ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

Research on the Full Cost Analysis Framework and Risk Assessment Method under the Guidance of Smart Mines

Yanmei He[1,2], *Yongping Tang*[1,2], *Ping Zheng*[3,*] *and Yan Tang*[4]

[1] Changsha Institute Of Mining Research Co., Ltd., Changsha 410012, Hunan, China
[2] National Engineering Research Center for Metal Mining, Changsha 410012, Hunan, China
[3] Department of Natural Resources, Hunan Vocational College of Engineering, Changsha 410151, Hunan, China
[4] Natural Resources Bureau of Ningyuan County, Yongzhou 425600, Hunan, China

SUMMARY: *In response to the problems such as scattered cost sources, complex process coupling, and insufficient dynamic response of traditional accounting methods during the operation of smart mines, this paper proposes a full-cost analysis framework and risk assessment method under the drive of smart mines. The research integrates multi-source heterogeneous data including production scheduling, equipment monitoring, energy measurement, material circulation, safety and environmental protection, and financial accounting, and establishes a mapping relationship of "resource consumption - operation activities - cost objects" to achieve the calculation of costs in all stages of mining, transportation, beneficiation, maintenance, and management. On this basis, by combining long short-term memory networks and random forest algorithms, a cost risk assessment model that takes into account both temporal fluctuation characteristics and static structural characteristics is established, and a low, medium, and high-level risk output mechanism is formed. Experimental results show that the accuracy rate on the test set reaches 88.3%, the recall rate is 84.9%, and the F1 value is 0.854, which is overall superior to a single model. The research shows that the computer simulation technology and multi-source information fusion technology can more accurately identify the cause path and diffusion mode of mining expenditure anomaly, which provides a certain theoretical support for the realization of fine management of intelligent mine.*

Povzetek: Študija razvija celovito ogrodje za analizo polnih stroškov in oceno tveganj v pametnih rudnikih. Združuje večvirove podatke o proizvodnji, opremi, energiji, materialih in financah ter uporablja LSTM in random forest modele. Rezultati kažejo izboljšano zaznavanje anomalij stroškov in napovedovanje tveganj, pri čemer ostajajo izzivi pri prenosu med scenariji, kakovosti podatkov in integraciji v realnem času.

KEYWORDS: *Smart Mine; Total Cost Analysis; Risk Assessment Multi-source data fusion*

1 Introduction

Under the circumstances of continuously increasing resource development intensity, increasingly complex ore body occurrence conditions, and continuously strengthening green and low-carbon constraints, the business focus of mining enterprises has shifted from solely pursuing production expansion to the coordinated optimization of efficiency, cost, and safety.

---

*Carina032000@163.com
https://doi.org/10.65102/is2026275

Especially in deep mining, complex geological environments, and the concurrent operation of large-scale mechanical and electrical systems, the various links such as mining, transportation, beneficiation, ventilation, drainage, filling, and equipment maintenance are coupled with each other, and the cost formation process no longer presents a static financial summary, but exhibits characteristics of multi-source data-driven, significant dynamic fluctuations, and a relatively long risk transmission chain [1-3]. If the traditional approach of combining departmental accounting, post-event statistical analysis, and experience judgment is still used, it can only reveal the cost results, but it is difficult to identify the key disturbing factors in the cost generation process and to respond promptly to potential risks.

In recent years, the accelerated penetration of technologies such as the Internet of Things, industrial internet, edge computing, digital twin, and artificial intelligence in mining scenarios has provided new technical support for the full cost analysis of mines. Relying on sensor networks, production scheduling systems, equipment monitoring platforms, energy management systems, and management databases, mines can continuously collect multi-dimensional information such as equipment load, operation duration, energy consumption level, material consumption, transportation efficiency, geological condition changes, and safety event records, gradually forming a heterogeneous data system covering the production end, management end, and operation end. Using the methods of data cleaning, feature extraction, correlation model construction and intelligent prediction, the cost information originally scattered in each business unit is integrated again, which reflects the internal relationship between workflow, equipment flow, energy consumption flow and dangerous source flow [4-6]. This means that the cost research in mines is no longer suitable for staying at the level of single cost item estimation, but should shift to the construction of an analysis framework covering the entire process, all elements, and the entire cycle.

Although there have been many discussions on mine cost estimation, capital expenditure prediction, supply chain management, and safety risk identification in existing studies, the existing achievements are mostly concentrated on a certain type of cost indicator, a certain stage of project evaluation, or a certain type of accident risk analysis, and the attention to the "cost-operation-risk" integration issue in the smart mine environment is still insufficient [7]. On one hand, most models place more emphasis on the numerical prediction of cost itself, and do not adequately handle the heterogeneity of data sources, time series fluctuations, and business correlations; on the other hand, risk assessment is often separated from cost analysis, and factors such as equipment failure, energy anomaly, production line fluctuations, and geological disturbances are not included in a unified cost risk transmission framework, resulting in limited model interpretability and application value.

Based on this, this paper conducts research on the full cost analysis framework driven by smart mines and its risk assessment methods. Based on the convergence of multiple data sources, a full cost analysis framework covering mining, transportation, beneficiation, auxiliary support, and management and operation links is constructed, and combined with machine learning and risk quantification methods, a cost risk model that takes into account dynamic identification and prediction assessment is established.

2 Relevant Studies

In recent years, the study of mine cost has risen from simple financial management to multiple levels of complexity involving production system, supply chain and intelligent decision-making. Botin et al. first proposed the cost management method with the economic development of mining industry as the goal, and believed that cost control should be carried out with the improvement of productivity, rather than just the post-settlement work [8]. In addition, Ozdemir et al. proposed the problem of how to optimize the overall operation process to achieve cost reduction in the open-pit step production environment [9]. Based on this, Zeng et al. further included mine supply chain management in the research scope, pointing out that there is a significant linkage relationship among procurement, transportation, inventory, and production planning, and that single-cost reduction cannot represent the optimal system cost [10]. These studies provided ideas for total cost analysis, but they mostly relied on static indicator calculations or local business optimization, with insufficient attention to real-time data integration, cross-system correlation calculations, and dynamic fluctuation identification.

With the advancement of smart mine construction, cost analysis has begun to exhibit clear characteristics of dataization and modelization. Majstorovic et al. believe that smart mining and intelligent manufacturing share a consistent underlying logic, with the core lying in equipment interconnection, process perception, and data-driven decision-making [11]. Nobahar et al. started from the digital twin system and pointed out that the operational status of the mining industry can be realized through online reconstruction through virtual-real mapping and continuous updates, providing computational support for cost tracking, equipment scheduling, and resource allocation [12]. Guo et al., Nourali et al., and Zhang et al. respectively used neural networks, support vector regression, and deep optimization algorithms to predict mining capital costs, indicating that artificial intelligence methods have strong adaptability in nonlinear cost estimation [13-15]. However, although the above literatures mainly focus on project capital expenditure or investment measurement, these works do not consider all cost elements in a series of processes from collection and screening to energy consumption, maintenance, security and so on, and do not study in detail how to deal with the integration of different kinds of data (such as industrial network data, scheduling data, financial data).

