

Damage prediction of shield tunnel structures driven by numerical simulation and machine learning

Yang Yu

Beihua University

Haishuang Guan

Beihua University

Shuang Gao

Beihua University

Chunbing Guo

Beihua University

Yanyu Meng

myugina@beihua.edu.cn

Beihua University

Research Article

Keywords: Shield tunnel, ultimate bearing capacity, failure mechanism, machine learning, prediction model, crack width

Posted Date: September 28th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11042831/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Damage prediction of shield tunnel structures driven by numerical simulation and machine learning

Yang Yu^{a,b}, Haishuang Guan^c, Shuang Gao^a, Chunbing Guo^a, Yanyu Meng^{a*}

a. School of Civil Engineering & Transportation, Beihua University, Jilin, 132013, China

b. Institute of Transportation Engineering, Zhejiang University, Hangzhou, 310058, China

c. School of Electrical and Information Engineering, Beihua University, Jilin, 132021, China

Abstract: Shield tunnels are designed as permanent transportation infrastructure with a service life of up to 100 years. To investigate the failure process and underlying mechanisms of shield tunnel structures and to ensure metro operational safety, this study proposes a damage prediction framework that integrates high fidelity numerical simulation with machine learning. First, a refined numerical model of a shield tunnel structure is developed by adopting a total strain crack model to represent concrete nonlinearity, while explicitly considering the effects of local manhole weakening, multiple contact interfaces, and embedded reinforcement. The reliability of the model is verified against available experimental results. Subsequently, failure simulations under ultimate bearing conditions are conducted to elucidate structural stress and deformation responses as well as reinforcement stress evolution, thereby revealing detailed damage characteristics of the tunnel structure. A particle swarm optimization algorithm is then employed to optimize key hyperparameters of a long short term memory network, including the number of hidden units, the optimal initial learning rate, and the L2 regularization coefficient. On this basis, a particle swarm optimization (PSO) long short-term memory (LSTM) damage prediction model is established. Training samples are generated from numerical simulations, enabling accurate prediction of crack width evolution in shield tunnels under ultimate loading conditions. The results indicate that the numerical model exhibits good agreement with experimental observations, demonstrating sufficient reliability. The structural failure process can be divided into three stages, namely the linear elastic stage, the local damage stage, and the global failure stage. The proposed PSO-LSTM model achieves a root mean square error (RMSE) of 0.1033 and a coefficient of determination (R^2) of 0.9487 in crack width prediction, outperforming backpropagation neural network (BP), convolutional neural network (CNN), and conventional LSTM models in both accuracy and stability. The proposed approach provides effective guidance for monitoring and protection of metro tunnels and contributes to improving operational safety.

Keywords: Shield tunnel; ultimate bearing capacity; failure mechanism; machine learning; prediction model; crack width

1. Introduction

Metro systems as a crucial component of urban rail transit, play an important role in alleviating traffic congestion and have been widely adopted in cities worldwide. Shield tunnels constructed in soft soil deposits are particularly vulnerable to disturbances induced by surrounding construction activities, such as surface surcharge loading [1], excavation of adjacent foundation pits [2], and seismic excitation [3]. These external actions can inevitably lead to various structural defects, including segment cracking and damage [4],

joint dislocation, and water leakage [5], and in severe cases may trigger global instability and structural failure [6][7][8]. To ensure the long term service performance of shield tunnel structures, it is therefore essential to clarify the failure process and mechanisms of tunnel segments under ultimate loading conditions, thereby safeguarding metro operational safety.

Extensive studies have been conducted worldwide on the service performance of shield tunnel structures, yielding substantial achievements. In terms of simplified analytical approaches, Huang et al. [9] proposed an analytical method for evaluating the longitudinal

Funding: Jilin Provincial Department of Education Project (20260701, JJKH20200043KJ)

Natural Science Foundation of Jilin Provincial Department of Science and Technology (YDZJ202501ZYTS423, 2019012015JH)

Jilin Provincial Development and Reform Commission Industrial Technology Research and Development Project (2024C006-1)

Outstanding Young Talents Training Project of Jilin Science and Technology Bureau (201831767, 20156416)

mechanical response of shield tunnels, in which soil structure interaction and joint shear resistance were taken into account, and the validity of the method was verified through finite element analysis. Chen et al. [10] derived the mechanical response of shield tunnels under opening conditions using a state space approach, and showed that an increase in opening ring displacement leads to a reduction in internal forces, while the influence of openings is closely related to the surrounding geological conditions. Feng et al. [11] investigated circumferential joints connected by inclined bolts, analyzed the nonlinear relationship between shear stiffness and joint dislocation at different bolt positions, and derived the deformation and internal force responses of shield tunnels subjected to surface surcharge loading, with validation provided by numerical simulations. Feng et al. [12] further proposed a new analytical model for multi ring shield tunnels, which accurately captures the nonlinear deformation behavior of stagger joint assemblies and was validated through full scale experiments. To overcome the limitations of analytical theories, numerical simulations and laboratory experiments have been widely employed to explore damage and failure mechanisms of shield tunnels. Zhu et al. [13] based on resilience theory, developed a refined numerical model to investigate the structural response characteristics and resilience evolution of shield tunnels with different burial depths under surface loading and unloading. The results indicated that tunnel resilience decreases rapidly with increasing horizontal convergence, whereas shortening response time and improving repair efficiency can effectively enhance structural resilience. Sun et al. [14] considered the discontinuous deformation characteristics of interfaces in steel plate reinforced shield tunnels and performed numerical analyses under ultimate bearing conditions, demonstrating that interface failure is the direct cause of reinforced structural failure. Liu et al. [15] proposed an improved lightweight epoxy bonded steel plate strengthening scheme and investigated the load bearing mechanism and failure process through full scale experiments and numerical simulations. Their results showed that structural strengthening increased the ultimate bearing capacity by 21 percent and exhibited ductile failure behavior. Yu et al. [16][17] based on the total strain crack model, established refined numerical

models for shield tunnels and double lining structures, and in combination with resilience theory, examined progressive failure characteristics and discussed the influence of strengthening vertical deformation on structural performance.

In recent years, machine learning based neural network prediction methods [18][19][20] have attracted increasing attention due to their strong self learning and data analysis capabilities, which enable them to uncover latent patterns in data and effectively address time series prediction problems. To date, several studies have applied neural networks to prediction tasks in tunnel engineering. Zhou et al. [21] developed a genetic algorithm-Bi-directional Long Short-Term Memory (GA-Bi-LSTM) based settlement prediction model for existing tunnels. Relying on field engineering data, the model was trained and tested using monitored settlement records and corresponding input parameters, providing a new approach for predicting settlement induced by newly constructed tunnels passing beneath existing ones. Sun et al. [22] employed a PSO optimized LSTM network to predict surface settlement at the portal of a mountain tunnel during construction and to assess the average collapse risk of the section. Lin et al. [23] based on an actual engineering project, adopted a physics informed neural network approach to predict the longitudinal deformation of a super large diameter shallow buried shield tunnel and conducted comparative analyses with five other algorithms. To address the so called black box issue inherent in many machine learning algorithms, Chen et al. [24] proposed a novel data and experience driven prediction framework to estimate ground settlement induced by shield tunnel excavation, achieving effective prediction of settlement curve characteristics. A review of existing studies indicates that neural network based prediction methods have been predominantly applied to settlement and deformation problems in tunnel engineering, whereas predictive research focusing on failure and damage evolution of shield tunnels under ultimate loading conditions remains limited. Therefore, it is necessary to develop an appropriate neural network based approach to predict shield tunnel failure processes, thereby contributing to the safe operation of metro systems.

In this study, a refined structural model of shield

tunnel segments is established based on the total strain crack model to investigate failure mechanisms under ultimate bearing conditions. An LSTM network is selected as the base predictive model, and a particle swarm optimization algorithm, which iteratively updates particle velocities and positions to identify the global optimum, is employed to optimize key network parameters. On this basis, a PSO optimized LSTM prediction model is developed to characterize damage evolution of shield tunnels under ultimate loading. The model is trained and tested using data validated by full scale experiments, and its performance in crack width prediction is comparatively evaluated against BP, CNN, and conventional LSTM models, thereby demonstrating the reliability of the proposed PSO-LSTM framework.

2. Methodology

An overview of the proposed methodology is illustrated in Fig. 1. The framework integrates numerical simulation and machine learning to develop a hybrid neural network model for predicting damage in shield tunnels. First, concrete nonlinearity is represented using

a total strain crack model. The effects of local manhole weakening, multiple contact interfaces, and embedded reinforcement are explicitly considered to establish a refined structural model of the shield tunnel, whose reliability is validated against existing experimental results (Step 1, Section 2.1). Subsequently, failure simulations under ultimate bearing conditions are performed to comprehensively analyze structural stress deformation behavior and reinforcement stress evolution, thereby obtaining detailed damage characteristics of the shield tunnel and extracting crack width as the key indicator for damage prediction (Step 2, Section 2.2). A particle swarm optimization algorithm is then employed to optimize critical hyperparameters of the long short term memory network, including the number of hidden units, the optimal initial learning rate, and the L2 regularization coefficient, leading to the development of a PSO optimized LSTM damage prediction model (Step 3, Section 3.4). Finally, datasets generated from numerical simulations are used to achieve accurate prediction of crack width evolution in shield tunnels under ultimate loading conditions (Step 4, Section 3.5).

