American Journal of Industrial and Business Management, 2019, 9, 2064-2077
https://www.scirp.org/journal/ajibm
ISSN Online: 2164-5175
ISSN Print: 2164-5167

Study on the Macro-Level Risk Assessment and Intelligent Line Selection for Overseas Railway Construction

Jing Lian[1], Jing Jin[2*], Zonghao Li[3]

[1]Department of Engineering Economics, Design Institute of China Railway Academy Co., Ltd, Chengdu, China
[2]China Academy of Railway Sciences, Beijing, China
[3]Department of Civil Engineering, Southwest Jiaotong University, Chengdu, China
Email: *lianjingswjtu@163.com, 273260107@qq.com, 278946257@qq.com

How to cite this paper: Lian, J., Jin, J. and Li, Z.H. (2019) Study on the Macro-Level Risk Assessment and Intelligent Line Selection for Overseas Railway Construction. *American Journal of Industrial and Business Management*, **9**, 2064-2077.
https://doi.org/10.4236/ajibm.2019.911136

Received: September 18, 2019
Accepted: November 26, 2019
Published: November 29, 2019

Abstract

In recent years, China has made overseas railway construction a key investment project. The primary task of overseas railway investment construction is to select railway routes. Taking some sections of the Belt and Road as an example, 15 representative risk indicators have been established based on the survey data. Based on the principal component analysis method, the risk assessment is carried out in 63 countries along the Belt and Road district, and finally the risk scores are sorted, and the reasonable high-speed rail lines are programmed through the ranking of risk scores.

Keywords

Railway Route Selection, Principal Component Analysis Method, Risk Assessment

1. Introduction

Since the establishment of "the Belt and Road" cooperation in 2013, China's commitment to overseas railway investment and construction can not only drive the economic development of China and neighboring countries, but also demonstrate China's economic strength and the development of high-speed rail technology.

The primary task of the construction of overseas high-speed railway is railway route selection. The design of railway route selection is the overall design of a railway line, which directly affects the railway transportation capacity, transportation quality and economic benefits of investment. Because the work load of constructing a new railway (especially the oversea railways) is very large, and the technology is complex and widely involved. Therefore, before planning a railway, in-depth investigation and research, survey and design work must be carried out, and an optimal solution should be selected from several comparable solutions in the end. The Belt and Road Initiative involves many countries, and each country has different conditions. Therefore, it is necessary to use a unified standard to conduct risk assessments for all countries, and there are many methods for risk assessment. Choosing the appropriate method has a key role in risk assessment [1].

In this paper, the principal component analysis method is finally used for evaluation. The advantages are as follows:

1) The principal component analysis can eliminate the correlation between evaluation indicators Because the principal component analysis forms the principal components that are independent of each other after transforming the original index variables, and the higher the degree of correlation between the indicators is proved, the better the principal component analysis is.

2) The principal component analysis can reduce the workload of indicator selection for other evaluation methods; it is difficult to eliminate the correlation between the evaluation indicators, so it takes a lot of effort to select the indicators. While the principal component analysis can eliminate the related influences, so it is relatively easy to select the indicators.

3) When there are more rating indicators, it is also possible to use a few comprehensive indicators instead of the original indicators for analysis while retaining most of the information. In the principal component analysis, the principal components are arranged in order of variance. When analyzing the problem, some of the principal components can be discarded, and only the principal components with larger pre and post variance are used to represent the original variables, thus reducing the computational workload.

4) In the comprehensive evaluation function, the weight of each principal component is its contribution rate, which reflects the proportion of the information of the primary component of the original data to the total amount of information, so that the determination of the weight is objective and reasonable, and it overcomes the defect of artificially determining the weight in some evaluation methods.

5) The calculation of this method is relatively standardized, which can be easily implemented on a computer, and can be done with specialized software.

2. Macro-Level Risk Indicators

From a macro perspective, the construction of overseas railways is closely related to the political, economic, and social development factors of each country. Therefore, when considering the risk assessment indicators for overseas railway line selection, the principles of data availability and authority are considered. From the World Bank and the National Bureau of Statistics of China, three general indicators are selected here, namely, political, economic, and social development-related specific factors to reflect the specific situation of each country [2].

In consultation with Dr. Tong Xinhao, Dr. Zeng Hailin and other experts (both professors in the railway industry from Southwest Jiaotong University) and based on the actual situation of countries along the railway, the macro-level risk assessment indicators are comprehensively selected of data from various authoritative databases on railway line selection and data on foreign project contracting and import and export in China, which are shown in Table 1.

3. Macro-Level Risk Assessment and Route Selection

3.1. Macro-Level Risk Assessment Principle [3] [4]

Principal component analysis is a multivariate statistical technique that transforms a set of possible correlation variables into a set of linearly uncorrelated variables by orthogonal transformation. The converted set of variables is called the principal component. The basic idea is to reduce the dimensionality of the original variable data to obtain several principal component integrated variables that are not related to each other instead of a large number of original variables, and these integrated variables carry most of the information in the original variables [3]. The first comprehensive variable selected is denoted as F_1, and F_1 has the largest $\mathrm{Var}(F_1)$, which means that F_1 contains the largest amount of information, and F_1 is called the first principal component. If the first principal component is insufficient to represent the information carried by the original p variables, then the second principal component F_2 is selected, and F_2 is independent of F_1 linear, and the mathematical language expression requires $\mathrm{Cov}(F_1, F_2) = 0$. By analogy, the third principal component and the fourth principal component can be constructed up to the No. p principal component.