In terms of risk assessment, existing results mainly focus on safety risks, accident risks, and investment uncertainty analysis. Savolainen discussed the handling of uncertainty in the value assessment of metal mine projects from the perspective of real options [16]; Tubis et al. systematically reviewed mining risk assessment methods, indicating that matrix methods, hierarchical methods, probability models, and expert judgment are still common approaches, but they all have insufficient dynamic responses [17]. Beeche attempted to analyze the risk distribution of underground coal mines using computational modeling methods [18], and Wang proposed a data-driven and model-driven equipment safety investment optimization method [19], and Du combined large language models with Bayesian networks for coal mine accident risk analysis [20]. This indicates that computer methods are changing the way mines identify risks, but most studies still limit risks to safety incidents or single equipment failures, and do not sufficiently explore the transmission mechanism, temporal evolution, and coupling relationship with production and operation status of cost risks.

Table 1: Comparison of Relevant Studies

Reference	Research Object	Main Methods	Focused Indicators	Limitations
Botín et al. [8]	Mine operations management	Cost management models	Economic sustainability, continuous improvement	Emphasizes management logic; lacks real-time data modeling
Ozdemir et al. [9]	Open-pit mining production systems	System-level optimization	Operating costs, production efficiency	Focuses on single-scenario operational optimization; limited coverage
Guo et al. [13]	Open-pit mining projects	ANN prediction	Capital cost	Focused on investment cost; does not cover full-process cost
Nourali et al. [14]	Mining projects	SVR prediction	Accuracy of capital cost estimation	Insufficient support for heterogeneous business data integration
Nobahar et al. [12]	Smart mine operation systems	Digital twin review	Real-time mapping, state prediction	Emphasizes system framework; lacks practical cost analysis methods
Tubis et al. [17]	Mining risk assessment	Systematic review	Risk identification and assessment pathways	Insufficient focus on dynamic cost risks
Wang et al. [19]	Digital twin coal mine enterprises	Data-driven + model-driven	Equipment safety investment optimization	Mainly targets safety investment; no full-cost framework established
Du et al. [20]	Coal mine accident scenarios	Large language model + Bayesian network	Accident risk analysis	Focuses on safety incidents; does not address cost propagation chain

Overall, the existing research has established certain foundations in the areas of mine cost estimation, construction of smart mine systems, and risk analysis. However, there are still significant gaps among these three aspects: cost research often lacks multi-source industrial data support; smart mine research focuses more on system architecture description; and risk research often separates from cost management. Based on this, this paper intends to construct a comprehensive cost analysis framework for the entire process, all elements, and dynamic evolution in the smart mine scenario. It will also combine multi-source data fusion, feature modeling, and risk assessment algorithms to enhance the uniformity and interpretability of cost identification, fluctuation diagnosis, and risk warning.

3 Methods

3.1 Acquisition and Preprocessing of Comprehensive Cost Data for Smart Mines

The data used in this paper is sourced from an integrated operation platform for smart mines, covering business segments such as mining, transportation, mineral processing, ventilation and drainage, equipment maintenance, safety and environmental protection, and comprehensive management. Unlike traditional cost studies that mainly rely on financial monthly reports and manual ledgers, this paper integrates data from the production execution system, equipment monitoring system, energy management system, material management system, and financial accounting system to form a multi-source heterogeneous dataset that corresponds to "operation process - resource consumption - cost result". The data time span is from January 2023 to December 2024. Considering that the original data contains both daily-granularity operating data and hourly-level monitoring sequences, this paper constructs training samples based on the sliding time window on the basis of daily-scale cost accounting. After cleaning, a total of 3120 valid samples were formed, among which daily-granularity indicators are used for cost aggregation and structural analysis, and hourly-level sequences are used to enhance the model's ability to identify short-term disturbances and risk accumulation processes. Finally, 46 core indicators were included, such as output, drilling footage, equipment startup rate, electricity consumption, fuel consumption, explosive consumption, spare parts replacement frequency, labor hours, outsourcing costs, safety investment, and environmental expenditures. To enhance computational stability, this paper also retains some hourly-level monitoring sequences and maps them to the daily-scale cost analysis window through time resampling. In the comprehensive cost perspective, the daily cost of the mine can be expressed as

$$\mathrm{C_t = C_t^m + C_t^e + C_t^l + C_t^r + C_t^s + C_t^o} \tag{1}$$

In the formula,$\mathrm{C_t^m}$ represents the material cost, $\mathrm{C_t^e}$ represents the energy cost,$\mathrm{C_t^l}$ represents the labor cost, $\mathrm{C_t^r}$ represents the maintenance and spare parts cost, $\mathrm{C_t^s}$ represents the safety and environmental protection cost, and $\mathrm{C_t^o}$ represents other management expenses. Considering the limited comparability of costs under different production volumes, this paper further constructs the unit output cost indicator:

$$\mathrm{U_t = \frac{C_t}{Q_t}} \tag{2}$$

Among them, $\mathrm{Q_t}$ represents the original ore processing volume or equivalent output volume on the t-th day. This processing can alleviate the problem that the simple total amount indicator is overly affected by scale fluctuations, making it easier for subsequent models to identify the true source of cost anomalies. Before entering the model, the original data need to undergo missing value repair, anomaly identification, scale unification, and feature construction. For continuous time series data formed by equipment sensors and energy meters, the missing values are completed using linear interpolation method:

$$\mathrm{\hat{x}_t = x_{t_1} + \frac{t - t_1}{t_2 - t_1}(x_{t_2} - x_{t_1})} \tag{3}$$

Among them, $\mathrm{t_1}$ and $\mathrm{t_2}$ represent the nearest valid moments before and after the missing point. For relatively discrete data such as financial entries and material requisitions, the same-cycle consolidation and mean imputation are carried out in accordance with business rules. The identification of outliers is verified by both the 3σ criterion and the business threshold, which not only eliminates noise such as sensor drift and repeated accounting by the system, but also retains the true abnormal fluctuations caused by equipment downtime, sudden increase in energy consumption, or concentrated maintenance, avoiding misjudging valuable risk signals as dirty data.

Due to the significant differences in the units of different indicators, this paper uses Z-Score standardization to process each dimension variable:

$$x_t' = \frac{x_t - \mu}{\sigma} \tag{4}$$

Among them, μ and σ represent the sample mean and standard deviation respectively. Based on this, further differential features, rolling mean features, and intensity features are extracted around the cost formation process, such as the change rate of electricity consumption per day, the consumption of explosives per unit drilling length, the maintenance cost per ton of ore, and the deviation of energy consumption 24 hours after equipment failure, etc., to enhance the model's ability to identify trend changes and local disturbances. At the same time, to ensure that the labels used for model training are consistent with the semantic of on-site management, this paper constructs the initial risk markers based on the deviation degree of unit output cost, the abnormal rate of electricity consumption per ton of ore, the growth amplitude of equipment downtime, and the sudden increase level of maintenance cost, and conducts verification by combining the dispatching logs, equipment maintenance records, and operation analysis ledgers. After multiple rounds of screening, the samples are classified into low, medium, and high risk states. The purpose of such processing is not to simply increase the statistical weight of outliers, but to retain as much as possible the real connection between cost changes and operation behavior so that the results of subsequent models can better meet the alarm needs of mine operation and management.

