International Journal of Railway Research, Vol. 10, No. 1, (2023), 84-92

Iran University of Science and Technology

International Journal of

Railway Research

ISSN: 2423-3838

Optimizing Railway Energy Consumption with Multiple-Phase Optimal Control Method

Fatemeh Bakhshi[1], Mohammad Ali Sandidzadeh[2*], Saeed Ebadollahi[3]

[1]MSc. Student, School of Railway Engineering, Iran University of Science & Technology

[2]Associate Professor, School of Railway Engineering, Iran University of Science & Technology

[3]Assistant Professor, School of Electrical Engineering, Iran University of Science & Technology

ARTICLE INFO

Article history:

Received: 08.04.2023

Accepted: 09.17.2023

Published: 09.20.2023

Keywords:

Energy Consumption

Train Speed Profile

Multiphase Optimal Control

Pseudo-spectral Method

GPOPS Software

A B S T R A C T

In many countries, railways, including intercity, urban railway (metro), tram, monorail, and other lines, have played a vital role in passenger and freight transportation in the past. The problem of optimizing energy and reducing air pollution has always been a very important issue in transportation systems. Given that trains are faster and more accessible for passenger and freight transportation and consume less energy compared to other public transportation vehicles, the use of railway lines has gained significant attention from many countries, individuals, and various companies. Nowadays, with the limitation of resources, the importance of optimal energy consumption has received a lot of attention from researchers, and various methods have been proposed in this regard. As is well known, trains travel on railway lines based on speed profiles and control tables designed by the interlocking system, which have a close relationship with energy consumption. In recent years, investment in providing methods for designing train speed profiles with the aim of optimizing energy consumption in railway transportation systems has increased. There are various methods for optimizing speed profiles, among which the use of optimal control theory can be mentioned. In this article, the multiple-phase optimal control theory has been used to design train speed profiles. First, the train movement is divided into several phases based on the route layout, and then for each phase, dynamic equations and cost functions have been written to optimize energy consumption. In the next step, the solution to this multiple-phase problem will be done using pseudo-spectral methods and the GPOPS software. Finally, the designed speed profile for trains has been used for the Tehran-Mashhad path in the Abardej-Kavir section, and the level of energy consumption and optimization using this method has been discussed and investigated.

1. Introduction

In recent years, investment in optimizing train speed profiles with the aim of reducing energy consumption in public transportation systems has greatly increased. Proper energy management in public transportation, especially railways and metros, not only leads to significant savings at a macro level but also allows for the optimization and updating of train schedules through intelligent movement. The effort to achieve an optimal speed profile is actually a type of optimization problem that seeks to reach an appropriate speed profile. One of the effective ways to achieve energy efficiency is to optimize train performance and subsequently optimize train speed profiles through movement

* Corresponding author

Email address: sandidzadeh@iust.ac.ir

[DOI: 10.22068/ijrare.329]

strategies. Even if each train saves a negligible amount of energy during operation, the total operating costs of the rail network are significantly reduced.

Research in this field has so far focused on using optimal control theory to find the optimal trajectory for the train (speed profiles in terms of position and position in time during travel), which results in safe and real-time train operations, passenger comfort, and reduced energy consumption [1]. The problem of optimizing train energy consumption is considered as an optimal control problem and is then addressed by using existing methods to solve this problem, resulting in the optimization of train speed, which leads to rail traffic efficiency. Specifically, this article examines two important aspects: train trajectory optimization and the scheduling problem for energy efficiency.

In fact, the problem of this article is an optimal control problem. To solve this, the movement of the train will first be modeled. To model the train movement, all factors affecting the movement of the train, including resistance forces, the effects of signaling systems on movement, and other relevant factors, are considered to obtain an accurate model of the train movement. In the next step, the optimal control problem will be written for the train movement. The optimal control problem for train movement is generally described as an energy consumption optimization problem when the train travels a predetermined distance from x_0 to x_f within a predetermined time interval from t_0 to t_f. Therefore, a cost function will be written, taking into account the boundary conditions of the problem along with all the constraints and limitations governing the problem.

