Research Article

Discrete Train Speed Profile Optimization for Urban Rail Transit: A Data-Driven Model and Integrated Algorithms Based on Machine Learning

Kang Huang1,2,3, Jianjun Wu1,2, Xin Yang1, Ziyou Gao1,2, Feng Liu4, and Yuting Zhu3,5

1 State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China
2 Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Ministry of Transport, Beijing Jiaotong University, Beijing 100044, China
3 School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China
4 Transportation Research Institute (IMOB), Hasselt University, Wetenschapspark 5, Bus 6, 3590 Diepenbeek, Belgium
5 Beijing Transport Institute, Beijing 100073, China

Correspondence should be addressed to Jianjun Wu; jjwu1@bjtu.edu.cn

Received 19 October 2018; Accepted 26 March 2019; Published 2 May 2019

Academic Editor: Hocine Imine

Copyright © 2019 Kang Huang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Energy-efficient train speed profile optimization problem in urban rail transit systems has attracted much attention in recent years because of the requirement of reducing operation cost and protecting the environment. Traditional methods on this problem mainly focused on formulating kinematical equations to derive the speed profile and calculate the energy consumption, which caused the possible errors due to some assumptions used in the empirical equations. To fill this gap, according to the actual speed and energy data collected from the real-world urban rail system, this paper proposes a data-driven model and integrated heuristic algorithm based on machine learning to determine the optimal speed profile with minimum energy consumption. Firstly, a data-driven optimization model (DDOM) is proposed to describe the relationship between energy consumption and discrete speed profile processed from actual data. Then, two typical machine learning algorithms, random forest regression (RFR) algorithm and support vector machine regression (SVR) algorithm, are used to identify the importance degree of velocity in the different positions of profile and calculate the traction energy consumption. Results show that the calculation average error is less than 0.1 kwh, and the energy consumption can be reduced by about 2.84% in a case study of Beijing Changping Line.

1. Introduction

In recent years, urban rail transit has developed rapidly around the world due to its high capacity, safety, superior energy performance, and reliable service with sufficient punctuality [1], which is becoming increasingly important for large cities development [2]. For example, 35 cities in China have urban rail transit with total length over 4750 km in 2017 [3]. According to the Web of China Rail Transit, there will be more than 50 cities operating urban rail transit in the next few years. In 2020, the total mileage of urban rail transit in China will be 6000 km, making the rail systems an important component of urban public transportation. Around the world, more and more cities are traveling oriented to public transportation. As shown in Figure 1 (which is from Global Cities Public Transit Usage Report of moovit), urban rail transit system has attracted much attention in recent years especially in some large cities and accounts for a high proportion of public transportation. However, the quick expansion of urban rail transit networks led to the problem of larger energy consumption. Taking Beijing rail transit as an example, in 2011, the total electric consumption of Beijing urban rail transit was 750 million kwh, and 470 million kwh was used for traction energy consumption, with the proportion as high as 55% which has attracted tremendous attention in recent years (Yin et al. [4]). In 2015, it reached 1.4 billion kwh, accounting for 40% of the total operating cost of the metro[5], which was equivalent to the annual electricity consumption of 730,000 households (annual electricity consumption of one household is based on 2016 BEIJING STATISTICAL
Learning is stable. For instance, the RFR and the SVR have been well applied in many fields, such as biology, medicine, economy, management and so on [8]. Therefore, energy saving has become an important issue in real train operating in order to reduce the operation cost and satisfy the requirement of environment protection.

To reduce the energy consumption in urban rail transit, a lot of models have been developed in recent years which mainly considered the train controlling between two stations based on the kinematic equations. There are three types in general, i.e., mathematical optimization models, simulation methods, and multiple linear regression, and neural network model based on the data. Although a lot of works had been done in optimizing speed profiles, existing methods have some limitations: (1) the mathematical optimization model in theoretical aspects has been sounded. However, the actual situation is often more complex, and the theory of optimization may not get a good performance when the actual facts are taken into consideration. (2) The establishment of the simulation model (e.g., agent-based simulation [7]) is complicated and costly. Further, there is a certain deviation between the simulation results and the actual measurement data. (3) The traction energy consumption and its influence factors are not linear, and the precision of the multiple linear regression model is limited. The neural network relies too much on the empirical information extracted from historical data. The phenomenon of overfitting is prone to occur, and the generalization ability may be hard to guarantee. Besides, it is easy to fall into the local optimum. In contrast, from view of the data-driven optimization on the basis of machine learning theories, the limitations could be avoided. Firstly, real-world data that contains the influences from actual factors can be utilized well. Secondly, machine learning has been well applied in many fields, which provides a method to study the existing information from data, acquire new information, and improve performance of data set. The process that utilizes input data (real-world profile) to obtain output data (energy consumption) is easier to be realized. Thirdly, machine learning is stable. For instance, the RFR and the SVR have stable performance in the data set, and they have been widely used in many fields, such as biology, medicine, economy, management and so on [8]. Therefore, it becomes possible to optimize the train speed profile in the urban rail transit system on the premise of verifying their effectiveness.

Main contributions of this research can be summarized as follows:

(1) A data-driven optimization model (DDOM) is proposed to optimize the speed profile in urban rail transit system. The traditional speed profile optimization model is easy to be analyzed in the theoretical aspects. In this paper, the train speed profile is optimized based on the view of discrete profile which can be applied in the practice easily.

(2) Based on actual data obtained by experimental measurements, a novel method of utilizing the machine learning algorithm to calculate the energy consumption of speed profile is proposed which can avoid considering longitudinal train dynamics. Besides, the calculation error of machine learning algorithm (RFR and SVR) on speed profile energy is verified.

(3) To solve the proposed model, an integrated heuristic optimization algorithm based on RFR and SVR is developed. In addition, comparison of real data, results show average 2.84% energy reduction.