3.2 Design of the Full Cost Analysis Framework for Smart Mines

In the context of smart mines, cost formation no longer simply amounts to a simple aggregation of financial results. Instead, it is a dynamic process driven by the mining organization, equipment operation, energy consumption, material circulation, personnel allocation, and safety and environmental protection investment. Influenced by the operation rhythm, rock conditions, equipment health status, and scheduling strategies, the same cost item often shows significant fluctuations in different time windows. If only static aggregation based on financial accounts is conducted, although it can reflect the expenditure results, it is difficult to explain where the cost originates from, why it fluctuates, and along which path the risks are transmitted. Based on this, this paper constructs a full cost analysis framework for smart mines, integrating industrial perception data, business management data, and financial accounting data into the same calculation link, achieving "traceable sources, calculable processes, and identifiable anomalies" for costs.

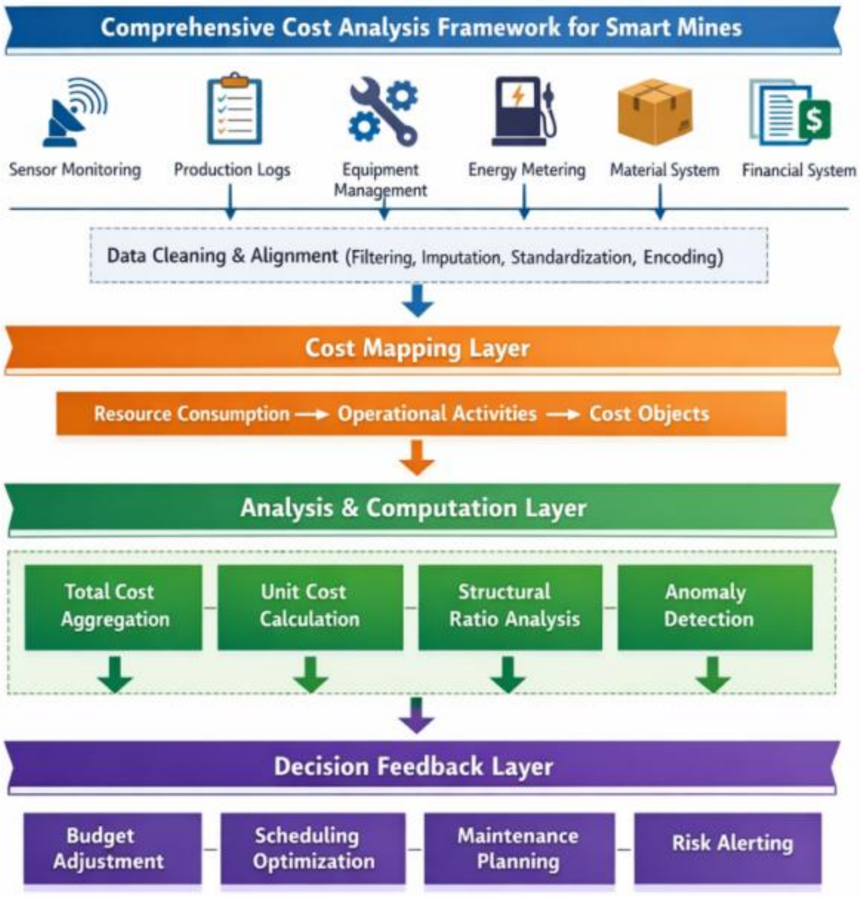

Figure 1: Schematic Diagram of the Design of the Total Cost Analysis Framework

Its framework structure is shown in Figure 1, which includes four parts: data input layer, cost mapping layer, analysis and calculation layer and decision feedback layer. The data input module mainly collects various information from different sources, including environmental monitoring data, production scheduling data, equipment running status data, energy consumption data, material collection data, financial data, etc. The cost mapping module establishes the relationship chain of "resource consumption-activity - cost object" based on work unit, equipment and process; the analysis and calculation layer completes the full cost aggregation, unit cost accounting, structural proportion analysis, and abnormal fluctuation identification; the decision feedback layer returns the analysis results to the scheduling system and the management operation end, used to support budget revision, equipment maintenance scheduling, and resource allocation optimization. The design adopts the concept of computer system hierarchy management and module decomposition, and unitizes the data distributed in different servers into a coherent cost area that can be calculated and tracked. The standardized access nodes, standardized coding convention and timestamp connection mode are used to ensure the corresponding relationship between production activities, resource consumption and financial records on the same evaluation dimension. This provides a more stable source of information for us in the later stage of the problem investigation, accountability and risk review. The total cost of the mine at time t can be expressed as

$$C_t = \sum_{k=1}^{m} c_{k,t} \tag{5}$$

Among them, $c_{k,t}$ represents the expenditure of the k-th type of cost item at time t, which includes costs such as materials, energy, labor, maintenance, safety and environmental protection, as well as management. Considering that costs do not occur in isolation, this paper further characterizes the activity cost as the product of resource consumption quantity and resource unit price:

$$c_{k,t} = \sum_{j=1}^{n} q_{j,t}\, p_{j,t} \tag{6}$$

In this formula, $q_{j,t}$ represents the consumption amount of the jth type of resource at time t, and $p_{j,t}$ represents the corresponding unit price. Through this expression, on-site data such as electricity consumption, fuel consumption, spare part replacement, and labor input can be directly mapped to cost values, thereby establishing a calculation relationship between production operation and financial results.

To enhance the analytical capability of the framework, this paper introduces the structural contribution degree indicator based on cost aggregation to describe the influence degree of each cost item on the total cost:

$$w_{k,t} = \frac{c_{k,t}}{C_t} \tag{7}$$

When the contribution degree of a certain type of cost suddenly rises abnormally within a short period of time, the system can conduct a cause-and-effect analysis by taking into account the equipment status, operation load and environmental parameters. In the decision feedback stage, the system does not merely output an over-limit warning for a certain cost item, but also simultaneously provides the source of the anomaly, related processes and possible transmission paths. For example, when both the tonnage of electricity consumption per unit and the downtime simultaneously increase while the maintenance cost has not significantly risen yet, the system will prioritize pointing the risk towards the decline in equipment efficiency or the mismatch of operating parameters; if the cost proportion of the transportation link and the outsourcing cost expand simultaneously, it is more likely to reflect deviations in transportation organization and resource allocation. It is designed to translate the results of analysis directly into management actions such as drawing adjustments, table checks, and budget changes, so that all cost structures can not only be calculated, but also used.

