning		0.429039	44.9862	622.213	84.75749	<b>0.0201</b>	0.0274	0.028	0.0272

<https://doi.org/10.1371/journal.pone.0316326.t003>

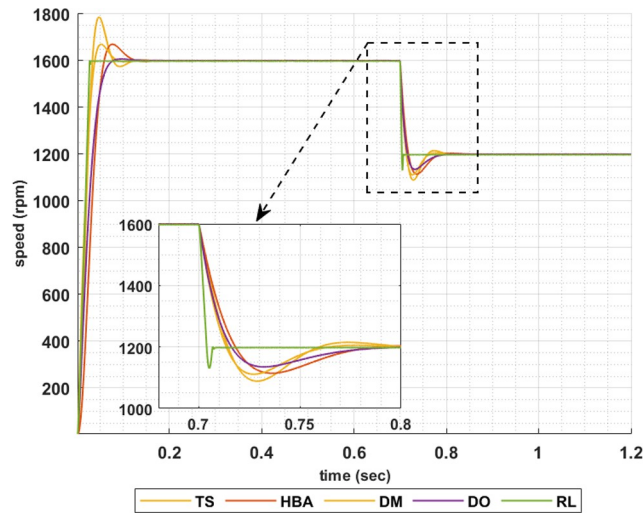


Fig 11. The motor speed responses when IAE error criteria is used.

<https://doi.org/10.1371/journal.pone.0316326.g011>

satisfy the cost function given in Eq (21) to minimize torque ripples and thus minimizing ripples in the motor speed.

### 7. Results

Several sets of results are obtained to prove the correctness of the suggested control methodology for the 5ph-IPMSM. The 5ph-IPMSM parameters are given in Table 2 [35].

The following set of results are obtained when the motor is driving a constant load torque whose value equals to the motor rated torque, 63.662 Nm, and the reference speed is 1600 rpm. In this operating condition the motor is operated at maximum torque per ampere. Figs 7–10 show the speed responses of the 5ph-IPMSM when the PI controllers gains are obtained using the metaheuristic optimization techniques (TS, HBA, DM, and DO) and the presented RL based TD3 algorithm when the four types of error criteria (IAE, ISE, ITAE and ITSE) are used in the optimization techniques respectively. From these figures it can be noticed that the

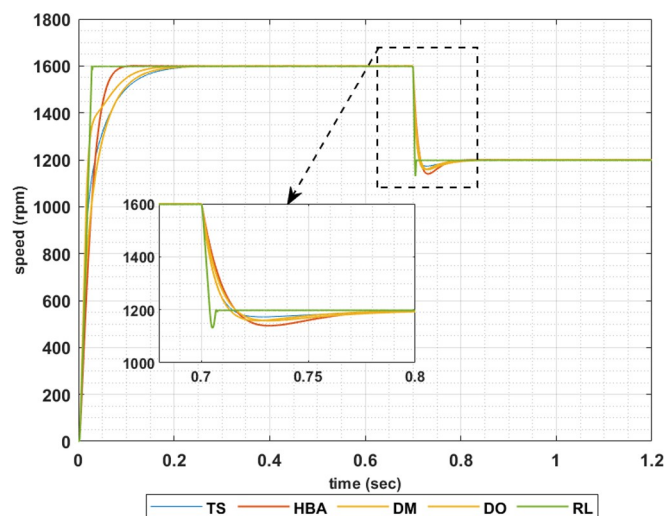


Fig 12. The motor speed responses when ISE error criteria is used.

<https://doi.org/10.1371/journal.pone.0316326.g012>

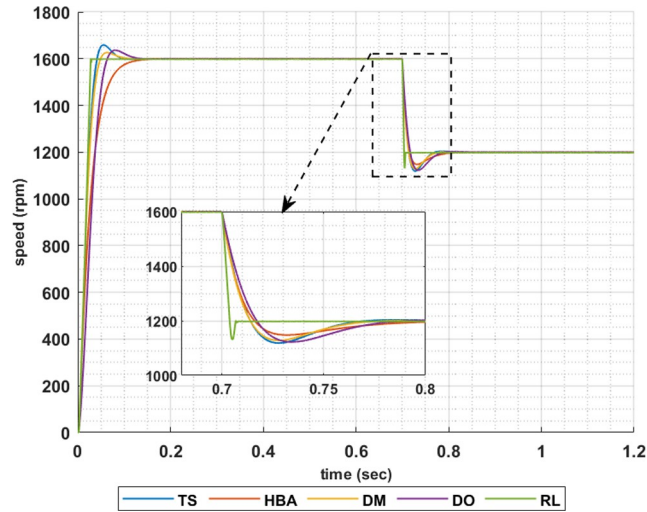


Fig 13. The motor speed responses when ITAE error criteria is used.

<https://doi.org/10.1371/journal.pone.0316326.g013>

speed response of the 5ph-IPMSM when the presented RL based TD3 algorithm is used, to obtain PI controllers gain, has the fastest time response and relatively low overshoot.

Table 3 summarizes the values of the PI controllers gains, settling time, rise time, maximum time and overshoot percentage for the different speed responses shown in Figs 7–10 obtained using the presented RL based TD3 algorithm and the different optimization techniques with the different error criteria. From this table, it can be noticed that the suggested RL based TD3 results in the minimum settling time and relatively low values for the rise time, max time and overshoot percentage. This proves the correctness of the suggested control methodology in improving the motor 5p-IPMSM speed response in the constant torque region.