Optimal control problems can also be expressed as multi-phase problems. That is, the time interval of the problems is formed by combining several small sequential time intervals. Each of these sub-intervals can be considered an independent optimization problem. However, there are connections between the state and control variables within the initial and final stages of each phase, which leads to the integration of the overall problem. These phases can be either fixed or free [2]. Using this method, the optimal control problem for train movement is transformed into a multi-phase optimal control problem. This is because the train path has various gradients and different speed constraints. Each phase of the optimal control problem for train movement has its own dynamic constraints, path constraints, boundary conditions, and cost function. The optimal trajectory or optimal speed profile will be obtained by minimizing the sum of the cost functions of each phase.

After expressing the optimal control problem for train movement as a multiple-phase optimal control problem, the next step is to solve it. There are various methods for solving this problem. For example, Pontryagin's maximum principle can be used to find optimal movement strategies. However, this method leads to the solution of differential and integral equations and requires a lot of time for analysis processes. In this article, a direct addressing method is used to solve the optimal control problem. First, the optimal control problem for the train is transformed into a nonlinear programming problem using discrete pseudo-spectral methods, and then a gradient-based optimization solver in the Gauss Pseudospectral Optimization Software (GPOPS) is used to solve the nonlinear programming problem.

So far, extensive research has been conducted on reducing energy consumption in rail transportation, among which the most important ones include the use of neural networks and genetic algorithms [3], fuzzy control [4], and optimal control theory [5].

Optimization algorithms for speed profiles not only reduce energy consumption during travel but also help achieve the timetable goals of train movements [6]. Optimizing the train speed profile is also called train trajectory optimization (TTO). The optimized trajectory is the basis for automatic train operation systems to control train movements, as well as the basis for guidance systems for train operators to provide recommendations for speed and control regimes so that the train can move safely and efficiently [7].

Methods for solving the problem of optimizing the train trajectory, and specifically the speed profile of the train, can be divided into two categories: indirect methods and direct methods. Indirect methods solve the problem indirectly by transforming the optimal control problem into a boundary value problem. Direct

[DOI: 10.22068/ijrare.329]

methods find the optimal solution by converting the continuous-time optimization problem into a nonlinear programming problem (NLP). Researchers who focus on indirect methods are interested in solving differential equations to a large extent, while researchers who focus on direct methods use optimization approaches to solve the problem. The Pontryagin maximum principle (PMP) is an example of an indirect method [8].

2. Literature review

Research conducted on the TTO issue aims to optimize the unique movement of a train. The Pontryagin maximum principle is widely used to analyze the optimal control strategy for achieving energy efficiency in train movement [9].

In 1980, Milroy conducted his doctoral thesis in the field of automatic control aspects of trains. This included new work in the area of train navigation strategies. Based on Pontryagin's maximum principle, he demonstrated that for short trips, such as those occurring in metro systems, an appropriate strategy for minimal energy movement consists of three stages shown in Figure 1 [5]:

- Initial maximum acceleration;
- Coasting; and
- Maximum braking.

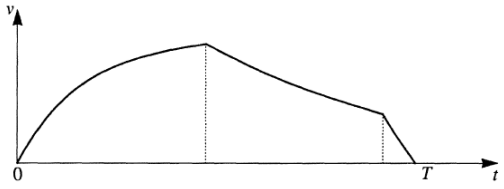

Figure 1. Optimal movement with three steps

After research conducted by Milroy, in 1982, research on the optimal motion strategy was continued by one of the graduates of the same college, Kim Taylor, and one of the students of the mathematics department, David Lee, under the supervision of Milroy. In 1982, they obtained four states for the problem of optimal control of a train on a long journey on a railroad without gradient, with sufficient complementary time for the trip, according to the principle of Pontryagin's maximum [10].

This motion strategy consists of the following stages and is shown in Figure 2:

- Initial maximum acceleration;
- Constant speed;
- Coasting; and
- Maximum braking.

