WILEY | Hindawi

Research Article

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

Kang Huang [1,2,3] **Jianjun Wu** [1,2] **Xin Yang,** [1] **Ziyou Gao,** [1,2] **Feng Liu,** [4] **and Yuting Zhu** [3,5]

[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

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
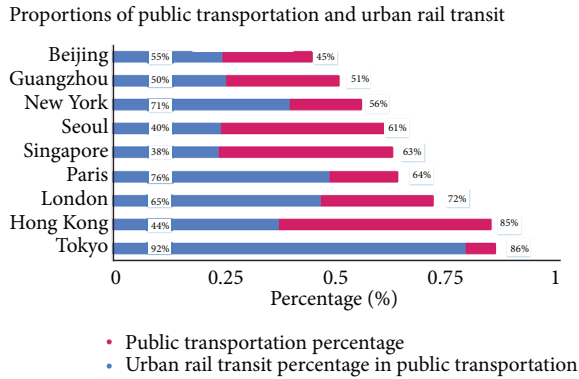
Figure 1: Proportions of public transportation and urban rail transit.

Figure 2: Overall framework.

YEARBOOK from Beijing statistical information website). In the European Union (EU), for instance, transport causes approximately 31% of total greenhouse gas (GHG) emissions. Within this sector, metropolitan transportation is responsible for about 25% of the total CO_2 emissions (González-Gil et al. [6]). 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
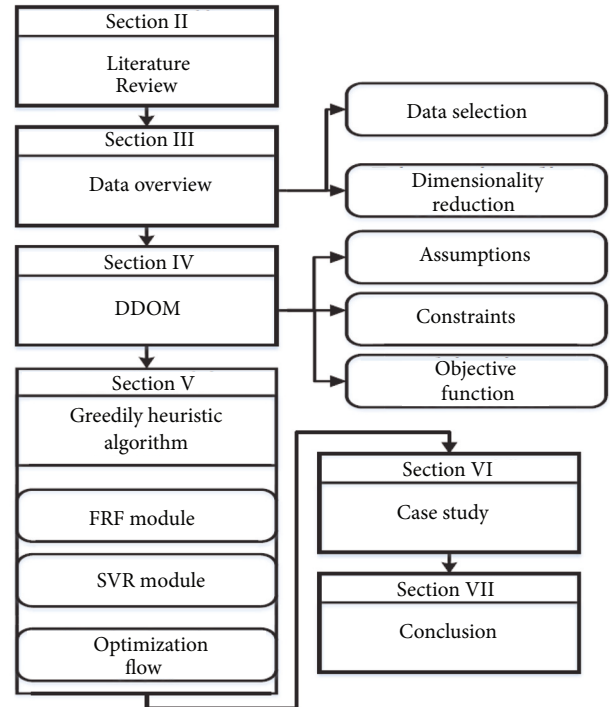
stable performance in the data set, and they have been widely used in many fields, such as biology, medicine, economy, managementm and so on [8] Therefore, it becomes possible to optimize the train speed profile in the urban rail transit system on the premise of verifying their effectiveness.

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.

View	Publication	Years	Model type	Objective	Energy consumption calculational method	Solution method
I	Cheng and Howlett. [9].	1993	Discrete control model	Energy consumption of profile	Empirical-formula, numerical integration	Optimize control
I	Howlett et al. [10].	1996	Continuous control model	Energy consumption of profile	Empirical-formula, numerical integration	Optimize control
I	Wong and Ho [11]	2004	Discrete control model	Energy consumption of profile	Genetic method	Genetic search
I	Albrecht et al. [12]	2013	Continuous control model	Energy consumption of profile	Empirical-formula, numerical integration	Optimize control
I	Albrecht, Howlett et al. [13, 14]	2016 a&b	Continuous control model	Energy consumption of profile	Empirical-formula, numerical integration	Optimize control
I	Yin et al. [15]	2014	Reinforcement learning	Energy consumption of profile	Empirical-formula, numerical integration Simulation platform	Dynamic programming
I& II	Nasri et al. [16]	2010	Simulation model	Energy consumption of timetable	Empirical-formula, numerical integration Simulation platform	Simulation
II	Sun et al. [17]	2013	MILP	Energy consumption of timetable	Empirical-formula, numerical integration	Genetic search
II	Yang et al. [18]	2015a	MILP	Energy consumption of whole line	Taking into consideration recovery energy	Genetic search
II	Li and Lo [19, 20]	2014 a&b	Integrated-operation model	Energy consumption of network	Empirical-formula, numerical integration	Genetic search
II	Canca and Zarzo [21]	2017	MILP	Energy consumption of whole line	Empirical-formula, numerical integration	Iterative algorithm and Python+Gurobi
II	Yin et al. [22]	2017	MILP	Energy consumption and the passenger waiting time	Empirical-formula, numerical integration	Lagrangian relaxation (LR)-based heuristic algorithm

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
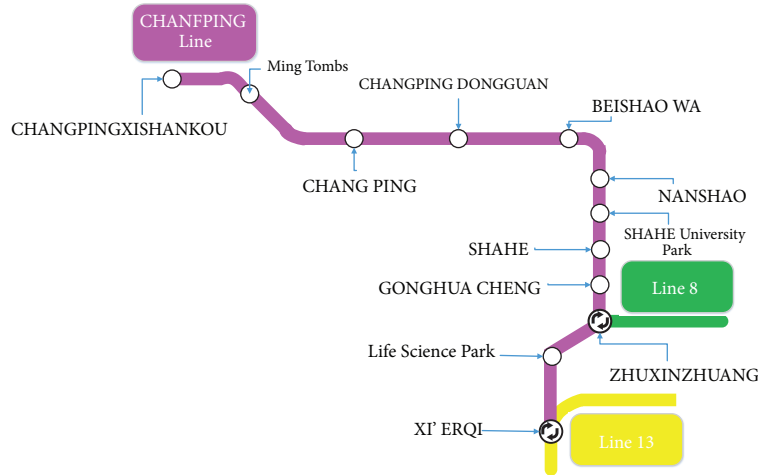
FIGURE 3: Illustration of the Changping Line.

TABLE 2: Overview of measurement characteristics.

Parameter	Unit	Resolution
Speed	km/h	0.001
Position	m	0.001
Time	s	0.2
Train weight	ton	1
Current slope	1‰	1
EBI speed	km/h	0.001
Station spacing	m	0.001
Expected acceleration of PID	(km/h)/s	1
Electric energy consumption	Kwh	1

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. Fernándeza 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.

3.2. Data Preprocessing

Symbols

> n:　number of section is discretized to.
>
> v_j^0:　jth speed point of original profile C_i.

TABLE 3: Part types of the original data.

