Iran University of Science and Technology

International Journal of

Railway Research

ISSN: 2423-3838

Reinforcement learning-based speed control of an electric train without utilizing its dynamics

Mahdi Soltani Nejad[1*], S. M. Mousavi G[1].

[1]School of Railway Engineering, Iran University of Science and Technology, Tehran, Iran

ARTICLE INFO

Article history:

Received: 01.12.2023

Accepted: 07.04.2024

Published: 14.04.2024

Keywords:

Electric Train

Speed Control

Railway

Reinforcement Learning

Model Free

A B S T R A C T

Automatic speed control of electric trains is always a matter of attention due to reasons such as safety, travel comfort, and, most importantly, preventing human errors. In order to achieve this goal, the dynamic models of the train and electric motor will be estimated and then simulated. Based on the simulated model and the desired objectives, the controller will be designed by an experienced engineer. During this process, the simulated state space models always encounter errors. Additionally, the controller design process will be conducted offline. Thus, the issue will be addressed by incorporating a state feedback controller and formulating the Bellman equation for reference signal tracking. In order to obtain the state controller parameters, the policy is initially evaluated and subsequently improved. This iterative process will continue until the termination conditions are met. In this process, there is no need for a dynamic model of the system, and the controller parameters will be obtained solely through interaction with the environment. Therefore, even with changes in the train dynamics, the controller will be updated online. The proposed method for determining the values of state feedback parameters will be juxtaposed with other artificial intelligence techniques, including particle swarm optimization, genetic algorithm, and bees algorithm. Evaluation metrics such as root mean square error, coefficient of determination (R-squared), and explained variance will be employed to assess the performance of these algorithms. The results obtained underscore the superior efficacy of the proposed method.

1. Introduction

The advent of electric trains revolutionized transportation, offering efficient, environmentally friendly, and often high-speed travel options. Central to the functionality of electric trains is the precise control of their speed, a critical aspect that ensures safety, efficiency, and passenger comfort [1]. Electric trains also produce fewer environmental pollutants [2]. Speed control mechanisms in electric trains have undergone significant advancements over the years, driven by technological innovations and the pursuit of optimal performance.

Electric train speed control refers to the sophisticated systems and mechanisms designed to regulate the velocity of trains powered by electric traction motors. Unlike traditional diesel-powered locomotives, electric trains rely on electricity for propulsion, typically obtained from overhead lines or third rails. This reliance on electrical power enables more precise and responsive speed control, allowing for smoother acceleration, deceleration, and maintenance of desired speeds.

The importance of speed control on electric trains cannot be overstated. It directly impacts factors such as energy efficiency, braking effectiveness, and passenger comfort. Efficient speed control systems not only optimize energy consumption but also contribute to the overall safety and reliability of train operations [3]. Additionally, in high-speed rail networks, precise speed control is essential for maintaining

*Corresponding author
Email address: soltaninejad.mahdi@yahoo.com

[DOI: 10.22068/ijrare.340]

schedules and ensuring timely arrivals and departures.

Over the years, various technologies have been developed to enhance electric train speed control. These include advanced propulsion systems, sophisticated onboard computers, and automated signaling systems. Furthermore, innovations in regenerative braking systems allow electric trains to recover and utilize energy during deceleration, further improving efficiency and reducing environmental impact [4].

This introduction sets the stage for exploring the intricacies of electric train speed control, delving into the principles, technologies, and challenges involved. As we delve deeper, we will uncover the mechanisms that govern the smooth and efficient movement of electric trains, driving progress in modern transportation systems.

The benefits of the railway system have been mentioned. In this system, speed control plays a significant role in preventing accidents, energy consumption, travel comfort, and traffic management. Human error in controlling train speed can result in significant disruptions and dreadful accidents within railway systems [3]. Such errors may arise from inappropriate actions by the train driver, train control system operator, or other human-related factors. Some examples of human errors in train speed control include:

1. **Speed Non-Compliance**: The train driver intentionally or inadvertently does not adhere to the prescribed speed limits for the track. This can lead to unforeseen accidents and be preventable [5].
2. **Loss of Focus**: The train control system operator, due to a lack of attention to their responsibilities, may incorrectly manage the train's speed. This loss of focus can be a result of fatigue, human shortcomings, or psychological issues [6].
3. **Decision-Making Errors**: In some instances, incorrect decisions are made by responsible individuals. This may involve making wrong decisions regarding speed, applying brakes when not necessary, or failing to make decisions in emergency situations.
4. **Use of Drugs or Alcohol**: The consumption of drugs or alcohol by drivers or train operators can lead to poor decision-making and a reduction in decision-making abilities.
5. **Fatigue**: Inappropriate fatigue can impair concentration and decision-making abilities. Fatigued drivers or train operators may fail to properly control the train's speed.

For the reduction of human errors in train speed control, modern tools and technologies, such as automatic train control (ATC) systems and train traffic management systems, are commonly utilized. These systems can effectively minimize errors made by train drivers and operators, ensuring the correct control of train speeds. Additionally, continuous training and safety education for individuals engaged in train control and traffic management are highly crucial.

Extensive research has been carried out in the field of automatic speed control for electric trains, and we will now explore some of these studies in further detail. In their study [7], Wang focused on the speed control of trains using a sliding mode proportional integral derivative (PID) controller. After analyzing and comparing various existing research methods, the neural network and sliding mode control techniques were carefully chosen and integrated into the train's speed control system. Jiaxin Li et al. propose a verification method for the simulation of a semi-physical train operation control theory. The method is based on an intelligent mobile robot [8]. In order to optimize the performance of the control system for Permanent Magnet Synchronous Motors (PMSM), an improved Model Predictive Control (MPC) scheme based on neural networks is proposed in [9]. In [10], the paper presents novel techniques aimed at reducing the occurrence of pole-slipping induced by control systems, along with a cost-effective detection and recovery scheme for magnetic drive-trains (MDTs). Boudallaa proposes the speed control of an asynchronous motor (AM) using the $H\infty$ Antiwindup design in his paper [11]. Paper [12] presents a fuzzy controller for

[DOI: 10.22068/ijrare.340]

tracking specified speed characteristics while maintaining a smooth ride for a passenger vehicle within a predetermined speed range. A genetic algorithm is employed to tune the performance of the fuzzy controller by adjusting its parameters, including scaling factors and membership functions. One of the simplest methods employed in [13] can be mentioned, where the genetic algorithm is utilized to estimate the control parameters of a PID controller for train speed control.