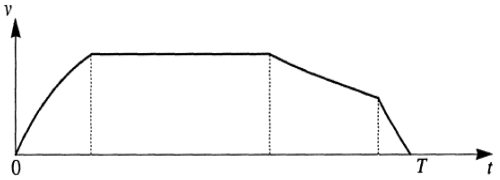

Figure 2. Optimal movement with four steps

After this research, according to the four modes obtained, the goal of most train control algorithms is to obtain optimal switching points between these modes [5].

In 1996, Howlett, in 2000, Khmelnitsky, and in 2003, Liu and Golovitcher concluded that for the movement of a train on a path with speed limitations and different gradients, the optimal control strategy is a sequence of states in which the transition points from one state to another depend on speed and gradient constraints [11-12].

In 2016, Albrecht et al. found that finding optimal switching points is a challenging problem, except in cases where the train's movement path has only one speed constraint and the crossing path is without gradients [13].

Another approach to solving the TTO problem is based on discretizing the continuous optimal control model into a nonlinear programming model, and then nonlinear programming solvers are directly employed to solve the problem. Recently, this approach has been applied to solve the train trajectory (TTO) problem and has shown some advantages and superiority compared to the PMP method [14].

Among the methods that have low computational time, direct methods are included. Direct methods convert the optimal control problem into a mathematical programming problem. Wang et al. in 2013 and Goverde et al. in 2016 redefined the problem as a multiple-

[DOI: 10.22068/ijrare.329]

phase optimal control problem and solved it using pseudo-spectral methods, in which the Legendre pseudo-spectral method was employed to discretize continuous optimal control problems into multiple-phase optimal control problems [15].

In this paper, each phase of the problem is solved independently. Then, each phase is connected through continuity conditions between the independent variables, states, and controls.

The advantages of this method are that if the optimal solution for a desired phase is not obtained for any reason, different methods can be implemented and compared. After discretization by the Legendre pseudo-spectral method, the problem is transformed into a nonlinear programming problem, which has a good structure. Finally, for the implementation and solution of this problem, the GPOPS software, based on MATLAB, has been used [16].

The direct method has a significant advantage over the indirect method due to not requiring an explicit integration process. For this reason, in this article, this method has been employed.

3. Train movement modeling

According to Newton's second law and according to the forces applied to the moving train, equation (1), which is the dynamic equation of train movement, is written as follows:

$$ma = \sum(Forces)$$
$$F_a = F(t) - (R(v) + G(x) + B(t)) = Ma \quad (1)$$

where Fa is the accelerating force, F(t) is the momentary traction force applied to the train, R(v), G(x), and B(t) are the Davis force, the resistance force due to the gradient of the path at location x, and the instantaneous braking force applied to the train. M is the mass of the train and a is the acceleration of the train.

According to Newton's second law mentioned above and also considering that the independent variable of this problem is time and the position and speed of the train are dependent variables, which are represented by x(t) and v(t), respectively, the movement of the train is written as equation (2).

$$\begin{cases} \dot{x}(t) = v(t) \\ \dot{v}(t) = \dfrac{1}{M}(F(t) - B(t) - R(v) - G(x)) \end{cases} \quad (2)$$

And here the simple and basic modeling of the train ends.

3.1. Multiple-phase movement and using the pseudo-spectral method

Since the goal is to solve the optimal control problem of multiple-phase train movement, the train motion model presented in the previous section must be rewritten as a multiple-phase optimal control problem. The advantage of this model is that it includes time, speed, and gradient limits precisely.

In the multiple-phase optimal control model, the train path is divided into phases. Each section of the complete path is called a phase, with each phase of the optimal control problem having its own cost function, dynamic model, path constraints, and specific boundary conditions. The complete path is obtained by connecting the phases through connection conditions.

The total cost function is the sum of the cost functions of each phase. Ultimately, the optimal trajectory can be obtained by minimizing the total cost function along with all its constraints and connection conditions.