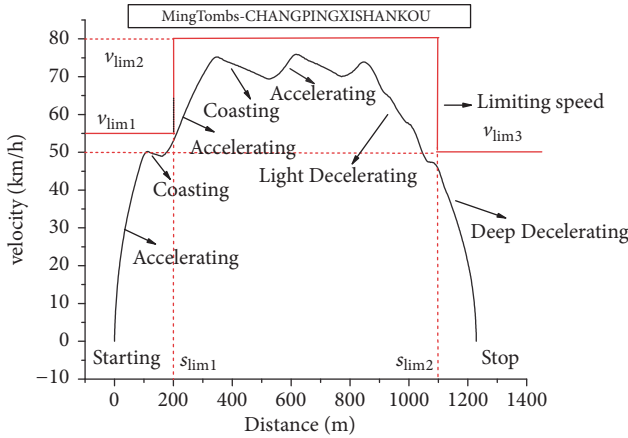| Time	Velocity(km/h)	Distance(m)
t_1^0	v_1^0	s_1^0
t_2^0	v_2^0	s_2^0
......	......	......
t_i^0	v_i^0	s_i^0
......	......	......
t_m^0	v_m^0	s_m^0

FIGURE 4: MingTombs-Changpingxishankou.

s_j^0: jth position point of original profile C_i.

t_j^0: jth time point of original profile C_i.

s_i: it is the ith displacement from the beginning of the urban rail transit section.

v_i: the speed at s_i.

t_i: the time at s_i.

∇t: the time interval used to record the speed and displacement data during train traction.

S_i^0:   distance set at a time interval ∇t. $S_i^0 = \{s_j^0 \mid j = 1, 2 \dots n\}$.

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 (showed 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: $0 \longrightarrow S_{lim1}$, $S_{lim1} \longrightarrow S_{lim2}$, $S_{lim2} \longrightarrow End$. Each part has its maximum speed limit.

Train running state form is shown in Table 3 (m: the number of data recorded on an original speed profile). 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 $\{s_0, \dots s_i \dots s_n \mid i = 0, 1, \dots n\}$. 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 ∇t interval, recorded as s_i^- and s_i^+. Sequence $\{s_0^- \dots s_i^- \dots s_n^-\}$, $\{s_0^+ \dots s_i^+ \dots s_n^+\}$ 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_i^-, 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 $\{v_0 \dots v_i \dots v_n\}$, 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.

3.3. Extraction of Training Data Set and Testing Data Set. After processing above, $\{v_i - s_i\}$, $\{v_i^+ - s_i^+ - t_i^+\}$, $\{v_i^- - s_i^- - t_i^-\}$ can be obtained. For example, let $s_i(i = 0, 1 \dots n)$ be with
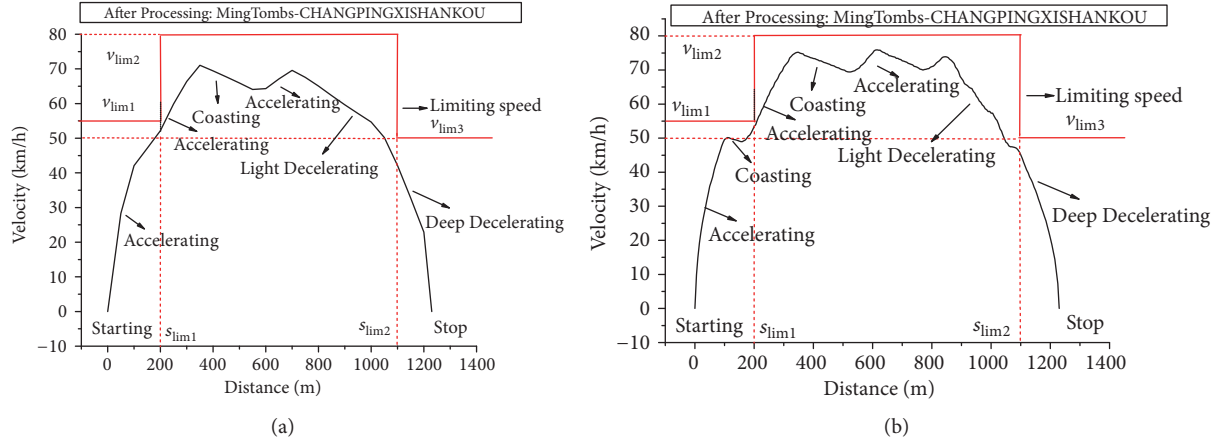
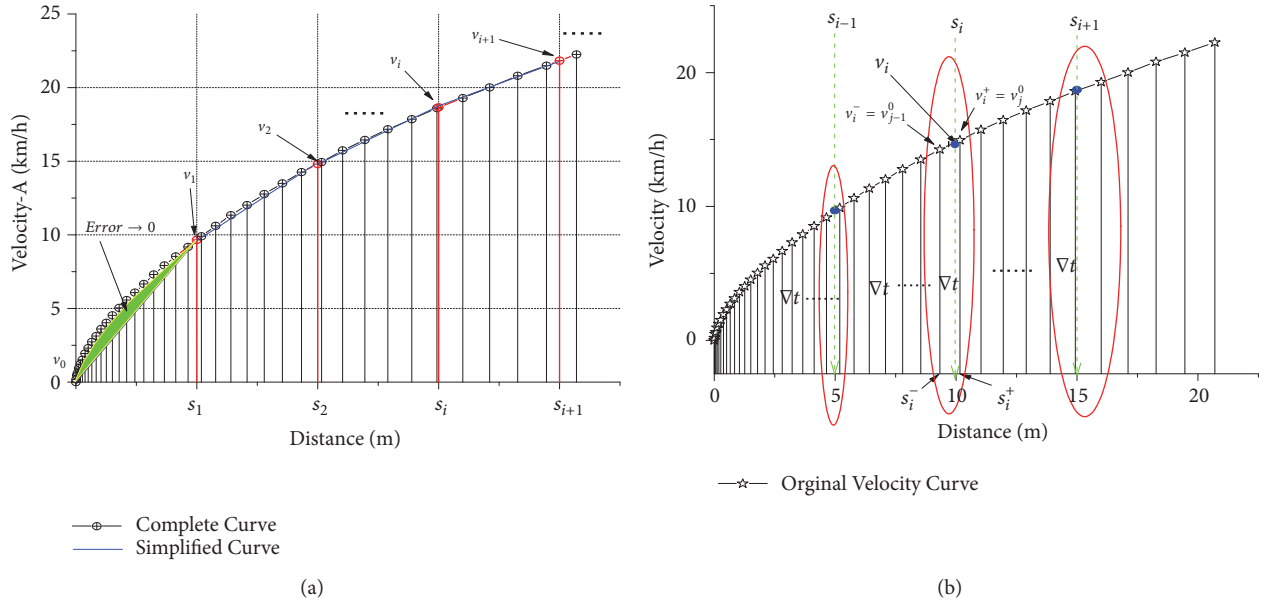FIGURE 5: Profiles description at different distance intervals. (a) 50m interval. (b) 5m interval.

