

Research Article

Prediction of Seepage Pressure Based on Memory Cells and Significance Analysis of Influencing Factors

Mengdie Zhao,^{1,2,3,4} Haifeng Jiang,^{1,2} Shoukai Chen ,^{1,2,3} and Yajing Bie^{1,2,3}

¹School of Water Conservancy, North China University of Water Resources and Electric Power, Zhengzhou 450045, China

²Henan Key Laboratory of Water Environment Simulation and Treatment, Zhengzhou 450045, China

³Collaborative Innovation Center of Water Resources Efficient Utilization and Protection Engineering, Zhengzhou 450045, China

⁴Henan Key Laboratory of Water Resources Conservation and Intensive Utilization in the Yellow River Basin, Zhengzhou 450046, China

Correspondence should be addressed to Shoukai Chen; man200177@163.com

Received 4 February 2021; Revised 3 March 2021; Accepted 17 March 2021; Published 12 April 2021

Academic Editor: Muhammad Ahmad

Copyright © 2021 Mengdie Zhao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Seepage analysis is always a concern in dam safety and stability research. The prediction and analysis of seepage pressure monitoring data is an effective way to ensure the safety and stability of dam seepage. With the timeliness of a change in a monitoring value and lag due to external influences, a RS-LSTM model written in Python is developed in this paper which combines rough set theory (RS) and the long- and short-term memory network model (LSTM). The model proposed calculates the prediction score of the seepage pressure of a dam experiencing multiple effects by preordering factor importance values to eliminate the interference of redundant factors. A case study shows that the water level, rainfall, temperature, and duration are all factors that affect the seepage pressure, and their importance values decrease successively. Thus, the seepage pressure of a dam can be predicted with a determination coefficient R^2 of 0.96. Compared with the recurrent neural network (RNN) model and BP neural network model, the training time of the RS-LSTM model proposed is 6.37 s, and the operation efficiency is 41% and 59% higher than that of the RNN and BP models, respectively. The mean relative error is also 3.00%, which is 50% lower than that of the RNN model and 31% lower than that of the BP model. Based on these results, this model has the advantages of fast computation speed and high accuracy in prediction.

1. Introduction

Reservoir dams are an important infrastructure that play an important role in the development of economies and societies. However, with the recent rapid increase in the construction of reservoir dams, many safety accidents have followed [1]. From 1954 to 2006, 3498 dams broke in China alone, and more than 93% of them were earth rock dam failures [2]. Based on the material characteristics and operation mechanism of earth rock dams, the primary failure modes are overtopping, landslide instability, and seepage damage [3]. Combined with relevant data, the number of dam failures caused by seepage failure in China is approximately 39% of the total [4]. Similarly, foreign survey data demonstrate that approximately 46% of earth rock dam

failures are caused by infiltration damage, ranking only second to overtopping failure (approximately 48%) [5–8]. Therefore, the influence of seepage failure on the safety of earth rock dams is self-evident, and its significance is second to none.

Seepage failure primarily refers to the reduction of soil effective stress and seepage deformation due to the movement of soil particles inside the dam due to the action of seepage pressure, which decreases the stability of the dam or even destroys the dam. Therefore, the establishment of adequate seepage safety monitoring is an essential prerequisite to ensure seepage safety: prototype observation data are used for analysis, and a seepage monitoring model is developed to monitor and evaluate the operational conditions of water conservancy projects.

For seepage instability, influencing factors primarily include water level change, rainfall, temperature change, and duration. For example, Xu et al. [9] considered that the seepage flow of dam bodies is primarily caused by the superposition of reservoir water level and rainfall, and the seepage head and reservoir water level tend to be positively correlated. Guangjin et al. [10] and Ashraf et al. [11] studied variations in the seepage head of a soil slope due to changing rainfall intensity and duration. Testing indicated that as rainfall intensity and duration increased, the deeper the influence of rainfall factors on soil, the greater the change in the seepage head and the greater the change in seepage pressure. Xu and Chai [12] investigated the influence of changes in temperature on the permeability coefficient and calculated water heads in seepage fields. Compared to the calculated results without a change in temperature, their findings show that the former is more accurate when calculating the distribution of water heads in seepage fields. However, most studies to date have focused on the influence of changes in influencing factors on seepage under single or a combination of conditions, which only verifies that there is a significant correlation between these factors and the seepage field, namely, the measured seepage pressure value, making it difficult to distinguish the correlation or importance between the two. Based on existing monitoring results of seepage pressure, a change in the measured seepage pressure over a long time period is typically complex and variable; thus, it is difficult to directly determine how seepage pressure changes over time. There also tends to be a long-time delay in response to changes in external conditions. Therefore, it is important to determine the effect of various factors and consider the dynamic change characteristics of seepage pressure monitoring for analyzing the change law of seepage pressure value of long sequence of dam and realizing the related prediction.

Currently, there are many methods to establish mathematical models to predict seepage monitoring data. Data fitting and analysis can be achieved via multiple regression; however, the accuracy of this method is restricted by the collinearity between input factors [13]. Predictions based on neural network models have also been reported; for example, Zhang et al. [14] used a genetic algorithm (GA) to optimize the weights and thresholds of a backpropagation neural network (BPNN), thus establishing the BPNN-GA seepage prediction model. Shi et al. [15], Rankovic et al. [16], and others used a variety of different neural network models, including radial basis neural network models, to achieve accurate predictions of dam seepage. These models provide various options to predict the measured seepage pressure using a neural network model; however, such static models are primarily based on correlations between the seepage pressure and associated factors and describe the lag in seepage pressure in response to changes in the factors via the averaging method using existing data. Also, these static models cannot describe the timeliness of the seepage pressure of a dam and thus cannot describe the memory function of the historical information hidden in the data, maintain the information, and output it to the current neuron for calculation. Thus, a predictive model developed based on dynamic changes in relevant factors will be useful for practical engineering applications.

A predictive model of seepage pressure is established in this paper that considers the effects of multiple factors and dynamic changes, integrates rough set theory and dynamic neural network technology, and examines the advantages of the dynamic model based on the distribution of error and the determination coefficient R^2 [17]. Model performance is characterized by comparing its structural characteristics and predictive results to those of other models based on engineering examples to provide references for the prediction of the seepage pressure value and engineering safety assessments.