3.3 Construction of Full Cost Risk Assessment Algorithm Model

The cost risk generated in the process of intelligent operation of mines is not caused by the sudden increase of a single cost item, but by the comprehensive effect of multiple factors such as production disturbance factors, equipment loss, abnormal energy consumption, resource changes and control deviation. Based on this feature, this article constructs a full cost risk assessment model of "time series feature extraction static feature fusion risk grading output" to improve the recognition ability of cost anomaly evolution process. The overall construction idea of the model is shown in Figure 2. The model input consists of two parts: one is dynamic features organized by time windows, including continuous sequences of total cost, unit cost, energy intensity, maintenance cost change rate, downtime, production fluctuations, etc; The other type is relatively stable static features, such as mining conditions, process types, equipment levels, transportation distances, and outsourcing structures. To eliminate dimensional differences in multi-source variables and achieve unified expression, first perform linear mapping on the input vector:

$$z_t = Wx_t + b \tag{8}$$

In the equation, x_t represents the original feature at time t, and z_t represents the unified representation after mapping. W and b are respectively the transformation parameters.

In the dynamic modeling stage, this paper uses the LSTM network to extract the cumulative effect and lag relationship in the cost sequence. Compared with the single-time-point discrimination method, LSTM is more suitable for depicting the change processes with time-memory characteristics such as maintenance lag, energy drift, and cost transmission. Given a time window of length w, the temporal representation output by the model is denoted as

$$h_t = \mathrm{LSTM}(z_{t-w+1}, z_{t-w+2}, \dots, z_t) \tag{9}$$

Among them, h_t represents the dynamic risk characterization at the end of the window. Subsequently, h_t is concatenated with the static feature vector s_t, and input into the random forest classifier to form an integrated discriminative result oriented towards cost risk. Random forest has good stability when dealing with high-dimensional, nonlinear and noisy data, and is also conducive to depicting the interaction between different cost factors. If there are M decision trees in the random forest, the outputs of each tree can be weighted and integrated based on the performance of the validation set, and the cost risk probability can be expressed as

$$p_t = \sum_{m=1}^{M} \omega_m \, g_m([h_t; s_t]), \sum_{m=1}^{M} \omega_m = 1 \tag{10}$$

In the formula, $g_m(\cdot)$ represents the output of the m-th tree, and ω_m is the weight determined based on the performance of the validation set. After this processing, the model not only retains the dynamic perception ability of time series encoding, but also enhances the classification robustness for cost anomalies in complex scenarios. To facilitate direct use by the management end, this paper further maps the output probability to three-level risk states:

$$R_t = \begin{cases} \text{Low risk}, & p_t < \theta_1 \\ \text{Medium risk}, & \theta_1 \le p_t < \theta_2 \\ \text{High risk}, & p_t \ge \theta_2 \end{cases} \tag{11}$$

Among them, θ_1 and θ_2 are determined through optimization using the validation set. Low risk indicates that the cost fluctuations are within a controllable range, medium risk indicates that local links have already deviated, and high risk implies that abnormal transmission has a tendency to spread, requiring triggering of early warnings and intervention. The threshold setting is not directly based on empirical constants, but is optimized based on the balance results of Precision and Recall for different risk levels on the validation set, so as to maintain a relatively stable coverage ability and false alarm control for high-risk samples. During model deployment, LSTM is responsible for rolling reading the dynamic sequence of the recent w time steps, RF receives the fused representation vector and completes rapid discrimination. The above logic process can be integrated into the stope scheduling system, the operation kanban board or the equipment operation and maintenance system to realize the dynamic monitoring of the price deviation value and output in different categories. The model not only evaluates the cost of the result statically, but also connects real-time monitoring, attribute fusion and risk level evaluation into a link for calculation, and further provides more real-time operators for the total factor cost control of smart mine.

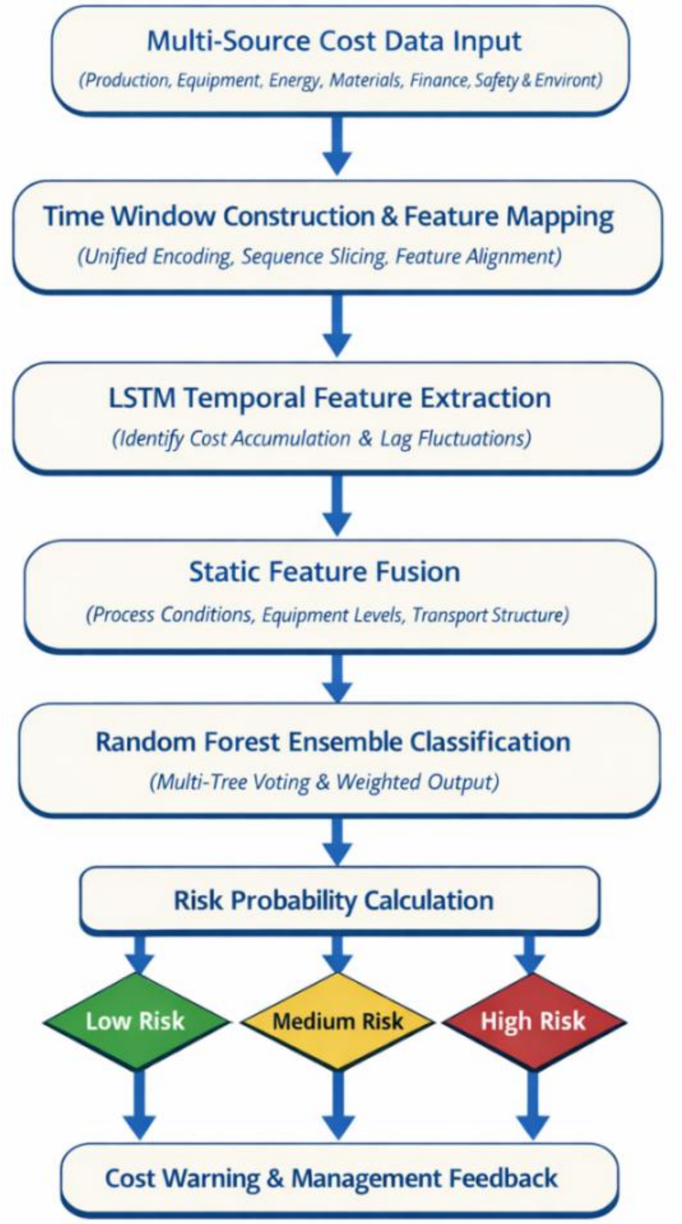

Figure 2: Construction Process of the Total Cost Risk Assessment Model

3.4 Model Training, Validation and Performance Evaluation

After constructing the facial cost risk assessment model, it is necessary to test the function of the model and evaluate its performance in many aspects. Considering that the cost data of smart mines has significant temporal continuity, this paper divides the samples into training set, validation set and test set in chronological order to avoid the information leakage problem caused by random sampling. In the training stage, sliding time windows are used to generate sequence samples, and the learning rate, time step, number of trees and risk classification thresholds are jointly tuned using the validation set. To prevent overfitting of the model on local fluctuations, the LSTM part introduces early stopping mechanism and Dropout strategy, while the random forest part enhances generalization stability by limiting the maximum depth and the number of samples in the minimum leaf node. The entire training process is completed within a unified feature space, enabling the dynamic cost sequence and static structural variables to collaborate in learning within the same calculation framework.