It was previously mentioned that there is a possibility of an error occurring in the control of the electric train speed by the human operator. Additionally, the automatic control of the train speed was studied using the usual methods. The aforementioned approaches require the use of a system's dynamic model for designing the controller. However, calculating the dynamics of the system is frequently accompanied by errors. As a result, this research aims to control the speed of the electric train using reinforcement learning. Therefore, by considering the state feedback controller and utilizing the Q-learning method, the control parameters will be calculated without the need for calculating the system's dynamic equations.

In Section 2, a general approach is presented for solving the problem. First, a dynamical system of the train is simulated, and then general explanations regarding state feedback control will be provided in Section 2.2. Section 2.3 will describe the fundamentals of reinforcement learning. Section 3 focuses on the design of the controller using reinforcement learning. Section 3.1 presents the equations related to state feedback control, and Section 3.2 elaborates on the equations for designing the controller using reinforcement learning. Moreover, in Sections 3.3 and 3.4, respectively, other artificial intelligence methods and evaluation metrics have been introduced. Section 4 is dedicated to the results, and Section 5 concludes the paper.

2. Approach

In this section, the dynamic model of the train is initially constructed to simulate the problem environment. Indeed, the dynamic model is utilized to apply a load to the electric motor. Finally, the state feedback controller and reinforcement learning will be studied for controlling the direct current motor.

2.1. Train dynamics

According to Newton's second law and the provided diagram, the dynamic equations of the train's motion will be represented by Equations 1-8 [14], [15].

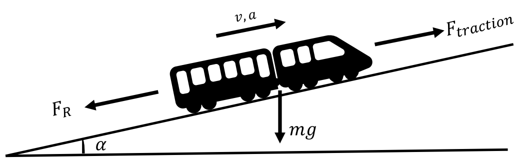


Figure 1. Schematic of the train's motion on an incline and the forces influencing its motion.

$$F_{Trac} - \sum F_R = M\frac{dv}{dt} \tag{1}$$

$$F_R = F_{rr} + F_{ar} + F_{gr} \tag{2}$$

$$F_{rr} = f_r M g \cos\alpha \tag{3}$$

$$F_{ar} = \frac{1}{2} C_w A \rho v^2 \tag{4}$$

$$F_{gr} = f_r M g \sin\alpha \tag{5}$$

$$F_{Trac} = f_r M g \cos\alpha + \frac{1}{2} C_w A \rho v^2 + f_r M g \sin\alpha + M\frac{dv}{dt} \tag{6}$$

$$T = \frac{F_{Trac} r}{4 n_c} \tag{7}$$

$$\omega_w = \frac{v}{r} \tag{8}$$

Using these equations, the motor load is determined. Figure 1 illustrates the forces applied to the train in motion.

2.2. State feedback control

[DOI: 10.22068/ijrare.340]

State Feedback Control is a control method in systems that adjusts the control input based on the system state information. It utilizes the state information of the system to regulate the control input [16]. In this method, the system's state is measured by sensors or other methods, and then the control input is applied to the system based on this state information. In state feedback control, a control weighting matrix called the state feedback gain matrix is used. This matrix is designed using various techniques, such as pole placement or the linear quadratic regulator (LQR) method. By appropriately setting the state feedback gain matrix, it is possible to achieve improved performance and more precise control of the system. By changing the values of the state feedback gain matrix, the system's characteristics can be altered, allowing for reaching desired outputs [17]–[19]. The state feedback control schematic is depicted in Figure 2.

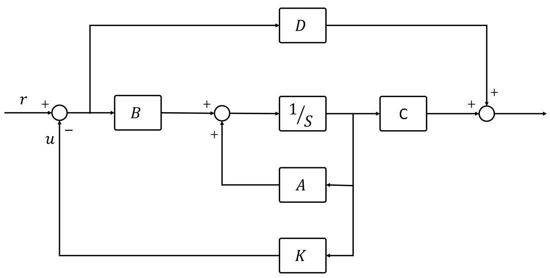


Figure 2. Schematic of a system with state feedback control.

2.3. Reinforcement learning

Reinforcement learning is a machine learning method that is based on the interaction of an agent with its environment. In this approach, an agent or machine is expected to learn how to act in order to achieve a specific goal through interacting with its environment. The primary objective of reinforcement learning is to improve and optimize the agent's actions in the environment. There are various methods for reinforcement learning, including algorithms such as Q-Learning, SARSA, and Temporal Difference [20]. Specifically, in this research, we will use reinforcement learning to control the speed of an electric train [21]. The process of controlling the speed of an electric train using reinforcement learning is illustrated in Figure 3. The actions of the agent are the control signals that are applied to the environment. Through interaction with the environment, the agent learns which control signal or action to apply in each state.

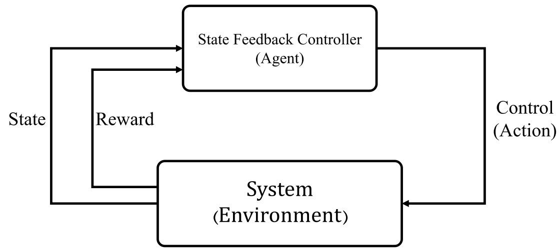


Figure 3. State feedback control process through reinforcement learning.

3. Train speed controller

In this section, the state space model of the electric train motor is first calculated. Taking into account the state feedback controller, the intended state space is considered the environment in the learning process. The agent, which is the system's controller itself, interacts with the environment to carry out the learning process, which will be further elaborated upon.

3.1. State space of an electric motor

In this section, the state space equations of the electric train motor are considered as in Equation 9. The specifications and values of the parameters used are provided in Table 1.

$$A = \begin{bmatrix} -R/L & 0 & -K_b/L \\ 0 & 0 & 1 \\ -K_t/J & 0 & -B/J \end{bmatrix}$$
$$B = \begin{bmatrix} 1/L \\ 0 \\ 0 \end{bmatrix} \tag{9}$$
$$C = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$$
$$D = 0$$

[DOI: 10.22068/ijrare.340]

Table 1. Parameters of an electric motor.

