Journal of Network Intelligence
Taiwan Ubiquitous Information
 ISSN 2414-8105
Volume **10**, Number **4**, November **2025**

Inversion of Surrounding Rock Creep Parameters Based on Improved Echo State Network

Cong-Cong Meng*

College of Road and Bridge Engineering
Jiangxi Vocational and Technical College of Communications
Nanchang 330013, P. R. China
mc1219@126.com

Hai-Long Liu

Department of Project Management
Jiangxi Province Expressway Asset Management Co., Ltd
Nanchang 330108, P. R. China
liuhlong8299@126.com

Kan Huang

National University of the Philippines
Manila 0900, Philippines
23585554@qq.com

*Corresponding author: Cong-Cong Meng
Received September 5, 2024, revised March 5, 2025, accepted August 24, 2025.

Abstract. *Traditional inversion methods for the creep parameters of the surrounding rock often suffer from the problems of low accuracy and limited generalisation ability, which make it difficult to accurately describe the creep characteristics of the rock mass under the complex engineering environment. With the continuous development of deep-buried tunnel engineering, there is an increasing demand for inversion methods that can effectively deal with nonlinear dynamic features and have good generalisation ability. In order to solve these problems, this work proposes an inversion method based on improved Echo State Network (ESN) for the creep parameters of surrounding rock. Firstly, by combining the improved ESN structure with Convolutional Neural Network (CNN), the ability of the model to capture the nonlinear dynamic features of the creep behaviour of the surrounding rock is improved. Secondly, a hybrid activation function strategy combining Leaky ReLU and tanh functions is adopted to enhance the nonlinear modelling capability of the model. In addition, regularisation and sparsity constraints are introduced to significantly improve the generalisation ability of the model. Finally, the applicability of the model in the case of limited data is improved by the idea of reserve pooling computation. The experimental results show that the improved ESN model outperforms the comparison model in several aspects. In short-term (30 days) and long-term (365 days) forecasts, the average relative errors of the improved ESN model are 6.2% and 10.2%, respectively. In terms of generalisation performance, the average prediction accuracy of the improved ESN model reached 87.4% on five typical rock types. In terms of computational efficiency and model complexity, the improved ESN model shows a good balance. The sensitivity analysis shows that the prediction error of the improved ESN model increases by no more than 5% over the range of rock body parameter variations, which is better than other models and proves its high robustness.*
Keywords: enclosure creep; parametric inversion; echo state network; convolutional neural network; deep learning

1. **Introduction.** Deep-buried tunnelling is widely used in the fields of water conservancy and hydropower, transportation, and mining, but it faces serious safety challenges during its construction and operation. The creep behaviour of the surrounding rock is one of the key factors affecting the long-term stability of deeply buried tunnels [1, 2], and accurate prediction and assessment of the creep characteristics of the surrounding rock is crucial to ensure the safety and durability of the tunnels. However, due to the complexity and uncertainty of the deep-buried environment, the traditional methods for obtaining creep parameters often suffer from the problems of low accuracy and low efficiency [3, 4], which are difficult to meet the needs of engineering practice.

Artificial intelligence techniques, especially machine learning and neural network techniques, show great potential for application in solving complex nonlinear problems. In the field of geotechnical engineering, these techniques can be used to process large amounts of monitoring data, identify deformation patterns under complex geological conditions, and perform parameter inversion and prediction [5]. Machine learning algorithms are able to learn from historical data and automatically identify potential patterns and laws of creep behaviour without the need to pre-assume specific mathematical models [6]. Neural network techniques, such as recurrent neural networks and convolutional neural networks, on the other hand, can efficiently process time-series data and extract key features, which is particularly useful for analysing long-term creep behaviour. These AI methods not only improve the accuracy and efficiency of parameter inversion, but also help engineers better understand complex geological conditions and predict potential risks. With the continuous improvement of computing power and the continuous optimisation of algorithms, the application of AI technology in geotechnical engineering will have a broader prospect [7, 8], which is expected to provide more reliable theoretical support and technical guarantee for the design, construction and operation of deep- buried tunnels.

1.1. **Related work.** The traditional methods of creep parameter inversion for surrounding rocks mainly include laboratory test method, field test method and inverse analysis method. Li et al. [9] proposed a parameter inversion method based on triaxial creep test, in which the creep parameters of rocks are obtained by controlling the creep process under different stress levels. It is shown that the method can better reflect the creep properties of rocks, but the sample size effect and the difference between the experimental conditions and the actual engineering environment may lead to the under-representation of the parameters. Yang et al. [10] developed an in situ flat plate load creep test method to invert the creep parameters by monitoring the rock deformation over a long period of time. This method overcomes the dimensional limitations of laboratory tests, but the test period is long, costly, and limited by field conditions. Wang et al. [11] proposed a creep parameter inversion method based on displacement inverse analysis, which uses the monitoring data during tunnel excavation to invert the rock mass parameters. This method can consider the actual engineering conditions, but the selection of the optimisation algorithm and the setting of the initial parameters have a large impact on the inversion results. Li et al. [12] proposed an improved inverse analysis method by combining the finite element method and genetic algorithm. The method improves the efficiency of parameter inversion, but the computational cost is still high when dealing with large-scale monitoring data. In general, the traditional methods have made some progress in parameter inversion, but they still face problems such as low accuracy, low efficiency, limited applicability, etc., and are difficult to meet the needs in complex engineering environments.

In recent years, the main direction of the research on the parameter inversion of surrounding rock creep is focused on the application of intelligent algorithms and the development of data- driven methods, in which the introduction of deep learning technology

provides new ideas for solving this complex problem. Chen et al. [13] constructed a creep model for weak surrounding rocks in tunnels and used the BPNN model for parameter inversion. Li et al. [14] developed a Bayesian Optimisation-based Random Forest (BO-RF) model for the inversion of the creep parameters of the surrounding rock. It is shown that the method performs well in feature extraction and parameter prediction, but has limitations in dealing with nonlinear dynamic features. Zhang et al. [15] proposed a creep parameter prediction model based on the long- Short-Term Memory Network (LSTM), which is able to effectively capture long-term dependencies in time-series data, but has some limitations in dealing with high-dimensional features. To address these problems, Echo State Network (ESN) [16, 17], as a special recurrent neural network with advantages of simple structure, fast training, and the ability to effectively handle time series data, is expected to be an effective tool for solving the problem of creep parameter inversion in enclosing rocks.The introduction of ESN may alleviate the limitations of the existing deep learning methods in terms of computational efficiency, data requirements, and model complexity for the The introduction of ESN may alleviate the limitations of the existing deep learning methods in terms of computational efficiency, data requirements and model complexity, and provide new ideas and methods for the inversion research of surrounding rock creep parameters.