This article uses related software to perform principal component analysis. The main steps are as follows:

1) Normalizing raw data.

In order to make the indicators comparable, the first thing is eliminating the different dimensions of each indicator and standardizing the indicators to obtain standardized data. This standardized process is actually doing the following transformation on the raw data X:

$$ZX_{ij} = \frac{x_{ij} - \overline{x_j}}{\sqrt{Var(x_j)}} \left(i = 1, 2, 3, \cdots, n, j = 1, 2, \cdots, p\right)$$

in which:

$$\overline{x_j} = \frac{\sum_{i=1}^{n} x_{ij}}{n}, \quad \sqrt{Var(x_j)} = \frac{\sum_{i=1}^{n}\left(x_{ij} - \overline{x_j}\right)^2}{n-1}$$

2) Calculating the correlation coefficient matrix of a normalized data matrix.

3) Typing the output result can directly obtain the eigenvalue and the corresponding eigenvector.

Table 1. Macro-level risk assessment indicators for overseas railway construction.

Number	Indicators	Indicator Description
1	Annual GDP growth	The annual growth rate of GDP is based on market prices based on local currency.
2	Inflation rate	The inflation rate, measured by the consumer price index, reflects the annual percentage change in the average consumer cost of purchasing a basket of goods and services.
3	China's foreign contracted project completed turnover	The amount of work performed by Chinese enterprises or other units in contracting overseas construction projects in the form of money completed during the reporting period.
4	The changes of exchange rate	The rate of change in the average annual exchange rate of the official exchange rate of each country compared to the average exchange rate of the previous year.
5	Power coverage	The percentage of the electricity coverage, that is, the percentage of the population with electricity supply as a percentage of the total population.
6	The difficult degree of company registration	By measuring the company's registration process complexity, the registration process for starting a company includes obtaining the necessary permits, certifications, and other procedures.
7	The rate of traffic accident	The number of deaths caused by road traffic in Shanghai per 100,000 populations.
8	Tax burden	The percentage of total tax out of the commercial profits. The total tax rate refers to the amount of tax and mandatory contributions that the enterprise should pay after deducting the deduction and tax exemption as part of the commercial profit.
9	The population density	The population per square kilometer is the number that the mid-year population divided by the land area (square kilometers).
10	Population growth rate	The t-year population growth rate refers to the medium-term population growth rate from t-1 to t.
11	Total population	Calculated according to the actual number of people, all residents are counted, regardless of their legal status or nationality.
12	Proportion of urban population	The percentage of the urban population out of the total population is collected and collated by the United Nations Population Division.
13	Establishing cooperative relations with China	From low to high, it is divided into diplomatic relations, partners, comprehensive partners, strategic partners, strategic partners, and comprehensive strategic partners, ranging from 1 to 6.
14	Political stability	Political stability measures people's perceptions of the possibility of political instability and politically motivated violence or terrorism.
15	Government efficiency	It reflects the public's perceptions of public service quality, the civil service system and the degree of neutrality, the quality of policy formulation and implementation, and the extent to which the government implements these policies.

4) Calculating the variance contribution rate and the cumulative contribution rate of each principal component.

The contribution rate of the main component F_i:

$$\alpha_i = \frac{\lambda_i}{\sum_{k=1}^{p} \lambda_k} \left(i = 1, 2, \cdots, p\right)$$

Cumulative contribution rate:

$$\beta = \frac{\sum_{k=1}^{m} \lambda_k}{\sum_{k=1}^{p} \lambda_k}$$

In the practical application of the principal component analysis method, the corresponding first, second, …, m-th main components to the $\lambda_1, \lambda_2, \cdots, \lambda_m$ ($m \le p$) are chosen, in which the, $\lambda_1, \lambda_2, \cdots, \lambda_m$ ($m \le p$) of the eigenvalues must be the those that cumulative contribution rate is higher than 60%.

5) Calculating the principal component coefficients and principal component scores.

Let the load matrix be A, then the coefficient of the principal component F_i is the square root of each load matrix divided by the square root of the corresponding principal component variance, and propose the coefficient of F_i is the matrix C_i, then the score of the principal component F_i is:

$$Z_i = C_i \times ZX_i$$

where ZX_i represents the i-th column of the matrix ZX of data normalization.

6) Comprehensive score assessment.

The first m principal components with cumulative contribution rate of 60% are selected, and take the variance contribution rate α_k ($k = 1, 2, \cdots, m$) as the weight to construct a linear combination as a comprehensive evaluation function:

$$R = \alpha_1 Z_1 + \alpha_2 Z_2 + \cdots + \alpha_m Z_m$$

From the above formula, the comprehensive score R of the evaluation can be obtained, and then the magnitude of the R value is calculated according to the value of each data of each evaluation object and these R values are comprehensively sorted, thereby obtaining a comprehensive evaluation of each object to be evaluated [5].

