

# Driving Strategy of Heavy Haul Train based on Support Vector Regression

Shuo Yang<sup>1, a</sup>, Xiaofeng Yang<sup>1, b</sup> and Zhengnan Lin<sup>1, c</sup>

<sup>1</sup>*School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100000, China.*

**Keywords:** Heavy Haul Train; Driving Strategy; Support Vector Regression (SVR).

**Abstract:** In order to reduce the labor intensity of heavy-haul train drivers in the downgrade section and near the split phase area, the paper analyzes the factors that affect the driving strategy of heavy-haul train in accordance with the manual driving strategy. In this paper, A control model of heavy-haul train electric braking force based on support vector regression (SVR) is proposed to control the electric braking force. With the manual driving records used as training data, Electric braking force and other information are extracted as output results and features to train the control model. By trial and error, parameters of the control model are adjusted to optimize the model. The results show that the control model in this paper is close to the manual driving in the same situation, which is positive for reducing the labor intensity of drivers in heavy-haul railway.

## 1 INTRODUCTION

In freight transportation, the heavy-haul railway has the advantages of large capacity, high efficiency and low transportation cost, which is of great significance to the "west to East Coal Transportation" project in China.

Due to the difficulty of heavy-haul train driving, the drivers need to give their whole attention to driving with no distractions for a long time; on the other hand, the heavy-haul line is long, the drivers need to drive without mistake for more than eight hours, which cause the high labor intensity of drivers.

In the relevant research of heavy-haul trains, Wang et al. (Xi Wang, et.al, 2018) had studied the air braking of heavy-haul trains on the long and steep downgrade to ensure the safety of heavy-haul trains in the implementation of air braking; Yu et al. (H. Yu, et.al, 2018) Put forward an intelligent optimization method based on particle swarm optimization (PSO) to generate driving strategy; Lin et al. (Xuan Lin, et.al, 2019) Analyzed the operation energy consumption through the maximum value principle, gave the method of energy saving through finding the time of "full air breaking"; Gao K et al. (Gao K, et.al, 2013) Designed a distributed controller to solve the control problem of multiple locomotives in the complex terrain and unreliable communication of the heavy-haul combined train.

In addition, some scholars control the train operation by improving PID algorithm, and Chang et al. (Chang C, et.al, 2017) Designed ATO controller by fuzzy differential evolution algorithm to optimize train operation. Hou et al. (Hou Zhongsheng, et.al, 2011) Used model-free adaptive control method to stop the train automatically when entering the station.

Shao (Shao, H, 2016) studied the method based on genetic algorithm (GA). This method has strong robustness, the dynamic and stability characteristics of the system have been greatly improved, and the PID parameters will change with the external interference.

Huang et al. (Huang Y, et.al, 2016) Designed a Back Propagation (BP) neural network to generate driving curve for heavy-haul trains based on genetic algorithm (GA). The reliability and feasibility of the method were verified by comparing with the actual driving curve.

Lu X. et al. (Lu Xiaohong, et.al, 2017) used fuzzy control to track the recommended speed of heavy-haul train and obtained a satisfactory result.

Bonissone et al. (Bonissone P P, et.al, 1996) generated a fuzzy controller to track the speed curve, and used genetic algorithm (GA) to optimize the performance of the fuzzy controller by adjusting the parameters of the fuzzy controller.

Qin et al. (Yufu Qin, et.al, 2014) Designed an error detection estimator for generating error detection residuals, and designed the driving strategy

of the train considering the error mode, which was verified in Da Qin railway.

This paper mainly focuses on the driving strategy of heavy-haul train near the phase separation area and in the downgrade section. First of all, by analyzing the influence of various line conditions on the driving strategy of heavy-duty train, the features that affect the driving strategy are constructed. Then, the training model is constructed by using the SVR method with the features and the electric braking force which is in the manual records. and the model is optimized by adjusting the parameters of SVR.

## 2 DATA PREPROCESSING

### 2.1 Feature Analyze

The information recorded by on-board recording system can be divided into train formation information and train status information.