2. Materials and Methods

2.1. Project Overview. A concrete-faced rockfill dam is investigated in the paper, with a project scale of Class A Large (1). The elevation of the dam crest is 413.8 m with a maximum dam height of 99.8 m, a dam crest width of 10 m, and a dam crest length of 540.46 m. An L-shaped reinforced concrete wave wall with a height of 4 m is set at the upstream side of the dam crest. The normal water level of the reservoir is 410 m, and the storage capacity is 11,076,000 m³. To monitor dam safety, a comprehensive prototype dam monitoring system that is composed of deformation monitoring, seepage monitoring, and environmental monitoring (e.g., rainfall and reservoir water level) and other equipment is installed at the reservoir. The seepage monitoring equipment uses buried and installed osmometers; four osmometers are installed in total with a data-collection interval of 1–3 d, as shown in Figure 1.

Considering osmometer PB4 as an example, changes in the data over the last 8 years from 2008 to 2015 are given in Figure 2. The seepage pressure measured fluctuates within a relatively stable range and varies with changes of water level, rainfall, temperature, and other factors; thus, the monitoring results are reasonable. The rainy season is concentrated in July and August every year and exhibits an annual cycle. The maximum rainfall during the monitoring period was 95.50 mm on July 8, 2012. The temperature change between years is relatively stable; however, the change within a given year is major, with the high value of approximately 37.8°C and the low value of -9.70°C. During the monitoring period, the change in the water level upstream and downstream is minor, and the water level upstream varies between 385 and 410 m. Compared to the periodic alterations in the rainfall and temperature, the change in the seepage pressure is more moderate and lags due to the influence of the level of seepage pressure at the front of the dam. For example, from January 2011 to March 2011, the water level of the reservoir gradually decreased by 4 to 5 m; however, the change in the seepage pressure started at the end of March 2011 and gradually decreased from approximately 0.1 to 0.3 m (see Figure 2(a)).

2.2. Methods

2.2.1. Rough Set Theory. Rough set (RS) theory is a theoretical method that can achieve the same classification ability of a knowledge base via attribute reduction to solve problems or apply classification rules [18, 19] and is typically

represented by a decision information table $S = \langle U, C \cup D, V, f \rangle$, where U is the domain or set of nonempty finite objects; $C \cup D$ is the attribute set; C is the conditional attribute; D is the decision attribute, in which $C \cap D = \emptyset$ is satisfied; V is the set of all attribute value domains; and $f: U \times R \rightarrow V$ represents the attribute value of each object in the domain on the corresponding attribute. The reduction process of the influencing factors of the real seepage pressure is described as follows:

- (1) The seepage pressure value domain U , which is composed of the seepage pressure and the related influencing factor data, is divided into two equivalent classes by the Pawlak attribute importance method, which is $\text{IND}(C) = \bigvee_{i,j,k,n \in N} \{\{C_i\}, \dots, \{C_j, \dots, C_k\}\}_n$ and $\text{IND}(D) = \bigvee_{i,j,k,n \in N} \{\{D_i\}, \dots, \{D_j, \dots, D_k\}\}_n$ respectively.
- (2) $\text{pos}_C(D)$ is calculated based on the dependence of the seepage pressure value D on the influencing factor set C , and the formula is as follows:

$$\text{pos}_C(D) = \bigcup_{X \in \text{IND}(D)} \underline{C}(X), \quad (1)$$

$$\text{pos}_C(D) = \frac{\sum_{i=1}^k |\text{pos}_C(D_i)|}{|U|}. \quad (2)$$

- (3) Removing the importance of a given influencing factor c_i from the seepage pressure D can be defined as

$$\text{pos}_{C-\{c_i\}}(D) = \frac{\sum_{i=1}^k |\text{pos}_{C-\{c_i\}}(D_i)|}{|U|}. \quad (3)$$

- (4) The attribute importance of the influencing factor c_i ($c_i \in C$) is

$$\sigma_{C D}(c_i) = \text{pos}_C(D) - \text{pos}_{C-\{c_i\}}(D). \quad (4)$$

Formula (4) represents the degree to which the single factor c_i affects the classification of the seepage pressure after factor removal: the larger the value $\sigma_{C D}(c_i)$ is, the greater the importance of the influencing factor is, and vice versa.

The seepage characteristics of dams are affected by multiple factors. However, due to the different importance of various factors regarding seepage, factors with little affect or irrelevant redundancy can be screened out by this method to facilitate the prediction of dynamic neural networks.

2.2.2. Neural Network Dynamic Model/Long- and Short-Term Memory Network Model. The RS method can be employed to screen the influencing factors of seepage pressure; however, it cannot obtain accurate predictions. Considering the timeliness of changes in seepage pressure and the lag between influencing factors, the long- and short-term memory network model (LSTM) is introduced [20]. The LSTM model, which is an improved form of recurrent neural networks (RNNs), is a deep learning model that can

process time series data [21]. Compared to the BP neural network model (BP model), the nodes between the hidden layers in LSTM are connected to each other; the adjacent nodes of the hidden layers are connected [15], which can describe the memory function, retain hidden information in the historical data, and output the data to the current neuron for data calculation. These nodes are also constantly updated with input data (see Figure 3). Compared to RNN, LSTM includes a memory unit to solve the problem of gradient disappearance or explosion during prediction [22] to express the hidden historical information in the data more accurately. The memory unit of the model consists of an input gate, forget gate, and output gate, which are utilized to control information transmission at different times. The input gate controls the strength of the new input heading into the memory unit and determines how many new memories will be merged with previous memories. The forget gate manages the strength of the memory unit to maintain the value at the last moment (i.e., selects or rejects historical information). If the forget gate is closed, no memory can pass; otherwise, all memories can come through. The output gate controls the strength of the output memory unit and determines the LSTM response to the outside world [14]. The descriptions in detail can be figured out in Figure 4.

