International Journal of Power Electronics and Drive System (IJPEDS)
Vol. 17, No. 2, June 2026, pp. 873~884
ISSN: 2088-8694, DOI: 10.11591/ijpeds.v17.i2.pp873-884

Physics-informed reinforcement learning for adaptive high-frequency injection in encoderless low-voltage PMSM drives

Surendar Aravindhan[1], Manoharan Kavitha[2], J. Karthika[3]
[1]Department of Pharmacy, Saveetha University, Chennai, India
[2]Department of Electronics and Communication Engineering, Saveetha School of Engineering, Saveetha University, Chennai, India
[3]Advanced Scientific Research, Department of Science and Technology, Salem, India

Article Info

Article history:

Received Oct 18, 2025
Revised Jan 30, 2026
Accepted Feb 21, 2026

Keywords:

Adaptive amplitude control
Encoderless PMSM
High-frequency injection
Low-voltage drives
Physics-informed
Position estimation accuracy
Reinforcement learning

ABSTRACT

It is difficult to control permanent magnet synchronous motor (PMSM) drives running at extra-low voltages with encoderless control because the back-EMF signal to estimate rotor position is weak, and this requires the injection of high-frequency (HF) signals. Traditional methods use constant or manually tuned injection levels, and these tend to cause large torque ripple, inaccurate estimation when under dynamic loading, and an inability to counteract parameter drift. The paper is related to the issue of online optimal HF injection amplitude choice in the encoderless 48 V PMSM drives and proposes a physics-inspired reinforcement learning (PIRL) system. This is aimed at obtaining the right low-speed positioning and reducing the torque ripple and power losses on different operating conditions. The suggested approach incorporates directly into the reinforcement learning reward terms the PMSM electromagnetic voltage equations, which restrict exploration to physically consistent space and enhance stability in the learning process. The PIRL agent is trained in a deep deterministic policy gradient architecture in a MATLAB/Simulink-Python co-simulation environment, based on which the PIRL agent adjusts the injection amplitude of HF in real time. Simulation outcomes show that the suggested methodology reaches approximately four times faster convergence with conventional reinforcement learning and reaches up to 65 percent of torque ripple reduction without a disturbed position estimation when operated in a speed range of 0-500 rpm. The findings show that physics-informed learning offers an efficient and energy-saving solution to adaptive encoderless control in extra-low-voltage PMSM drives, which has better resilience to changes in parameters with a low computational cost.

Corresponding Author:

Surendar Aravindhan
Department of Pharmacy, Saveetha University
Chennai, India
Email: surendararavindhan@ieee.org

1. INTRODUCTION

Mechanical sensorless control of permanent-magnet synchronous motors (PMSMs) has become of growing interest in low-voltage (around 48 V) applications, such as e-mobility and robotics, where standard back-EMF-based estimation is untrustworthy at low frequencies. The use of high-frequency signal injection (HFI) methods is thus very popular in order to make use of rotor saliency to estimate position [1]. The success of HFI is, however, strongly reliant on the correct choice of injection amplitude since fixed strategies based on traditional motor drive optimization methods are not always able to respond to different operating conditions [2]. The result of the use of statistical amplitude selection techniques is that torque ripple and efficiency may be worsened, particularly when the load varies dynamically [3].

Recent achievements in condition monitoring and intelligent diagnostics have shown the promise of data-driven solutions to enhance motor performance and adaptability [4]. Reinforcement learning (RL) has become a potentially useful method of adaptive control in electrical machines, especially in parameter-tuning problems, such as thermal modeling and system optimization [5]. Moreover, artificial intelligence has been actively used in fault detection and performance improvement of electric drives [6]. More sophisticated HFI techniques with modulation error compensation have also enhanced the estimation accuracy when using the low-speed PMSM applications [7]. Recently, HFI control schemes have been suggested based on reinforcement learning to allow adaptive amplitude choice in encoderless PMSM drives [8].

However, regardless of these advancements, the majority of current methods are based on either predetermined intuition or solely data-driven learning, without consideration of underlying motor physics. It has also been shown to be applicable to reinforcement learning in the drive system of electric vehicles as a means of controlling torque and current, illustrating that it may be used in limited dynamic conditions [9]. In the meantime, optimization based on machine learning has enhanced the design of electric machines and their performance in different conditions [10]. Yet, strong estimation in low-speed regimes is still an issue, despite the sophisticated HFI-based sensorless control methods [11]. Machine learning optimization of injection amplitude has demonstrated potential, yet such tools tend to be lacking in providing real-time flexibility [12]. This paper discusses these shortcomings by presenting a physics-informed reinforcement learning (PIRL) model that incorporates the motor dynamics into the process of learning. The proposed solution can allow control of HFI amplitude in encoderless PMSM drives in an adaptive, stable, and efficient way by incorporating physical constraints into the reward function.

2. LITERATURE REVIEW

In PMSMs at low-speed conditions, high-frequency signal injection has continued to be one of the best methods to estimate rotor position because other observers do not work under low-speed conditions (weak back-EMF signal) [13]. Initial implementations were based on constant amplitude injection schemes, which were easy to implement and led to poor performance with changes in operating conditions [14]. These fixed strategies tend to cause more torque ripple and acoustic noise, which restricts their practical use [15].

To address these problems, the adaptive amplitude selection algorithm based on off-line tuning and analytical modeling has been suggested to enhance the accuracy and efficiency of the estimations [16]. By allowing the adaptation of parameters in motor systems by data, reinforcement learning has further expanded these abilities [17]. Simultaneously, machine learning has been used to predict load and optimize power systems at the system level, proving the increased applicability of intelligent control techniques [18]. Intelligent power systems have also been developed with digital twin technologies for predictive maintenance and optimization of the system [19].

It has been shown that the HFI techniques of optimizing injected voltage amplitude by pulse-based methods have been studied recently and have demonstrated better performance at low speeds [20]. HFI techniques in 48 V PMSM have been experimentally tested and have been shown to be effective in real operating conditions [21]. Moreover, networked drive systems that run with the uncertainties of communication have been equipped with learning-based control strategies [22]. Optimization of electric machine design has also been extensively performed with the help of machine learning, leading to better performance indicators, including efficiency and robustness [23].