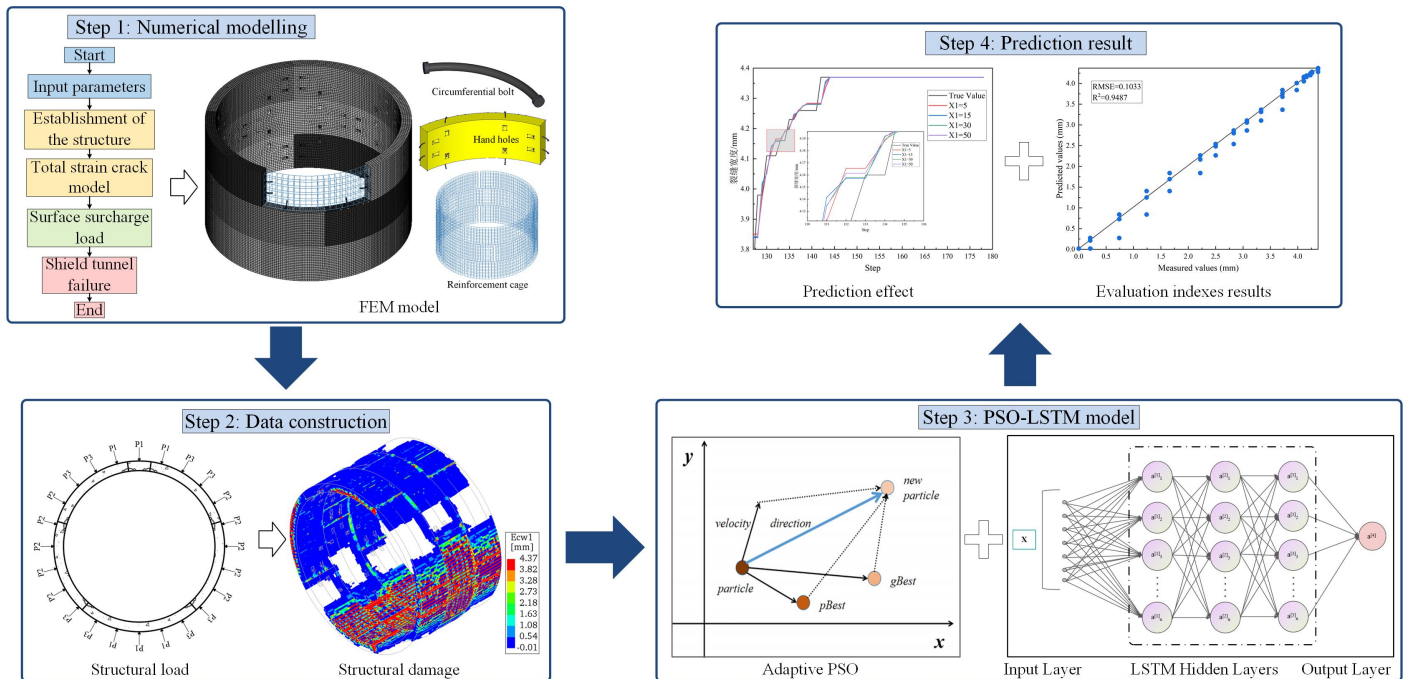


Fig. 1. Flowchart of this study

2.1 Refined model

2.1.1 Model development

The numerical model was established based on the

actual dimensions of the Shanghai Metro tunnel project. The shield tunnel segment has an outer diameter of 6.2 m, an inner diameter of 5.5 m, a thickness of 0.35 m, and a

ring width of 1.2 m. Staggered assembly was adopted. The segmental lining consists of one key segment with a central angle of 16° , two adjacent segments with central angles of 65° , two standard segments with central angles of 65° , and one invert segment with a central angle of 84° , as shown in Fig. 2. Adjacent segments are connected by 12 circumferential curved bolts of type M30. The contribution of reinforcement is explicitly considered in the model. The diameters of the inner and outer main longitudinal rebars are 22 mm and 16 mm, respectively, while the stirrups have a diameter of 10 mm. Reinforcement is simulated using embedded truss elements L6TRU, which are embedded within the concrete segments. Both the bolts and the segments are modeled using three dimensional solid elements HX24L, and the interaction between bolts and segments is described using a Coulomb friction model. The contact behavior between adjacent segment joints is simulated using interface elements, which can accurately capture the sliding and rotational behavior of longitudinal joints.

The normal and tangential stiffnesses of the interface are taken as $8 \times 10^8 \text{ N/m}^3$ and $5 \times 10^{11} \text{ N/m}^3$, respectively. Based on the above assumptions, a refined three ring staggered assembled shield tunnel segment model is established, as shown in Fig. 3.

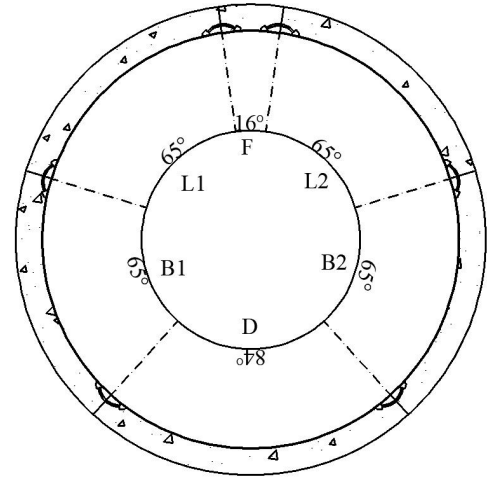


Fig. 2 Segmental lining configuration of the shield tunnel

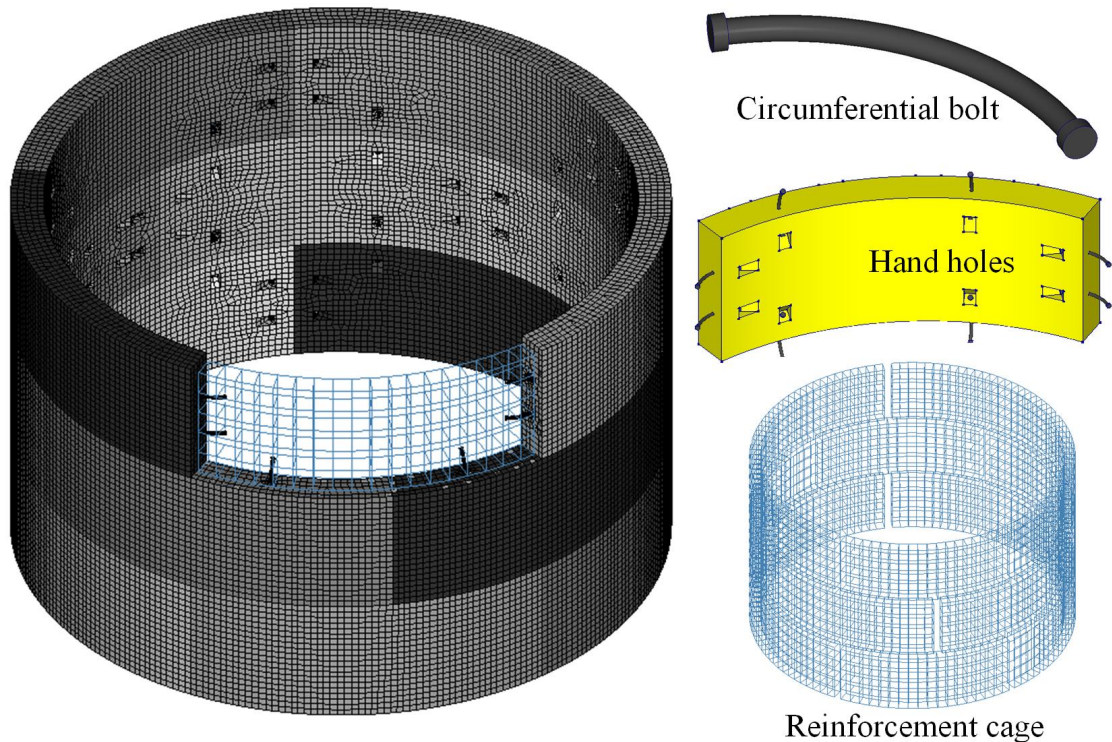


Fig. 3 Refined analytical model of the shield tunnel segment structure

2.1.2 Constitutive model and material parameters

The concrete segments are simulated using the total strain crack model, in which stress evolution is primarily

described as a function of strain. This model is capable of accurately capturing the nonlinear behavior of concrete under combined tension, compression, and

shear. The core concept of the model is to consider cracking strain ε_{cr} and elastic strain ε_e in a unified framework, while the stress magnitude is determined according to the crack orientation. The strain vector of an element ε_{xyz} expressed in the global Cartesian coordinate system xyz can be written as:

$$\varepsilon_{xyz}^{t+\Delta t} = \varepsilon_{xyz}^t + \Delta \varepsilon_{xyz} \quad (1)$$

Based on the strain transformation matrix T , the transformed strain vector ε_{nst} is expressed as:

$$\varepsilon_{nst}^{t+\Delta t} = T_{i+1}^{t+\Delta t} \varepsilon_{xyz} \quad (2)$$

According to the relationship between the crack orientation and the principal stress direction, the total strain crack model can be classified into the fixed crack model and the rotating crack model. In the rotating crack model, the principal stress direction of the element governs the strain transformation matrix T , which can be expressed as:

$$T = T(\varepsilon_{nst}^{t+\Delta t}) \quad (3)$$

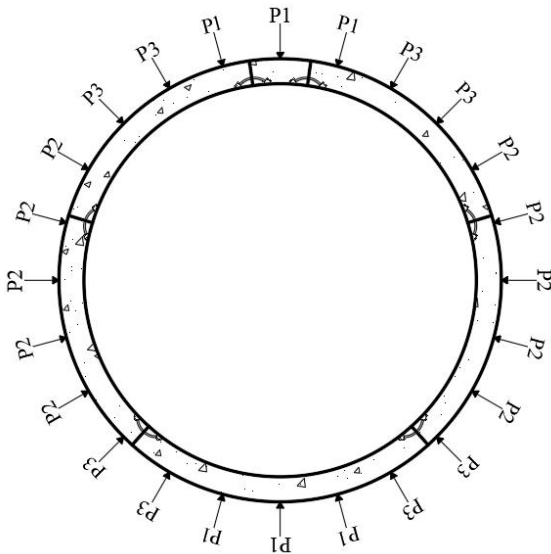
For the fixed crack model, the stress vector $\sigma_{nst}^{t+\Delta t}$ in the crack coordinate system is given by:

$$\sigma_{nst}^{t+\Delta t} = f(\varepsilon_{nst}^{t+\Delta t}, \theta) \quad (4)$$

where θ represents the threshold value of the included angle between different cracks.

Finally, the stress vector in the element coordinate system can be obtained as:

$$\sigma_{xyz}^{t+\Delta t} = T_{i+1}^{t+\Delta t} \sigma_{nst} \quad (5)$$



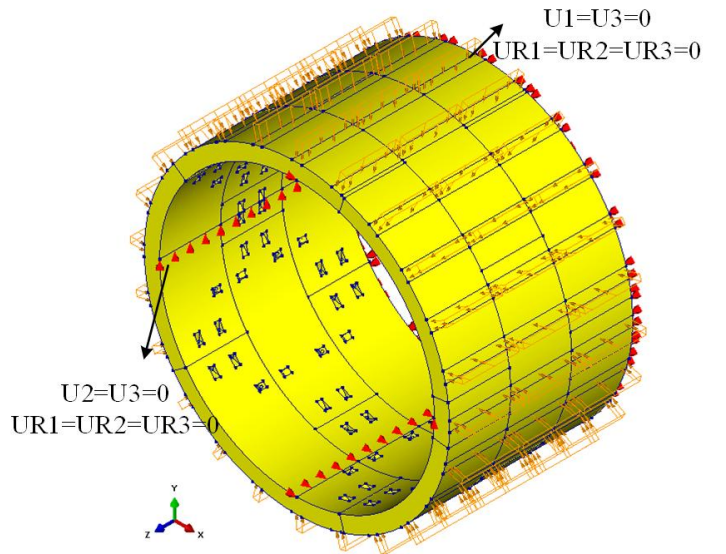
In this study, the crack propagation behavior is described using the rotating crack model. The material parameters adopted for concrete and reinforcement are summarized in Table 1. The embedded reinforcement and connecting bolts are modeled using a Von-Mises plastic hardening constitutive law. The reinforcement is simulated with embedded truss elements, and the thickness of the concrete cover is taken as 50 mm.

