frontiers | Frontiers in Materials

TYPE Original Research
PUBLISHED 07 January 2026
DOI 10.3389/fmats.2025.1737888

OPEN ACCESS

EDITED BY
Mario Milazzo,
University of Pisa, Italy

REVIEWED BY
Sana Ullah,
George Mason University, United States
Sandesh Patil,
D Y Patil International University, India

*CORRESPONDENCE
Yueying Zhang,
lhzlhz1028@163.com

RECEIVED 02 November 2025
REVISED 29 November 2025
ACCEPTED 05 December 2025
PUBLISHED 07 January 2026

CITATION
Liang H, Zhang Y, Xiang C, Fei S, Wei H, Han B and Zhang A (2026) Physics-informed neural network–transformer for dual-objective prediction and mix optimization of backfill materials.
Front. Mater. 12:1737888.
doi: 10.3389/fmats.2025.1737888

Physics-informed neural network–transformer for dual-objective prediction and mix optimization of backfill materials

Huizhen Liang, Yueying Zhang*, Chengmi Xiang, Shanshan Fei, Han Wei, Bei Han and Aijun Zhang

School of Economics and Management, Xi'an Kedagaoxin University, Xian, China

To address the issues of traditional backfill material mix design relying on experience and low efficiency, this study proposes a physics-informed neural network (PINN)–transformer method that integrates physical constraints. A dual-task prediction framework is constructed considering material strength and slump, embedding strength development monotonicity, convexity constraints, and slump rheological principles into model training to improve the accuracy and physical reasonableness of the prediction results. Experimental results show that this method improves the mean absolute error (MAE) metric by 6.0% compared to the transformer in strength prediction and improves the slump prediction MAE metric by 6.5%. A multi-objective mix optimization system is established based on prediction results and economic analysis, proposing three optimization strategies adapted to different engineering requirements. This method breaks through the limitations of traditional empirical design and provides efficient and reliable technical support for scientific mix design and engineering decision-making regarding mine backfill materials.

KEYWORDS

backfill materials, dual-objective prediction, engineering economic analysis, multi-objective mix optimization, physics-informed neural network, transformer

1 Introduction

Backfill materials are widely used in mining, underground engineering, and foundation reinforcement, with their performance directly related to engineering safety and economic considerations. Promoting backfill mining technology faces challenges, including large investments and high costs (Liu et al., 2020), along with the disposal pressure of industrial solid waste such as coal gangue (Chang et al., 2022). Optimizing the material mix has become key to achieving safety, environmental protection, and economic benefits.

Current mix optimization faces multiple engineering technical difficulties. First, backfill material components are complex (such as cementitious materials, aggregates, additives, and other multi-phase mixtures), with nonlinear synergistic or antagonistic effects existing between components (Ghirian and Fall 2013) that make it difficult for traditional methods to accurately quantify their impact on core performance.

Second, material performance exhibits significant temporal evolution characteristics and is subject to multiple dynamic factor interferences, requiring mix designs to balance static parameter optimization with dynamic process adaptation (Wu et al., 2020). Third, increasingly stringent engineering requirements for mix economy and environmental friendliness further increase the complexity of multi-objective optimization, making it difficult for traditional experimental methods or single models to achieve global optimal solutions (Zhao et al., 2024).

Traditional backfill material mix design mainly relies on empirical analogy methods (Belem and Benzaazoua, 2008), experimental optimization methods (such as orthogonal experiments (Chen et al., 2022) and response surface methodology (Cihangir et al., 2022)), and numerical simulation methods (Liu et al., 2021). With the development of artificial intelligence technology, the limitations of these traditional methods have become increasingly apparent. Empirical analogy methods are highly subjective, experimental optimization methods have long cycles and high costs, and numerical simulation methods depend on accurate constitutive relationships. Therefore, new methods combining data-driven and physical constraints have emerged, aiming to discover complex patterns that are difficult to reveal by traditional methods with higher efficiency and lower cost, thus achieving global optimization of mix design.

Among existing machine learning methods, gradient boosting regression trees, extreme learning machines, and other models perform well in specific scenarios but have obvious limitations. Research by Qi et al. (2018) demonstrates that although machine learning methods such as regression trees (RTs), random forests (RFs), and gradient boosting regression trees (GBRTs) are widely used in backfill strength prediction, their application has not reached peak potential, lacking more robust technical support. Related research shows that ensemble learning methods based on GBRTs demonstrate good performance in predicting the uniaxial compressive strength of backfill materials with hyperparameter tuning through particle swarm optimization (Taraghi et al., 2023). However, such models mostly rely on large amounts of labeled data and do not fully incorporate physical laws of material mechanics, resulting in limited generalization capability in data-scarce or condition-changing scenarios.

Physics-informed neural networks (PINNs) can introduce physical constraints into the deep learning process, ensuring that prediction results conform to the physical laws of materials. Wang J. D. et al. (2025) analyzed buried pipeline deformation under permanent ground displacement based on PINN, integrating residuals of governing physical equations into the loss function, thus significantly reducing dependence on large datasets and demonstrating excellent generalization and parameter inversion capabilities. Related research by Tang et al. (2025) demonstrates that physics-informed neural networks embedding physical constraints in the loss function integrate mechanical principles into the model training process, and the results indicate that the model has good prediction accuracy, robustness, and generalization capability. These studies demonstrate that combining physical information with machine learning can provide reliable and effective methods for evaluating mine backfill material performance. However, physics-informed neural networks mainly focus on embedding physical constraints and face challenges in handling temporal evolution processes and multi-stage reaction mechanisms of material mix.

The transformer model, with its powerful feature extraction and sequence modeling capabilities, has shown significant advantages in material design and multi-variable optimization. Xu et al. (2023) proposed a transformer-based polymer property prediction model, with rigorous experiments on 10 polymer property prediction benchmarks demonstrating its superior performance, indicating that the transformer can facilitate rational material design. Bo et al. (2023) proposed a transformer-based anomaly detection model for filtering abnormal data of perceived object real states in mine backfill operations, with results showing that the model is suitable for data validity detection in scenarios with increasing perceived objects in mine backfill sensing systems. However, the purely data-driven nature of the transformer has inherent limitations: a lack of physical constraint embedding mechanisms, an inability to ensure that prediction results conform to the physical laws and mechanical principles of materials themselves, and a rapid decline in generalization capability when data distributions shift or samples are scarce, making it difficult to meet the reliability requirements for engineering applications (Michaloglou et al., 2025).

Addressing the respective limitations of existing PINN and transformer research, this study achieves key breakthroughs in their fusion technology, constructing a PINN–transformer model with both physical consistency and strong feature extraction capability. The transformer utilizes its self-attention mechanism to accurately capture long-range dependencies and high-dimensional nonlinear mappings between backfill material multi-component mix parameters, environmental factors, and performance indicators, solving problems that exist in traditional PINN in temporal modeling and high-dimensional feature extraction. Conversely, PINN effectively guides the model to find physically consistent solutions by introducing physical equations as soft constraints into the loss function, compensating for the transformer's lack of physical constraints. The combination of PINN and transformer theoretically has significant complementarity and synergistic effects, effectively compensating for the deficiencies of single methods in complex experiments, resolving the contradiction between the poor generalization capability of purely data-driven models and the insufficient adaptability of purely physical models, and providing scientific and reliable technical support for backfill material mix design.

