The Relationship Between Speed and Energy Consumption of High-Speed Train

Jun-hua REN* and Feng LIU
Beijing Jiaotong University, Beijing, China
*Corresponding author

Keywords: High-speed train, Speed, Energy consumption, Self-organization Data Mining.

Abstract. In order to explore the relationship between speed and energy consumption of high-speed train, comprehensively understand the train energy consumption and determine the reasonable operating speed, the energy consumption automatic identification detection system of high-speed train is designed, a data collection method on high-speed train energy consumption is developed. Furthermore, the running data for a Passenger Dedicated train is collected along with the high-speed railway joint debugging and testing. By means of the Self-organization Data Mining, the mathematical model to describe the relationship among energy consumption, running speed and distance is constructed. The running parameters of different distance from the starting point of the high-speed train are analyzed using the mathematical model, and the optimizing driving speed of the low energy consumption is obtained when the high-speed train is at different operating points. The present work provides a common method for analyzing the relationship between speed and energy consumption of the train, and also offers an approach for optimizing the driving strategy of high-speed train.

Introduction

The energy consumption in the high-speed train operating is related to train speed, moving distance, line condition, stop times, distance between stations, operation status and manipulation skill of driver and so on[1,2,3,4]. In ordinary circumstances, the line of high-speed railway is more straight, its curve radius and line slope change small. For specific high-speed railway, the line conditions, running status, stop times and the distance between stations are usually fixed, so their effect on energy consumption of high-speed train can be seen as a constant. Therefore, the present work does not consider the influence of line conditions on the train energy consumption, and the speed and distance are the main factor affecting energy consumption of train. Especially, studying the relationship between speed and energy consumption of high-speed train has an important guiding significance to fully understand the high-speed train energy consumption situation, determine the reasonable operating speed, guide the construction and operations of high-speed rail, protect environmental and so on. BEK B H, ORGENSEN M W J and ORENSON S C S proposed a formula for estimating energy consumption of the high-speed train based on the average speed (technical speed) and the average distance between stops, and the constant in the formula is corrected by the specific line[5]. But BEK B H et al gived the energy consumption calculation model based on the whole line, and the average speed (technical speed) and the distance between stops are calculated based on the mean value of train running in the whole range. Because of the distance difference between stations, train maximum and average velocities are different in the running process, i.e., the state difference of train in the operation between stations is bigger so that there is a big error in actual use the model. Tie-Cheng WANG improved BEK B H et al model using calculating each stop interval, but as the Beijing-Tianjin inter city high-speed rail for an example, the results show that the error is still larger[6]. Xiong KANG studied the simulation and calculation technology of high-speed train traction, and provided useful reference for high-speed railway engineering construction[7], but the guiding significance to the high-speed train actual operation is limited. Rui-fang MOU and Qin-Jie XIAO established energy optimization control model on given interval running time, distance and speed as the constraints of train through
analyzing the running process of high-speed train[8], which provided theoretical guidance for the optimization high-speed train control scheme. Jian-qiang LIU, Yuan-le WEI and Hu HU established the calculation model of minimum energy consumption in the timing constraints of high-speed train based on modern optimal control theory and considering the braking energy feedback[9]. At present, the research on the relationship between energy consumption and speed of high-speed train is still imperfect.

The energy consumption of high-speed train is influenced by many factors, which not only presents a certain rule, but also has certain randomness. To find the rules and quantitative relation, the data acquisition, processing and analysis of a large operating parameters must be carried from massive data. The superposition of various factors impacting high-speed train energy consumption presents a nonlinear. Studying the relationship between speed and energy consumption of high-speed train must use nonlinear method. The self-organizing data mining often can get satisfactory results based on the observation and statistics of the independent and dependent variables without presetting any specific form of parameters and model, and automatically finding the function relationship between the datas by the computer. Therefore, with the help of high-speed train power consumption display system, the high-speed train operation energy consumption data recognition monitoring system is designed, data acquisition and processing technology is chosen. The relationship model running speed and energy consumption of high-speed train is established for the specific line through the self-organizing data mining method, and the analysis is carried, which provides general methods and scientific reference for determining the high-speed train running speed, forming the optimal driving mode and effective saving energy.