Tab.1. Material parameters

Material	Type	Young's modulus/M Pa	Poisson's ratio	Tensile stress/M Pa	Compressive stress/MPa
Segmental lining	C50	3.45×10^4	0.2	1.89	23.1
Rebar	HRB335	2×10^5	0.3	335	—
Joint bolts	M30	2.1×10^5	0.2	400	—

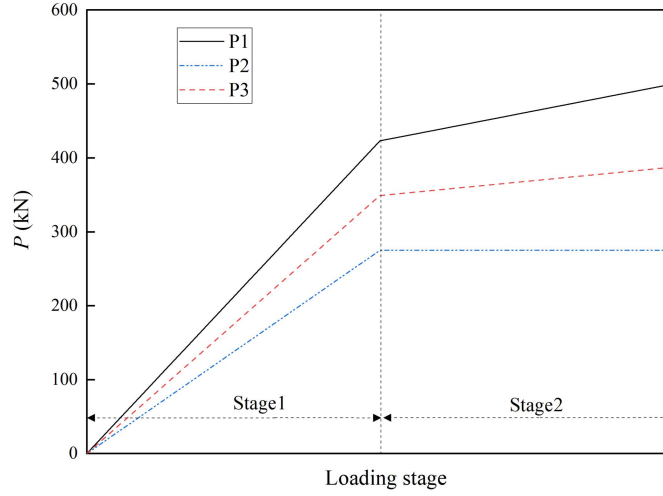
2.1.3 Loading procedure

The loading scheme and loading process are illustrated in Fig. 4. The numerical loading protocol follows the full scale experimental setup and includes a total of 24 loading points. These points are arranged at angular intervals of 15° about the tunnel center and are subjected to radial loads directed toward the tunnel axis, applied by hydraulic jacks and denoted as P1, P2, and P3. Load P1 is applied at the tunnel crown and invert, comprising six loading points in total. Load P2 is applied at the haunch on both sides of the tunnel, with ten loading points. The remaining points are subjected to load P3.



(a) Load grouping

(b) Loading configuration



(c) Loading stages

Fig. 4. Loading scheme for failure analysis of the shield tunnel

As shown in Fig. 4(c), the loading process is divided into two stages. In the first stage, the loads P1, P2, and P3 increase linearly from zero. Load P2 is increased to 275 kN and then held constant, while P1 and P3 are increased to $1.54 \times P2$ and $0.5 \times (P1 + P2)$, respectively. In the second stage, P1 continues to increase linearly until the maximum load is reached, corresponding to structural bearing failure, while P2 remains constant. Throughout the entire loading process, P3 increases linearly and is maintained at $0.5 \times (P1 + P2)$.

2.2 Analysis of Numerical Results

2.2.1 Tunnel deformation

Fig. 5 compares the load-displacement curves obtained from the numerical model based on the total strain crack model with those from the full scale experimental test. As shown in Fig. 5, during the initial loading stage, a high level of agreement is observed between the experimental and numerical load convergence responses. With increasing load, the experimental displacement data exhibit a certain degree of fluctuation. This behavior is mainly attributed to the progressive damage of the concrete material and the existence of contact gaps between the loading apparatus and the shield tunnel in the experiment, which are difficult to fully capture in numerical simulations. Overall, the numerical results are in good agreement with the experimental data, indicating that the adopted

constitutive model can reasonably represent the mechanical behavior and deformation characteristics of the shield tunnel structure, and is suitable for further refined analysis of structural responses.

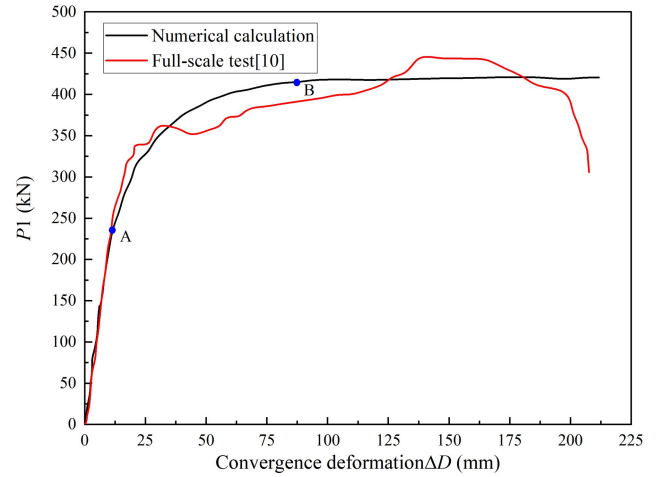


Fig. 5. Load-displacement curve

The mechanical response of the shield tunnel can be divided into three distinct stages. The first stage is the linear elastic stage, corresponding to segment OA on the load-displacement curve. The associated load range is from 0 to 230 kN, with a deformation range of 0 to 11 mm. In this stage, the tunnel structure behaves elastically, and the initial structural stiffness remains relatively high. The second stage is the elastic plastic damage stage,

corresponding to segment AB. The load increases from 230 to 420 kN, while the deformation ranges from 11 to 83 mm. During this stage, the structural stiffness gradually decreases, cracks within the shield tunnel progressively initiate and propagate, and the internal bolts and reinforcing bars begin to enter the yielding state. The third stage is the failure stage, corresponding to loads greater than 420 kN and deformations exceeding 83 mm. In this stage, the structural stiffness decreases progressively toward zero, through cracks develop within the tunnel lining, and the damage level becomes severe. In addition, most of the internal bolts and reinforcing bars yield, multiple plastic hinges form, and the structure ultimately undergoes global failure. Based on the above analysis, the ultimate bearing capacity of the structure can be defined as the load corresponding to point B. The simulated ultimate bearing capacity of the shield tunnel is 420 kN, which is close to the value obtained from the full scale experimental test. Such significant convergence deformation is unlikely to occur under normal service conditions of shield tunnels and is generally associated with extreme catastrophic scenarios, such as severe surface surcharge loading.

2.2.2 Damage evolution of the tunnel

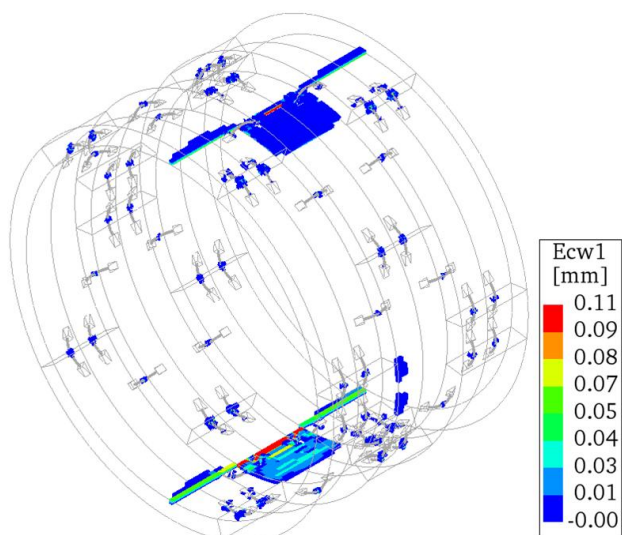
Fig. 6 illustrates the damage evolution of the shield tunnel during loading, where ECW1 in the legend denotes crack width. As shown in Fig. 6, when the tunnel is subjected to external loading, pronounced tensile cracks first develop in tensile zones due to the relatively low tensile strength of concrete. With increasing load, cracks initially appear on the inner sides of the crown and invert. Owing to stress concentration in the bolts, noticeable local damage also develops at the segment joint and bolt connection regions.

As shown in Fig. 6(a), when the applied load reaches approximately 170 kN, microcracks emerge on the inner sides of the crown and invert, indicating a transition of the structure from a linear elastic state to an elastoplastic state. At this stage, the maximum crack width is 0.11 mm. With further loading, both the number and depth of cracks at the crown and invert gradually increase. As shown in Fig. 6(b), when P1 reaches about 230 kN, cracks begin to appear on the outer side of the haunch, while the number of cracks on the inner sides of

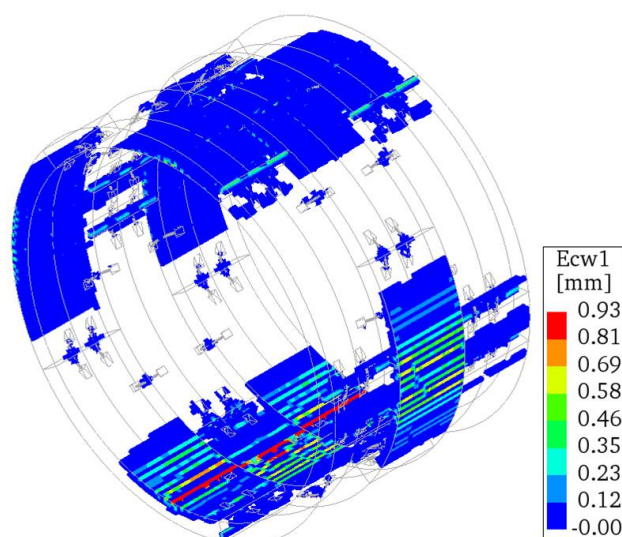
the crown and invert continues to grow, with the maximum crack width increasing to 0.93 mm. As loading proceeds, both the number and width of cracks in tensile regions increase continuously. As shown in Fig. 6(d), when P1 reaches approximately 370 kN, multiple cracks have developed in the tensile zones, the segments suffer extensive damage, and most cracks have propagated through the section. At this stage, the maximum crack width reaches 2.1 mm. As shown in Fig. 6(e), when P1 reaches about 420 kN, large damaged regions form in all tensile zones of the tunnel, and the number of cracks along the outer edges of both haunch further increases. Severe damage is observed at the joints near the haunch and crown, where plastic hinges begin to form. The maximum crack width at this stage is 3.16 mm. As shown in Fig. 6(f), at structural failure, four plastic hinges form in the tensile regions of the tunnel. Throughout the loading process, tensile failure dominates the damage pattern. In addition, local collapse occurs at segment joints, and the maximum crack width reaches 4.37 mm.

2.2.3 Internal forces and stresses of the tunnel

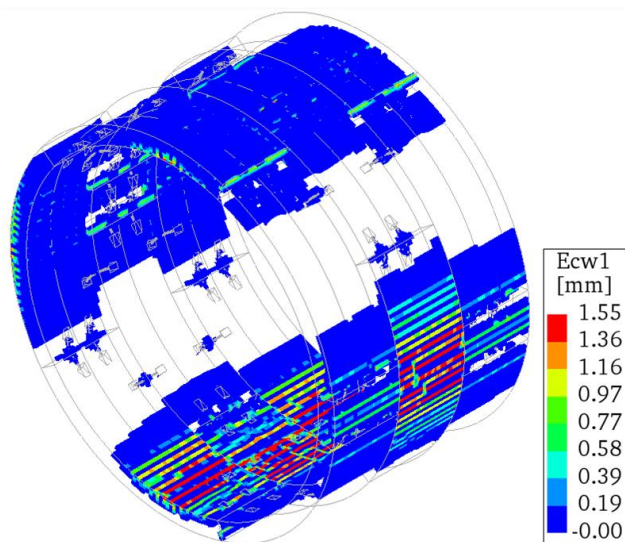
Fig. 7 presents the distributions of axial force and bending moment in the middle ring segments obtained from the numerical analysis. Angles of 0° and 180° correspond to the crown and invert, respectively, while 90° and 270° denote the left and right haunch. As shown in Figs. 7(a) and 7(b), the axial force in the segment lining is compressive around the entire ring. In contrast, the bending moment distribution indicates tensile regions at the inner sides of the crown and invert and at the outer sides of the left and right haunch. This response implies outward expansion at the haunch and inward contraction at the crown and invert. The bending moment and axial force acting on the lining exhibit axial symmetry along the 0° – 180° axis. With increasing external load, the axial forces at the crown, haunch, and invert increase progressively. In terms of uniformity, the bending moment distribution is relatively even, whereas the axial force distribution is nonuniform at the left and right haunch. This indicates the presence of stress concentration at the haunch, which adversely affects the load bearing capacity of the lining structure.



