Deep Reinforcement Learning Algorithm based PMSM Motor Control for Energy Management of Hybrid Electric Vehicles

S. MUTHURAJAN
St. Peter's Institute of Higher Education and Research, Deemed to be University, Chennai, INDIA

RAJAJI LOGANATHAN
ARM College of Engineering and Technology, Chennai, INDIA

R. RANI HEMAMALINI
St. Peter's Institute of Higher Education and Research, Deemed to be University, Chennai, INDIA

Abstract: Hybrid electric vehicles (HEV) have great potential to reduce emissions and improve fuel economy. The application of artificial intelligence-based control algorithms for controlling the electric motor speed and torque yields excellent fuel economy by reducing the losses drastically. In this paper, a novel strategy to improve the performance of an electric motor-like control system for Permanent Magnet Synchronous Motor (PMSM) with the help of a sensorless vector control method where a trained reinforcement learning agent is used and provides accurate signals which will be added to the control signals. Control Signals referred to here are direct and quadrature voltage signals with reference quadrature current signals. The types of reinforcement learning used are the Deep Deterministic Policy Gradient (DDPG) and Deep Q Network (DQN) agents. Integration and implementation of these control systems are presented, and results are published in this paper. The advantages of the proposed method over the conventional vector control strategy are validated by numerical simulation results.

Key-Words: PMSM, Deep Learning, Reinforcement Learning, Intelligence Control, Agent.

Received: June 26, 2022. Revised: January 6, 2023. Accepted: February 13, 2023. Published: March 7, 2023.

1 Introduction

Hybrid electric vehicles have come into vogue due to the limitations of the total IC engine vehicles and fully electric vehicles. Even today, most of the vehicles in India are powered by internal combustion (IC) engines using petrol or diesel fuel for commercial and long travel purposes, [1]. The engine power capacity used in these vehicles caters to the power required for maximum speed. In hybrid electric vehicles, Permanent Magnet Synchronous Motor (PMSM) plays a vital role due to its compact size, lower torque ripple, comfortable cooling technique, high ratio of torque to volume, and efficiency in HEV and EV vehicles. However, many motor control algorithms have been developed for PMSM motors in the past including classic PID control, adaptive control, predictive control, robust control, fuzzy and neuro-fuzzy, artificial neural networks, and advanced intelligent control algorithms. The reinforcement Learning

algorithm which is one of the most important Machine learning algorithms that contribute to the development of intelligent control systems is used for PMSM control have been utilized in HEVs, [2], [3], [4], [5], [9], [10]. Reinforcement learning is characterized by the fact that the accurate mathematical model of the motor does not need to be given as input. In this paper, the mathematical model of the PMSM is considered by assuming the set of simplifications of the parameters in line with the classical control algorithms. Signals representing the state of the system are given as actions to the motor with the involvement of reward optimization. A reward is also consisting of characteristics of signals which are added as a driven process into the control strategy.

In this research, adaptive sensorless stator field-oriented control (SFOC) technique for PMSM motor with deep reinforcement learning agent for creation, training, and testing have been analyzed and discussed the same with results in this paper. Deep reinforcement learning (DRL) proposed also contributes towards improving and contributes towards the improvement of the performance of the PMSM control system. DDPG and DQN agents are used in the proposed deep reinforcement algorithm and an improved variant of this algorithm called the Twin-Delayed Deep Deterministic Policy Gradient (TD3) agent is mainly used for the presented SFOC control system. TD3 algorithm is an effective algorithm with an optimized process for precise estimation of parameters. The advantages of the PMSM control system using deep reinforcement learning are analysed with the help of real-time data in the Matlab/Simulink platform, [6], [7], [8], [9]. The main contributions presented in this paper are (i) Proposed Deep Reinforcement Learning, (ii) Methods of optimizing the control signals for the PMSM based on the SFOC control technique with the different deep reinforcement learning agents (iii) the Behaviour of proposed PMSM motor control strategy in hybrid electric vehicle and (iv) Analysis of real-time results. The results of real-time simulations are presented in this paper and the conclusion and ideas for further approaches have also been presented.