1.2. **Motivation and contribution.** Existing inversion methods for the creep parameters of surrounding rock usually face the problems of poor accuracy and limited generalisation capability. Traditional laboratory and field test methods are difficult to fully reflect the rock creep characteristics in complex engineering environments, while inversion methods based on numerical analysis are often constrained by the limitations of optimisation algorithms. Existing methods still need to be improved in dealing with nonlinear dynamic features and improving model generalisation ability. In particular, when dealing with complex surrounding rock creep behaviour, how to effectively capture its nonlinear features and improve the performance of the model has become an urgent problem to be solved.

In order to solve the above problems, this paper proposes a parameter inversion method based on Improved Echo State Network (ESN) for perimeter rock creep, which significantly improves the accuracy and generalisation ability of parameter inversion. The main innovations and contributions of this work include:

1. Aiming at the limitations of existing methods in dealing with nonlinear dynamic features, this paper proposes an improved ESN structure incorporating convolutional neural network (CNN). This innovative structure is able to extract and utilize the spatio-temporal features in the input data more effectively, which significantly improves the model's ability to characterise the creep behaviour of the surrounding rock. In particular, the method demonstrates significant advantages in capturing complex nonlinear dynamic features.
2. In order to enhance the nonlinear modelling ability of the model, this paper proposes a hybrid activation function strategy that combines Leaky ReLU and tanh functions. This improvement not only effectively alleviates the problem of gradient vanishing in traditional ESNs, but also enhances the model's ability to express complex nonlinear relationships, thus improving the accuracy of parameter inversion.
3. To address the problem that ESN models are prone to overfitting when dealing with high-dimensional data, regularisation and sparsity constraints are introduced in this paper. The application of these techniques significantly improves the generalisation ability of the model, which enables the inversion method to better adapt to different

rock types and engineering conditions, and enhances the performance of the model on unseen data.

2. Fundamentals of rock creep.

2.1. **Creep theory.** Creep is a key phenomenon in rock mechanics [18] and describes the process of continuous deformation of rocks with time under constant stress. This time-dependent deformation has significant implications for the long-term stability of deep rock engineering. In this section, the basic concepts and characteristics of creep theory and its importance in rock mechanics will be elaborated.

Creep is defined as the phenomenon of increasing strain with time in a material under constant stress [19]. For rock materials, significant creep deformation may occur even at lower stress levels. The main characteristics of rock creep include time-dependence, i.e., deformation develops continuously over time rather than instantaneously; nonlinearity, as evidenced by the fact that strain versus time is usually nonlinear; the existence of a stress threshold, i.e., the creep effect is not significant when the stress is below a certain critical level; and temperature sensitivity, i.e., the creep process is accelerated by an increase in temperature.

The typical rock creep process can be divided into three stages:

1. Instantaneous elastic deformation: the elastic deformation produced instantaneously after loading.
2. A creep: also known as the deceleration creep stage. The rate of deformation decreases with time, and the curve is convex.
3. Secondary creep: also known as the steady state creep stage. The rate of deformation is approximately constant and the curve is approximately linear.
4. Three creep: also known as the accelerated creep stage. The rate of deformation gradually increases, eventually leading to destruction.

Creep strain can be expressed as a function of time [20]:

$$\varepsilon(t) = \varepsilon_0 + \varepsilon_1(t) + \varepsilon_2(t) + \varepsilon_3(t), \tag{1}$$

where ε_0 is the instantaneous elastic strain, and $\varepsilon_1(t)$, $\varepsilon_2(t)$, and $\varepsilon_3(t)$ are the primary, secondary and tertiary creep strains, respectively.

The creep behaviour of rocks is affected by a number of factors. Firstly, the stress level plays a significant role in creep deformation, and the higher the stress, the more obvious the creep phenomenon. Secondly, different rock types exhibit different creep behaviour. Temperature is also an important factor, and generally speaking, an increase in temperature accelerates the creep process. In addition, the water content of the rock is also important, and an increase in water content tends to exacerbate the creep effect. Changes in confining pressure can also affect creep behaviour, and usually an increase in confining pressure can inhibit creep deformation. Finally, structural features within the rock, such as structural defects like microfractures and pores, also have an important influence on the creep behaviour.

Creep rate is an important parameter describing the creep process and is defined as the strain increment per unit time:

$$\dot{\varepsilon} = \frac{d\varepsilon}{dt}, \tag{2}$$

where $\dot{\varepsilon}$ is the creep rate, ε is the strain and t is the time. In different creep stages, the creep rate shows different characteristics of change: the rate of the first creep stage decreases, the rate of the second creep stage is nearly constant, and the rate of the third creep stage increases.

Understanding and predicting the creep behaviour of rocks is critical for many rock engineering applications, such as long-term stability assessment of deep tunnels, settlement prediction of high-rise buildings and settlement prediction of high-rise buildings and dam foundations. Settlement prediction of high-rise buildings and dam foundations. In summary, creep theory provides an important foundation for understanding and predicting the time- dependent deformation of rocks. Equation (1) and Equation (2) allow us to quantitatively describe the strain development and rate changes during creep.

2.2. **Creep models.** In order to describe and predict the creep behaviour of rocks more accurately, various creep models have been proposed by researchers. These models are designed to portray the creep behaviour of rocks under different conditions through mathematical expressions. In this section, several commonly used creep models, including empirical and rheological models, are introduced. Empirical models are mathematical expressions based on the fitting of a large amount of experimental data, they are usually simple and intuitive, but lack clear physical meaning [21]. The following empirical models are commonly used. The power function model is one of the simplest models of creep and is expressed as

$$\varepsilon(t) = At^n \tag{3}$$

where A and n are model parameters. This model is mainly used to describe the primary creep phase. The logarithmic function model is commonly used to describe the primary and secondary creep phases, and is expressed as follows.

$$\varepsilon(t) = a + b\log(t) \tag{4}$$

where a and b are model parameters.

The exponential function model is suitable for describing the three creep phases and is expressed as follows.

$$\varepsilon(t) = A\left(1 - e^{-\alpha t}\right) + Bt \tag{5}$$

where B are model parameters.