Since the speed and gradient constraints along the railway path are changing, the complete train path is divided into several sections based on the points where the gradient and speed change. Therefore, each phase of the train movement has its own specific speed constraint and, as a result, its own specific line resistance (gradient resistance) and train resistance (Davis resistance). In this section, the equations of the multiple-phase optimal control problem are written.

It is assumed that the general problem has P phases and the start and end times, the control variables, and the state of the p-th phase are indicated by $t_0^{(p)}, t_f^{(p)}, u^p$, and $x^{(p)}$, respectively, where $p = 1, 2, ..., P$. In this case, the cost function is given by equation (3).

[DOI: 10.22068/ijrare.329]

$$\min J=\sum_{p=1}^{P}[\phi^{(p)}(x^{(p)}(t_0^{(p)}),t_0^{(p)},x^{(p)}(t_f^{(p)}),t_f^{(p)})+\int_{t_0^{(p)}}^{t_f^{(p)}}\varphi^{(p)}(x^{(p)}(t),u^{(p)}(t),t)\,dt] \quad (3)$$

where the conditions for connecting the phases to each other are in the form of relation (4).

$$t_f^{(p)}=t_0^{(p+1)},\ x^{(p)}(t_f^{(p)})=x^{(p+1)}(t_0^{(p+1)}),\ p=1,...,P-1 \quad (4)$$

The location, speed, traction force, and braking force of the n-th train at time t in phase p are determined by equations (5), respectively.

$$x_n^{(p)}=x_n^{(p)}(t),\ v_n^{(p)}=v_n^{(p)}(t),\ F_n^{(p)}=F_n^{(p)}(t),\ B_n^{(p)}=B_n^{(p)}(t) \quad (5)$$

The cost function of the problem of reducing energy consumption is presented as equation (6).

$$\min\sum_{n\in N}\sum_{p\in P}\int_{t_0^{(p)}}^{t_f^{(p)}}u_p^2(t)\,dt,\ \mathrm{n}=\{1,2,...,N\},\ p=\{1,2,...,P\} \quad (6)$$

Now, to use the pseudo-spectral method, first a new independent variable is defined instead of the independent variable t, and the relationship between the new independent variable and t is given by equation (7) [17-18].

$$t=\frac{t_f^{(p)}-t_0^{(p)}}{2}\tau+\frac{t_f^{(p)}+t_0^{(p)}}{2} \quad (7)$$

Now, the multiple-phase optimal control problem, which was written in terms of the variable t, is rewritten in terms of the variable $\tau=[-1,1]$, and equation (8) is obtained.

$$\min J=\sum_{p=1}^{P}[\phi^{(p)}(-1),t_0^{(p)},x^{(p)}(1),t_f^{(p)})+\frac{t_f^{(p)}-t_0^{(p)}}{2}\int_{-1}^{1}\varphi^{(p)}(x^{(p)}(\tau),u^{(p)}(\tau),\tau;t_0^{(p)},t_f^{(p)})\,d\tau] \quad (8)$$

Finally, the phase connection conditions to each other are written as equation (9).

$$\begin{cases}t_f^{(p)}=t_0^{(p+1)}\\ x_n^{(p)}(t_f^{(p)})=x_n^{(p+1)}(t_0^{(p+1)}),n\in\mathrm{N},\mathrm{p}\in\{1,...,P-1\}\\ v_n^{(p)}(t_f^{(p)})=v_n^{(p+1)}(t_0^{(p+1)})\end{cases} \quad (9)$$

4. Simulations and Results

In this section, energy optimization will be performed for two motion modes, single-phase and three-phase, using the method and equations described in the previous section, and then they will be compared.

4.1. Single-phase motion

This section considers a problem in which two trains move according to a pre-designed timetable, which is provided in Table (1). The first and third nodes are the origin and the destination stations, respectively, and the second node is the intersection point [15].

D and A, respectively, indicate the time of departure and arrival of the train.

Table 1. Timetable for two trains in single-phase movement mode

Train Number	D time from node 1	A time at node 3
1	08:00	08:35
2	08:05	08:40

Since the travel time for both trains is the same, both trains have the same specifications, and their departure times are far enough apart that the two trains do not interact with each other, so the limit of separating the two trains can be omitted. As a result, the same optimal control strategies are considered for both trains.