The framework of this paper is shown in Figure 2.

2. Literature Review

During last years, many studies have focused on the energy-efficiency analysis of train traction; Scheepmaker et al. [23]...
Table 1: Some typical publications about energy-efficient.

<table>
<thead>
<tr>
<th>View</th>
<th>Publication</th>
<th>Years</th>
<th>Model type</th>
<th>Objective</th>
<th>Energy consumption calculation method</th>
<th>Solution method</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Howlett et al. [10]</td>
<td>1996</td>
<td>Continuous control model</td>
<td>Energy consumption of profile</td>
<td>Empirical-formula, numerical integration</td>
<td>Optimize control</td>
</tr>
<tr>
<td>I</td>
<td>Albrecht et al. [12]</td>
<td>2013</td>
<td>Continuous control model</td>
<td>Energy consumption of profile</td>
<td>Empirical-formula, numerical integration</td>
<td>Optimize control</td>
</tr>
<tr>
<td>I</td>
<td>Albrecht, Howlett et al. [13, 14]</td>
<td>2016 a&amp;b</td>
<td>Continuous control model</td>
<td>Energy consumption of profile</td>
<td>Empirical-formula, numerical integration</td>
<td>Optimize control</td>
</tr>
<tr>
<td>I &amp; II</td>
<td>Nasri et al. [16]</td>
<td>2010</td>
<td>Simulation model</td>
<td>Energy consumption of timetable</td>
<td>Empirical-formula, numerical integration, Simulation platform</td>
<td>Simulation</td>
</tr>
<tr>
<td>II</td>
<td>Sun et al. [17]</td>
<td>2013</td>
<td>MILP</td>
<td>Energy consumption of timetable</td>
<td>Empirical-formula, numerical integration</td>
<td>Genetic search</td>
</tr>
<tr>
<td>II</td>
<td>Yang et al. [18]</td>
<td>2015a</td>
<td>MILP</td>
<td>Energy consumption of whole line</td>
<td>Empirical-formula, numerical integration</td>
<td>Genetic search</td>
</tr>
<tr>
<td>II</td>
<td>Yin et al. [22]</td>
<td>2017</td>
<td>MILP</td>
<td>Energy consumption and the passenger waiting time</td>
<td>Empirical-formula, numerical integration</td>
<td></td>
</tr>
</tbody>
</table>

I: speed profiles/driving strategy; II: energy-efficient timetable.

summarized and gave a review from two aspects, (1) optimizing the speed profiles and driving strategies to reduce the energy consumption (e.g., Howlett [24, 25]; Albrecht et al. [12]; Scheepmaker and Goverde [26]; Yang et al. [18, 27]; Tian et al. [28]; Sun et al. [17]; Yang et al. [29]) and (2) optimizing the timetable by means of utilization of regenerative energy with minimum energy consumption (e.g., Chevrier et al. [30]; Li and Lo [19, 20]; Wang and Goverde [31]; Wang et al. [32]; Zhao et al. [33]). Some typical publications about energy-efficient research are listed in Table 1. In essence, energy consumption is related to the train traction process. It is a fundamental work to improve the speed profiles. Over the past 25 years, the challenges in the train speed profile optimization have resulted in a variety of analysis frameworks. (1) Mathematical optimization models. The modern theory of optimal train control was developed during the years 1992-2014 by the Scheduling and Control Group (SCG) at the University of South Australia in a collection of papers. For example, Howlett and Cheng [9] built a discrete control model and confirmed the fundamental optimality of the accelerate-coast-brake strategy for energy-efficient train operation. On the basis of the Pontryagin maximum principle, if no energy is recovered during braking, then it becomes an optimal switching strategy. Wong and Ho [11] showed that a genetic algorithm was more robust in calculational processes. After reformulating the necessary conditions for optimal switching, Howlett et al. [34] proposed a less general model that the optimal switching points for each steep section can be found by minimizing an intrinsic local energy function. Albrecht et al. [13] used the Pontryagin principle to find necessary conditions on an optimal strategy and showed that a strategy of optimal type uses only a limited set of optimal control modes, Maximum Power, HoldP (Hold using Power), Coast, HoldR (Hold using Regenerative braking), and Maximum Brake. Albrecht et al. [14] developed general bounds on the position of optimal switching points and proved that an optimal strategy always exists. And an intrinsic local energy minimization principle for determination of optimal switching points was established, which shows that the optimal strategy is unique. Huang et al. [35] proposed an integrated approach for the energy-efficient driving strategy and timetable which was solved by a particle swarm optimization (PSO) algorithm. Yang et al. [36] employed an energy-efficient through the Taylor approximation. They
transformed the train scheduling problem using a nonconvex formation into a quadratic formation and search the solution by a PSO method. (2) Simulation method. Yin et al. [15] built an ITO (intelligent train operation) simulation platform on the basis of the multiple-point-mass train model that the platform consists of four parts, i.e., the Input Module, the Algorithm Module, the Train Module, and the Output Module. (3) Multiple linear regression model and neural network model based on the data. Fernández et al. [37] modeled electric trains energy consumption using neural networks, providing a reliable estimation of the consumption along a specific route when being fed with input data such as train speed, acceleration, or track longitudinal slope.

Big data analytics (BDA) has increasingly attracted a strong attention of analysts, researchers and practitioners in railway transportation and engineering filed [38]. From a data-driven view, this paper mainly focuses on how to obtain the optimal speed profile based on well-developed machine learning algorithms. There are still seldom researches aiming at optimal speed profile by this proposed method.