**Build Relationship Model between Speed and Energy Consumption of High-speed Train**

**Overview Method of Self-organizing Data Mining**

The basic idea of self-organizing data mining considers that any functional relation between independent and dependent variables can be expressed by Kolmogorov- Gavbor polynomial (1):

$$y = a_0 + \sum_{i=1}^{m} a_i x_i + \sum_{i=1}^{m} \sum_{j=1}^{m} a_{ij} x_i x_j + \sum_{i=1}^{m} \sum_{j=1}^{m} \sum_{k=1}^{m} a_{ijk} x_i x_j x_k + \cdots$$

(1)

As long as the data and calculation are enough, the formula (1) coefficient can be fitted, and get its function trajectory expression. However, completely determining $a_0, a_i, \ldots$parametric values is unrealistic. The number of terms increases sharply with the increasing of the times and the variable number, huge dimension disaster will produce. Model (1) can not be built directly along with the calculation instability.

The self-organizing data mining method proposed by Ukraine scholar Ivakhnenko can solve above problem through the multi-layer self-organizing structure[10,11]. Its basic idea is to construct mathematical models based on the theory of organisms evolution. A series of active neurons are generated by cross combining each input unit of the system, and each neuron has the function of selecting optimal transfer. Then, a number of neurons closest to target variables are selected from the generation neurons. The selected neurons combine strongly to produce new neurons again. Repeating such a dominant genetic, competitive survival and evolution process, the optimal model is built until the new neurons are not better than the previous generation’s.

Firstly, started from input variables $x_1, x_2, \ldots x_m$ (Such as $x_1$—speed, $x_2$—distance, $x_3$—slope, $x_4$—drag coefficient, $x_5$—stop times, etc), the regression equation (2) is calculated for each pair of inputs $x_i, x_j$ and output $y$ (energy consumption):

$$y = a + bx_i + cx_j + dx_i^2 + ex_j^2 + fx_i x_j$$

(2)

The m (m-1)/2 higher order variables will be produced and replace the original m variables $x_1, x_2, \ldots x_m$ to estimate $y$ (output variable). A criterion is used to evaluate each equation after
finding these regression equations from a set of input and output observations, the best regression equations is selected and preserved. A set of (assumed numbers $m$) the best estimate $y$ of the quadratic equation (each estimate only depends on two independent variables) is got. The observation values of the second generation input variables are generated used by each of newly obtained regression equation, and then instead of the original observation value $x_1, x_2, \ldots, x_m$.

And the same method as above, the quadratic regression equation is calculated for these new input variables, the regression equations with $m(m_1-1)/2$ new variables estimating $y$ will be obtained. The optimal variables are choiced, the third generation input variables are selected to replace the second generation, and the quadratic regression equation is constructed by combining the input variables of the third generations. This process is continued until estimation of regression equation begins to decline than the previous equations. The best one in quadratic polynomials of the last generation is picked out after the stopping of successive constructing regression equation process. Then complex Ivakhnenko polynomials (1) will be obtained when reverse algebraic substitution is made.

**Relationship Model between Speed and Energy Consumption of High-speed Train**

The data acquisition is the prerequisite and foundation for analysing relationship between speed and energy consumption of high-speed train. At present, the CRH2, CRH380A, CRH380AL and other types of high-speed train have power consumption graphical display interface in China, but their electric quantity display is a cumulative value, and its display precision is 1 kw·h. Which provides a basic conditions for studying the dynamic relationship between speed and energy consumption of high-speed train. However, the researchers can not control the energy consumption informatin of high-speed train and analyse and process them because the information can not be stored and transmitted automatically. Therefore, a data automatic identification and monitoring system was developed by drawing support from the passenger dedicated line coupling test. The display table was video recorded in whole high-speed train running, and then off line processed. The captured video image data was recognized. The dynamic change data of the energy consumption with the speed and distance was recorded and extracted in the train running process. The data was transformed into a digital form and storied in a corresponding storage file. At the same time, the video frames of energy consumption, speed and distance not correctly identified were stored separately, and then were added to the collected data set after manual reviewing.

