Research on the Indicator System of Overseas Interest Security Risks for Energy Enterprises

Hanxiong Xiao *

State Grid Energy Research Institute, Beijing, China

*xiaohanxiong@sgeri.sgcc.com.cn

Abstract. To address the complex safety risks faced by Chinese energy enterprises in overseas projects, this study constructs a literature database, text database, and case database, and identifies 12 core risk categories through text mining and grounded theory. Using R-type clustering and entropy weight methods, a hierarchical risk indicator system covering macro and meso levels is established, with key weights assigned to indicators such as natural environment restrictions and economic stability. The research provides a scientific tool for risk assessment and early warning, supporting enterprises in safeguarding overseas interests and promoting high-quality overseas development.

Keywords: Overseas energy enterprises; Safety risk; Indicator system; Entropy weight method; R-type clustering

1 Introduction

With the deepening of globalization and the accelerating pace of Chinese energy enterprises' "going global" strategy, overseas projects have become an important pillar for enterprises to expand market space and optimize resource allocation[1,2]. To address these issues, this study constructs a comprehensive risk database through literature review, text mining, and case analysis, systematically identifies key risk factors, and establishes a scientific and operable safety risk indicator system. The research results are expected to provide practical decision support for energy enterprises to prevent overseas investment risks and safeguard the security of overseas interests.

2 Identification of Safety Risk Factors in Overseas Projects of Energy Enterprises

2.1 Risk Data Collection and Database Construction

To systematically sort out the investment risks of overseas projects of energy enterprises in both academic and practical fields, the basic information database for such risks established in this section mainly includes three modules: a literature database, a text database, and a case database.

D. Magni et al. (eds.), *Proceedings of the 2026 3rd International Conference on Applied Economics, Management Science and Social Development (AEMSS 2026)*, Advances in Economics, Business and Management Research 389,
https://doi.org/10.2991/978-94-6239-672-2_9

1.Construction of Literature Library. We conducted searches on a large number of foreign journals, with the keywords mainly limited to "Overseas energy projects, Energy investment, Risk management, Energy security, International energy cooperation, Foreign energy projects, Energy sector risks, Risk mitigation strategies". The journals mainly come from Energy Policy, Energy Economics, Renewable & Sustainable Energy Reviews, Applied Energy and Journal of Cleaner Production Waiting for high-quality journals. We imported the abstract set into the database and constructed a structured literature library focusing on the investment risks of overseas energy projects.

2.Construction of Text Library. This section analyzes 180 "Country (Region) Guidelines for Foreign Investment Cooperation" provided on the website of the Ministry of Commerce, and extracts the "Main Risks and Preventive Measures" chapter from each guideline. The search found that the content related to overseas project investment risks in the guide includes: foreign exchange, job requirements, laws and regulations, environmental policies, government changes, labor compensation, etc. By extracting and integrating the above content into the original text data statements, a new txt document is created to form a text library.

3.Construction of the Case Database. Based on information of overseas investment projects by Chinese enterprises disclosed on official websites such as China Energy Network, cases whose titles include nouns such as "overseas", "investment", "risk", "safety", and "energy" and whose full text is related to overseas energy investment projects were retrieved. The full text of relevant cases was extracted to form the case database for risks of overseas energy investment projects.

2.2 Establishment of the Investment Risk List

1.Extraction of High-Frequency Investment Risk Words from the Literature and Text Databases Based on Text Mining. The extraction of high-frequency investment risk words based on text mining mainly relies on data resources from the literature and text databases. Through natural language processing (NLP) technology, a large amount of text data is analyzed to extract high-frequency words related to investment risks, supporting risk identification and assessment. Python was used for document segmentation and high-frequency word statistics to obtain a series of structured high-frequency word data.

2.Extraction of High-Frequency Investment Risk Factors from the Case Database Based on Grounded Theory. Grounded theory was used to fully analyze the cases of overseas energy project investment risks in the case database, identify high-frequency investment risk factors, which were mutually verified and supplemented with the high-frequency investment risk word list obtained from text mining[3]. The specific coding process includes: open coding, axial coding, and selective coding.