The aforementioned function equation of propagation in each mechanism during LSTM training is described as follows [23–25]:

$$\begin{aligned} f_t &= g(W_{f1}C_{\text{LSTM}}^t + W_{f2}D_{\text{LSTM}}^{t-1} + b_f), \\ I_t &= g(W_{I1}C_{\text{LSTM}}^t + W_{I2}D_{\text{LSTM}}^{t-1} + b_I), \\ \tilde{S}_t &= \tan h(W_{c1}C_{\text{LSTM}}^t + W_{c2}D_{\text{LSTM}}^{t-1} + b_S), \\ S_t &= f_t \times S_{t-1} + I_t \times \tilde{S}_t, \\ O_t &= g(W_{o1}C_{\text{LSTM}}^t + W_{o2}D_{\text{LSTM}}^{t-1} + b_o), \\ h_t &= O_t \times \tan h(S_t), \end{aligned} \quad (5)$$

where I_t , f_t , and O_t are vectors describing the input gate, forget gate, and output gate of the model at time t , respectively; C_{LSTM}^t is the set of influencing factors of the seepage pressure after RS reduction at time t ; D_{LSTM}^{t-1} is the seepage pressure at time $t-1$, representing the hidden historical information of LSTM; W_1 is the connection weight between the input layer and the hidden layer; W_2 is the connection weight between the hidden layer and the output layer; b_f , b_I , b_o , and b_S are bias terms that correspond to each structure, respectively; S_t is the vector of the memory unit at time t and is only used for memory units to forget old information and add new information; \tilde{S}_t is a new candidate vector value created for \tanh , which is the hyperbolic tangent function; and g is an activation function, which can map real numbers to $[0, 1]$, where 1 indicates that all information in the unit at the previous time is reserved and 0 represents that all information in the unit at the previous time is discarded.

2.2.3. RS-LSTM Model Structure. Dam seepage is complicated and typically affected by many factors. To demonstrate

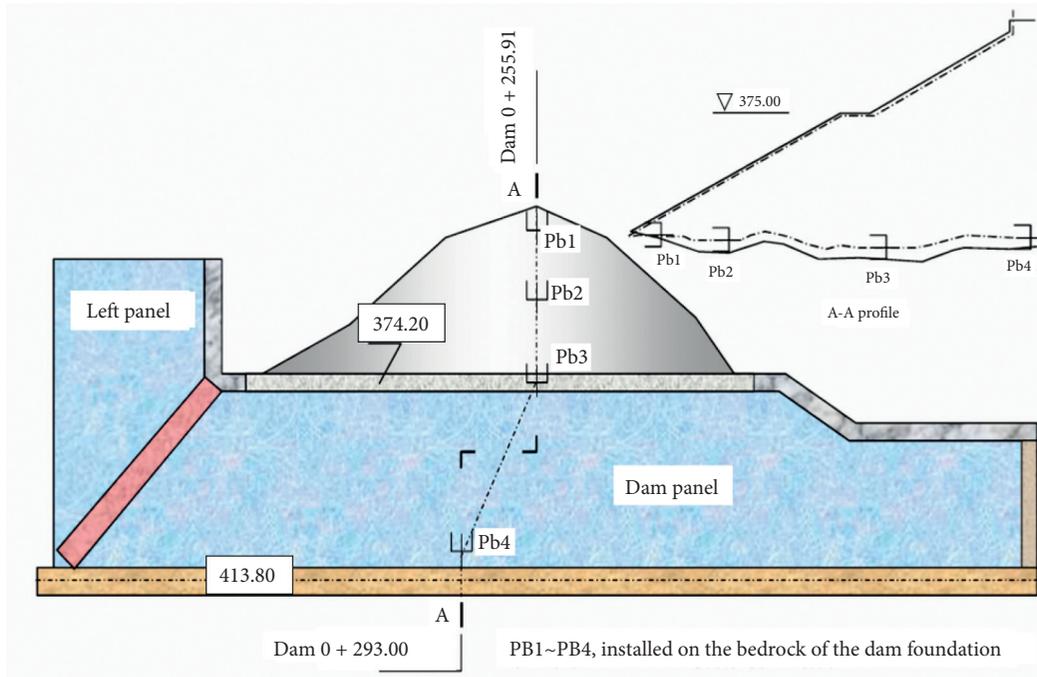


FIGURE 1: Arrangement of seepage pressure measuring instruments for dam foundations.

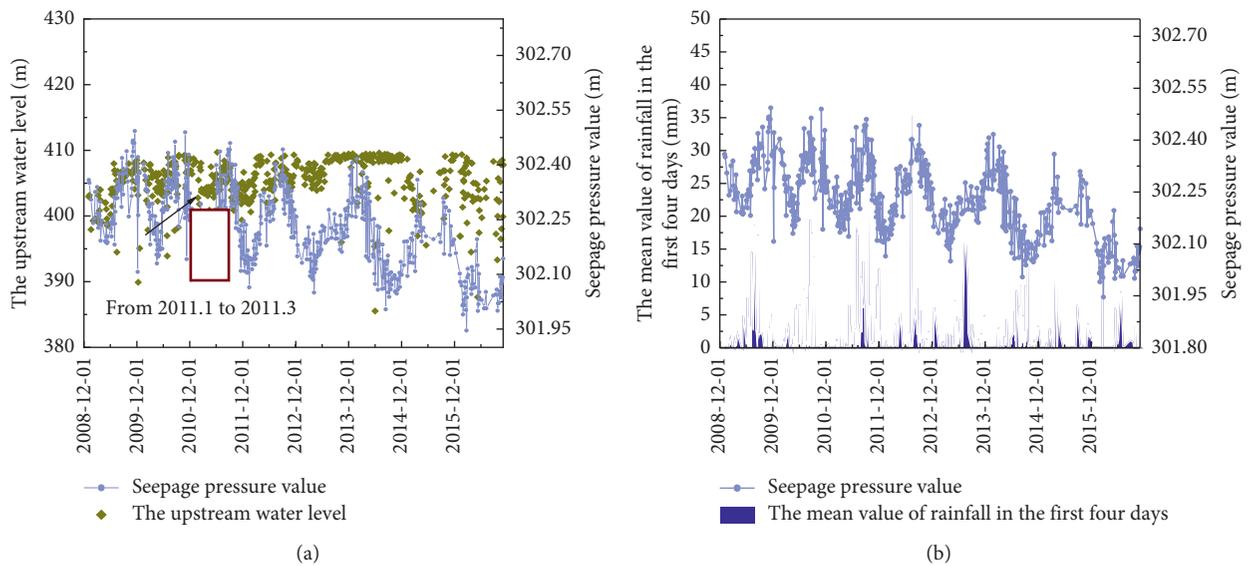


FIGURE 2: Continued.