2 Theoretical foundation

2.1 PINN method

A PINN is an innovative method that embeds physical constraints into deep learning frameworks, achieving accurate modeling and prediction of complex physical systems by incorporating known physical equations as regularization terms into the neural-network training process (Antonion et al., 2024). In research regarding optimal backfill material mix, PINN effectively guides the model to capture complex nonlinear relationships between components by introducing physical equations from

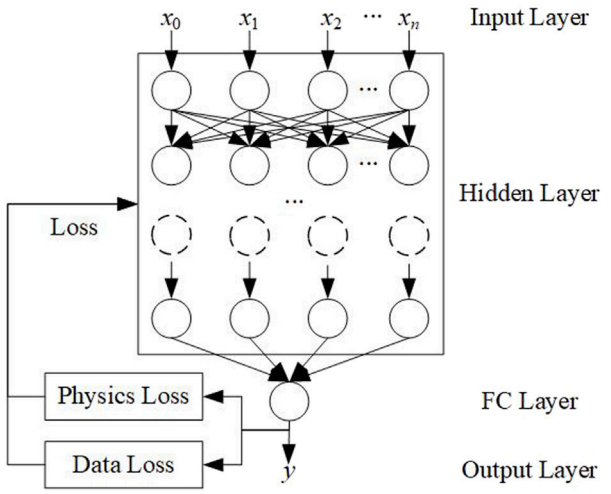

FIGURE 1
PINN architecture principle diagram.

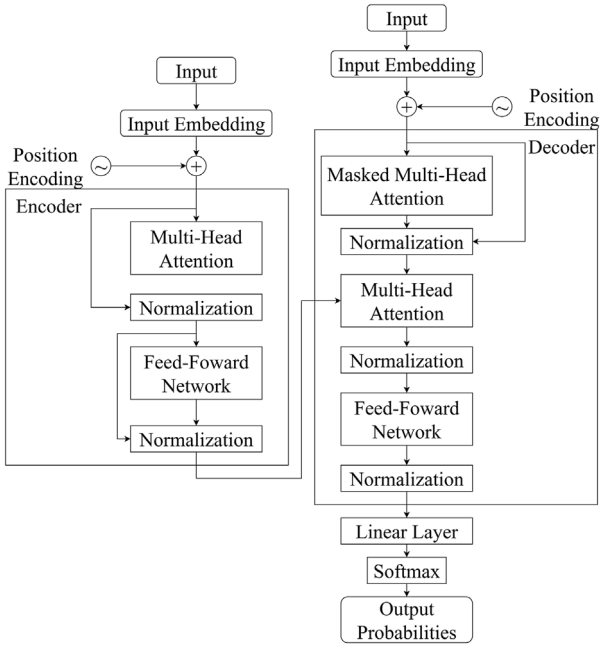

FIGURE 2
Transformer model architecture diagram.

material mechanics, fluid mechanics, and other related fields, such as physical constraints. Its architecture diagram is shown in Figure 1.

The advantage of this method lies in its ability to fuse domain knowledge with data-driven learning, thus improving model generalization capability and physical reasonableness of the prediction results in data-limited scenarios, which is especially suitable for handling multi-physics coupling problems involved in backfill processes. However, the PINN method also has certain limitations. First, its performance strongly depends on the accuracy and completeness of the introduced physical equations. When the physical models have simplified assumptions, parameter deviations, or boundary conditions that are difficult to accurately quantify, their errors will directly transfer to the prediction results, affecting model reliability (Fernández de la Mata et al., 2023). Second, the training process needs to simultaneously optimize multiple losses, including data error, physical residuals, and boundary constraints, with weight settings that lack unified criteria, easily causing gradient conflicts or training instability, thereby reducing model convergence efficiency. Third, compared to purely data-driven methods, PINNs have higher computational costs, especially in high-dimensional scenarios, complex partial differential equations, or fine mesh solving conditions, thus potentially facing convergence difficulties or significantly increased computational resource consumption.

2.2 Transformer method

The transformer is a deep learning architecture based on self-attention mechanisms, achieving efficient processing of sequence data through multi-head attention and positional encoding. In backfill material mix research, the transformer can treat different mix parameters as a sequence of elements, capturing complex interaction relationships between components through self-attention mechanisms and identifying key mix factors affecting backfill effectiveness. This model excels at handling dynamic characteristics of backfill material performance evolution over time, and it is capable of modeling temporal dependencies of various parameters throughout the entire process from initial preparation to solidification (Song et al., 2025). The transformer model architecture diagram is shown in Figure 2.

The model adopts an encoder–decoder structure, where the encoder is responsible for extracting deep representations of input features, and the decoder generates predicted outputs based on encoded information. The core of the transformer is the self-attention mechanism, which consists primarily of three parts: query, key, and value. The calculation formula is provided in Equation 1 (Ghojogh and Ghodsi, 2020):

$$\text{Attention}(Q,K,V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \tag{1}$$

where Q represents the query vector, K represents the key vector, V represents the value vector, and d_k represents the dimension of the key vector. Equation 1 scales attention scores by dividing by $\sqrt{d_k}$ and then performs weighted summation using the value vector V to obtain the output of each attention head. Then, the outputs of h attention heads are concatenated to obtain the final multi-head attention output with the calculation formulas provided in Equations 2, 3:

$$\text{MultiHead}(Q,K,V) = \text{Concat}(\text{h}_1,\ldots,\text{h}_n)W^O, \tag{2}$$

$$\text{h}_i = \text{Attention}\left(QW_i^Q, KW_i^K, VW_i^V\right), \tag{3}$$

where h_i represents the ith self-attention head.

The main advantages of the transformer are reflected in its powerful parallel computing capability and long-distance dependency capture ability; it is capable of simultaneously

considering the global impact of multiple mix proportion factors, performing excellently when dealing with high-dimensional mix proportion optimization problems. However, this method also has limitations, including high computational complexity and a large number of parameters, which lead to expensive training costs and require substantial amounts of labeled data to achieve optimal performance (Antonion et al., 2024). In addition, the "black box" nature of the transformer makes its interpretability relatively poor in the field of backfill materials, making it difficult to directly reveal the physical mechanisms between mix proportion parameters and material performance, which limits its application in engineering practice to some extent.

3 Backfill material mix proportion analysis based on PINN–transformer

3.1 Data collection

In experiments, the main components of backfill materials are coal gangue, cement, and fly ash, with water (H_2O) as the main liquid component. This study uses coal gangue as the backfill aggregate and ordinary Portland cement as the main cementitious component, supplemented with fly ash as the auxiliary cementitious material (Wang Z. et al., 2025). The selected coal gangue is first crushed to below 50 mm using a jaw crusher and then finely crushed using an impact crusher, with the overall particle size finally controlled within 16 mm according to the densest packing theory. The cement used was 32.5 strength-grade ordinary Portland cement, supplemented with class-F fly ash as the auxiliary cementitious material. The fly ash is obtained after flue gas capture and special treatment from Yulin Power Plant. Class-F fly ash is well suited for backfill applications requiring long-term stability and corrosion resistance due to its low calcium and high silicon–aluminum composition. Adding appropriate amounts of fly ash can effectively improve the rheological properties of backfill slurry and optimize the microstructure, while significantly enhancing late-stage strength, reducing hydration heat, and drying shrinkage risks, with significant benefits in engineering economy, environmental friendliness, and solid-waste resource utilization (Shao et al., 2020).

To investigate the effects of fly ash content and slurry mass concentration on the performance of backfill materials, samples were prepared according to the mix proportions listed in Table 1. Mass concentration is defined as the percentage of solid components (fly ash, cement, and coal gangue) to the total slurry mass. In the experiments, each component was first weighed and then fed into a mortar mixer, with water added according to designed mass concentrations (72%, 74%, 76%, 78%, and 80%) for mixing until the slurry reached a uniform state. This concentration range was determined based on engineering practical experience and pre-experimental results, ensuring that the slurry had the necessary flowability for pumping while systematically examining its transition from the pumpable state to the flowability critical state.