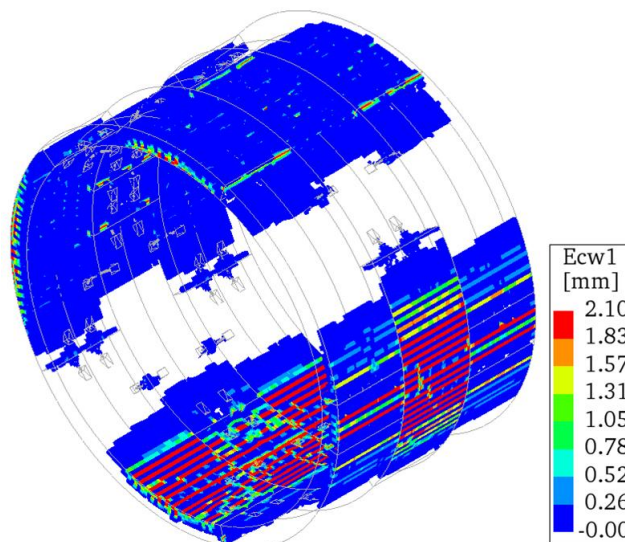
(a) 170 kN, 0.13% D



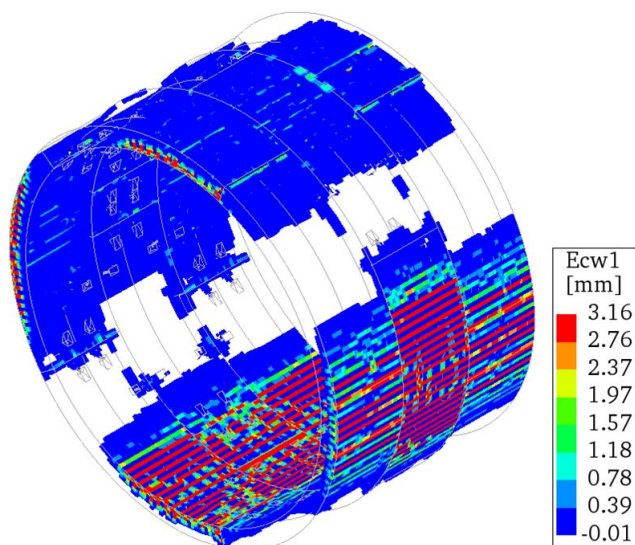
(b) 230 kN, 0.17% D



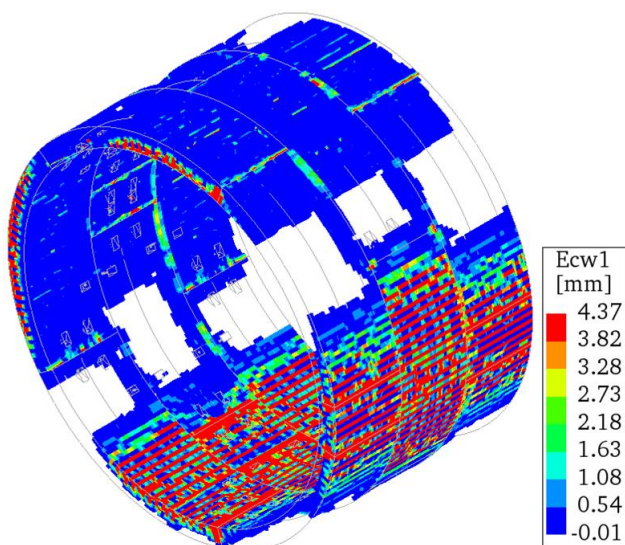
(c) 330 kN, 0.41% D



(d) 370 kN, 0.62% D



(e) 420 kN, 1.33% D



(f) Failure

Fig. 6. Damage evolution of shield tunnel

Overall, the evolution of structural damage leads to pronounced changes in internal forces. As shown in Fig. 7(a), a sudden variation in axial force occurs at the 90° and 270° positions when P_1 reaches 330 kN, followed by a continuous increase with further loading. This behavior can be attributed to changes in the contact state of nearby

joints. At $P_1=330$ kN, multiple cracks initiate and propagate through the joints and haunch regions, reducing the effective load carrying area. The reduction in the effective section directly alters the axial force distribution in the segments, ultimately resulting in the observed abrupt change in axial force.

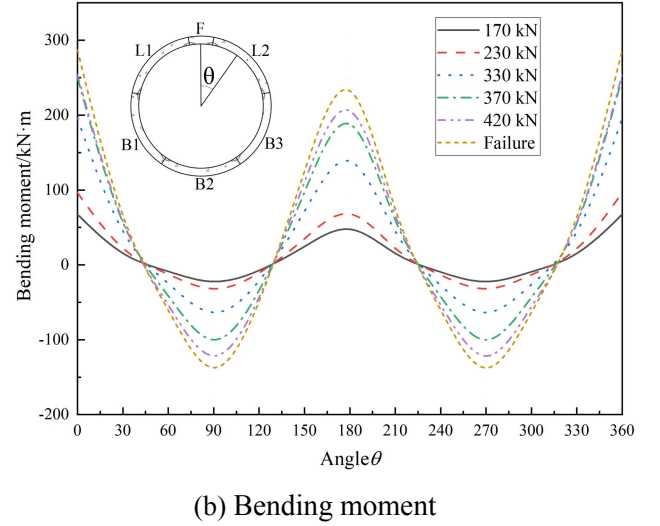
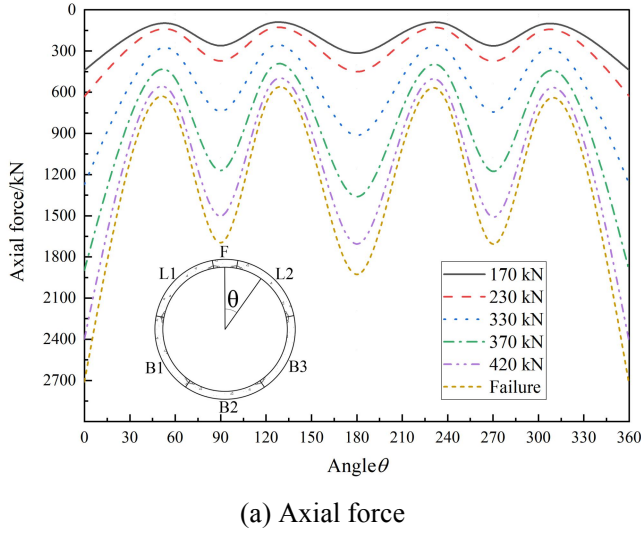


Fig. 7. Internal forces of shield tunnel

The stress evolution of the internal reinforcement in the shield tunnel during the loading process was further analyzed. Fig. 8 presents the stress contours of the embedded reinforcement in the three ring segment model. The internal reinforcement is mainly subjected to tensile stress, whereas the external reinforcement experiences compressive stress, and the reinforcement stress increases progressively with increasing load. By comparing Fig. 8 with the segment damage contours shown in Fig. 6, it can be observed that the locations of segment cracking coincide well with the zones of maximum tensile stress concentration in the reinforcement. As shown in Fig. 8(a), before the initiation of segment cracking, the stress distribution in the reinforcement remains relatively continuous. When P_1 reaches 230 kN, multiple cracks appear in segment D, leading to stress redistribution in both the segments and

the reinforcement, as illustrated in Fig. 8(b). This redistribution results in a pronounced local tensile stress concentration in the reinforcement at the invert, with a maximum stress of 98.84 MPa, which remains below the yield strength. With further loading, the number of cracks in the tensile zones of the tunnel increases, and the corresponding stress concentration regions in the internal reinforcement expand accordingly. As shown in Fig. 8(e), when P_1 reaches 420 kN, several distinct stress concentration zones develop in the reinforcement within the tensile regions, with a maximum tensile stress of 295.78 MPa. As illustrated in Fig. 8(f), at structural failure extensive tensile stress concentration occurs in the reinforcement, and the maximum tensile stress reaches 335 MPa, indicating that the reinforcement has yielded.