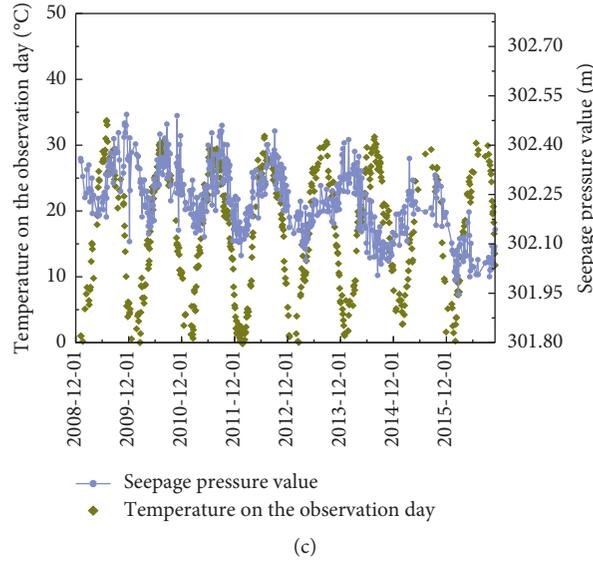


FIGURE 2: Long series change curve of permeability pressure value. (a) Relationship between the upstream water level and the seepage pressure. (b) Change in the seepage pressure due to rainfall. (c) Change in the osmotic pressure due to temperature.

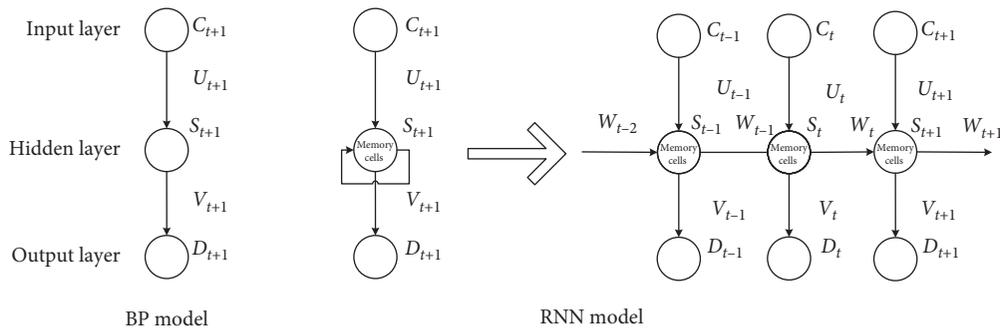


FIGURE 3: Static and dynamic model comparison.

the superior performance of the two methods proposed in data analysis, RS reduction is adopted before the LSTM prediction and analysis, and the RS-LSTM model structure is constructed as illustrated in Figure 5 to achieve the analysis and prediction of dam seepage in the operating period under the influence of multiple factors.

Based on the RS analysis results, an LSTM prediction model that considers various influencing factors is built to predict the seepage pressure. Two hidden layers are selected, and the number of hidden layer neurons is determined by the following formula:

$$l \leq \sqrt{m+n} + \alpha, \quad (6)$$

where l is the number of nodes in the hidden layers; m is the number of nodes in the output layers; n is the number of nodes in the input layers; and α is the adjustment constant between 1 and 10. To determine the best prediction model structure after multiple calculations and training, training was found to be most effective when the number of nodes in

the hidden layers is 10; thus, the neural network topology structure is 6-10-1, which is displayed in Figure 6.

Considering PB4 as an example, the sample data contain water level, temperature, rainfall, duration, and monitoring values and have a total of 8,673 groups, including 1,239 groups of seepage pressure monitoring values. A total of 196 groups of data (2% of the total data set) were selected as prediction samples, 28 groups of data were selected as output samples for prediction, and the remaining were selected as training samples. The number of iterations was set to 5,000, the error range was set to 0.001, and the correction factor was set to 0.02.

3. Results and Discussion

3.1. Application Results and Analysis of RS. Based on existing literatures [9–12], seven influencing factors are considered for prediction and analysis in this study, including the upstream water level, downstream water level and temperature on the observation day, the mean rainfall in the first

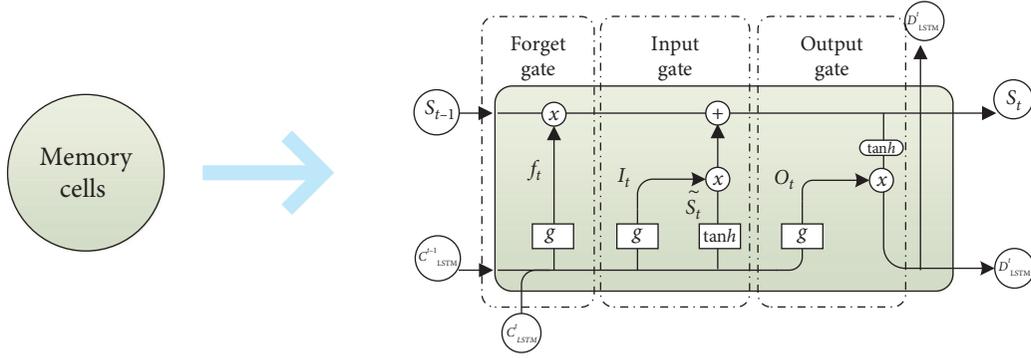


FIGURE 4: Structural model diagram of the LSTM memory unit.

