Journal of Transportation Technologies, 2023, 13, 731-745
https://www.scirp.org/journal/jtts
ISSN Online: 2160-0481
ISSN Print: 2160-0473

Establishment and Optimization of Status Assessment Variables for Heavy Haul Railway Line Service Performance

Changfan Zhang[1], Wendong Kong[1], Zhongmei Wang[1*], Lin Jia[1], Shou Chen[2]

[1]College of Railway Transportation, Hunan University of Technology, Zhuzhou, China
[2]Shenzhen Beauty Star Co. Ltd., Shenzhen, China
Email: zcf@hut.edu.cn, kongwd166@gmail.com, *wangzhongmei@hut.edu.cn, jialin@hut.edu.cn, chens@beautystar.cn

How to cite this paper: Zhang, C.F., Kong, W.D., Wang, Z.M., Jia, L. and Chen, S. (2023) Establishment and Optimization of Status Assessment Variables for Heavy Haul Railway Line Service Performance. *Journal of Transportation Technologies*, **13**, 731-745.
https://doi.org/10.4236/jtts.2023.134034

Received: September 16, 2023
Accepted: October 23, 2023
Published: October 26, 2023

Abstract

In order to address the issues of complex system structure and variable selection difficulty for the current heavy haul railway line status evaluation system, a three-category and three-layer heavy-haul line status evaluation variable set construction and reduction optimization method is proposed. Firstly, the status of heavy haul railway line is analyzed, and an initial set of evaluation variables affecting the line status is constructed. Then, based on the association rule and the principal component analysis method, key variables are extracted from the initial variable set to establish the evaluation system. Finally, this method is verified with actual data of a line. The results show that the service performance of heavy haul railway line can still be evaluated accurately when the evaluation variables are reduced by 60% in the proposed method.

Keywords

Set of Variables, Key Variables, Heavy Haul Railway Line, Association Rule, Principal Component Analysis

1. Introduction

Heavy haul railway lines are one of the important channels for transporting bulk goods. Heavy-haul trains are featured with heavy axles, long marshaling, and high transportation efficiency [1], and the heavy haul railway lines are therefore under heavy impact. With the operation of trains year by year, the rate of structural deterioration on the line accelerates, the damages of the line increase significantly, as does the maintenance cost. For the purpose of achieving a high cost effectiveness for the heavy haul railway transportation, in addition to strengthening the line, it is necessary to improve the maintenance method of the line structure. Traditionally, the line is maintained according to fixed periodic schedules [2], by which excessive and insufficient maintenance are both possible. Therefore, in order to reasonably evaluate the condition of the line, it is urgent to replace the "scheduled maintenance" to "status-based maintenance".

In recent years, many scholars have done some research on the condition evaluation of heavy-haul railway lines. The fatigue performance and cumulative damage of the bottom structure of heavy-duty railway tunnels were studied [3], and a nonlinear fatigue cumulative damage model was proposed to evaluate the bottom structure of heavy-duty railway tunnels. Literature [4] carried out condition evaluation and performance prediction for railway bridges by combining experimental modal analysis [5] with actual structural response to load. Literature [6] established a comprehensive evaluation model for subgrade sinkhole based on the collapse filling method, collapse equilibrium method, stability coefficient method, and Procrustes analysis method. The above studies are mainly focused on the analysis with a small number of variables or the analysis of partial structures. It is hard for them to reflect the comprehensive state of heavy haul railway lines. Literature [7] raised two indicators, *i.e.* crack interstitial volume (CIV) and debonding interstitial volume (DIV) based on entropy weight method-fuzzy analytic hierarchy process [8] [9]. It evaluated comprehensively the quality condition of the line through the quality indicators. Literature [10] demonstrated a line geometry status evaluation method by integrating multiple variables of track geometry and train response through analytic hierarchy process and fuzzy comprehensive evaluation. Literature [11] came up with a comprehensive evaluation indicator of the track geometric state based on detected geometric data, and analyzed the sensitivity of the weight of the index parameters by the perturbation method. Literature [12] constructed line quality evaluation indicators by integrating the data from line geometric condition inspection and structural condition inspection. In terms of variable selection and system construction, however, the above researches are weak. It is hard for them to correctly reflect the status of the railway line.

As the variables of heavy haul railway line are mostly from the comprehensive inspection vehicles and manual patrols, it is less efficient to evaluate all variables. Because of the correlation between variables, it is necessary to select key variables that are representative and reflect the state of the heavy haul railway line. Currently, most of the researches at home and abroad mainly focus on the state evaluation method and system construction, while the research on the key variable system for the overall evaluation of heavy haul railway line is few.

To address this issue, and to go deeper on the accurate evaluation of heavy haul railway line service performance, this paper first builds a variable set of heavy haul railway line evaluation system based on the analysis on statuses of the heavy haul railway line. Then, this variable set is refined for optimization based on association analysis and principal component analysis. This process is to solve the problems of variable redundancy and overlapping.