FIGURE 6: Dimension reduction process of velocity profile. (a) Get the $\{v_0 \ldots v_i \ldots v_n\}$. (b) Error in simplified profile.

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 ∇t. $V_i^0 = \{v_j^0 \mid j = 1, 2 \ldots n\}$.

S_i^0: distance set at a time interval ∇t. $S_i^0 = \{s_j^0 \mid j = 1, 2 \ldots n\}$.

T_i^0: time set with a time interval of ∇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_t: energy consumption of urban rail transit traction under running time of t.

TABLE 4: Part of the velocity series after being processed.

i	$v_i^-(km/h)$	$s_i^-(m)$	$t_i^-(\nabla t)$	$v_i(km/h)$	$s_i(m)$	$v_i^+(km/h)$	$s_i^+(m)$
1	0	0	0	0	0	0	0
2	9.18	4.64	21	9.657	5	9.9	5.19
3	14.256	9.33	28	14.811	10	14.94	10.16
4	18.612	14.928	34	18.659	15	19.296	16
5	21.492	19.462	38	21.824	20	22.248	20.698
6	24.408	24.646	42	24.593	25	25.128	26.042
7	26.568	28.954	45	27.042	30	27.252	30.468
8	28.764	33.622	48	29.371	35	29.484	35.26
9	30.888	38.652	51	31.442	40	31.608	40.408
10	33.012	44.04	54	33.419	45	33.804	45.918
11	35.172	49.786	57	35.25	50	35.892	51.78
12	36.576	53.812	59	36.991	55	37.296	55.884
13	37.944	57.992	61	38.617	60	38.664	60.14
14	40.068	64.554	64	40.197	65	40.716	66.816
15	41.436	69.118	66	41.709	70	42.156	71.46
16	42.804	73.838	68	43.134	75	43.488	76.254
17	44.172	78.708	70	44.528	80	44.856	81.2
18	45.576	83.732	72	45.897	85	46.224	86.3
19	46.944	88.908	74	47.213	90	47.592	91.552
20	48.204	94.23	76	48.368	95	48.78	96.94

TABLE 5: Data format of training and testing set.

Serial number	s_0	s_1	...	s_{n-1}	s_n	Time	Energy consumption
1	v_0^1	v_1^1	...	v_{n-1}^1	v_n^1	t^1	E^1
2	v_0^2	...	...	...	...	t^2	E^2
...	...	...	...	...	...	...	...
q-1	v_0^{q-1}	....	...	...	...	...	E^{q-1}
q	v_0^q	v_1^q	...	v_{n-1}^q	v_n^q	t^n	E^q

$v_{S_i}^{min}$: minimum speed limit corresponding to s_i.

$v_{S_i}^{max}$: maximum speed limit corresponding to s_i.

a_{min}: minimum acceleration limit in operational section.

a_{max}: maximum acceleration limit in operational section.

T_{min}: minimum time limit in operational section.

T_{max}: maximum time limit in operational section.

Assumption. During the process of $s_i^- \longrightarrow s_i \longrightarrow s_i^+$, 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.

$$\frac{\left((v_i^+)^2 - v_i^2\right)}{(s_i^+ - s_i)} = 2a_i \quad i = 1 \ldots n \tag{1}$$

$$\frac{\left(v_i^2 - (v_i^-)^2\right)}{(s_i - s_i^-)} = 2a_i \quad i = 1 \ldots n \tag{2}$$

$$a_i = \frac{(v_i^+ - v_i^-)}{\nabla t} \quad i = 1 \ldots n \tag{3}$$

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

$$v_i = \sqrt{2a_i(s_i - s_i^-) + (v_i^-)^2}$$
$$= \sqrt{\frac{2(v_i^+ - v_i^-)(s_i - s_i^-)}{\nabla t} + (v_i^-)^2} \quad i = 1 \ldots n \tag{4}$$

or

$$v_i = \sqrt{(v_i^+)^2 - 2a_i(s_i^+ - s_i)}$$
$$= \sqrt{(v_i^+)^2 - \frac{2(v_i^+ - v_i^-)(s_i^+ - s_i)}{\nabla t}} \quad i = 1 \ldots n \tag{5}$$

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.

$$v_{s_i}^{min} < v_{s_i} < v_{s_i}^{max} \quad i = 1 \ldots n \qquad (6)$$

$v_{s_i}^{max}$ and $v_{s_i}^{min}$ 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_{min} and a_{max} are determined by actual empirical parameters, and $a_{max} > 0, a_{min} < 0$.

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

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

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).

$$t \in \left[T_{min}, T_{max}\right] \qquad (9)$$

where T_{min} and T_{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.

$$s_n = S_0 \qquad (10)$$

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...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).

$$\min E = \min_t \left\{ E_t \left(\{v_i - s_i\} \mid i = 0, 1 \ldots n \right) \right\}$$
$$t \in \left[T_{min}, T_{max}\right], \left(\{v_i - s_i\} \mid i = 0, 1 \ldots n \right) \in C^* \qquad (11)$$

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
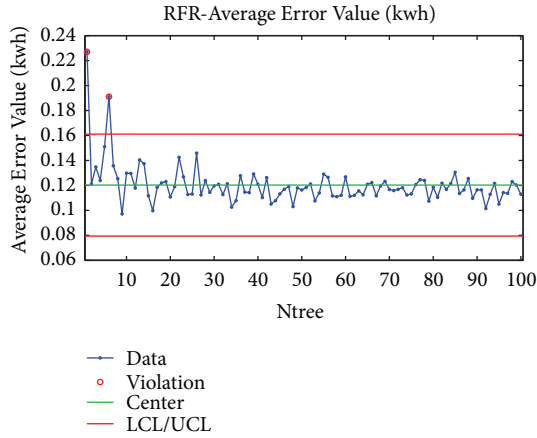
FIGURE 7: The error values at different Ntrees.

Mtry=2 or 3, the error is minimal. Therefore, the optimal parameter combination used in this paper is Mtry=2 or 3 and Ntree≥50. By using the FR algorithm, the traction energy consumption evaluation average error is less than 0.1kwh and within range of 1%.

In addition to the high precision evaluation ability, we also get importance degrees of the velocity in different displacements during the traction energy consumption of the urban rail transit. We can find that the speed at which position is more significant to the energy consumption in a section, which indicates contributions to energy consumption of pairs space-speed. For instance, in the section of MingTombs-Changpingxishankou, section length is 1230 m, the importance degrees at different positions are shown in Figure 9.