Since the proportion of "high cost risk" samples in the actual operation of mines is usually lower than that of normal samples, the data distribution is significantly unbalanced. If only relying on the overall accuracy rate, the model is prone to bias towards the majority class, resulting in insufficient identification of critical risk periods. Based on this, this paper introduces category weights in the training stage, giving higher error penalties to high-risk samples. The weighted loss function can be expressed as

$$L = -\sum_{i=1}^{N} \alpha_{y_i} \left[y_i \log(\hat{y_i}) + (1 - y_i) \log(1 - \hat{y_i})\right] \tag{12}$$

Among them, y_i represents the true label, $\hat{y_i}$ represents the predicted probability, and α_{y_i} represents the corresponding category weight. This practice can improve the attention of minority abnormal samples, and make the cost warning results better meet the needs of smart mine risk control.

The evaluation indicators are precision, recall and f1-score. The accuracy rate is a measure of the total judgment and it is calculated as

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{13}$$

Among them, TP, TN, FP, and FN represent true positive examples, true negative examples, false positive examples, and false negative examples respectively. The recall rate characterizes the model's ability to cover the truly high-risk samples:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{14}$$

In the scenario of cost control in mines, failing to identify high-risk situations often implies delayed maintenance, uncontrolled energy consumption, or continuous expansion of abnormal expenditures. Therefore, the recall rate has strong management implications. At the same time, if the model solely focuses on achieving high recall rates, it may lead to excessive false alarms, thereby weakening the credibility of the early warning results. Based on this consideration, this paper further uses the F1 value to comprehensively measure the precision rate and recall rate:

$$F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{15}$$

Overall, accuracy can reflect the overall classification effect, recall rate pays more attention to the risk capture ability, and F1 value helps to balance false alarms and missed alarms. When these three are combined, they can comprehensively evaluate the practicality and robustness of the model in the full cost risk identification of intelligent mines, and also provide a unified basis for the analysis of subsequent experimental results. Besides the above core indicators, this paper also takes Precision and AUC as auxiliary evaluation indicators. Precision is used to measure the credibility of the warning results, and AUC is used to depict the overall ability of the model to distinguish risk samples from non-risk samples under different thresholds. Combining the core indicators with the auxiliary indicators can more completely observe the discrimination boundary and stable performance of the model under complex working conditions, and also makes the subsequent model comparison results more interpretable.

4 Results and Discussion

4.1 Experimental Results and Analysis

To verify the effectiveness of the comprehensive cost analysis framework and risk assessment model proposed in this paper, a comprehensive analysis was conducted on the distribution of cost characteristics, model discrimination results, risk early identification capabilities, and stability under complex working conditions based on the test set. The overall results show that the cost changes in the smart mine scenario are not isolated fluctuations but are continuously coupled with equipment failures, energy consumption load, material consumption, and transportation organization. After jointly inputting multi-source business data, time series features, and static structure information into the model, the recognition accuracy of cost anomalies and the interpretability of management have significantly improved.

(1) Analysis of Comprehensive Cost Characteristics Distribution

The statistical results of the main cost eigenvalues are shown in Table 2. It can be seen from the table that the mean square deviations of daily total cost, unit consumption, ore yield, power consumption and spare parts cost are large, indicating that mine cost is greatly affected by production process and machine operation state. The mean square error of equipment failure downtime and the corresponding spare parts cost is large. It also means that costs can be unpredictable not only because they keep rising, but also because they can be quickly revealed by sudden shutdowns, emergency repairs, or energy overruns. The variation range of transportation cost proportion is also large, reflecting that the transportation distance, vehicle utilization rate, and loading and unloading connection efficiency will directly change the cost structure. In contrast, the fluctuation of safety and environmental protection investment is relatively smooth, but it still forms a linkage with other indicators during high-risk periods and cannot be simply regarded as a stable item.

Table 2: Statistical Results of Main Cost Characteristics

Feature	Mean	Standard Deviation	Minimum	Maximum
Daily Total Cost / 10,000 CNY	128.6	18.4	96.2	176.5
Unit Output Cost / CNY$\cdot$t^{-1}	83.7	11.5	61.4	109.8
Electricity Consumption per Ton of Ore / kWh$\cdot$t^{-1}	18.9	2.6	14.3	25.1
Equipment Downtime due to Failures / h$\cdot$d^{-1}	2.4	1.3	0.3	6.8
Maintenance Spare Parts Cost / 10,000 CNY$\cdot$d^{-1}	9.7	3.1	4.2	18.6
Explosives and Material Consumption Cost / 10,000 CNY$\cdot$d^{-1}	22.5	4.9	14.8	36.2
Safety and Environmental Protection Investment / 10,000 CNY$\cdot$d^{-1}	6.1	1.4	3.5	10.2
Transportation Cost Ratio / %	21.8	4.2	14.9	31.7

From the perspective of data form, there is no simple linear relationship among the cost variables. Samples with higher unit output costs often simultaneously show an increase in downtime, an increase in tonnage power consumption, and an increase in maintenance costs; while during periods of high production and stable equipment operation, although the total cost may not be the lowest, the unit cost is usually more convergent. This indicates that merely using the total financial amount for judgment can easily conceal the true source of cost risks, and by synchronously modeling operations, equipment, and financial information using computer methods, it is possible to more clearly depict the process of risks transmitting from "operational anomalies" to "cost anomalies".

(2) Comparison of recognition effects of different models

Under the same training set and test set conditions, this paper compared the proposed model with methods such as LR, SVM, XGBoost, LSTM, and RF. The results are shown in Table 3. Compared with linear models and single classification models, the fusion model achieved better results in terms of accuracy, recall rate, F1 value, and AUC. Specifically, the accuracy of the model in this paper reached 88.3%, the F1 value reached 0.854, and the AUC reached 0.92. This indicates that the model balances the recognition ability and false alarm control when capturing high-risk samples. Although LSTM has advantages for solving the time-dependent cost problem, it may not fully capture some fixed features in some cases; On the other hand, although random forests can deal with highly complex high-dimensional data well, they cannot effectively describe the cumulative effect of cost fluctuations. However, when the two are combined, the model can not only ensure its time series characteristics, but also contain various static information such as production process, equipment level, logistics conditions, etc., which improves the reliability of the prediction results.