As mentioned earlier, although the operating energy consumption of high-seep train is related to train speed, moving distance, line condition, stop times, distance between stations, operation status and manipulation skill of driver and so on, the main factors affecting its energy consumption are moving speed and distance for specific high-speed railway line. Therefore, we focus on speed and distance impact on high-seep train energy consumption. The datas of high-speed train energy consumption in a passenger dedicated line joint debugging and testing are taken as an example. The high-speed train has 8 units, the energy consumption data in running was videoed from 9:11’13” to 9:44’50”one day, moving total time 33’37” and distance 179km. The total traction power consumption was 3007 kw·h, the renewable power 247 kw·h was subtracted, actual electricity consumption was 2760 kw·h. The number of collected and conversion from total data was 1764 items. Because the parameters of automatic identification system was read faster, the accuracy of high-speed train electric meter was low(1 kw·h as unit), and the power consumption change in electric meter was relatively slow, which leaded a lot recorded datas of the automatic identification were same. Therefore, in accordance with the growth of electricity consumption, the array of the maximum speed, distance and corresponding electricity cumulative value were extracted when the electricity display same and speed shows different, the remaining records were deleted. The array of the speed, distance and electricity corresponding maximum power was extracted when the speed displayed same and the electricity shows different, the rest records were deleted. An array with the maximum running distance was extracted when the electric power and speed display same, the remaining records ware deleted. So this way, the 270 different energy consumption records with
speed and distance changing were retained. The traction energy consumption subtracting the renewable’s was the high-speed train actual energy consumption at the time. The actual energy consumption with time gradually accumulated was taken as dependent variable, and the moving speed and distance were taken as the independent variables at the same time. Then iterative relation model (3) was obtained with the aid of Knowledge Mining software tools using the self organizing data mining method [10-11]:

\[
\begin{align*}
E &= 677.4z_{21} - 3133.0z_{21}z_{22} + 1202.0z_{21}^2 + 1892.0z_{22}^2 + 1643 \\
&\quad + 1.035z_{12} - 0.1499z_{12}z_{12}^2 - 0.02956z_{12}^2 \\
&\quad + 0.2912v - 92.26 \\
&\quad + 0.02478s - 0.00004768s^2 - 1.399
\end{align*}
\]

(3)

In which, \(E\) — high-speed train energy consumption (kw \cdot h); \(v\) — high-speed train speed (km/h); \(s\) — moving distance of high-speed train (km); \(z_{11}, z_{12}, z_{21}, z_{22}\) — intermediate simulation units.

Self organizing data mining model shows a good fitting effect, and the average error between the calculated value and the actual observation data was 0.0254. The comparison of model simulation data with the actual measured data shows in Figure 1.

The iteration is carried from backward forward for model (3). The model (4) reflecting the change of energy consumption and moving speed and distance of high-speed train on the line is obtained:

\[
E(s, v) = 0.00755857v^4 + (0.00652344s - 0.00012552s^2 - 9.94733272)v^3 + (4902.936657 - 6.31941148s + 0.01356687s^2 - 0.0000054165s^3 + 0.0000000521101s^4)v^2 + (2037.184958 - 4.79446457s^2 + 0.00336589s^3 - 0.00000323822s^4 - 1072686.649)v + (0.00050287s^4 - 0.52255372s^3 + 556.2882181s^2 - 218539.4374s + 8789669.65).
\]

(4)

Figure 1. Comparison of the self organizing mining simulation datas with the measured datas.

Analysis on the Relationship Model between Speed and Energy Consumption of High-speed Train

Model (4) is somewhat complex, which distance and speed variables are up to biquadratic. It is very difficult analyzing the relationship among speed, distance and energy consumption with it comprehensively. But the change of energy consumption caused by the speed at a certain distance can be analysed. For example, the speed change from the initial point 100km is analysed. For this purpose, order \(s=100\) in the model (4), the single variable function model (5) is got by iteration.
calculation:

\[ E(s, v) = E(100, v) = 0.00755857v^3 - 9.42050872v^2 + 4401.76881v^2 - 913870.7309v + 71133341.37. \]  \hspace{1cm} (5)

The graph of model (5) is shown in Figure 2. It can be seen that the high-speed train energy consumption declines at less 302 km/h and 310km/h~320km/h, and it ascends at 306 km/h~310km/h and over 320km/h. Taking the derivative of model (5) gets formula (6):

\[ E'(100, v) = 0.03023428v^3 - 28.26152616v^2 + 8803.53762v - 913870.7309. \]  \hspace{1cm} (6)