5.1.2. Support Vector Machine Regression (SVR) Algorithm Module. Support vector machine (SVM) algorithm is from statistical learning theory (SLT), which is based on the structural risk minimization principle that can avoid excessive learning problems and ensure the generalization ability of the model. In essence, it can solve the convex quadratic programming problem and avoid falling into the local minimum. It can be applied not only to classification problems but also to the case of regression [42]. Therefore, it can be divided into support vector classification (SVC) and support vector regression (SVR). Because of its solid theoretical foundation and its complete theoretical derivation, support vector machine is an effective tool in dealing with small samples, nonlinear, local issues. In this paper, it is applied to calculate the energy consumption based on real data.

Before using the SVR, the first step requires the determination of the kernel functions. The second step is to optimize parameters corresponding to different kernel functions. In this paper, three typical kernel functions are verified: radial basis kernel function (RBF), linear kernel function (LINEAR), and polynomial kernel function (POLY).

(1) For RBF, calibration parameters include C penalty factor and *Gamma* value. As shown in Figure 10(a), convergence rate of RBF is very fast. When $C \geq 20$, the error will drop to a lower level. As $C \geq 100$, the average error of traction energy

consumption can reach about 0.1kwh. The best combination of parameters is $C \geq 30$, and *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 $\{v_i - s_i\}(i = 0, 1 \ldots n)$ should be found. Velocity v_i in every position can be in a range, and the number of $\{v_i - s_i\}(i = 0, 1 \ldots n)$ 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
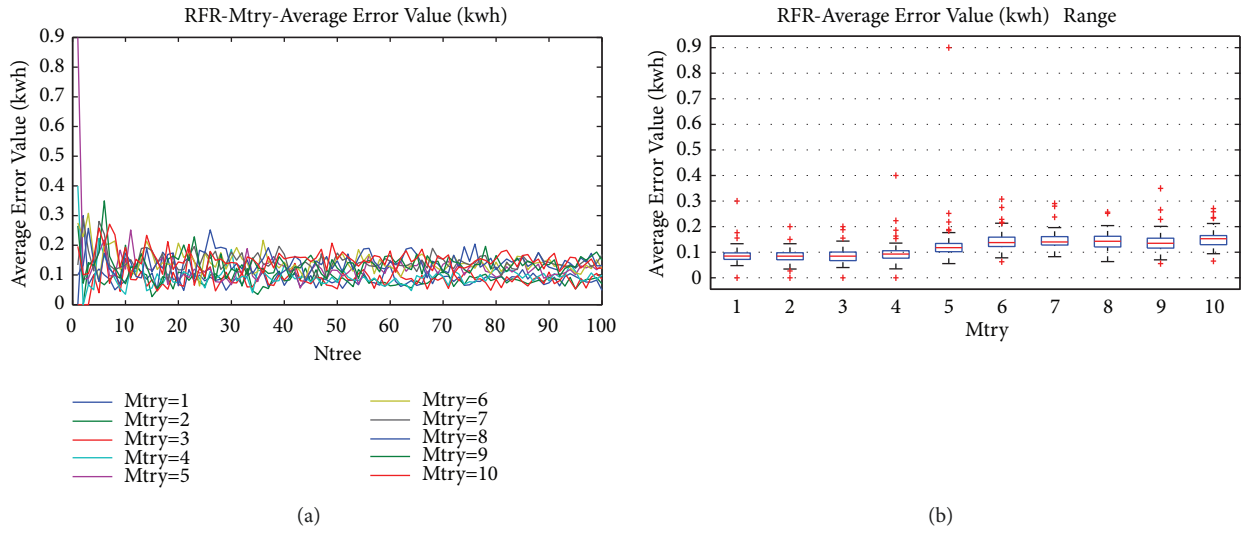
FIGURE 8: Convergence process and errors in RFR. (a) Errors in different Mtrys. (b) Convergence range.
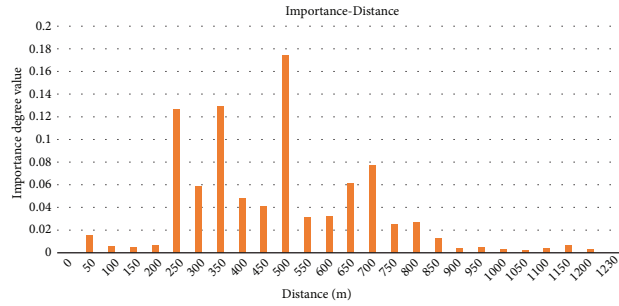

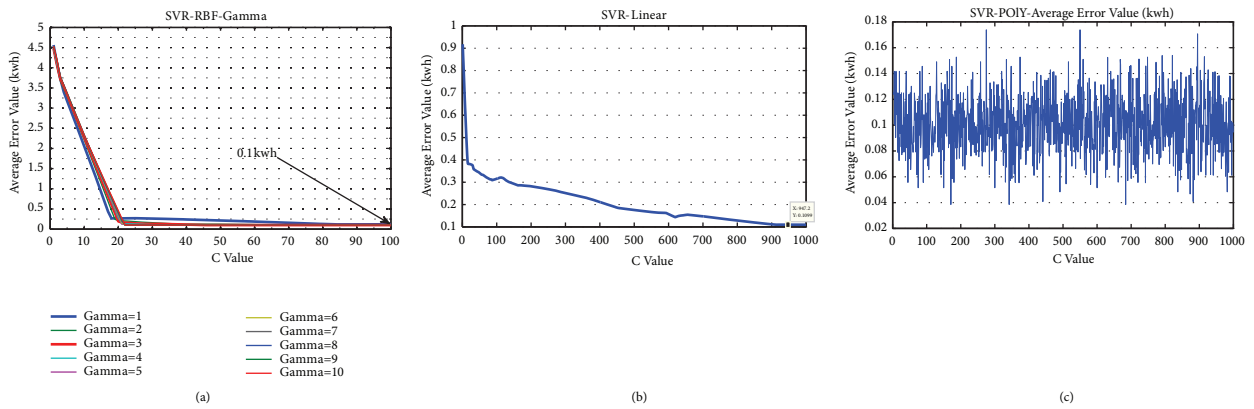FIGURE 9: Importance of velocity at different locations in the section.

FIGURE 10: Convergence process in different kernel functions. (a) SVR-RBF-Gamma. (b) SVR-LINEAR. (c) SVR-POLY.

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.