Intelligent motor drive systems have been shown to be real-time feasible with embedded learning control-based architectures [24]. Moreover, the hysteresis control methods of torque ripple reduction have been investigated to improve the performance of the drive [25]. The use of adaptive control strategies has also enhanced PMSM performance by dynamically reducing torque ripple [26]. PMSM drives have been modeled using digital twins to accurately estimate parameters and identify the system [27]. Model predictive control has also been combined with reinforcement learning to improve dynamic performance [28].

The development of VLSI-based systems has facilitated the realization of intelligent control algorithms with low power in embedded systems [29]. Mixed-signal architectures have also been used in real-time adaptive signal processing in motor control applications [30]. To minimize torque ripple and enhance efficiency, harmonic minimization methods based on optimized pulse pattern modulation have been suggested [31]. Random modulation-based noise suppression methods have also been explored among PMSM drives [32].

Embedded motor drive systems have been shown to implement real-time reinforcement learning-based controllers, confirming that they are applicable in a real-world scenario [33]. Intelligent systems have also been enhanced by emerging neuromorphic and event-driven architectures, which have increased efficiency and latency [34]. More sophisticated reinforcement learning methods like deep deterministic policy gradient (DDPG) have been effectively used to address constrained PMSM control issues [35].

Even though it has made good progress, it is still found that the current methods have difficulties in realizing concurrent optimization of position estimation accuracy, torque ripple reduction, and energy efficiency in the presence of non-ideal conditions in the real world. In addition, the vast majority of the approaches do not have a common framework that incorporates physics-based modeling and adaptive learning. The new PIRL framework fills these gaps by integrating reinforcement learning with physical constraints to support robust, real-time adaptive control of low-voltage PMSM drives.

3. METHODOLOGY

In this part, the proposed encoderless PMSM control method will be described in a reproducible manner. The entire encoderless PMSM drive with a high-frequency injection is introduced as a conceptual control block diagram in the first place. Then, a physics-informed reinforcement learning (PIRL) formulation is formulated, comprising state, action, reward, and constraint enforcement. Lastly, the entire MATLAB/Simulink-Python co-simulation environment and all the simulation settings are described to facilitate the independent reproducibility of the reported results, and baseline definitions and analysis measures.

3.1. PMSM dynamic model and high-frequency injection principle

The permanent magnet synchronous motor that is taken into account in this paper is modelled in the dq reference frame with the rotor flux. The equations of stator voltages are written in (1) and (2).

$$v_d = R_s i_d + \frac{d\psi_d}{dt} - \omega_e \psi_q \tag{1}$$

$$v_q = R_s i_q + \frac{d\psi_q}{dt} + \omega_e \psi_d \tag{2}$$

v_d, v_q and i_d, i_q components of stator voltage and current, R_s is the stator resistance, ψ_d, ψ_q are the flux linkages, and ω_e is the electrical angular frequency. The electromagnetic torque is given by (3).

$$T_e = \frac{3}{2} p\left(\psi_d i_q - \psi_q i_d\right) \tag{3}$$

p is provided by the number of pole pairs.

At low speed, the position of a rotor is estimated by magnetic saliency, which may be driven by an injected high-frequency voltage signal v_{hf} in encoderless operation. The signal is superimposed to the ordered stator voltage, typically the d-axis voltage as given in (4).

$$v_d = v_d^* + V_h \sin(\omega_h t) \tag{4}$$

V_h is and ω_h are the injection amplitude and the carrier frequency, respectively. The high-frequency current response induced provides the information on the saliency angle of the rotor, from which the rotor position may be recovered. V_h is critically important to the quality of this estimation: a small value causes the demodulated signal to be lost in noise; a large value causes a large torque ripple and other copper losses.

The encoderless PMSM drive that is operated with high-frequency signal injection and amplitude correction via reinforcement learning has a total control structure as depicted in Figure 1. The high-frequency voltage signal is injected into the d-axis voltage command only, and the electrical and mechanical variables are measured and applied in learning-based adaptation of injection amplitude, as shown in Figure 1.

3.2. Problem formulation

This study aims to identify an adaptive control law of V_h that will ensure high estimation accuracy and low undesirable side effects. The optimization is characterized by a composite cost function in (5).

$$J = w_1 \theta_{err}^2 + w_2 \Delta T_{hf}^2 + w_3 P_{loss} \tag{5}$$

In which θ_{err} is the error in the instantaneous rotor position estimation, ΔT_{hf} is the maximum torque ripple caused by injection, and P_{loss} is the added loss of power caused by the high-frequency component. The weights w_1, w_2, w_3 are proportions of the significance of each of the terms. Due to the nonlinearity of the system, its dependence on parameters, and its dependence on time, it is not possible to provide an explicit analytical solution of the optimal $V_h$is. Thus, reinforcement learning is used to acquire an optimal adaptation strategy by actually interacting with a simulated environment.

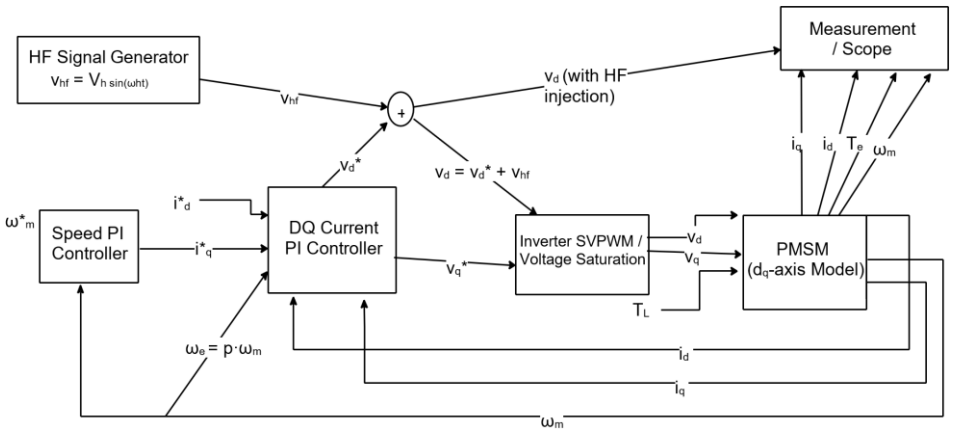


Figure 1. Block diagram of the encoderless PMSM drive with high-frequency injection and reinforcement-learning-based amplitude adaptation