2. Establishment of Status Evaluation Variables for Heavy Haul Railway Line Service Performance

2.1. Status Analysis of Heavy Haul Railway Line

Bearing trains with heavy weight axle, the heavy haul railway line subjects to a larger load, which affects its regularity [13]. The direct results are the changes of the geometric dimensions [14], such as deviated track height and gauge by different degrees, leading to stress concentration on local track structure. Consequently, rail surface fatigue defects would be aggravated [15], such as spalling and scaling [16]. As the connecting part of rails, the fasteners [17] would break, lose or shift under the lateral force of the wheel [18], aggravating the geometric state degradation of the track and causing safety issues; Placed under rails, the sleeper [19] is to support and maintain the geometric shape and position between the rails. Under the impact of the train and the open air environment, the wooden sleepers are prone to aging, corrosion, breakage, etc. As the under-rail structure, the ballast-subgrade [20] is influenced by train operation [21], its geographical location and natural environment [22], therefore prone to damages such as smudging, ballast compaction, and settlements [23] [24]. This will cause the rails to change geometry, leading to a loss of regularity. It can even result in derailment accidents [25]. Furthermore, with an increase in the total mass of the train [26], the quantity and growth rate of rail damages ascend nonlinearly [27]. As a consequence, the rail service life will be reduced. Both the rail status and train operation quality will be affected by those damages. Hence, the status of heavy haul railway line is not the state of a single structure, but the combined condition of various foundation structures.

2.2. Variable Selection of Heavy Haul Railway Line Evaluation

According to the above discussion, the heavy haul line state is mainly composed of three kinds of status: rail geometry status, track structure status and operating status. The rail geometry includes elevation, gauge, alignment, level, warp, curvature, TQI, etc.; the track structure status involves rail system, subgrade ballast system, connecting part system, sleeper system, etc. The main defects of them affecting the rail status concerns the rail surface defects of the rail system, fastener missing of the connecting part system, broken sleepers of the sleeper system, deviated ballast bed thickness and ballast bed settlement of the subgrade ballast system For the operating status, the main indicator is the total mass of passing trains.

2.3. Construction of the Variable Set for Heavy Haul Railway Line Status Evaluation

Considering the requirements of TG/GW102-2019 "Rules for repair of general speed railway lines" of China, "Comprehensive inspection vehicle of Shuo-Huang heavy haul railway line" books and so on, a set of variables is constructed for the status evaluation of heavy haul railway line. It takes 3 kinds of status into account, *i.e.* the rail geometry, the track structure and the operating system. And it has 3 layers, *i.e.* status, system, variable, as shown in **Table 1**.

Table 1. Initial variable set for heavy haul railway line evaluation.

	Status	System	Variables
Status evaluation variables for heavy haul railway line	Rail geometry	Geometry system	Left/right elevation, gauge, left/right rail alignment, level, warp, curvature, superelevation, rail quality indicators, change rate of rail quality variables, change rate of curvature, carbody verticality, horizontal acceleration, curve radius
	Rail structure status	Rail system	Spalling and chipping, top surface scratch, squat damage, corrugation wear, flow of rail head or working edge, rail joint malocclusion on the top or inner side, rail joint gap, rail vertical wear, rail side wear, tread scratch, rail head wear, scaling, rail head horizontal crack, rail head vertical crack, rail head transverse damage, rail-web transition horizontal crack, rail web horizontal crack, rail web oblique crack, rail web vertical split, wire guide hole crack, screw hole transverse/longitudinal crack, rail foot transverse/longitudinal crack, abnormal light band
		Connecting part system	Fastener missing, fastener skewed, backing plate missing, fastener buried; rubber gasket skewed, broken, deformed, severely worn; spike heads falling off, severely rusted
		Sleeper system	Breakage, split, corrosion, sleeper deflection, sleeper chipping, sleeper crack, number of sleepers, sleeper spacing
		Subgrade ballast system	Thickness, ballast bed smudgy, ballast pulverization, ballast collapse, ballast bed smudging, ballast bed deformation, ballast bed compaction, subgrade bed subsidence and settlement, landslide, collapse and rockfall, weathering and spalling, sinkhole, subgrade bed mud pumping, bank erosion, flooded subgrade, poor drainage, frost damage, sand damage, snow damage, debris flow
	Rail operating status	Operating system	Total mass of passing trains

3. Optimization of Status Evaluation Variables for Heavy Haul Railway Line Service Performance

Table 1 demonstrates the whole evaluation system of 71 variables for the heavy haul line. The variables are interrelated and have different degrees of importance. For example, in the rail system, spalling and rail side wear have a great impact on the service performance of heavy haul railway line; while the collapse, debris flow, sand damage and other damages in the subgrade ballast system have little influence. Therefore, in order to reduce the complexity of the entire evaluation system, it is necessary to refine the variables for the purpose of optimization.

3.1. Quantification of Variables by Association Rule

At present, the deduction score system stipulated by industry standards is used to evaluate all variables in China. The basic evaluation variables for heavy haul railway line are mostly derived from inspections. Not only there is no unified standard for such inspection, but also some variables have little impact on the status of heavy haul railway line. Thus, evaluating all variables is not necessary and will reduce the efficiency. In this paper, key variables are selected out for the status evaluation of heavy haul railway line. In order to evaluate the line more reasonably, the association rule is applied. According to the definition of support and confidence in the association rule, the four levels of state evaluation of heavy haul railway line over the years, namely Level I, Level II, Level III and Level IV, are quantified into a set of 4-dimensional arrays to build a quantification matrix.