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

$I_{(k)}-$:   in ascending order, the speed index value corresponding to the kth importance degree.
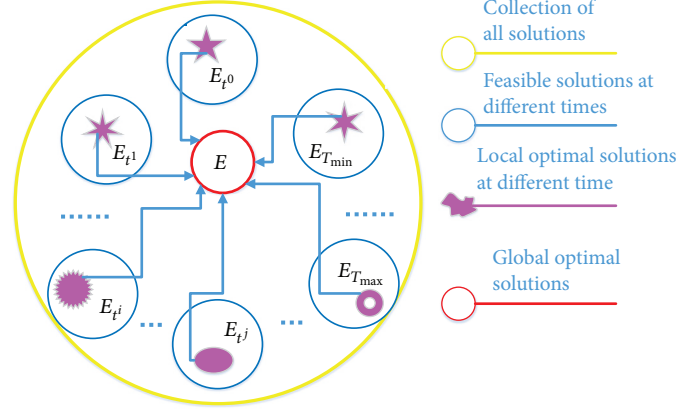
FIGURE 11: Distribution of solutions.

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 $e_k^+(1 \le k \le K)$, we can get $e_1^+ \ge e_2^+ \ldots \ge e_k^+ \ldots \ge e_K^+$. Then, in ascending order, similarly, the K speed sequences $\{v_{k^-} - s_{k^-}\}$ of the previous m% are selected, and get $e_1^- \le e_2^- \ldots \le e_k^- \ldots \le e_K^-$.

Step 2. Initialize the operation time t of the urban rail transit train, and set $t_0 = T_{min}$. According to the minimum and maximum time in the data, T_{min}, T_{max} are determined, and discretized unit of time is ∇t. Then let $k = 1, r = 0$.

Step 3. In the case of $t = t_0 + r * \nabla t(r = 0, 1, 2 \ldots r_{max}) \in [T_{min}, T_{max}]$, we choose the minimum energy speed profile C_t^{min} from the data set, and begin to adjust the velocity sequence. The adjustment process is as follows: assume that the $e_k^+\{k = 1, 2 \ldots K\}$ importance degree corresponds to $\{v_i - s_i\}$, then adjusted speed v_i is $v_i^\wedge = v_i + g * \sigma(g = g_{min} \ldots 0, 1, 2 \ldots g_{max})$ ($g_{min}, g_{max}, v_{i,min}^*$, and $v_{i,max}^*$ should meet acceleration constraints and speed constraints). To ensure the train can reach the station, displacement change caused by adjusting v_i is $\nabla s_i^\wedge$ (in formula (12)), which must be offset by another displacement change ∇s_j^- (in formula (13)) in different positions. As shown in Figure 12, we choose the speed v_j at ($e_k^-\{k = 1, 2 \ldots K\}$ corresponds to v_j) to offset the displacement change.

Step 4. Then, we can get a new profile after adjustment of v_i and 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 $g = g_{max}$ and get the minimum energy consumption $E_{t,k}^{min}$ during the adjustment process and the corresponding speed $v_i^\wedge$. Then, let $v_i = v_i^\wedge$ and $v_j = v_j^\wedge$.

Formulas (12) and (13) show the calculation of $\nabla s_i^\wedge$ and ∇s_j^- where velocity changes are $\nabla v_i^\wedge$ and ∇v_i^-. To ensure the
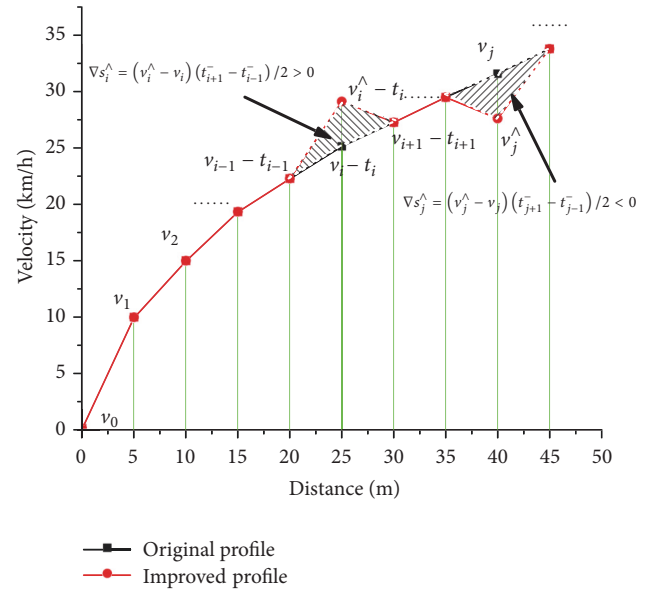

FIGURE 12: Explanation of changes of velocity and displacement.

balance of displacement, let $\nabla s_i^\wedge = \nabla s_j^-$.

$$\nabla s_i^\wedge = \nabla v_i^\wedge * \frac{(t_i^- - t_{i-1}^-)}{2} + \nabla v_i^\wedge * \frac{(t_{i+1}^- - t_i^-)}{2}$$
$$= \nabla v_i^\wedge * \frac{(t_{i+1}^- - t_{i-1}^-)}{2} = \frac{(v_i^\wedge - v_i)(t_{i+1}^- - t_{i-1}^-)}{2} \quad (12)$$

$$\nabla s_j^- = \nabla v_j^- * \frac{(t_j^- - t_{j-1}^-)}{2} + \nabla v_j^- * \frac{(t_{j+1}^- - t_j^-)}{2}$$
$$= \nabla v_j^- * \frac{(t_{j+1}^- - t_{j-1}^-)}{2} = \frac{(v_j^- - v_j)(t_{j+1}^- - t_{j-1}^-)}{2} \quad (13)$$

Step 5. If $k = K$, then go to Step 6; if $k = k + 1$, repeat Step 3.

Step 6. If $t = T_{max}$, then go to Step 7; if $r = r+1$, repeat Step 3.