After mixing, a part of the slurry was immediately taken for slump testing to characterize its flow performance. The remaining slurry was used to prepare mechanical performance specimens: the mixed slurry was poured into 70.7 mm^3 triple plastic molds, demolded, and numbered after 24 h of room-temperature curing and then placed in a standard curing box (temperature 20 °C ± 2 °C; relative humidity 95% ± 1%) for subsequent curing. Curing ages were set at 7, 14, 28, and 56 days to completely characterize the entire process from early strength development to long-term stability. Among these, 7 and 14 days represent the early strength development stages, which help evaluate material strength growth rate and early stability, providing data support for rapid underground backfill support; 28 days is the commonly used standard age for cement-based materials, serving as a baseline reference for backfill material strength evaluation; 56 days is the long-term strength observation stage, which is used to evaluate strength evolution and stability of backfill during long-term service, thereby verifying the long-term reliability of the mix schemes (Ercikdi et al., 2014). After each curing age, compressive strength was measured according to standard test methods to ensure that the obtained data can truly reflect the performance characteristics of backfill materials under different mix proportions and curing conditions.

To construct the PINN–transformer dual-task prediction model, this study systematically collected material strength and slump data through backfill material experiments with different proportions, as designed in Table 1. Specifically, the fly ash content was set at seven levels: 0%, 5%, 10%, 15%, 20%, 25%, and 30%, with each level corresponding to five cement contents, and the coal gangue content was determined using a mass balance. Each mix proportion was measured at five curing time points, with three independent repeated trials. Slump indicators were measured at different mass concentrations ranging from 72% to 80%, and each measurement was repeated three times using the same mix proportions and curing times. The experimental dataset scale reached 900, providing sufficient samples for model training and validation, which was capable of truly reflecting the nonlinear relationship between the material mix and performance evolution over time. Figure 3 shows the backfill material strength at different mix proportions and curing times under the same mass concentration (76%), along with slump data at the same mix proportion under different mass concentrations.

On analyzing Figure 3, based on the strength curves, it can be observed that strength increases with prolonged curing time for all mix proportions; based on the slump comparison, a higher concentration leads to a smaller slump and decreased flowability. The two figures reflect the contradictory relationship between strength and workability. Simply increasing or decreasing mass concentration cannot simultaneously meet engineering strength requirements and construction convenience needs; therefore, finding optimal mix proportions that balance material performance and construction performance is particularly important and urgent.

3.2 PINN–transformer model construction

To predict the mechanical performance of backfill materials under different mix proportions and curing times, this study designs a PINN–transformer model integrating physical constraints. PINN–transformer utilizes the transformer's global feature modeling capability to capture complex relationships between mix parameters while also using PINN's physical constraint mechanism

TABLE 1 Experimental proportion table.

Mass concentration	Fly ash content	Cement content	Coal gangue content
72%~80%	0%	6%, 8%, 10%, 12%, and 14%	86%, 88%, 90%, 92%, and 94%
	5%		81%, 83%, 85%, 87%, and 89%
	10%		76%, 78%, 80%, 82%, and 84%
	15%		71%, 73%, 75%, 77%, and 79%
	20%		66%, 68%, 70%, 72%, and 74%
	25%		61%, 63%, 65%, 67%, and 69%
	30%		56%, 58%, 60%, 62%, and 64%

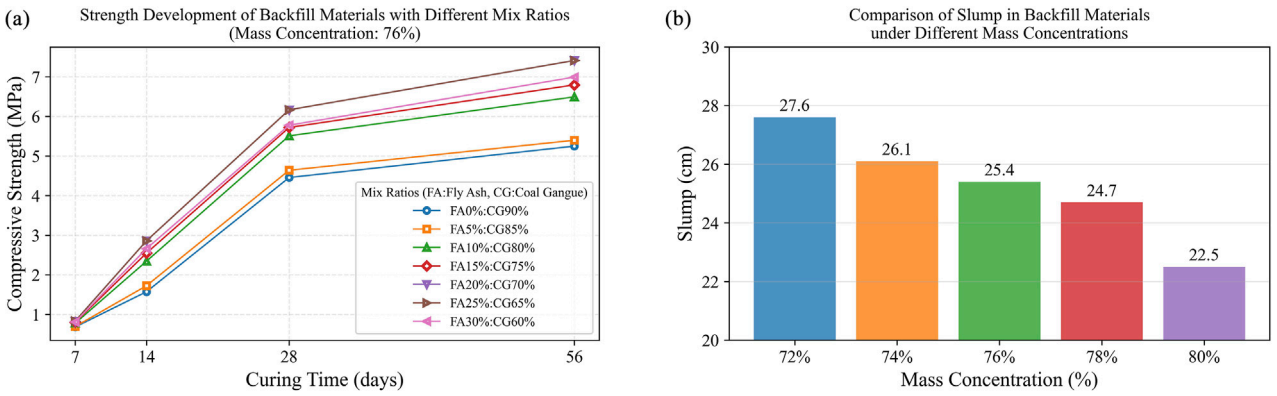

FIGURE 3
Comparison of the backfill material performance under different conditions. **(a)** Strength variation in backfill materials with different mix proportions and curing times under the same mass concentration. **(b)** Slump variation in backfill materials with identical mix proportions under different mass concentrations.

to ensure reasonableness and interpretability of the prediction results, thereby achieving high-precision prediction of backfill material performance. Model inputs include the cement content, coal gangue content, fly ash content, water content, and curing time. There are two outputs: one is backfill material strength (related to the cement content, coal gangue content, fly ash content, water content, and curing time), and the other is backfill material slump (related to the cement content, coal gangue content, fly ash content, and water content).

The overall framework of the dual-task PINN–transformer model constructed in this study is shown in Figure 4.

The model is a deep network consisting of multiple transformer encoder blocks connected in series. During training, it is optimized as a whole by both physical constraint loss and data loss. It mainly consists of four primary parts, namely, the input layer, transformer encoder module, PINN constraint layer, and output layer, with the functions of each part as follows.

The input feature vector $X = \left[x_c, x_g, x_f, x_w, t\right]^T$ includes the cement mass fraction x_c, coal gangue mass fraction x_g, fly ash mass fraction x_f, water mass fraction x_w, and curing time t as raw features. Data are processed through Z-score normalization, converting multi-dimensional data into standardized vector representations that are suitable for neural network processing, where curing time t is normalized as a continuous input variable, facilitating the model's understanding of time sequence and intervals.

The model utilizes N-layer transformer encoders, with each layer containing a multi-head self-attention mechanism and a feedforward neural network. Through positional encoding, it captures the sequence characteristics of mix proportion parameters, uses attention weight matrices to learn interaction relationships and global dependencies between different mix proportion components, and ensures training stability through residual connections and layer normalization, thereby extracting deep nonlinear feature representations of mix proportion parameters.

Physical laws that have been verified over time in the field of materials science are directly integrated into the loss function of the neural network in the form of mathematical constraints. According to scientific research evidence, the strength development of backfill materials follows specific temporal evolution laws. In the early curing stage (0 days–28 days) (Ghojogh and Ghodsi, 2020), strength development mainly follows hyperbolic function laws because cement hydration reactions are relatively active, resulting in rapid strength growth; however, as the hydration components