Parameter	Value	Unit
R	0.5	Ohm
L	0.003	H
B	0.01	$N.m.s$
J	0.0167	$kg.m^2$
K_t	0.8	$N.m/Amp$
K_b	0.8	$V/rad/s$

The equations for simulating the environment have been introduced. To ensure confidence, the step response of the system is depicted in Figure 4, and the controllability matrix has been computed, indicating that the system is controllable.

$$cntrb = \begin{bmatrix} 3.33 \times 10^2 & -5.55 \times 10^4 & 5.001 \times 10^6 \\ 0 & 0 & 1.596 \times 10^4 \\ 0 & 1.596 \times 10^4 & -2.67 \times 10^6 \end{bmatrix}$$

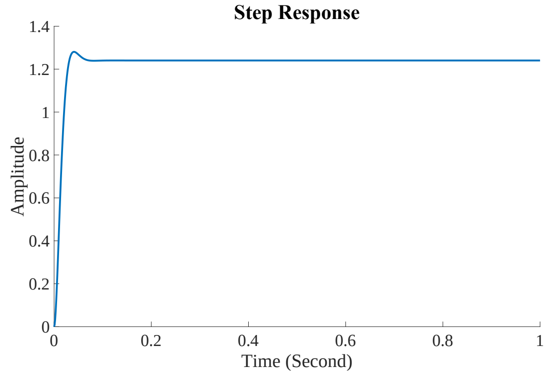


Figure 4. Step response of the system.

3.2. Controller

The general overview of control systems controlled by closed-loop controllers is illustrated in Figure 5. In order to tune the controllers, the system dynamics are typically computed, and the controller is designed accordingly. However, in many systems, the dynamic model is either unavailable or its estimation is prone to errors. Moreover, the process of controller design and its implementation is often performed offline. To design an optimal adaptive controller without the need for a dynamic model (model-free) and online, the Q-Learning technique can be utilized. By having the Q-function, the optimal policy can be determined without relying on the system model. The function $Q_h(x_k, u_k)$ specifies the value associated with taking action u_k in state x_k. In other words, if the system is in state x_k, the controller performs action u_k, and this control action represents how optimal the control is. Equation 10 illustrates how $Q_h(x_k, u_k)$ is computed.

$$Q_h(x_k, u_k) = r(x_k, u_k) + \gamma V_h(x_{k+1}) \qquad (10)$$

$$h^*(x_k) = \underset{u}{\operatorname{argmin}}(Q^*(x_k, u)) \qquad (11)$$

In Equation 11, $h^*(x_k)$ indicates that for a certain range of values for u, Q becomes optimal. By taking the derivative of Equation 10 with respect to u_k, in other words, for which action the maximum reward is obtained, the optimal policy can be computed. The optimality equation for the Q-function, known as the

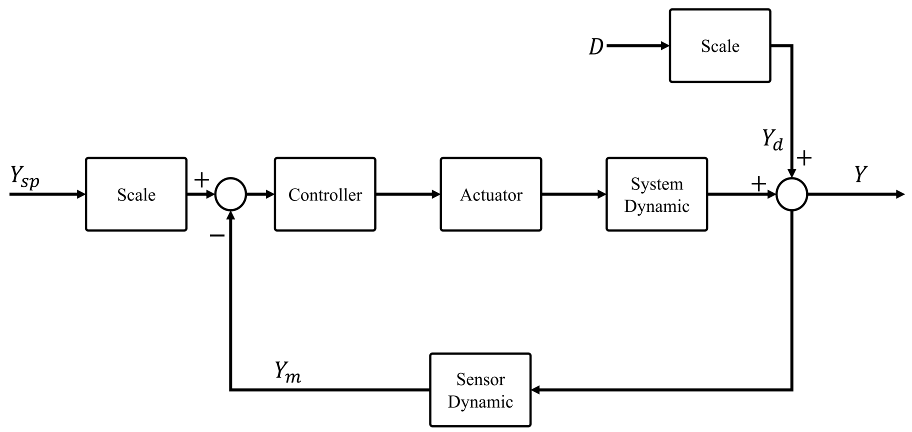


Figure 5. Controller diagram

[DOI: 10.22068/ijrare.340]

Bellman equation, is expressed by the following equation:

$$Q^*(x_k,u_k) = r(x_k,u_k) + \gamma V_h(x_{k+1}) \quad (12)$$

The Bellman equations for the Linear Quadratic Regulation (LQR) problem are defined as a second-order function of the state and action, as follows:

$$Q_k(x_k,u_k) = x_k^T Q x_k + u_k^T R u_k$$

$$+x_{k+1}^T P x_{k+1} = x_k^T Q x_k + u_k^T R u_k \quad (13)$$

$$+(Ax_k + Bu_k)^T P (Ax_k + Bu_k)$$

$$Q_k(x_k,u_k)$$

$$= \begin{bmatrix} x_k \\ u_k \end{bmatrix}^T \begin{bmatrix} Q + A^T PA & B^T PA \\ A^T PB & R + B^T PB \end{bmatrix} \begin{bmatrix} x_k \\ u_k \end{bmatrix} \quad (14)$$

$$= Z_k^T H Z_k$$

$$\bar{H}^T \bar{Z}_k = x_k^T Q x_k + u_k^T R u_k + \bar{H}^T \bar{Z}_{k+1} \quad (15)$$

$$\bar{H}^T \bar{Z}_k - \bar{H}^T \bar{Z}_{k+1} = x_k^T Q x_k + u_k^T R u_k \quad (16)$$

$$Q_k(x_k,u_k) = \begin{bmatrix} x_k \\ u_k \end{bmatrix}^T \begin{bmatrix} H_{xx} & H_{xu} \\ H_{ux} & H_{uu} \end{bmatrix} \begin{bmatrix} x_k \\ u_k \end{bmatrix} \quad (17)$$

$$\frac{\partial Q_k(x_k,u_k)}{\partial u_k} = 0 \quad (18)$$

$$H_{ux} x_k + H_{uu} u_k = 0$$

$$\rightarrow u_k = -(H_{uu})^{-1} H_{ux} x_k \quad (19)$$

Equation 13 represents the Bellman optimality equation, where the state equations for x_{k+1} are substituted. This equation is then simplified in Equation 14 to obtain the form shown in Equation 15. In this context, $\bar{H}$ refers to $Vec(H)$, and $\bar{Z}$ denotes the Kronecker product. After simplification in Equations 16 and 17, the derivative with respect to u can be taken, and the parameter values for K can be determined. In this case, where H includes the parameters of the system's dynamic equations, it is obtained through interaction with the environment and the technique of least squares.