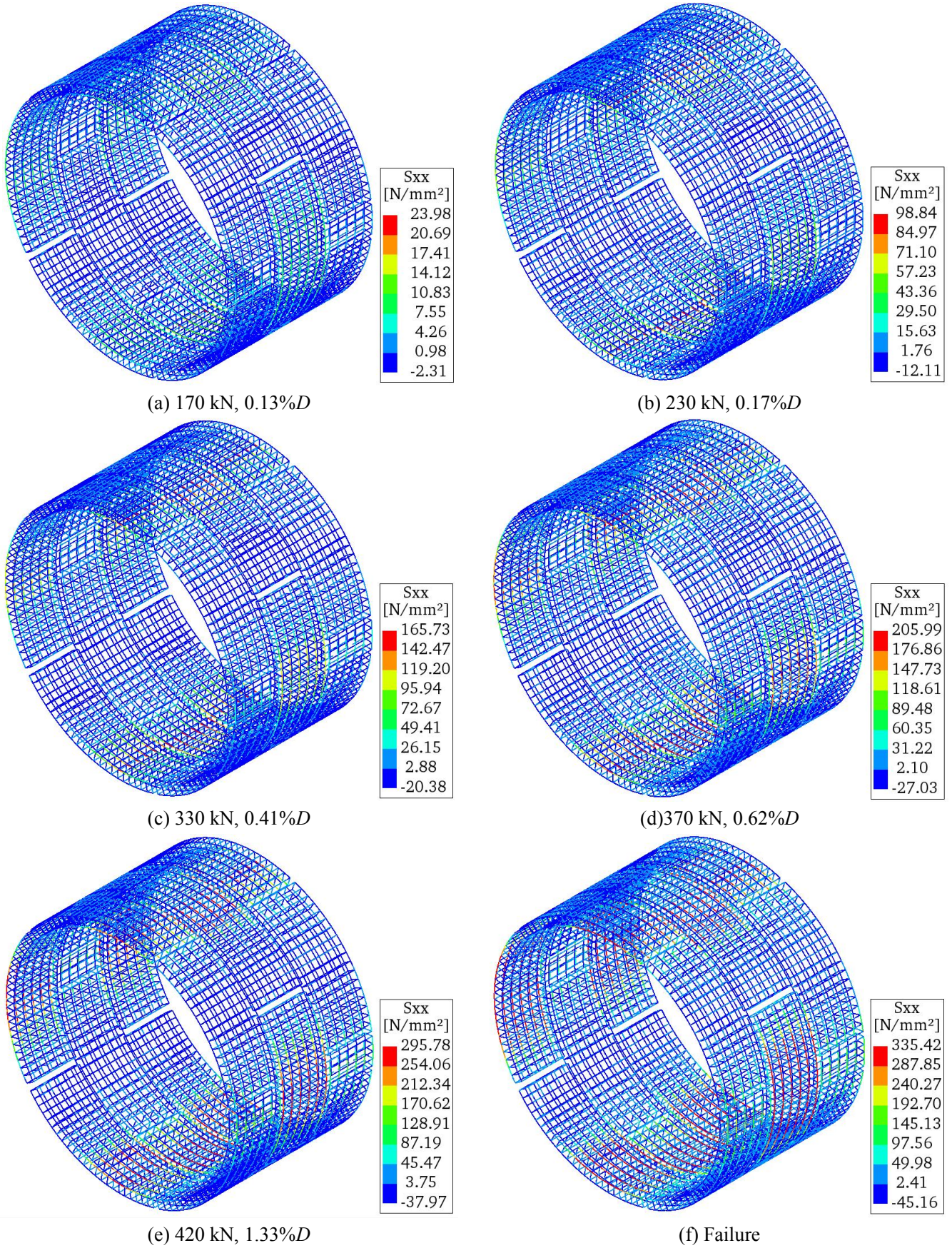


Fig. 8. Stress distribution of internal reinforcement in the shield tunnel

3. Shield Tunnel Damage Prediction

3.1 LSTM Neural Network

The service condition of a shield tunnel progressively deteriorates under the sustained action of ultimate loading, and this degradation process can be regarded as a time series that reflects the evolutionary characteristics of structural deterioration. Under ultimate loading, a long term dependency exists between the crack width of the shield tunnel and its structural condition. Therefore, a LSTM neural network is adopted in this study to perform time series analysis and to estimate the structural health state of the shield tunnel. The crack width data together with their temporal sequence are used as inputs for training and predicting

the degradation behavior of the tunnel structure. The LSTM neural network is a type of recurrent neural network (RNN), and its typical architecture is illustrated in Fig. 9. An LSTM network consists of multiple homogeneous memory cells, each of which can store information over long time spans by updating its internal state. In Fig. 9, module A represents a structure composed of three identical cells. Each cell mainly includes four core components: an input gate, a forget gate, an output gate, and a cell state. These components work collaboratively to enable the network to effectively capture long term dependencies in time series data, thereby accurately representing the damage evolution of shield tunnels under ultimate loading.

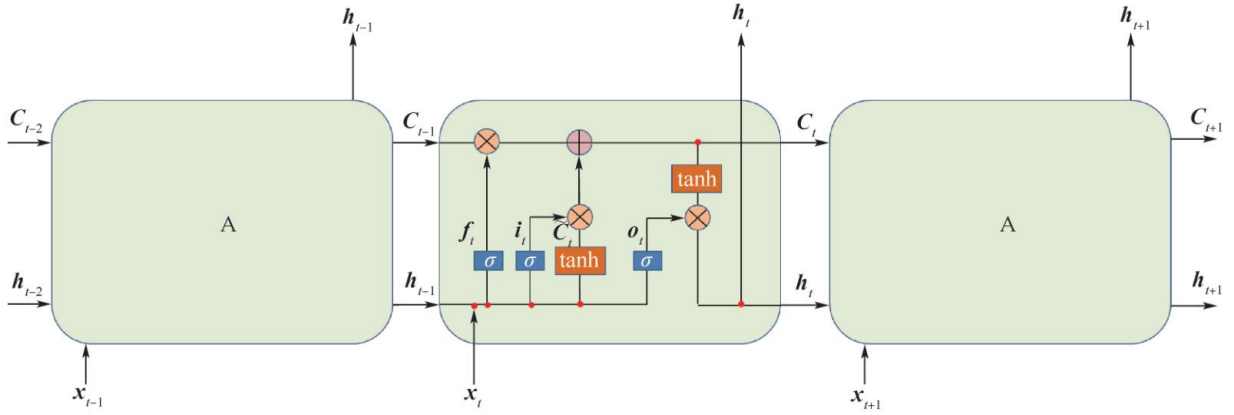


Fig. 9. Structure of an LSTM cell

In the LSTM network, the forget gate is used to discard redundant information, the input gate determines the key information to be stored in the cell state, and the output gate controls the information to be passed to the next layer. Through dynamic updating of the memory cell state, these three gates selectively add or remove information, allowing the LSTM network to retain and update critical features over long time horizons. This mechanism effectively alleviates the gradient vanishing problem commonly encountered in conventional RNNs. The data processing procedure of the LSTM network can be summarized as follows.

According to Eq. (6), the forget gate f_t discards unnecessary information from the previous cell state c_{t-1}

$$f_t = \sigma(w_f x_t + u_f h_{t-1} + b_f) \quad (6)$$

The input vector i_t and the $\tilde{c}_t$ cell state are updated using Eqs. (7) and (8), respectively.

$$i_t = \sigma(w_i x_t + u_i h_{t-1} + b_i) \quad (7)$$

$$\tilde{c}_t = \tanh(w_c x_t + u_c h_{t-1} + b_c) \quad (8)$$

The cell state c_t is then updated according to Eq. (9).

$$c_t = f_t c_{t-1} + i_t \tilde{c}_t \quad (9)$$

Finally, the network output h_t is generated based on the output gate o_t and the updated cell state c_t , as shown in Eqs. (10) and (11).

$$o_t = \sigma(w_o x_t + u_o h_{t-1} + b_o) \quad (10)$$

$$h_t = o_t \tanh(c_t) \quad (11)$$

In these equations, w denotes the input weight matrix, and u represents the recurrent weight matrix. The symbols f , i and o correspond to the forget gate, input gate, and output gate, respectively, while b is the bias

term. The activation function σ and the hyperbolic tangent function $\tanh$ are defined in Eqs. (12) and (13).

$$\sigma(x) = \frac{1}{1 + \exp(-x)} \quad (12)$$

$$\tanh(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)} \quad (13)$$

During LSTM training, the learning rate is a critical hyperparameter that directly influences the learning efficiency, convergence speed, and accuracy of the objective function. An appropriate learning rate enables the objective function to converge to an optimal solution within a reasonable number of training epochs. An excessively large learning rate may lead to oscillations during training, whereas an overly small learning rate can result in slow convergence or even overfitting. To enhance the predictive performance of the LSTM neural network, the PSO algorithm is employed in this study to optimize the learning rate, ensuring that the model converges to an optimal solution efficiently and stably during training.

3.2 Particle Swarm Optimization Algorithm

The PSO algorithm, also referred to as the swarm intelligence or bird flocking optimization algorithm, belongs to the class of evolutionary algorithms. It starts from a set of randomly generated solutions and iteratively approaches the global optimum. During each iteration, candidate solutions are evaluated using a fitness function, and the current best solution is used to guide the global search direction. Owing to its simple implementation, fast convergence, and relatively high accuracy, PSO has attracted extensive attention in both academic research and engineering practice, demonstrating excellent performance in a wide range of practical applications. The core concept of PSO lies in social information sharing, which is inspired by the collective foraging behavior of bird flocks. In the algorithm, each particle represents a potential solution in the search space and is characterized by two fundamental attributes. The particle position denotes its current state, while the particle velocity determines the direction and magnitude of its movement in the next iteration. Based on these principles, the global optimization process of PSO in a D dimensional search space can be summarized

into the following steps.

First, the particle swarm is initialized by defining parameters such as the population size N , and the termination criteria. The initial velocity v_i and position x_i of each particle i are randomly generated, as expressed in Eqs. (14) and (15).

$$v_i = (v_{i1}, v_{i2}, \dots, v_{iD}) \quad i = 1, 2, \dots, N \quad (14)$$

$$x_i = (x_{i1}, x_{i2}, \dots, x_{iD}) \quad i = 1, 2, \dots, N \quad (15)$$

Second, the fitness value $F[i]$ of each particle i is calculated according to the fitness function.

Third, the global best solution and the personal best solution of each particle are determined using Eqs. (16) and (17),

$$p(i) = \begin{cases} F[i] & F[i] < p(i) \\ p(i) & \text{others} \end{cases} \quad i = 1, 2, \dots, N \quad (16)$$

$$g = \begin{cases} F[i] & F[i] < g \\ g & \text{others} \end{cases} \quad i = 1, 2, \dots, N \quad (17)$$

Where $p_{\text{best}}(i)$ denotes the best solution achieved by particle i and g represents the global best solution of the swarm. Fourth, the velocity and position of each particle are updated according to Eqs. (18) and (19).

$$v_{id}^{k+1} = wv_{id}^k + c_1(p_{id}^k - x_{id}^k) + c_2s_{id}^k(g^k - x_{id}^k) \quad (18)$$

$$x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1} \quad (19)$$

In these equations, r and s are random constants, $k = 1, 2, \dots, H$ denotes the iteration index with H representing the maximum number of iterations, c_1 and c_2 are acceleration coefficients, and w is the inertia weight.

The above steps are repeatedly executed until the termination criterion is satisfied, at which point the algorithm stops.

3.3 Optimization of the LSTM Neural Network Using PSO

Under ultimate loading conditions, the crack width of shield tunnels is influenced by the coupled effects of multiple factors, resulting in a pronounced nonlinear evolution process with instability and periodic uncertainty. To achieve high accuracy prediction of crack width, an LSTM based prediction model is developed in

this study, taking advantage of the strong capability of LSTM networks in time series modeling. In LSTM models, several key hyperparameters directly determine the network structure and learning capacity. Meanwhile, crack width time series data exhibit strong complexity and nonstationary characteristics. To enhance the adaptability and generalization performance of the model with respect to crack evolution patterns, the PSO algorithm is introduced to optimize the hyperparameters of the LSTM model. By integrating an adaptive PSO algorithm with the LSTM network, a PSO-LSTM prediction model for shield tunnel crack width under ultimate loading conditions is established. The model development and implementation are carried out on the MATLAB platform. The overall workflow of the PSO-LSTM model is illustrated in Fig. 10, and the detailed implementation steps are described as follows.