Table 3: Comparison Results of Different Models' Performance

Model	Accuracy/%	Precision	Recall/%	F1-score	AUC
LR	76.4 ± 1.3	0.741 ± 0.02	72.5 ± 1.6	0.732 ± 0.02	0.79
SVM	79.1 ± 1.1	0.768 ± 0.02	75.0 ± 1.3	0.758 ± 0.02	0.82
XGBoost	82.4 ± 0.9	0.801 ± 0.02	78.2 ± 1.0	0.791 ± 0.02	0.86
LSTM	84.0 ± 0.8	0.816 ± 0.01	80.1 ± 0.9	0.808 ± 0.01	0.88
RF	84.7 ± 0.7	0.824 ± 0.01	80.8 ± 0.8	0.814 ± 0.01	0.88
Proposed Model（LSTM+RF）	88.3 ± 0.6	0.861 ± 0.01	84.9 ± 0.7	0.854 ± 0.01	0.92

Finally, in conclusion, the improvement of model performance is not caused by a single reason, but from the results caused by the overall combined representation of cost information and feature fusion technology. Cost anomaly identification can not be solved by a single financial problem, which is essentially an intelligent processing process of the enterprise, and it needs to combine the enterprise industry in the time dimension and multi-layer modeling at the structure level to achieve. Further analysis of the distribution of error samples shows that the main errors of different models are concentrated in two boundary conditions. First, when the unit cost increases significantly while the total cost does not change. Second, when the downtime increases but the maintenance cost has not been fully revealed, the risk signal shows the characteristics of periodicity and lag. For this type of data, it is easy to underestimate the risk with constant parameters alone, and may pay too much attention to the noise in the short term based only on the time trend. The predicted value of the hybrid model on these two types of samples is closer to the true label, which means that they can better grasp the law of cost anomaly generation and propagation path.

(3) Typical Sample Risk Probability Analysis

In order to better show the results of the model response to the cost anomaly sample, we select some of the 15 adjacent samples in the test set for analysis. According to the table, when the true label is 1(there is a cost risk), that is, at the 4th, 8th, 11th, and 14th samples, our model gives a risk value much higher than that of LSTM or RF alone. In addition, the model also suppresses some "false positive points" that "look abnormal but do not really produce abnormal", such as points numbered 5, 9, 12, etc., which indicates that it can not only detect dangerous peaks, but also filter out some false peak points, achieving a good balance effect.

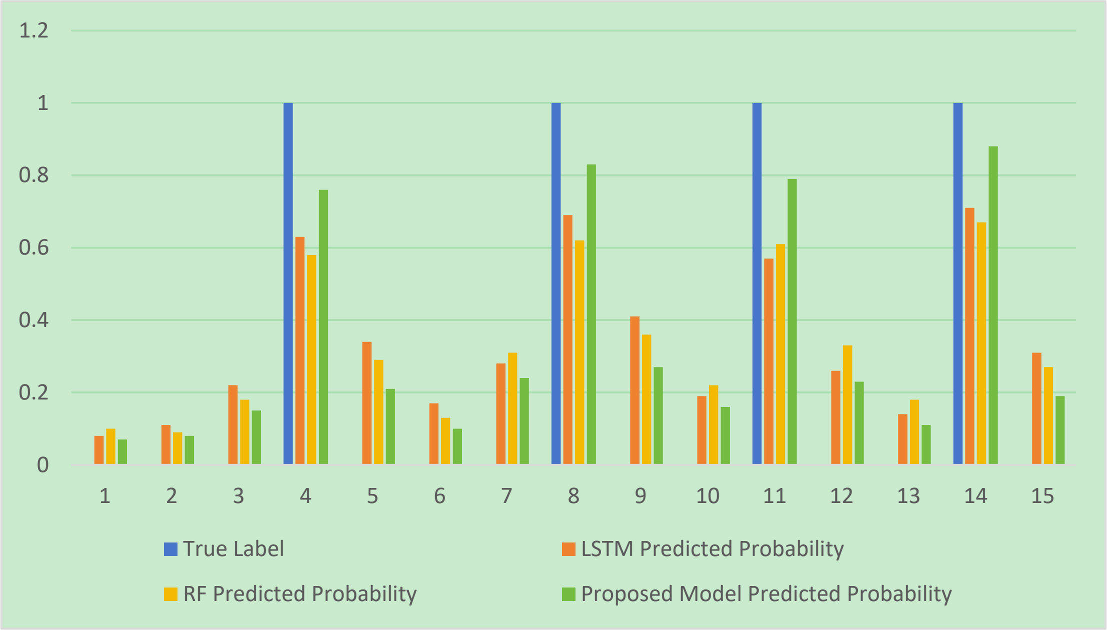

Figure 3: Comparison of Risk Probability of Typical Samples

This result indicates that the integrated model not only can quickly raise the score when risks actually occur, but also can maintain a low probability output during non-risk periods, avoiding the cost control system from being disturbed by a large number of invalid alerts. For mine operation management, this is very crucial because too many false alarms will undermine the trust of on-site departments in the warning results, while missed alarms will miss the intervention opportunity.

(4) Risk Classification and Early Warning Effect Analysis

After mapping the output probabilities to three levels of risk (low, medium, and high), the recognition success rate of high-risk samples in the test set reached 94.7%, the recognition success rate of medium-risk samples was 88.5%, and the stability rate of discrimination for low-risk samples was 91.2%. From a management perspective, high-risk samples typically correspond to scenarios such as sudden equipment shutdown, abnormal increase in energy consumption, concentrated occurrence of maintenance costs, or significant aggravation of transportation bottlenecks. If such samples are missed in the identification, the cost deviation will often be magnified in the subsequent several production cycles. The model in this paper has a higher recall rate in the high-risk range, indicating that it is more suitable for serving the goal of "fewer missed detections and traceability" in the refined management of mines.

In addition, through the analysis of feature contribution, it can be found that the unit output cost, equipment failure downtime, tonne of ore electricity consumption, maintenance spare parts cost, and transportation cost proportion are the five most significant variables affecting the risk score. This ranking has strong engineering rationality: the unit output cost reflects the cost result, downtime and energy consumption reveal process abnormalities, maintenance and transportation correspond to the cost diffusion paths. After the model incorporates these variables into the same discrimination space, it can better restore the formation logic of cost risks, rather than remaining at the surface level of statistical correlation.

(5) Performance changes under different prediction lead times

To evaluate the forward-looking identification ability of the model, the changes in indicators under prediction lead times of 1 day, 3 days, 7 days, and 14 days were further compared. Figure 4 Line chart can be drawn according to the table below. The results show that the Precision, Recall and F1 values all slowly decrease with the extension of the prediction lead time, but remain at a high level in the 7-day window as a whole. The performance of 1-3 days window is the most stable, indicating that the model has a strong response ability to short cycle cost risk. When the prediction span is extended to 14 days, the accuracy of the model decreases due to the influence of production plan adjustment, external supply fluctuation and on-site random disturbance, but it still has an acceptable warning value.