3. Data Analysis and Preprocessing

3.1. Data Overview. During the operation of the subway, the most widely used power is electricity. Some are used for the consumption of facilities in the train, such as air conditioning, lighting, etc. The rest is for traction of metro trains. Our data resources are formed by urban rail transit train running state and corresponding energy consumption, which are derived from Changping Line of Beijing urban rail transit. The operation section of Changping Line is from the Xi’erqi station to the Changpingxishankou station, with operating mileage of 31.9 kilometers and total of 12 stations opened (as illustrated in Figure 3). In order to accurately capture the actual traction power consumption during the operation of the subway, we installed sensors and computers on the train. The total energy consumption and the energy consumptions of various electrical appliances in the train are both recorded. Then, the total consumption is subtracted from the electrical energy consumed by the electrical appliances, and the rest is the energy consumed by the traction of the subway train. The provided data covers running stage of 4 months. There are two circle running tests every night in the up and down direction. The types of recorded data are showed in Table 2.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Unit</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed</td>
<td>km/h</td>
<td>0.001</td>
</tr>
<tr>
<td>Position</td>
<td>m</td>
<td>0.001</td>
</tr>
<tr>
<td>Time</td>
<td>s</td>
<td>0.2</td>
</tr>
<tr>
<td>Train weight</td>
<td>ton</td>
<td>1</td>
</tr>
<tr>
<td>Current slope</td>
<td>%</td>
<td>1</td>
</tr>
<tr>
<td>EBI speed</td>
<td>km/h</td>
<td>0.001</td>
</tr>
<tr>
<td>Station spacing</td>
<td>m</td>
<td>0.001</td>
</tr>
<tr>
<td>Expected acceleration of PID</td>
<td>(km/h)/s</td>
<td>1</td>
</tr>
<tr>
<td>Electric energy consumption</td>
<td>Kwh</td>
<td>1</td>
</tr>
</tbody>
</table>

3.2. Data Preprocessing

Symbols

\( n \): number of section is discretized to.
\( v^0_j \): \( j \)th speed point of original profile \( C_j \).

Figure 3: Illustration of the Changping Line.
three speed limiting sections: section of the urban rail transit. In this section, there are limits on speed at different locations in each time. Normally, differences in track conditions are caused by construction and geological reasons. In the high-speed zone; the third stage is the deceleration process of the urban rail transit train. Taking MingTombs-Changpingxishankou of the down direction, for instance, as shown in Figure 4, a uniform interval of 50 m and 5 m is selected for discrete process. In Figure 5(a), the speed profile record number drops to 26, getting 26 control points during the train traction, respectively, in Figure 5(b), speed profile record number is 247, and the density of control points is higher.

Using these recorded data, we can draw out the running process of the urban rail transit train. Taking MingTombs-Changpingxishankou of the down direction, for instance (shown in Figure 4), the train operation process is divided into three stages. The first stage is accelerating until approaching the maximum speed limit; the second stage is fluctuating in the high-speed zone; the third stage is the deceleration braking until the train stops. Normally, differences in track conditions are caused by construction and geological reasons. There will be limited speed at different locations in each section of the urban rail transit. In this section, there are three speed limiting sections: \( \rightarrow S_{lim1}, S_{lim1} \rightarrow S_{lim2}, S_{lim2} \rightarrow \text{End} \). Each part has its maximum speed limit.

A speed profile has three elements, speed, time, and distance. The time interval between records in the table is 0.2 seconds. However, the running time between two stations varies from almost one to several hundred seconds. This means that a speed profile may be made up of thousands of records. We need to calculate the energy consumption from the profile, that is to say, to find the relationship between energy consumption and the thousands of data records, which is the so-called “high-dimensional” data in statistics.

Although machine learning algorithms under the back of big data are suitable for dealing with high-dimensional data, for extremely high-dimensional situations, large amounts of data are needed as training sets, and calculation precision is hard to be gained [39]. Therefore, we choose dimensionality reduction for the limitation of data quantity. Not only can the algorithm achieve good training effect, but also the accuracy of the original high-dimensional data can be reserved.

Process of reducing the dimension is as follows: (1) The section length \( S_0 \) can be obtained from records, then \( S_0 \) is divided into \( n \) small sections (the uniform segmentation method is chosen in this paper). Thus, the \((n+1)\) points are represented by \( \{s_0, s_1, \ldots, s_n | i = 0, 1, \ldots, n\} \). Clearly, \( s_0 = 0, s_n = S_0 \) (section total length). Taking MingTombs-Changpingxishankou of the down direction, for instance, as shown in Figure 5, a uniform interval of 50 m and 5 m is selected for discrete process. In Figure 5(a), the speed profile record number drops to 26, getting 26 control points during the train traction, respectively, in Figure 5(b), speed profile record number is 247, and the density of control points is higher.

(2) Find the latter and previous positions of \( s_i \) in original profile within \( Vt \) interval, recorded as \( s_i^- \) and \( s_i^+ \). Sequence \( \{s_0^- \ldots s_i^- \ldots s_n^-, s_0^+ \ldots s_i^+ \ldots s_n^+\} \) is obtained.

(3) In the original velocity profile, we can get the velocity and time corresponding to the \( s_i^- \) and \( s_i^+ \), recorded as \( V', V_i^- \), \( t_i^- \) and \( t_i^+ \). In the small section from \( s_i^- \) to \( s_i^+ \), the train is assumed to be in a uniformly accelerated state. As shown in Figure 6, by using \( v_i^-, v_i^+, s_i^- \) and \( s_i^+ \), the \( v_i \) can be obtained. Therefore, we can get the \( \{v_0 \ldots v_i \ldots v_n\} \), where \( v_0 = v_n = 0 \). Figure 6(a) indicates speed profile can be represented by fewer points. Figure 6(b) shows error between the simplified profile and original one could be ignored when compared the whole length of section.