To show the correctness of the presented control methodology in case of sudden change in the reference speed, the above results are obtained when the desired speed is suddenly varied from 1600 rpm to 1200 rpm. Figs 11–14 show the speed responses for the sudden change in

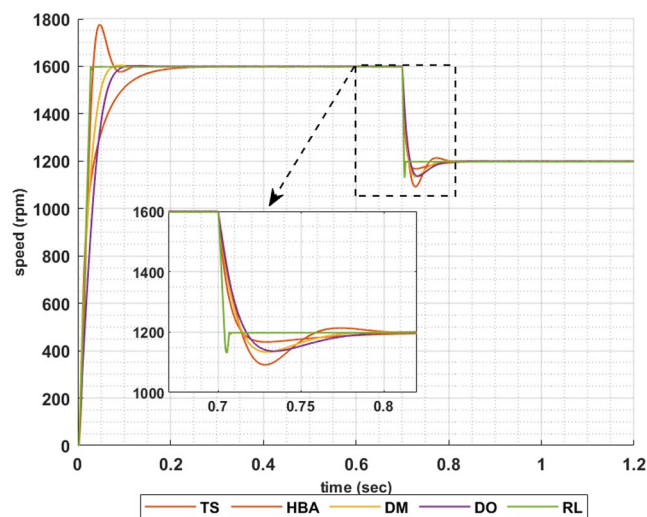


Fig 14. The motor speed responses when ITSE error criteria is used.

<https://doi.org/10.1371/journal.pone.0316326.g014>

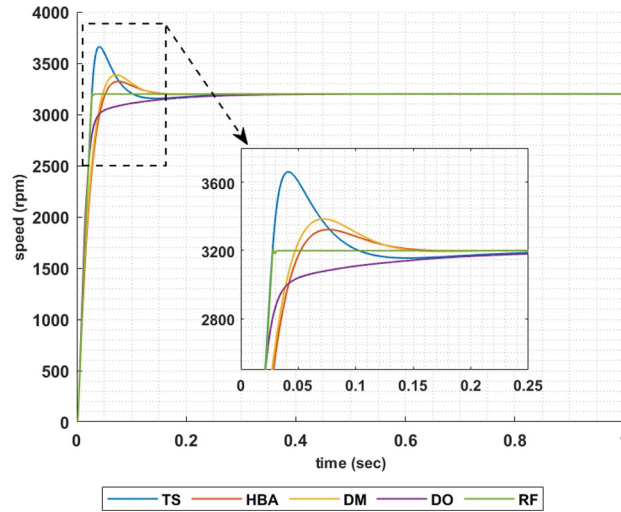


Fig 15. The motor speed responses when IAE error criteria is used.

<https://doi.org/10.1371/journal.pone.0316326.g015>

the motor reference speed using the metaheuristic optimization techniques and the presented RL based TD3 algorithm for the four types of error criteria. From these figures it can be noticed that the speed response of the 5ph-IPMSM when the presented RL based TD3 algorithm is used, also, has the fastest time response and relatively low overshoot.

Another set of results are obtained when the motor is driving a constant load torque whose value equals to 35.8 Nm, and the reference speed is 3200 rpm. In this operating condition the motor is operated in field weakening mode of operation, i.e constant power region. Figs 15–18 show the speed responses of the 5ph-IPMSM when the PI controllers gains are given using the metaheuristic optimization techniques and the presented RL based TD3 approach when the four types of error criteria are used in the optimization techniques. From these figures it can be noticed that the speed response of the 5ph-IPMSM when the presented RL based TD3

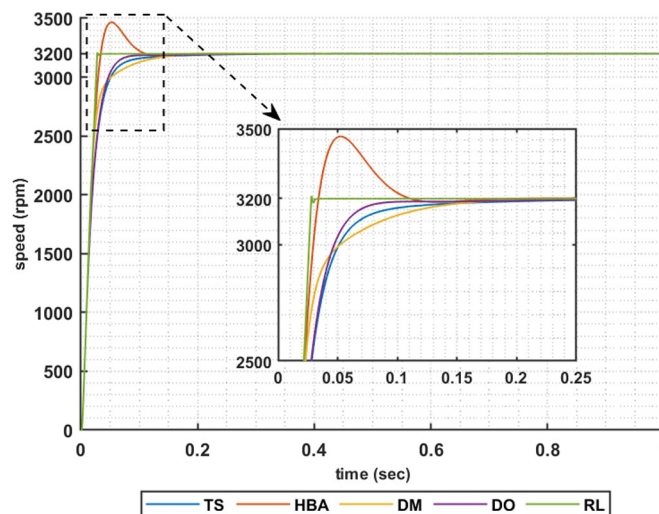


Fig 16. The motor speed responses when ISE error criteria is used.