The concept of association rule is to find an interrelated relationship between things and factors, reflecting the interdependence and relevance between them. It is defined that $I=\{i_1,i_2,i_3,\cdots,i_m\}$ is a finite item set composed of M items to be studied, and $T=\{T_1,T_2,T_3,\cdots,T_N\}$ is the transaction data table, where $T_i=\{i_1,i_2,i_3,\cdots,i_k\}\subset I$, called k-item set. X and Y are subsets that appear in finite itemset and k itemset.

The association rule has two basic measures: support and confidence. Support S is defined as the probability of X and Y appearing simultaneously in a transaction. It is calculated by ratio of transactions with X and Y item appearing simultaneously in the sample data set I to total transactions. It reflects the probability of X and Y appearing simultaneously. The equation is:

$$S(X\to Y)=\frac{|T(X\vee Y)|}{|T|} \tag{1}$$

where: $|T(X\vee Y)|$ indicates transactions contain X and Y at the same time. $|T|$ represents the total number of transactions.

Confidence C of the association rule is used to indicate the dependence of the consequent item Y to the preceding item X. It is calculated as the ratio of Y appearing in the transactions with X item, as follows:

$$C(X\to Y)=\frac{|T(X\vee Y)|}{|T(X)|} \tag{2}$$

where: $\left|T\left(X \vee Y\right)\right|$ means the quantity of transactions with both *X* and *Y*. And $\left|T\left(X\right)\right|$ indicates number of transactions with *X*. Confidence of association rule is a relative variable that measures the accuracy of association rule. The higher this value, the higher the dependent possibility of *Y* to *X*.

Confidence can be used to quantify the correlation between each basic variable and the operation status of heavy haul railway line. The higher the confidence, the stronger the correlation between the variable and the actual operation. Taking the quantification of the initial variables of the rail according to the state statistics over the years as an example, there is:

1) Transaction database I = {Operation and maintenance status of heavy haul railway line}

2) Itemset $X_{i,j}$ = {damage occurrence on the *j*th variable of the *i*th system}

3) Itemset Y_i = {damage occurrence in the *i*th system}

By Equation (1), the support of each variable can be calculated first:

$$S\left(X_{i,j} \to Y_i\right) = P\left(X_{i,j} \cup Y_i\right) = \frac{\sigma\left(X_{i,j} \cup Y_i\right)}{|I|} \times 100\% \tag{3}$$

where, $X_{i,j}$ and Y_i are elements of itemset X and Y. $S\left(X_{i,j} \to Y_i\right)$ is the support of $X_{i,j}$ and Y_i appearing simultaneously. $P\left(X_{i,j} \cup Y_i\right)$ is the conditional possibility of I containing both $X_{i,j}$ and Y_i. $\sigma\left(X_{i,j} \cup Y_i\right)$ is the support count of $X_{i,j}$ and Y_i. Based on support, confidence C of the variable is obtained according to Equations (2) and (3). The calculation formula is as follows:

$$C\left(X_{i,j} \to Y_i\right) = \frac{P\left(X_{i,j} \cup Y_i\right)}{P\left(X_{i,j}\right)} = \frac{\sigma\left(X_{i,j} \cup Y_i\right)/|I|}{\sigma\left(X_{i,j}\right)/|I|} = \frac{\sigma\left(X_{i,j} \cup Y_i\right)}{\sigma\left(X_{i,j}\right)} \times 100\% \tag{4}$$

where, $C\left(X_{i,j} \to Y_i\right)$ indicates the confidence of $X_{i,j}$ and Y_i appearing at the same time. It is an element of the matrix *C*. $P\left(X_{i,j}\right)$ gives the possibility of I containing $X_{i,j}$. $\sigma\left(X_{i,j}\right)$ is the support count of $X_{i,j}$. Its meaning is the same as defined by Equation (3). Taking the first three variables of the rail (spalling and chipping, top surface scratches, squat damage) as an example, the overall operation and maintenance status of the rail is expressed as *I*, the 3 variables are given as $X_{2,1}$, $X_{2,2}$, $X_{2,3}$, and the occurrence of rail damage is recorded as Y_2.

In this paper, the operation and maintenance status data of a certain line from 2015 to 2019 is used.

According to its statistics, there 483 times of damage recordings in total for its Level III state, including 107, 208 and 338 times of spalling, corrugation wear and rail side wear, respectively. While for the rail system state, the recorded damages 234 times, 15, 36 and 103 times for spalling, corrugation wear and rail side wear, respectively.

According to Equation (3), there is $|Y_2| = 234$, $\sigma X_{2,1} = 107$, $\sigma X_{2,2} = 208$, $\sigma X_{2,3} = 338$, $\sigma X_{2,1} \cup Y_2 = 15$, $\sigma X_{2,2} \cup Y_2 = 36$, and $\sigma X_{2,3} \cup Y_2 = 103$.