2 Deep Reinforcement Learning Algorithm

A deep reinforcement learning algorithm is mainly used for a system where only a minimum of information is available. DRL algorithm is applied to closed-loop control of motors to execute the tasks without explicitly using complex programming. The learning process in the DRL algorithm is based on the set of decisions made which will be helpful to extend the cumulative reward. In this algorithm, the deep deterministic policy gradient is used in which an off-policy reinforcement learning method is deployed and it is highly suitable for HEVs. This gradient is a model-free, online, and flexible one. DDPG is an actor-critic agent that computes an optimal policy and maximizes the long-term reward. DDPG agent is highly suitable for HEV application. Figure 1 shows the generic schematic of the deep reinforcement learning algorithm where observation and reward are the input signals to the policy update, [9], [10], [11], [12].

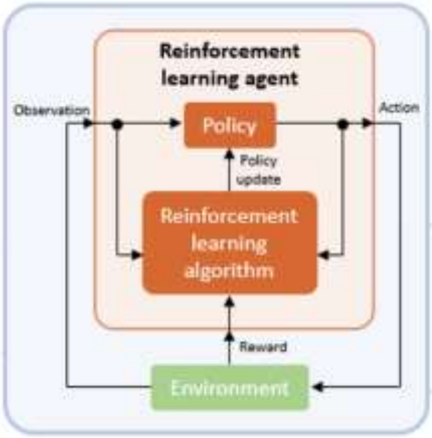

Fig. 1: Block diagram of Deep Reinforcement Learning Algorithm

The environment gets the action from policy and gives the observation which consists of a set of predefined signals deriving the process, and the reward is the output of the environment and represents the success rate. The action is represented by the control variables of the closed-loop control system. Observations represent signals visible to the agent and they are found in the form of measured signals, their rates of change, and associated errors. Usually, the reward is created as part of the continuous actions in the form of a sum, the square of the error of the signals of the present, and the square of the past actions. Weight bias is given to these terms and the same is determined by the problem statement. In motor control, the reward is expressed as a function to reduce the steady-state error. The policy is a component of an agent that implements the learning algorithm and it represents the way of actions associated with the observations it is described by a function with configurable parameters. In the case of a motor control application, the policy is the same as the operating mode of the control system. The optimal policy is determined by the configured learning algorithm with the help of continuous functional parameter configuration. These parameters are associated with the policy depending on cumulative reward maximization. The environment consists of physical devices, reference & actual signals and steady-state errors, filters, disturbances, measurement noise, and A/D and D/A converters, [9],[10].

Important steps of the RL Programming are:

(i) Problem Identification – Learning agents and policy are defined and the process integration is initiated

(ii) Creation of Process model as the environment – Dynamic model of physical

systems and the interface between the subsystems to be defined
(iii) Creation of rewards in DRL – Reward in the form of mathematical equations has been defined to measure the output of the task assigned
(iv) To train the agent – Training of the agent is to be done to accomplish a policy as per reward, algorithm and the process followed.
(v) Policy deployment – Integration of agent and control system of HEV. In this step, auto code generation is playing an important role where executable code with reference to the target embedded platform is generated from the Simulink models. DDPG is usually an agent used in the continuous system and TD3 is a subvariant reagent from DDPG considered for simulation in this research work. It is an actor-critical agent meant for long-term reward maximization, [9].

The improved variant of the DDPG agent is a continuous system and is used in this research work. This agent calculates the long-term maximization award. Training considered in this work has the following phases:

(i) For the observation, the present state is S and the action is “, where, N is the stochastic noise level.
(ii) After the execution of A, rewards R and S’ are calculated. S’ is the next state observation
(iii) The experience formulated as (S, A, R, S’) and which is stored in the next step
(iv) “(S_i, A_i, R_i, S_i’)” are randomly generated, [9]
(v)

$$\boldsymbol{y_i = R_i + \delta.\min(Q'_k(S'_k, clip(\mu'(S'_k|\theta_\mu) + (\varepsilon)|\theta_{Q'k}))} \quad (1)$$