3.2. Instance Application

According to the overseas railway macro-level risk assessment indicators established in Section 2, here are 15 indicators in total, the indicator data are from the World Bank (https://www.worldbank.org/) and the China National Bureau of Statistics. In order to increase the reliability of the data, this paper selects 63 countries (All the Belt and Road project routes and participating countries) as samples, because there is no direct relationship between the risks, and the difference between the dimensions Large, it is necessary to standardize the various risk data, and use the unified standard to judge, so the original data is first standardized, and the standardized data is shown in **Table 2** [6].

The normalized data is used in the dimension reduction factor analysis, and the data is subjected to KMO and Bartley test. If the result of KMO value is greater than 0.6 and the significance of the Bartley test is less than 0.01, principal component analysis or factor analysis can be performed. Because the amount of samples are huge, so on the basis of principal component analysis, the rotation of the factors is actually rotating the factor load matrix, which can simplify the structure of the factor load matrix, so that the square of the element of each column or row in the load matrix is polarized to 0 and 1, through the factor rotation (actually coordinate rotation), it makes each original variable have a close relationship between as few factors as possible, so the actual meaning of the factor solution is easier to explain [7].

Then, the factor analysis tool is used for dimensionality reduction. Based on the principal component analysis, the maximum variance method is used to perform the factor rotation, and the result **Table 3** is obtained.

It can be seen from the above test results that the KMO value is greater than 0.6 and the significance is less than 0.01, so it is suitable for principal component analysis or factor analysis.

As can be seen from **Table 4**, the first five principal components contain nearly 66% of the information, and it can be considered that these five principal components contain most of the information of the original elements.

The load matrix after the rotation of these five principal components is shown in **Table 5**, the coefficient indicating the risk of each component, generally greater than 0.5 - 0.6, is attributed to the component.

The above data was processed to obtain **Table 6**, where the gray shading marks were the portions with coefficients greater than 0.58.

Name each principal component according to the data marked in the below table.

The fifth item (power coverage rate) and the 12th item of urbanization rate in F_1 have large coefficients. These two indicators are related to the level of urban development. Therefore, F_1 is called "the main component of urban modernization level";

The 14th (political stability) and 15th (government efficiency) factors in F_2 are relatively large. Both of these indicators are related to the political situation, so F_2 is called "the main component of the political environment";

The third item in F_3 (the turnover of China's foreign contracted projects) and the 13th (the establishment of cooperative relations with China) have a large coefficient. These two indicators are related to bilateral cooperation, so F_3 is called "bilateral with China". The main component of the partnership;

The coefficient of item 10 (population growth rate) in F_4 is relatively large, and F_4 is called "the main component of population development trend";