According to Equation (4), confidence of the variable spalling can be calculated as $C_{2,1}$.

$$C_{2,1}=\frac{P\left(X_{2,1}\cup Y_2\right)}{P\left(X_{2,1}\right)}\times 100\%=\frac{15/483}{107/483}=14\%$$

The confidence values of corrugation wear $C_{2,2}$ and rail side wear $C_{2,3}$ can be obtained in the same way. And the results were 17.31% and 30.18% respectively. This calculation method was then applied also to confidence of other variables of the rail system. Combining the statistic databases of Level I, Level II, Level III, and Level IV, confidence values of rail systems were calculated, as shown in Figure 1.

The variables appeared in Figure 1 are listed out in the following Table 2 with their sequence number.

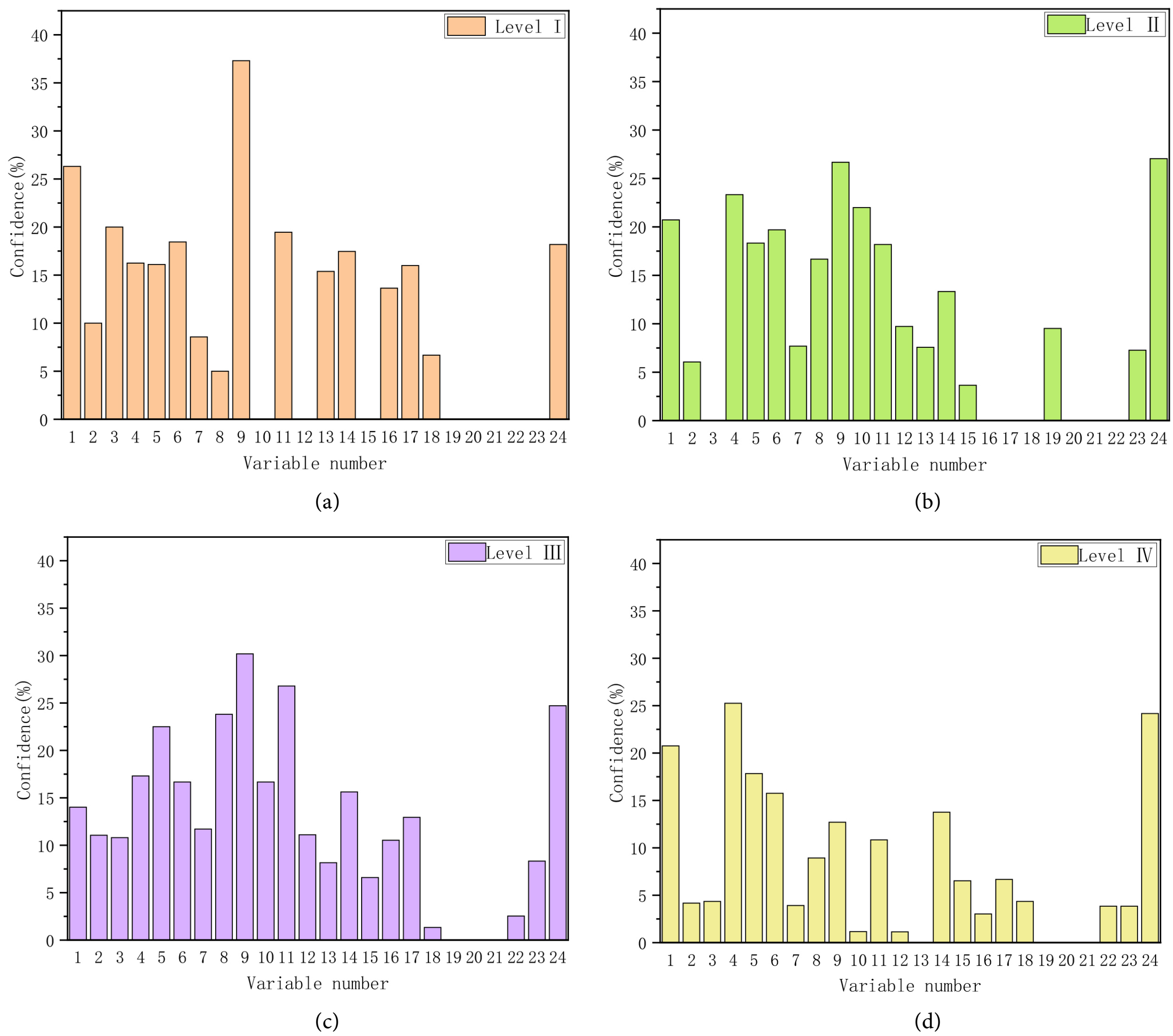

Figure 1. Confidence of rail system. (a) Level I status confidence; (b) Level II status confidence; (c) Level III status confidence; (d) Level IV status confidence.

Table 2. Variables of rail system.