The main target value function given in equation (1) is as below:

- Sum of experience reward (R_i) and the minimum discounted feature reward

3 Optimization of SFOC Control Strategy for PMSM with Deep Reinforcement Learning

SFOC is an efficient method of controlling the PMSM motor and is effectively integrated with a deep reinforcement learning algorithm.
The following equations represent the dynamics of PMSM:

$$\boldsymbol{\frac{di_d}{dt} = -\frac{R_s}{L_s}i_d + \frac{L_q}{L_d}n_p\omega i_q + \frac{1}{L_d}v_d} \quad (2)$$

$$\boldsymbol{\frac{di_q}{dt} = -\frac{R_s}{L_s}i_q + \frac{L_d}{L_q}n_p\omega i_d - \frac{\lambda_0}{L_q}n_p\omega + \frac{1}{L_q}v_q} \quad (3)$$

$$\boldsymbol{\frac{d\omega}{dt} = \frac{3}{2}\frac{n_p}{J}(\lambda_0 i_q + (L_d - L_q)i_d i_q) - 1/J\, T_L - B/J\, \omega} \quad (4)$$

The above equations represent the dynamics of PMSM in the ‘d-q’ reference frame. Field-oriented control of the PMSM motor along with the deep learning algorithm is shown in Figure 2. DRL with TD3 agent learns the behaviour of the PMSM control system as given in Figure 2 is analysed in this paper. It provides the reference signals as the three control inputs to the cascade control system (i_{qref}, v_{dref}, v_{qref}) after the training phase, so that the improved control system will have better performance. These three control signals are reference quadrature current, reference direct voltage, and reference quadrature voltage respectively.

TD3 agent is trained with 300 episodes for PMSM control and the number of steps per episode is 100. The sampling time taken for every agent is 10^{-4}s. The training phase for the agent gets stopped when the cumulative average reward is greater than -150 for 100 consecutive episodes or after 300 episodes initially set training episodes have elapsed. During the simulation, learning needs to be improved to get the best training and for this purpose, Gaussian noise intersects the signals received and the same is transmitted by the agent.