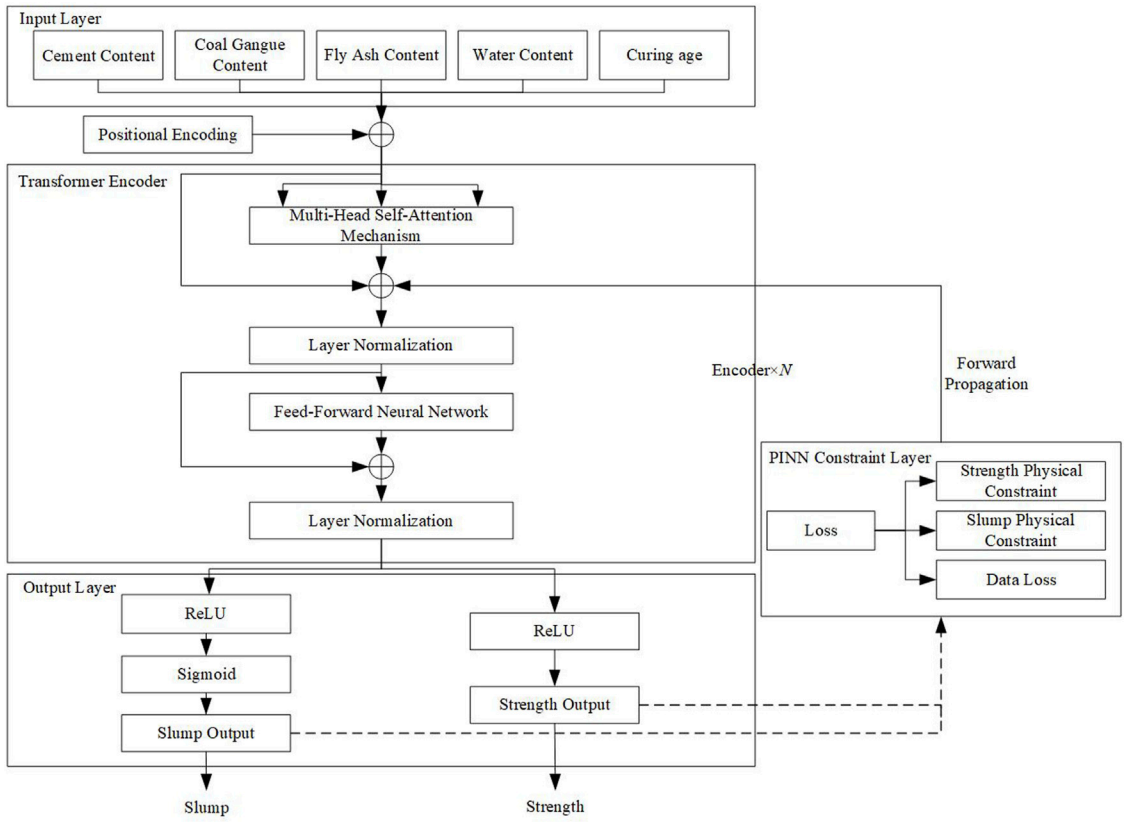

FIGURE 4
Overall framework of the PINN–transformer model.

are gradually consumed, the growth rate slows down. For long-term strength development after 28 days, research shows that the modified exponential function model is more applicable, reflecting the characteristic that strength tends to stabilize in the later stage and conforming to the physical mechanism of long-term cement hydration. The formula is shown in Equation 4:

$$\sigma(t) = \begin{cases} \dfrac{\sigma_{28} \cdot t}{a + t}, 0 \le t \le 28 \\ \sigma_{\infty}\left[1 - e^{-\beta(t-28)^n}\right] + \sigma_{28}, t > 28 \end{cases}, \tag{4}$$

where σ_{28} is the 28-day standard strength, $a = f(x_c, x_g, x_f, x_w)$ is the development rate parameter related to the mix proportion, σ_{∞} is the long-term ultimate strength increment, β is the later-stage development rate coefficient, and n is the shape parameter (usually taken as 0.5–1.0).

1. The strength development of backfill materials must meet basic physical rationality requirements.
   a. First is the monotonicity constraint, meaning that strength can only increase with time and will not decrease, which must satisfy Equation 5:

$$\frac{\partial \sigma}{\partial t} \ge 0. \tag{5}$$

This constraint term is implemented by adding a penalty term to the loss function. When the predicted strength growth rate is negative, this term produces a large loss, forcing the model to correct the prediction result. The monotonicity constraint loss is shown in Equation 6:

$$L_{monotonic} = \max\left(0, -\frac{\partial \sigma}{\partial t}\right)^2. \tag{6}$$

   b. During the curing stage, strength growth should also satisfy the convexity constraint; that is, the growth rate should gradually decrease, as shown in Equation 7:

$$\frac{\partial^2 \sigma}{\partial t^2} \le 0, t \le 28. \tag{7}$$

The convexity constraint loss is shown in Equation 8:

$$L_{convex} = \max\left(0, \frac{\partial^2 \sigma}{\partial t^2}\right)^2. \tag{8}$$

2. As an important indicator of backfill material performance, slump prediction needs to follow the basic laws of fluid mechanics (Wang et al., 2024). Based on the rheological properties of backfill materials, slump is mainly affected by the water content and particle composition:

$$s = s_0 \cdot x_w^{\alpha_1} \cdot \exp\left(-\alpha_2 x_c - \alpha_3\left(x_g + x_f\right)\right) \cdot \left(1 + \alpha_4 x_f\right), \quad (9)$$

where s_0 is the reference slump, α_1 reflects the positive effect of water content, α_2 embodies the hindering effect of cement fineness on fluidity, α_3 characterizes the overall hindering effect of solid particles, and α_4 characterizes the water-reducing effect of fly ash.

The corresponding constraint loss is shown in Equation 10:

$$L_{rheology} = \left\| s_{pred} - s_0 x_w^{\alpha_1} e^{-\alpha_2 x_c - \alpha_3\left(x_g + x_f\right)}\left(1 + \alpha_4 x_f\right) \right\|^2, \quad (10)$$

where λ_i represents the weight coefficients for each constraint term, with the optimal configuration determined through hyperparameter optimization.

Through the PINN physical constraint system, the model can not only accurately fit training data but, more importantly, can also ensure that prediction results conform to the basic laws of materials science throughout the entire parameter space (Uddin et al., 2025), enabling the model to more deeply understand the essential laws of backfill material mix proportion design.

The output layer is the final decision-making module of the model, adopting a multi-task learning architecture to simultaneously predict two key performance indicators of backfill materials: strength and slump. This layer maps the high-dimensional feature representations extracted by the transformer encoder to specific engineering performance parameters. Based on the differences in task characteristics, the output layer designs two parallel prediction branches. The strength prediction branch receives complete feature representations, including mix proportion information and temporal information. Since strength development is an obvious time-dependent process, this branch needs to learn the strength evolution laws of materials at different curing times. The branch adopts a two-layer fully connected network structure, with the first layer using the ReLU activation function for nonlinear transformation and the second layer directly outputting strength prediction values. Since slump is an instantaneous performance indicator that does not change with time, the slump prediction branch automatically learns to filter time-related information. Finally, to ensure that prediction values are within physically reasonable ranges, the output layer uses a sigmoid activation function and multiplies it by the maximum slump value for constraint.

The PINN–transformer model combines the advantages of transformers in complex feature modeling and sequence dependency mining with PINN's strengths in physical law constraints and interpretability, and it is capable of improving the physical rationality and generalization ability of the model while ensuring prediction accuracy. It can improve prediction accuracy for the relationship between backfill material strength/slump while ensuring physical interpretability, providing a reliable model foundation and quantification means for subsequent material mix proportion optimization and engineering applications.

3.3 Evaluation indicators

To evaluate the performance of the PINN–transformer model in predicting the mechanical properties of backfill materials, this study selects MAE (mean absolute error) and R^2 (coefficient of determination) as the primary indicators, rather than RMSE and MAPE. This is because RMSE is easily affected by a few large errors, and MAPE may produce extreme percentage errors at low-value stages; in comparison, MAE can directly reflect the average deviation between the predicted and true values and is not sensitive to outliers; R^2 can quantify the model's ability to explain the total variation of target variables. Therefore, in this dual-task prediction framework, MAE and R^2 are used to evaluate the prediction accuracy of compressive strength and slump, respectively, comprehensively reflecting model performance on different performance indicators. The calculation formulas are as follows:

MAE reflects the average absolute deviation between the predicted and true values, with the calculation formula:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \quad (11)$$

where y_i is the true value of the ith sample, $\hat{y}_i$ is the model prediction value, and n is the total number of samples. MAE has relatively low sensitivity to outliers and can more robustly reflect the overall prediction error level of the model. The smaller its value, the closer the model prediction results are to actual values.