Seq. No.	Variable
1	Spalling and chipping
2	Top surface scratch
3	Squat damage
4	Corrugation wear
5	Flow of rail head or working edge
6	Rail joint malocclusion on the top or inner side
7	Rail joint gap
8	Rail vertical wear
9	Rail side wear
10	Rail head wear
11	Scaling
12	Rail head horizontal crack
13	Rail head vertical crack
14	Rail head transverse damage
15	Rail-web transition horizontal crack
16	Rail web horizontal crack
17	Rail web oblique crack
18	Rail web vertical split
19	Wire guide hole crack
20	Screw hole transverse crack
21	Screw hole longitudinal crack
22	Rail foot transverse crack
23	Rail foot longitudinal crack
24	Abnormal light band

3.2. Extraction of Key Variables by Principal Component Analysis

Too many variables will increase the difficulty and complexity of status evaluation to a certain extent. In order to reasonably extract key variables, principal component analysis is used to extract variables and remove redundant variables, so as to establish key variables for state evaluation. The extracting of key variables by principal component analysis is carried out as follows:

Step 1: Convert confidence values in **Figure 1** into a matrix for standard orthogonalization. A correlation coefficient matrix $R = XX^{\mathrm{T}}$ is constructed and the singular value decomposition is carried out, and then the eigenvalues and eigenvectors of the correlation coefficient matrix are obtained. The eigenvalues are ranked according to their size and expressed $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_p \ge 0$. $\alpha_1, \alpha_2, \cdots, \alpha_p$ is the eigenvector corresponding to the eigenvalue.

Step 2: Determine the number of principal components. The cumulative contribution rate C_{Ri} of the eigenvalue λ_i is defined as:

$$C_{Ri} = \frac{\sum_{k=1}^{i} \lambda_k}{\sum_{k=1}^{p} \lambda_k} \left(i = 1, 2, \cdots, n \right) \tag{5}$$

The eigenvalues with the cumulative contribution rate in 85% - 95% are selected. Provided that the number of them is *m*, it is defined that *m* eigenvalues can represent the information of the original *p* basic variables, *i.e.* $F = \left(F_1, F_2, \cdots, F_m \right)^{\mathrm{T}}$

$$F = AX = \left(\sqrt{\lambda_1}\alpha_1, \sqrt{\lambda_2}\alpha_2, \cdots, \sqrt{\lambda_m}\alpha_m \right)^{\mathrm{T}} \cdot \left(X_1, X_2, \cdots, X_m \right)^{\mathrm{T}} \tag{6}$$

where: $\left(\sqrt{\lambda_1}\alpha_1, \sqrt{\lambda_2}\alpha_2, \cdots \sqrt{\lambda_m}\alpha_m \right)^{\mathrm{T}}$ is the factor loading matrix, $\alpha_1, \alpha_2, \cdots, \alpha_m$, and $\lambda_1, \lambda_2, \cdots, \lambda_m$ is the eigenvector corresponds to eigenvalue.

Step 3: Calculate the comprehensive score of principal components. By weighted summation of principal components whose cumulative contribution rate meets the requirements, the comprehensive score is calculated as follows:

$$\begin{aligned} \hat{F} = \omega F = \omega AX &= \left(\sqrt{\lambda_1}\alpha_1, \sqrt{\lambda_2}\alpha_2, \cdots, \sqrt{\lambda_m}\alpha_m \right)^{\mathrm{T}} \cdot \left(X_1 X_2, \cdots, X_m \right)^{\mathrm{T}} \\ &= \left(\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_m\alpha_m \right)^{\mathrm{T}} \left(X_1, X_2, \cdots, X_p \right)^{\mathrm{T}} \end{aligned} \tag{7}$$

$\omega = \left(\omega_1, \omega_2, \cdots, \omega_m \right)$ is the weight of the principal component in the comprehensible score. The weight vector *H* of basic variables is:

$$H = \left(h_1, h_2, \cdots, h_p \right) = \left(\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_m\alpha_m \right)^{\mathrm{T}} \tag{8}$$

According to the above steps, the confidence data of the rail system was input into MATLAB operation, and the eigenvalues obtained after standardization, orthogonalization and singular value decomposition of the matrix were as follows:

$$\lambda_1 = 3.02, \lambda_2 = 0.498, \lambda_3 = 0.349, \lambda_4 = 0.133, \lambda_5 = 0, \lambda_6 = 0, \cdots, \lambda_{24} = 0$$

According to Equation (6), the contribution rates of the 1st - 4th principal components were 75.495%, 12.449%, 8.733% and 3.324% respectively. The cumulative contribution rate of the 1st and 2nd principal components was 87.944%, which is in the 85% to 95% confidence range required by the algorithm. Therefore, the eigenvector corresponding eigenvalue λ_1, λ_2 can be used to calculate the weight of each variable in the comprehensive score. Nine variables with a weight greater than 0.5 were extracted as the key evaluation variables of rail condition, such as spalling and chipping, flow of rail head or working edge, rail joint malocclusion on the top or inner side, rail vertical wear, rail side wear, scaling, rail head transverse damage, abnormal light band and screw hole transverse crack, as shown in **Figure 2**. The principal component analysis method was then applied to the whole initial variable system, and key variables of each sub-system were extracted. A total of 30 key variables were obtained. Finally, the evaluation system with key variables for heavy haul railway line was established, and shown in **Table 3**.