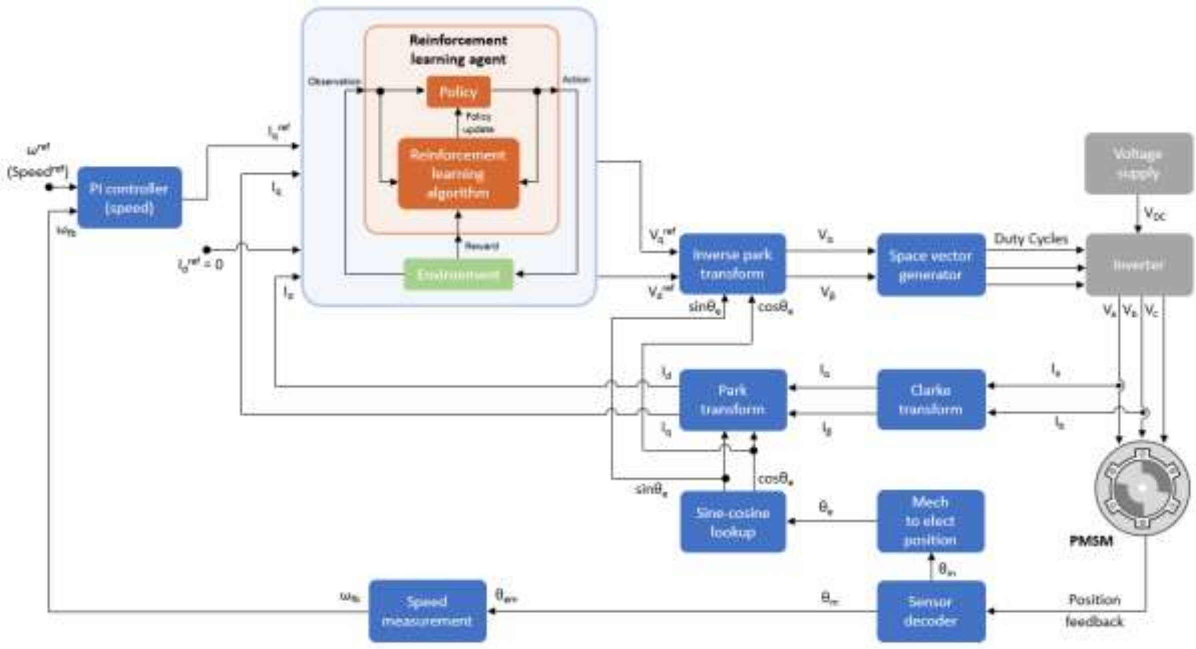

Fig. 2: Block diagram of HEV with DRL Algorithm

4 Deep Reinforcement Learning with Inner Current Control Loop

The inner loop is working based on the current control operation with the TD3 agent as shown in Figure 2. Once the learning is done, the TD3 agent will provide the command/reference signals for the voltage control signals v_d and v_q. Figure 3 shows the Simulink implementation of the proposed deep reinforcement learning for both the inner current control loop (Torque Control) and outer voltage control loop (Speed Control). In the inner current control loop, the observation signals are i_d, i_q, i_{derror}, and i_{qerror}. To start with, the deep neural network is created with two inputs and one output. The total training time for this case is 7:12:5.

5 Deep Reinforcement Learning with outer Speed control loop

Figure 3 shows the Matlab/Simulink implementation diagram of PMSM control for outer-loop speed/voltage control using a TD3 agent. In this case, the command/reference signal from the TD3 agent is added to the control signal i_{qref}. The observations consist of the signals and error signals such as ω, ω_{error}, i_d, i_q, i_{derror}, and i_{qerror}. In this phase, the total training time taken is 2:54:45.

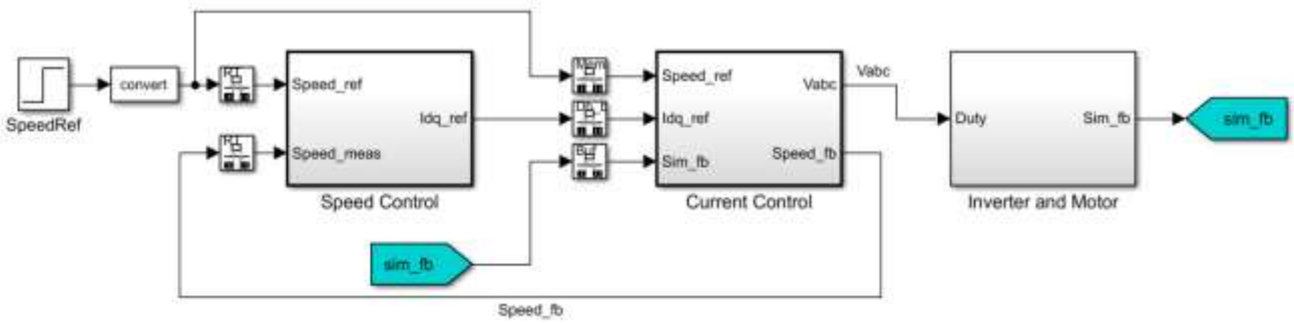

Fig. 3: PMSM Motor Control with outer Speed loop

In Figure 3, the speed controller and current controller are cascaded in which the current controller feeds the speed feedback to the speed controller to derive the PWM pulses (uniform duty cycle). Motor torque is estimated from stator current components like i_d and i_q and these components are to be compared with the desired components such as i_d^* and i_q^* respectively. Figure 4 shows the Rewards r1,r2,r3, and r4 during the deep learning training process, and episode rewards are shown in Figure 5. Episode number information, average results, training options, and final results are shown in Figure 5.