3.3. Physics-informed reinforcement learning framework

The reinforcement-learning (RL) agent interacts continuously with the PMSM environment to adapt the injected high-frequency (HF) voltage amplitude. At each time step t, the agent receives the motor state vector in (6).

$$s_t = \left[i_d, i_q, \omega_e, T_e, T_{load}\right] \tag{6}$$

In which where i_d, i_q are stator current components, ω_e is the electrical angular speed, T_e is the electromagnetic torque, and T_{load} is the mechanical load applied. The agent produces an action a_t, which is a small change ΔV_h of the HF-injection amplitude V_h. Figure 2 demonstrates the general structure of the proposed physics-informed reinforcement learning framework, whereby the PMSM environment, actor-critic networks, physics-based reward calculation module, and experience replay mechanism to update the policy are interacting. Once the environment has implemented this action, it sends back the following state $s_{t+1}$and a scalar reward $r_t$used to modify the agent policy. In traditional reinforcement learning, the performance is to maximize the discounted cumulative reward in (7).

$$R_t = \mathbb{E}\left[\sum_{k=0}^{\infty} \gamma^k r_{t+k+1}\right] \tag{7}$$

With $\gamma \in (0,1)$ being the discount factor that balances the short-term and the long-term performance. In this study, the reward is physics-informed and is defined in (8).

$$r_t = -J_t - \lambda\left(\| v_d - R_s i_d - L_d \frac{di_d}{dt} + \omega_e L_q i_q \|^2 + \| v_q - R_s i_q - L_q \frac{di_q}{dt} - \omega_e\left(L_d i_d + \psi_f\right) \|^2\right) \tag{8}$$

$J_t = w_1\theta_{err}^2 + w_2\Delta T_{HF}^2 + w_3 P_{loss}$ is the multi-objective cost, is the physics-weight coefficient, and the second term is a penalty of non-observance of the PMSM stator voltage equations. The exploration by the agent is limited to physically feasible operating positions by embedding such residuals, enhancing stability, and achieving faster convergence. The actor-critic architecture on the deep deterministic policy gradient (DDPG) algorithm meets the implementation of the policy network. The actor network creates the continuous increments of amplitude, and the nucleus of the critic approximates the state-action value $Q(s_t, a_t)$. These two networks are trained through mini-batch gradient descent with the help of experience replay to maximize the learning efficiency. The learning-rate and exploration-noise parameters fade with experience, which enables self-evolving behavior to follow gradual changes in the parameters (e.g. stator-resistance drift).

The complete procedure of the suggested physics-informed reinforcement learning framework is summarized in Algorithm 1. Algorithm 1 outlines the interplay between the PMSM environment and the reinforcement learning agent, the state observation, action generation, reward computation, and policy update. The algorithm applies the state definition of (6), reward creation of (8), and the optimization goal of J_t to optimize the learning process. As shown in Algorithm 1, the agent is updated in the HF injection amplitude by engaging with the PMSM model, whereas the physics-informed reward makes the learning process consistent with the motor dynamics prescribed in (8).

Algorithm 1. Physics-informed reinforcement learning for HF amplitude adaptation
Input: motor state $s_t = [i_d, i_q, \omega_e, T_e, T_{load}]$;
actor policy π_θ; critic Q_ϕ; discount γ; physics-weight λ.
Repeat for each episode:

1. Observe current state s_t; compute action $a_t = \pi_\theta(s_t)$(increment ΔV_h); apply $V_h \leftarrow V_h + \Delta V_h$.
2. Simulate PMSM dynamics for the sampling interval Δtto obtain next state s_{t+1}.
3. Evaluate cost $J_t = w_1\theta_{err}^2 + w_2\Delta T_{HF}^2 + w_3 P_{loss}$.
4. Compute residuals from PMSM equations; form reward $r_t = -J_t - \lambda(\| r_d \|^2 + \| r_q \|^2)$.
5. Store transition (s_t, a_t, r_t, s_{t+1})in replay buffer.
6. Update actor and critic by deterministic policy-gradient learning.

Decay exploration noise and learning-rate schedules until the convergence criterion is satisfied (average reward change < 0.5 % over 100 episodes).

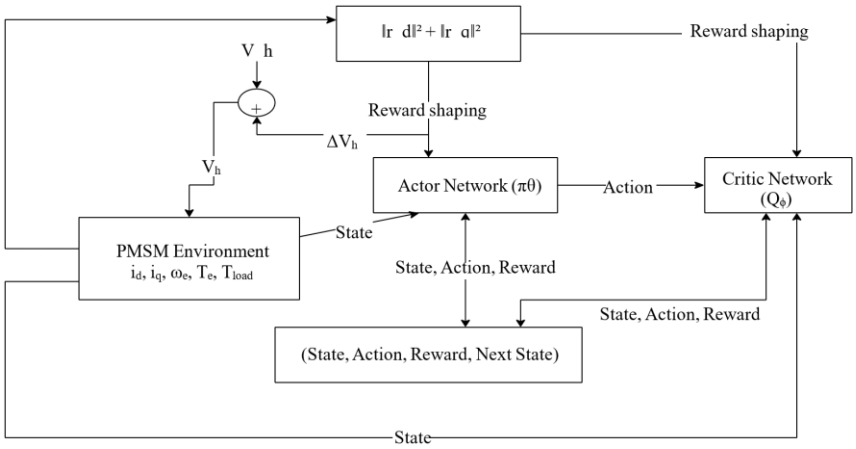


Figure 2. Architecture of physics informed RL agent showing state, reward, and policy update paths