<https://doi.org/10.1371/journal.pone.0316326.g016>

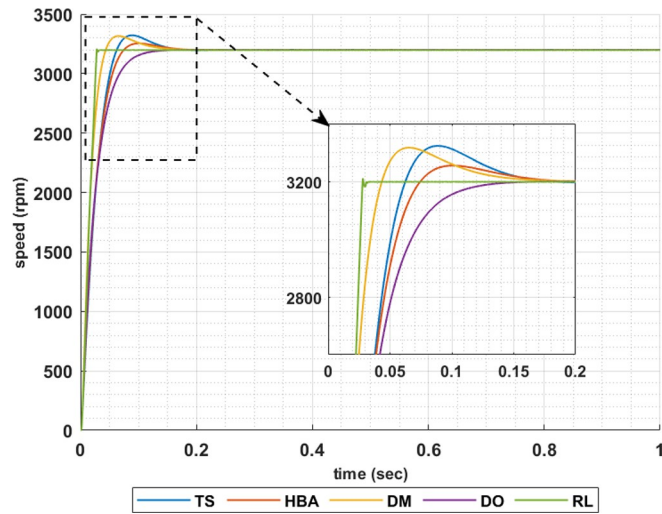


Fig 17. The motor speed responses when ITAE error criteria is used.

<https://doi.org/10.1371/journal.pone.0316326.g017>

algorithm is used, to obtain PI controllers gain, has the fastest time response and relatively low overshoot.

Table 4 summarizes the values of the PI controllers gains, settling time, rise time, maximum time and overshoot percentage for the different speed responses shown in Figs 15–18 obtained using the presented RL based TD3 algorithm and the different optimization techniques with the different error criteria when the motor is operated in constant power region. It can be noticed from Table 4 that the proposed RL based TD3 gives the minimum values for the settling time and maximum time and relatively low values for the rise time and overshoot percentage. This again proves the correctness of the suggested control methodology in improving the motor 5p-IPMSM speed response in the constant power region.

The above results are obtained when the desired speed is suddenly varied from 3200 rpm at 35 Nm load torque to 4000 rpm at 28.6 Nm load torque, constant power region. Figs 19–22 show the speed responses for the sudden change in the motor reference speed with the

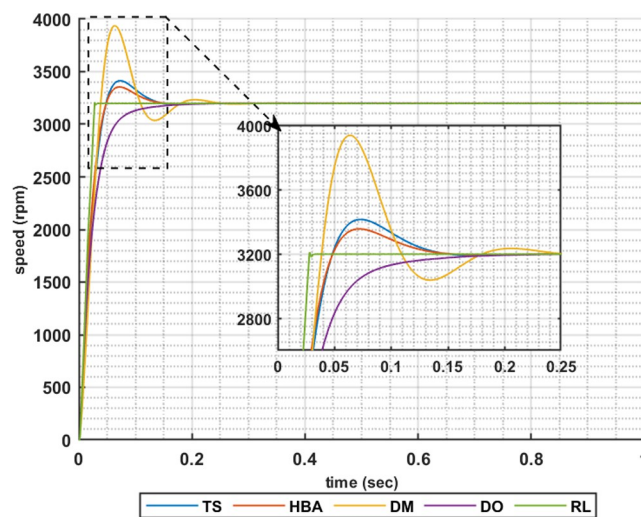


Fig 18. The motor speed responses when ITSE error criteria is used.

<https://doi.org/10.1371/journal.pone.0316326.g018>

**Table 4. Summary of the comparison between the RL based TD3 and the metaheuristic optimization techniques at speed 3200 rpm.**

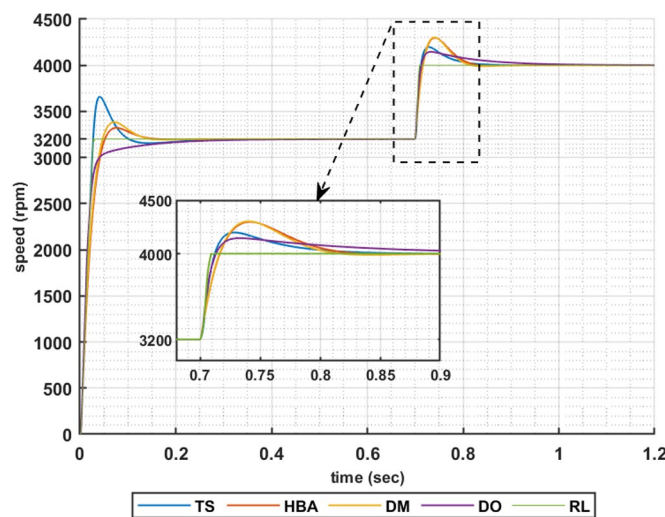
Technique	Error criteria	kp1	ki1	kp2	ki2	Settling Time (s)	Rise Time(s)	Max. time (s)	Overshoot Percentage
Transit Search	IAE	0.5163	20.3512	74.0624	21.8978	0.0694	<b>0.0171</b>	0.0392	6.1898
	ISE	0.3243	15.5619	28.2828	10.0919	0.094	0.0406	0.5532	0.0165
	ITSE	0.2319	29.2639	200	15.4351	0.1177	0.0332	0.0737	6.7256
	ITAE	0.2225	23.2968	11.1122	51.3965	0.1232	0.0412	0.089	3.7255
Honey Badger Algorithm	IAE	0.24717	26.27137	152.6986	40.78611	0.1084	0.0336	0.0761	3.8731
	ISE	0.28545	31.4349	49.3307	49.5886	0.094	0.0247	0.0582	6.0532
	ITSE	0.24746	27.6937	40.4038	47.3758	0.1093	0.0322	0.073	4.816
	ITAE	0.23101	20.807	40.3527	24.0813	0.0665	0.0446	0.1022	1.7359
Dwarf Mongoose	IAE	0.23817	30.74569	233.662	159.5167	0.1128	0.0319	0.072	5.829
	ISE	0.31232	17.13106	105.5177	128.8104	0.0983	0.0403	0.5237	0.0082
	ITSE	0.16397	54.8639	39.9551	24.7158	0.1621	0.0279	0.0633	22.6599
	ITAE	0.32712	24.3881	61.3568	1.64838	0.0895	0.0274	0.0687	2.6998
Dandelion-Optimizer	IAE	0.55772	9.49771	107.7051	0.03886	0.1657	0.0276	0.5936	<b>0</b>
	ISE	0.3086	17.4093	42.1558	13.2945	0.0732	0.038	0.5829	0.0097
	ITSE	0.26063	16.2196	9.54845	59.3708	0.104	0.0494	0.4713	0.0091
	ITAE	0.26049	16.0258	17.9417	10.0006	0.0941	0.0524	0.2483	0.0445
Reinforcement Learning		0.20445	14.5061	88.9718	20.7561	<b>0.0272</b>	0.0205	<b>0.0281</b>	0.329