R^2 represents the proportion of total variation in the target variable explained by the model, with the calculation formula:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}, \quad (12)$$

where $\bar{y}$ is the mean of true values. R^2 has a value range of [0,1], with values closer to 1 indicating better model fitting. When R^2 approaches 1, it indicates that the model can well explain the variability of the data.

4 Experimental results and analysis

4.1 Operating environment and parameter introduction

The research uses Python 3.8 and the PyTorch-GPU 2.0.1 deep learning framework for model development, with CUDA 11.7 for GPU acceleration. The dataset is divided into training, validation, and test sets in a 7:2:1 ratio to effectively evaluate the model's generalization performance.

During training, the Adam optimizer is used with an initial learning rate of 0.001, combined with the learning rate decay strategy to improve convergence stability. The batch size is set to 32, and the maximum number of training epochs is set to 500, with an early stopping strategy to prevent overfitting. All experiments in this study are run in GPU-accelerated environments to ensure the training efficiency of the joint transformer and PINN framework model.

Physical constraint weight was determined using a hybrid strategy. Before training, a small-scale grid search determines the initial range of physical constraint weights. Specifically, for three types of physical loss terms—strength monotonicity constraint, convexity constraint, and slump rheological constraint—candidate weight values [0.1, 0.5, 1.0, and 2.0] are set, respectively, constructing $4 \times 4 \times 4 = 64$ weight combinations. To control computational

TABLE 2 Optimal parameters for physics-informed loss function weights.

Weight parameter	λ_1	λ_2	λ_3
Weight value	1.00	0.63	0.74

costs, each configuration is only trained for 50 epochs, with 3-fold cross-validation to evaluate comprehensive performance on the validation set. The reason for choosing 3-fold cross-validation is that, on one hand, the dataset scale is limited, and each fold needs to ensure sufficient training samples to guarantee model convergence and evaluation reliability; on the other hand, 3-fold cross-validation can provide stable model performance estimates at a reasonable computational cost. Compared to 5-fold or 10-fold cross-validation, it can significantly reduce training times and total computational load, thus avoiding wastage of resources due to repeated calculations. After comparing 64 experimental results, the weight configuration with the highest comprehensive score is selected as the initial value. Experiments found that when the strength monotonicity constraint weight is $\lambda_1 = 1.0$, $\lambda_2 = 0.5$, and $\lambda_3 = 1.0$, the model achieves optimal prediction performance while ensuring physical reasonableness. This combination was selected as the initial weight configuration for subsequent training.

An adaptive weight adjustment mechanism is adopted during training to further optimize the balance of physical constraints. Considering that the strength monotonicity constraint is a hard physical law (violation of which means unreasonableness), its weight λ_1 remains fixed at 1.0 throughout training, while the weights of the convexity constraint and rheological constraint are set as learnable parameters, optimized together with the network parameters. In a specific implementation, an uncertainty weighting method is introduced, expressing λ_2 and λ_3 in the form of inverse log variance, i.e., $\lambda_i = 1/(2\sigma_i^2)$, where σ_i serves as learnable parameters of the neural network and is automatically updated through backpropagation. Additionally, to prevent weight degradation or explosion, interval constraints [0.01 and 10.0] are applied to λ_2 and λ_3 after each parameter update. The training process adopts a three-stage curriculum learning strategy: the first 100 epochs focus on data fitting (physical loss weights decay to 0.5-times the initial values), the middle 200 epochs gradually strengthen physical constraints (weights linearly recover to the initial values), and the final stage (if early stopping is not triggered) further amplifies physical constraint weights to 1.5-times the initial values, strengthening the model's adherence to materials science laws. Throughout the training process, the gradient norms of each loss term are monitored in real-time to ensure that gradient contributions of data loss and physical loss remain of the same order of magnitude, avoiding training imbalance.

The optimal parameter configuration obtained through experiments is shown in Table 2:

Using the optimal parameter configuration, model loss changes and learning rate scheduling during 500 training rounds are shown in Figure 5.

As shown in Figure 5, the learning rate scheduling strategy basically matches the loss value changes. The training process shows obvious phased characteristics. At approximately 380 rounds, the training loss value reaches a convergence state, indicating that the model has basically learned the main features of the data.

4.2 Experimental results

To verify the effectiveness of each component, ablation experiments were conducted, and the results are shown in Tables 3, 4.

The complete PINN–transformer model demonstrates good performance in both prediction tasks. In the strength prediction task, MAE is only 1.09 MPa, with R^2 reaching 0.945. In the slump prediction task, MAE is 0.68 cm, with R^2 of 0.921, which is superior to that of other configurations. Through ablation experiments, it can be concluded that the transformer's global attention mechanism has advantages in capturing complex nonlinear relationships between mix proportion parameters. PINN constraints not only improve prediction accuracy but, more importantly, ensure the physical rationality of prediction results, thus avoiding model outputs that violate the basic principles of material mechanics.

To further verify the robustness and reliability of model prediction performance, residual analysis was performed on strength and slump prediction results, and the results are shown in Figure 6.

From Figure 6, it can be observed that most residuals are concentrated around 0, showing an approximately symmetric normal distribution without obvious bias trends, indicating that the model has no systematic errors. Additionally, as material mix parameters change, residuals show no obvious increasing trend, indicating that the model maintains high prediction stability under different material mix conditions.

Generally speaking, backfill material strength requirements are 5 MPa–8 MPa, with an optimal slump range of 22 cm–25 cm. The trained PINN–transformer model can predict strength from curing ages ranging from 7 days to 56 days under different mix proportions. The strength prediction diagram for a mass concentration of 76% with backfill material of 15% fly ash, 10% cement, and 75% coal gangue is shown in Figure 6.

Figure 7 shows the prediction effect of the PINN–transformer model on material strength development. Backfill material strength increases rapidly in the early stage, reflecting the active period of hydration reactions, and grows slowly in the later stage; overall, it exhibits a monotonically increasing trend. The prediction curve retains the monotonically increasing characteristic due to physical constraints while demonstrating the model's accurate fitting ability, indicating that the PINN–transformer model provides high-precision predictions while maintaining physical consistency.

Partial slump prediction results for different mix proportions at different concentrations are shown in Figure 7.

Figure 8 shows the comparison between the experimental and predicted values of slump under different material mix proportions. From the figure, it can be observed that the model has high prediction accuracy with a reasonable distribution of prediction point errors. As shown in Figure 8, with constant cement content, as the fly ash proportion and solid mass concentration increase, the slump shows a significant downward trend, conforming to the rheological properties of backfill materials, thus verifying the model's reliability.

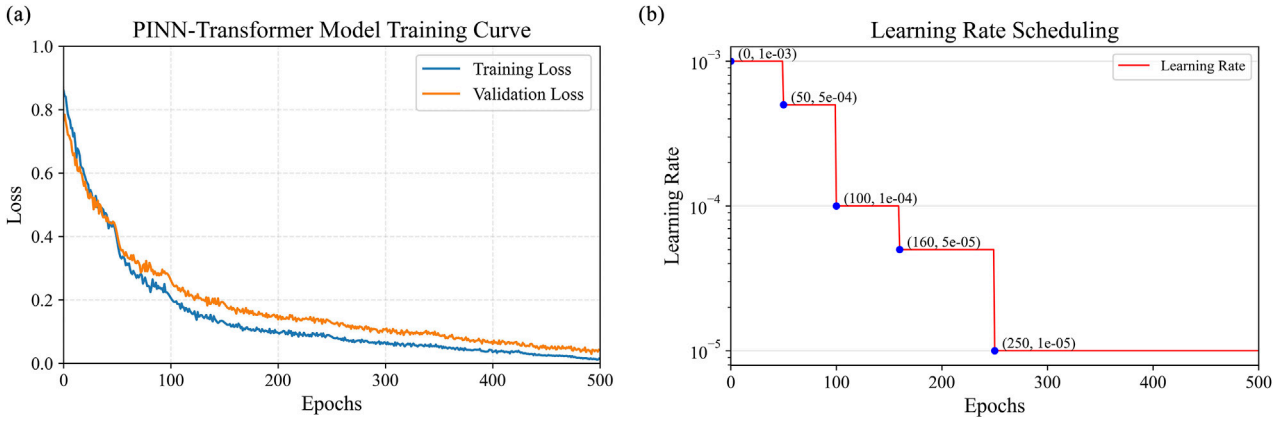

FIGURE 5
Training process diagram. **(a)** Model loss evolution diagram. **(b)** Learning rate scheduling diagram.