The pseudocode for executing a controller to control the speed of an electric train is provided in Table 2:

Table 2. Proposed algorithm.

```
Initializations:
Itr : total iterations
M : must bigger than H̄s parameters
K : state feedback
% Calculate T, B1, C1 and Q1 For
% constructing Bellman equations
T = [A 0; 0 1]
B1 = [B; 0]
C1 = [C  -1]
Q1 = C1^T Q C1
% Main Loop to Calculate K
For i=1:Itr-1
    ϕ : consider as empty
    ψ : consider as empty
    For k=1:M
        u_k = -K x_k
        x_{k+1} = T x_k + B1 u_k
        u_{k+1} = -K x_{k+1}
        Attach ϕ as row :
            Z̄([x_k; u_k]) - γZ̄([x_{k+1}; u_{k+!}])
        Attach ψ as row :
            x_k^T Q1 x_k + u_k^T R u_k
    End
    H̄ = (ϕ^T ϕ)^{-1} ϕ^T ψ
    % Compute H from H̄
    % Compute H_ux and H_uu from H
    K = (H_uu)^{-1} H_ux
    % Termination Criteria
    If ‖(K_j - K_{j-1})‖ < 10^{-5} :
        Break
    End
End
```


[DOI: 10.22068/ijrare.340]

3.3. Machine learning methods

In this section, machine learning algorithms will be reviewed to find the parameters of the feedback controller.

3.3.1. Particle swarm optimization (PSO)

Particle swarm optimization (PSO) is a computational technique used for optimization problems. It is inspired by the social behavior of birds flocking or fish schooling. In PSO, a population of candidate solutions, called particles, moves through the search space to find the optimal solution. Each particle adjusts its position based on its own experience (local best) and the collective experience of the swarm (global best). Through this iterative process, particles converge towards the optimal solution [22].

3.3.2. Genetic algorithm (GA)

Genetic algorithm (GA) is a heuristic optimization technique inspired by the process of natural selection and genetics. It operates by mimicking the principles of evolution, where candidate solutions, often represented as chromosomes or strings of parameters, undergo selection, crossover, mutation, and reproduction to produce offspring solutions iteratively. The fitness of each solution determines its likelihood of being selected for reproduction, with fitter solutions having a higher chance of passing their genetic material to the next generation. Over successive generations, the population evolves, and through the process of natural selection, increasingly better solutions are generated [23].

3.3.3. Bee Algorithm (BA)

The Bee Algorithm is a nature-inspired optimization algorithm that mimics the foraging behavior of honeybee colonies. In this algorithm, three main types of bees are simulated: employed bees, onlooker bees, and scout bees. Employed bees exploit the search space by visiting neighboring solutions; onlooker bees select solutions based on the information shared by employed bees; and scout bees explore new areas of the search space [24].

The algorithm starts with an initial population of solutions, represented as food sources, and iteratively improves these solutions to find the optimal solution. Employed bees search for food sources in their vicinity and evaluate their quality using a fitness function. Onlooker bees choose food sources to visit based on the waggle dance information shared by employed bees. Scout bees randomly search for new food sources when the employed and onlooker bees cannot find better solutions.

Through the collaboration between employed, onlooker, and scout bees, the algorithm efficiently explores the search space and converges towards the optimal solution. The Bee Algorithm has been successfully applied to various optimization problems, including engineering design, scheduling, and data clustering.

3.4. Metrics

There are several metrics commonly used to evaluate algorithms, depending on the nature of the problem being solved and the goals of the optimization process. Some of the key metrics mentioned for the problem include the following:

3.4.1. Root mean squared error (RMSE)

Root mean squared error (RMSE) is a widely used metric for evaluating the accuracy of predictive models, particularly in regression analysis. It measures the average magnitude of the differences between predicted values and observed values.

Mathematically, RMSE is calculated as the square root of the average of the squared differences between predicted and observed values:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad (20)$$

Where:

- n is the number of observations;

[DOI: 10.22068/ijrare.340]

- y_i is the observed value for the i^{th} observation; and

- $\hat{y}_i$ is the predicted value for the i^{th} observation.

RMSE provides a measure of how accurately the model's predictions match the observed data. It is preferred because it penalizes larger errors more heavily than smaller errors, making it sensitive to outliers and providing a more comprehensive assessment of model performance.

Lower RMSE values indicate better model performance, with a value of 0 indicating perfect predictions (i.e., the model exactly matches the observed data). However, RMSE values should be interpreted in the context of the specific problem and the scale of the target variable.

3.4.2. R-squared (R^2)

The R-squared (R^2) metric, also known as the coefficient of determination, is a statistical measure used to evaluate the goodness of fit of a regression model to the observed data. It quantifies the proportion of the variance in the dependent variable that is explained by the independent variables in the model. Mathematically, R^2 is calculated as:

$$R^2 = 1 - \frac{SS_{res}}{SS_{total}} \tag{21}$$

Where:

- SS_{res} is the sum of the squares of the residuals (the differences between the observed and predicted values); and

- SS_{total} is the total sum of squares, which measures the total variance of the dependent variable around its mean.

R^2 can take values between 0 and 1. A value of 1 indicates that the regression model perfectly fits the data, explaining all the variability in the dependent variable. On the other hand, a value of 0 indicates that the model does not explain any of the variability in the dependent variable, and its predictions are equivalent to simply using the mean of the dependent variable.

Interpretation of R^2:

- A high R^2 value (close to 1) indicates that the model explains a large proportion of the variability in the dependent variable and is considered a good fit.