### Table 3: Part types of the original data.

<table>
<thead>
<tr>
<th>Time</th>
<th>Velocity(km/h)</th>
<th>Distance(m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_1^0 )</td>
<td>( v_1 )</td>
<td>( s_1^0 )</td>
</tr>
<tr>
<td>( t_2^0 )</td>
<td>( v_2 )</td>
<td>( s_2^0 )</td>
</tr>
<tr>
<td>( \ldots )</td>
<td>( \ldots )</td>
<td>( \ldots )</td>
</tr>
<tr>
<td>( t_i^0 )</td>
<td>( v_i )</td>
<td>( s_i^0 )</td>
</tr>
<tr>
<td>( \ldots )</td>
<td>( \ldots )</td>
<td>( \ldots )</td>
</tr>
<tr>
<td>( t_m^0 )</td>
<td>( v_m )</td>
<td>( s_m^0 )</td>
</tr>
</tbody>
</table>

"Table 3: Part types of the original data."
a uniform interval of 5 m, and part of results are shown in Table 4.

The speed profile sequence \(\{v_i - s_i\}, i = 1, 2, \ldots, n\) and the traction energy consumptions of each sequence \(E\) are extracted. And the data is shown in Table 5 (q: number of processed data records). Then, to eliminate dimension, the data is normalized. The extracted data is divided into two parts. 80% is as the training set, and 20% is as the test set.

4. Formulation

In this section, a data-driven optimization model (DDOM) is proposed to optimize the urban rail transit traction energy consumption, which discretizes velocity profile and describes the relation between velocity profile and energy consumption as a complex mapping-relation.

### 4.1. Symbols and Assumptions

**Parameters**

- \(V_i^0\): velocity set at a time interval \(\forall t, V_i^0 = \{v_j^0 \mid j = 1, 2 \ldots n\}\).
- \(S_i^0\): distance set at a time interval \(\forall t, S_i^0 = \{s_j^0 \mid j = 1, 2 \ldots n\}\).
- \(T_i^0\): time set with a time interval of \(\forall t, T_i^0 = \{t_j^0 \mid j = 1, 2 \ldots n\}\).
- \(C^*\): set of processed speed profiles, and \(\{v_i - s_i\} \mid i = 0, 1 \ldots n\} \in C^*\).
- \(a_i\): the acceleration at \(s_i\).
- \(E_i\): energy consumption of urban rail transit traction under running time of \(t\).
Table 4: Part of the velocity series after being processed.

<table>
<thead>
<tr>
<th>i</th>
<th>$v_i' (\text{km/h})$</th>
<th>$s_i' (\text{m})$</th>
<th>$t_i' (\text{Vt})$</th>
<th>$v_i (\text{km/h})$</th>
<th>$s_i (\text{m})$</th>
<th>$v_i' (\text{km/h})$</th>
<th>$s_i' (\text{m})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>9.18</td>
<td>4.64</td>
<td>21</td>
<td>9.657</td>
<td>5</td>
<td>9.9</td>
<td>5.19</td>
</tr>
<tr>
<td>4</td>
<td>18.612</td>
<td>14.928</td>
<td>34</td>
<td>18.659</td>
<td>15</td>
<td>19.296</td>
<td>16</td>
</tr>
<tr>
<td>6</td>
<td>24.408</td>
<td>24.646</td>
<td>42</td>
<td>24.593</td>
<td>25</td>
<td>25.128</td>
<td>26.042</td>
</tr>
<tr>
<td>7</td>
<td>26.568</td>
<td>28.954</td>
<td>45</td>
<td>27.042</td>
<td>30</td>
<td>27.252</td>
<td>30.468</td>
</tr>
<tr>
<td>8</td>
<td>28.764</td>
<td>33.622</td>
<td>48</td>
<td>29.371</td>
<td>35</td>
<td>29.484</td>
<td>35.26</td>
</tr>
<tr>
<td>9</td>
<td>30.888</td>
<td>38.652</td>
<td>51</td>
<td>31.442</td>
<td>40</td>
<td>31.608</td>
<td>40.408</td>
</tr>
<tr>
<td>10</td>
<td>33.012</td>
<td>44.04</td>
<td>54</td>
<td>33.419</td>
<td>45</td>
<td>33.804</td>
<td>45.918</td>
</tr>
<tr>
<td>11</td>
<td>35.172</td>
<td>49.786</td>
<td>57</td>
<td>35.25</td>
<td>50</td>
<td>35.892</td>
<td>51.78</td>
</tr>
<tr>
<td>12</td>
<td>36.576</td>
<td>53.812</td>
<td>59</td>
<td>36.991</td>
<td>55</td>
<td>37.296</td>
<td>55.884</td>
</tr>
<tr>
<td>13</td>
<td>37.942</td>
<td>57.992</td>
<td>61</td>
<td>38.617</td>
<td>60</td>
<td>38.664</td>
<td>60.14</td>
</tr>
<tr>
<td>14</td>
<td>40.068</td>
<td>64.554</td>
<td>64</td>
<td>40.197</td>
<td>65</td>
<td>40.716</td>
<td>66.816</td>
</tr>
<tr>
<td>15</td>
<td>41.436</td>
<td>69.118</td>
<td>66</td>
<td>41.709</td>
<td>70</td>
<td>42.156</td>
<td>71.46</td>
</tr>
<tr>
<td>16</td>
<td>42.804</td>
<td>73.838</td>
<td>68</td>
<td>43.134</td>
<td>75</td>
<td>43.488</td>
<td>76.254</td>
</tr>
<tr>
<td>17</td>
<td>44.172</td>
<td>78.708</td>
<td>70</td>
<td>44.528</td>
<td>80</td>
<td>44.856</td>
<td>81.2</td>
</tr>
<tr>
<td>18</td>
<td>45.576</td>
<td>83.732</td>
<td>72</td>
<td>45.897</td>
<td>85</td>
<td>46.224</td>
<td>86.3</td>
</tr>
<tr>
<td>19</td>
<td>46.944</td>
<td>88.908</td>
<td>74</td>
<td>47.213</td>
<td>90</td>
<td>47.592</td>
<td>91.352</td>
</tr>
<tr>
<td>20</td>
<td>48.204</td>
<td>94.23</td>
<td>76</td>
<td>48.368</td>
<td>95</td>
<td>48.78</td>
<td>96.94</td>
</tr>
</tbody>
</table>

Table 5: Data format of training and testing set.