TABLE 3 Comparison of strength attenuation results in backfill materials.

Model configuration	Transformer	PINN constraint	Positional encoding	MAE (MPa)	R^2
Complete model	✓	✓	✓	1.09	0.945
Without PINN constraint	✓	×	✓	1.16	0.938
Without transformer	×	✓	×	1.42	0.912
Without positional encoding	✓	✓	×	1.13	0.941

TABLE 4 Comparison of slump attenuation results in backfill materials.

Model configuration	Transformer	PINN constraint	Positional encoding	MAE (MPa)	R^2
Complete model	✓	✓	✓	0.68	0.921
Without PINN constraint	✓	×	✓	0.62	0.913
Without transformer	×	✓	×	0.76	0.893
Without positional encoding	✓	✓	×	0.61	0.961

Model performance comparison results are shown in Tables 5, 6.

Based on the experimental results in Tables 6, 7, the analysis shows that in the backfill material strength prediction task, models show obvious performance gradients. For the PINN–transformer model, in terms of strength prediction, its MAE indicator is improved by 31.0%, 60.5%, 6.0%, and 23.2% compared to that of LSTM, ANN, transformer, and PINN, respectively. Compared to the transformer, R^2 improved from 0.938 to 0.945. In terms of slump prediction, the improvement in MAE reached 31.8%, 51.3%, 6.5%, and 23.7%, respectively. Compared to the transformer, R^2 improved from 0.913 to 0.921. LSTM's sequence modeling characteristics do not match the modeling requirements of mix proportion parameters, resulting in an MAE of 1.58 MPa. Traditional ANN has the smallest number of parameters due to its concise network structure, but it is limited by local connection characteristics and cannot fully capture global dependency relationships between mix proportion parameters. The pure transformer architecture significantly improves prediction performance through the global attention mechanism, verifying its advantages in modeling multi-variable relationships. Traditional PINN methods, although introducing physical constraints, have limited feature extraction capabilities based on fully connected layers, resulting in decreased prediction accuracy. The PINN–transformer model performs optimally on all evaluation indicators, indicating that the introduction of physical constraints significantly improves the model's prediction accuracy and generalization ability. The constraint mechanism of PINNs and the transformer's global modeling capability form an effective synergistic effect, ensuring both prediction accuracy and physical rationality of the results. This indicates that the PINN–transformer model can accurately

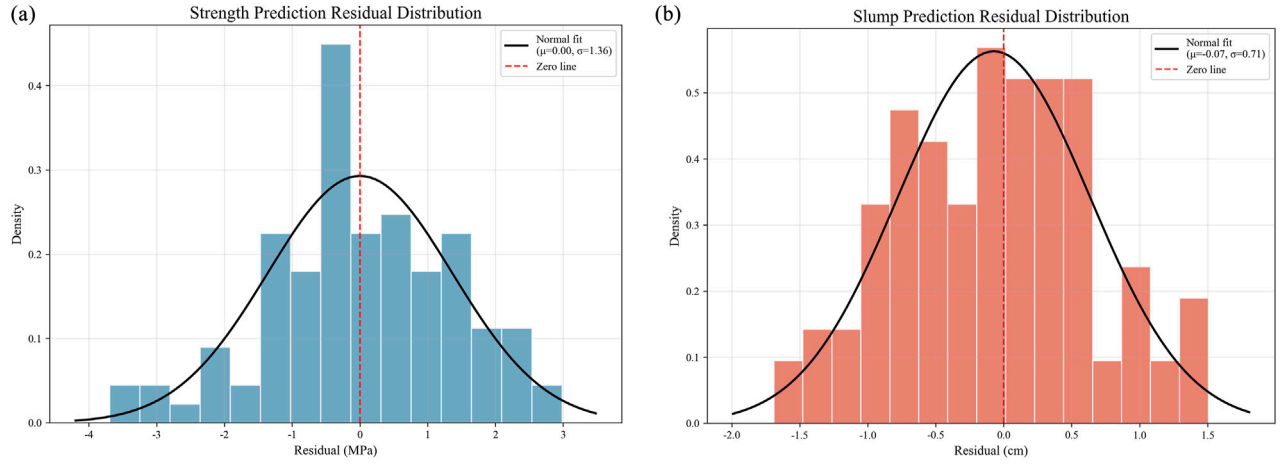

FIGURE 6
Prediction residual plot. **(a)** Strength residual plot. **(b)** Slump residual plot.

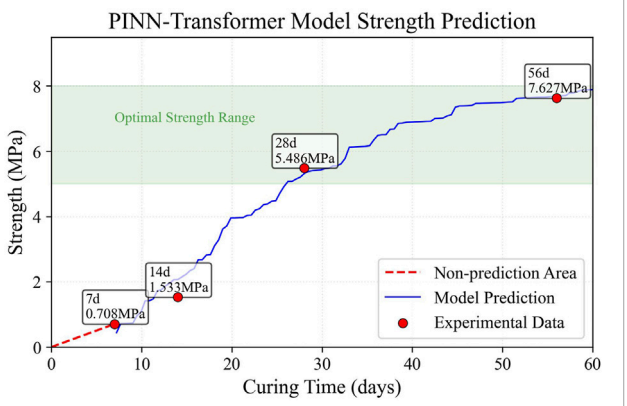

FIGURE 7
Backfill material strength prediction diagram at different curing times.

predict backfill material performance, providing reliable workability guidance for field construction.

4.3 Economic analysis of backfill materials

After obtaining full-age strength, slump, and unit cost data predicted by the PINN–transformer model, the core objective of economic analysis in this study is to establish a multi-objective optimization screening system under the premise of strictly satisfying the hard indicators of material performance. The strength indicator is set at 5 MPa–8 MPa, mainly based on the load-bearing requirements of mine backfill support and the standard range of backfill body compressive capacity in previous engineering practice, where 5 MPa is the minimum safe load-bearing lower limit, ensuring that the backfill body does not fail during construction and early service stages, while 8 MPa is the recommended upper limit in actual engineering to avoid raw material waste and increased construction costs caused by excessive material strength (Belem and Benzaazoua, 2008). The slump indicator is set at 22 cm–25 cm, referencing current pumping process requirements for slurry flowability. This range ensures that the slurry has good pumpability and construction adaptability while avoiding pumping blockage from too low flowability or settlement and segregation risks from too high flowability (Sivakugan et al., 2006). Based on this, a multi-objective optimization framework is constructed to screen the optimal mix schemes with the lowest unit cost and shortest possible curing period, thereby achieving coordination of safety, construction adaptability, and economic benefits and providing a scientific and reliable decision-making basis for mine backfill material design. Based on the field research of the Yulin area building materials market in 2025, combined with the latest price trend analysis of national fly ash and cement industries, prices are as follows: fly ash, 110 yuan/ton; cement, 397 yuan/ton; coal gangue, 6 yuan/ton; and water, 2 yuan/ton. The curing cost per ton of material is time cost, with no specific value estimate, but generally, the shorter the time cost, the better. Ideally, the cost of each ton of backfill material and its corresponding curing age can be calculated.