four days, and the time-dependent components θ , $e^{1+\theta}$, and $\ln(1+\theta)$. These factors are recorded as X_1 – X_7 , respectively, in the influencing factor set C , in which $C_1: X_2$ – X_7 ; $C_2: X_1$, X_3 – X_7 ; $C_3: X_1$ – X_2 , X_4 – X_7 ; $C_4: X_1$ – X_3 , X_5 – X_7 ; $C_5: X_1$ – X_4 , X_6 – X_7 ; $C_6: X_1$ – X_5 , X_7 ; and $C_7: X_1$ – X_6 . The seepage pressure is set as decision attribute D . After the normalization and discretization of C and D , the decision sequence is developed. The discretization is achieved by the method of equidistance; thus, after the normalization of the decision sequence value, they are discretized into five levels, i.e., $[0, 0.2)$, $[0.2, 0.4)$, $[0.4, 0.6)$, $[0.6, 0.8)$, and $[0.8, 1]$, which are recorded as 1, 2, 3, 4, and 5, respectively. After merging similar values, 100 groups of data to be deduced are obtained and recorded as U_1 – U_{100} (see Figure 7(a) for details; according to the decision attribute D , it can be divided into normal data and abnormal data). These groups of data can be reduced using formulae (1)–(3), and the results are presented in Figure 7(b), which show that the importance of X_6 is 0; X_1 , X_2 , X_3 , X_4 , X_5 , and X_7 are all necessary attributes, and their importance order can be obtained. The importance of the water level upstream and downstream is the highest, scoring 0.31 and 0.29, respectively; the influence of the rainfall and temperature components follows closely, scoring 0.25 and 0.24; and the influence of time-dependent components θ and $\ln(1+\theta)$ is small, only scoring 0.12 and 0.14.

Considering X_1 as an example to describe the decision reduction process, the equivalence classes of conditional attributes C and D are $\text{IND}(D) = \{\{U_1, U_2, \dots, U_{37}\}, \{U_{38}, U_{39}, \dots, U_{100}\}\}$ and $\text{IND}(C) = \{\{U_1\}, \{U_2\}, \dots, \{U_{100}\}\}$, respectively, which are determined by combining collective knowledge based on $\text{IND}(C) \subseteq \text{IND}(D)$, where $\text{pos}_C(D) = U$. With $\text{IND}(C_1)$ and setting $\{U_{32}, U_{60}, U_{95}\}, \{U_7, U_{55}\}, \{U_{37}, U_{98}\}$, etc., a total of 31 sets of datasets are not subsets (as shown in Figure 7(b), the horizontal axis is the collection number that does not belong to $\text{IND}(D)$); therefore, $\text{pos}_C(D) \neq \text{pos}_{C_1}(D)$. X_1 is necessary in C relative to D , and the importance of X_1 to C relative to D can be determined based on equation (4), yielding a result of 0.31.

Based on this analysis, the factors that affect seepage pressure do not include X_6 because its importance is 0; thus,

the relative reduction of the decision sequence can be obtained.

3.2. Prediction Results and Analysis. Figure 8 reveals the prediction results of the seepage pressure, and Table 1 makes a comparison of the model prediction results. From the results, the neural network method is shown to predict the seepage pressure accurately. The determination coefficient R^2 of the RS-LSTM, RNN, and BP models is 0.97, 0.89, and 0.83, respectively, and the mean relative errors are 3.00%, 6.08%, and 9.85%, respectively, all of which are within 10%; thus, the prediction accuracy requirements are satisfactory. Comparing the prediction results of the three models, the RS-LSTM model is found to have the highest prediction accuracy, followed by the RNN model and the BP model. The calculation accuracy of the former is 2.03 times and 3.28 times that of the RNN and BP models, respectively, based on the mean relative error.

During the operation of the RS-LSTM model, the importance ranking of influencing factors is determined by RS calculation, which provides decision support for the selection of the input layer and weight setting of LSTM; eliminates the interference of redundant factors; and markedly improves the operation efficiency. In this case, prediction can be completed in 6.37 s and 278 iterations; the training error convergence process is shown in Figure 9, and the average operation efficiency is 41% and 59% higher than that of the RNN model and BP model, respectively. Additionally, both the LSTM and RNN models consider historical information hidden in the data, which can maintain correlations between the seepage pressure at different times and use them to describe the dynamic characteristics of the seepage pressure, which the BP model cannot describe.

The realization process considers historical seepage pressure values. The output of the hidden layer in the model has a connection weight from the previous hidden layer to the current layer. When the seepage pressures are output, the current water level of the upstream, rainfall, and other factors will affect the current output through the memory unit with the output layer status and will again affect the

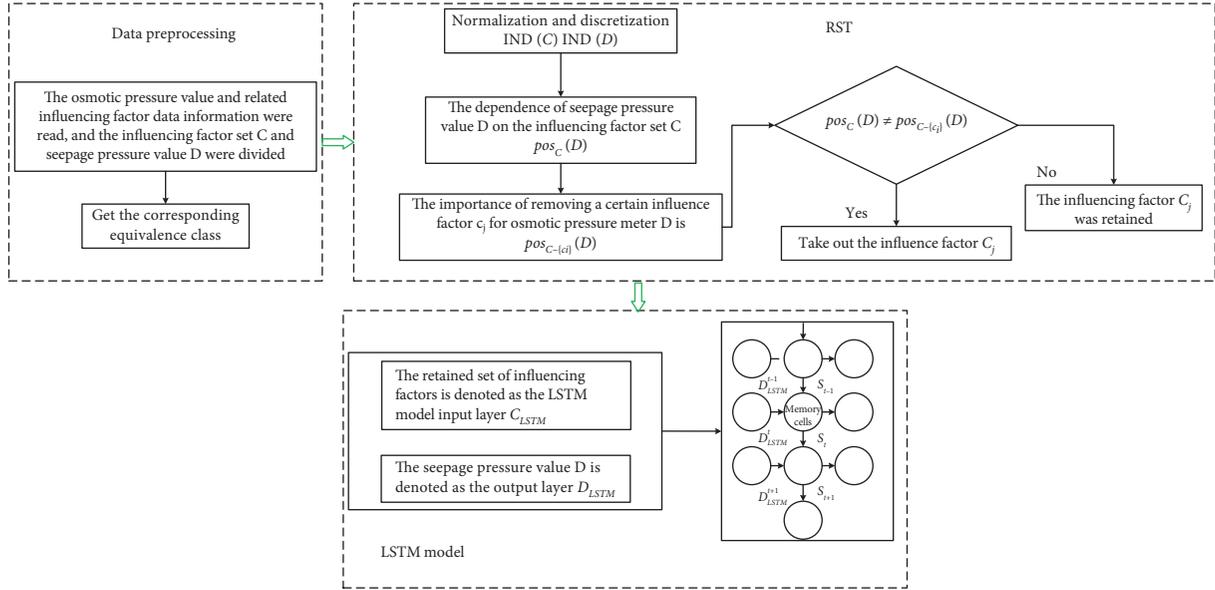


FIGURE 5: Prediction flowchart of the RS-LSTM model.