Train formation information includes: total train weight, number of vehicles, train length, effective load and so on. These data record the information related to the train body.

The train status information includes: speed, kilometer mark, front signal status, distance from front signal, time, pipe pressure, handle position, traction force, electric braking force. These data record the status of the train and operation of the driver.

In the recorded data, the size of electric braking force is taken as the result, and other data are included in the training features.

### 2.2 Feature Supplement

For the purpose of more accurate prediction of the driving strategy of the train, the line features near the running position of the characteristic train are supplemented from the static data of the line.

When the heavy-haul train is running on the downgrade, the slope in front of the train is very important. The slope value 200m, 400m, 600m, ..., 1600m in front of the train at 1600m is included in the training features.

The electric braking force of the train is related to the distance to the phase separation area, so the distance between the train and the nearest phase separation area in the driving process (passing through is positive, not passing through is negative) is also added as a set of training features.

Other information such as curvature of the current position curve of the train, direction of the current

position curve of the train and current speed limit are also used as training features.

## 3 ALGORITHM FLOW

### 3.1 Feature Analyze

When loss function of support vector machine (SVM) is changed to make SVM used in regression analysis, it is called support vector regression (SVR).

The problem of SVR can be expressed as follows: for given training data  $D = [(x_1, y_1), \dots, (x_N, y_N)]$ , we hope to get a regression model  $f$ , to make the predicted value  $f(\mathbf{x})$  as close as possible to the actual value  $y$ .

Because the trained samples are not necessarily linearly separable, they can be mapped to high-dimensional feature space by nonlinear mapping  $\mathbf{x} \rightarrow \phi(\mathbf{x})$ . In this case, the model corresponding to the hyperplane divided in the feature space can be expressed as:

$$f(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x}) + b \quad (1)$$

In formula (1),  $\mathbf{w}$  is the normal vector of the hyperplane, and  $b$  is the displacement term, which determines the distance between the hyperplane and the origin.

Assuming that the most tolerable deviation is  $\varepsilon$ , the problem can be written as follows:

$$\min \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^m l_{\varepsilon}(f(\mathbf{x}_i) - y_i) \quad (2)$$

Where  $C$  is the penalty factor and  $l_{\varepsilon}$  is the insensitive loss function:

$$l_{\varepsilon}(x) = \begin{cases} 0, & |x| < \varepsilon \\ |x| - \varepsilon, & |x| \geq \varepsilon \end{cases} \quad (3)$$

The solution of the original problem can be expressed as follows:

$$f(\mathbf{x}) = \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) \kappa(\phi(\mathbf{x}_i), \phi(\mathbf{x})) + b \quad (4)$$

The calculation method of parameter  $b$  is:

$$b = y_i + \varepsilon - \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) \kappa(\phi(\mathbf{x}_i), \phi(\mathbf{x}_i)) \quad (5)$$

$\kappa(\phi(\mathbf{x}_i), \phi(\mathbf{x}_i))$  is the kernel function. The parameter  $\alpha$  can be got by calculating the average value of the trained samples satisfying condition  $0 < \alpha_i < C$ .

### 3.2 Feature Analyze

In order to accurately evaluate the deviation degree of electric braking force error which is used in this paper. The evaluation indexes of commonly used prediction models are as follows:

1. Mean Average Error (MAE)

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

2. Root Mean Square Error (RMSE)

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

Among the above evaluation indexes,  $n$  is the number of test data;  $y_i$  and  $\hat{y}_i$  are the actual and predicted values of the group  $i$  of test data respectively.