<https://doi.org/10.1371/journal.pone.0316326.t004>

corresponding load torque utilizing the metaheuristic optimization techniques and the presented RL based TD3 algorithm for the four types of error criteria. From these figures it can be noticed that the speed response of the 5ph-IPMSM when the presented RL based TD3 algorithm is used, also, has the superior speed response.

### 8. Conclusion

In this study, a new RL control algorithm based TD3 algorithm is suggested to obtain gains of two cascaded PI controllers in a 5ph-IPMSM drive system. The purpose of this algorithm is to optimize the 5ph-IPMSM speed response either in constant torque region or constant power region. The PI controllers gains are also obtained using four recent metaheuristic optimization



**Fig 19. The motor speed responses when IAE error criteria is used.**

<https://doi.org/10.1371/journal.pone.0316326.g019>

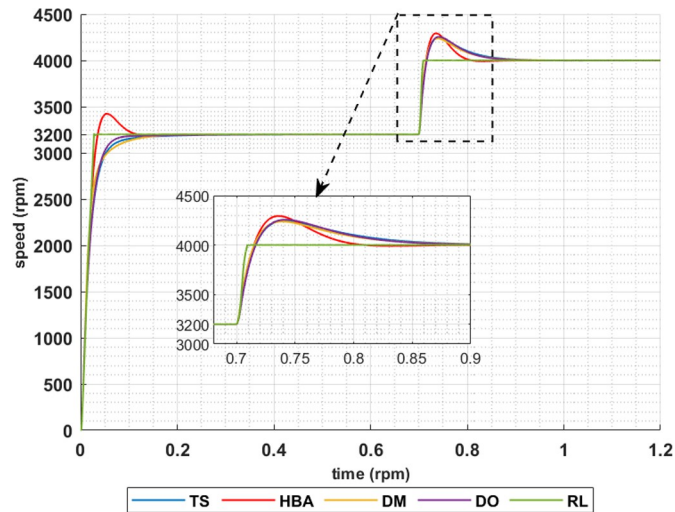


Fig 20. The motor speed responses when ISE error criteria is used.

<https://doi.org/10.1371/journal.pone.0316326.g020>

techniques to be compared with the proposed algorithm. The most recent metaheuristic optimization techniques used are Transit Search, Honey Badger Algorithm, Dwarf Mongoose, and Dandelion-Optimizer optimization techniques. MATLAB Simulink package is utilized to obtain simulation results to validate the proposed algorithm. It can be concluded from the results that the suggested control RL algorithm based TD3 algorithm results in improved motor speed response compared with the metaheuristic optimization techniques with the fastest response and with relatively lower overshoot either when the 5ph-IPMSM is operated either in the constant torque region or the constant power region. However, the MHOs are more easily to be implemented and have more simple computations compared with the RL. As a future work, the experimental implementation of the proposed methodology to be achieved. In addition to this, the proposed control methodology of the drive system can be investigated with the utilization in the electric vehicles, robotics and renewable energy systems.

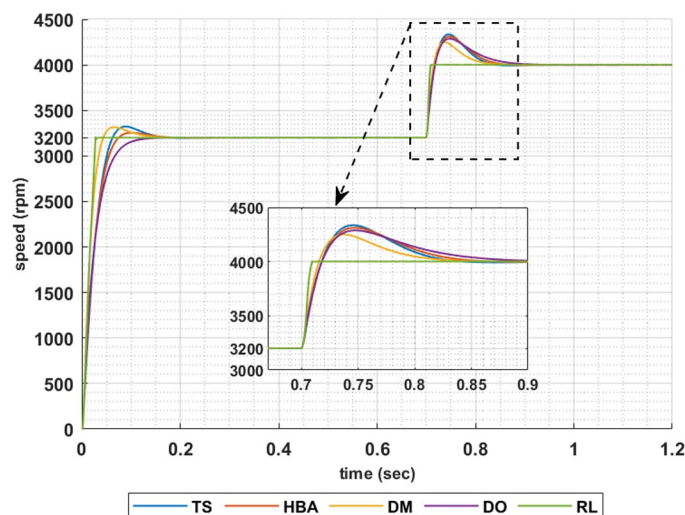


Fig 21. The motor speed responses when ITAE error criteria is used.