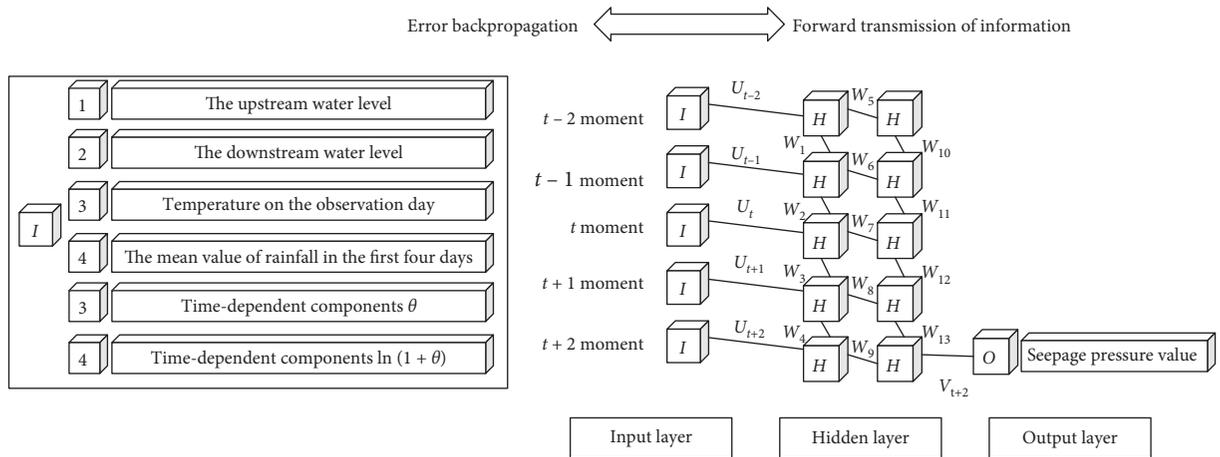


FIGURE 6: Network topology of the RS-LSTM prediction model.

output of the seepage pressure in the next time step through the memory unit with the current output layer status. This cycle then repeats.

Compared to the RNN model, the LSTM model comprises an input gate, forget gate, and output gate of the memory unit, which can select, forget, and output the current input x_t (namely, the influencing factors at the current moment), the output values of the hidden status h_{t-1} at the last moment (namely, the seepage pressure values of the hidden layer at the last moment), and the maximum reservation and output of the important hidden information, respectively, to markedly improve prediction accuracy. The specific process is described as follows. When data C_{LSTM}^t and D_{LSTM}^{t-1} pass through the forgetting gate, the information that must be discarded in the data is determined based on the setting range $[0,1]$ of the model and is determined as the information retention level f_t . After the input gate, the data in this study consist

of two parts: the sigmoid layer is primarily used to describe the renewal of the numerical range and the tanh layer is employed to create the new candidate vector-valued \tilde{C}_t . These layers are combined with the forgotten door to generate a new vector-valued S_t . For the next time step, the memory cell seepage pressure is predicted, and the output includes S_t . The output value of the current time is then determined.

The primary advantages of the RS-LSTM include the following three points. Firstly, the RS-LSTM considers historical information. By storing, writing, or reading information through the memory unit, each unit can determine which piece of information is going to be transmitted through the door switch and when it is allowed to be read, written, or cleared and can more intelligently describe the prediction and analysis of time series data. Secondly, the RS-LSTM yields high prediction accuracies and fast convergence speeds with fewer iterations and avoids local minima.

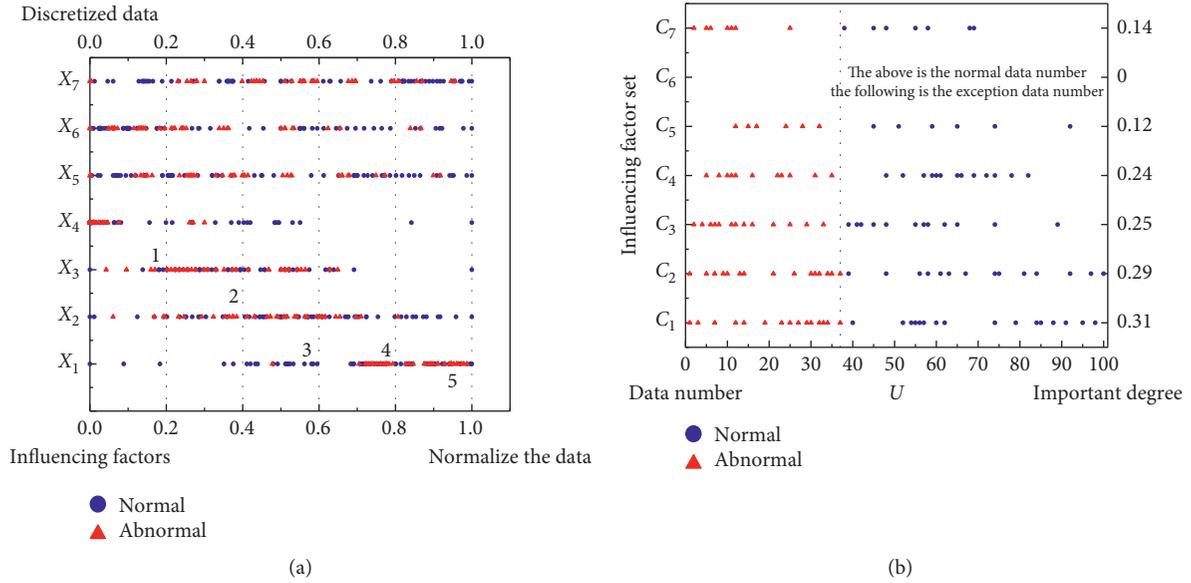


FIGURE 7: Rough set theory decision making and attribute reduction classification. (a) Decision sequence of influencing factors of the seepage pressure water level. (b) Reduction classification results of influencing factors of the seepage pressure water level.