Model (6) has three real roots \( v_1 = 320.4639 \), \( v_2 = 311.2306 \) and \( v_3 = 303.0566 \). Namely \( E' \) equals zero when \( v \) is 320.4639km/h, 311.2306km/h and 303.0566km/h respectively. Taking \( v_1 = 320.4639 \), \( v_2 = 311.2306 \) and \( v_3 = 303.0566 \) into the model (5) calculated respectively obtains \( E(100,320.4639) = 1954.356 \), \( E(100,311.2306) = 2005.094248 \) and \( E(100,303.0566) = 1968.44 \), which minimum value is \( E(100,320.4639) = 1954.356 \). That is to say, about 320.4639km/h \( \approx 320 \) km/h is a optimization speed at 100 km from starting point for the coupling test high-speed train. The graph of Model (6) is shown in Figure 3.

We can use the same method to find the saving energy operating speed when the high-speed train runs at the other point of the line using the model (4). For example, the optimal operating speed equals 331 km/h, 325 km/h, 318 km/h and 316km/h approximately when \( s \) is 50 km, 75 km, 125 km and 150 km from starting point respectively. In other words, The running at speed about 331 km/h, 325 km/h, 318 km/h and 316km/h is all optimal saving energy speed when high-speed train is 50km, 75 km, 125 km and 150 km from the starting point respectively. So the speeds of saving energy in different distance form the optimal speed set, then the set becomes a optimal driving strategy for the train on the high-speed railway line.

In order to calculate the energy consume of the high-speed train with optimal driving speed in the whole section, using same method, calculating optimal driving speed is 315km/h when distance is 179km from the starting point. The whole process energy consumption is \( E(179,315) = 2502.5 \) ( kw\cdot h) if \( S=179km \) and \( v=315km/h \) are substituted into the model (3) or (4). Which saves electricity energy 257.5 kw\cdot h comparing with the actual consumption energy 2760 kw\cdot h.

**Figure 2.** Model (5) function graph.  
**Figure 3.** Model (6) function graph.

**Summary and Conclusions**

In summary, studying the relationship between speed and energy consumption of high-speed train can adopt the following steps:
The data acquisition system is developed with the aid of the high-speed train energy consumption display system, and the data of the energy consumption, speed, and distance of the train are collected.

The collected running data of the energy consumption, speed, and distance of the high-speed train are cleaned. The data of no repetition energy consumption with corresponding velocity and distance are used to establish a relationship model among energy consumption, speed, and distance.

Using cleaned data of energy consumption, speed, and distance, the relationship model among energy consumption, speed, and distance of the high-speed train is established by self-organization data mining method.

The relationship model of energy consumption with speed and distance of the high-speed train is analyzed, and the minimum energy consumption driving speed of different distances is calculated.

The driving speed set of lowest energy consumption of the high-speed train at different distances forms the optimal driving strategy in this section.

The research shows that there is an optimum operating speed making energy consumption lowest for high-speed trains from the point of view of energy consumption. The main purpose is to establish the relationship between running speed and energy consumption of high-speed trains and provide the general analysis method with the help of passenger dedicated line joint debugging and testing. This method can be used for studying each type of high-speed train in order to confirm its relationship between speed and energy consumption. The research provides a scientific guide for determining reasonable driving speed and formulating optimal driving strategy.

But there are no automatic storage & gained system about speed, distance, and energy consumption operation state on high-speed trains in China, which is difficult to get real-time running state data of high-speed trains. The method of real time video and offline processing is used for data acquisition and analysis in this research. With the rapid development of information and communication technology, it is possible that people get a lot of high-speed train operating data and realize effective transmission. It is suggested that high-speed train manufacturing factories should improve high-speed train equipment and come equipped with automatic data storage & gained system in order to provide basic conditions for researchers studying the relationship between energy consumption and speed of high-speed trains.

Acknowledgement

The author would like to thank Beijing Jiaotong University high-speed railway network management engineering research center to support this study by national 863 plan project from the education Ministry, also thank China Railway Science Research Institute cooperation and coordination in high-speed train line joint debugging and testing.

References