<https://doi.org/10.1371/journal.pone.0316326.g021>

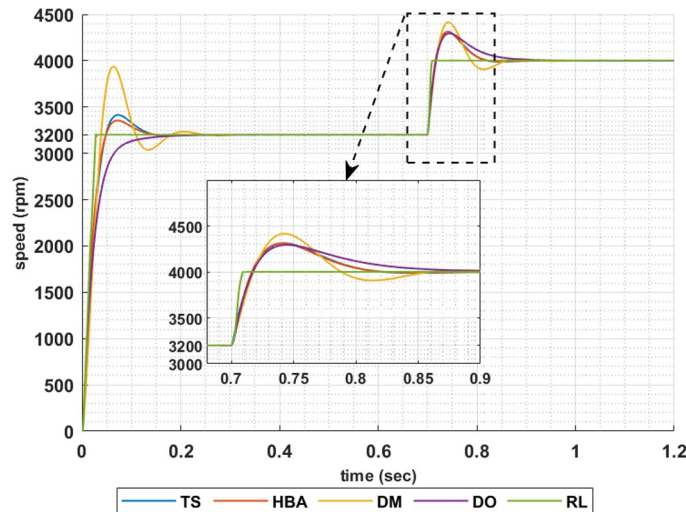


Fig 22. The motor speed responses when ITSE error criteria is used.

<https://doi.org/10.1371/journal.pone.0316326.g022>

## Author Contributions

**Conceptualization:** Ahmed M. Hassan.

**Formal analysis:** Ahmed M. Hassan, Hani Attar, Mohamed Eladly Metwally.

**Funding acquisition:** Jafar Ababneh, Hani Attar, Tamer Shamseldin.

**Investigation:** Ahmed M. Hassan, Jafar Ababneh, Tamer Shamseldin, Ahmed Abdelbaset.

**Methodology:** Ahmed M. Hassan, Mohamed Eladly Metwally.

**Resources:** Ahmed Abdelbaset.

**Software:** Mohamed Eladly Metwally.

**Visualization:** Jafar Ababneh, Hani Attar, Tamer Shamseldin, Ahmed Abdelbaset, Mohamed Eladly Metwally.

**Writing – original draft:** Ahmed M. Hassan.

**Writing – review & editing:** Ahmed M. Hassan, Mohamed Eladly Metwally.

## References

1. Khanh1 PQ, Anh HPH. Hybrid Optimal Fuzzy Jaya Technique for Advanced PMSM Driving Control. *Electrical Engineering*. 2023; 105, 3629–3646. <https://doi.org/10.1007/s00202-023-01911-6>.
2. Mercorelli P. Control of Permanent Magnet Synchronous Motors for Track Applications. *Electronics*. 2023; 12 (15), 1–20. <https://doi.org/10.3390/electronics12153285>.
3. Boumegouas MKB, Ilten E, Kouzi1 K, Demirtas M, M'hamed B. Application of a Novel Synergetic Observer for PMSM in Electrical Vehicle. *Electrical Engineering*; 2024. <https://doi.org/10.1007/s00202-024-02297-9>
4. Long L, Sun T and Liang J. Five Phase Permanent Magnet Synchronous Motor Decoupled Model with Dual Frame Frequency Adaptive Flux Observer. *Elsevier, Energy Reports*. 2020; 6, 1403–1408. <https://doi.org/10.1016/j.egy.2020.11.010>.
5. Tian B, Lu R and Hu J. Single Line/Phase Open Fault-Tolerant Decoupling Control of a Five-Phase Permanent Magnet Synchronous Motor under Different Stator Connections. *Energies*. 2022; 15 (9), 1–18. <https://doi.org/10.3390/en15093366>.