The first step of economic analysis is to establish clear performance admission standards. All mix proportion schemes that are compared must simultaneously meet the following two rigid conditions: 1. strength criterion: predicted compressive strength at the specified curing age (such as 28 days) must not be lower than 5.0 MPa; 2. workability criterion: predicted slump must strictly be within the range of 22 cm–25 cm. Exceeding this range is considered to compromise pumpability and is therefore not allowed.

Then, a multi-objective optimization decision-making framework is constructed. For qualified schemes that pass the initial screening, their quality evaluation needs to be incorporated into a comprehensive decision-making framework. The primary optimization objective is cost minimization: directly comparing the unit material costs (yuan/ton) of each scheme. Based on the premise of equivalent performance, the lower the cost, the better. The secondary objective is the shortest curing period: when cost differences between schemes are not significant, priority is given to

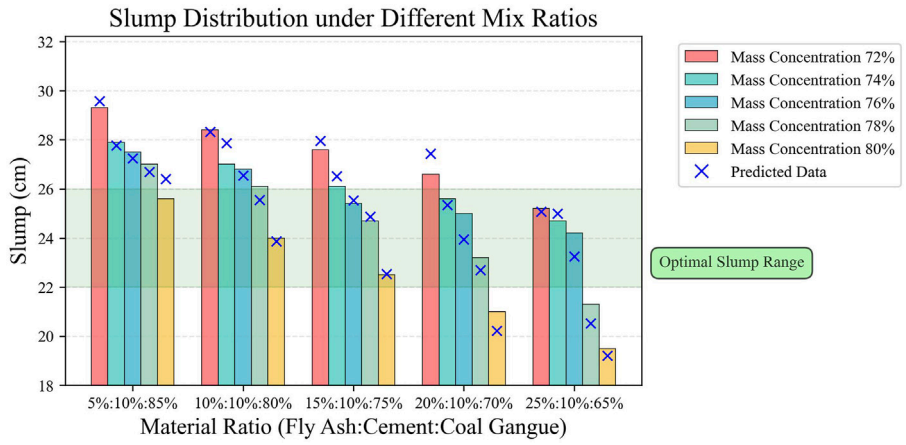

FIGURE 8
Slump prediction results' diagram for backfill materials at different concentrations and proportions.

TABLE 5 Performance comparison of different models (backfill material strength).

Model	Model parameter	MAE (MPa)	R^2
LSTM	1.8 M	1.58	0.913
ANN	0.5 M	2.76	0.864
Transformer	2.0 M	1.26	0.929
PINN	1.2 M	1.83	0.890
PINN–transformer	2.1 M	1.09	0.945

TABLE 6 Performance comparison of different models (backfill material slump).

Model	Model parameter	MAE (cm)	R^2
LSTM	1.8 M	0.85	0.897
ANN	0.5 M	1.19	0.873
Transformer	2.0 M	0.64	0.906
PINN	1.2 M	1.07	0.881
PINN–transformer	2.1 M	0.68	0.921

schemes that can meet strength design requirements at a shorter curing age.

Cost deconstruction of the qualified schemes reveals that, given the extremely low unit prices of coal gangue and water, the total backfill material cost is mainly dominated by cement (397 yuan/ton) and fly ash (110 yuan/ton) usage. Therefore, under the constraints of meeting the final strength and construction workability, it is necessary to partially replace expensive cement with lower-cost fly ash as much as possible. At a fixed concentration, slump shows a trend of first slightly increasing and then decreasing with increasing fly ash proportion (ball bearing effect at low dosage and water absorption effect at high dosage). After strictly limiting slump to the "optimal workability range" of 22 cm–25 cm and analyzing mix proportion data, it is found that at solid mass concentration of approximately 76%, slumps of most mix proportions can meet backfill requirements. Within this range, after analyzing the relationship between strength and cost, Table 7 lists the representative high cost-effective mix proportions and their economic indicators screened for different age strength targets under the premise of meeting slump requirements.

Cost comparison clearly shows that strategy 3, maximizing the use of inexpensive fly ash (110 yuan/ton) to replace expensive cement (397 yuan/ton), becomes the most cost-optimal choice, and its predicted strength and slump can meet engineering requirements. If engineering requires robust and rapid development of slurry strength, strategy 1 should be considered as it may provide a good balance between cost and performance. The selection of backfill mix proportions is essentially a decision-making process of trade-offs between "material cost," "time cost," and "engineering requirements." Mines need to select corresponding baseline mix proportions based on specific production plans, mine bed value, and safety specifications, using prediction models for fine-tuning to maximize the overall project economic benefits.

5 Conclusion

5.1 Application prospects

The PINN–transformer model proposed in this study has significant engineering application potential, especially with broad application prospects in mine backfill operations. By combining physical constraints with data-driven modeling, this model can make accurate predictions in complex nonlinear relationships of multi-dimensional parameters, providing a scientific basis for the optimal mix design of backfill materials. In the future, as technology

TABLE 7 Economic analysis of optimal proportions based on different age strength targets.

Strategy	Recommended proportion range (solid concentration: 76%)	Key age predicted strength (MPa)	Slump (cm)	Cost level (RMB/ton of slurry)
1	Cement: 11%–12% Fly ash: 15%–18% Coal gangue: 70%–74%	14 d: 2.7–3.5 28 d: 5.8–6.4 56 d: 7.3–8.1	23.2–24.7	49.58–54.93
2	Cement: 13%–14% Fly ash: 10%–12% Coal gangue: 74%–77%	14 d: 3.5–4.5 28 d: 7.0–7.8 56 d: 8.2–9.0	22.6–24.1	51.57–56.12
3	Cement: 10%–11% Fly ash: 18%–20% Coal gangue: 69%–72%	14 d: 1.8–2.5 28 d: 5.2–5.8 56 d: 6.0–7.0	23.9–24.7	48.98–53.54

matures and application fields expand, the model may be widely applied in the following aspects:

1. Through accurate material performance prediction, the model can help mining enterprises optimize backfill material mix design, reduce resource waste, and improve the economy and reliability of backfill operations.
2. Although this study mainly focuses on mine backfill, this method has good extensibility and can be applied to other material fields in the future, such as building materials and chemicals, especially in scenarios involving complex physical and chemical characteristics and multi-objective optimization.
3. By reducing trial and error and resource waste in experimental processes, this model can promote the green and sustainable development of mine backfill operations.

5.2 Summary and outlook

This paper proposes a PINN–transformer method integrating physical constraints for backfill material mix optimization. The model achieves effective fusion of physical constraints and data-driven modeling, improving the accuracy of strength and slump prediction under different backfill material mix proportions. Through the synergistic effect of physical constraints and global attention mechanisms, this method can accurately capture complex nonlinear relationships between material mix, curing conditions, and performance indicators. Experimental results show that the PINN–transformer model performs excellently in strength and slump prediction (strength MAE 1.09 MPa, R^2 0.945; slump MAE 0.58 cm, R^2 0.921), verifying its reliability in multi-objective performance prediction tasks. By embedding physical constraints into the transformer's sequence modeling process, the model can simultaneously capture high-dimensional coupling relationships between material components, curing conditions, and performance indicators. The synergistic effect of physical constraints and global attention mechanisms provides a scientific basis for optimal backfill material mix design, thereby effectively reducing raw material consumption and experimental costs and improving the economy and reliability of mine backfill operations.

Research confirms that with accurate predictions from the PINN–transformer model, quantitative optimization from empirical mix to data-driven can be achieved, but certain limitations still exist. The model's generalization capability in high-dimensional mix spaces or extreme conditions needs further verification; current physical constraints mainly cover strength monotonicity, convexity, and slump rheological laws, without fully covering complex chemical reactions and construction environment factors, while the experimental dataset scale is relatively limited, which may affect prediction accuracy for extreme mix proportions.