Open Coding. Open coding mainly achieves concept classification, clarifies the attributes and dimensions of investment risks, and sorts out the names and categories of risk events. In the conceptualization analysis, if a risk event clearly describes that the risk factor leads to risk results and risk effects, the risk factor will be conceptually extracted. In the categorization analysis, if multiple concepts lead to the same event, the concepts can be merged and categorized.

Axial Coding. The core task of axial coding is to connect various categories through causal analysis. Its main principle is: if various categories lead to the same type of risk results and risk effects, these categories are considered to belong to the same category with high correlation, and then categories with high correlation are merged to form main categories, thereby sorting out the mutual relationships between main categories and various categories. This chapter sorted out 77 concepts and 52 categories, forming 28 main categories.

Selective Coding. Selective coding further sorts out the relationships between categories on the basis of axial coding and identifies core categories among the numerous discovered categories. Among the 28 main categories, through re-integration and refinement of the relationships between corresponding categories, 12 core categories were summarized: political environment, institutional environment, economic environment, socio-cultural environment, natural environment, industry environment, market environment, enterprise environment, project operation, project production, project resources, and project system. Under their combined influence, they exert risk impacts on enterprises' overseas energy project investments. The core categories, main categories, and their sub-categories are shown in Table.

Theoretical Saturation Test. Theoretical saturation can be understood as the sufficiency of a theory. Even if new data is continuously added, no new categories will be formed to increase the density of existing concepts, that is, the "concept density" reaches saturation. This study used 20 cases of enterprises' overseas energy project investments for the theoretical saturation test. Coding of the case data did not reveal new concepts or relationships. It can be seen that the extracted categories are sufficiently rich, thus judging that theoretical saturation has been achieved.

By merging the high-frequency words obtained from the literature and text databases with the risk factors obtained from the case database, a list of investment risks for overseas energy projects was formed in Table 1.

Table 1. List of investment risks for overseas energy projects

Main Risks	Risk Categories	Risk Factors
A Political Environment	A1 Political Stability	A11 Regime Change, A12 Degree of International Political Stability
	A2 Policy Clarity	A21 Policy Change, A22 Policy Transparency
B Institutional Environment	B1 Degree of Rule of Law Construction	B11 Incomplete Legislation
	B2 Administrative Standardization	B21 Administrative Efficiency, B22 Harsh Law Enforcement Against Foreign Investors
	B3 Government Performance Capacity	B31 Government Credit Level
C Economic Environment	C1 Macroeconomic Stability	C11 Host Country's Debt Situation, C12 Inflation, C13 Exchange Rate Fluctuation
	C2 Cross-Border Trade Costs	C21 High Commodity Inspection Fees, C22 Risk of

		Complex Customs Clearance Procedures
	C3 Foreign Financial Protectionism	C31 Trade Protectionism, C32 Tax Risks
D Socio-Cultural Environment	D1 Social Security	D11 Public Health Incidents, D12 Social Security Risks
	D2 Cultural Differences	D21 Cultural Distance
	D3 Infrastructure Completeness	D31 Backward Transportation Infrastructure in the Host Country, D32 Medical Risks
E Natural Environment	E1 Ecological Restriction Degree	E11 Natural Environment Restrictions, E12 Strict Environmental Protection Requirements
F Industry Environment	F1 Industry Development Level	F11 Backward Industry Development, F12 Industry Structure Adjustment
	F2 Industry Order Stability	F21 Malicious Industry Competition, F22 Industry Association Restrictions
G Market Environment	G1 Market Barrier Degree	G11 Market Access Restrictions, G12 Backward Market Development, G13 Public Acceptance, G14 Education Level of Local Residents
	G2 Market Information Changes	G21 Risk of Lack of Market Information
H Enterprise Environment	H1 Enterprise Public Relations Management	H11 Social Responsibility, H12 Public Opinion Risks
	H2 Enterprise Operation Management	H21 Company Equity Structure Risks, H22 Operation Costs, H23 Restricted Sales Channels, H24 Insurance Invalidity Risks
	H3 Human Resource Management	H31 Labor Dispute Risks, H32 Labor Shortage, H33 Employees' Negative Work Attitude, H34 Difficulties in Employee Training
	H4 Supply Chain Management	H41 Risk of Delivery Delays
I Project Operation	I1 Project Management	I11 Decision-Making Error Risks, I12 Project Team Stability, I13 High Project Exit Risks
	I2 Financial Management	I21 Payment Capacity Risks, I22 Financing Difficulties, I23 Capital Chain Breakage Risks
J Project Production	J1 Production Management	J11 High Engineering Acceptance Standards, J12 Product Value Risks
	J2 Construction Management	J21 Human-Induced Accident Risks, J22 Trade Union Influence Risks
K Project Resources	K1 Asset Management	K11 Land Disputes, K12 Equipment Risks, K13 Data Security
	K2 Technology Management	K21 Technology Applicability Risks
L Project System	L1 Legal Compliance Management	L11 Intellectual Property Disputes, L12 Contract Performance Risks