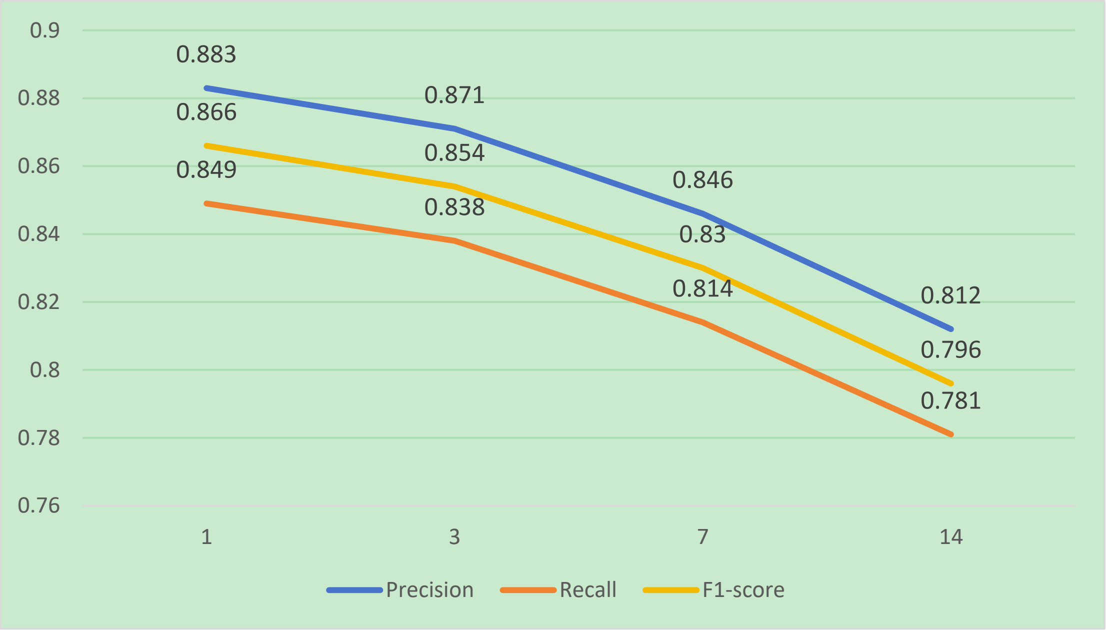

Figure 4: Changes in Model Performance under Different Forecast Lead Times

This evolution law is also in line with the operation characteristics of the mine cost system itself. In the short term, the mechanical state, energy consumption level and workload and other factors show good inertia, so the model is easy to learn from historical sequence information. However, with the change of production tasks, the fluctuation of material prices and the allocation of outsourcing jobs, more uncertainties will be brought. This will increase the difficulty of prediction in advance, but the integrated model can still maintain an F1 value of approximately 0.80 within the 14-day warning cycle, indicating that it does not cause obvious instability due to the longer period.

(6) Stability observation under complex conditions

In addition, the proposed method still has strong robustness in the presence of partial missing observations, adding noise and class imbalance. After 10% of the sensor observations are randomly masked, the decrease of F1 score is less than 0.02. After adding a certain amount of Gaussian noise, the accuracy decreases less than 0.002. In the case of fewer low-risk samples, the recall rate of this model fluctuates around 0.8 by using the time window resampling method after class weighting. The above conclusions prove that the fusion of diversified cost information does not reduce the expressiveness of the model, and because more input pathways are added, other features are relied on to maintain the decision-making function when some routes are blocked.

In summary, the experimental results show that the analysis of the total cost of intelligent minerals based on the total number of financial cases or single type of service indicators can not reflect the process and route of risk change. However, the use of computer technology to integrate various service data such as production, machining, energy consumption and materials can improve the accuracy of risk identification, and secondly can deepen the perception effect of managers. This has laid a good empirical foundation for us to deeply analyze its advantages, applicable conditions and development direction.

4.2 Discussion

Combined with the experimental results, it can be seen that the full-cost risk assessment model constructed in this paper shows clear comprehensive advantages in the smart mine scenario. Compared with the LSTM model alone, the F1 score of the fusion model was increased from 0.808 to 0.854, and the recall rate was increased from 80.1% to 84.9%. Compared with the single RF model, the F1 value is increased from 0.814 to 0.854, and the recall rate is increased from 80.8% to 84.9%. This indicates that the comprehensive cost risk identification is not suitable to be understood as a single financial classification task, but is more closely related to an industrial intelligent analysis problem driven by time-dependent relationships, structural constraints, and operational disturbances. LSTM can capture the time-dependent relationships formed by cost accumulation, equipment deterioration, and energy drift, while RF is better at handling static variables such as process conditions, transportation structures, and equipment grades. After the features are connected at the feature level, the model's ability to identify complex cost anomalies has significantly improved.

Compared with the existing research, the research in this paper not only improves the evaluation index, but also expands a new research perspective. Most of the existing cost analysis of the mineral industry focuses on the prediction of investment cost, the cost estimation of a single process or financial analysis. The research on risk mainly focuses on safety risk or equipment damage, and there is no unified calculation method for these two levels. In this paper, we try to integrate the data of production line monitoring, equipment maintenance, energy consumption measurement, material flow tracking and financial reconciliation, so that the cost fluctuation can be found continuously according to the path of "operation deviation - resource overconsumption - cost increase - risk level determination". Therefore, this model does not simply replace manual decision-making, but transforms the originally scattered business information into a measurable, comparable and traceable risk expression on the computer. From the perspective of practical application, this approach also has the possibility of system integration. The full-cost analysis system can rely on ERP, MES, equipment maintenance platform and energy management and control platform for interface docking. The risk level formed can be used as an early warning signal in the business analysis cockpit, and can also be used to guide the adjustment of repair procedures and repair systems, annual plan and material preparation plan. In a word, the calculation results can not only be used as theoretical numerical evaluation, but also be applied to the daily production management of mining enterprises to form a closed-loop mechanism of "monitoring-identifiing-feedback-correct ion".

However, the method in this paper still has some limitations: because of the high quality requirements of historical samples, the model will learn inaccurate parameters when there are fewer high-risk samples or the labeling of high-risk samples is unstable. The equipment configuration, transportation link and pricing mode of different mines are also quite different. This will cause some instability of the model in different scenarios. Secondly, the reasoning time is short, and the efficiency is affected by multi-source information processing, cleaning and synchronization. If the interface of the real system is not uniform, the online effect will be limited. Future research can further explore the stability and continuity of the model in complex production environments. The research work of multi-mine transfer learning, few-sample risk identification and online update of digital twin should be strengthened.