Therefore, this problem becomes an optimal control problem for a single train with one phase, in which the train travels a route of 80 km in 2100 seconds.

Imagine that the positions of the nodes are in order at $X_1=0$, $X_2=40km$, and $X_3=80km$. It is worth mentioning that in this problem, there is no stop at the intermediate node because

[DOI: 10.22068/ijrare.329]

it is assumed that both trains move according to a predetermined timetable and neither of them suffers any delays. Therefore, there is no need for overtaking at the intermediate node. The maximum speed that a train can have is 180 kilometers per hour, and its mass is assumed to be 600 tons. The traction force of the train is a maximum of 140 kilonewtons, and it can brake with a maximum force of 200 kilonewtons.

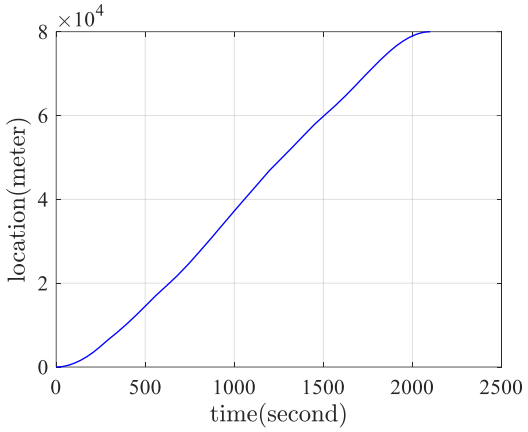

Figure 3. Diagram of train location in single-phase motion mode

By employing the pseudo-spectral method described in the previous section, the location diagram obtained in the single-phase motion of the train is illustrated in Figure 3.

Furthermore, the speed diagram with respect to the train's position is depicted in Figure 4.

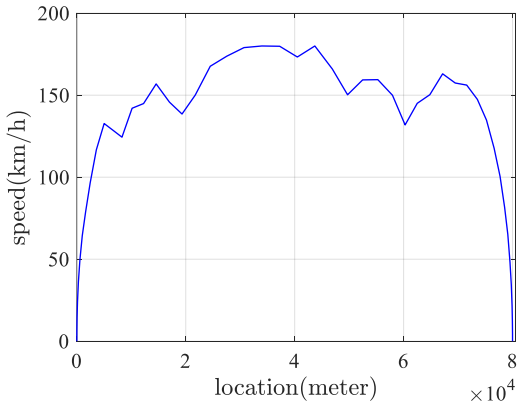

Figure 4. Train speed diagram in single-phase motion mode

Finally, the diagram related to the train input (gas or brake), or, in other words, the train control diagram in the single-phase mode using the pseudo-spectral method, is drawn in Figure 5. The amount of energy consumption in this case is equal to $1.7950\times10^{8}\,J$.

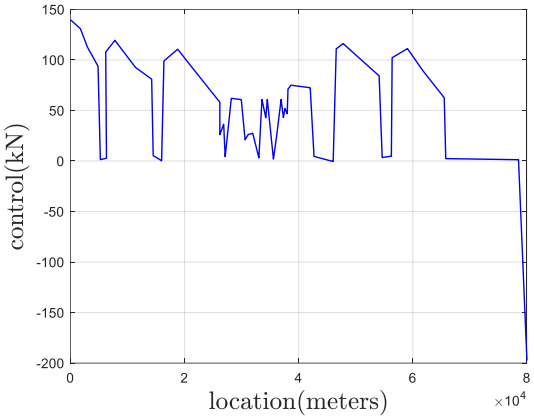

Figure 5. Train control diagram (optimal input) in single-phase motion mode

4.2. Three-phase motion

In this scenario, the motion of a single train with three stations is simulated, where the middle node in the previous section is considered the second station. According to the equation 2N(S-1), where S is the total number of nodes and N is the total number of trains, the number of events in this scenario is four, therefore this problem is a three-phase problem. The first phase involves the train's movement from the origin station to the intermediate station, at a distance of 40 kilometers from the origin station. After reaching the second station, the train stops for 50 seconds and then continues its movement until it reaches the destination station (the third station) after covering a distance of 80 kilometers. Therefore, the optimal control model for this problem, along with its constraints, is as follows [15]:

$$\text{minimize } J = \int_{t=0}^{t=1025} u_1^2(t)dt + \int_{t=1025}^{t=1075} u_2^2(t)dt + \int_{t=0}^{t=2100} u_3^2(t)dt \tag{10}$$