### 3.3 Model Training and Optimizing

---

Algorithm: Training model and optimizing

---

1. Use the original data  $D$  and the line data  $S$  to construct the feature  $X$ , according to the method in Chapter 2.
  2. Extract the control data  $Y$  of heavy-haul train from the original data  $D$ . The extracted dataset is  $[X \mid Y]$ .
- 

---

3. The dataset  $[X \mid Y]$  is divided into  $[X_{train} \mid Y_{train}]$  and  $[X_{test} \mid Y_{test}]$  according to the ratio of 4:1. Specify the parameter penalty factor  $C$  and kernel function  $\kappa(\phi(\mathbf{x}_i), \phi(\mathbf{x}_i))$  in SVR, and then train  $[X_{train} \mid Y_{train}]$ . Suppose the training result is  $Y_{train} = f(X_{train})$ .

4.  $\hat{Y}_{test} = f(X_{test})$  is calculated by using the control model, and then MAE and RMSE of the model are calculated. Repeat step 3 and select the one with the minimum MAE and RMSE as the optimal control model  $f$ . The result  $f$  is the control model.

5. Select some testing data, and compare the predicted value  $\hat{Y}_{test}$  with the real value  $Y_{test}$ .

---

## 4 ALGORITHM FLOW

Firstly, according to the data processing method in Chapter 2, 10426 data of Shuo Huang railway operation records are selected as training original data. The method in Chapter 3 is used for optimization. Bring equation (6) into the kernel function of Chapter 3, and optimize the training model by adjusting  $\gamma$  of equation (6) and penalty factor  $C$  of Chapter 3.

$$\kappa(x_1, x_2) = e^{-\gamma \|x_1 - x_2\|^2} \quad (6)$$

The debugging range of  $\gamma$  in equation (6) and penalty factor  $C$  in Chapter 3 is shown in .

Table 1: SVR debugging scope.

Debugging value	Debugging scope
Coefficient of kernel function $\gamma$	[10-4,10-1]
penalty factor $C$	500, 800, 1000

MAE and RMSE are as follows:

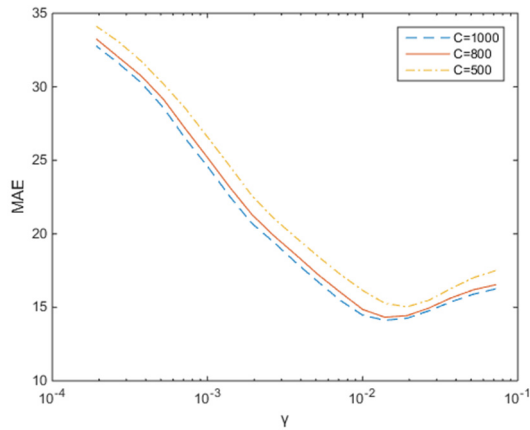


Figure 1: Optimization of the MAE.

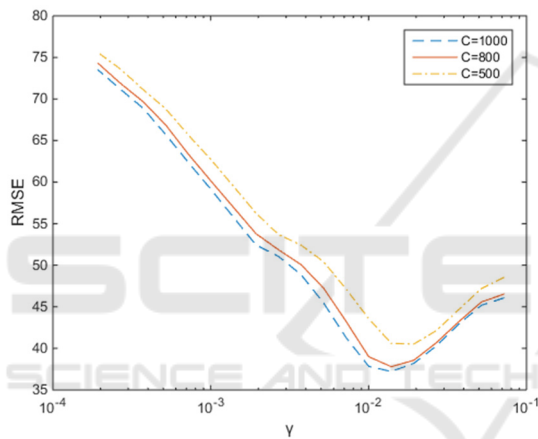


Figure 2: Optimization of the RMSE.

From Figure 1 and Figure 2, when  $C = 1000$  and  $\gamma = 0.0138$ , the absolute mean error (MAE) and root mean square error (RMSE) can be minimized. At this time,  $MAE = 14.1$  (KN),  $RMSE = 37.2$  (KN). Therefore,  $C = 1000$ ,  $\gamma = 0.0138$  is taken as the result of parameter optimization.

After parameter optimization, some manual driving records running near the kilometer mark 140-185 are selected in order to reflect the prediction results intuitively. The output braking force of the control model with the selected parameters is compared with the actual value. The train load 5000 ton with 66 vehicle pulled by SS4B electric locomotive.