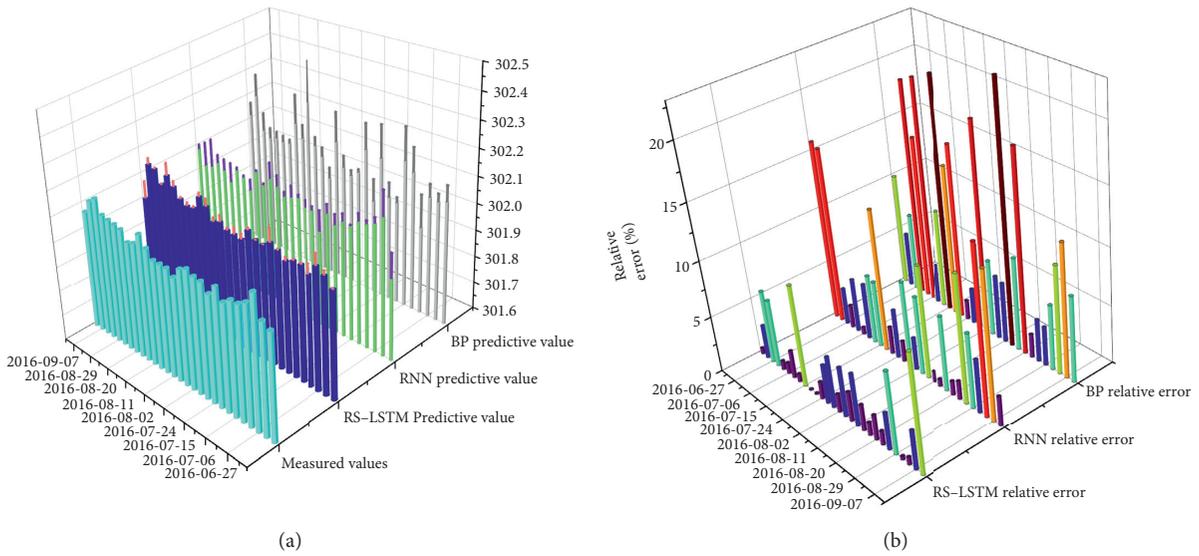


FIGURE 8: Comparison between prediction and measurement of different models. (a) Comparison diagram of the seepage pressure of each model. (b) Relative error comparison diagram.

TABLE 1: Comparison of model prediction results.

Model	Absolute error (m)				Relative error (%)				Iteration time (s)	Iteration times (times)	R^2
	Maximum value	Minimum value	Median	Mean value	Maximum value	Minimum value	Median	Mean value			
RNN model	0.109	0.004	0.031	0.040	15.84	0.59	4.85	6.08	10.85	775	0.89
BP model	0.156	0.011	0.051	0.064	22.55	1.67	7.83	9.85	15.55	1344	0.83
RS-LSTM model	0.054	0.000	0.015	0.020	9.54	0.03	2.34	3.00	6.37	278	0.97

Remarks: absolute error = |measured value of seepage pressure - predicted value of seepage pressure|; relative error = absolute error/water head \times 100%.

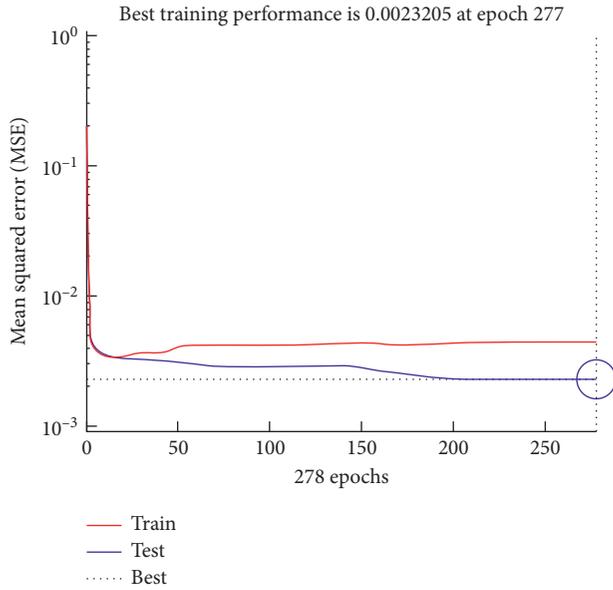


FIGURE 9: Training error convergence process diagram.

Finally, the RS-LSTM mitigates correlations between a single factor and the seepage pressure during seepage, such as the increase in the water level increasing the seepage pressure.

4. Conclusion

In this paper, the RS-LSTM model structure is constructed using engineering examples, the significances of different influencing factors on the seepage pressure are investigated using rough set theory, and the influence of different models on the prediction of the seepage pressure is described by comparing average relative errors. The specific conclusions of this study are as follows:

- (1) Written in Python, the RS theory and LSTM model are integrated to develop the RS-LSTM model in this paper. Examples elucidate that this model considers changes in measured seepage pressure during seepage and their correlations with external factors and synchronously describes the importance ranking of the influencing factors of seepage. Meanwhile, this model can eliminate the influence of redundant factors and predict dam seepage monitoring accurately and effectively. The model can provide the corresponding and the theoretical support for seepage safety during the operation of dam projects in the future.
- (2) A practical example demonstrates that the primary influencing factors of seepage pressure include the water level upstream, the water level downstream, the temperature, the average rainfall in the first four days, and the time-dependent components θ and $\ln(1 + \theta)$. The importance scores of these factors are 0.31, 0.29, 0.25, 0.24, 0.12, and 0.14, respectively, and the time-dependent component $e^{1+\theta}$ is redundant.

- (3) The prediction results demonstrate that neural networks can be used to predict seepage pressures in dams, and the determination coefficient R^2 of the RS-LSTM model proposed in this paper can reach 0.96. Compared to the traditional neural network model, the RS-LSTM model yields a higher prediction accuracy, and the accuracy of the RS-LSTM model is 2.03 and 3.28 times that of the RNN model and BP model, respectively.