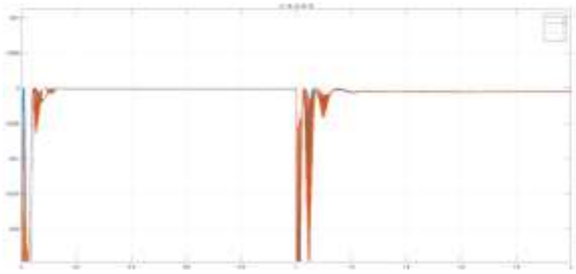

Fig. 4: Rewards during the training of DRL

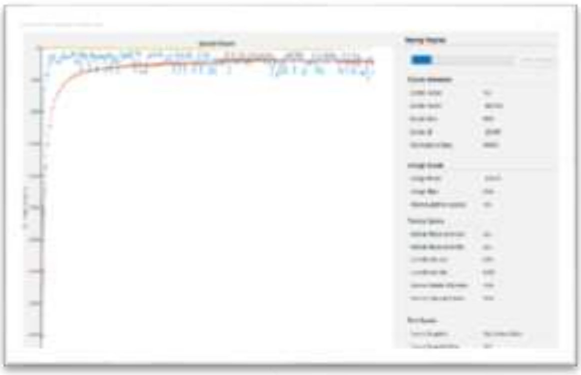

Fig. 5: DRL Training Progress

6 Results and Analysis

TD3 approximates reward from the environment for I_{dref} and I_{qref} was taken from the PI controller model and actions such as actual I_d and I_q using the representations such as speed and torque values of the closed-loop control system. Agent TD3 tunes the feedback current and voltage values delivered from FOC which would influence the given reference current values using the actor representation. Both DDPG and its agent TD3 are using the same structure in the proposed Simulink model. The DDPG agent maximizes the Q value and the actor-network network is used to estimate the action such as feedback values of current and voltage. Since the TD3 uses the value of Q to update the policy and the resulting policy may be suboptimal and accumulating training errors may lead to different behavior. The TD3 algorithm is an extension of DDPG with improvements that make it more robust by preventing over-estimation of Q values, [13].

Figure 6 shows the proposed HEV with PMSM control based on the DRL algorithm. The main contribution of the DRL algorithm is to optimize the control loop of PMSM in HEV to achieve the energy management cycle. Inputs to DRL blocks are actual I_d and I_q, the reference value of I_d and I_q, and actual speed with reference speed. Action from the DRL algorithm is the voltage which has the components such as direct and quadrature values. So, based on the training process proposed in this paper, learning takes place, and action is generated.

Figure 7 shows the speed tracking of the PMSM motor in HEV where the actual speed of the motor is following the reference and met the target value. The target value of PMSM speed is set to 600 rpm and the actual speed attains the set value by the combination of the TD3 agent and SFOC algorithm. Figure 8 shows the current waveforms such as i_d & i_{dref} and i_q and i_{qref}. Both the i_d and i_q values are tracking the reference respective values and placing 90° apart.