The comparison of predicted and actual values is shown in Figure 3. It can be seen from the figure that in the same railway line situations, the driving strategy based on SVR proposed in this paper is close to the manual driving operation.

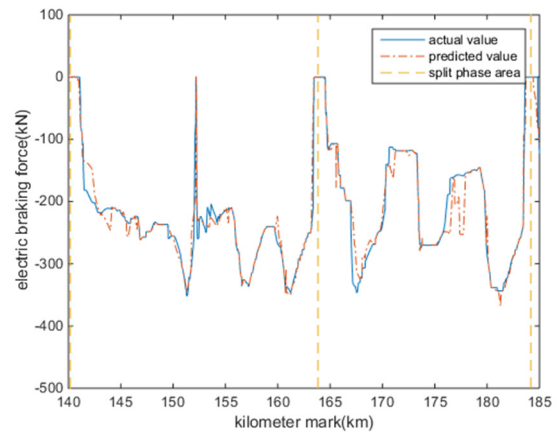


Figure 3: Comparison of predicted the actual values.

## 5 CONCLUSIONS

In this paper, the driving strategy of heavy-haul train in the downgrade section and near the split phase area is studied. The control model of electric braking force is built by training the manual driving records with SVR method. After parameter optimization, the MAE and RMSE of the predicted value and the actual value output are 14.1kN and 37.2kN respectively. Therefore, the control model based on SVR in this paper can be used to reduce the labor intensity of the heavy-haul train drivers with acceptable error in the downgrade section and near the split phase area.

## REFERENCES

- Bonissone P P, Khedkar P S, Chen Y. Genetic algorithms for automated tuning of fuzzy controllers: a transportation application. *Proceedings of IEEE 5th International Fuzzy Systems*. IEEE. New Orleans, LA, 1996.9.11, p. 674-680.
- Chang C, Guo G, Wang J, et al. Study on longitudinal force simulation of heavy-haul train. *Vehicle System Dynamics*. Vol. 55(2017) No.4, p. 571-582.
- Gao K, Huang Z W, Wang J, et al. Decentralized control of heavy-haul trains with input constraints and communication delays. *Control Engineering Practice*. Vol. 21(2013), p. 420-427.
- Hou Zhongsheng, Yi Wang, Chengkun Yin, et al. Terminal iterative learning control based station stop control of a train. *International Journal of Control*. Vol. 84(2011) No. 7, p. 1263-1274.
- Huang Y, Tan L, Chen L, et al. A neural network driving curve generation method for the heavy-haul train. *Advances in Mechanical Engineering*. Vol. 8(2016) No. 5.

- H.Yu,Y.Huang, M.Wang. Research on operating strategy based on particle swarm optimization for heavy haul train on long down slope. 2018 21st International Conference on Intelligent Transport at ion Systems (ITSC). Maui, HI. 2018.11.4, p. 2735-2740.
- Lu Xiaohong, Zheng Muhuo, Lin Hongquan, et al. Research on and Implementation of Intelligent Control Algorithm for Heavy Haul Train. Journal of the China Railway Society. Vol. 39(2017) No. 1 p. 11-18.
- Shao, H. Research on PID control parameter tuning based on genetic algorithm. Wireless Internet Technology. (2016) No.21 p. 111-112.
- Xi Wang, Shukai Li, Tao Tang, et al. Intelligent operation of heavy haul train with data imbalance: A machine learning method. Knowledge-Based Systems. Vol. 163(2018), p. 36-50.
- Xuan Lin, Qingyuan Wang, Qiangqiang Liu, et al. On Periodic Braking of Freight Trains on Long Steep Downhill Section. Journal of the China Railway Society. Vol. 41(2019) No. 01, p. 50-58.
- Yufu Qin, Jun Peng, Xiaoyong Zhang. A robust fault estimation scheme for Heavy-haul trains equipped with ECP brake systems. 26th Chinese Control and Decision Conference (CCDC). Changsha, China, 2014.5.31, p. 2831-2836.