- A low R^2 value (close to 0) suggests that the model does not explain much of the variability in the dependent variable and may not be useful for prediction.

It is important to note that R^2 is a relative measure and should be interpreted in the context of the specific problem and compared with the R^2 values of alternative models. Additionally, R^2 can be influenced by the number of independent variables in the model, so caution should be exercised when comparing models with different numbers of predictors.

3.4.3. Explained variance

The explained variance metric quantifies the proportion of variance in the dependent variable that is explained by the independent variables in a regression model. It is often used in conjunction with regression analysis. Mathematically, the explained variance is calculated as:

$$Explained\ Variance = 1 - \frac{Var(y-\hat{y})}{Var(y)} \tag{22}$$

Where:

- $Var(y-\hat{y})$ is the variance of the residuals (the differences between the observed and predicted values); and

- $Var(y)$ is the variance of the dependent variable.

The explained variance metric provides a measure of how well the independent variables in the model account for the variability in the dependent variable. Like R^2, it ranges between 0 and 1, with higher values indicating that a larger proportion of the variance in the dependent variable is explained by the model.

Interpretation of explained variance:

[DOI: 10.22068/ijrare.340]

$$H = \begin{bmatrix} \bar{H}(1) & 0.5\bar{H}(2) & 0.5\bar{H}(3) & 0.5\bar{H}(4) & 0.5\bar{H}(5) \\ 0.5\bar{H}(2) & \bar{H}(6) & 0.5\bar{H}(7) & 0.5\bar{H}(8) & 0.5\bar{H}(9) \\ 0.5\bar{H}(3) & 0.5\bar{H}(7) & \bar{H}(10) & 0.5\bar{H}(11) & 0.5\bar{H}(12) \\ 0.5\bar{H}(4) & 0.5\bar{H}(8) & 0.5\bar{H}(11) & \bar{H}(13) & 0.5\bar{H}(14) \\ 0.5\bar{H}(5) & 0.5\bar{H}(9) & 0.5\bar{H}(12) & 0.5\bar{H}(14) & \bar{H}(15) \end{bmatrix} \quad (23)$$

- A high explained variance value (close to 1) suggests that the independent variables in the model effectively explain the variability in the dependent variable.

- A low explained variance value (close to 0) indicates that the model does not explain much of the variability in the dependent variable, and its predictions are not reliable.

The explained variance metric is particularly useful for assessing the predictive power of regression models and determining how well they capture the underlying patterns in the data. It complements other evaluation metrics such as R² and mean squared error (MSE) and helps researchers and practitioners gauge the overall performance of regression models.

4. Results

The Q-Learning algorithm mentioned reached the final solution after nine iterations. The termination condition for this algorithm is the change in the values of K for two consecutive iterations. In such a way that if the norm of the difference between two consecutive K values becomes less than 10^{-5}, the algorithm terminates. Considering that the value of H is obtained from the least squares technique in each iteration of the algorithm, the number of calculated parameters for x_k and u_k should be greater than the number of parameters in $\bar{H}$ in each iteration. After calculating the vector $\bar{H}$ using the least squares technique, it should be transformed into the matrix H. Given the problem, the matrix H is obtained as Equation 23.

More details about the parameter values in the reinforcement learning algorithm are given in Table 3.

Table 3. Specifications of Reinforcement Learning (RL).

Stopping Criteria	$\|K_j - K_{j-1}\|$
Last Utility	1487
R	1
Q	$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 200 & -200 \\ 0 & 0 & -200 & 200 \end{bmatrix}$

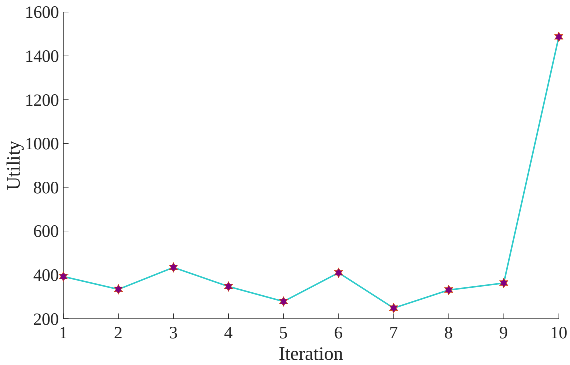


Figure 6. Utility - iteration in the RL algorithm.

Figure 6 shows the utility value for each episode. Now, based on the obtained values of H, the values of H_{ux} and $H_{uu}$are derived. In this specific problem, these values are explained in the given equation. In these equations, H_{ij} represents the element in the ith row and jth column of matrix H. The method for obtaining the control signal for this problem is given in Equation 24.

$$H_{ux} = [H_{51} \quad H_{52} \quad H_{53} \quad H_{54}]$$
$$H_{uu} = H_{55} \quad (24)$$
$$u_k = -(H_{uu})^{-1} H_{ux} x_k$$

[DOI: 10.22068/ijrare.340]

To calculate $\bar{Z}$, the Kronecker product is used, and the method for calculating it for this problem is given in Equation 25.

$$\bar{Z} = \begin{bmatrix} z_1{}^2 \\ z_1 z_2 \\ z_1 z_3 \\ z_1 z_4 \\ z_1 z_5 \\ z_2{}^2 \\ z_2 z_3 \\ z_2 z_4 \\ z_2 z_5 \\ z_3{}^2 \\ z_3 z_4 \\ z_3 z_5 \\ z_4{}^2 \\ z_4 z_5 \\ z_5{}^2 \end{bmatrix} \tag{25}$$

Based on the initial explanation in this section, the algorithm has reached the solution after nine iterations. Figure 8 represents the speed profile of the train, and it is expected that the controller will track this profile. In this figure, the speed profile of the train is designed for three phases: acceleration, constant speed, and braking.

The values of the state feedback parameters (K) and the convergence of the solution are shown in Figure 7. The values K_1, K_2, K_3, and K_4 have converged to their final values in iterations 6, 2, 3, and 8, respectively. The initial values of K are taken as the vector [0.3 1.3 0.75 0.1]. Furthermore, the discount factor γ is 0.8. The values of R and Q are 1 and 200, respectively. However, changing these values will have an impact on the final solution.