Step 7. Get all the energy consumption $E_{t,K}^{min}, t \in [T_{min}, T_{max}]$. Then, $MinE = min_t E_{t,K}^{min}, K = m * n/100, t \in [T_{min}, T_{max}]$.
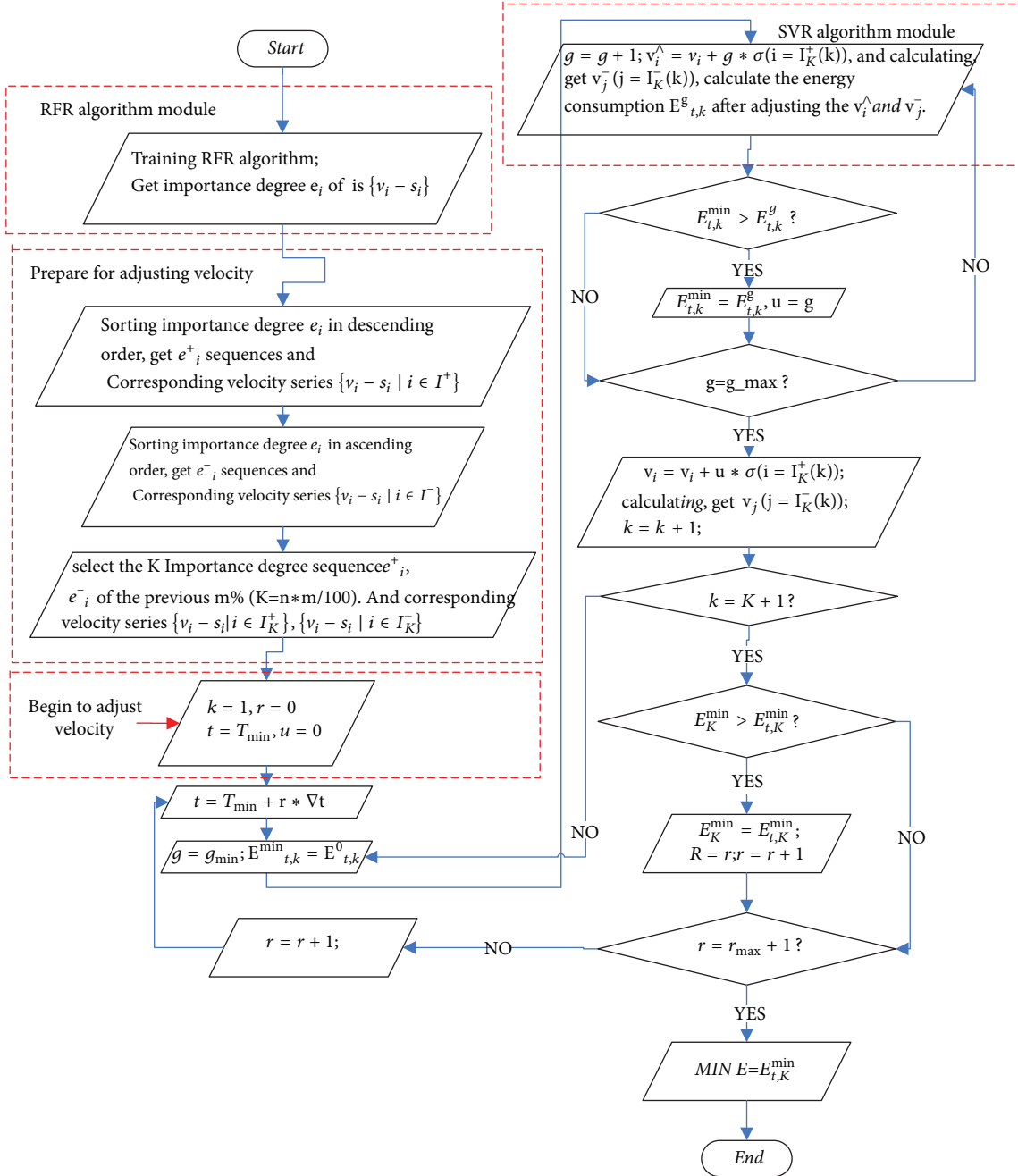
FIGURE 13: Algorithm flow.

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

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

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 5 m, and let $v_0 = v_{246} = 0$, $s_0 = 0, s_{246} = 1230$. The
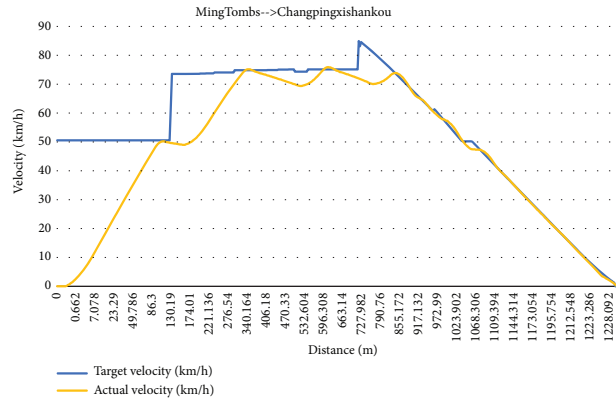
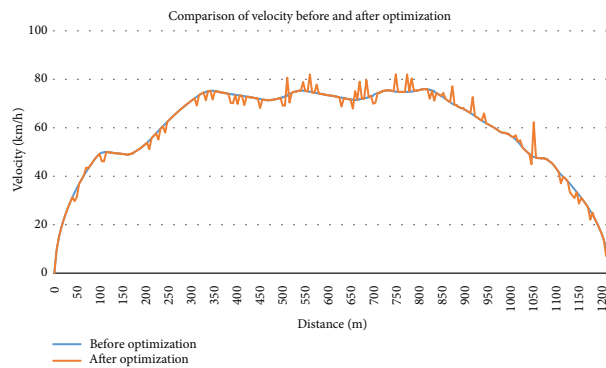Figure 14: Train operation state.
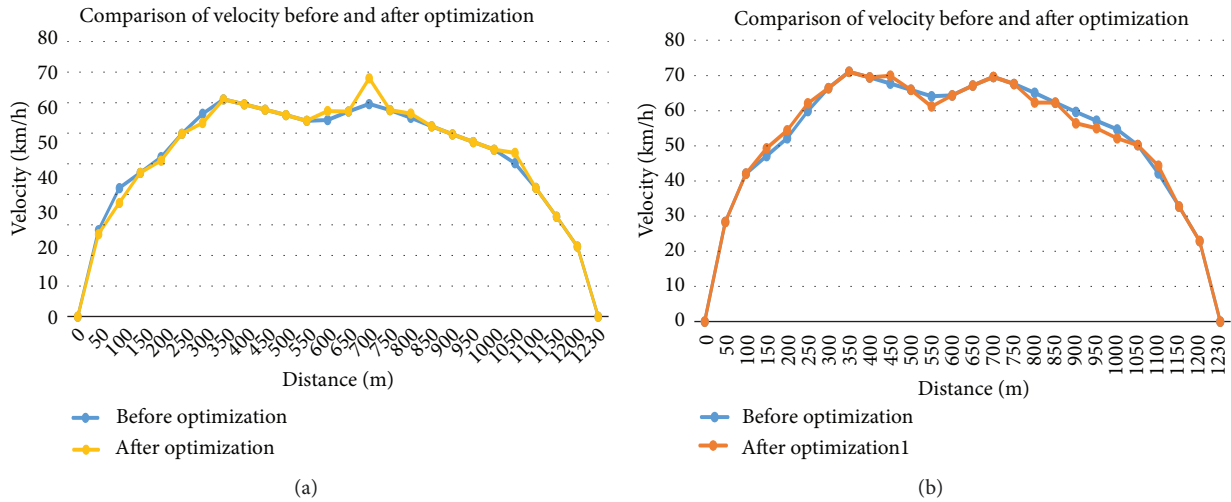

Figure 15: Optimization result with small intervals.

(a)

(b)

Figure 16: Optimization results with big intervals. (a) m=50%. (b) m=100%.