As part of the future study, the proposed model could be utilized for datasets in other dams. To further enhance the performance of the proposed model, it can be attempted to integrate a variety of seepage pressure gauges to achieve multi-point and multi-factor pressure prediction.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Jiang Haifeng developed the RS-LSTM model and prepared the manuscript. Zhao Mengdie, Shoukai Chen, and Bie Yajing contributed to conceptualization, provided idea and suggestion for the study, and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

The authors would like to thank the North China University of Water Resources and Electric Power. This research was funded by the National Key R&D Plan (2018YFC0406900) and the National Natural Science Foundation of China (52009045).

References

- [1] W. Ge, H. Sun, H. Zhang et al., "Economic risk criteria for dams considering the relative level of economy and industrial economic contribution," *Science of the Total Environment*, vol. 725, Article ID 138139, 2020.
- [2] J. Li and D.-y. Sun, "Statistics of dam failures in China and analysis on failure causations," *Water Resources and Hydro-power Engineering*, vol. 40, no. 12, pp. 124–128, 2009.
- [3] S.-h. Jiang and Z.-w. Fan, "Earth-rockfill dam safety classification and risk rate assessment on flood control," *Journal of Hydraulic Engineering*, vol. 39, no. 1, pp. 35–40, 2008.
- [4] W. Ge, X. Wang, Z. Li et al., "Interval analysis of loss of life caused by dam failure," *Journal of Water Resources Planning and Management*, vol. 147, no. 1, Article ID 04020098, 2021.
- [5] S.-s. Chen, *The Failure Mechanism of Earth-Rock DAMS and the Simulation of Dam Failure Process*, China Water & Power Press, Beijing, China, 2012.

- [6] Z. Li, Y. Zhang, J. Wang et al., "Impact evaluation of geomorphic changes caused by extreme floods on inundation area considering geomorphic variations and land use types," *Science of the Total Environment*, vol. 754, Article ID 142424, 2020.
- [7] W. Ge, Y. Qin, Z. Li et al., "An innovative methodology for establishing societal life risk criteria for dams: a case study to reservoir dam failure events in China," *International Journal of Disaster Risk Reduction*, vol. 49, Article ID 101663, 2020.
- [8] M. Zhang, R. Fell, and M. Spannagle, "The statistics of embankment dam failures and accidents," *Canadian Geotechnical Journal*, vol. 37, no. 5, pp. 1000–1024, 2000.
- [9] J. Xu, W. Wei, H. Bao et al., "Failure models of a loess stacked dam: a case study in the Ansai Area (China)," *Bulletin of Engineering Geology and the Environment*, vol. 79, no. 12, 2019.
- [10] W. Guangjin, K. Jingwen, D. Chao et al., "Study on tailings dam over-topping failure model test and break mechanism under the rainfall condition," *Tehnicki Vjesnik*, vol. 24, no. 6, pp. 1897–1904, 2017.
- [11] M. Ashraf, M. Ismail, S. M. Ng, and A. Ismail, "Parametric study of horizontal drains for slope stability measure: a case study in Putrajaya, Malaysia," *KSCE Journal of Civil Engineering*, vol. 21, pp. 2162–2167, 2017.
- [12] Z.-g. Xu and J.-r. Chai, "Numerical analysis of stable seepage field in rock mass fracture network under the influence of temperature," *Journal of Xi'an Petroleum University (Natural Science Edition)*, vol. 02, pp. 169–172, 2007.
- [13] C. Xiang, L. Qingquan, Z. Zhiwei et al., "Research on a seepage monitoring model of a high core rockfill dam based on machine learning," *Sensors*, vol. 18, no. 9, p. 2749, 2018.
- [14] X. Zhang, X. Chen, and J. Li, "Improving dam seepage prediction using back-propagation neural network and genetic algorithm," *Mathematical Problems in Engineering*, vol. 2020, no. 11, 8 pages, Article ID 1404295, 2020.
- [15] Z. Shi, C. Gu, E. Zhao, and B. Xu, "A novel seepage safety monitoring model of CFRD with slab cracks using monitoring data," *Mathematical Problems in Engineering*, vol. 2020, no. 12, pp. 1–13, Article ID 1641747, 2020.
- [16] V. Rankovic, A. Novakovic, N. Grujovic et al., "Predicting piezometric water level in dams via artificial neural networks," *Neural Computing & Applications*, vol. 24, no. 5, pp. 1115–1121, 2014.
- [17] B. Yang, K. Yin, and J. Du, "A model for predicting landslide displacement based on time series and long and short-term memory neural network," *Chinese Journal of Rock Mechanics and Engineering*, vol. 37, no. 10, pp. 2334–2343, 2018.
- [18] Z. Pawlak, "Rough set theory and its applications to data analysis," *Cybernetics and Systems*, vol. 29, no. 7, pp. 661–688, 1998.
- [19] J. Zhao, J.-m. Liang, Z.-n. Dong, D.-y. Tang, Z. Liu et al., "Accelerating information entropy-based feature selection using rough set theory with classified nested equivalence classes," *Pattern Recognition*, vol. 107, 2020.
- [20] S. Hochreite and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [21] I. Good fellow, Y. Bengio, A. Courville et al., *Deep Learning*, MIT press, Cambridge, MA, USA, 2016.
- [22] S. Zhang, Y. Wang, M. Liu et al., "Data-based line trip fault prediction in power systems using LSTM networks and SVM," *IEEE Access*, vol. 6, pp. 7675–7686, 2017.
- [23] Y. Guo and X. Lai, "Water level prediction of Lake Poyang based on long short term memory neural network," *Journal of Lake Sciences*, vol. 32, no. 3, pp. 865–876, 2020.
- [24] J. Zhanga, Y. Zhub, X. Zhanga et al., "Developing a Long short-term memory (LSTM) based model for predicting water table depth in agricultural areas," *Journal of Hydrology*, vol. 561, pp. 918–929, 2018.
- [25] D. Zhang, G. Lindholm, and H. Ratnaweera, "Use long short-term memory to enhance Internet of Things for combined sewer overflow monitoring," *Journal of Hydrology*, vol. 556, pp. 409–418, 2018.