Table 2. Standardized data.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Albania	−0.736	−0.522	−0.523	0.149	0.432	−0.682	−0.193	0.14	−0.201	−0.961	−0.29	−0.063	−0.972	0.484	−0.202
Afghanistan	0.534	−0.069	−0.528	−0.378	−1.619	−1.353	−0.152	0.925	−0.254	1.218	−0.107	−1.527	0.713	−2.55	−1.783
United Arab Emirates	0.222	−0.482	0.607	0.507	0.432	−0.615	−0.544	−1.308	−0.197	−0.065	−0.252	1.281	0.152	1.079	1.746
Oman	0.761	−0.621	−0.046	2.29	0.432	−0.615	2.211	−0.714	−0.292	3.251	−0.282	0.903	−1.534	0.972	0.242
Azerbaijan	−0.889	−0.348	−0.513	0.564	0.432	−0.95	−0.907	0.506	−0.19	−0.001	−0.25	−0.194	−0.972	−0.474	−0.567
Egypt	−0.209	1.163	−0.265	0.329	0.424	0.323	−0.58	0.797	−0.212	0.644	0.249	−0.742	1.275	−1.433	−1.066
Estonia	−0.474	−0.623	−0.502	0.156	0.432	−1.151	−0.837	1.206	−0.275	−0.907	−0.299	0.422	−1.534	1.024	1.333
Pakistan	0.37	0.362	−0.14	0.17	0.002	1.664	−0.197	0	−0.062	0.599	0.817	−0.951	0.713	−2.542	−0.99
Bahrain	0.172	−0.34	−0.5	−0.957	0.432	−0.012	−0.772	−1.449	1.477	1.103	−0.299	1.436	−0.972	−0.921	0.659
Republic of Belarus	−1.74	3.974	−0.496	0.12	0.432	−0.548	0.554	1.39	−0.259	−0.766	−0.251	0.857	1.275	0.361	−0.891
Bulgaria	−0.737	−0.814	−0.489	0.568	0.432	−0.012	−0.895	−0.446	−0.24	−1.278	−0.264	0.728	−0.41	0.403	0.205
Bosnia and Herzegovina	−0.82	−0.916	−0.533	1.501	0.432	1.664	0.175	−0.744	−0.237	−1.383	−0.286	−0.903	−1.534	−0.131	−0.643
Poland	−0.464	−0.751	−0.135	−1.554	0.432	−0.682	−0.372	0.542	−0.183	−0.862	−0.081	0.089	1.275	1.196	0.958
the Kingdom of Bhutan	0.776	0.529	−0.537	0.198	0.073	0.323	0.04	0.171	−0.285	0.151	−0.302	−0.958	−1.534	1.271	0.52
East Timor	0.269	0.039	−0.535	−0.27	−3.106	−0.682	0.297	−1.625	−0.223	0.766	−0.3	−1.238	−0.41	−0.011	−1.53
Russian Federation	−1.304	0.817	1.489	0.339	0.432	−0.749	0.595	1.11	−0.297	−0.704	0.548	0.731	1.275	−0.724	−0.373
the Philippines	1.231	−0.335	0.507	0.2	−0.563	3.005	−1.197	0.726	0.032	0.289	0.297	−0.682	0.152	−0.885	0.111
Georgia	0.141	−0.564	−0.516	−1.013	0.43	−1.688	−0.552	−1.238	−0.241	−1.231	−0.285	−0.241	−1.534	−0.232	0.662
Kazakhstan	−0.097	0.681	−0.022	0.419	0.429	−0.347	1.215	−0.293	−0.299	0.172	−0.203	−0.259	1.275	0.085	−0.332
Montenegro	−0.808	−0.597	−0.533	−0.607	0.425	−0.347	−0.331	−0.825	−0.26	−0.788	−0.303	0.254	−0.972	0.612	0.189
Kyrgyzstan	0.351	0.077	−0.411	0.216	0.419	−1.017	0.101	−0.298	−0.275	0.581	−0.272	−1.097	0.152	−0.668	−1.083
Cambodia	1.479	−0.315	−0.441	0.355	−4.111	0.994	0.566	−0.884	−0.219	0.289	−0.215	−1.812	1.275	0.259	−1.036
Czech Republic	−0.775	−0.632	−0.281	0.298	0.432	0.323	−0.972	1.279	−0.17	−0.723	−0.244	0.683	0.152	1.383	1.284
Qatar	0.035	−0.36	−0.315	0.375	0.432	0.323	0.97	−1.618	−0.096	2.308	−0.292	1.935	0.152	1.422	1.204
Kuwait	−0.513	−0.262	−0.239	0.214	0.432	1.463	1.109	−1.503	−0.089	1.935	−0.284	1.892	−0.972	0.289	−0.161
Croatia	−1.444	−0.72	−0.504	0.493	0.432	0.323	−0.703	−0.981	−0.231	−1.304	−0.282	0.013	−0.41	0.948	0.795
Latvia	−0.42	−0.748	−0.502	0.238	0.432	−1.017	0.277	0.189	−0.274	−1.49	−0.295	0.415	−1.534	0.779	1.239
Laos	1.663	−0.172	−0.473	0.144	−1.069	0.323	0.028	−0.525	−0.277	0.105	−0.267	−0.96	1.275	0.617	−0.791
Israel	−0.157	−0.773	−0.261	−0.289	0.432	−0.816	−1.747	−0.431	0.078	0.523	−0.257	1.596	−0.41	−0.884	1.691
India	1.365	0.596	1.161	0.2	−1.069	1.932	0.317	1.674	0.13	−0.01	7.463	−1.238	0.152	−0.945	−0.122
Indonesia	0.686	0.206	0.972	0.239	0.166	1.463	−0.009	−0.166	−0.165	−0.004	1.225	−0.237	1.275	−0.304	−0.222
Jordan	−0.465	−0.451	−0.451	0.021	0.422	−0.012	1.305	−0.324	−0.204	1.886	−0.253	1.192	0.152	−0.368	0.056
Vietnam	0.946	0.083	1.405	0.37	0.419	0.659	1.134	0.499	−0.007	−0.063	0.248	−1.198	1.275	0.417	−0.167
Armenia	−0.08	−0.322	−0.531	0.119	0.429	−1.017	0.126	−0.775	−0.204	−0.596	−0.29	0.191	−0.972	0.058	−0.155
Iraq	1.582	−0.294	0.005	−0.006	0.401	0.726	0.224	−0.123	−0.223	1.342	−0.093	0.515	0.152	−2.172	−1.572