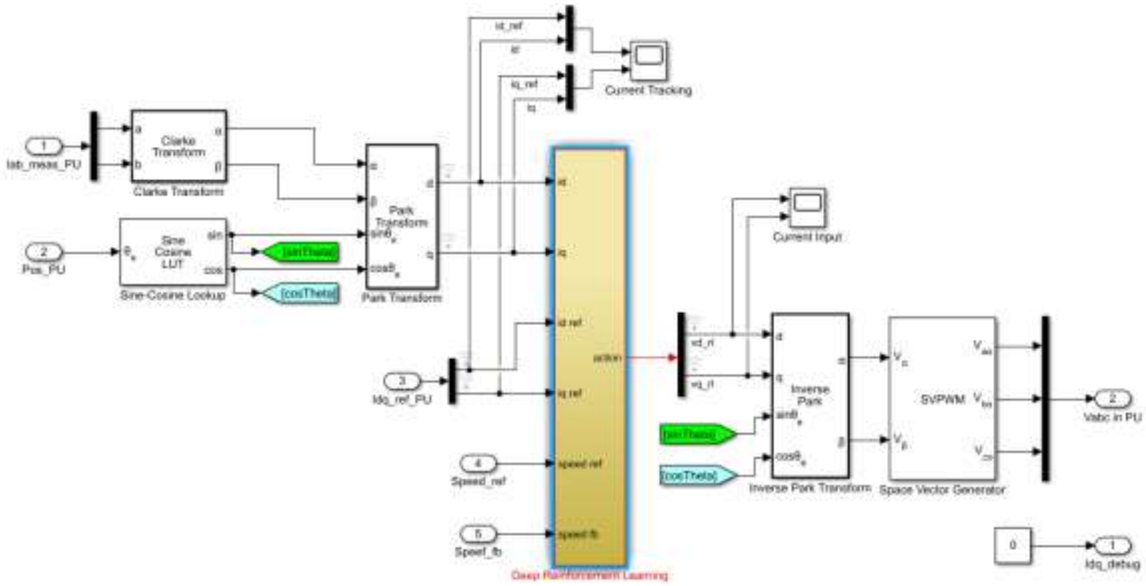

Fig. 6: Simulink Diagram of Proposed DRL-based PMSM for HEV

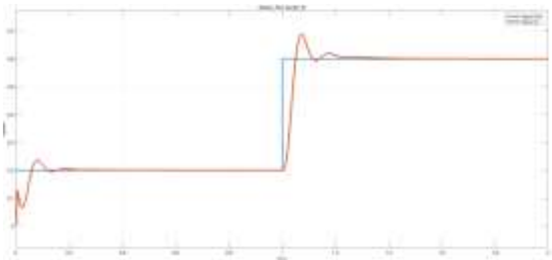

Fig. 7: Speed of PMSM

Figure 9 shows the voltage profile V_d and V_q with DRL learning and these components are also following their respective values closely and 90° phase apart. Figure 10 shows the logic analyzer which shows the digital and analog signals for the input and output characteristics of the proposed HEV control system. From this figure, it is observed that analog and digital signals with reference to the sampling time, time offset, time span values are well within the limit as per the given speed control system parameters.