<table>
<thead>
<tr>
<th>Serial number</th>
<th>$s_0$</th>
<th>$s_1$</th>
<th>...</th>
<th>$s_{n-1}$</th>
<th>$s_n$</th>
<th>Time</th>
<th>Energy consumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$v_0'$</td>
<td>$v_1'$</td>
<td>...</td>
<td>$v_{n-1}'$</td>
<td>$v_n'$</td>
<td>$t_1'$</td>
<td>$E_1$</td>
</tr>
<tr>
<td>2</td>
<td>$v_0'$</td>
<td>$v_1'$</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>$t_2'$</td>
<td>$E_2$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>q-1</td>
<td>$v_{q-1}'$</td>
<td>$v_q'$</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>$t_q'$</td>
<td>$E_{q-1}$</td>
</tr>
<tr>
<td>q</td>
<td>$v_q'$</td>
<td>$v_q'$</td>
<td>...</td>
<td>$v_{n-1}'$</td>
<td>$v_n'$</td>
<td>$t_n'$</td>
<td>$E_q$</td>
</tr>
</tbody>
</table>

$V^\text{min}_{Si}$: minimum speed limit corresponding to $s_i$.
$V^\text{max}_{Si}$: maximum speed limit corresponding to $s_i$.
$a^\text{min}$: minimum acceleration limit in operational section.
$a^\text{max}$: maximum acceleration limit in operational section.
$T^\text{min}$: minimum time limit in operational section.
$T^\text{max}$: maximum time limit in operational section.

Assumption. During the process of $s_i' \rightarrow s_i \rightarrow s_i^{+}$, because the interval is small enough, it is assumed that the train is in uniform acceleration. According to the theorem of velocity–displacement relationship in physics, the quadratic function can be given.

\[
\frac{(v_i'^2 - v_i^2)}{(s_i^{+} - s_i)} = 2a_i \quad i = 1 \ldots n
\]  
\[
a_i = \frac{(v_i'^2 - v_i^2)}{Vt} \quad i = 1 \ldots n
\]

Derived by formulas (1)-(3), we get the velocity sequence $\{v_0 \ldots v_n\}$ as follows:

\[
v_i = \sqrt{2a_i (s_i - s_i^{'-}) + (v_i^{'-})^2} 
\]

or

\[
v_i = \sqrt{(v_i'^2 - 2a_i (s_i^-' - s_i))} 
\]
4.2. Train Operation Constraints. During the running state from one station to a neighboring station, some constraints should be satisfied.

Speed limit (SL) constraints: the speed limit of the section at $s_i$ should be satisfied.

\[
v_{\text{min}}^i < v_i < v_{\text{max}}^i \quad i = 1 \ldots n \tag{6}
\]

$v_{\text{min}}^i$ and $v_{\text{max}}^i$ are determined by the actual speed limit of the section.

Acceleration constraints: in order to satisfy the comfort of passengers on the train, the acceleration needs to be kept in a suitable range. As shown in formula (7)-(8), $a_{\text{min}}$ and $a_{\text{max}}$ are determined by actual empirical parameters, and $a_{\text{max}} > 0, a_{\text{min}} < 0$.

\[
\left( \frac{(v_{i+1}^2 - v_i^2)}{2(s_{i+1} - s_i)} \right) = a_i \in [a_{\text{min}}, a_{\text{max}}] \quad i = 1 \ldots (n-1) \tag{7}
\]
\[
\left( \frac{(v_i^2 - (v_{i-1})^2)}{2(s_i - s_{i-1})} \right) = a_i \in [a_{\text{min}}, a_{\text{max}}] \quad i = 1 \ldots n \tag{8}
\]

Train operation time constraints: transportation efficiency also should be taken into account. Therefore, the train running time $t$ also needs to be within a certain range as shown in formula (9).

\[
t \in [T_{\text{min}}, T_{\text{max}}] \tag{9}
\]

where $T_{\text{min}}$ and $T_{\text{max}}$ are determined by the service level and operational condition.

Train operation distance constraints: to ensure that the train can reach the station accurately, the total displacement of the train in the section must be equal to the length of the section.

\[
s_n = S_0 \tag{10}
\]

4.3. Objective Function. When the section running time of train is $t$, the corresponding energy consumption is $E_t$, which has a complicated relationship with the sequence of velocity points. That is, $E_t(s_0-v_0) \ldots (s_i-v_i) \ldots (s_n-v_n) i=0,1 \ldots n$. The optimization of urban rail transit speed profile is to minimize the energy consumption under the condition of satisfying transportation task, and the objective function of data-driven optimization model (DDOM) is showed in (11).

\[
\min \ E = \min \ t \ E_i(\{v_i - s_i\} \mid i = 0, 1 \ldots n) \tag{11}
\]

\[
t \in [T_{\text{min}}, T_{\text{max}}], (\{v_i - s_i\} \mid i = 0, 1 \ldots n) \in C^*
\]

5. A Greedily Heuristic Algorithm for Model

In this section, firstly two energy consumption calculation methods based on machine learning algorithm are introduced. Then, by analysis the characters of them, an integrated optimization flow is developed with a combination of their merits.

5.1. Energy Consumption Calculation Based on Machine Learning Algorithm. From the view of data-driven method, urban rail transit train runs within each section and produces a traction speed profile that corresponds to an energy consumption value. Although the factors affecting the energy consumption of each train are not only related to the speed profile, the external factors are determined once the operational section is fixed. Moreover, the transmission characteristic of the train is determined when the type of train is selected; then the energy consumption is only related to the speed profile during the traction process. Therefore, the speed profile becomes the key to the energy consumption of train traction.