Continued

Iran	−0.8	3.378	0.414	0.206	0.42	0.726	2.697	0.845	−0.257	0.007	0.164	0.7	1.275	−0.879	−0.569
Brunei	−2.132	−0.895	−0.492	0.182	0.432	1.329	−0.813	−1.631	−0.227	0.149	−0.304	0.882	0.152	1.526	1.281
Ukraine	−2.746	2.134	−0.321	−6.352	0.43	−0.28	−0.091	1.258	−0.228	−1.066	−0.039	0.525	0.152	−1.238	−0.756
Singapore	−0.176	−0.598	1.326	−0.13	0.432	−1.353	−1.825	−1.062	7.413	−0.061	−0.274	1.971	−0.41	1.753	2.859
Hungary	−0.696	−0.574	−0.325	−1.019	0.432	−0.347	−0.858	1.098	−0.198	−1.028	−0.249	0.597	−0.972	1.052	0.707
Slovenia	−1.18	−0.715	−0.476	−0.506	0.432	−1.017	−0.911	−0.143	−0.204	−0.765	−0.295	−0.431	−1.534	1.301	1.337
Tajikistan	1.368	0.274	−0.489	0.193	0.415	−0.883	0.407	3.129	−0.245	0.706	−0.256	−1.523	0.152	−0.806	−1.258
Thailand	−0.116	−0.619	1.263	−0.221	0.407	−0.213	2.272	−0.376	−0.173	−0.586	0.1	−0.399	1.275	−0.932	0.37
Turkey	0.791	0.755	−0.015	0.302	0.432	0.122	−1.135	0.578	−0.205	0.285	0.158	0.701	0.152	−1.306	0.351
Nepal	0.017	0.869	−0.497	0.256	−1.005	−0.012	0.305	−0.26	−0.108	−0.016	−0.137	−1.913	−0.41	−0.858	−1.254
Serbia	−1.291	−0.089	−0.523	0.036	0.421	−0.481	−1.258	0.386	−0.225	−1.179	−0.265	−0.15	1.275	0.294	0
Saudi Arabia	−0.07	−0.283	0.989	0.247	0.432	2.134	1.285	−1.333	−0.291	0.912	−0.12	1.166	1.275	−0.259	0.174
Sri Lanka	0.706	0.065	−0.439	0.266	−0.248	0.323	0.354	1.684	0.026	−0.148	−0.182	−1.925	0.152	−0.078	−0.172
Slovakia	−0.457	−0.711	−0.39	0.572	0.432	0.122	−0.776	1.364	−0.194	−0.747	−0.275	−0.242	−0.972	1.304	1.095
Macedonia	−0.519	−0.663	−0.532	0.035	0.432	−0.883	−1.417	−1.645	−0.224	−0.769	−0.295	−0.075	−0.972	−0.038	0.032
Mongolia	1.451	0.846	−0.39	0.146	−1.226	−0.347	0.603	−0.617	−0.304	0.414	−0.289	0.635	1.275	0.983	−0.597
Bangladesh	1.178	0.425	−0.236	0.188	−2.529	0.659	−0.401	0.119	0.92	−0.048	0.65	−1.166	0.713	−1.152	−1.029
Myanmar	1.548	0.284	−0.195	0.629	−3.421	2.201	0.579	0.009	−0.226	−0.19	0.004	−1.174	1.275	−0.846	−1.745
Moldova	−0.105	0.345	−0.534	0.301	0.432	−0.548	−0.397	0.515	−0.183	−0.869	−0.286	−0.653	−0.972	0.096	−0.734
Lebanon	−0.696	−0.539	−0.48	−0.829	0.432	0.323	0.575	−0.206	0.252	2.151	−0.273	1.389	−1.534	−1.585	−0.589
Lithuania	−0.272	−0.704	−0.494	−1.69	0.432	−0.816	0.485	0.718	−0.26	−1.602	−0.29	0.373	−1.534	1.151	1.277
Romania	−0.223	−0.624	−0.428	0.127	0.432	−0.347	−0.821	0.599	−0.22	−1.163	−0.189	−0.197	−0.41	0.424	−0.197
Maldives	0.605	−0.143	−0.532	−0.07	0.413	−0.347	−2.164	−0.209	1.073	0.88	−0.304	−0.63	−0.41	0.535	−0.452
Malaysia	0.595	−0.429	1.868	0.259	0.427	−0.079	1.13	0.486	−0.213	0.301	−0.125	0.762	1.275	0.403	1.263
Korea	−0.365	−0.617	6.127	0.122	0.432	−1.487	0.158	0.015	0.212	−0.481	−0.004	1.135	0.713	0.473	1.451
South Africa	−0.89	0.263	0.779	0.311	−0.842	−0.079	2.133	−0.312	−0.26	0.12	0.021	0.291	1.275	0.13	0.413
Turkmenistan	2.203	−1.189	0.067	0.275	0.432	1.195	0.35	−1.208	−0.294	0.417	−0.274	−0.413	0.152	0.332	−1.452
Uzbekistan	1.819	3.258	0.03	0.444	0.432	−0.816	−0.752	1.316	−0.233	0.348	−0.121	−1.065	0.152	−0.198	−1.014

The coefficient of item 7 (traffic accident rate) in F_5 is relatively large, and F_5 is called "main component of traffic safety index".

In summary, the five main components of national line selection risk are shown in Table 7.

Next, each column of the load matrix is divided by the square root of the variance of the corresponding principal component, and the coefficient of each principal component is obtained, and the matrix is denoted as A; then the variance matrix is normalized, which are the Weights of each principal component, this is regarded as matrix B.

Matrix B:

Table 3. KMO and Bartley test.

KMO	Sampling suitability	0.613
Bartlett's sphericity test	Chi-square obtained last time	304.631
	Degree of freedom	105
	Significant	0.000

Table 4. Total variance interpretation.