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 ∗ 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.
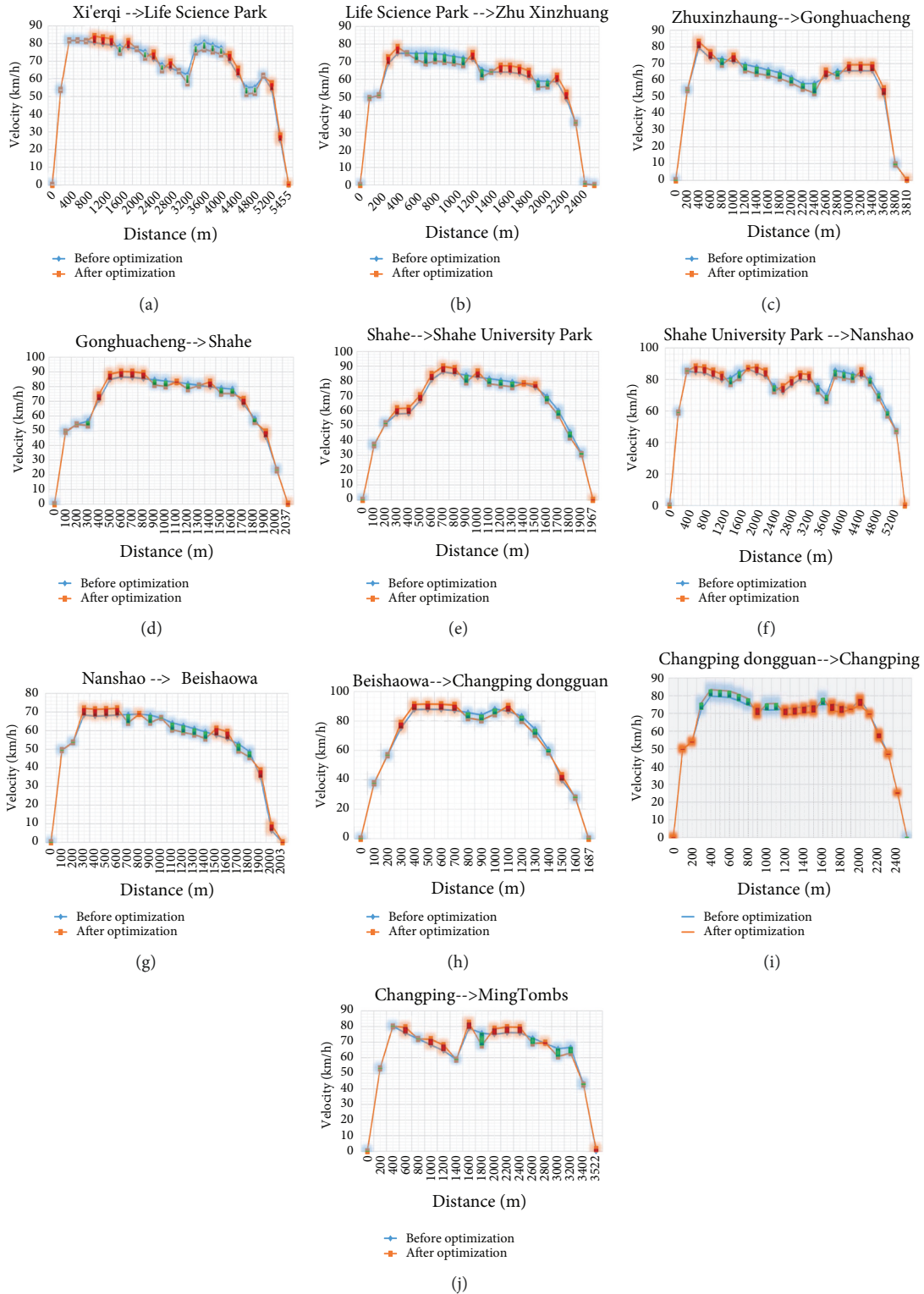
Figure 17: The obtained profiles in different sections. Section (a)–(j) are listed in Table 6.

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.

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

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 on 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

References

[1] X. Guo, J. Wu, J. Zhou, X. Yang, D. Wu, and Z. Gao, "First-train timing synchronization using multi-objective optimization in urban transit networks," *International Journal of Production Research*, 2018.

[2] L. Kang, X. Zhu, H. Sun, J. Wu, Z. Gao, and B. Hu, "Last train timetabling optimization and bus bridging service management in urban railway transit networks," *OMEGA - The International Journal of Management Science*, vol. 74, no. 1, pp. 31–44, 2018.

[3] X. Yang, H. Yin, J. Wu, Y. Qu, Z. Gao, and T. Tang, "Recognizing the critical stations in urban rail networks: an analysis method based on the smart-card data," *IEEE Intelligent Transportation Systems Magazine*, vol. 11, no. 1, pp. 29–35, 2019.

[4] J. Yin, Y. Wang, T. Tang, J. Xun, and S. Su, "Metro train rescheduling by adding backup trains under disrupted scenarios," *Frontiers of Engineering Management*, vol. 4, no. 4, pp. 418–427, 2017.

[5] T. Tang and J. Xun, "Research on energy-efficient driving strategy in Beijing Yizhuang line," *Journal of BeijingJiaoTong University*, vol. 40, no. 4, pp. 20–24, 2016.

[6] A. González-Gil, R. Palacin, P. Batty, and J. P. Powell, "A systems approach to reduce urban rail energy consumption," *Energy Conversion and Management*, vol. 80, pp. 509–524, 2014.

[7] H. Yin, J. Wu, Z. Liu, H. Yin, Y. Qu, and H. Sun, "Optimizing the release of passenger flow guidance information in urban rail transit network via agent-based simulation," *Applied Mathematical Modelling*, vol. 72, no. 8, pp. 337–355, 2019.

[8] R. Genuer, J.-M. Poggi, C. Tuleau-Malot, and N. Villa-Vialaneix, "Random forests for big data," *Big Data Research*, vol. 9, no. 3, pp. 28–46, 2017.

[9] J. X. Cheng and P. Howlett, "A note on the calculation of optimal strategies for the minimization of fuel consumption in the control of trains," *IEEE Transactions on Automatic Control*, vol. 38, no. 11, pp. 1730–1734, 1993.

[10] P. Howlett, "Optimal strategies for the control of a train," *Automatica*, vol. 32, no. 4, pp. 519–532, 1996.

[11] K. Wong and T. Ho, "Coast control for mass rapid transit railways with searching methods," *IEE Proceedings - Electric Power Applications*, vol. 151, no. 5, pp. 365–376, 2004.

[12] A. R. Albrecht, P. G. Howlett, P. J. Pudney, and X. Vu, "Energy-efficient train control: from local convexity to global optimization and uniqueness," *Automatica*, vol. 49, no. 10, pp. 3072–3078, 2013.