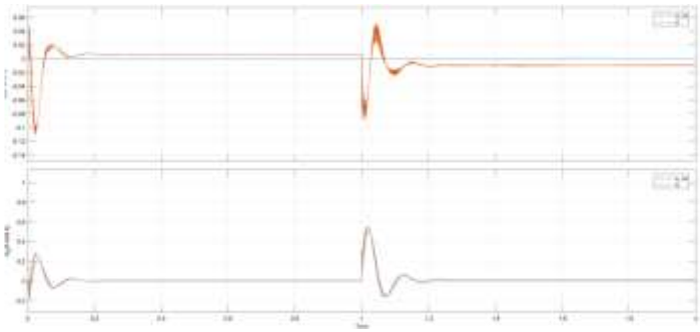

Fig. 8: Current Control

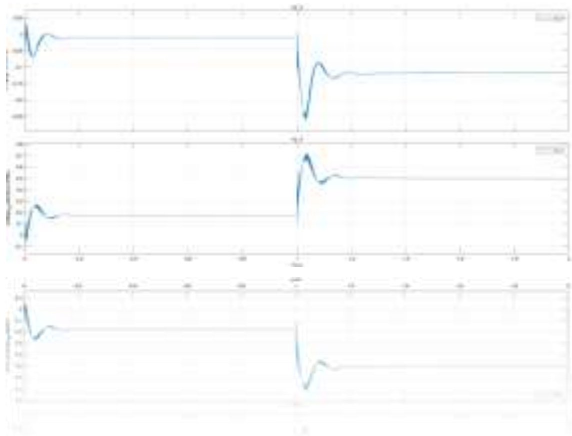

Fig. 9: Voltage profile

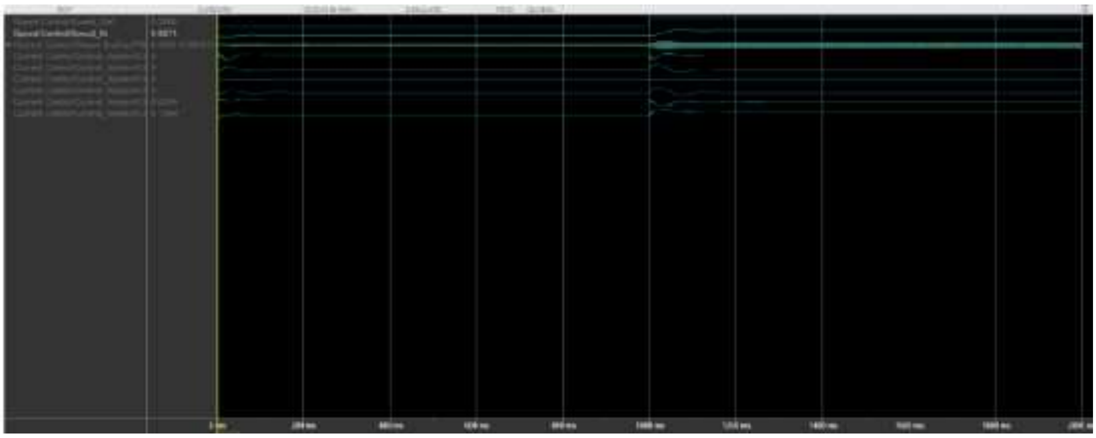

Fig. 10: Logic Analyzer

Figure 11 shows the output results with reference to the performance of hybrid electric vehicles. The following waveforms are captured from the HEV scope:

(i) Drive cycle (mph)
(ii) Engine and Motor Speed Comparison (RPM)
(iii) Engine and Motor Torque Comparison (Nm)
(iv) Battery Current (A)
(v) Battery State of Charge (%)
(vi) Fuel Consumption (g/kWh)

Vehicle velocity varies from 0 to 60 mph as the drive cycle and engine speed is in line with the drive cycle profile. Motor speed is also tried to follow the driving schedule as shown in the figure. Engine speed and motor speed vary from 0 to 3500 rpm whereas torque delivered from the motor and engine varies from 0 to 200 Nm. In most cases, the engine and motor are dividing the power delivered by contributing the torque. From the battery current waveform, it is understood that currently varies from -50A to +50A as per the drive schedule and torque profile of HEV, and hence based on the DRL control mechanism of the motor battery current is also tuned to minimize the losses and hence it is proven that the solution presented in this paper is energy efficient one by using a deep learning algorithm. In Figure 11, the battery state of charge is also shown and it is evident that charging is controlled within the band of 70%. Fuel consumption is also controlled and 40% of fuel saving is proven from the result.

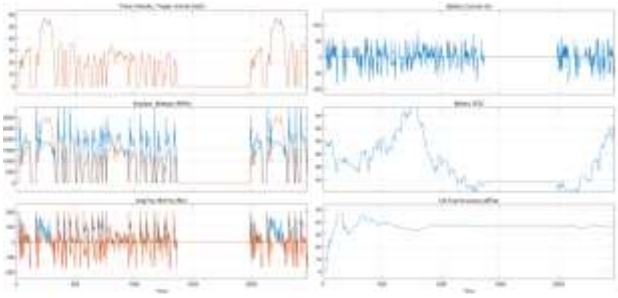

Fig. 11: HEV Vehicle Parameters

Hence benefits of the proposed algorithm with a speed control of the PMSM motor are given below:

(i) Motor and engine speed profiles are inline with the HEV speed profile
(ii) Steady-state error of the control system is less than 1% and the efficiency of the control accuracy is more than 99.5%
(iii) Percentage of Fuel saving is improved by 40%
(iv) Accuracy of estimation of the state of charge (SoC) and state of health (SoH) are improved by 99.9%
(v) High torque-to-speed ratio
(vi) Very high torque/volume ratio

7 Conclusions

In this paper, the SFOC-type control structure for the PMSM for HEV energy management is presented which shows improved performance using the deep reinforcement learning algorithm. Comparison results are thus presented for a case where a deep reinforcement learning agent is properly trained to provide the reference/ command signals that are added to actual control signals v_d, v_q, and i_{qref}. The main objective of this research work is to improve the performance of the HEV by incorporating the novel control technique of the PMSM motor in order to save the energy that is monitored by an energy management system.