Rheological models are based on a combination of mechanical components that have a well-defined physical meaning. These models usually consist of a combination of basic elements such as springs (for elasticity), dampers (for viscosity) and sliders (for plasticity). Several commonly used rheological models are described below. The Maxwell model consists of a spring and a damper in series, and its intrinsic equations are.

$$\dot{\varepsilon} = \frac{\dot{\sigma}}{E} + \frac{\sigma}{\eta} \tag{6}$$

where $\dot{\varepsilon}$ is the strain rate, $\dot{\sigma}$ is the stress rate, E is the modulus of elasticity, and η is the coefficient of viscosity. Maxwell's model can describe both instantaneous elastic deformation and infinite creep. The Kelvin model consists of a spring and a damper connected in parallel, and its intrinsic equations are.

$$\sigma = E\,\varepsilon + \eta\,\dot{\varepsilon} \tag{7}$$

The Kelvin model can describe delayed elastic deformation, but not instantaneous elastic deformation and infinite creep. The Burgers model is a tandem combination of the Maxwell and Kelvin models with the following intrinsic equations.

$$\varepsilon(t) = \frac{\sigma_0}{E_1} + \frac{\sigma_0 t}{\eta_1} + \frac{\sigma_0}{E_2}\left(1 - e^{-E_2 t/\eta_2}\right) \tag{8}$$

where σ_0 is the constant stress; E_1 and η_1 are the modulus of elasticity and coefficient of viscosity of the Maxwell unit; E_2 and η_2 are the modulus of elasticity and coefficient of viscosity of the Kelvin unit, respectively. The Burgers model is able to describe the three-stage creep behaviour of the rock better. The generalised Kelvin model, which consists of multiple Kelvin cells connected in series, provides a more accurate description of the complex creep behaviour.

$$\varepsilon(t) = \frac{\sigma_0}{E_0} + \sum_{i=1}^{n} \frac{\sigma_0}{E_i}\left(1 - e^{-E_i t/\eta_i}\right) \tag{9}$$

where E_0 is the instantaneous elastic modulus; E_i and η_i are the modulus of elasticity and coefficient of viscosity of the i-th Kelvin cell respectively; n is the number of Kelvin cells. These creep models provide us with mathematical tools to describe the creep behaviour of rocks. In practical applications, it is necessary to choose the appropriate model according to the specific rock type and engineering conditions. In the next section, we will discuss the scope of application of these models and model identification methods to lay the foundation for the subsequent parameter inversion.

2.3. **Selection of Creep Models.** We have introduced several commonly used creep models. However, each model has its specific scope of application and limitations. Selection of an appropriate model is crucial for an accurate description of the creep behaviour of rocks. In this section, we will discuss the scope of application of different creep models and introduce the methods of model identification.

Empirical models are widely used in engineering practice due to their simplicity and flexibility [22]. However, they usually can only describe a particular stage of the creep process. For example, power function models are mainly suitable for describing the primary creep stage. Logarithmic function models are suitable for describing primary and secondary creep stages. Exponential function models can fit the third creep stage better. The main advantage of these models is that they have few parameters and are simple to compute. However, they lack clear physical significance and are difficult to reflect the internal mechanical mechanisms of rocks.

The rheological model is based on the combination of mechanical components, which has a clear physical meaning and can describe the creep behaviour of rocks more comprehensively [23]. Maxwell's model is suitable for describing the transient elastic deformation and infinite creep, but cannot reflect the delayed elastic deformation. Kelvin's model can describe the delayed elastic deformation, but cannot reflect the transient elastic deformation and the steady state creep. Burgers' model can describe the three-stage creep behaviour of rocks better than the Kelvin model, which is suitable for many rock types. The Burgers model can better describe the three-stage creep behaviour of rocks and is applicable to many rock types. The generalised Kelvin model can describe the complex creep behaviour more accurately by increasing the number of Kelvin cells, but at the same time it increases the complexity of the model and the number of parameters. Among many creep models, the Burgers model was chosen as the main creep model in this study for several reasons:

(1) Completeness: The Burgers model is capable of describing the three stages of rock creep (primary, secondary and tertiary creep). As shown in Equation (8), the Burgers model incorporates transient elastic deformation (σ_0/E_1), steady-state creep ($\sigma_0 t/\eta_1$) and delayed elastic deformation. This completeness allows the Burgers model to provide a more comprehensive description of the creep behaviour of rocks.

(2) Physical significance: As a rheological model, each parameter of the Burgers model has a clear physical significance. This helps us to understand the mechanical mechanism inside the rock and provides a physical explanation for the parameter inversion results.

(3) Applicability: The Burgers model has been widely used to describe the creep behaviour of a wide range of rock types, including soft, hard and salt rocks. This wide applicability makes our results more generalisable.

(4) Balance of complexity: Compared with the simple Maxwell model or Kelvin model, the Burgers model can more accurately describe the complex creep behaviour. Meanwhile, compared with the generalised Kelvin model, the Burgers model has fewer parameters, which is conducive to improving the efficiency and stability of the inversion.

(5) Engineering practice: In deep rock engineering, such as tunnels and underground repositories, the Burgers model has been shown to be effective in predicting the long-term deformation behaviour of rocks. This is consistent with the engineering background of this study.

(6) Advantage of numerical simulation: The Burgers model can be more easily integrated into numerical simulation software such as finite element or finite difference, which facilitates subsequent engineering applications.

Considering the above factors, the Burgers model becomes the preferred model for describing the rock creep behaviour in this study. However, we also recognise that there is no single model that can perfectly describe the creep behaviour of rocks in all cases. In practical applications, we still need to verify the applicability of the Burgers model according to specific rock types and engineering conditions, and consider other models or improve the Burgers model when necessary. Next, we will develop a new creep parameter inversion method based on the Burgers model combined with the improved ESN method. This method aims to improve the accuracy and efficiency of parameter inversion and provide a more reliable theoretical basis for the design and safety assessment of deep rock engineering.

3. **Improved ESN model.**

3.1. **ESN model characteristics and basic structure.**

Echo State Networks are a special type of recurrent neural network originally proposed by Jaeger in 2001. ESN models have a unique structure and training methodology that makes them perform well when dealing with problems such as time series prediction and system identification. The most important feature of the ESN model is the reserve pool with large-scale stochastic connections [24]. The core of ESN is a reserve pool (Reservoir) containing a large number of neurons. The connection weights between these neurons are randomly generated and remain fixed during training. The states of neurons in the reservoir are considered as non- linear transitions of the input signal and "echoes" of historical information. This enables ESNs to effectively capture the dynamic nature of time-series data. Compared with traditional recurrent neural networks, the training process of ESN is much simpler and more efficient. Only the weights of the output layer need to be trained, which can usually be done by linear regression methods. Due to the stochastic and high-dimensional nature of the reserve pool, ESNs usually have better generalisation ability and can adapt to various non-linear dynamic systems. The basic structure of the ESN consists of three main parts [25]: the input layer, the reserve pool, and the output layer, which is shown in Figure 1.

Input layer: external signals are introduced into the network. At time step t, the input vector is denoted as $u(t) \in \mathbb{R}^{N_u}$, where N is the input dimension.