5 Conclusion

In order to deal with the complex cost composition of the interval in the above intelligent mine operation scene, the close cross of related factors, and the difficulty of traditional accounting methods to adapt to the requirements of dynamic response, this paper proposes a cost analysis system that runs through the whole process of operation, full elements, and multi-dimensional data fusion, and proposes to apply LSTM-RF method to realize cost risk assessment. In this paper, multi-source heterogeneous data such as production planning and scheduling, equipment operation status monitoring, energy consumption measurement, logistics process, safety and environment control and financial accounting are incorporated into the unified computing process to realize the whole process perception and dynamic identification from static one-way cost accounting. The abnormal cost fluctuation can be dynamically described in the form of "operation deviation -- resource consumption anomaly -- cost increase -- risk assessment classification".

The experimental results show that the model we constructed is superior to the single model in terms of accuracy, recall, F1 value and AUC indicators, and has better stability in identifying dangerous samples. Therefore, the combination of time series construction function and static feature discrimination function can more effectively capture the lag, accumulation and structural characteristics of mine cost fluctuation. Compared with the post-analysis method relying solely on financial information, the method in this paper relies more on computer technology to solve the problems of data storage, feature fusion, risk prediction and early warning response, so as to improve the degree of refinement and prediction level of mining operation and management. Of course, there are some limitations of our study: the model requires high historical sample size, data quality, and connected integrity of the system. Therefore, its application will be limited in the case of new mines, insufficient samples and non-uniform multi-system standards. At the same time, there are differences in the production process, machinery and equipment and cost accounting caliber of different mines, so the migration ability of the model still needs to be further verified. In the future, we can also explore the problems of small data learning methods, knowledge transfer methods, online incremental update mechanism relying on digital twin, and interpretability enhancement for decision-making process. On this basis, we can improve our cost risk identification and early warning system, and realize the direction of smart mining area construction from "visual" to "calculable".

Acknowledgements

This study was supported by the Economic Demonstration Project of the Shuitabixi Antimony Mine in Canada and the Construction Project for the Tailings Treatment of Lithium-Tin Polymetallic Ore Mining and Dressing at Weilastuo Mining Co., Ltd.

References

[1] Osei R, Frimpong S, Venkat A. Human-machine collaboration in mining: A critical review of emerging frontiers of intelligence systems in the mining industry[J]. The Extractive Industries and Society, 2025, 24: 101746. https://doi.org/10.1016/j.exis.

2025.101746

[2] Mahmoudi Kouhi R, Tariq Najeeb A, Taherdangkoo R, et al. A survey study on the adoption and perception of artificial intelligence in the mining industry[J]. Discover Applied Sciences, 2025, 7(7): 713.https://doi.org/10.1007/s42452-025-07342-1

[3] Zheng X, Nguyen H, Bui X N. Exploring the relation between production factors, ore grades, and life of mine for forecasting mining capital cost through a novel cascade forward neural network-based salp swarm optimization model[J]. Resources Policy, 2021, 74: 102300. https://doi.org/10.1016/j.resourpol.2021.102300

[4] O'Brien T M, Alonso E. A simplified rare earth element mining project cost estimator-a new tool for evaluating future mine supply[J]. Mineral Economics, 2025, 38(2): 237-252. https://doi.org/10.1007/s13563-024-00467-6

[5] Losaladjome Mboyo H, Huo B, Mulenga F K, et al. Distribution of operating costs along the value chain of an open-pit copper mine[J]. Applied Sciences, 2025, 15(3): 1602. https://doi.org/10.3390/app15031602

[6] Domínguez C R, Martínez I V, Peña P M P, et al. Analysis and evaluation of risks in underground mining using the decision matrix risk-assessment (DMRA) technique, in Guanajuato, Mexico[J]. Journal of Sustainable Mining, 2019, 18(1): 52-59. https://doi.org/10.1016/j.jsm.2019.01.001

[7] Buchholz P, Schumacher A, Al Barazi S. Big data analyses for real-time tracking of risks in the mineral raw material markets: implications for improved supply chain risk management[J]. Mineral Economics, 2022, 35(3): 701-744. https://doi.org/10.1007/s13563-022-00337-z

[8] Botín J A, Vergara M A. A cost management model for economic sustainability and continuos improvement of mining operations[J]. Resources Policy, 2015, 46: 212-218. https://doi.org/10.1016/j.resourpol.2015.10.004

[9] Ozdemir B, Kumral M. A system-wide approach to minimize the operational cost of bench production in open-cast mining operations[J]. International Journal of Coal Science & Technology, 2019, 6(1): 84-94.https://doi.org/10.1007/s40789-018-0234-1

[10] Zeng L, Liu S Q, Kozan E, et al. A comprehensive interdisciplinary review of mine supply chain management[J]. Resources policy, 2021, 74: 102274. https://doi.org/10.1016/j.resourpol.2021.102274

[11] Majstorovic V, Simeunovic V, Miskovic Z, et al. Smart Manufacturing as a framework for Smart Mining[J]. Procedia CIRP, 2021, 104: 188-193. https://doi.org/10.1016/j.procir.2021.11.032

[12] Nobahar P, Xu C, Dowd P, et al. Exploring digital twin systems in mining operations: A review[J]. Green and Smart Mining Engineering, 2024, 1(4): 474-492. https://doi.org/10.1016/j.gsme.2024.09.003

[13] Guo H, Nguyen H, Vu D A, et al. Forecasting mining capital cost for open-pit mining

projects based on artificial neural network approach[J]. Resources Policy, 2021, 74: 101474. https://doi.org/10.1016/j.resourpol.2019.101474

[14] Nourali H, Osanloo M. Mining capital cost estimation using Support Vector Regression (SVR)[J]. Resources Policy, 2019, 62: 527-540. https://doi.org/10.1016/j.resourpol.2018.10.008

[15] Zhang H, Nguyen H, Bui X N, et al. Developing a novel artificial intelligence model to estimate the capital cost of mining projects using deep neural network-based ant colony optimization algorithm[J]. Resources Policy, 2020, 66: 101604. https://doi.org/10.1016/j.resourpol.2020.101604

[16] Savolainen J. Real options in metal mining project valuation: Review of literature[J]. Resources Policy, 2016, 50: 49-65.https://doi.org/10.1016/j.resourpol.2016.08.007

[17] Tubis A, Werbińska-Wojciechowska S, Wroblewski A. Risk assessment methods in mining industry—a systematic review[J]. Applied Sciences, 2020, 10(15): 5172. https://doi.org/10.3390/app10155172

[18] Beeche C A, Garcia M A, Leng S, et al. Computational risk modeling of underground coal mines based on NIOSH employment demographics[J]. Safety science, 2023, 164: 106170. https://doi.org/10.1016/j.ssci.2023.106170

[19] Wang Y, Wang L, Wang H, et al. Data-Driven and Model-Driven Integration Approach for Optimizing Equipment Safety Investment in Digital Twin Coal Mining Enterprises[J]. Applied Sciences, 2024, 14(23): 11101. https://doi.org/10.3390/app142311101

[20] Du G, Chen A. Coal mine accident risk analysis with large language models and bayesian networks[J]. Sustainability, 2025, 17(5): 1896. https://doi.org/10.3390/su17051896