The dynamic equations related to the first and second phases are:

$$\begin{cases} \dot{x}^{(p)} = \dot{v}^{(p)}, \; p=1,3 \\ \dot{v}^{(p)} = \dfrac{1}{M}(F^{(p)}(t) - B^{(p)}(t) - R(v^{(p)}) - G(x^{(p)})), \; \text{p=1,3} \end{cases} \tag{11}$$

[DOI: 10.22068/ijrare.329]

with the following path restrictions:

$$\begin{cases} 0 \le F^{(p)} \le 140 \\ 0 \le B^{(p)} \le 200, \ \text{p=1,3} \\ 0 \le v^{(p)} \le 180 \end{cases} \quad (12)$$

The dynamic equations of the second phase and boundary conditions are:

$$\begin{cases} \dot{x} = 0 \\ \dot{v} = 0 \end{cases}, v_n^{(p)} = 0, F_n^{(p)} = 0, \quad (13)$$
$$B_n^{(p)} = 0$$

The conditions for connecting phases to each other are:

$$\begin{cases} t_f^{(p)} = t_0^{(p+1)} \\ x_n^{(p)}(t_f^{(p)}) = x_n^{(p+1)}(t_0^{(p+1)}), n \in \mathrm{N}, \mathrm{p} \in \{1, \ldots, P-1\} \\ v_n^{(p)}(t_f^{(p)}) = v_n^{(p+1)}(t_0^{(p+1)}) \end{cases} \quad (14)$$

Now, in MATLAB software, the optimal movement strategy is obtained using the pseudo-spectral method.

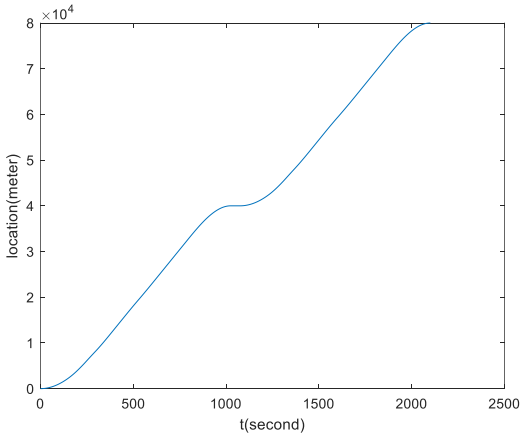

Figure 6. The location time graph

Figure 6 clearly demonstrates the three-phase motion. The first phase is the movement to the second station; the second phase is the stop at the second station; and the third phase is the movement to the final station. The optimal speed profile is obtained as follows:

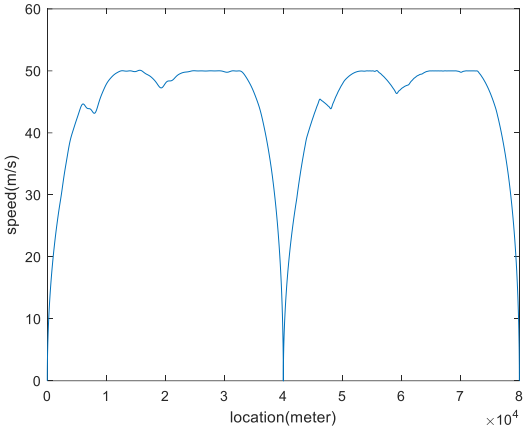

Figure 7. The speed location graph

As indicated, the train has come to a stop at the 40-kilometer mark, which means it has reached the station. The control graph is as follows:

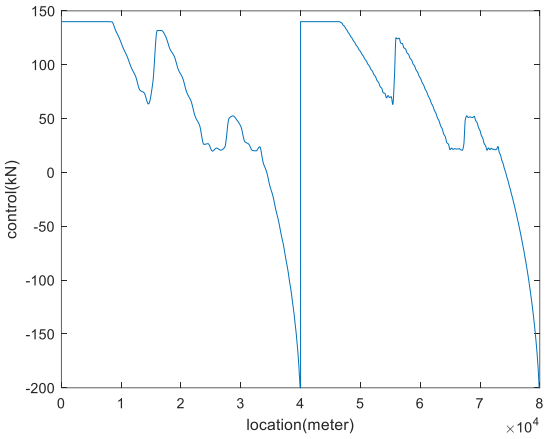

Figure 8. The control location graph

4.3. Optimal movement on the Abardej-Kavir railway line

The Abardej and Kavir stations are located on the Tehran-Mashhad railway line. To achieve optimal train movement between these two stations, real data such as maximum allowable speed, dispatch times, and the gradient of this path have been used. The designed optimal speed and control are shown in the following figures:

[DOI: 10.22068/ijrare.329]

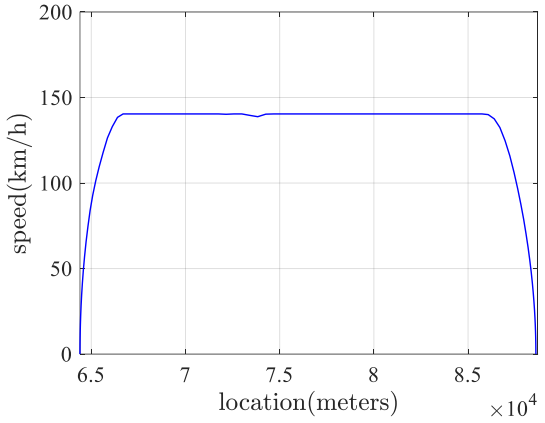

Figure 9. The speed location graph

According to the designed speed profile, it is evident that using the pseudo-spectral method, the train motion consists of four phases: maximum acceleration, constant speed, coasting, and maximum braking. In fact, based on the four optimized motion regimes extracted from the PMP principle, which leads to energy optimization, the train moves. It is worth mentioning that this speed profile has only been designed for a time duration of 32.12 seconds, which is significantly reduced compared to other lengthy and complicated methods.

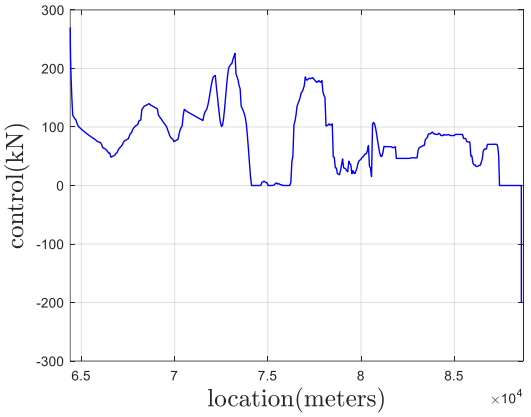

Figure 10. The control location graph

Using the pseudo-spectral method, an energy consumption level of $204.1452\times10^{6}\,J$ has been obtained, which has resulted in a 15% energy saving compared to the optimization performed through previous methods such as reducing the maximum speed.

5. Conclusion

In this article, the optimization method of pseudo-spectral, which is among the direct collocation methods, was used to generate a speed profile with the goal of optimizing energy consumption. Initially, an optimal control model of train motion was obtained, and then the train motion was divided into multiple phases based on path and speed constraints. For each phase, an optimal control model along with its specific constraints was formulated. After writing the equations related to multi-phase optimal control, this problem was discretized using the pseudo-spectral method and transformed into a nonlinear programming (NLP) problem.