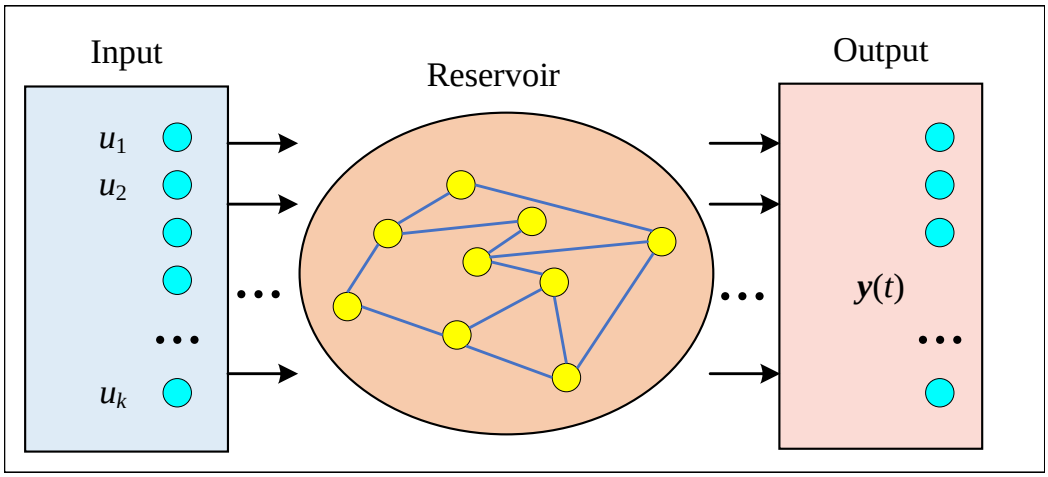

Figure 1. Schematic diagram of the basic structure of the ESN.

Reserve pool: contains a large number of randomly connected neurons. At time step t, the reserve pool state vector is denoted as $x(t) \in \mathbb{R}^{N_x}$, where N is the reserve pool size. The reserve pool state update equation is:

$$x(t) = f\big(W_{\text{in}}\, u(t) + W\, x(t-1) + W_{\text{fb}}\, y(t-1)\big) \tag{10}$$

where matrix,W_{in} W_{fb} is the input weight matrix, W is the reserve pool internal connection weight is the feedback weight matrix, and f is the activation function (usually a hyperbolic tangent function).

Output layer: maps the state of the reserve pool to the desired output. At time step t, the output vector is represented as computational equation is: $y(t) \in \mathbb{R}^{N_y}$ where N_y is the output dimension. The output computational equation is:

$$y(t) = W_{\text{out}}\,[\,x(t); u(t)\,] \tag{11}$$

where W_{out} is the output weight matrix, and $[x(t); u(t)]$ denotes the connection of the reserve pool states and the current inputs into a vector. The training process of ESN mainly consists of the following steps [26]: (1) Random initialisation: randomly generate the input weight matrix W_{in}, the reserve pool connection weight matrix W and the feedback weight matrix W_{fb}.

(2) State collection: use training data to drive ESN and collect reserve pool states.

(3) Output weight calculation: using the collected reserve pool states and target outputs, the output weight matrix W_{out} is calculated by linear regression methods (e.g. ridge regression).

These features of ESN models make them particularly suitable for dealing with time series prediction and system identification problems. However, there are still some limitations of the conventional ESN model when applied to complex rock creep parameter inversion problems. In the next subsections, we introduce several improvement strategies to enhance the performance of ESN models in creep parameter inversion.

3.2. **Structure optimisation with CNN..** In order to improve the performance of the ESN model when dealing with the inversion problem of rock creep parameters, we propose a structural optimisation method incorporating a CNN. This structure optimisation aims to enhance the model's ability to extract spatio-temporal features from the input data.

CNNs are known for their excellent performance in image processing and time series analysis. In our improved model, CNN is mainly used to preprocess the input data and

extract key features from it. First, we add a CNN module before the input layer of ESN. This CNN module contains multiple convolutional and pooling layers. The convolution operation can be represented as:

$$z_l = f\big(W_l * z_{l-1} + b_l\big) \tag{12}$$

where z_l is the feature map of the l-th layer, W_l is the convolution kernel, b_l is the bias term, $*$ denotes the convolution operation, and f is the activation function. Secondly, spatio-temporal feature extraction is achieved by designing appropriate convolutional kernels. CNN can capture both temporal and spatial features of the input data. This is particularly effective for dealing with multi-dimensional creeping data. The improved CNN-ESN structure is shown in Figure 2.

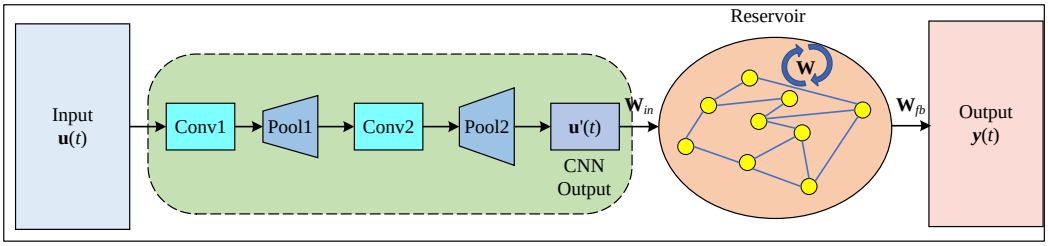

Figure 2. Schematic diagram of CNN-ESN structure

In the CNN-ESN structure, the raw input data $u(t)$ is first processed through the CNN module:

$$u'(t) = \mathrm{CNN}\big(u(t)\big) \tag{13}$$

The processed feature $u'(t)$ is then fed into the ESN reserve pool:

$$x(t) = f\big(W_{\mathrm{in}}\, u'(t) + W\, x(t-1) + W_{\mathrm{fb}}\, y(t-1)\big) \tag{14}$$

The introduction of CNN enables the model to automatically learn and extract key features from the input data, reducing the need for manual feature engineering. Through CNN's convolution and pooling operations, the dimensionality of the data input to the ESN reserve pool can be effectively reduced and the computational efficiency can be improved. CNN's feature extraction capability can help the model to better deal with noisy and anomalous data, and improve the robustness of the model.

3.3. **Improvement of the activation function in the reserve pool.** In addition to the structural optimisation, we have also improved the activation function in the ESN reserve pool to improve the nonlinear modelling capability and stability of the model. Traditional ESN models usually use the hyperbolic tangent function (tanh) as the activation function [27]. Although the tanh function performs well in many applications, the following problems may exist when dealing with complex nonlinear systems such as rock creep: (1) Saturation problem: When the input value is large, the tanh function tends to saturate, causing the gradient to disappear. (2) Limited nonlinear expressivity: a single tanh function may not adequately capture complex nonlinear relationships. To solve the above problem, we introduce Leaky ReLU function in the reserve pool:

$$f(x) = \begin{cases} x, & \text{if } x > 0, \\ \alpha\, x, & \text{otherwise}, \end{cases} \tag{15}$$

where α is a small positive constant (e.g. 0.01).