In this paper, two typical machine learning algorithms (RFR and SVR) are introduced, where RFR is utilized to get velocity points’ importance degrees in different positions, which can be responsible for obtaining these pairs space-speed with a major contribution to the energy consumption. And, SVR is employed to calculate the energy consumption of the profile. The programming environment is Python 3 and its machine learning module is scikit-learn.

5.1.1. Random Forest Regression (RFR) Algorithm Module. Random forest is a kind of ensemble learning algorithm, which uses multiple trees to train and predict a classifier, and also can be used for regression [40]. Based on decision trees combined with aggregation and bootstrap ideas, random forests were introduced by Breiman in 2001, which added an additional layer of randomness to bagging. In addition to constructing each tree using a different bootstrap sample of the data, random forests change how the classification or regression trees are constructed. They are a powerful nonparametric statistical method allowing consideration in a single and versatile framework regression problem [41]. The random forest optionally produces two additional pieces of information: a measure of the importance of the predictor variables and a measure of the internal structure of the data (the proximity of different data points between one and another). In this paper, we can take advantages of this module to get velocity points’ importance degree in different positions which can be used in heuristic solution process for model.

Evaluation and Analysis of RFR. In the utilization of RFR algorithm, two important parameters should be calibrated: the number of split attributes (Mtry) and number of decision trees (Ntree). For simplicity, the enumeration method is used to traverse the two parameters. The convergence process is shown in Figure 7 over ten experiments. We can see that, when Ntree≥50, the average error is close to 0.1kwh. For different Mtrys, errors are shown in Figure 8(a), and there is an acceptable convergence range in Figure 8(b). When the...
calculate the energy consumption based on real data. In this paper, it is applied to vector machine is an effective tool in dealing with small nonlinear, local issues. Support vector regression (SVR) is a solid theoretical foundation and its complete theoretical derivation, support vector machine (SVM) algorithm is from statistical learning theory (SLT), which is based on the structural risk minimization principle that can avoid excessive statistical learning theory (SLT), which is based on the structural risk minimization principle that can avoid excessive

can reach about 0.1kwh. The best combination of parameters is \( C \geq 30 \), and \( \text{Gamma} = 3 \).

(2) For LINEAR, calibration parameter is \( C \) penalty factor. As shown in Figure 10(b), the convergence is slow. When \( C \geq 900 \), the average error of traction energy consumption also can reach about 0.1kwh, which means that it will take a little longer time to reach minimum errors.

(3) For POLY, calibration parameter is \( C \) penalty factor. As shown in Figure 10(c), average error is fluctuating up-down at 0.1kwh and not stable, which fails to achieve better convergence results.

Comparing the performance of the three kernel functions, average error of the RBF kernel function is the best, which means that the traction energy consumption can be calculated under the optimal parameter conditions.

5.1.3. Analysis of the Two Machine Learning Algorithms. For RFR algorithm, stable performance is in the data set, and the evaluation results are satisfactory. At the same time, the more momentous point is that the importance degrees of the velocity points in different positions can be sorted, which will be a valid guiding to the optimization control of the speed profile. For example, we can adjust the speed with high importance degree in the speed profile optimization process. As for the SVR algorithm, although the performance is not good in some kernel conditions, the ability to calculate in the RBF kernel function is also serviceable enough. For optimizing the speed profile of an urban rail transit train, we should find a speed profile that is not less than the existing energy consumption or is even lower than the existing energy consumption. However, the RFR algorithm has a fatal flaw: random forest cannot make the output beyond the range of data set, which may lead to overfitting in modeling of some specific data with noise. Therefore, the design of urban rail transit speed profile optimization algorithms could be beneficial to the combination virtues of the SVR and RFR.

5.2. Optimization Process. Form the view of discrete train speed profile optimization, the key problem is how to design a method to get a more energy-efficient profile; thus a group of combinations \( \{v_i - s_j; (i = 0, 1 \ldots n) \} \) should be found. Velocity \( v_i \) in every position can be in a range, and the number of \( \{v_i - s_j; (i = 0, 1 \ldots n) \} \) combinations will be beyond imagination. It is necessary to discretize the speed changing value. Thus, there should be a step size used for the speed adjustment. A simple and effective step size is the unit from recording instrument (in our experiment, it is 0.001km/h).

Further, a heuristic process can be proposed to reduce the combinations: we can utilize important degree from RFR to adjust the velocity with fixed order. Then, energy-saving profile will be easier to get by the heuristic process. As shown in Figure 11, in one operation section, of the real-world data, there are many profiles under the same running time but with different energy consumptions. Under every running time condition, we can try to find a satisfactory profile at this fixed running time. Then, the best of them with different fixed running time is taken as the optimal solution. Based
on this, we develop an integrated greedily heuristic algorithm combined with RFR and SVR.

**Parameters**

\( I^+ \): set of index values corresponding to the speed at which the importance degree is arranged in descending order.

\( I^- \): set of index values corresponding to the speed at which the importance degree is arranged in ascending order.

\( \Gamma \): set of index values corresponding to the speed at which the importance degree is arranged in ascending order.

\( I_{(k)}^+ \): in descending order, the speed index value corresponding to the \( k \)th importance degree.

\( I_{(k)}^- \): in ascending order, the speed index value corresponding to the \( k \)th importance degree.
Step 1. In the case of optimal parameters, random forest regression (RFR) Algorithm Module (Section 5.1.1) is used to obtain the importance degree of speed series \( \{v_i - s_i\} \). Then, sort them (because the importance degrees of \( \{v_0 - s_0\}, \{v_n - s_n\} \) are zero, they are excluded) in descending order. And the \( K \) speed sequences \( \{v_k - s_k\} \) of the previous m% \((K = n * m/100)\) are selected. For the corresponding importance degree \( e_{k} \), we can get \( e_{1} \geq e_{2} \ldots \geq e_{K} \geq e_{K+1} \ldots \). Then, in ascending order, similarly, the \( K \) speed sequences \( \{v_k - s_k\} \) of the previous m% are selected, and get \( e_{1} \leq e_{2} \ldots \leq e_{K} \).