The next step was implementing the problem in GPOPS software and solving the discretized problem using NLP solvers. The results obtained from adopting this method indicate that it has a significantly higher convergence speed compared to other optimization methods and also leads to more effective energy savings.

The weakness of this method is its relatively low accuracy compared to other methods. For example, when a part of the path has a negative gradient, the traction force does not decrease sufficiently to prevent the train from accelerating in the coasting phase. However, increasing the number of time nodes and mesh points can largely overcome this weakness.

[DOI: 10.22068/ijrare.329]

References

[1] Z. Tan, S. Lu, K. Bao, S. Zhang, C. Wu, J. Yang, and F. Xue, Adaptive partial train speed trajectory optimization, Vol.11, no.12, (2018), pp.3302.

[2] CL. Darby, AV. Rao, A mesh refinement algorithm for solving optimal control problems using pseudospectral methods, Proceedings of the AIAA (2009).

[3] L. Wei, L. Qunzhan, T. Bing, Energy saving train control for urban railway train with multi-population genetic algorithm, In 2009 International Forum on Information Technology and Applications, Vol. 2, pp. 58-62. IEEE, 2009.

[4] M. Khanbaghi, and R. P. Malhame, Reducing travel energy costs for a subway train via fuzzy logic controls In Proceedings of 1994 9th IEEE International Symposium on Intelligent Control, pp. 99-104. IEEE, 1994.

[5] PG. Howlett, IP. Milroy, PJ. Pudney, Energy-efficient train control, Vol.2, no. 2 (1994), pp.193-200.

[6] H. Ye, R. Liu, A multiphase optimal control method for multi-train control and scheduling on railway lines, Vol. 93 (2016), pp.377-393.

[7] P. Wang, RMP. Goverde, Multi-train trajectory optimization for energy-efficient timetabling, Vol.272, no. 2 (2019), pp.621-635.

[8] R. Bulirsch, E. Nerz, HJ. Pesch, O. Stryk, Combining direct and indirect methods in optimal control, Range maximization of a hang glider in Optimal control, (1993), pp. 273-288, 1993.

[9] K. Bao, S. Lu, F. Xue, Z. Tan, Optimization for train speed trajectory based on Pontryagin's Maximum Principle, In 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), (2017), pp. 1-6.

[10] IP. Milroy, Aspects of automatic train control, PhD diss., Loughborough University (1980).

[11] E. Khmelnitsky, On an optimal control problem of train operation, Vol.45, no. 7 (2000), pp.1257-1266.

[12] RR. Liu, IM. Golovitcher, Energy-efficient operation of rail vehicles, Vol.37, no. 10 (2003), pp. 917-932.

[13] A. Albrecht, P. Howlett, P. Pudney, X. Vu, P. Zhou, The key principles of optimal train control—Part 2: Existence of an optimal strategy, the local energy minimization principle, uniqueness, computational techniques, Vol.94 (2016), pp.509-538.

[14] H. Ye, R. Liu, Nonlinear programming methods based on closed-form expressions for optimal train control, Vol.82, (2017), pp.102-123.

[15] A. Rao, Extension of a pseudospectral legendre method to non-sequential multiple-phase optimal control problems, In AIAA Guidance, Navigation, and Control Conference and Exhibit, pp. 5634. (2003).

[16] CL. Darby, WW. Hager, AV. Rao, Direct trajectory optimization using a variable low-order adaptive pseudospectral method, Vol.48, no. 3 (2011), pp.433-445.

[17] D. Garg, MA. Patterson, C. Francolin, CL. Darby, GT. Huntington, WW. Hager, AV. Rao, Direct trajectory optimization and costate estimation of finite-horizon and infinite-horizon optimal control problems using a Radau pseudospectral method, Vol. 49, no. 2 (2011), pp. 335-358.

[18] MA. Patterson, AV. Rao, GPOPS-II: A MATLAB software for solving multiple-phase optimal control problems using hp-adaptive Gaussian quadrature collocation methods and sparse nonlinear programming, Vol.41, no. 1 (2014), pp.1-37.

[DOI: 10.22068/ijrare.329]