The Leaky ReLU function has the following advantages: (1) Avoiding gradient vanishing: for negative inputs, Leaky ReLU still has a small gradient, avoiding complete gradient vanishing. (2) Computationally efficient: compared to the tanh function, Leaky ReLU is simpler and more efficient. (3) Non-saturation: the Leaky ReLU is not saturated throughout the entire domain of definition, which facilitates the propagation of the gradient. To further enhance the nonlinear modelling capability of the model, we adopt a hybrid activation function strategy. Specifically, we use both tanh and Leaky ReLU functions in the reserve pool:

$$x(t) = \left[f_1\big(W_{\text{in},1}\, u'(t) + W_1\, x_1(t-1)\big) \;;\; f_2\big(W_{\text{in},2}\, u'(t) + W_2\, x_2(t-1)\big) \right] \tag{16}$$

where f_1 is the tanh function; f_2 is the Leaky ReLU function; and $\mathbf{x}_1$ and x_2 are subsets of the reserve pool states using different activation functions, respectively. This strategy of mixing different activation functions has several advantages. Firstly, it enhances the nonlinear representation of the network, which is able to capture more complex nonlinear relationships by combining multiple activation functions. Second, this strategy helps to balance the stability and flexibility of the model, for example, the tanh function provides a stable output range, while the Leaky ReLU increases the flexibility of the model when dealing with negative values. Finally, the hybrid strategy allows the model to better adapt to different types of input data and nonlinear relationships, thus improving the overall adaptability.

3.4. **Introducing regularisation and sparsity constraints.** In the previous subsections, we enhanced the performance of the ESN model by combining CNNs and improved activation functions. However, in order to further improve the generalisation ability and robustness of the model, especially when dealing with complex rock creep parameter inversion problems, we introduce regularisation and sparsity constraints. Regularisation is a commonly used technique to prevent overfitting. In our improved ESN model, we mainly use two regularisation methods: (1) L2 regularisation
L2 regularisation is achieved by adding a penalty term for the sum of squared weights to the loss function:

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^{N} \big(y_i - \hat{y}_i\big)^2 + \lambda \left\| W_{\text{out}} \right\|_2^2 \tag{17}$$

where N is the number of samples, y_i is the true output, $\hat{y}_i$ is the model predicted output, λ is the regularisation parameter, and $\|W_{\text{out}}\|_2^2$ is the L2 paradigm of the output weight matrix. (2) Early Stopping
Early Stopping decides when to stop training by monitoring performance on the validation set to prevent overfitting. We periodically evaluate the performance of the model on the validation set during training and stop training when the performance no longer improves. Sparsity constraints can help models learn more efficient feature representations and improve model interpretability. We introduce two sparsity constraints in the ESN model: (1) Input sparsity
By limiting the number of input connections, we can reduce the complexity of the model and improve its robustness to input perturbations. This is achieved by applying a masking operation to the input weight matrix W_{in}:

$$W_{\text{in}} = W_{\text{in}} \odot M_{\text{in}} \tag{18}$$

where $\odot$ denotes elementwise multiplication and M_{in} is a binary mask matrix where most elements are 0 and only a few are 1. (2) Reserve pool sparsity

By increasing the sparsity of the reserve pool connections, we can reduce the complexity of the reserve pool and improve the generalisation of the model. We use a similar masking operation:

$$W = W \odot M_{\text{res}} \tag{19}$$

where M_{res} is a sparse mask applied to the reserve pool connection weight matrix.

Combining the above regularisation and sparsity constraints, our final optimisation objective can be expressed as:

$$\min_{W_{\text{out}}} \frac{1}{N} \sum_{i=1}^{N} \left(y_i - \hat{y}_i\right)^2 + \lambda_1 \left\|W_{\text{out}}\right\|_2^2 + \lambda_2 \left\|W_{\text{out}}\right\|_1 \tag{20}$$

where λ_1 and λ_2 are the parameters for L2 regularisation and L1 regularisation (promoting sparsity), respectively. Based on the above optimisation objectives, we adjusted the training algorithm of ESN accordingly:

Step 1: Randomly initialise W_{in} and W and apply sparse mask.

Step 2: State update using improved activation function and CNN preprocessed input data.

Step 3: Calculate W_{out} using a regularised linear regression method.

Step 4: Regularly evaluate the model performance on the validation set and stop training early if necessary.

Step 5: Fine-tune the regularisation parameters λ_1 and λ_2 based on the validation set performance.

By introducing these regularisation and sparsity constraints, our improved ESN model is able to show stronger generalisation ability and robustness when dealing with complex rock creep parameter inversion problems. These improvements enable the model to better cope with various challenges that may be encountered in real engineering, such as limited training data, noise interference, and so on.

4. **Inversion of perimeter rock creep parameters based on improved ESN model.**

4.1. **Establishment of numerical model.**

In order to accurately simulate the creep behaviour of the surrounding rock and to provide a reliable basis for parameter inversion, we take a deep- buried diversion tunnel as the research object and establish a corresponding numerical model. The research object is a hydropower station diversion tunnel, the tunnel depth is about 500m, and the surrounding rock is mainly gneiss. The tunnel section is in the shape of a city gate, and the excavation size is 9.51m×10.62m (width×height), as shown in Figure 3.

Considering the complexity and computational efficiency of practical engineering, we make the following assumptions: (1) The surrounding rock material is a continuous, homogeneous, isotropic medium. (2) The simplified tunnel section is an equivalent circular shape with the diameter taken as 8.5m. (3) The initial stress field is assumed to be a hydrostatic pressure state. (4) The creep behaviour of the surrounding rock follows the Burgers model. (5) Neglecting the stress release effect during excavation, the tunnel is assumed to be completed by instantaneous excavation. Based on the above assumptions, we used FLAC3D software to build a two-dimensional axisymmetric model. The main features of the model are as follows: the size of the model is 100 metres in the radial direction and 1 metre in the axial direction. In terms of meshing, an inhomogeneous mesh is used in the radial direction, in which the mesh is encrypted near the tunnel and divided into 100 cells in total, while in the axial direction it is divided into 1 cell. Horizontal