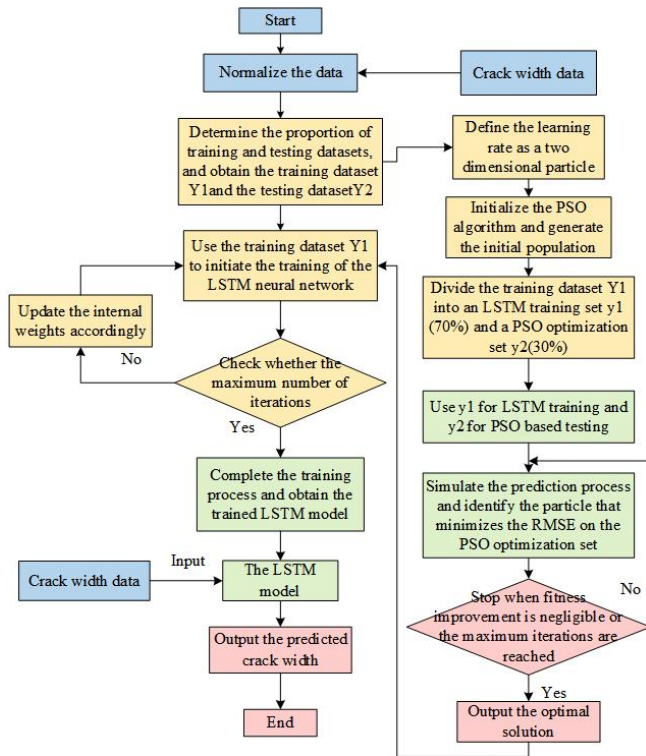


Fig. 10. Flowchart of the PSO-LSTM framework

(1) The original dataset is preprocessed and then divided into a training set and a testing set for model training and performance evaluation.

(2) The parameters of the PSO algorithm are initialized, including swarm size, particle dimensionality, maximum number of iterations, velocity and position

bounds, and termination criteria. Meanwhile, a feasible range of the LSTM learning rate is specified based on prior experience.

(3) During each iteration, the LSTM model is trained using the training dataset, while the PSO optimization dataset is used for validation. The LSTM prediction process is simulated, and within the particle swarm search, the particle that yields the minimum prediction error is selected as the current optimal solution.

(4) Once the termination criteria are satisfied, the global optimal solution obtained during the PSO search is output. The corresponding parameters are adopted as the optimal learning rate for the LSTM model and are used for subsequent crack width prediction to generate the final results. To quantitatively evaluate the predictive performance of the proposed model, the root mean square error (RMSE) and the coefficient of determination (R^2) are employed as evaluation metrics. The formulations of RMSE and R^2 are given as follows.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2} \quad (20)$$

$$R^2 = \frac{\sum_{i=1}^N (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (21)$$

3.4 Dataset construction and model training

3.4.1 Data Source

To comprehensively assess the catastrophic risk of shield tunnels and reduce potential hazards during future operation, this study focuses on the evolution of crack width in shield tunnel segments under ultimate bearing conditions. Fig. 11 presents the variation of crack width obtained from finite element simulations as a function of the analysis steps. As shown, the crack width exhibits a pronounced step like evolution during the loading process. This behavior is primarily attributed to the brittle nature of concrete. When the elastic strain energy accumulated in the segments exceeds the surface energy required for crack initiation, crack propagation occurs. The driving force for crack growth is the release of elastic strain energy stored in the segment concrete, and the entire process is characterized by repeated cycles of

energy accumulation and release, resulting in the stepwise increase in crack width. Ultimately, the crack width reaches a maximum value of 4.37 mm.

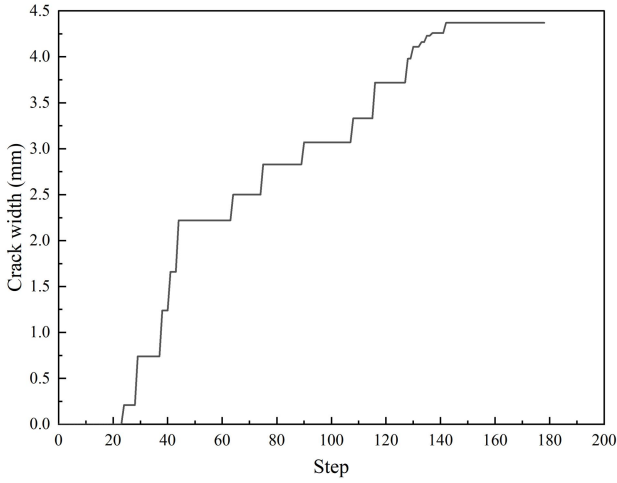


Fig. 11. Sample data of crack width evolution

In this study, the PSO-LSTM algorithm is employed to predict the evolution trend of crack width in shield tunnels. To verify the effectiveness and superiority of the proposed model, a systematic evaluation is conducted in terms of accuracy, precision, and reliability. The primary objectives are to elucidate the damage evolution mechanism of shield tunnels and to predict operational risks. Emphasis is placed on providing a scientific basis for risk management strategies through a robust and reliable predictive model.

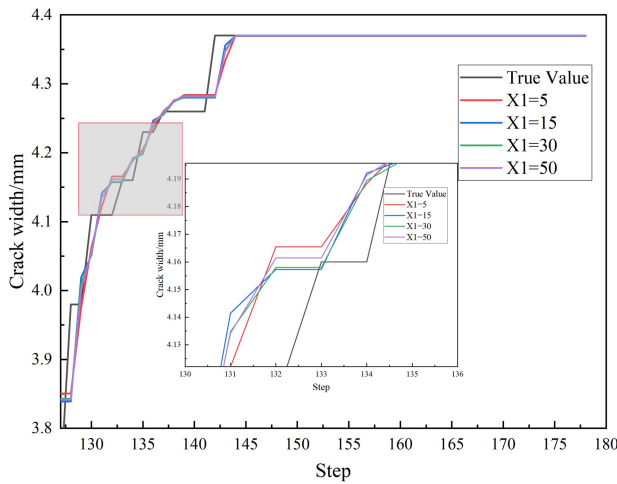


Fig. 12. Factor X1 sensitivity analysis

3.4.2 Parameter sensitivity analysis

To improve the prediction of shield tunnel crack width evolution, a hyperparameter sensitivity analysis was conducted for the proposed model and used to guide model configuration and prediction. The dataset was divided into training and prediction subsets to ensure sufficient learning while maintaining the reliability of the predicted results. The primary objective of the prediction is to assess the structural safety of the tunnel under its current operating condition, identify potential catastrophic risks, and support decisions on whether timely repair or reinforcement is required.

To enhance the predictive capability of the PSO-LSTM model for tunnel crack width, a systematic parameter sensitivity analysis was performed. This analysis was based on finite element simulation data validated by experimental results and focused on quantifying the influence of key hyperparameters on model performance, thereby identifying an optimal initial parameter configuration. Four dominant factors were examined: the number of particles in the particle swarm optimization algorithm X1, the number of PSO iterations X2, the number of LSTM training iterations X3, and the learning rate decay factor X4. The sensitivity analysis results are illustrated in Figs. 12 to 15 and summarized in Table 2, providing a quantitative basis for evaluating the influence of machine learning parameters on damage evolution risk prediction in shield tunnels.

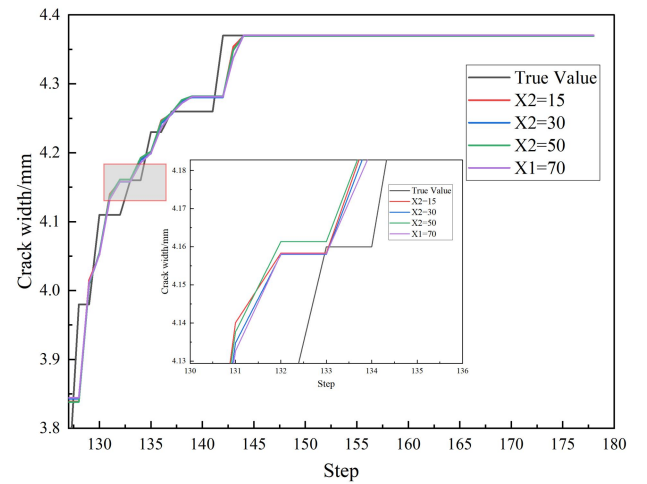


Fig. 13. Factor X2 sensitivity analysis

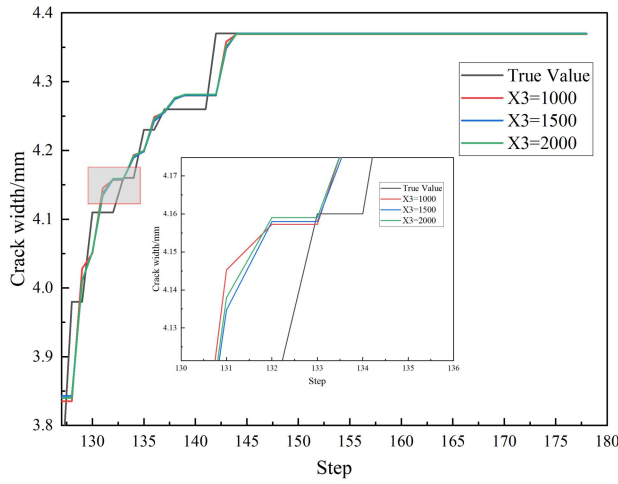


Fig. 14. Factor X3 sensitivity analysis

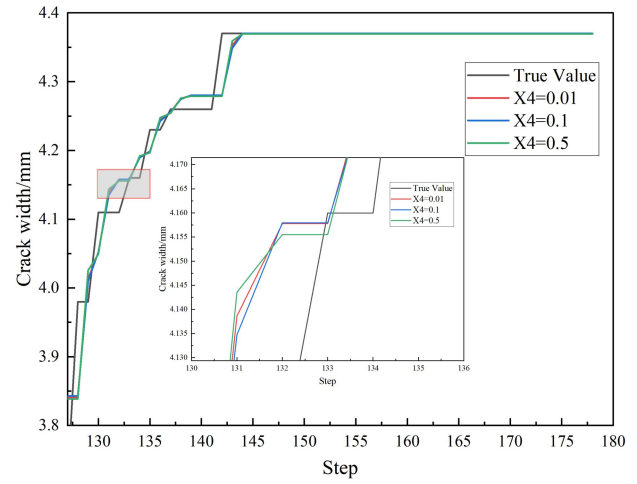


Fig. 15. Factor X4 sensitivity analysis

Tab.2. Sensitivity analysis comparison table

Group	X_1	X_2	X_3	X_4	RMSE	R^2
1	5	30	1500	0.1	0.1459	0.9420
	15	30	1500	0.1	0.1033	0.9487
	30	30	1500	0.1	0.4552	0.9402
	50	30	1500	0.1	0.2432	0.9408
2	15	15	1500	0.1	0.1064	0.9391
	15	50	1500	0.1	0.5247	0.9399
	15	70	1500	0.1	0.1099	0.9401
3	15	30	1000	0.1	0.1795	0.9370
	15	30	2000	0.1	0.2151	0.9398
4	15	30	1500	0.01	0.5239	0.9395
	15	30	1500	0.5	0.1480	0.9377

The results indicate that increasing the particle number X_1 does not necessarily improve prediction accuracy, and excessive particles may even degrade model performance. An optimal configuration of 15 particles was therefore identified. The number of PSO iterations X_2 has a pronounced effect on prediction accuracy. A value of 50 iterations resulted in relatively poor performance, whereas 30 iterations significantly improved prediction accuracy, leading to the selection of 30 as the optimal iteration number. In contrast, the number of LSTM iterations X_3 and the learning rate decay factor X_4 exerted only minor influence on both prediction accuracy and computational cost. These findings suggest that effective optimization of the PSO algorithm plays a dominant role in enhancing the predictive performance of the LSTM model.