3 Construction of the Index System

Based on the established list of investment risks for overseas energy projects, this chapter initially extracts risk transmission indicators and optimizes the index system to form a reasonable set of risk transmission indicators for overseas energy project investments.

3.1 Methods and Data Sources

R-type clustering method is a clustering analysis method for evaluation indicators. Its core is to group indicators based on their similarity, making the correlation between indicators within groups strong and between groups weak. It is commonly used for dimensionality reduction and optimization of multi indicator systems. Generally, the degree of similarity between variables is quantified by calculating statistical measures such as Pearson correlation coefficient. The Pearson correlation coefficient ranges from -1 to 1, and the closer the absolute value is to 1, the higher the degree of linear correlation between variables; The closer it is to 0, the lower the degree of linear correlation[4].

To ensure the scientificity and rigor of data sources, the original data of investment risk transmission indicators are obtained from authoritative reports and official database data. Taking the data source of political risks in overseas energy project investments as an example, the degree of international political stability is measured by the national-level geopolitical risk index, reflecting the political stability of the host country and its surrounding regions. A higher value indicates a higher risk level. Regime change is measured by data from the World Bank's WGI database, reflecting the frequency of fundamental changes in the ruling regime of a country or region. A lower value indicates more frequent changes in the host country and a higher risk level. Administrative efficiency is measured by "Government effectiveness" in the WGI[5,6].

Since micro risks are mainly related to the projects themselves, only macro risks and meso risks are considered when evaluating the risks faced by overseas investment projects.

3.2 Macro Index System

This paper selects 33 countries and regions around the world as research objects, with a data time span of 2015-2023. For a small amount of missing data, the median is used for imputation. The indicators are normalized, and the correlation coefficients between each pair of the 23 macro indicators are calculated according to the aforementioned theories and models. As is shown in Figure 1, only the correlation between "government performance capacity", "incomplete legislation", and "administrative efficiency" exceeds 0.7. From the perspective of indicator representativeness, indicators are selected by first using the systematic clustering method to group indicators into a certain number of categories, and then selecting typical indicators from each category as the selected indicators. The indicator with a larger average value of correlation indices with other indicators is used as the typical indicator.

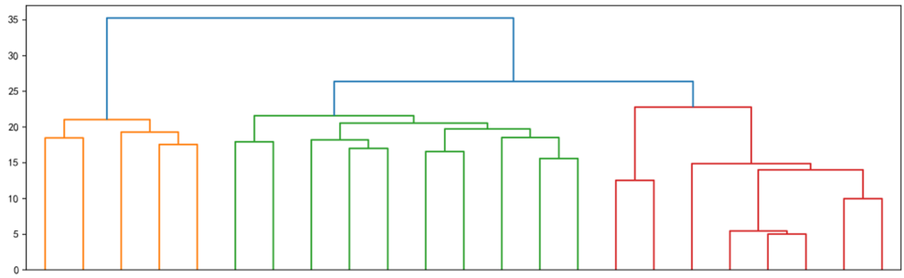

Fig. 1. Dendrogram of cluster analysis between macro indicators

According to the dendrogram of cluster analysis between macro indicators, it is found that the "government performance capacity", "incomplete legislation", and "administrative efficiency" indicators have the highest similarity. Reviewing the data sources of these three indicators, they all belong to the WGI database. These three indicators are now merged and represented by "government governance effectiveness", which is synthesized using the average value of these three indicators.