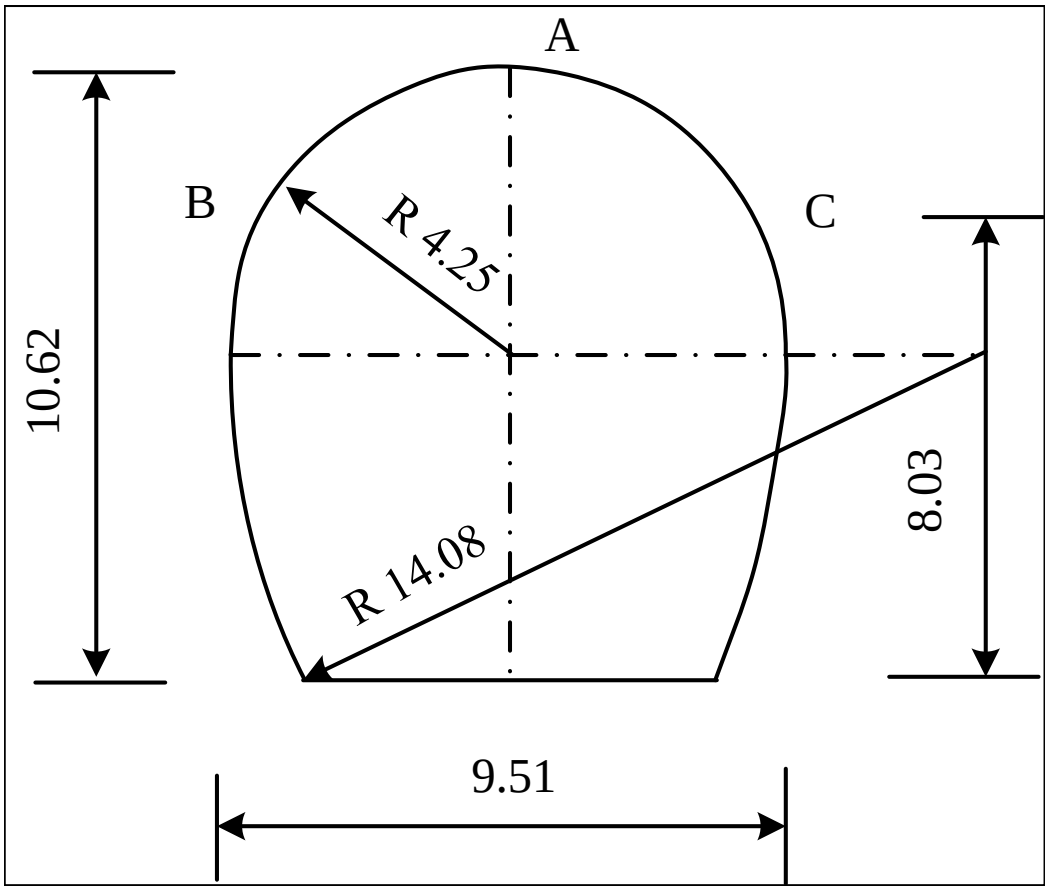

Figure 3. Schematic Cross-Section of the Tunnel

displacement constraints were imposed on the left boundary (tunnel axis), hydrostatic pressure boundary conditions were set on the right boundary, vertical displacement constraints were imposed on the upper and lower boundaries, and plane strain conditions were satisfied on the front and rear boundaries. According to the engineering geological investigation report and indoor test data, the mechanical parameters of the surrounding rock are determined as shown in Table 1.

Table 1. Mechanical parameters of the enclosing rock

Parameters	Numerical value
Density ρ (kg/m^3)	2700
Modulus of elasticity E (GPa)	15
Poisson's ratio ν	0.25
Cohesion c (MPa)	2.0
Angle of internal friction φ (°)	45
Tensile strength σ_t (MPa)	1.5

For the parameters of the Burgers creep model, we will determine them in the subsequent parameter inversion process. The initial values can be estimated based on similar engineering experience or laboratory test results. Considering the deeply buried nature of the tunnel, the initial stress field is in hydrostatic pressure state:

$$\sigma_v = \sigma_h = \gamma H \tag{21}$$

where σ_v and σ_h are the vertical and horizontal stresses, respectively; γ is the rock mass gravity (taken as 26.5 kN/m^3); and H is the burial depth (500 m). Therefore, the initial stress is:

$$\sigma_v = \sigma_h = 26.5 \times 500 = 13.25 \text{ MPa} \tag{22}$$

The numerical simulation was carried out in the following steps: first, an initial stress field was established to reflect the natural stress state in the formation; then, the transient excavation of the tunnel was simulated; then, the Burgers creep model was activated to describe the creep behaviour of the rock material over time; and finally, the long-term creep process was simulated, with the total time set to 1000 days, to assess the stability of the tunnel and the behaviour of the surrounding rock mass on a long time scale. behaviour of the surrounding rock mass on a long-time scale. During the simulation process, we will record the displacement and stress changes with time at key points around the tunnel, and these data will be used for subsequent parameter inversion and model validation. In order to verify the accuracy of the numerical model, we will take the following measures: firstly, we will carry out elasticity validation, i.e., we will simplify the model into an elastic problem and compare the results with the analytical solution; secondly, we will carry out mesh sensitivity analysis to check the convergence of the simulation results by changing the mesh density; lastly, we will compare the results of numerical simulation with the actual on-site monitoring data, and we will select the key indexes, such as the subsidence of the vault of the tunnel and the internal displacement of the enclosing rock. simulation results are compared with the actual field monitoring data to ensure the validity and accuracy of the model. Through these validation steps, we can ensure that our numerical model can accurately simulate the creep behaviour of the surrounding rock and provide a reliable basis for subsequent parameter inversion.

4.2. **Creep parameter inversion.**

4.2.1. *Mathematical description of the inversion problem.* The objective of the inversion of the envelope creep parameters is to find a set of Burgers model parameters that minimise the difference between the model predictions and the measured data. Mathematically this can be expressed as:

$$\min_{\boldsymbol{\theta}} \ J(\boldsymbol{\theta}) = \sum_{i=1}^{N} w_i \left[y_i - f_i(\boldsymbol{\theta}) \right]^2 \tag{23}$$

where $\boldsymbol{\theta} = [E_1, \eta_1, E_2, \eta_2]$ is the parameter vector of the Burgers model to be inverted, $J(\boldsymbol{\theta})$ is the objective function, N is the number of observation points, w_i is the weighting coefficient, y_i is the i-th measured value of the observation point, and $f_i(\boldsymbol{\theta})$ is the corresponding model prediction.

4.2.2. *Application of the improved ESN model to parameter inversion.* Using the improved ESN model proposed in Section 3, we can directly construct the mapping relationship between creep parameters and tunnel displacements:

$$y = F_{\text{CNN-ESN}}(\boldsymbol{\theta}, t) \tag{24}$$

where y is the displacement vector, t is the time vector, and $F_{\text{CNN-ESN}}$ denotes the improved ESN model. The numerical model developed in Section 4.1 was used to generate a large number of displacement-time curves for different combinations of creep parameters. The dataset was divided into a training set, a validation set and a test set. Use the training set data to train the improved ESN model, where the inputs are creep parameters

and time and the output is displacement. Use the validation set to adjust the model hyperparameters such as reserve pool size, sparsity, etc. Inversion process. Use the trained ESN model to directly predict the creep parameters corresponding to the given measured displacement data.

4.2.3. *Inversion Algorithm Flow.* The complete inversion algorithm flow is as follows: Step 1: Data Preprocessing. Collect the measured tunnel displacement-time data. Normalise the data. Step 2: ESN model application. Input the measured displacement-time data into the trained ESN model. The ESN model directly outputs the predicted creep parameters $\boldsymbol{\theta}^*$. Step 3: Parameter fine-tuning. A small fine-tuning can be done on the basis of the parameters output from the ESN to further improve the accuracy. The fine-tuning can be achieved by minimising the objective function:

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} \sum_{i=1}^{N} w_i \left[y_i - F_{\text{CNN-ESN}}(\boldsymbol{\theta}, t) \right]^2 \tag{25}$$

A simple gradient descent method can be used here as we already have a good initial estimate. Step 4: Output. Output the final creep parameter $\boldsymbol{\theta}^*$.