[13] A. Albrecht, P. Howlett, P. Pudney, X. Vu, and P. Zhou, "The key principles of optimal train control—Part 1: Formulation of the model, strategies of optimal type, evolutionary lines, location of optimal switching points," *Transportation Research Part B: Methodological*, vol. 94, pp. 482–508, 2016.

[14] A. Albrecht, P. Howlett, P. Pudney, X. Vu, and P. Zhou, "The key principles of optimal train control—Part 2: Existence of an optimal strategy, the local energy minimization principle, uniqueness, computational techniques," *Transportation Research Part B: Methodological*, vol. 94, pp. 509–538, 2016.

[15] J. Yin, D. Chen, and L. Li, "Intelligent train operation algorithms for urban rail transit by expert system and reinforcement learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 6, pp. 2561–2571, 2014.

[16] A. Nasri, M. Fekri Moghadam, and H. Mokhtari, "Timetable optimization for maximum usage of regenerative energy of braking in electrical railway systems," in *International Symposium on Power Electronics, Electrical Drives, Automation and Motion*, pp. 1218–1221, Pisa, Italy, 2010.

[17] H. Sun, J. Wu, H. Ma, X. Yang, and Z. Gao, "A bi-objective timetable optimization model for urban rail transit based on the time-dependent passenger volume," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 2, pp. 604–615, 2019.

[18] X. Yang, A. Chen, J. Wu, Z. Gao, and T. Tang, "An energy-efficient rescheduling approach under delay perturbations for metro systems," *Transportmetrica B: Transport Dynamics*, vol. 7, no. 1, pp. 386–400, 2019.

[19] X. Li and K. Lo Hong, "An energy-efficient scheduling and speed control approach for metro rail operations," *Transportation Research Part B: Methodological*, vol. 64, pp. 73–89, 2014.

[20] X. Li and H. K. Lo, "Energy minimization in dynamic train scheduling and control for urban rail transit rail operations," *Transportation Research Part B: Methodological*, vol. 70, no. 1, pp. 269–284, 2014.

[21] D. Canca and A. Zarzo, "Design of energy-Efficient timetables in two-way railway rapid transit lines," *Transportation Research Part B: Methodological*, vol. 102, pp. 142–161, 2017.

[22] J. Yin, L. Yang, T. Tang, Z. Gao, and B. Ran, "Dynamic passenger demand oriented metro train scheduling with energy-efficiency and waiting time minimization: Mixed-integer linear programming approaches," *Transportation Research Part B: Methodological*, vol. 97, pp. 182–213, 2017.

[23] G. M. Scheepmaker, R. M. Goverde, and L. Kroon, "Review of energy-efficient train control and timetabling," *European Journal of Operational Research*, vol. 257, no. 2, pp. 355–376, 2017.

[24] P. G. Howlett, I. P. Milroy, and P. J. Pudney, "Energy-efficient train control," in *Advances in Industrial Control*, Springer, London, UK, 1995.

[25] P. Howlett, "A new look at the rate of change of energy consumption with respect to journey time on an optimal train journey," *Transportation Research Part B: Methodological*, vol. 94, pp. 387–408, 2016.

[26] G. M. Scheepmaker and R. M. P. Goverde, "The interplay between energy-efficient train control and scheduled running time supplements," *Journal of Rail Transport Planning and Management*, vol. 5, no. 4, pp. 225–239, 2015.

[27] X. Yang, X. Li, B. Ning, and T. Tang, "A survey on energy-efficient train operation for urban rail transit," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 1, pp. 2–13, 2016.

[28] Z. Tian, P. Weston, N. Zhao, S. Hillmansen, C. Roberts, and L. Chen, "System energy optimisation strategies for metros with regeneration," *Transportation Research Part C: Emerging Technologies*, vol. 75, pp. 120–135, 2017.

[29] S. Yang, J. Wu, X. Yang, F. Liao, D. Li, and Y. Wei, "Analysis of energy consumption reduction in metro system using rolling stop-skipping patterns," *Computers & Industrial Engineering*, vol. 127, no. 1, pp. 129–142, 2019.

[30] R. Chevrier, P. Pellegrini, and J. Rodriguez, "Energy saving in railway timetabling: a bi-objective evolutionary approach for computing alternative running times," *Transportation Research Part C: Emerging Technologies*, vol. 37, pp. 20–41, 2013.

[31] P. Wang and R. M. P. Goverde, "Multi-train trajectory optimization for energy efficiency and delay recovery on single-track railway lines," *Transportation Research Part B: Methodological*, vol. 105, pp. 340–361, 2017.

[32] L. Wang, L. Yang, Z. Gao, and Y. Huang, "Energy-saving operation approaches for urban rail transit systems," *Frontiers of Engineering Management*, vol. 4, no. 4, pp. 408–417, 2017.

[33] N. Zhao, C. Roberts, S. Hillmansen, Z. Tian, P. Weston, and L. Chen, "An integrated metro operation optimization to minimize energy consumption," *Transportation Research Part C: Emerging Technologies*, vol. 75, pp. 168–182, 2017.

[34] P. G. Howlett, P. J. Pudney, and X. Vu, "Local energy minimization in optimal train control," *Automatica*, vol. 45, no. 11, pp. 2692–2698, 2009.

[35] Y. Huang, H. Yu, J. Yin et al., "An integrated approach for the energy-efficient driving strategy optimization of multiple trains by considering regenerative braking," *Computers & Industrial Engineering*, vol. 126, pp. 399-400, 2018.

[36] S. Yang, J. Wu, X. Yang, H. Sun, and Z. Gao, "Energy-efficient timetable and speed profile optimization with multi-phase speed limits: theoretical analysis and application," *Applied Mathematical Modelling*, vol. 56, no. 4, pp. 32–50, 2018.

[37] P. M. Fernández, C. G. Román, and R. I. Franco, "Modelling electric trains energy consumption using neural networks," *Transportation Research Procedia*, vol. 18, pp. 59–65, 2016.

[38] F. Ghofrani, Q. He, R. M. P. Goverde, and X. Liu, "Recent applications of big data analytics in railway transportation systems: A survey," *Transportation Research Part C: Emerging Technologies*, vol. 90, pp. 226–246, 2018.

[39] R. S. Michalski, I. Bratko, and M. Kubat, "Machine learning and data mining methods and application," *ACM SIGKDD Explorations Newsletter*, vol. 2, no. 2, pp. 110–114, 2004.

[40] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[41] A. Liaw and M. Wiener, "Classification and regression by random forest," *R News*, vol. 23, no. 23, pp. 18–22, 2002.

[42] D. Basak and S. Pal, "Support vector regression," *Statistics and Computing*, vol. 11, no. 10, pp. 203–224, 2007.