3.3 Meso Index System

This paper selects 34 foreign enterprises with power cooperation projects with Chinese enterprises around the world as research objects, with a data time span of 2015-2023. For a small amount of missing data, the group mean imputation method is used to process the unreported indicator data in an enterprise's annual report, that is, the average value of the industry in the country is used for imputation. Linear interpolation is used to process indicator data for a single year (the commonly missing year is 2023), because such data has a time series nature and changes year by year. Finally, the mean imputation method is used to process stable values, such as the number of employees and the proportion of the largest shareholder.

The correlation coefficients between each pair of the 20 macro indicators are calculated. The heatmap of the correlation coefficients between indicators is shown in Figure 2.

Analysis of the heatmap of correlation coefficients between original indicators shows that no correlation between indicators exceeds 0.7.

Then, from the perspective of indicator representativeness, indicators are selected by first using the systematic clustering method to group indicators into a certain number of categories, and then selecting typical indicators from each category as the selected indicators. The indicator with a larger average value of correlation indices with other indicators is used as the typical indicator.

In the dendrogram, the horizontal axis represents the names of each indicator, and the vertical axis represents the clustering distance, which measures the degree of similarity between indicators. The smaller the distance, the higher the similarity between indicators. According to the dendrogram of cluster analysis between meso indicators, the correlation between each indicator is within an acceptable range.

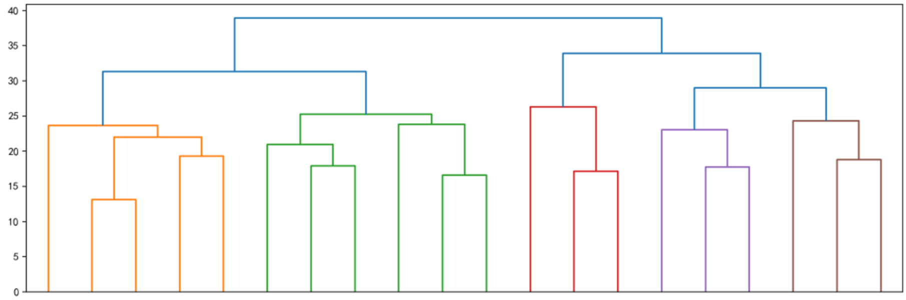

Fig. 2. Dendrogram of cluster analysis between meso indicators

4 Risk Value Calculation and Network Node Quantification

4.1 Steps for Calculating Weights of Overseas Energy Project Investment Risk Indicators

Entropy weighting method is an objective weighting method. The basic principle of entropy weight method is to measure the size of information through the degree of information disorder. The more information conveyed by an indicator, the greater the information it represents in the decision-making process, and therefore the weight value assigned to it should be higher. This method eliminates subjective weighting bias and can fully tap into the value of the data itself. The specific steps are as follows:

1.Construction of the Judgment Matrix.

Assuming that the values of N risk indicators in M years are r_{ij}, the judgment matrix is as follows:

$$R=(r_{ij})_{m\times n}(i=1,2,3\ldots,m)(j=1,2,3\ldots,n)$$

2.Normalization Processing. To improve the comparability of risk indicators, normalization processing is the first step of original data processing. The specific formula is as follows:

$$b_{ij}=\frac{r_{ij}-r_{min}}{r_{max}-r_{min}}$$

3.Calculation of Entropy Weight.

$$H_j=-\frac{\left(\sum_{i=1}^{m} f_{ij}\ln f_{ij}\right)}{\ln m}$$

Among them, $f_{ij}=\frac{b_{ij}}{\sum_{i=1}^{m} b_{ij}}$.

4.Determination of Indicator Entropy Weights.

$$w_j = \frac{(1 - H_j)}{\left(n - \sum_{j=1}^{n} H_j\right)}$$

4.2 Determination of Weights of Overseas Energy Project Investment Risk Indicators

Based on the principle of the entropy weight method, this chapter uses the quantitative values of overseas energy project investment risk indicators from 33 sample countries during 2015-2023 to determine the weights of macro and meso indicators.