NO	Initial eigenvalue			Extracting the sum of squared loads			Sum of squared rotational loads		
	Total	Percentage of variance	Cumulative %	Total	Percentage of variance	Cumulative %	total	Percentage of variance	Cumulative %
1	3.584	23.890	23.890	3.584	23.890	23.890	2.765	18.431	18.431
2	2.002	13.350	37.240	2.002	13.350	37.240	2.106	14.038	32.470
3	1.745	11.634	48.874	1.745	11.634	48.874	1.790	11.933	44.402
4	1.423	9.488	58.362	1.423	9.488	58.362	1.712	11.415	55.817
5	1.145	7.636	65.998	1.145	7.636	65.998	1.527	10.182	65.998
6	0.991	6.610	72.608						
7	0.947	6.310	78.918						
8	0.707	4.715	83.633						
9	0.616	4.104	87.737						
10	0.485	3.234	90.972						
11	0.383	2.555	93.527						
12	0.341	2.273	95.800						
13	0.274	1.827	97.627						
14	0.213	1.421	99.048						
15	0.143	0.952	100.000						

Table 5. Component matrix after the rotation.

Risk number	Components				
	1	2	3	4	5
1	−0.784	0.044	0.141	0.173	−0.052
2	0.042	−0.717	0.212	−0.147	0.211
3	0.127	0.121	0.835	0.031	−0.061
4	−0.418	0.530	0.151	0.183	0.290
5	0.714	0.128	0.018	−0.038	0.004
6	−0.488	−0.028	0.256	0.202	0.172
7	−0.017	−0.143	0.361	0.340	0.680
8	−0.015	−0.386	0.199	−0.666	0.082
9	0.076	0.075	0.196	0.213	−0.817
10	−0.235	−0.176	0.066	0.815	0.053
11	−0.366	−0.134	0.536	−0.253	−0.156
12	0.727	0.168	0.180	0.479	−0.147
13	−0.209	−0.347	0.597	0.041	0.246
14	0.336	0.732	−0.026	−0.136	−0.004
15	0.574	0.608	0.221	−0.036	−0.339

Table 6. Coefficients with coefficients greater than 0.6 in the load matrix.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
F_1	−0.784	0.042	0.127	−0.418	0.714	−0.488	−0.017	−0.015	0.076	−0.235	−0.366	0.727	−0.209	0.336	0.574
F_2	0.044	−0.717	0.121	0.530	0.128	−0.028	−0.143	−0.386	0.075	−0.176	−0.134	0.168	−0.347	0.732	0.608
F_3	0.141	0.212	0.835	0.151	0.018	0.256	0.361	0.199	0.196	0.066	0.536	0.180	0.597	−0.026	0.221
F_4	0.173	−0.147	0.031	0.183	−0.038	0.202	0.340	−0.666	0.213	0.815	−0.253	0.479	0.041	−0.136	−0.036
F_5	−0.052	0.211	−0.061	0.290	0.004	0.172	0.680	0.082	−0.817	0.053	−0.156	−0.147	0.246	−0.004	−0.339

Table 7. 5 main components of national line selection risk.

Main components	Names of the main components
F_1	Main component of urban modernization
F_2	Main component of the political environment
F_3	The main component of bilateral cooperation with China
F_4	Main component of population development trend
F_5	Traffic safety index main component

$$B = \begin{pmatrix} 0.27929 \\ 0.21273 \\ 0.18081 \\ 0.17293 \\ 0.15424 \end{pmatrix}$$

Matrix A:

$$A = \begin{pmatrix} -0.47136 & 0.03028 & 0.10537 & 0.13193 & -0.04228 \\ 0.02501 & -0.49417 & 0.15823 & -0.11215 & 0.17072 \\ 0.07624 & 0.08342 & 0.62446 & 0.02401 & -0.04938 \\ -0.25111 & 0.36531 & 0.11270 & 0.13971 & 0.23499 \\ 0.42956 & 0.08841 & 0.01350 & -0.02911 & 0.00320 \\ -0.29322 & -0.01949 & 0.19150 & 0.15453 & 0.13921 \\ -0.01022 & -0.09820 & 0.27012 & 0.25967 & 0.55058 \\ -0.00918 & -0.26565 & 0.14840 & -0.50938 & 0.06616 \\ 0.04578 & 0.05147 & 0.14674 & 0.16255 & -0.66155 \\ -0.14158 & -0.12123 & 0.04950 & 0.62279 & 0.04328 \\ -0.22008 & -0.09239 & 0.40058 & -0.19332 & -0.12596 \\ 0.43733 & 0.11563 & 0.13456 & 0.36634 & -0.11936 \\ -0.12579 & -0.23893 & 0.44631 & 0.03156 & 0.19908 \\ 0.20211 & 0.50417 & -0.01946 & -0.10390 & -0.00303 \\ 0.34499 & 0.41885 & 0.16485 & -0.02726 & -0.27457 \end{pmatrix}$$

Let the data after standardization be matrix X, then the composite score of the sample countrie: $S = X \times A \times B$.

The results are generally irregular, for example, South Korea is 1.6243, and there is a negative number in the score, which only indicates that the score is lower than the average. Since this score does not intuitively judge the country's risk, it is converted into a familiar percentage system score [8].

Suppose:

$$M = f(S) = \frac{1}{1+\exp(-S)}$$

Then control the score between (0,1), then suppose:

$$R = M \times 60 + 40$$

The score can be converted to a percentile, as shown in Table 8:

Table 8. Risk percentage scores and rankings of each country.