Based on the provided values, the output of the controlled speed by the mentioned algorithm is shown in Figure 7. As evident, the speed control has been performed effectively. Using the obtained values, the performance of RL can be evaluated, which will be expressed as Equation 25:

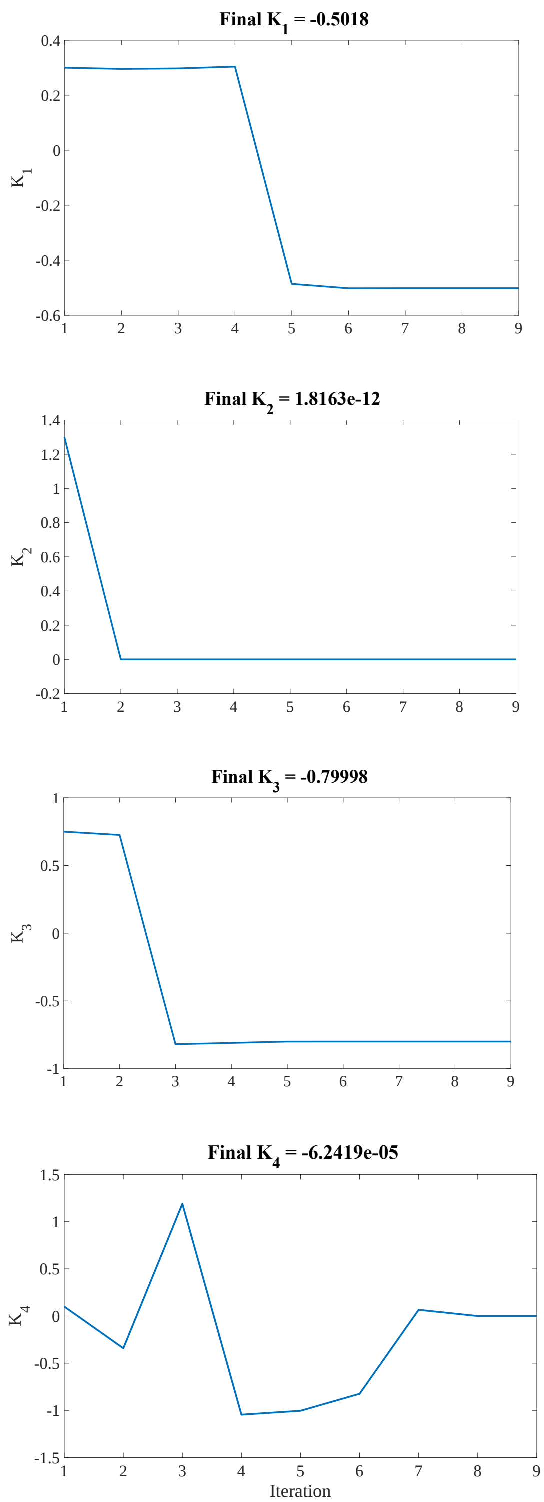


Figure 7. Training process of state feedback controller parameters.

[DOI: 10.22068/ijrare.340]

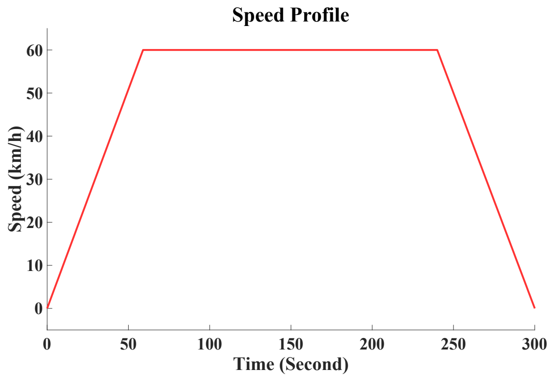
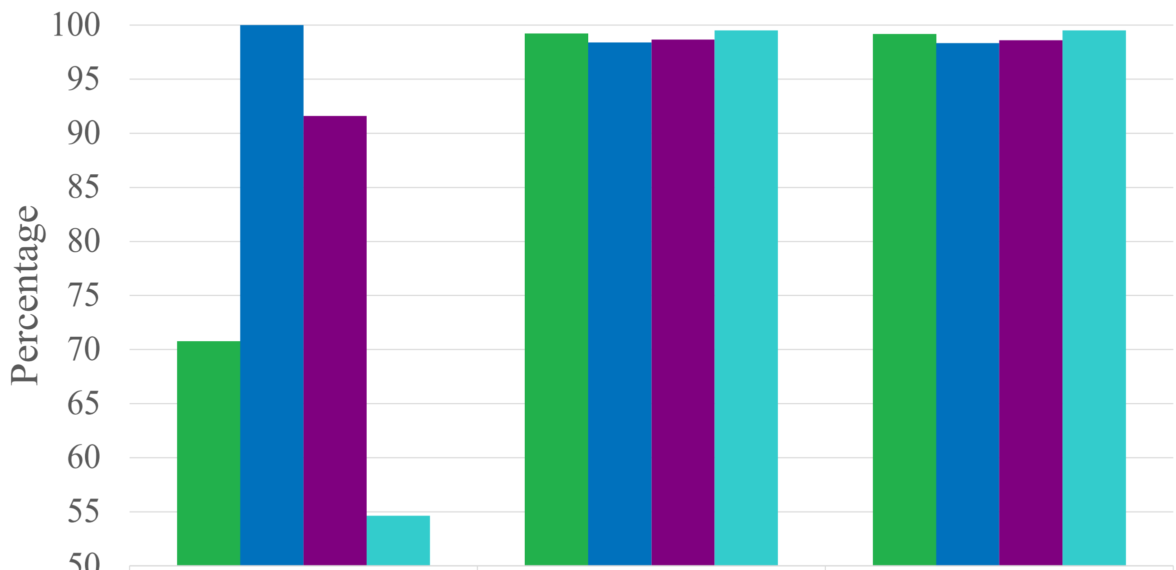


Figure 8. Sample of an electric train speed profile.

$$u_k = Kx_k$$
$$x_{k+1} = Ax_k + Bu_k \qquad (25)$$
$$y_k = Cx_k + Du_k$$

Additionally, the capability of other algorithms to find feedback parameters for speed control of the train has also been depicted in this figure. As can be observed from Figure 10, methods such as GA, PSO, and BA exhibit oscillations. These oscillations are significantly higher at the beginning of the train movement, but overall, when transitioning phases, oscillations occur. Among the algorithms, GA exhibits the highest oscillations, while PSO has the lowest. As seen, the RL algorithm controls the train speed without oscillation.