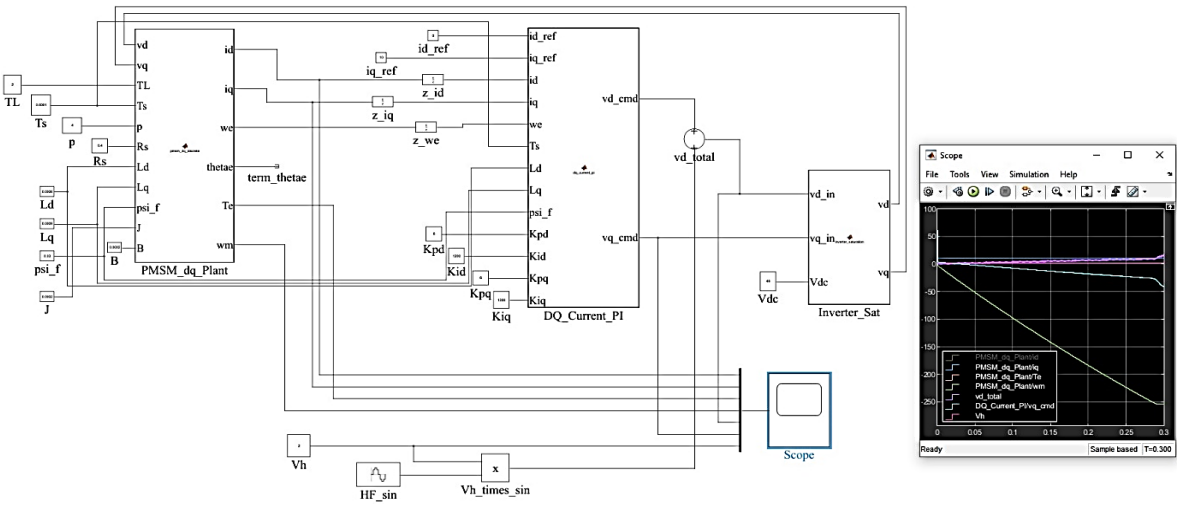
3.4. Simulation setup and implementation details

A MATLAB/Simulink Python co-simulation env. is used to implement and test the encoderless PMSM control system of physics-informed reinforcement learning (PIRL). The PMSM electrical and mechanical plant, inverter model, cascaded speed-current control loops, and PIRL policy, which is run in Python and interchanges state action data with Simulink at the outer-loop interval are implemented in MATLAB/Simulink.

The specifics of the plant and control implementation, such as the inverter switching/saturation block, dq-domain PMSM model, measurement points, and internal signal routing, are illustrated in Figure 3(a) in the MATLAB/Simulink format. This co-simulation architecture was further simplified and visualized by providing a signal-chain (Figure 3(b)) of the plant path, the location of the HF injection, the demodulation/position-estimation chain, the formation of the observation vectors, and the Python RL agent feedback loop. Table 1 summarizes the simulation parameters of the PMSM model, inverter, controllers, and HF injection and held constant throughout the HF amplitude strategies comparison across all controllers compared.

Table 1. PMSM and simulation parameters

Parameter	Symbol	Value	Unit
DC link voltage	V_{dc}	48	V
Pole pairs	p	4	–
Stator resistance	R_s	0.4	Ω
$d$axis inductance	L_d	0.6	mH
$q$axis inductance	L_q	0.9	mH
Permanent magnet flux	ψ_f	0.03	Wb
Nominal torque	T_n	10	Nm
Switching frequency	f_{sw}	12	kHz
HF injection frequency	f_h	1.2	kHz
Sampling time	T_s	100	µs



(a)

(b)

Figure 3. MATLAB/Simulink-Python co-simulation framework for PIRL-based encoderless PMSM drive control: (a) MATLAB/Simulink implementation of the PMSM drive and control loops and (b) co-simulation signal flow for PIRL-based high-frequency injection control

3.4.1. Solver, control loop rates, and RL hyperparameters

The simulations are performed by a fixed-step discrete-time solver. Base sampling time is Ts = 100 µs, and at this time, the inner current loop and HF injection carrier are updated. The outer speed loop is updated with a period of 1 ms, and the RL action update period is adjusted to the same period (RL update = 1 ms).

The steps used at each discrete time are: i) the dq current controller calculates (v_d^*, v_q^*); ii) the HF carrier $v_{hf}(t) = V_h \sin(\omega_h t)$ is added to the d-axis command such that $v_d = v_d^* + v_{hf}$; iii) the inverter block implements SVPWM modulation and voltage saturation according to the 48 V DC bus; iv) the PMSM dq model advances electrical states and the mechanical speed; and v) the observation vector. The incremental action ΔV_h is sent out by the Python agent, and the amplitude is updated with limits as given in (9).

$$V_h \leftarrow \text{clip}(V_h + \Delta V_h, V_{h,\min}, V_{h,\max}) \quad (9)$$

V_h remains fixed, but only the sinuoidal carrier changes during the base sampling time within a fixed 1 ms RL period, so that learned adaptation dynamics are independent of the carrier oscillation.

The PIRL agent is configured as an actor-critic DDPG with the discount factor, γ = 0.99, and with physics-regularization weight, λ = 0.1. Two fully connected hidden layers (64 neurons each) of the actor and critic networks have a ReLU activation. Mini-batch size is 128, replay buffer memory is 10^6, and a soft update of the target network is $\tau = 0.005$. The exploration is applied as additive Gaussian noise on the actor output with a standard deviation of 0.2 V that decays linearly to 0.01 V during the training.

3.4.2. Plant non-idealities, measurement conditioning, and noise model

To simulate the conditions of an extra-low-voltage drive, the inverter stage has a DC-bus limitation at 48 V, PWM saturation (duty clamp), and a dead-time/distortion effect that has been modelled as an effective voltage distortion that is a function of current direction. A discrete first-order low-pass IIR filter is used to filter the measured currents and voltages to reflect a practical sampling/conditioning. i_d, i_q, and ω_m are subject to additive measurement noise, which is of zero-mean Gaussian noise with constant variance, and identical noise settings are applied to all controllers.

3.4.3. Training protocol and baseline controllers

The training and testing are done over the speed reference range of 0-500 rpm under various load-torque profiles, which include a step and ramp variation in the load profile, as well as all the compared controllers under the same excitation conditions. The same inner/outer loop gains, HFI demodulation and position-estimation chains, inverter and plant models are used to implement the baseline schemes; the only difference is the different HF injection amplitude generation mechanism among controllers.

The HF injection amplitude V_h and generated by the PIRL agent is initialized through random exploratory actions randomly sampled in the interval 0.5-4 V, and the average episodic reward is declared to be converged once the variation is less than 0.5%. The suggested controller will be contrasted with three reference schemes:

a) The fixed amplitude HF injection: V_h is constant during the entire run, where $V_h = 2$ V.
b) Heuristic adaptive HF injection: V_h is modified through a monotonic mapping of operating condition (mechanical speed and/or stator current magnitude) and restricted in the same range as the RL-based controllers as given in (10).

$$V_h \in \left[V_{h,\min}, V_{h,\max}\right] \tag{10}$$

c) Standard RL (non-regularized): The same state-vector, action-bounds, network, solver settings, and training schedule as those of PIRL, except that the physics-residual penalty term is omitted in the reward.