3.4.3 Determination of training and prediction sets

When applying the PSO-LSTM model to predict crack width evolution in shield tunnels, the criteria for constructing the training and testing sets are critical. This process must account for the characteristics of crack width time series data, model complexity, and practical engineering requirements. The present study focuses on short term prediction, aiming to evaluate model performance on both the training and testing sets and to examine the influence of different data split ratios on predictive accuracy. Commonly used split ratios include 8:2, 7:3, and 6:4.

Based on the comparison of RMSE values calculated using Equation (20), the 7:3 split ratio yields the most favorable prediction performance. Accordingly, a two stage data partitioning strategy was adopted, with 70 percent of the data used for training and 30 percent reserved for testing. This scheme achieves a suitable

balance between prediction accuracy and data utilization, making it well suited for practical engineering applications. After determining the optimal split ratio, this configuration was applied to establish the training and prediction framework of the PSO-LSTM model.

3.4.4 Model and hyperparameter settings

In constructing the PSO-LSTM model, the swarm size was set to 15 particles and the maximum number of PSO iterations was fixed at 30. The learning rate was defined within the range of 0.01 to 0.1, the number of hidden layer neurons was varied from 10 to 120, and the L2 regularization coefficient ranged from 1×10^{-5} to 1×10^{-6} . These parameters jointly form the foundation of the PSO-LSTM model, with global optimization performed using the Adam algorithm. To enhance the reliability of the prediction results, the PSO-LSTM model was further compared with a back propagation neural network, a convolutional neural network, and a conventional LSTM model to identify the most effective prediction approach. For a fair comparison, all four models were configured with identical numbers of input units, output units, and training epochs, namely 1, 1, and 1500, respectively. The crack width values from analysis steps 1 to 127 were used as input, while the crack width values from steps 128 to 181 were used as prediction targets.

3.5 Training and prediction

Based on Equations (18) and (19), together with the hyperparameters optimized by the PSO algorithm, the evolution of fitness during the optimization process was

obtained, as illustrated in Fig. 16. The fitness value decreases progressively with increasing PSO iterations and eventually stabilizes, indicating good convergence of the optimization process. This curve provides a dynamic framework for parameter adaptability and offers deeper insight into the behavior and performance characteristics of the proposed model.

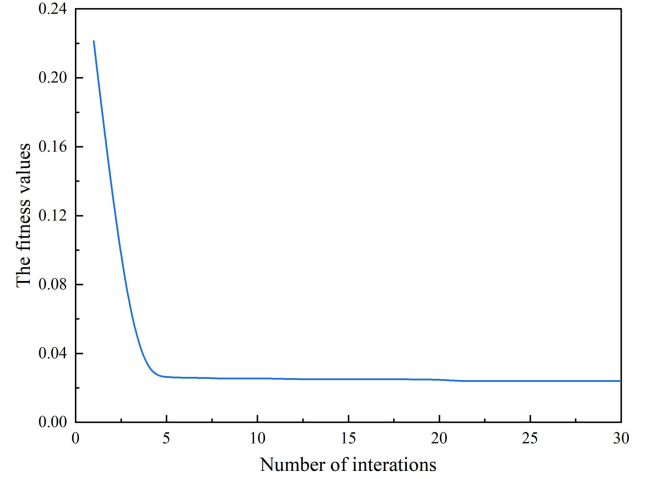
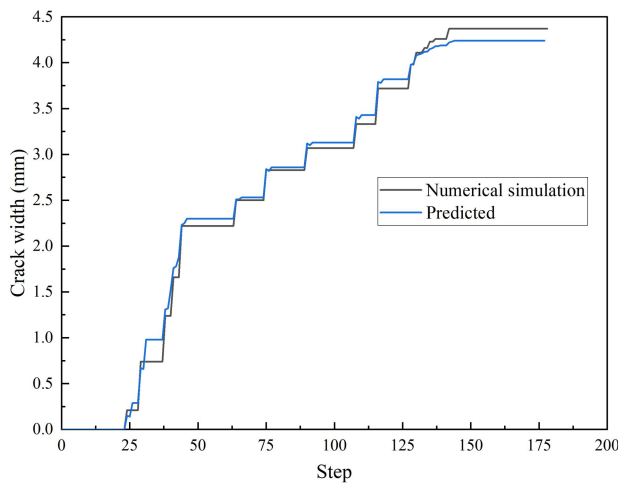
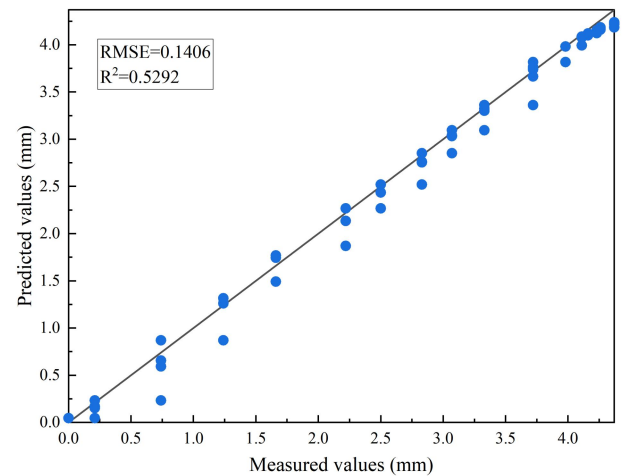


Fig. 16. Iteration of fitness values

During the training phase, experimentally validated finite element crack width data were adopted as the training dataset to train four prediction models, namely the back propagation neural network, CNN, LSTM, and PSO-LSTM. Each model was iteratively adjusted to achieve optimal performance. To further examine the differences between the prediction results and the numerical simulation data, evaluation index plots for the four models were generated, as shown in Figs 17 to 20.