Step 2. Initialize the operation time \( t \) of the urban rail transit train, and set \( t_0 = T_{\text{min}}. \) According to the minimum and discretized unit of time \( T \text{ and } }', t \}_{\text{max}} \), we choose the minimum energy speed profile \( V_{i,\text{min}} \) from the data set, and begin to adjust the velocity sequence. The adjustment process is as follows: assume that the \( e_{k} \) importance degree corresponds to \( \{v_i - s_i\} \), then adjusted speed \( v_i = v_i + g * \sigma(g = g_{\text{min}} \ldots 0, 1, 2 \ldots g_{\text{max}}) \) \((g_{\text{min}}, g_{\text{max}}, v_{i,\text{min}}, v_{i,\text{max}} \) should meet acceleration constraints and speed constraints). To ensure the train can reach the station, displacement change caused by adjusting \( v_i \) is \( \Delta s_i \) (in formula (12)), which must be offset by another displacement change \( \Delta s_j \) (in formula (13)) in different positions. As shown in Figure 12, we choose the speed \( v_j \) at \( (e_{k} \text{ for } k = 1, 2 \ldots K) \) corresponds to \( v_j \) to offset the displacement change.

Step 4. Then, we can get a new profile after adjustment of \( v_i, v_j \). Support vector machines regression algorithm (SVR) module (Section 5.1.2) is used to calculate the energy consumption. We adjust the velocity until \( g = g_{\text{max}} \) and get the minimum energy consumption \( E^{\text{min}}_{t,K} \) during the adjustment process and the corresponding speed \( v_i \). Then, let \( v_i = v_i^* \) and \( v_j = v_j^* \).

Formulas (12) and (13) show the calculation of \( \Delta s_i^* \) and \( \Delta s_j^* \) where velocity changes are \( v_i^\Delta \) and \( v_j^\Delta \). To ensure the balance of displacement, let \( \Delta s_i^* = \Delta s_j^* \).

\[
\forall s_i^* = \forall v_i^\Delta * \frac{(t_j - t_{j-1})}{2} + \forall v_i^\Delta * \frac{(t_{j+1} - t_j)}{2} - \frac{(v_i - v_j)(t_{j+1} - t_{j-1})}{2}
\]

\[
\forall s_j^* = \forall v_j^\Delta * \frac{(t_j - t_{j-1})}{2} + \forall v_j^\Delta * \frac{(t_{j+1} - t_j)}{2} - \frac{(v_j - v_i)(t_{j+1} - t_{j-1})}{2}
\]
Finally, algorithm flow is shown in Figure 13.

6. Numerical Experiment

6.1. Section Parameters

Section Parameters

Sectional length ($s_n$): 1230m.

Speed limits (SL): (1) $0 - 200m$, $SL = 60km/h$ (2) $200m - 1100m$, $SL = 80km/h$ (3) $1100m - 1230m$, $SL = 50km/h$.

Acceleration: $a_{\text{max}} = -a_{\text{min}} = 1.5m/s^2$.

Operation time: $T_{\text{min}} = 95.4(s)$, $T_{\text{max}} = 103.4(s)$.

We take Changping Line MingTombs-Changpingxishankou section of down direction as a numerical experiment to explain the optimization process, and the section parameters are listed as above. And there are two cases in different intervals. A complete operation state is showed in Figure 14.

6.2. Optimization Result

Case 1. $s_i (i = 0, 1 \ldots n)$ is set as an uniform interval of 5m, and let $v_0 = s_{246} = 0$, $s_0 = 0$, $s_{246} = 1230$. The
operation time is 103.4s. The results after optimization are shown in Figure 15. We can see that the optimal profile is not smooth. It suddenly increases or decreases in some places. Apparently, the availability of the optimized profile is not enough.

Case 2. \( s_i (i = 0, 1 \ldots n) \) is set as an uniform interval of 50 m, and let \( v_0 = v_{26} = 0, s_0 = 0, s_{26} = 1230 \). Figure 16 shows the optimal results when \( m = 50\% \) (showed in Figure 16(a)) and \( m = 100\% \) (showed in Figure 16(b)). In this case, the operation time is also 103.4s. The optimized energy consumption can be reduced by 0.65 kwh. We can see that the speed profile is much smoother than Case 1 with rate of energy reduction is 3.1\% \((0.65/21 \times 100\%)\). In Figure 16(a), for \( m=50\% \), after optimization, the acceleration stage is slightly flat. However, in Figure 16(b), when \( m=100\% \), whole speed profile is flatter compared to the original profile, and it is more valuable in practice.
Operation sections with different distances should not have the same discrete interval. For longer section, the interval could be bigger. For example, distance of Xi’erqi-Life Science Park is 5455 m, and interval could be 200 m. In addition, the comparison of profile before and after optimization is shown in Figures 17(a)–17(j). Optimization results of other operation sections are listed in Table 6. We can see that, in some section, the maximum energy saving
Table 6: Optimization results of other sections.