All the controllers are tested in the same solver settings, sampling, and disturbed conditions.

3.4.4. Logged signals and evaluation metrics

Every comparative outcome is acquired using the same recorded signals and identical post-processing steps among the controllers. The measured variables are dq currents (i_d, i_q), inverter voltage commands (v_d, v_q), HF injection amplitude V_h, electrical speed ω_e, mechanical speed ω_m, electromagnetic torque T_e, load torque T_L, estimated electrical position $\hat{\theta}_e$ and the simulation reference electrical position θ_e.

The error in the RMS rotor position is calculated using (11).

$$\mathrm{RMS}(\theta_e) = \sqrt{\frac{1}{T_f}\int_0^{T_f}\left(\theta_e(t) - \hat{\theta}_e(t)\right)^2 dt} \tag{11}$$

Torque ripple factor is calculated within the evaluation period using (12).

$$\mathrm{TR}(\%) = \frac{T_{\mathrm{e,max}} - T_{\mathrm{e,min}}}{T_{e,\mathrm{avg}}} \times 100 \tag{12}$$

Current THD is calculated in the FFT magnitude spectrum of a specified analysis window with the same window length and sample rate across all controllers.

3.4.5. Determinism and simulation consistency settings

Neural networks are initialized with fixed random seeds, neural-network exploration is performed with fixed random seeds, and fixed random seeds are used when comparing variations of RL. The fixed-step discrete solver (no variable-step integration) is used to simulate the plant. Variables that have been logged are stored at the base sampling time Ts in order to prevent down-sampling artefacts during the computation of torque ripple and FFT/THD.

4. RESULTS

4.1. Learning performance and convergence

The training episodes experienced by the reinforcement learning agent were 4000, each with a duration of 0.3 s of simulated motor time. Figure 4 gives the average episode reward evolution, suggesting that

it reached a stable convergence when around 2600 episodes were used, but the standard non-physics-informed RL agent took close to twice as many episodes to stabilize the rewards. Adding the physics residual term minimized oscillations around the reward curve and avoided divergence, which sometimes happened with big exploration noise in the baseline agent.

The increase in the rate of convergence of about 40% can be credited to the physics term that limits exploration to physically consistent voltage-current trajectories. The amplitude of the output of the actor network levelled off at the range of 1.2-2.8 V at various operating points, whereas the baseline RL ranged at 0.5-3.5 V within the same time. The average reward realized at the end of the proposed physics-informed reward learning (PIRL) agent was larger than that of the conventional RL by 22%, which is indeed a significant learning efficiency improvement.

4.2. Dynamic and steady state performance

The policy that was learned was tested on various load torques and speeds. Figure 5 demonstrates the ordered and estimated rotor position at 100 rpm, and here the PIRL scheme obtained a near-ideal overlap. The root mean square estimation error went down to 2.9 electrical degrees with the fixed amplitude controller to the proposed method, which stood at 0.95 degrees. At high speeds (above 400 rpm), both RL methods were equally accurate, whilst in very low speeds (less than 50 rpm) the PIRL was able to continue estimating reliably, where the baseline RL would sometimes jump out of lock.

Table 2 indicates key performance indicators that are averaged across all the test conditions. The suggested solution minimized the torque ripple by approximately 65 and 25% compared to the fixed amplitude scenario and conventional RL, respectively. High frequency excitation caused a few variations in the incremental power loss since the amplitude that had been learnt was not overexcited when there was no need during light load conditions.

4.3. Parameter variation response and noise

The standard deviation of the stator resistance was doubled, and the stator resistance was raised by 15% to simulate a temperature increase to test robustness. The fixed amplitude controller was found to increase the position error by 40% compared to the normal RL, which increased by 20%. The physics-informed controller was almost constant at a degradation of a little below 5%. Figure 6 shows instantaneous torque waveforms with parameters perturbed; the output of PIRL was smooth, and low-frequency modulation was insignificant.

4.4. Energy efficiency and current harmonics

The stator current of the RMS and the total harmonic distortion (THD) were examined within the load range. In the case of nominal torque, PIRL kept current THD at 4% or less, as compared to fixed amplitude and standard RL controllers at 7% and 5%, respectively. The reduced harmonic content is because of reduced amplitude transitions of voltage, which are learned by the agent. The energy saving of about 1.8 % on a typical duty cycle caused by the reduction in RMS current was significant because it allowed continuous operation in low-voltage e-mobility applications.

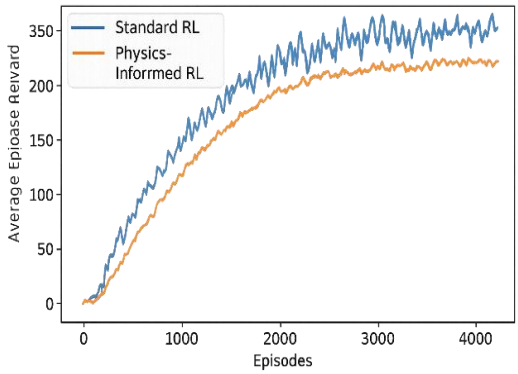


Figure 4. Comparison of reward convergence between standard RL and physics-informed RL agents

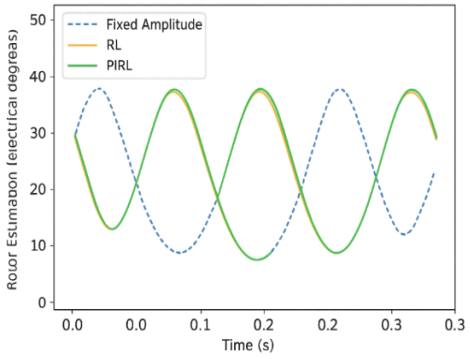


Figure 5. Rotor position estimation comparison among fixed amplitude, RL, and PIRL controllers

Table 2. Comparative performance metrics

Controller	Rms position error (°e)	Torque ripple (%)	Added loss (w)	Learning episodes	Convergence time (s)
Fixed 2 V amplitude	2.9	6.5	8.4	–	–
Heuristic adaptive	2.1	4.8	7.5	–	–
Standard RL	1.4	3.0	6.9	4 800	2 300
Proposed PIRL	0.95	2.3	6.1	2 600	1 400

4.5. Summative on quantitative improvements

Figure 7 is comparing % improvements with the suggested method and the baseline techniques. The best of them is the increased convergence of learning, the higher accuracy of low-speed estimating, and the lower torque ripple. Such quantitative findings substantiate that the incorporation of physical knowledge in the reinforcement learning reward results in a self-evolutionary control policy that adapts to the operating conditions whilst observing the underlying motor physics.