4.2.4. *Validation of inversion results.* In order to validate the reliability of the inversion results, firstly, we use the k-fold cross-validation method to assess the generalisation ability of the model. Secondly, sensitivity analysis is performed to assess the sensitivity of the inversion results to factors such as noise in the observed data. Next, a numerical simulation validation is performed by substituting the parameters obtained from the inversion into the numerical model described in Section 4.1 and comparing the agreement between the simulation results and the measured data. At the same time, a physical significance check is performed to ensure that the parameters obtained by inversion are within a reasonable physical range. Finally, the prediction performance is assessed by predicting the future deformation based on the inversion parameters and comparing them with the new monitoring data.

Through these validation steps, we can ensure the accuracy and reliability of the inversion results. This direct parameter inversion method based on the improved ESN model has the advantages of high computational efficiency and does not require complex optimisation algorithms, which is particularly suitable for dealing with practical engineering problems with large amounts of monitoring data.

4.3. **Analysis of numerical simulation results.** In the previous content, we established the numerical model of the deep-buried diversion tunnel and proposed the parameter inversion method based on the improved ESN model. Now, the numerical simulation results will be analysed in detail. The parameter inversion method for perimeter rock creep based on the improved ESN model shows significant advantages in the numerical simulation. We compared the method with four commonly used parameter inversion models: the BPNN model, the BO- RF model, the LSTM model, and the traditional ESN model. The comparison tests were performed on different rock types and time scales. The results are analysed as follows

Short-term prediction accuracy refers to the average relative error of the model in predicting perimeter rock deformation over a 30-day period. Long-term prediction accuracy, on the other hand, refers to the average relative error of the model in predicting perimeter rock deformation over 365 days (i.e., 1 year). The generalisation performance, on the other hand, is used to measure the average prediction accuracy of the model over different rock types. As shown in Table 2, the improved ESN model proposed in this paper achieves significant improvement in both short-term and long-term prediction accuracy.

Table 2. TABLE 2. Comparison of the performance of different models in the inversion of creep parameters in the surrounding rock

Performance indicators	**BPNN**	**BO-RF**	**LSTM**	**ESN**	**Improved ESN**
Short-term forecasting accuracy (average relative error)	8.5 %	7.2 %	6.8 %	7.5 %	6.2 %
Long-term prediction accuracy (mean relative error)	13.5 %	11.2 %	10.8 %	11.5 %	10.2 %
Generalisation performance (average accuracy for different rock types)	75 %	82 %	84 %	81 %	87.4 %
Calculation time (relative)	1.0	1.5	1.8	0.8	1.2
Number of model parameters (in millions)	5.2	8.7	7.5	3.8	6.3
Training data requirements (relative values)	1.0	1.3	1.5	0.9	1.1

In short-term forecasting, the average relative error of the improved ESN model is 6.2%, which is better than all other compared models. In long-term prediction, the average relative error of the improved ESN model is 10.2%, which is also better than the other models. This improvement is mainly due to the improved ESN structure combined with CNN, which enables the model to capture the nonlinear dynamic characteristics of the creep behaviour of the surrounding rock more effectively. The improved ESN model shows excellent generalisation ability in the tests of different rock body types. We selected five typical rock bodies for testing, including sandstone, shale, granite, limestone and mudstone. These tests aim to verify the predictive performance of the model under different geological conditions.

Table 3. Prediction Accuracy of Different Models on Various Rock Types

Rock type	BPNN	BO-RF	LSTM	ESN	Improved ESN
Limestone	78 %	84 %	86 %	83 %	89 %
Shale	72 %	80 %	82 %	79 %	86 %
Granite	80 %	85 %	87 %	84 %	90 %
Limestone	76 %	82 %	84 %	81 %	88 %
Shale	69 %	79 %	81 %	78 %	84 %
Average	75 %	82 %	84 %	81 %	87.4 %

As shown in Table 3, the improved ESN model exhibits the highest prediction accuracy on all rock types, reaching 87.4% on average. This result demonstrates the effectiveness

of introducing regularisation and sparsity constraints, allowing the model to better adapt to different geological conditions. In terms of computation time and the number of model parameters, the improved ESN model still maintains a good computational efficiency while maintaining high performance. Although its computation time is slightly higher than that of traditional ESN, it is significantly lower than that of BO-RF and LSTM models. The moderate number of model parameters indicates a good balance. The training data requirement of the improved ESN model is slightly higher than that of the traditional ESN, but lower than that of the BO-RF and LSTM models. This indicates that the model maintains high performance while relying relatively little on training data, which is important for practical engineering applications where data acquisition is difficult.

In addition, we conducted a sensitivity analysis of the model's performance under different rock parameters. The results show that the improved ESN model shows good robustness with an increase in prediction error of no more than 5% over a range of $\pm 20\%$ change in elastic modulus and $\pm 30\%$ change in cohesion coefficient. In contrast, the increase in prediction error of other models under the same conditions is generally higher than 6%. Overall, the numerical simulation results confirm the superiority of the improved ESN model in the problem of inversion of creep parameters in the surrounding rock. The method not only improves the short-term and long-term prediction accuracy, but also shows excellent generalisation ability over different rock types, while maintaining good computational efficiency and moderate model complexity. These improvements provide a more reliable tool for the prediction and parameter inversion of the creep behaviour of surrounding rock in deep- buried tunneling projects, which helps to improve the safety and efficiency of engineering design and construction.

5. **Conclusions.** In this paper, a parameter inversion method for envelope creep based on improved ESN is proposed, which effectively solves the limitations of the traditional inversion model in dealing with nonlinear dynamic features and long-term prediction. By introducing the CNN structure, the model is able to extract the spatio-temporal features in the surrounding rock creep data more accurately, which significantly improves the accuracy of the parameter inversion. In addition, the improved ESN model utilises a hybrid activation function strategy and regularisation constraints to further enhance the ability to capture the complex creep behaviour, ensuring the stability and robustness of the inversion results. The following conclusions can be drawn from the experiments conducted on multiple rock types and different time scales: (1) The introduction of CNN structures can significantly improve the ability of ESN models in capturing nonlinear dynamic features, especially when dealing with long-term creep behaviour. (2) Improved ESN models have higher computational efficiency and lower training data requirements than traditional ESNs and other deep learning models while maintaining high accuracy. (3) The improved ESN strategy combining hybrid activation functions and regularisation constraints achieves the best balance between prediction accuracy and generalisation ability, and is the optimal parameter inversion method recommended in this paper. The experimental data in this paper are mainly from five typical rock types (sandstone, shale, granite, limestone and mudstone), which covers common rock types but may still limit the generalisation ability of the model. Future work should consider introducing more perimeter rock creep data under complex geological conditions to validate the applicability of the model in a wider range of engineering environments. In addition, the combination of the improved ESN model with other advanced optimisation algorithms can be explored to further improve the efficiency and accuracy of parameter inversion.