Nations	Score	Ranking	Nation	Score	Ranking
Korea	90.12	1	Albania	70.12	33
Qatar	87.87	2	Macedonia	69.52	34
Oman	85.98	3	Indonesia	69.43	35
Singapore	85.20	4	Belarus	69.12	36
United Arab Emirates	83.99	5	Bosnia	68.82	37
Kuwait	83.41	6	Turkmenistan	68.79	38
Brunei	83.38	7	Georgia	68.54	39
Malaysia	83.36	8	Bhutan	68.18	40
Saudi Arabia	83.17	9	Serbia	68.06	41
South Africa	79.54	10	Romania	66.94	42
Jordan	79.42	11	Turkey	66.52	43
Thailand	77.55	12	Azerbaijan	65.92	44
Latvia	76.13	13	Iraq	64.21	45
Croatia	76.01	14	Kyrgyzstan	63.77	46
Czech Republic	75.95	15	Moldova	63.62	47
Israel	75.52	16	Laos	63.14	48
Russian	74.66	17	Maldives	62.80	49
Iran	74.27	18	Philippines	61.26	50
Bahrain	73.79	19	Egypt	60.43	51
Kazakhstan	73.72	20	Sri Lanka	59.59	52
Bulgaria	73.71	21	East Timor	58.62	53
Lithuania	73.67	22	Pakistan	58.33	54
Poland	73.34	23	Ukraine	58.25	55
Slovenia	73.28	24	Uzbekistan	57.17	56
Estonia	73.09	25	Nepal	57.14	57
Montenegro	72.97	26	Cambodia	56.35	58
Lebanon	72.78	27	Tajikistan	56.08	59
Slovakia	71.97	28	India	54.56	60
Vietnam	71.79	29	Myanmar	54.43	61
Armenia	71.61	30	Bangladesh	53.84	62
Hungary	71.41	31	Afghanistan	52.02	63
Mongolia	70.47	32			

3.3. Line Selection Scheme [9]-[13]

In response to the planned plan for the passage of the Asia-Europe Railway, experts proposed several planning options, of which the two most important programs in the Middle East [9] (**Figure 1** from China Map Network [4]) are:

1) Channel plan one: China-Kyrgyzstan-Tajikistan-Afghanistan-Iran, finally arrived in Germany;

2) Channel plan two: China-Kazakhstan-Kyrgyzstan-Uzbekistan-Turkmenistan-Iran, finally arrived in Germany.

According to the risk score rankings in **Table 8**, it can be seen that in the plan 1, Tajikistan and Afghanistan rank at the bottom, while in the plan 2, although Uzbekistan and Kyrgyzstan are ranked lower. However, compared with the plan 1, the advantages are obvious; the total score of the country of plan 2 is much higher than that of plan 1, so the plan 2 was chosen as the Middle East connectivity scheme.

4. Conclusions

This paper studies the macro-level risk assessment and line selection of overseas railways based on Principal Component Analysis (PCA). Through the analysis of principal component characteristics, the original 15 risk indicators are reduced in dimension reduction, and five principal components that can represent the main risk factors are obtained. The scores of the load weights of the five principal