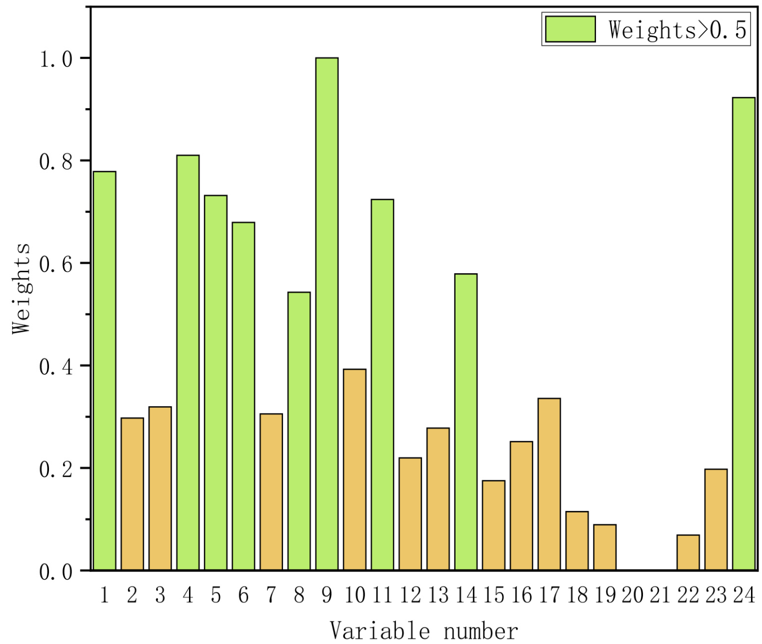

Figure 2. Weights of rail system variables.

Table 3. Key variable set for heavy haul railway line status evaluation.

System	Variable
Geometry system	Left/right elevation, gauge, left/right track alignment, level, warp, change rate of curvature, track quality indicators, change rate of track quality variables
Rail system	Spalling and chipping, flow of rail head or working edge, rail joint malocclusion on the top or inner side, rail vertical wear, rail side wear, scaling, rail head transverse damage, abnormal light band, screw hole transverse crack
Connecting part system	Fastener missing, fastener skewed, fastener buried; rubber gasket skewed
Sleeper system	Sleeper deflection, sleeper chipping, sleeper crack, number of sleepers, sleeper spacing
Subgrade ballast system	Thickness, ballast bed smudgy, ballast bed smudging
Operating system	Total mass of passing trains

4. Case Verification

The proposed method is verified with data of a 200 meter long railway line section. With multiple damages on and under the rail, this status of this section was determined to be Level III by experts. In accordance with TG/GW 102-2019 "Rules for repair of general speed railway lines" of china, the status of heavy haul railway line is to be evaluated with the accumulative deduction score method. Based on the final score, the statuses are to be classified into four levels, *i.e.* Level 1, Level 2, Level 3 and Level 4, as shown in **Table 4**. In order to verify the effectiveness of the proposed key variable set, the scores were obtained and compared to the actual operation condition for the initial variable set and key variable set. Part of the measured data of this line section was normalized and shown in **Table 5**. And the variables and final overall deduction scores are given in **Table 6**.

According to the provisions of the repair rules on the overall evaluation of the line, the overall score Z of the line section can be obtained by:

$$Z = 100 - \sum_{i=1}^{n} w_i S_i \tag{9}$$

where w_i is the weight of each system, and S_i is the deduction score of each system. The deduction score of the rail section was calculated with the deduction score of each system and given in **Table 6**.

According to the evaluation results in **Table 6** and **Table 7**, there is slight difference between the deduction scores of the key variable system and the initial variable system, however, the evaluation levels of are the same, and consistent with the actual level. This means that after reducing the quantity of variables, the key variable system of the heavy haul railway line can still stand for the overall status of each system, and correctly represent the overall status of the line. The proposed method can improve the evaluation efficiency, reduce the evaluation cost, and thus has practical significance.

Table 4. Classification of line status.

Level I	Level II	Level III	Level IV
$Z \geq 90$	$90 > Z \geq 80$	$80 > Z \geq 60$	$Z < 60$

Table 5. Partial data.

Status variable	Value
Elevation	9 mm/3 positions
Gauge	−8 mm/2 places
TQI	15.1
Side wear	3 positions
Abnormal light band	2 positions
Spalling and chipping	4 positions
Rail gap	2 positions

Table 6. Deduction scores of evaluation variables for heavy haul railway line.