Numerical simulations were used to demonstrate the superiority of control systems using deep reinforcement learning and subsequent work exploring optimization possibilities associated with implementing deep reinforcement learning on PMSM controllers for HEVs. Proposed algorithm is proven successful in Matlab/Simulink platform but has not yet been implemented in real-time passenger vehicles and that needs to be done in a real-time vehicle to show the performance of upcoming versions. Moreover, the suggested algorithm may also be suggested for core electric vehicles (EV) and to suggest the rugged energy management system.

References:

[1] Hannan, Mahammad A., F. A. Azidin, and Azah Mohamed. "Hybrid electric vehicles and their challenges: A review." *Renewable and Sustainable Energy Reviews* 29 (2014): 135-150.

[2] Liu, Xiangdong, Hao Chen, Jing Zhao, and Anouar Belahcen. "Research on the performances and parameters of interior PMSM used for electric vehicles." *IEEE Transactions on Industrial Electronics* 63, no. 6 (2016): 3533-3545.

[3] Li, Yaohua, Dieter Gerling, Jian Ma, Jingyu Liu, and Qiang Yu. "The Comparison of Control Strategies for the interior PMSM drive used in the Electric Vehicle." *World Electric Vehicle Journal* 4, no. 3 (2010): 648-654.

[4] Sheela, A., M. Suresh, V. Gowri Shankar, Hitesh Panchal, V. Priya, M. Atshaya, Kishor Kumar Sadasivuni, and Swapnil Dharaskar. "FEA based analysis and design of PMSM for electric vehicle applications using magnet software." *International Journal of Ambient Energy* 43, no. 1 (2022): 2742-2747.

[5] Luo, Yefeng, Jianqi Qiu, and Cenwei Shi. "Fault detection of permanent magnet synchronous motor based on deep learning method." In *2018 21st International Conference on Electrical Machines and Systems (ICEMS)*, pp. 699-703. IEEE, 2018.

[6] Dhulipati, Himavarsha, Eshaan Ghosh, Shruthi Mukundan, Philip Korta, Jimi Tjong, and Narayan C. Kar. "Advanced design optimization technique for torque profile improvement in six-phase PMSM using supervised machine learning for direct-drive EV." *IEEE Transactions on Energy Conversion* 34, no. 4 (2019): 2041-2051.

[7] Bilal, Ahmad, Asad Waheed, and Muhammad Hassan Shah. "A Comparative Study of Machine Learning Algorithms for Controlling Torque of Permanent Magnet Synchronous Motors through a Closed Loop System." In *2019 Second International Conference on Latest trends in Electrical Engineering and Computing Technologies (INTELLECT)*, pp. 1-6. IEEE, 2019.

[8] Song, Zhe, Jun Yang, Xuesong Mei, Tao Tao, and Muxun Xu. "Deep reinforcement learning for permanent magnet synchronous motor speed control systems." *Neural Computing and Applications* 33, no. 10 (2021): 5409-5418.

[9] Nicola, Marcel, Claudiu-Ionel Nicola, and Dan Selișteanu. "Improvement of PMSM sensorless control based on synergetic and sliding mode controllers using a reinforcement learning deep deterministic policy gradient agent." *Energies* 15, no. 6 (2022): 2208.

[10] Nicola, Marcel & Nicola, Claudiu. (2021). Improvement of PMSM Control Using Reinforcement Learning Deep Deterministic Policy Gradient Agent. 1-6. 10.1109/Ee53374.2021.9628371.

[11] Savant, Rushank, Abhiram Ajith Kumar, and Aditya Ghatak. "Prediction and analysis of permanent magnet synchronous motor parameters using machine learning algorithms." In *2020 Third International Conference on Advances in Electronics, Computers and Communications (ICAECC)*, pp. 1-5. IEEE, 2020.

[12] Traue, Arne, Gerrit Book, Wilhelm Kirchgässner, and Oliver Wallscheid. "Toward a reinforcement learning environment toolbox for intelligent electric motor control." *IEEE Transactions on Neural Networks and Learning Systems* (2020).

[13] MathWorks. (2023). MATLAB - MathWorks. www.mathworks.com [Accessed on 03/02/2023].

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The authors equally contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.