Furthermore, in Figure 10, the algorithms used in this paper were compared with each other using the metrics introduced in Section 3. However, for better comparison, the metrics were presented as percentages. Evidently, the RL algorithm outperformed others in all metrics. More detailed information is provided in Figure 9.



	RMSE	R^2	Explained Variance
PSO	70.7624395	99.201	99.18
BA	100	98.404	98.342
GA	91.5827369	98.661	98.611
RL	54.6436378	99.523	99.513

Figure 9. Comparison of the algorithms utilized based on the introduced metrics in terms of percentage.

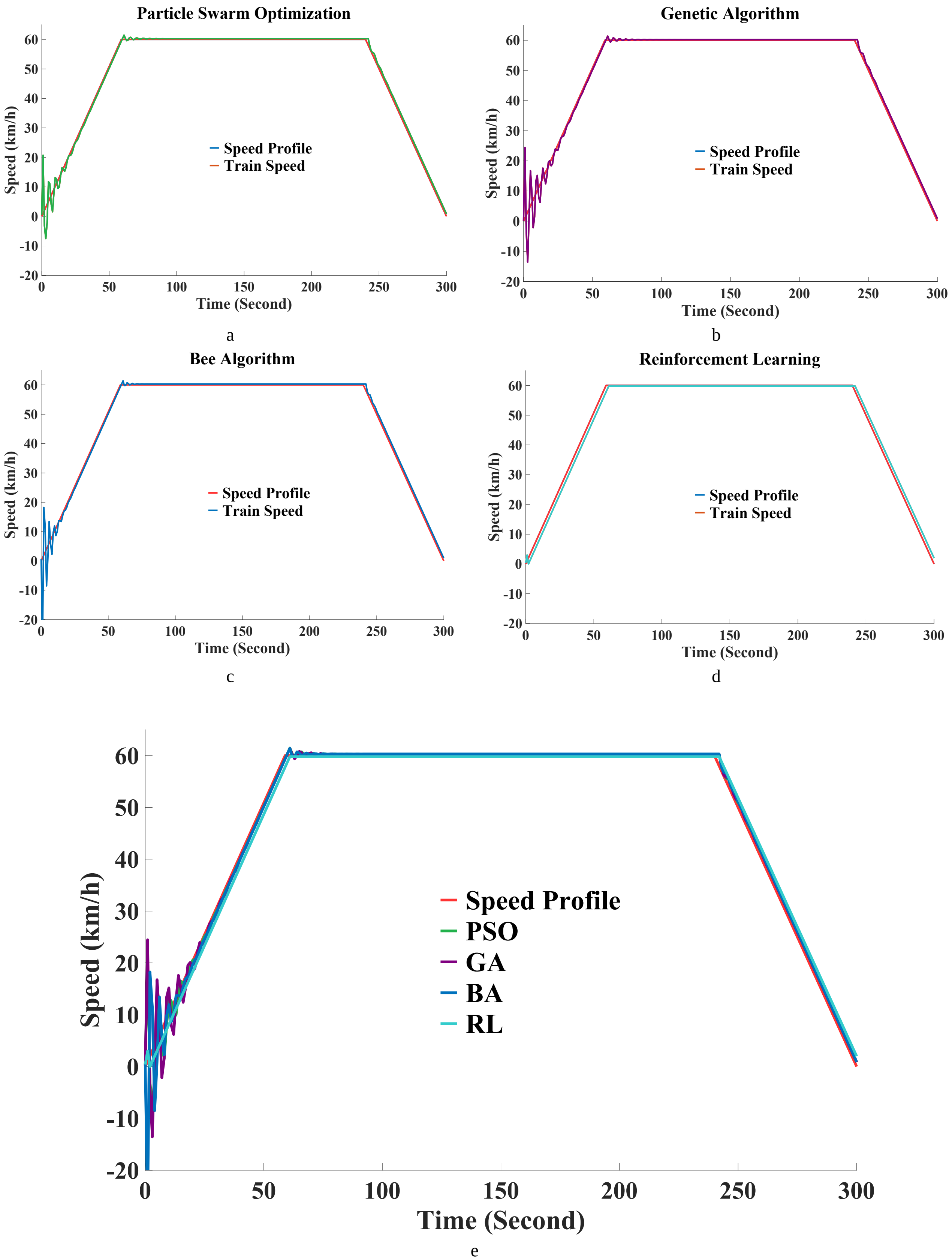


Figure 10. Speed control of electric trains considering the three-phase set point: acceleration, constant speed, and braking. a) PSO, b) GA, c) BA, d) RL, and e) comparison of all mentioned algorithms.

[DOI: 10.22068/ijrare.340]

5. Conclusion

For reasons such as fatigue, disregarding speed limits, and lack of focus, automatic speed control of trains is of great importance. Therefore, various methods are suggested for speed control, including controllers such as proportional integral derivative (PID), fuzzy control, or the use of neural networks as speed controllers. Most controllers require the dynamic model of the system for tuning and design purposes. Additionally, they typically operate offline, meaning they don't require real-time interaction with the system during control. In this research, automatic speed control of an electric train was performed using reinforcement learning (RL). In this method, there is no need for a dynamic model of the electric train system. The controller learns and performs speed control in an online manner, interacting with the training environment. As investigated in this research, a state feedback controller was used, which was trained using reinforcement learning. It successfully tracked the two-phase speed profile of the electric train. Moreover, other artificial intelligence methods were used for comparison with the introduced method. Using metrics such as root mean squared error (RMSE), R-squared, and explained variance, it was shown that the mentioned method has superior performance. For example, in the RMSE metric, the RL algorithm had 46% superior performance compared to the Bee algorithm (BA).