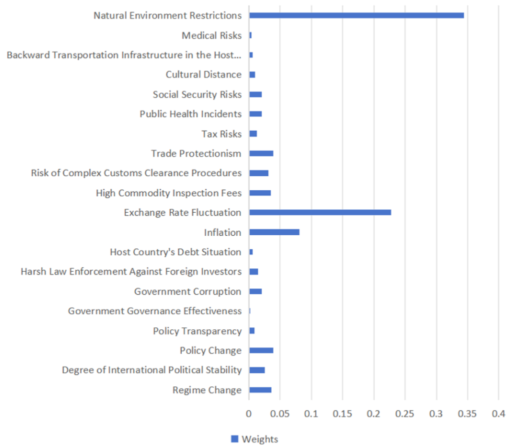

Fig. 3. Weights of macro indicators

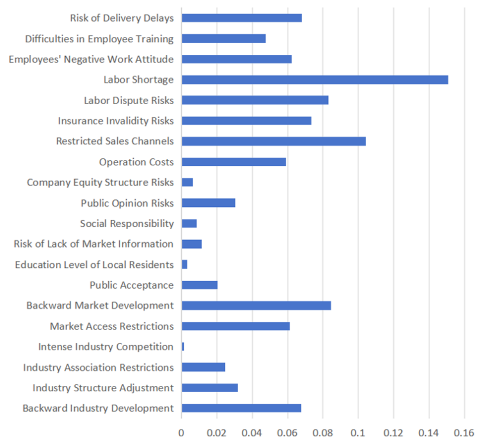

Fig. 4. Weights of meso indicators

1.Weights of Tertiary Indicators. In the original indicator system, some indicators are negative indicators, that is, the larger their values, the lower the risk level reflected. To facilitate the unified processing of all indicators, these negative indicators need to be processed. This study uses the range method to positiveize negative indicators. The weights is shown in Figures 3 and 4.

2.Weight of primary indicators.

As is shown in Table 2, the weight of each primary macro indicator is obtained by summing up the weights of its corresponding tertiary macro indicators.

Table 2. Weights of primary macro indicators

Indicator	Weight
Political Environment	0.110229
Systemetical Environment	0.040802
Economical Environment	0.43109
Social and Cultural Environment	0.060491
Natural Environment	0.363339

As is shown in Table 3, the weight of each primary meso indicator is obtained by summing up the weights of its corresponding tertiary meso indicators.

Table 3. Weights of primary meso indicators

Indicator	Weight
Industry Environment	0.125422
Market Environment	0.180958
Corporation Environment	0.69362

5 Conclusion

This study systematically identified the safety risk factors of overseas energy projects from 12 core dimensions, including politics, economy, and natural environment, and constructed a hierarchical indicator system covering macro and meso levels using R-type clustering and entropy weight methods. The weight calculation results clarify the key risk factors that affect overseas energy investment, such as natural environment limitations, exchange rate fluctuations, and enterprise operation management, and indicate that it is necessary for enterprises to adopt differentiated strategies for different risk factors. This indicator system provides a standardized tool for risk assessment and early warning of overseas energy projects, which can be applied not only at the enterprise level, but also to future research or policy formulation in the energy sector. Future research can further expand the sample size, integrate micro risk indicators, and combine machine learning methods to improve the dynamic adjustability and practicality of the indicator system, better supporting the high-quality development of Chinese energy enterprises' overseas business.

Acknowledgements

This work is sponsored by State Grid Corporation of China's technology project “Research on the Monitoring and Early Warning Method for Corporate Overseas Interest Security Based on the Overall National Security Concept (1400-202457285A-1-1-ZN)”.

References

1. Howell L D, Chaddick B. Models of political risk for foreign investment and trade: An assessment of three approaches[J]. The Columbia Journal of World Business, 1994, 29(3): 70-91
2. Alon I, Gurumoorthy R, Mitchell M C, et al. Managing micro political risk: A cross-sector examination[J]. Thunderbird International Business Review, 2006, 48(5): 623–642.
3. Robock S H. Political risk: Identification and assessment[J]. Columbia Journal of world business, 1971, 6(4): 6-20.
4. Boddewyn, Jean and Crao, F. The Political Game in World Business.[J].Columbia Journal of World Business, 1972,Vol.7,No.1:45-61.
5. GaoYan Q. Political risk distribution of Chinese outward foreign direct investment[J]. International Journal of Emerging Markets, 2021, 16(6): 1202-1227.
6. Busse M, Hefeker C. Political risk, institutions and foreign direct investment[J]. European journal of political economy, 2007, 23(2): 397-415.