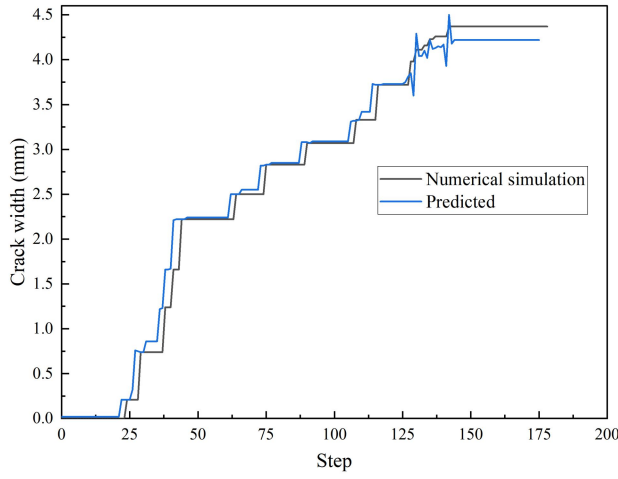


(a) Comparison of prediction results

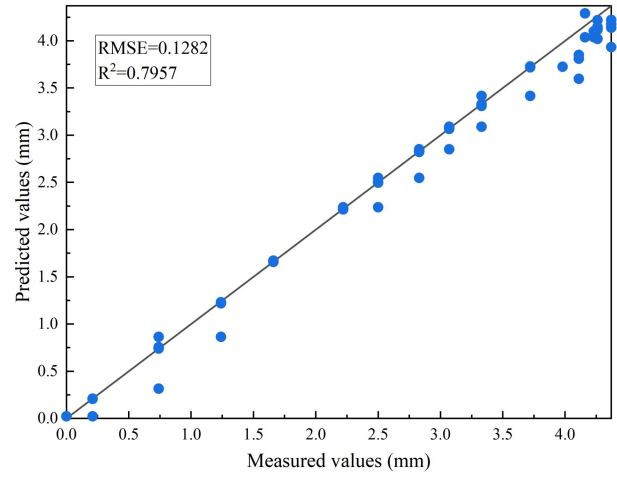


(b) Evaluation indexes results

Fig. 17. BP model

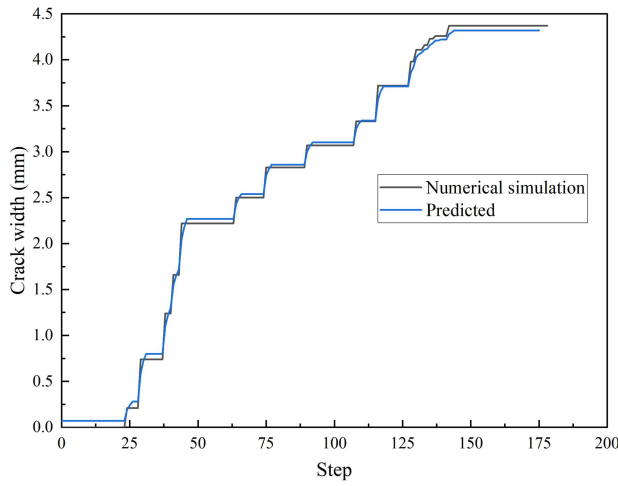


(a) Comparison of prediction results

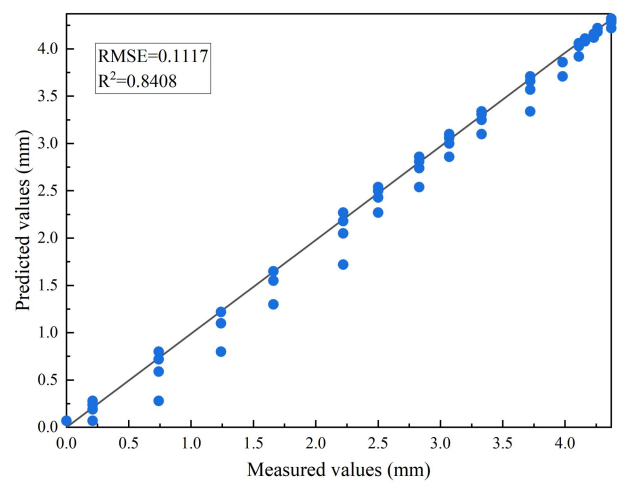


(b) Evaluation indexes results

Fig. 18. CNN model

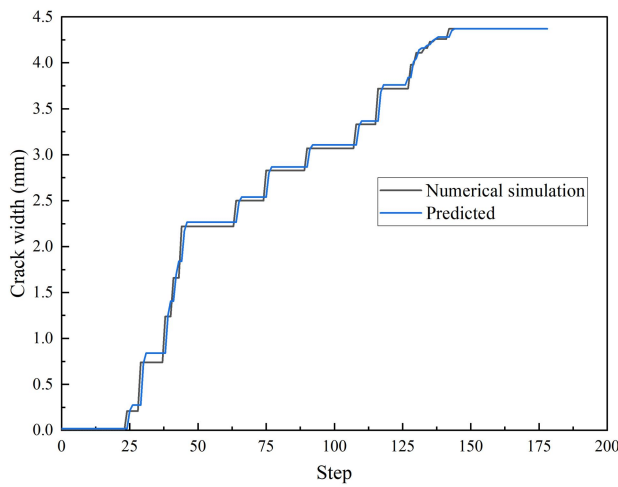


(a) Comparison of prediction results

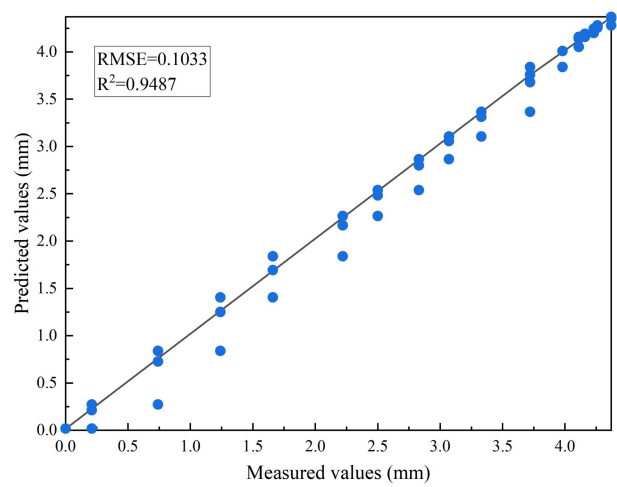


(b) Evaluation indexes results

Fig. 19. LSTM model



(a) Comparison of prediction results



(b) Evaluation indexes results

Fig. 20. PSO-LSTM model

The results indicate that all four models produce predictions generally consistent with the numerical simulation results, although clear performance differences exist. Among them, the proposed PSO-LSTM model exhibits the highest level of agreement with the numerical results, demonstrating a superior capability to capture the complex dynamic characteristics of damage evolution in shield tunnels. These findings confirm that the PSO-LSTM model can deliver predictions highly consistent with actual behavior, thereby providing timely and effective support for the operation and maintenance of shield tunnels. The comparative analysis indicates that the back propagation neural network exhibits pronounced discrepancies among evaluation metrics between the training set and the test set, suggesting the occurrence of overfitting during training and resulting in unsatisfactory predictive performance on the test data. In contrast, the CNN, LSTM, and PSO-LSTM models show good agreement between predicted values and reference data, with no evident overfitting behavior. Notably, the PSO-LSTM model achieves the best performance in predicting crack width for existing tunnels, with an RMSE of 0.1033 and an R^2 squared value of 0.9487, outperforming all other models considered. These results demonstrate that the PSO-LSTM model not only provides high fitting accuracy but also delivers superior predictive capability. It can reliably capture the damage evolution of shield tunnels under ultimate loading conditions, thereby offering a robust and credible basis for tunnel operation, and maintenance.

4. Conclusions

Based on the total strain crack model, a refined structural model of shield tunnel segments was developed to investigate the failure mechanism of shield tunnels under ultimate loading conditions. A machine learning framework that integrates particle swarm optimization with a long short term memory network was proposed and validated for predicting the damage evolution of shield tunnels at the ultimate bearing state. The main conclusions are summarized as follows.

(1) The failure responses of the shield tunnel structure simulated using the total strain crack model show good agreement with experimental observations. The model accurately captures the damage

characteristics of tunnel segments and effectively resolves millimeter scale damage features of shield tunnel linings.

(2) The structural failure process of the tunnel can be divided into three distinct stages, namely the linear elastic stage, the localized damage stage, and the global failure stage. The load thresholds corresponding to the linear elastic and localized damage stages are 230 kN and 420 kN, respectively, with associated crack widths of 0.11 mm and 3.16 mm, and convergence deformations of $0.17\%D$ and $1.33\%D$. During the global failure stage, the crack width increases to 4.37 mm, accompanied by pronounced and widespread tensile stress concentration in the reinforcement, indicating that the reinforcing steel has reached its yield strength.

(3) The PSO-LSTM model is capable of accurately capturing the damage evolution of shield tunnel structures under ultimate loading conditions. By optimizing key LSTM hyperparameters through the PSO algorithm, including the number of hidden layer neurons, the learning rate, and the L2 regularization coefficient, the predictive accuracy of the model is significantly enhanced, providing a reliable basis for tunnel operation, and maintenance.

(4) In the prediction of shield tunnel crack width, the PSO-LSTM model achieved an RMSE of 0.1033 and an R^2 squared value of 0.9487, demonstrating higher predictive accuracy and superior stability compared with the BP neural network, CNN, and conventional LSTM models.

References:

- [1] Hung La, Toan Nguyen-Minh, Tan Nguyen. Optimized prediction of tunnel stability using advanced machine learning and an ANN-based analytical expression. *Tunnelling and Underground Space Technology*, 2025, 164: 106778.
- [2] Zhang Dongming, Chen Hui, Xue Yadong, Zhang Jinzhang. Vulnerability analysis of shield tunnel segmental lining disturbed by adjacent deep excavations in soft soils. *Tunnelling and Underground Space Technology*, 2026, 167: 107033.
- [3] Jiasuo Qi, Xu Zhao, Jingqi Huang, M.Hesham El Nagggar, Xiuli Du, Mi Zhao. Seismic performance and fragility analysis for shield tunnel segments .

-
- Tunnelling and Underground Space Technology, 2026, 167: 107044.
- [4] Liu Jianwen, Shi Chenghua, Gong Chenjie, Lei Mingfeng, Wang Zuxian, Peng Zhu, Cao Chengyong. Investigation of ultimate bearing capacity of shield tunnel based on concrete damage model. *Tunnelling and Underground Space Technology*, 2022, 125: 104510.
- [5] Gong Chenjie, Xie Chaoran, Zhu Hehua, Yan Zhiguo. Multi-criteria evaluation model of rubber sealing gasket for shield tunnel joint. *Tunnelling and Underground Space Technology*, 2026, 168: 107127.
- [6] Yu Yang, Wang Jinchang. Crack mechanism of shield tunnel under excavation of foundation pit. *AIP Advances*, 2025, 15(2), 0249750.
- [7] Kan Huang, Yiwei Sun, Hongzhi Cui, Xiangsheng Chen, Dongmei Zhang, Bin Huang, Zhongkai Huang, Binlin Gan. Analysis of bearing capacity characteristics and resilience enhancement mechanism in shield tunnel segments based on fracture energy and modulus degradation. *Tunnelling and Underground Space Technology*, 2026, 167: 106952.
- [8] Linlin Yang, Wenbo Yang, Yang Zhou, Zefeng Chen, Chaofan Yao. An experimental study of the influence of local segment damage on dynamic response characteristics of shield tunnels. *Tunnelling and Underground Space Technology*, 2025, 166: 106976.
- [9] Weiming Huang, Yanwei Zang, Jinchang Wang, Changbao Liu, Zhongxuan Yang, Rongqiao Xu, Huajian Fang. General Solution for Longitudinal Response of Shield Tunnel Considering the Effects of Joints and Soil Shear Resistance. *International Journal for Numerical and Analytical Methods in Geomechanics*, 2025, 49: 948–960.
- [10] CHEN Zhangpeng, ZANG Yanwei, YANG Zhongxuan, XU Rongqiao, GONG Xiaonan, WANG Jinchang, LIU Changbao, YAN Jiaji. Analytical Solution of Tunnel Linings Subjected to Opening Condition and Its Application. *China Journal of Highway and Transport*, 2024, 37(7): 189-203.
- [11] Haolan Feng, Dongming Zhang, Hao Bai, Yelu Zhou, Hongwei Huang. Analytical solution for shield tunnel structural behavior considering nonlinear shear stiffness of circumferential joint. *Tunnelling and Underground Space Technology*, 2026, 168: 107138.
- [12] DongLin Feng, HuaiNa Wu, Jelena Ninić, SiQi Yang, RenPeng Chen. An analytical model of multi-ring segmental lining with nonlinear joint interactions. *Tunnelling and Underground Space Technology*, 2026, 169: 107249.
- [13] Zhu Ming, Chen Xiangsheng, Xia Changqing, Wang Chen, Bao Xiaohua. Resilience evolution of shield tunnel structures underground surcharge. *Chinese Journal of Geotechnical Engineering*, 2024, 46(1), 35–44.
- [14] Sun Yazhen, Yu Yang, Wang Jinchang, Wang Longyan. Interface failure of segmental tunnel lining strengthened with steel plates based on fracture mechanics. *Frontiers of Structural and Civil Engineering*, 2024, 18(1), 137–149.
- [15] Xian Liu, Jianyu Hong, Zhen Liu, Ba Trung Cao. Mechanical behavior and failure mechanisms of shield tunnel linings reinforced with lightweight epoxy bonded-bolted steel plates. *Tunnelling and Underground Space Technology*, 2026, 168: 107137.
- [16] Yu Yang, Wang Jinchang, Hu Xinzan. Damage Evolution and Failure Mechanism of Segmental Tunnel Lining. *Civil Engineering Journal-Tehran*. 2025, 11(8): 3108-3125.
- [17] Yang Yu, Jinchang Wang, Chuantao Cheng. Fracture and resilience evolution of double-layer lining structure of shield tunnel. *Engineering Computations*, 2025, 42(5):1771-1790.
- [18] Yifan Chen, Haibin Zhang, Xiang Shen, Xiangsheng Chen, Dong Su, Jiuqi Wu. A novel deep learning-based identification technology of cutting pile states during super-large diameter shield tunnelling. *Tunnelling and Underground Space Technology*, 2025, 164: 106836.
- [19] Xianlong Wu, Xiaohua Bao, Jun Shen, Xiangsheng Chen. Seismic resilience assessment of shield tunnels under contact loss defects using a hybrid neural network model driven by machine learning and numerical simulation. *Tunnelling and Underground Space Technology*, 2025, 164: 106801.
- [20] Changyu Wang, Zude Ding, Yuhang Shen, Wenyun Ding, Yongfa Guo. Prediction of surface settlement caused by tunneling with ARMA based time-series

decomposition. *Tunnelling and Underground Space Technology*, 2025, 164: 106770.

- [21] Zhou Zhong, Zhang Junjie, Ding Haohui, Li Fan. Settlement prediction model of shield tunnel under-crossing existing tunnel based on GA-Bi-LSTM. *Chinese Journal of Rock Mechanics and Engineering*, 2023, 42 (1): 224-234.
- [22] Sun YZ, Lin K, Wang JC, Zhu FB, Wang LY, Lu LH. Risk assessment of mountain tunnel entrance collapse based on PSO-LSTM surface settlement prediction. *Engineering, Construction and Architectural Management*, 2025. DOI10.1108/ECAM-08-2024-1047.
- [23] XingTao Lin, QiRui Lin, RuiDi Chen, Xiangsheng Chen, Dong Su. Physics-informed neural network for predicting longitudinal deformation in super-large-diameter shallow-buried shield tunnel during construction. *Tunnelling and Underground Space Technology*, 2026, 169: 107259.
- [24] RuiDi Chen, Xing-Tao Lin, Xiangsheng Chen, Hui Zeng. Data-and experience-driven prediction method of twin shield tunneling-induced complicated settlement curve and its evolution. *Tunnelling and Underground Space Technology*, 2026, 164: 106824.