6. Dharmasena S, Choi S. Model Predictive Control of Five-Phase Permanent Magnet Assisted Synchronous Reluctance Motor. 2019 IEEE Applied Power Electronics Conference and Exposition (APEC), Anaheim, CA, USA. 2019; 1885–1890. <https://doi.org/10.1109/APEC.2019.8722267>
7. Tufenkci S, Alagoz BB, Kavuran G, Yeroglu C, Herencsar N, Mahata S. A Theoretical Demonstration for Reinforcement Learning of PI Control Dynamics for Optimal Speed Control of DC Motors by using Twin Delay Deep Deterministic Policy Gradient Algorithm. *Expert Systems With Applications*. 2023; 213, 1–16. <https://doi.org/10.1016/j.eswa.2022.119192>.
8. Yu F, Zhang X and Wang S. Five-phase Permanent Magnet Synchronous Motor Vector Control Based on Harmonic Eliminating Space Vector Modulation. 2005 International Conference on Electrical Machines and Systems, Nanjing, China. 2005; 1, 392–396. <https://doi.org/10.1109/ICEMS.2005.202554>
9. Hosseyni A, Trabelsi R, Mimouni MF and Iqbal A. Vector controlled five-phase permanent magnet synchronous motor drive. 2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE), Istanbul, Turkey. 2014; 2122–2127. <https://doi.org/10.1109/ISIE.2014.6864945>
10. Kamel T, Abdelkader D and Said B. Vector Control of Five-Phase Permanent Magnet Synchronous Motor Drive. 2015 4th International Conference on Electrical Engineering (ICEE), Boumerdes, Algeria. 2015; 1–4. <https://doi.org/10.1109/INTEE.2015.7416853>
11. Zong ZL, Wang K and Zhang JY. Control Strategy of Five-Phase PMSM Utilizing Third Harmonic Current to Improve Output Torque. 2017 Chinese Automation Congress (CAC), Jinan, China. 2017; 6112–6117. <https://doi.org/10.1109/CAC.2017.8243878>
12. Parsa L and Toliyat HA. Sensorless Direct Torque Control of Five-phase Interior Permanent Magnet Motor Drives. *IEEE Transactions on Industry Applications*. 2007; 43 (4), 2007. <https://doi.org/10.1016/j.jestch.2021.02.004>.
13. Guo L, and Parsa L. Model Reference Adaptive Control of Five-phase IPM Motors Based on Neural Network. 2011 IEEE International Electric Machines & Drives Conference (IEMDC), Niagara Falls, ON, Canada. 2011; 563–568. <https://doi.org/10.1109/IEMDC.2011.5994871>
14. Li J, Du B, Zhao T, Cheng Y, and Cui S. Sensorless Control of Five-Phase Permanent-Magnet Synchronous Motor Based on Third-Harmonic Space. *IEEE Transactions on Industrial Electronics*. 2022; 69 (8), 7685–7695. <https://doi.org/10.1109/TIE.2021.3111579>
15. Mehedi F, Nezli L, Mahmoudi MO and Taleb R. Robust Speed Control of Five-Phase Permanent Magnet Synchronous Motor using Super-Twisting Sliding Mode Control. *J. Ren. Energies*. 2017; 20 (4), 649–657. <https://doi.org/10.54966/jreen.v20i4.657>
16. Cao B, Grainger BM, Wang X, Zou Y, Reed GR and Mao ZH. Direct Torque Model Predictive Control of a Poly-Phase Permanent Magnet Synchronous Motor with Current Harmonic Suppression and Loss Reduction. 2018 IEEE Applied Power Electronics Conference and Exposition (APEC), San Antonio, TX, USA. 2018; 2460–2464. <https://doi.org/10.1109/APEC.2018.8341362>
17. Han X, Wang H, Zhang Z, Fan Y and Wang W. An Improved Finite Control Set Model Predictive Current Control Strategy for Five-Phase PMSMs. 2019 4th Asia Conference on Power and Electrical Engineering, ACPEE 2019, IOP Conf. Series: Materials Science and Engineering. 2019; 486, 1–6. <https://doi.org/10.1088/1757-899X/486/1/012043>
18. Li G, Hu J, Li Y, and Zhu J. An Improved Model Predictive Direct Torque Control Strategy for Reducing Harmonic Currents and Torque Ripples of Five-Phase Permanent Magnet Synchronous Motors. *IEEE Transactions on Industrial Electronics*. 2019; 66 (8), 5820–5829. <https://doi.org/10.1109/TIE.2018.2870359>
19. Cao B, Grainger BM, Wang X, Zou Y, Reed G, and Mao ZH. Direct Torque Model Predictive Control of a Five-Phase Permanent Magnet Synchronous Motor. *IEEE Transactions on Power Electronics*. 2021; 36 (2), 2346–2360. <https://doi.org/10.1109/TPEL.2020.3011312>
20. Zhao W, Wang H, Tao T, and Xu D. Model Predictive Torque Control of Five-Phase PMSM by Using Double Virtual Voltage Vectors Based on Geometric Principle. *IEEE Transactions on Transportation Electrification*. 2021; 7 (4), 2635–2644. <https://doi.org/10.1109/TTE.2021.3063193>
21. Bahar ST, Omar RG. Torque Ripple Alleviation of a Five-Phase Permanent Magnet Synchronous Motor using Predictive Torque Control Method. *International Journal of Power Electronics and Drive Systems (IJPEDS)*. 2022; 13 (4), 2207–2215. <https://doi.org/10.11591/ijpeds.v13.i4.pp2207-2215>
22. Huang W, Huang Y and Xu D. Model-Free Predictive Current Control of Five-Phase PMSM Drives. *Electronics*. 2023; 12 (23), 1–16. <https://doi.org/10.3390/electronics12234848>.
23. Rajanikanth P, Parvathy ML and Thippiripati VK. Enhanced Model Predictive Current Control-Based Five-Phase PMSM Drive. *IEEE Journal of Emerging and Selected Topics in Power Electronics*. 2024; 12 (1), 838–848. <https://doi.org/10.1109/JESTPE.2023.3342653>