Acknowledgment. This work is supported by the Science and Technology Research Project of Jiangxi Provincial Department of Education (No. GJJ2205221): Research and Engineering Application of the Comprehensive Evaluation System for Preventive Maintenance of Bridges Based on the Whole Life Cycle.

REFERENCES

[1] W. Cai, H. Zhu, and W. Liang, "Three-dimensional stress rotation and control mechanism of deep tunneling incorporating generalized Zhang–Zhu strength-based forward analysis," *Engineering Geology*, vol. 308, 106806, 2022.

[2] Z. Chen, Y. Zhang, J. Li, X. Li, and L. Jing, "Diagnosing tunnel collapse sections based on TBM tunneling big data and deep learning: A case study on the Yinsong Project, China," *Tunnelling and Underground Space Technology*, vol. 108, 103700, 2021.

[3] C. Iasiello, J. C. G. Torralbo, and C. T. Fernández, "Large deformations in deep tunnels excavated in weak rocks: Study on Y-Basque high-speed railway tunnels in northern Spain," *Underground Space*, vol. 6, no. 6, pp. 636–649, 2021.

[4] Z. Sun, D. Zhang, Q. Fang, G. Dui, and Z. Chu, "Analytical solutions for deep tunnels in strain-softening rocks modeled by different elastic strain definitions with the unified strength theory," *Science China Technological Sciences*, vol. 65, no. 10, pp. 2503–2519, 2022.

[5] W. Frenelus, H. Peng, and J. Zhang, "Creep behavior of rocks and its application to the long-term stability of deep rock tunnels," *Applied Sciences*, vol. 12, no. 17, 8451, 2022.

[6] K.-K. Phoon, and W. Zhang, "Future of machine learning in geotechnics," *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, vol. 17, no. 1, pp. 7–22, 2023.

[7] P. Zhang, Z.-Y. Yin, and Y.-F. Jin, "Machine learning-based modelling of soil properties for geotechnical design: review, tool development and comparison," *Archives of Computational Methods in Engineering*, pp. 1–17, 2022.

[8] B. Satipaldy, T. Marzhan, U. Zhenis, and G. Damira, "Geotechnology in the Age of AI: The Convergence of Geotechnical Data Analytics and Machine Learning," *Fusion of Multidisciplinary Research, An International Journal*, vol. 2, no. 1, pp. 136–151, 2021.

[9] D.-W. Li, R.-H. Wang, and J.-G. Fan, "The nonlinear creep characteristics and parameter inversion of soft rock," *Journal of China Coal Society*, vol. 36, no. 3, pp. 388–392, 2011.

[10] W. Yang, Q. Zhang, S. Li, and S. Wang, "Estimation of in situ viscoelastic parameters of a weak rock layer by time-dependent plate-loading tests," *International Journal of Rock Mechanics and Mining Sciences*, vol. 66, pp. 169–176, 2014.

[11] Z. Wang, X. Chen, X. Xue, L. Zhang, and W. Zhu, "Mechanical parameter inversion in sandstone diversion tunnel and stability analysis during operation period," *Civil Engineering Journal*, vol. 5, no. 9, pp. 1917–1928, 2019.

[12] G. Li, Y. Hu, Q.-B. Li, T. Yin, J.-X. Miao, and M. Yao, "Inversion method of in-situ stress and rock damage characteristics in dam site using neural network and numerical simulation—a case study," *IEEE Access*, vol. 8, pp. 46701–46712, 2020.

[13] C. Chen, T. Li, C. Ma, H. Zhang, J. Tang, and Y. Zhang, "Hoek-Brown Failure Criterion-Based Creep Constitutive Model and BP Neural Network Parameter Inversion for Soft Surrounding Rock Mass of Tunnels," *Applied Sciences*, vol. 11, no. 21, 10033, 2021.

[14] K. Li, L. Pan, and Y. Wang, "Random forest-based modelling of parameters of fractional derivative concrete creep model with Bayesian optimization," *Materials and Structures*, vol. 55, no. 8, 215, 2022.

[15] X. Zhang, C. Zhu, M. He, M. Dong, G. Zhang, and F. Zhang, "Failure mechanism and long short-term memory neural network model for landslide risk prediction," *Remote Sensing*, vol. 14, no. 1, 166, 2021.

[16] T. Kim, and B. R. King, "Time series prediction using deep echo state networks," *Neural Computing and Applications*, vol. 32, no. 23, pp. 17769–17787, 2020.

[17] H. Hu, L. Wang, and S.-X. Lv, "Forecasting energy consumption and wind power generation using deep echo state network," *Renewable Energy*, vol. 154, pp. 598–613, 2020.

[18] T.-Y. Wu, H. Li, S. Kumari, and C.-M. Chen, "A Spectral Convolutional Neural Network Model Based on Adaptive Fick's Law for Hyperspectral Image Classification," *Computers, Materials & Continua*, vol. 79, no. 1, pp. 19–46, 2024.

[19] Y. Ma, Y. Peng, and T.-Y. Wu, "Transfer learning model for false positive reduction in lymph node detection via sparse coding and deep learning," *Journal of Intelligent & Fuzzy Systems*, vol. 43, no. 2, pp. 2121–2133, 2022.

[20] F. Zhang, T.-Y. Wu, J.-S. Pan, G. Ding, and Z. Li, "Human motion recognition based on SVM in VR art media interaction environment," *Human-centric Computing and Information Sciences*, vol. 9, 40, 2019.

[21] F. Wu, J. Chen, and Q. Zou, "A nonlinear creep damage model for salt rock," *International Journal of Damage Mechanics*, vol. 28, no. 5, pp. 758–771, 2019.

[22] A. Singh, C. Kumar, L. G. Kannan, K. S. Rao, and R. Ayothiraman, "Estimation of creep parameters of rock salt from uniaxial compression tests," *International Journal of Rock Mechanics and Mining Sciences*, vol. 107, pp. 243–248, 2018.

[23] M. Wang, and M. Cai, "A grain-based time-to-failure creep model for brittle rocks," *Computers and Geotechnics*, vol. 119, 103344, 2020.

[24] L. Grigoryeva, and J.-P. Ortega, "Echo state networks are universal," *Neural Networks*, vol. 108, pp. 495–508, 2018.

[25] C. Gallicchio, A. Micheli, and L. Pedrelli, "Design of deep echo state networks," *Neural Networks*, vol. 108, pp. 33–47, 2018.

[26] J. Long, S. Zhang, and C. Li, "Evolving deep echo state networks for intelligent fault diagnosis," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 7, pp. 4928–4937, 2019.

[27] J. Viehweg, K. Worthmann, and P. Mäder, "Parameterizing echo state networks for multi-step time series prediction," *Neurocomputing*, vol. 522, pp. 214–228, 2023.