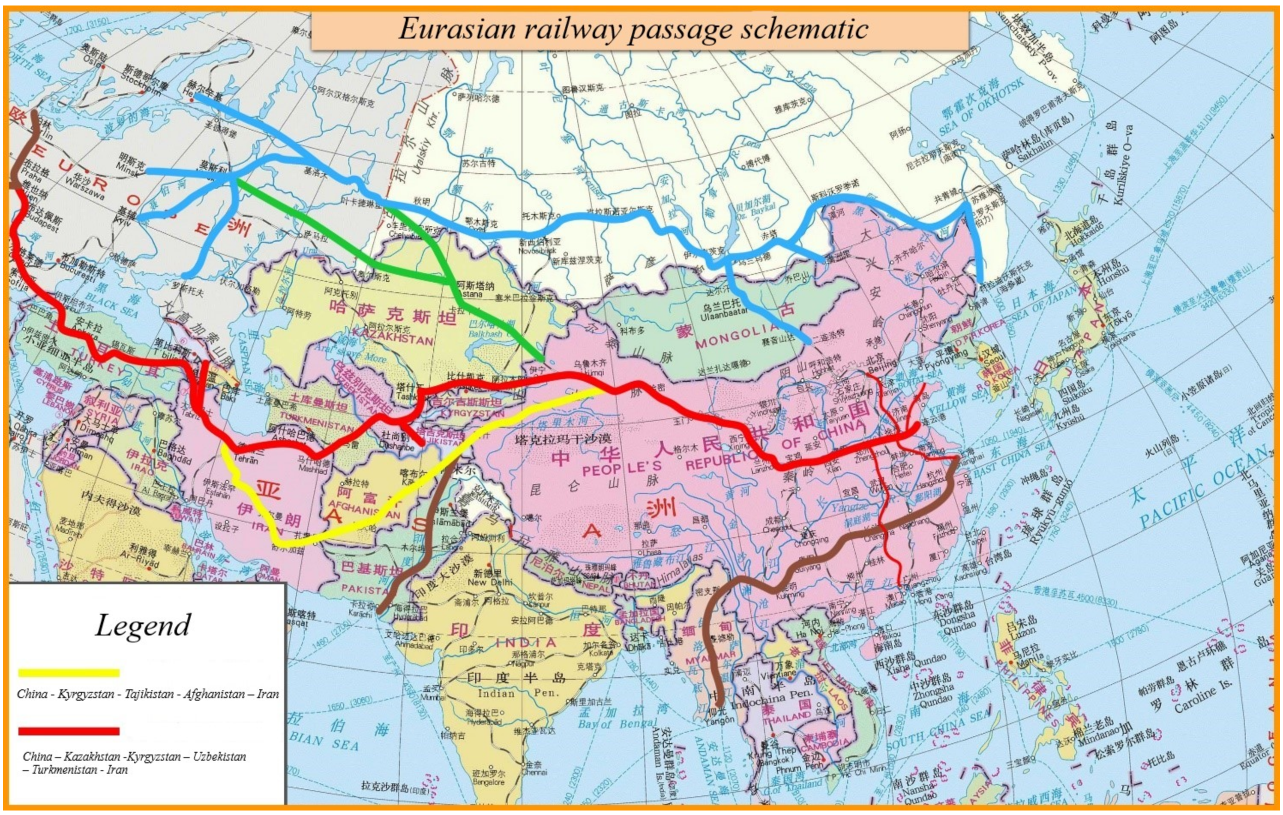

Figure 1. Eurasian railway passage schematic [4].

components are calculated by each country, and the scores are calculated. The risk score is specified, and then through the ranking order, the risk level of the candidate line can be quickly and clearly determined, and the corresponding line selection result is obtained.

The principal component analysis method can be determined from the size of the information sample and the system effect of the sample included in the indicator, avoiding the arbitrariness of the expert scoring and subjective judgment, and the risk rankings of each country can be visually seen through the results. All these indicate that this is a more practical method for railway risk assessment and route selection.

This article lacks a control experiment and will use other risk assessment methods in subsequent studies to compare the results with principal component analysis to further determine the accuracy of the method.

Acknowledgements

This study is supported in part by China Railway Eryuan Engineering Group Co., Ltd Scientific Research Project (NO. KYY2017069 (17-17)); Sichuan Provincial Science and Technology Support Project (NO. 2019JDRC0133); 2017-2019 Young Elite Scientist Sponsorship Program by CAST (YESS); 2018 Sichuan Provincial Ten Thousand Program Project.

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 no conflicts of interest regarding the publication of this paper.

References

[1] Zhu, Y. (2009) Innovation and Practice on Railway Location Concept. *Journal of Railway Engineering Society*, **26**, 1-5.

[2] Liu, W.M. (2002) Research on Main Engineering Economics Problems of Chinese High-Speed Passenger-Only Guided Transport System. Southwest Jiaotong University, Chengdu.

[3] Yang, H. (2013) The Application of Clustering Analysis Based on SPSS in Industry Statistical Data. Jilin University, Changchun.

[4] Jin, J. (2019) Research on Risk Evaluation of International Railway Corridors Construction among Countries along "The Belt and Road". China Academy of Railway Sciences, Beijing.

[5] Jia, R.N. (2015) The Study on the Country Risk Assessment and Control Based on Bank. University of International Business and Economics, Beijing.

[6] Wang, J.P. and Wang, N. (2016) Financial Risk Assessment of China's Energy Listed Companies—Based on Principal Component Analysis. *Friends of Accounting*, **11**,

60-66.

[7] Huang, A., Yang, L.A. and Du, T. (2014) Comprehensive Assessment of Soil Nutrients Based on PCA. *Arid Zone Research*, **31**, 819-825.

[8] Chen, P. (2014) Principal Component Analysis and Its Application in Feature Extraction. Shanxi Normal University, Xi'an.

[9] Cui, Y.P. and Wu, J.Y. (2017) Study on Increasing Operation Quality of China Railway Express by Using Asia-Europe Railway Transpor Corridor. *Railway Transport and Economy*, 39, 68-72.

[10] Yang, C.W., Li, Z.H., Guo X.Y., Yu, W.Y., Jin, J. and Zhu, L. (2019) Application of BP Neural Network Model in Risk Evaluation of Railway Construction. *Complexity*, **2019**, Article ID: 2946158. https://doi.org/10.1155/2019/2946158

[11] Jin, J., Li, Z.H., Zhu, L. and Yang, C.W. (2019) A Research on Risk Assessment of China Railway "Go-Global" Project Construction. *Railway Transport and Economy*, **41**, 86-91.

[12] Jin, J., Zhu, L., Li, Z.H., Tong X.H. and Yang, C.W. (2019) Application of Variable Structure of BPNN in Risk Evaluation of Overseas Railway Construction in Target Countries. *Journal of the China Railway Society*, **40**, 7-12.

[13] Jin J., Li Z.H., Zhu L., Tong X.H. and Yang C.W. (2019) Application of BP Neural Network in Risk Evaluation of Railway Construction. *Journal of Railway Engineering Society*, **36**, 106-112.