Variable		Unit	Deduction criteria	Deducted score	Total deduction score of initial variables	Total deduction score of key variables
Geometry system	Elevation/mm	Unit	8	3	9	9
	Gauge/mm		+8/−6	3	6	6
	TQI		≥15	40	40	40
Total deduction score for geometry system					55	55
Rail system	Rail gap	position	More than three consecutive	3	6	0
	Spalling and chipping		Depth > 2 mm	4	16	16
	Abnormal light band		Length > 3 cm	3	6	6
	Top surface scratch		Depth > 2 mm	3	6	0
	Side wear		>6 mm	4	12	12
Total deduction score of rail system					46	34
Connecting part	Fastener missing	piece	Number of missing	2	8	8
	Spike head falling-off		Falling off	2	4	0
	Rubber gasket skewed		Skew	2	2	2
	Rubber gasket broken		Break off	2	2	0
Total deduction score of connecting part system					16	10
Sleeper system	Break off	piece	Broken	2	2	0
	Sleeper chipping		Chipping off	2	2	2
	Number of sleepers		Less than specified number	2	4	4
	Sleeper gap		Larger than structural gap	2	4	4
Total deduction score of sleeper system					12	10
Subgrade ballast	Thickness	Centimetre	>4/5, 3/5 - 4/5	4	4	4
	Ballast bed smudging	Hole	Ballast bed smudging	2	4	4
	Ballast bed compaction		Compaction	2	6	0
	Ballast bed smudgy	position	Dirty	2	4	4
Total deduction score of subgrade ballast system					18	12
Total mass	Total mass of passing trains	Mass	Passing mass	20	20	20
Total deduction score of operating system					20	20

Table 7. Line overall score and evaluation table.

	Line overall score	Line evaluation	Actual level
Initial variable system	78	Level III	Level III
Key variable system	72	Level III	

5. Conclusions

1) Comprehensive analysis is used to integrate interrelated single variables and comprehensive state quantities, and an indicator set for the overload line status evaluation system is established. The evaluation system for the overload line consists of 3 states, 6 systems, and 30 variables.

2) The use of association rules and principal component analysis to optimize the variable system is more simplified and representative than before, and reduces the complexity of variables on the premise of ensuring the accuracy of the results.

3) The proposed method is preliminarily verified with the data of a certain line section, considering the complex operating environment of the heavy haul lines, the selection of variables and evaluation specifications need to be continuously revised in combination with the actual line in the future. Currently, it can provide valuable reference for the subsequent "status-based maintenance".

Funding Statement

This work was supported by the National Key Research and Development Program of China (2021YFF0501101), the National Natural Science Foundation of China (62173137, 52172403, 62303178).

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Stone, D.H. (1993) Rail Developments and Requirements for Heavy Haul Railways. In: *Rail Quality and Maintenance for Modern Railway Operation*: *International Conference on Rail Quality and Maintenance for Modern Railway Operation Delft*, Springer, Dordrecht, 15-26. https://doi.org/10.1007/978-94-015-8151-6_2

[2] Larsson-Kråik, P.O. (2009) Managing the Wheel-Rail Interface: Railway Infrastructure Maintenance in a Severe Environment: The Swedish Experience. In: Lewis, R. and Olofsson, U., Eds., *Wheel-Rail Interface Handbook*, Woodhead Publishing, Sawston, 633-652. https://doi.org/10.1533/9781845696788.2.634

[3] Liu, C., Wu, B., Li, R., Wang, F. and Tang, Q. (2022) Fatigue Test on Heavy Haul Railway Tunnel Bottom Structure with Base Cavity. *Frontiers in Earth Science*, **10**, Article ID: 870710. https://doi.org/10.3389/feart.2022.870710

[4] Moyo, P. and Tait, R. (2010) Structural Performance Assessment and Fatigue Analysis of a Railway Bridge. *Structure and Infrastructure Engineering*, **6**, 647-660. https://doi.org/10.1080/15732470903068912

[5] Schwarz, B.J. and Richardson, M.H. (1999) Experimental Modal Analysis. *CSI Reliability Week*, **35**, 1-12.

[6] Xiao, Z.Q., Xu, C.Y. and Zheng, N. (2015) Research on Comprehensive Evaluation Model for Heavy Haul Railway Subgrade Sinkhole. *Railway Standard Design*, No. 5, 11-15+23. (in Chinese)

[7] Ren, J.J., Zhang, Q., Zhang, Y.C., Wei, K., Zhang, K.Y., Ye, W.L. and Zhang, Y.

(2023) Evaluation of Slab Track Quality Indices Based on Entropy Weight-Fuzzy Analytic Hierarchy Process. *Engineering Failure Analysis*, **149**, Article ID: 107244. https://doi.org/10.1016/j.engfailanal.2023.107244

[8] de FSM Russo, R. and Camanho, R. (2015) Criteria in AHP: A Systematic Review of Literature. *Procedia Computer Science*, **55**, 1123-1132. https://doi.org/10.1016/j.procs.2015.07.081

[9] Emrouznejad, A. and Ho, W. (2017) Fuzzy Analytic Hierarchy Process. CRC Press, Boca Raton. https://doi.org/10.1201/9781315369884

[10] Ma, S. (2020) Research on Assessment Method of Track Geometry State Based on Vehicle Response. PhD Thesis, Beijing Jiaotong University, Beijing. (In Chinese) https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CDFDLAST2021&filename=1021555247.nh

[11] Shen, J.F., Xu, Y.D., Li, H.F. and Zhong, C.Y. (2015) Weight Sensitivity Analysis of Parameters of Track Geometry Comprehensive Index. *Journal of Tongji University* (*Natural Science*), No. 11, 1709-1714. (In Chinese)