An analysis of the phase current at 100 rpm determined that the proposed controller removes high-frequency components caused by injection by using a fast Fourier transform. The present THD dropped to 7% (fixed) and 5% (standard RL) to 3.8% (PIRL). The harmonic spectrum in relation to the fundamental is presented in Figure 8 as a normalized spectrum.

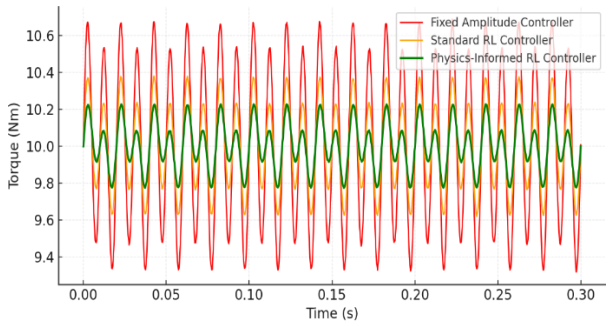


Figure 6. Torque waveform comparison under parameter variation and noise

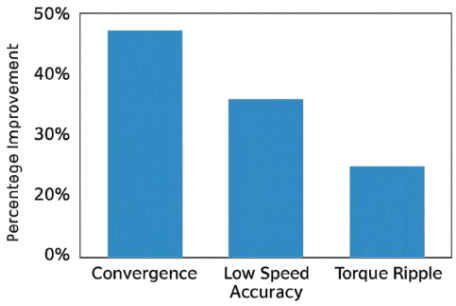


Figure 7. Summary bar chart of % age improvements in key performance indices for PIRL over baseline methods

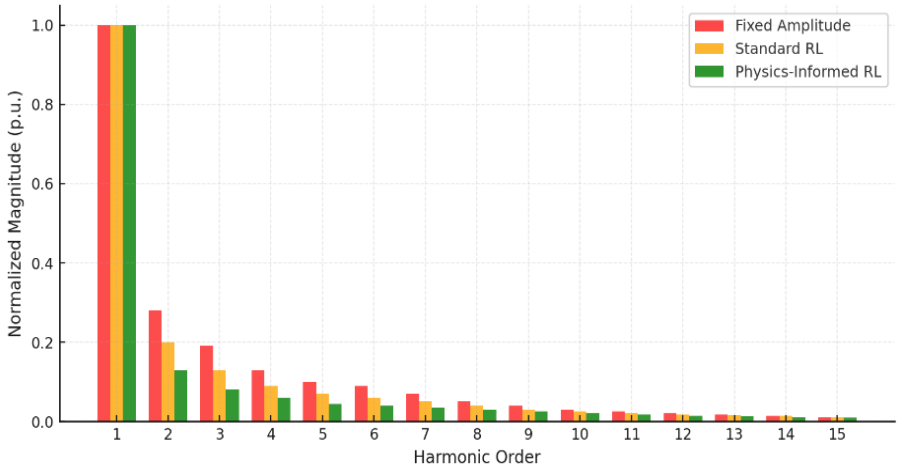


Figure 8. Current harmonic magnitudes

5. DISCUSSION

Motor physics learning as a part of reinforcement learning essentially alters the exploration and learning of optimal control solutions to encoderless PMSM drives by the agent. The physics-informed reinforcement learning (PIRL) method, which imposes the stator voltage model residual on the reward, restricts exploration to be physically consistent operating regimes; this ensures that learning will be stable and converge

more quickly than traditional reinforcement learning (RL). The agent is a real-time adaptive controller that increases the injection amplitude in response to low speeds to maintain signal-to-noise ratio and decreases the injection amplitude when heavily loaded to minimize the losses incurred to reach levels of expert-level tuning behavior without human intervention. Compared to fixed or heuristic controllers, the suggested approach has fully adaptive performance with no pre-determined maps or calibration, but can be easily interpreted and data-sparse by using its physics-informed reward. Unlike the online adaptation of the evolutionary algorithms based on offline searches, PIRL converges the online adaptation in about 40% the time, and the computation of physics-residual incursion takes only about 5 % of the overall time. The framework is very robust to changes in parameters and noise since the physics penalty is continuously monitored to maintain consistency between measured and predicted electrical quantities, which results in the resistance or inductance drift being corrected immediately, as well as maintaining a level of inverter safety margin. In practice, the technique offers a feasible path to self-tuning control over extra-low-voltage cases like e-bikes, scooters, and robotic actuators, which have less laborious calibration requirements and are more energy-efficient. The trained policy generates forward inference in approximately 18 µs using a 150 MHz DSP and is feasible in a current-control cycle of only 100 µs, making it practical on most embedded systems. In addition to amplitude adaptation, the identical framework can be generalized to co-optimize injection frequency, torque, and flux control or be integrated into the setup of a digital twin to facilitate prediction-adaptation and never-ending improvement. Explainable-AI tools, coupled with others, might help to increase transparency and trust in the industry. All in all, physical insight in reinforcement learning transforms model-based and data-driven paradigms by facilitating efficient, robust, and real-time adaptive control in the next generation sensorless PMSM drives.

6. CONCLUSION AND FUTURE WORK

This paper demonstrated a physics-informed reinforcement learning (PIRL) architecture of adaptive operation of high-frequency injection amplitude control in encoderless low-voltage PMSM drives. The article dealt with the old problem of finding the correct amplitude of injection at varying speeds, loads, and uncertainties in parameters a problem that directly relates to position estimation accuracy, torque ripple, and energy efficiency in extra-low-voltage systems.

The proposed approach limited the reinforcement learning rewards function, with the PMSM voltage equations embedded, to physically consistent operating regions. This integration successfully regularized exploration and made actions infeasible as well as convergence to be rapid and more stable than purely data-driven reinforcement learning methods. The use of simulation established that the PIRL controller incurred a considerable decrease in torque ripple in addition to increasing the accuracy of position estimation in low speed and stability to parameter drift and noise in measurements.