24. Guo B, Xia C and Han JF. NN-Based Model Predictive Direct Speed Control of PMSM Drive Systems. 2014 International Conference on Machine Learning and Cybernetics, Lanzhou. 2014; 163–168. <https://doi.org/10.1109/ICMLC.2014.7009111>
25. Kannan R, Gayathriy N, Natarajan M, Sankarkumary RS, Iyerz LV and Karz NC. Selection of PI Controller Tuning Parameters for Speed Control of PMSM using Biogeography based Optimization Algorithm. 2016 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES), Trivandrum, India. 2016; 1–6. <https://doi.org/10.1109/PEDES.2016.7914260>
26. Frijet Z, Zribi A and Chtourou M. Adaptive Neural Network Internal Model Control for PMSM Speed Regulation. *Journal of Electrical Systems*. 2018; 14 (2), 118–126.
27. Santos JLT, Mejia OA, Sanchez EP and Cortez RS. Parameter Tuning of PI Control for Speed Regulation of a PMSM Using Bio-Inspired Algorithms. *Algorithms*. 2019; 12 (3), 1–21. <https://doi.org/10.3390/a12030054>.
28. Venkatesan S, Kamaraj P, Priya MV. Speed Control of Permanent Magnet Synchronous Motor using Neural Network Model Predictive Control. *Journal of Energy Systems*. 2020; 4 (2), 71–82. <https://doi.org/10.1109/PEDES56012.2022.10080344>
29. Lin Y, Hu H, Chang Y and Wei H. Permanent Magnet Synchronous Motor Vector Control Based on BP Neural Network. *Journal of Physics: Conference Series*, The 5th International Conference on Mechanical, Electric, and Industrial Engineering (MEIE 2022) 24/05/2022–26/05/2022 Online. 2022; 2369, 1–7. <https://doi.org/10.1109/TCYB.2019.2897653>
30. Liu H, Zhang H, Zhang H, Chen G. Model Reference Adaptive Speed Observer Control of Permanent Magnet Synchronous Motor Based on Single Neuron PID. *Journal of Physics: Conference Series*, 2022 5th International Conference on Advanced Algorithms and Control Engineering (ICAACE 2022) 20/01/2022–22/01/2022, Sanya, China. 2022; 2258. <https://doi.org/10.1088/1742-6596/2258/1/012052>
31. Mao H, Tang X and Tang H. Speed Control of PMSM Based on Neural Network Model Predictive Control. *Transactions of the Institute of Measurement and Control*. 2022; 44 (14), 2781–2794. <https://doi.org/10.1177/014233122210862>.
32. Nguyen TT, Tran HN, Nguyen TH, and Jeon JW. Recurrent Neural Network-Based Robust Adaptive Model Predictive Speed Control for PMSM With Parameter Mismatch. *IEEE Transactions on Industrial Electronics*. 2023; 70 (6), 6219–6228. <https://doi.org/10.1109/TIE.2022.3198255>
33. Li H, Liu Z, Shao J. A Model Predictive Current Control Based on Adeline Neural Network for PMSM. *Journal of Electrical Engineering & Technology*. 2023; 18, 953–960. <https://doi.org/10.1007/s42835-022-01324-8>.
34. Wang B, Zhu J, Li Z, and Li Y. An Optimization Algorithm Used in PMSM Model Predictive Control. *Institute of Electronics, Information and Communication Engineers*. 2024; 21 (2), 1–6. <https://doi.org/10.1587/elex.20.20230444>.
35. Boubertakh H, Tadjine M, Glorennec PY, Labiod S. Tuning Fuzzy PD and PI Controllers using Reinforcement Learning. *ISA Transactions*. 2010; 49 (4), 543–551. <https://doi.org/10.1016/j.isatra.2010.05.005> PMID: 20605021
36. Qin Y, Zhang W, Shi J, Liu J. Improve PID Controller Through Reinforcement Learning. 2018 IEEE CSAA Guidance, Navigation and Control Conference (CGNCC), Xiamen, China. 2018; 1–6. <https://doi.org/10.1109/GNCC42960.2018.9019095>
37. Kushwaha A, Gopal M. Reinforcement Learning-Based Controller for Field-Oriented Control of Induction Machine. In: Bansal J., Das K., Nagar A., Deep K., Ojha A. (eds) *Soft Computing for Problem Solving*. *Advances in Intelligent Systems and Computing*, Springer, Singapore. 2019; 817. [https://doi.org/10.1007/978-981-13-1595-4\\_58](https://doi.org/10.1007/978-981-13-1595-4_58).
38. Guan Z and Yamamoto T. Design of a Reinforcement Learning PID controller. 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK. 2020; 1–6. <https://doi.org/10.1109/IJCNN48605.2020.9207624>
39. Book G, Traue A, Balakrishna P, Brosch A, Schenke M, et al. Transferring Online Reinforcement Learning for Electric Motor Control From Simulation to Real-World Experiments. *IEEE Open Journal of Power Electronics*. 2021; 2, 187–201. <https://doi.org/10.1109/OJPEL.2021.3065877>
40. Song Z, Yang J, Mei X, Tao T and Xu M. Deep Reinforcement Learning for Permanent Magnet Synchronous Motor Speed Control Systems. *Neural Computing and Applications*. 2021; 33, 5409–5418. <https://doi.org/10.1007/s00521-020-05352-1>.
41. Sun Q, Du C, Duan Y, Ren H and Li H. Design and Application of Adaptive PID Controller Based on Asynchronous Advantage Actor–Critic Learning Method. *Wireless Networks*. 2021; 27, 3537–3547. <https://doi.org/10.1007/s11276-019-02225-x>.
42. Wang CS, Guo CW, Tsay DM and Perng JW. PMSM Speed Control Based on Particle Swarm Optimization and Deep Deterministic Policy Gradient under Load Disturbance. *Machines*. 2021; 9 (12), 1–19. <https://doi.org/10.3390/machines9120343>.