<table>
<thead>
<tr>
<th>Section name</th>
<th>Minimum energy consumption of actual data (KWh)</th>
<th>After optimization (KWh)</th>
<th>Net energy saving (KWh)</th>
<th>Energy saving (%)</th>
<th>Section length (m)</th>
<th>interval (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xi'erqi-Life Science Park</td>
<td>28</td>
<td>26.94</td>
<td>1.06</td>
<td>3.79</td>
<td>5455</td>
<td>200</td>
</tr>
<tr>
<td>Life Science Park-Zhuxinzhuang</td>
<td>19</td>
<td>18.44</td>
<td>0.56</td>
<td>2.95</td>
<td>2405</td>
<td>100</td>
</tr>
<tr>
<td>Zhuxinzhuang-Gonghuacheng</td>
<td>19</td>
<td>18.36</td>
<td>0.64</td>
<td>3.39</td>
<td>3810</td>
<td>200</td>
</tr>
<tr>
<td>Gonghuacheng-Shahe</td>
<td>20</td>
<td>19.13</td>
<td>0.87</td>
<td>4.35</td>
<td>2037</td>
<td>100</td>
</tr>
<tr>
<td>Shahe-Shahe University Park</td>
<td>22</td>
<td>20.88</td>
<td>1.12</td>
<td>5.08</td>
<td>1967</td>
<td>100</td>
</tr>
<tr>
<td>Shahe University Park-Nanshao</td>
<td>30</td>
<td>29.45</td>
<td>0.55</td>
<td>1.83</td>
<td>5364</td>
<td>200</td>
</tr>
<tr>
<td>Nanshao-Beishaowa</td>
<td>14</td>
<td>13.55</td>
<td>0.45</td>
<td>3.21</td>
<td>2003</td>
<td>100</td>
</tr>
<tr>
<td>Beishawa-Changping dongguan</td>
<td>16</td>
<td>15.66</td>
<td>0.34</td>
<td>2.13</td>
<td>1687</td>
<td>100</td>
</tr>
<tr>
<td>Changping dongguan-Changping</td>
<td>22</td>
<td>21.58</td>
<td>0.42</td>
<td>1.91</td>
<td>2439</td>
<td>100</td>
</tr>
<tr>
<td>Changping-MingTombs</td>
<td>39</td>
<td>38.56</td>
<td>0.44</td>
<td>1.13</td>
<td>3522</td>
<td>200</td>
</tr>
<tr>
<td>MingTombs-Changpingxishankou</td>
<td>21</td>
<td>20.35</td>
<td>0.65</td>
<td>3.10</td>
<td>1230</td>
<td>50</td>
</tr>
<tr>
<td>Total</td>
<td>250</td>
<td>242.9</td>
<td>7.1</td>
<td>2.84</td>
<td>31964</td>
<td>-</td>
</tr>
<tr>
<td>Average value</td>
<td>22.73</td>
<td>22.08</td>
<td>0.65</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

is 5.08% (in the section Shahe to Shahe University Park), which is a good performance. And, for a 31.9km length with 12 stations train line, energy saving is 2.84%. The improvement may look modest when compared with previous researches (most claim saving energy above 4%). However, our improvement is compared with a real-world result that had already been imposed with an optimal control (traditional train optimal control with the basis of Pontryagin maximum principle). There is an ATO (automatic train system, which is equipped with optimal control) in Beijing Changping Line and Yizhuang Line. Yizhuang Line and Changping Line have some similar features, train type, number of organized group, passenger intensity, power supply mode, and so on. A well-designed method in real world that is applied into Yizhuang Line can achieve average saving energy blow 3% from the operator’s statement. Therefore, the improvement based on an ATO profile which makes it look modest is reasonable. Besides, for different section, there are different improvements. The results may be triggered by many factors, like different section external environments (radius of curve, slope, air humidity, and so on). The optimized control effects in different sections are key to the room for improvement. If the room for improvement is limited, the real improvement may be also limited. Therefore, there is no quantitative result to illustrate the different improvements in each section.

7. Conclusion

Reducing train traction energy consumption is one of the efficient ways to cut energy cost in urban rail transit systems. And to protect the environment, the optimization of urban rail transit traction energy conservation has been a significant task in urban rail transit operation and management. The traction energy consumption of a single train is related to the speed profile between stations. When energy-efficient profiles are applied in every section, there will be a positive effect on reducing energy consumption of the urban rail transit system. Therefore, train speed profile optimization is a fundamental work.

In this paper, the speed profile optimization problem is discretized, and the decision variables of the speed profile become a series of space-speed points. From this viewpoint, a data-driven urban rail transit train speed profile optimization model (DDOM) is proposed to describe the relationship between profiles and energy consumption. Two machine learning algorithms, namely, random forest regression (RFR) and support vector regression (SVR), are taken into account. RFR is applied to get the important degree of velocity in positions, and the degree is utilized as heuristic information to decide the optimization order of velocity in different positions. SVR is used to calculate energy consumption of profiles with a high accuracy (95%). Combined with the advantages of the two algorithms, an integrated heuristic greedy optimization algorithm is developed to solve the model, which can reduce energy consumption by 2.84%. In some theory research, energy conservation percentage is higher than our results. However, few are verified based on the real-world data. Furthermore, our methods may be quite simple and can be applied to practice easily.

Nevertheless, because the data samples are far from enough, when adjusting velocity in different positions to get a new profile in the optimization process, range of velocity
change is limited. There is still some room for an improvement on the basis of the optimization results. Although there are many different views, the data-driven method is new to the problem, and applying machine learning algorithms to the field of energy saving in urban rail transit is the innovation. Future research can be focused on the following areas. Firstly, a further improved algorithm for a different heuristic strategy could be studied. For instance, based on the data machine learning method, the regenerative electricity consumption in the braking process may be reused in the trains from neighboring sections. Thus, instead of optimizing one single train speed profile in each section separately, train speed profiles from neighboring sections should be taken into account. Secondly, in the urban rail transit networks, if power supply in the network nodes (transfer stations) is transmitted from the same transformer substation, the energy-saving optimization of trains can be extended to the urban rail transit network.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by the China National Funds for Distinguished Young Scientists (71525002), National Nature Science Foundation of China (71890972/71890970, 71771018, and 71621001), and Beijing Municipal Natural Science Foundation (L181008).

References