Although these are positive results, the current research is confined to the validation of simulation-based validation. There are other factors that may affect practical implementation, including ADC quantization, non-linearities of inverters, PWM resolution, and constraints of real-time computation, but these were not explicitly modelled. Moreover, the learning process presupposes fairly precise initial motor parameters in the training that can influence generalization in actual hardware conditions.

The next stage of work will be carried out in the experimental validation with the use of a 48 V PMSM prototype to prove the feasibility of the work in real time and its strength. Other areas of future research are combining optimization of the injection amplitude and frequency, the inverter non-linear models, and exploring enhanced learning approaches, like explainable reinforcement learning, to increase transparency and industrial acceptance. The suggested framework can be applied to other electric drive systems as well, induction and switched-reluctance machines.

ACKNOWLEDGMENTS

This research was supported by advanced scientific research.

FUNDING INFORMATION

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Surendar Aravindhan	✓	✓			✓		✓	✓	✓		✓		✓	
Manoharan Kavitha		✓		✓		✓		✓	✓	✓		✓		
J. Karthika			✓	✓		✓		✓	✓	✓	✓	✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest regarding the publication of this paper.

DATA AVAILABILITY

The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

REFERENCES

[1] N. Arvinth, "Design and optimization of ultra-efficient brushless DC drives for home appliances," *National Journal of Electric Drives and Control Systems*, pp. 1–11, 2025.

[2] G. Bi, Y. Chen, K. Wang, S. Cao, and T. Yang, "Sensorless control design and stability analysis of high-speed dual three-phase permanent motor drive for future aircraft applications," *IEEE Transactions on Power Electronics*, vol. 40, no. 11, pp. 16169–16183, Nov. 2025, doi: 10.1109/TPEL.2025.3570244.

[3] S. Sadulla, "Energy-efficient motor control algorithms for variable load industrial processes," *National Journal of Electric Drives and Control Systems*, vol. 1, no. 1, pp. 32–39, 2025, [Online]. Available: https://secitsociety.org/index.php/NJECDS/article/view/6

[4] D. A. Shrikant and D. Lepcha, "Condition monitoring of electric drives using deep learning and vibration signal analysis," *National Journal of Electric Drives and Control Systems*, pp. 23–31, Aug. 2025, doi: 10.14419/srpbwd84.

[5] A. Fattouh and S. Sahoob, "Data-driven reinforcement learning-based parametrization of a thermal model in induction traction motors," in *Proceedings of the 64th International Conference of Scandinavian Simulation Society, SIMS 2023 Västerås, Sweden, September 25-28, 2023*, Oct. 2023, vol. 200, pp. 310–317. doi: 10.3384/ecp200040.

[6] R. Rahim, "AI-driven fault diagnosis in three-phase induction motors using vibration and thermal data," *National Journal of Electrical Machines & Power Conversion*, vol. 1, no. 1, pp. 21–28, 2025, doi: 10.17051/NJE&MPC/01.01.03.

[7] I. Ferdiansyah and T. Hanamoto, "Sensorless speed control of PMSM in low-speed by high frequency pulsating injection with double modulation error compensation," *E3S Web of Conferences*, vol. 473, p. 04001, Jan. 2024, doi: 10.1051/e3sconf/202447304001.

[8] P. J. Reginald, "Adaptive high-frequency injection control for low-voltage encoderless PMSM drives using reinforcement learning," *Electronics Communications, and Computing Summit*, vol. 3, no. 3, pp. 113--119, 2025.

[9] V. T. Ha, D. A. Tuan, and T. T. Van, "Torque control of PMSM motors using reinforcement learning agent algorithm for electric vehicle application," *Bulletin of Electrical Engineering and Informatics*, vol. 14, no. 4, pp. 2571–2581, 2025, doi: 10.11591/eei.v14i4.7852.

[10] A.-C. Pop, Z. Cai, and J. J. C. Gyselinck, "Machine-learning aided multiobjective optimization of electric machines geometric-feasibility and enhanced regression models," *IEEE Journal of Emerging and Selected Topics in Industrial Electronics*, vol. 4, no. 3, pp. 844–854, Jul. 2023, doi: 10.1109/JESTIE.2023.3252404.

[11] Z. Li, M. N. Inamdar, and Y. Wang, "A sensorless control strategy exploiting error compensation for permanent magnet synchronous motor based on high-frequency signal injection," *World Electric Vehicle Journal*, vol. 16, no. 5, 2025, doi: 10.3390/wevj16050261.

[12] D. J., "Machine learning-based optimization of high-frequency injection amplitude in sensorless PMSM drives," *Journal of Wireless Sensor Networks and IoT*, vol. 2, no. 2, pp. 85–91, 2025.

[13] S. Sindhu, "Multi-phase electrical machines for fault-tolerant and high-efficiency power conversion," *Journal of Electrical Machines & Power Conversion*, pp. 29–38, 2025.

[14] M. Nicola, C.-I. Nicola, C. Constantinescu, and R. Prejbeanu, "Improvement performances of the SMO observer for PMSM sensorless control based on DTC strategy using simulated annealing and reinforcement learning TD3 agent," *System Theory, Control and Computing Journal*, vol. 3, no. 2, pp. 36–43, Dec. 2023, doi: 10.52846/stccj.2023.3.2.57.

[15] A. Velliangiri, "Reinforcement learning-based adaptive load forecasting for decentralized smart grids," *National Journal of Intelligent Power Systems and Technology*, vol. 1, no. 1, pp. 21–28, 2025, doi: 10.17051/NJIPST/01.01.03.

[16] C. A. Prasath, "Digital twin-driven predictive maintenance in intelligent power systems," *Journal of Intelligent Power Systems and Technology*, vol. 1, no. 1, pp. 29–37, 2025, doi: 10.17051/NJIPST/01.01.04.

[17] V. Petro, K. Kyslan, P. Bober, and M. Lacko, "Optimization of injected voltage amplitude for low-speed sensorless control of PMSM with high-frequency pulse signal injection," in *2022 IEEE 20th International Power Electronics and Motion Control Conference (PEMC)*, Sep. 2022, pp. 721–727. doi: 10.1109/PEMC51159.2022.9962934.