43. Yeom K. Model Predictive Control and Deep Reinforcement Learning Based Energy Efficient Eco-driving for Battery Electric Vehicles. *Energy Reports*. 2022; 8 (12), 34–42. <https://doi.org/10.1016/j.egy.2022.10.040>.
44. Traue A, Book B, Kirchgassner W, and Wallscheid O. Toward a Reinforcement Learning Environment Toolbox for Intelligent Electric Motor Control. *IEEE Transactions on Neural Networks and Learning Systems*. 2022; 33 (3), 919–928. <https://doi.org/10.1109/TNNLS.2020.3029573> PMID: 33112755
45. Nicola M and Nicola CI. Improvement of Linear and Nonlinear Control for PMSM Using Computational Intelligence and Reinforcement Learning. *Mathematics*. 2022; 10 (24), -34. <https://doi.org/10.3390/math10244667>.
46. Yin F, Yuan X, Ma Z and Xu X. Vector Control of PMSM Using TD3 Reinforcement Learning Algorithm. *Algorithms*. 2023; 16 (9), 1–20. <https://doi.org/10.3390/a16090404>.
47. Sanjines UA, Anthony Maisincho Jivaja, Victor Asanza, Leandro L. Lorente Leyva and Diego H. Peluffo Ordóñez. Adaptive PI Controller Based on a Reinforcement Learning Algorithm for Speed Control of a DC Motor. *Biomimetics*. 2023; 8 (5), 1–26. <https://doi.org/10.3390/biomimetics8050434>.
48. Adil Najem, Moutabir A, Rafik M, Ouchatti A. Comparative Study of PMSM Control Using Reinforcement Learning and PID Control. 2023 3rd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), Mohammedia, Morocco. 2023; 1–5. DOI: [10.1109/IRASET57153.2023.10153024](https://doi.org/10.1109/IRASET57153.2023.10153024)
49. Mei P, Karimi HR, Xie H, Chen F, Huang C and Yang S. A Deep Reinforcement Learning Approach to Energy Management Control with Connected Information for Hybrid Electric Vehicles. *Engineering Applications of Artificial Intelligence*. 2023; 123, 1–15. <https://doi.org/10.1016/j.engappai.2023.106239>.
50. Fayyazi M, Abdoos M, Phan D, Golafrouz M, Jalili M, et al. Real-Time Self-Adaptive Q-Learning Controller for Energy Management of Conventional Autonomous Vehicles. *Expert Systems With Applications*. 2023; 222, 1–14. <https://doi.org/10.1016/j.eswa.2023.119770>.
51. Parsa L and Toliyat HA. Five-Phase Permanent-Magnet Motor Drives. *IEEE Transactions on Industry Applications*. 2005; 41 (1), 30–37. <https://doi.org/10.1109/TIA.2004.841021>
52. Parsa L, Kim N and Toliyat HA. Field Weakening Operation of High Torque Density Five-Phase Permanent Magnet Motor Drives. *IEEE International Conference on Electric Machines and Drives*, 2005., San Antonio, TX, USA; 2005; 1507–1512. <https://doi.org/10.1109/IEMDC.2005.195920>
53. Mirrashid M and Naderpour H. Transit Search: An Optimization Algorithm Based on Exoplanet Exploration. *Results in Control and Optimization*. 2022; 7, 1–37. <https://doi.org/10.1016/j.rico.2022.100127>.
54. Hashim FA, Houssein EH, Hussain K, Mabrouk MS and Al-Atabany W. Honey Badger Algorithm: New metaheuristic algorithm for solving optimization problems. *Mathematics and Computers in Simulation*. 2022; 192, 84–110. <https://doi.org/10.1016/j.matcom.2021.08.013>.
55. Agushaka JO, Absalom E, Ezugwu AE, and Abualigah L. Dwarf mongoose optimization algorithm. *Computer Methods in Applied Mechanics and Engineering*. 2022; 391, 1–38. <https://doi.org/10.1016/j.cma.2022.114570>.
56. Zhao S, Zhang T, Shilin M and Chen M. Dandelion Optimizer: A Nature-Inspired Metaheuristic Algorithm for Engineering Applications. *Engineering Applications of Artificial Intelligence*. 2022; 114, 1–20. <https://doi.org/10.1016/j.engappai.2022.105075>.
57. Lillicrap TP, Hunt JJ, Pritzel A, Heess N, Erez T and et al. Continuous Control with Deep Reinforcement Learning. 4th International Conference on Learning Representations (ICLR). 2016; 1–14. <https://doi.org/10.48550/arXiv.1509.02971>.
58. Fujimoto S, Hoof H and Meger D. Addressing Function Approximation Error in Actor-Critic Methods. *Proceedings of the 35th International Conference on Machine Learning, PMLR*. 2018; 80, 1587–1596. <https://doi.org/10.48550/arXiv.1802.09477>
59. Silver D, Lever G, Heess N, Degris T, Wierstra D and Riedmiller M. Deterministic Policy Gradient Algorithms. 31st International Conference on Machine Learning, ICML. 2014; 1, 605–619. <https://dl.acm.org/doi/10.5555/3044805.3044850>.
60. Dankwa S, and Zheng W. Twin-Delayed DDPG: A Deep Reinforcement Learning Technique to Model a Continuous Movement of an Intelligent Robot Agent. *ICVISP 2019: Proceedings of the 3rd International Conference on Vision, Image and Signal Processing*. 2019; 66, 1–5. <https://dl.acm.org/doi/abs/10.1145/3387168.3387199>.
61. Sucar LE, Morales EF and Jesse Hoey J. *Decision Theory Models for Applications in Artificial Intelligence: Concepts and Solutions*. (Book) IGI Global, 1st edition; 2011. <https://doi.org/10.4018/978-1-60960-165-2>