List of symbols

F_{Trac}	Traction force
F_R	Resistive force
M	Mass of train
v	Speed of train
F_{rr}	Rolling resistive force
F_{ar}	Aerodynamics drag force.
F_{gr}	The gradient force arises from the rail's slope or inclination.
f_r	Rolling resistance coefficient
g	Acceleration due to gravity (9.81 m/s^2)
C_w	Drag coefficient
A	Projected frontal area of the vehicle/train
n_c	Numbers of cars of the train
r	Radius of the train's wheel
T	Torque of the train's car
R	Electric resistance
L	Electric inductance
B	Motor viscous friction constant
J	Moment of inertia of the rotor
K_t	Motor torque constant
K_b	Electromotive force constant
Q^*	Optimal Q function
Q_k	Q-function
x_k	State k
u_k	Control signal k
V_h	Value function
Q	Symmetric matric
R	Symmetric matric
P	Solution of the LQT
r	Reward
K_j	State feedback parameters at j^{th} generation
K_{j-1}	State feedback parameters at $(j-1)^{th}$ generation
Greek symbols	
α	Inclination angle
ρ	Air density
γ	Discount factor
ω_w	Angular velocity

References

[1] Rail Safety and Standards Board Limited, "Annual Safety Performance Report," *Rail Saf. Stand. Board*, no. July 2016, p. 53, 2015.

[2] X. Guo, W. Sun, S. Yao, and S. Zheng, "Does high-speed railway reduce air pollution along highways? —— Evidence from China," *Transp. Res. Part D Transp. Environ.*, vol. 89, p. 102607, Dec. 2020, doi: 10.1016/j.trd.2020.102607.

[3] H. Pham and H. Wang, *Springer Series in Reliability Engineering*. 2006. doi: 10.1007/978-1-4471-4588-2.

[4] S. Menicanti, M. di Benedetto, D. Marinelli, and F. Crescimbini, "Recovery of Trains' Braking Energy in a Railway Micro-Grid Devoted to Train plus Electric Vehicle Integrated Mobility," *Energies*, vol. 15, no. 4, p. 1261, Feb. 2022, doi: 10.3390/en15041261.

[5] G. S. Larue and A. Naweed, "Evaluating the effects of automated monitoring on driver

non-compliance at active railway level crossings," *Accid. Anal. Prev.*, vol. 163, p. 106432, Dec. 2021, doi: 10.1016/j.aap.2021.106432.

[6] A. J. Filtness and A. Naweed, "Causes, consequences and countermeasures to driver fatigue in the rail industry: The train driver perspective," *Appl. Ergon.*, vol. 60, pp. 12–21, Apr. 2017, doi: 10.1016/j.apergo.2016.10.009.

[7] Z. Wang and D. Ou, "Parameter Adaptive Research of Automatic Train Control Algorithm Based on Sliding Mode PID," *Transp. Res. Rec.*, 2023, doi: 10.1177/03611981231182707.

[8] J. Li, X. Pei, H. Liu, S. Su, T. Tang, and T. Hou, "A Novel Train Operation Simulation System Based on Intelligent Mobile Robot and ROS Communication Network," *Proc. 34th Chinese Control Decis. Conf. CCDC 2022*, no. 62103035, pp. 97–102, 2022, doi: 10.1109/CCDC55256.2022.10033650.

[9] H. Mao, X. Tang, and H. Tang, "Speed control of PMSM based on neural network model predictive control," *Trans. Inst. Meas. Control*, vol. 44, no. 14, pp. 2781–2794, 2022, doi: 10.1177/01423312221086267.

[10] X. Liao, C. Bingham, and T. Smith, "Speed Control of Magnetic Drive-Trains with Pole-Slipping Amelioration," *Energies*, vol. 15, no. 21, pp. 1–14, 2022, doi: 10.3390/en15218148.

[11] A. Boudallaa, M. Chennani, D. Belkhayat, and K. Rhofir, "Vector Control of Asynchronous Motor of Drive Train Using Speed Controller H∞," *Emerg. Sci. J.*, vol. 6, no. 4, pp. 834–850, 2022, doi: 10.28991/ESJ-2022-06-04-012.

[12] P. P. Bonissone, P. S. Khedkar, and Y. Chen, "Genetic algorithms for automated tuning of fuzzy controllers: a transportation application," in *Proceedings of IEEE 5th International Fuzzy Systems*, IEEE, 1996, pp. 674–680.

[13] S. K. Suman and V. K. Giri, "Speed control of DC motor using optimization techniques based PID Controller," *Proc. 2nd IEEE Int. Conf. Eng. Technol. ICETECH 2016*, no. March, pp. 581–587, 2016, doi: 10.1109/ICETECH.2016.7569318.

[14] C. J. Goodman, "Overview of electric railway systems and the calculation of train performance," in *IET Professional Development Course on Electric Traction Systems*, IEE, 2008, pp. 1–24. doi: 10.1049/ic:20080503.

[15] D. Crolla and B. Mashadi, *Vehicle powertrain systems*. John Wiley & Sons, 2011.

[16] E. D. Sontag, *Mathematical control theory: deterministic finite dimensional systems*, vol. 6. Springer Science & Business Media, 2013.

[17] Q. Gao and H. R. Karimi, *Stability, control and application of time-delay systems*. Butterworth-Heinemann, 2019.

[18] S. G. Braun, D. J. Ewins, S. S. Rao, and A. W. Leissa, "Encyclopedia of vibration: Volumes 1, 2, and 3," *Appl. Mech. Rev.*, vol. 55, no. 3, pp. B45–B45, 2002.

[19] M. A. Christodoulou, "Analysis and synthesis of singular systems," *Ph. D. Thesis*. Department of Electrical Engineering, Democritus University of Thrace Xanthi …, 1984.

[20] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[21] S. E. Li, *Reinforcement learning for sequential decision and optimal control*. Springer, 2023.

[22] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, IEEE, pp. 1942–1948. doi: 10.1109/ICNN.1995.488968.

[23] M. Mitchell, *An introduction to genetic algorithms*. MIT press, 1998.

[24] D. Karaboga, "An idea based on honey bee swarm for numerical optimization," Technical report-tr06, Erciyes university, engineering faculty, computer …, 2005.

[DOI: 10.22068/ijrare.340]