Future research can further introduce multi-scale physical constraint mechanisms and uncertainty quantification methods to improve model robustness, adopt adaptive loss weight strategies to optimize the training process, and combine actual engineering cases with cost–benefit analysis to further verify the engineering applicability and economic feasibility of the method. This research provides a feasible path for the selection of backfill materials from empirical mix to intelligent data-driven optimization, thus having important reference value for efficient, safe, and green development of mine backfill engineering.

Data availability statement

The datasets presented in this article are not readily available because the data that support the findings of this study are available from the corresponding author upon reasonable request. Requests to access the datasets should be directed to Yueying Zhang lhzlhz1028@163.com.

Author contributions

HL: Data curation, Formal Analysis, Investigation, Writing – original draft. YZ: Conceptualization, Methodology, Supervision, Writing – original draft. CX: Investigation, Software, Visualization, Writing – original draft. SF: Investigation, Software, Visualization, Writing – review and editing. HW: Software, Writing – review and

editing. BH: Software, Writing – review and editing. AZ: Writing – review and editing.

Funding

The author(s) declared that financial support was not received for this work and/or its publication.

Acknowledgements

The authors would like to thank the Xi'an Kedagaoxin University.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Antonion, K., Wang, X., Raissi, M., and Chen, Y. (2024). Machine learning through physics-informed neural networks: progress and challenges. *Acad. J. Sci. Technol.* 9, 46–49. doi:10.54097/b1d21816

Belem, T., and Benzaazoua, M. (2008). Design and application of underground mine paste backfill technology. *Geotech. Geol. Eng.* 26, 147–174. doi:10.1007/s10706-007-9167-y

Bo, L., Yang, S., Liu, Y., Wang, Y., and Zhang, Z. (2023). Research on the data validity of a coal mine solid backfill working face sensing system based on an improved transformer. *Sci. Rep.* 13, 1. doi:10.1038/s41598-023-38365-6

Chang, J. W., Du, G. J., Du, J. L., and Shi, X. L. (2022). Current situation of the comprehensive utilization of coal gangue in China and the related problems and recommendations. *China Environ. Prot. Ind.* 8, 13–17.

Chen, X., Shi, X., Zhou, J., Bürger, R., Wang, Y., Wang, S., et al. (2022). Optimization of parameters for rheological properties and strength of cemented paste backfill blended with coarse aggregates. *Minerals* 12, 374. doi:10.3390/min12030374

Cihangir, F., Akyol, Y., Gu, X., Yilmaz, E., Fang, K., and Jiang, H. (2022). Strength analysis and optimization of alkali activated slag backfills through response surface methodology. *Front. Mater.* 9, 844608. doi:10.3389/fmats.2022.844608

Ercikdi, B., Yılmaz, T., and Külekci, G. (2014). Strength and ultrasonic properties of cemented paste backfill. *Ultrasonics* 54, 195–204. doi:10.1016/j.ultras.2013.04.013

Fernández de la Mata, F., Gijón, A., Molina-Solana, M., and Gómez-Romero, J. (2023). Physics-informed neural networks for data-driven simulation: advantages, limitations, and opportunities. *Phys. A Stat. Mech. Appl.* 610, 128415. doi:10.1016/j.physa.2022.128415

Ghirian, A., and Fall, M. (2013). Coupled thermo-hydro-mechanical-chemical behaviour of cemented paste backfill in column experiments. Part I: physical, hydraulic and thermal processes and characteristics. *Eng. Geol.* 164, 195–207. doi:10.1016/j.enggeo.2013.01.015

Ghojogh, B., and Ghodsi, A. (2020). Attention mechanism, transformers, BERT, and GPT: tutorial and survey. *arXiv*. arXiv:2009.14794.

Liu, J. G., Li, X. W., and He, T. (2020). Application status and prospect of backfill mining in Chinese coal mines. *J. China Coal Soc.* 45, 141–150.

Liu, G., Li, L., Yang, X., and Guo, L. (2021). Stability analyses of side-exposed backfill considering mine depth and extraction of adjacent stope. *Int. J. Min. Sci. Technol.* 31, 307–318.

Michaloglou, A., Papadimitriou, I., Gialampoukidis, I., Vrochidis, S., and Kompatsiaris, I. (2025). Physics-informed neural networks in materials modeling and design: a review. *Arch. Comput. Methods Eng.* (in press). doi:10.1007/s11831-025-10448-9

Qi, C., Fourie, A., Chen, Q., and Zhang, Q. (2018). A strength prediction model using artificial intelligence for recycling waste tailings as cemented paste backfill. *J. Clean. Prod.* 183, 566–578. doi:10.1016/j.jclepro.2018.02.154

Shao, X., Wang, L., Li, X., Fang, Z., Zhao, B., Tao, Y., et al. (2020). Study on rheological and mechanical properties of aeolian sand-fly ash-based filling slurry. *Energies* 13, 1266. doi:10.3390/en13051266

Sivakugan, N., Rankine, R. M., Rankine, K. J., and Rankine, K. S. (2006). Geotechnical considerations in mine backfilling in Australia. *J. Clean. Prod.* 14, 1168–1175. doi:10.1016/j.jclepro.2004.06.007

Song, X., Chen, K., Bi, Z., and Niu, Q. (2025). Transformer: a survey and application. Available online at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5211988 (Accessed September 27, 2025).

Tang, P., Mo, S., Deng, N., Li, Z., and Lu, C. (2025). A physics-informed machine learning framework for predicting the failure pressure of buried PVC pipelines with axial cracks. *Eng. Fail. Anal.* 180, 109896. (in press). doi:10.1016/j.engfailanal.2025.109896

Taraghi, P., Li, Y., Yoosef-Ghodsi, N., Fowler, M., Kainat, M., and Adeeb, S. (2023). "Response of buried pipelines under permanent ground movements: physics-informed deep neural network approach," in Proceedings of the ASME 2023 Pressure Vessels and Piping Conference, Atlanta, GA, USA (New York, NY: ASME), V002T03A020. doi:10.1115/PVP2023-106201

Uddin, M. N., Alavi, A. H., and Zhang, C. (2025). Advancements in physics-informed neural networks for laminated composites: a comprehensive review. *Mathematics* 13, 17. doi:10.3390/math13010017

Wang, X., Li, J., Zhang, S., Gao, H., and Liu, W. (2024). A CFD-based numerical model for predicting the slump and slump flow of fresh concrete from a rheological perspective. *Constr. Build. Mater.* 453, 139052.

Wang, J. D., Huang, L., Liang, Y., Dong, J., and Sun, W. (2025a). Transfer learning enhanced physics-informed neural network for buried pipeline deformation analysis under permanent ground deformation. *Comput. Geotech.* 189, 107630. doi:10.1016/j.compgeo.2025.107630

Wang, Z., Tian, G., Zhang, J., Wang, X., and Li, M. (2025b). Hydration properties of activated "coal gangue-fly ash-cement" ternary compound cementitious material. *J. Min. Sci. Technol.* 10, 738–747.

Wu, D., Zhao, R., Xie, C., and Liu, S. (2020). Effect of curing humidity on performance of cemented paste backfill. *Int. J. Min. Metall. Mater.* 27, 1046–1053. doi:10.1007/s12613-020-1970-y

Xu, C., Wang, Y., and Barati Farimani, A. (2023). TransPolymer: a transformer-based language model for polymer property predictions. *Npj Comput. Mater.* 9, 64. doi:10.1038/s41524-023-01016-5

Zhao, X., Yang, K., He, X., Wei, Z., Zhang, J., and Yu, X. (2024). Mix proportion and microscopic characterization of coal-based solid waste backfill material based on response surface methodology and multi-objective decision-making. *Sci. Rep.* 14, 5556. doi:10.1038/s41598-024-56028-y