[12] Xu, W.C., Zhong, C.Y., Xu, Y.D. and Shen, J.F. (2017) Research on Track Quality Evaluation Index of High Speed Railway Ballastless Track. *Journal of Shijiazhuang Tiedao University* (*Natural Science Edition*), No. 1, 52-57. (In Chinese)

[13] Naeimi, M., Zakeri, J.A., Esmaeili, M. and Shadfar, M. (2015) Influence of Uneven Rail Irregularities on the Dynamic Response of the Railway Track Using a Three-Dimensional Model of the Vehicle-Track System. *Vehicle System Dynamics*, **53**, 88-111. https://doi.org/10.1080/00423114.2014.998243

[14] Sadeghi, J. (2010) Development of Railway Track Geometry Indexes Based on Statistical Distribution of Geometry Data. *Journal of Transportation Engineering*, **136**, 693-700. https://doi.org/10.1061/(ASCE)0733-947X(2010)136:8(693)

[15] Wang, W.J., Guo, H.M., Du, X., Guo, J., Liu, Q.Y. and Zhu, M.H. (2013) Investigation on the Damage Mechanism and Prevention of Heavy-Haul Railway Rail. *Engineering Failure Analysis*, **35**, 206-218. https://doi.org/10.1016/j.engfailanal.2013.01.033

[16] Yue, B., Wang, Y., Min, Y., Zhang, Z., Wang, W. and Yong, J. (2019) Rail Surface Defect Recognition Method Based on AdaBoost Multi-Classifier Combination. 2019 *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference* (*APSIPA ASC*), Lanzhou, 18-21 November 2019, 391-396. https://doi.org/10.1109/APSIPAASC47483.2019.9023096

[17] Liu, J., Huang, Y., Zou, Q., Tian, M., Wang, S., Zhao, X. and Ren, S. (2019) Learning Visual Similarity for Inspecting Defective Railway Fasteners. *IEEE Sensors Journal*, **19**, 6844-6857. https://doi.org/10.1109/JSEN.2019.2911015

[18] Dai, X., Peng, Y., Wang, K.C., Yang, E., Li, J.Q. and Ding, S. (2017) Railway Fastener Detection Method Based on 3D Images. In: *First International Conference on Rail Transportation* 2017, American Society of Civil Engineers, Reston, 938-946. https://doi.org/10.1061/9780784481257.095

[19] Abadi, T., Pen, L.L., Zervos, A. and Powrie, W. (2019) Effect of Sleeper Interventions on Railway Track Performance. *Journal of Geotechnical and Geoenvironmental Engineering*, **145**, Article ID: 04019009. https://doi.org/10.1061/(ASCE)GT.1943-5606.0002022

[20] Giunta, M., Bressi, S. and D'Angelo, G. (2018) Life Cycle Cost Assessment of Bitumen Stabilised Ballast: A Novel Maintenance Strategy for Railway Track-Bed. *Construction and Building Materials*, **172**, 751-759. https://doi.org/10.1016/j.conbuildmat.2018.04.020

[21] Li, D. and Selig, E.T. (1995) Evaluation of Railway Subgrade Problems. *Transportation Research Record*, **1489**, 17.

[22] Ngamkhanong, C., Goto, K. and Kaewunruen, S. (2022) Rail Infrastructure Systems and Hazards. In: Calçada, R. and Kaewunruen, S., Eds., *Rail Infrastructure Resilience*, Elsevier, Amsterdam, 97-109. https://doi.org/10.1016/B978-0-12-821042-0.00010-1

[23] Yavna, V., Shapovalov, V., Okost, M., Morozov, A., Ermolov, Y. and Kochur, A. (2022) Modeling of Long-Term Train Loads Impacts on Subgrade Soils: A Review. *International Journal of Transportation Science and Technology*, **12**, 729-752. https://doi.org/10.1016/j.ijtst.2022.06.005

[24] Li, Y., Liu, H., Wang, S., Jiang, B. and Fischer, S. (2023) Method of Railway Subgrade Diseases (Defects) Inspection, Based on Ground Penetrating Radar. *Acta Polytechnica Hungarica*, **20**, 199-211. https://doi.org/10.12700/APH.20.1.2023.20.14

[25] Pagaimo, J., Magalhães, H., Costa, J.N. and Ambrosio, J. (2022) Derailment Study of Railway Cargo Vehicles Using a Response Surface Methodology. *Vehicle System Dynamics*, **60**, 309-334. https://doi.org/10.1080/00423114.2020.1815810

[26] Baranovskyi, D., Myamlin, S.S. and Kebal, I. (2022) Increasing the Carrying Capacity of the Solid-Body Rail Freight Car. *Advances in Science and Technology Research Journal*, **16**, 219-225. https://doi.org/10.12913/22998624/149935

[27] Dai, Y.B. (2020) The Influence of Gross Weight of Trains Passing through Beijing-Guangzhou Railway on Rail Damage. *China Railway*, No. 4, 74-80. (In Chinese)