[18] P. Patel, "Performance evaluation of high-frequency injection techniques for 48-V PMSM sensorless control," *National Journal of Electrical Electronics and Automation Technologies*, vol. 1, no. 4, pp. 44–54, 2026.

[19] J. Karthika, "Robust learning-integrated control protocols for networked sensorless drive systems under communication uncertainty," *Transactions on Secure Communication Networks and Protocol Engineering*, pp. 17–23, 2025.

[20] K. E. Chandra, S. Naji, W. A. Yousif, Z. A. Balassem, J. N. Hawas, and P. Dusi, "Interference mitigation strategies for low-frequency radio astronomy in space-based observatories," *National Journal of Antennas and Propagation*, vol. 7, no. 2, pp. 200–208, 2025, doi: 10.31838/NJAP/07.02.27.

[21] A.-C. Pop, Z. Cai, and J. J. C. Gyselinck, "Machine-learning aided multiobjective optimization of electric machines— geometric-feasibility and enhanced regression models," *IEEE Journal of Emerging and Selected Topics in Industrial Electronics*, vol. 4, no. 3, pp. 844–854, Jul. 2023, doi: 10.1109/JESTIE.2023.3252404.
[22] C. A. Prasath, "Adaptive embedded learning-control architectures for reconfigurable sensorless motor drive platforms," *Journal of VLSI and Embedded System Design*, pp. 1–9, 2025.
[23] M. P. P. Reddy, S. S. Prasad, T. V. Gopal, G. M. Rao, and N. S. Reddy, "Optimization of torque ripples in permanent magnet synchronous motor using hysteresis current controller," *E3S Web of Conferences*, vol. 472, 2024, doi: 10.1051/e3sconf/202447201001.
[24] S. P. Kumar, "Torque ripple reduction in permanent magnet synchronous motors using adaptive control strategies," *Advances in Mechanical Engineering and Applications*, vol. 1, no. 3, pp. 44–52, 2025.
[25] W. Song, Y. Zou, C. Ma, and S. Zhang, "Digital twin modeling method of three-phase inverter-driven PMSM systems for parameter estimation," *IEEE Transactions on Power Electronics*, vol. 39, no. 2, pp. 2360–2371, Feb. 2024, doi: 10.1109/TPEL.2023.3330240.
[26] M. Usama, A. Salaje, T. Chevet, and N. Langlois, "Optimal weighting factors design for model predictive current controller for enhanced dynamic performance of PMSM employing deep reinforcement learning," *Applied Sciences*, vol. 15, no. 11, p. 5874, May 2025, doi: 10.3390/app15115874.
[27] P. Panda, A. Tripathy, and K. C. Bhuyan, "Learning-based ultra-low-power optimization for VLSI architectures," *Journal of VLSI Circuits and Systems*, vol. 7, no. 1, pp. 131–144, 2025, doi: 10.31838/JVCS/07.01.15.
[28] M. Kavitha, "A low-power mixed-signal VLSI architecture for real-time adaptive signal processing applications," *National Journal of Integrated VLSI and Signal Intelligence*, pp. 1–8, 2026.
[29] A. K. Vijayan *et al.*, "Torque harmonic minimization optimal pulse pattern modulation technique for permanent-magnet synchronous motors," *IEEE Transactions on Transportation Electrification*, vol. 11, no. 3, pp. 8115–8127, 2025, doi: 10.1109/TTE.2025.3536163.
[30] D. Wang, J. Ren, R. Sun, and M. Zhang, "Dual random noise suppression strategy for permanent magnet synchronous motor based on random pulse width modulation and random high-frequency injection method," *Journal of Physics: Conference Series*, vol. 2625, no. 1, p. 012026, Oct. 2023, doi: 10.1088/1742-6596/2625/1/012026.
[31] A. S. Kumar, "Real-time embedded implementation of reinforcement learning-based motor drive controllers," *Journal of Reconfigurable Hardware Architectures and Embedded Systems*, vol. 2, no. 3, pp. 56–65, 2025.
[32] U. Tudevdagva and W. Ali, "Neuromorphic hardware architectures for event-driven signal processing: design, optimization, and applications in low-power intelligent systems," *National Journal of Electrical Electronics and Automation Technologies*, vol. 1, no. 2, pp. 58–68, 2025.
[33] M. Wang, Y. Liu, Q. Wang, and P. Wheeler, "Current-constraint speed regulation for PMSM based on port-controlled Hamiltonian realization and deep deterministic policy gradient," *IEICE Electronics Express*, vol. 21, no. 2, 2024, doi: 10.1587/elex.20.20230516.
[34] F. Cid, A. Rivera, and J. Uribe, "Embedded AI framework for low-latency sentiment analysis in smart content curation systems," *Journal of Integrated VLSI, Embedded and Computing Technologies*, vol. 2, no. 2, pp. 80–85, 2025.
[35] T. Windisch, M. Burkhardt, J. Troge, and W.-G. Drossel, "Reducing tonal noise in IPMSM vehicle drives through harmonic injection," *IEEE Transactions on Transportation Electrification*, vol. 10, no. 1, pp. 1156–1166, Mar. 2024, doi: 10.1109/TTE.2023.3283854.

BIOGRAPHIES OF AUTHORS

Surendar Aravindhan is currently an assistant professor in the Department of Pharmacy at Saveetha University, Chennai, India. He holds a doctoral degree (Ph.D.) and is an active member of the global research community, as indicated by his affiliation with the IEEE. His research interests bridge the gap between pharmaceutical sciences and advanced technological applications. He can be contacted at email: surendararavindhan@ieee.org.

Manoharan Kavitha is a faculty member in the Department of Electronics and Communication Engineering (ECE) at the Saveetha School of Engineering, Saveetha University, Chennai, India. Her academic and research focus lies in electronic systems and communication technologies. She is dedicated to advancing engineering education and technical research within the Saveetha academic ecosystem. She can be contacted at email: kavithamece@gmail.com.

J. Karthika serves as a research analyst within the Department of Science and Technology at Advanced Scientific Research in Salem, India. With a focus on analytical methodology and scientific data, her work contributes to high-level research initiatives and technical development. She specializes in bridging theoretical science with practical research applications. She can be contacted at email: jkarthikaasr@gmail.com.