Study on Dynamic Prediction Model of Small Reservoir Safety Risk

HUANG Huaai¹, FENG Jingchun¹, SONG Zhe¹ and ZHANG Ke¹

¹Hohai University, No. 8 Focheng West Road, Jiangning District, Nanjing, Jiangsu, China

FENG Jingchun, feng.jingchun@163.com

Abstract. This paper comprehensively uses multi-source data such as small reservoir operation management information and mountain flood disaster investigation information to transform the disaster risk prediction of small reservoirs into disaster risk assessment of easy-to-observe indicators. On this basis, combined with the data of mountain flood investigation and reservoir operation management data, through the expert consultation method, the disaster risk state transition matrix and the disaster risk output matrix are studied and established to construct the Hidden Markov Model of mountain flood disaster risk prediction. Finally, combined with a reservoir in Zhejiang Province, the hidden Markov model constructed in this paper is verified. The verification results show that the model constructed in this paper is feasible and can effectively forecast the dynamic prediction of small reservoir safety risks. The research results of this paper lay the foundation for perfecting the safety risk management theory of small reservoirs and realize the dynamic prediction of small reservoir safety risks.

1. Introduction
In order to make up for the lack of monitoring information of a single small reservoir, and to improve the initiative of disaster risk prediction, comprehensive use of multi-source data such as small reservoir operation management information and mountain flood disaster investigation information. The reservoir flood risk prediction of small reservoirs should be transformed into reservoir risk assessment under the conditions of easy-to-observe indicators, to construct a data-driven reservoir risk active prediction model. Firstly, the probability of occurrence of flash flooding factors in small reservoirs is taken as a potential state variable, and the characteristics of engineering diseases and climate that are easy to observe in the disaster mode are used as external variables; Then, combined with the mountain flood survey data, reservoir operation management data and expert consultation, the disaster risk state transition matrix and the disaster risk output matrix are constructed to construct the hidden Markov model for mountain flood disaster risk prediction; Through the easy-to-observe reservoir instability and explicit variables such as geological disasters in the reservoir area, the probability of various types of disaster risk is proactively predicted, providing necessary information for downstream mountain flood warning.

2. Identification of safety risk status and safety factors of small reservoirs

2.1. Analysis of safety risk status of small reservoirs
The safety risk factors of reservoir dam breaks are divided into 5 categories and 16 small classes, namely, dam (51.5%), quality problems (38.5%), improper management (4.2%), other reasons, and unknown reasons [1].

2.2. Identification of safety risk factors for small reservoirs
There are many factors that endanger the safety of earth-rock dams, including internal, external, natural and human factors. Here, the type of earth-rock dam failure is divided into three main types, namely, the collapse of the top, the leakage and the landslide.

3. Construction of hidden Markov model for safety risk of small reservoirs
3.1. Small reservoir safety risk status and observation layer

3.1.1. Small reservoir safety risk status. According to the previous analysis, the dam break of small reservoirs can be divided into three main safety risk states: flood top, leakage and landslide. The three states of flood top, leakage and landslide may lead to dam failure. Therefore, the safety risk of the reservoir is divided into four main states: overtopping, leakage, landslide and dam failure. The transfer relationship between the safety risk status of small reservoirs is shown in the figure below.

![Figure 1 Small reservoir safety state transfer diagram](image)

3.1.2. Small reservoir safety risk status observation layer. According to the various types of safety risk obtained in the previous article and their observable indicators, the observation layer of various safety risk states of small reservoirs is obtained including observation level of overtopping risk status, observation layer of leakage safety risk status, observation layer of the landslide safety risk status and observation layer of the dam safety risk status.

3.2. HMM model construction of small reservoir safety risk
Hidden Markov Model (HMM) is described by a quintuple \([\pi, \lambda, A, B]\) as:

\[
\lambda = (N, M, A, B, \pi)
\]  

(1)  

The initial state probability matrix can be obtained from the past time security risk level. The initial safety risk status and its probability can be obtained by referring to the dam safety identification result or according to the method described above, and can be described as:
\[ \pi = \pi_i \]
\[ \pi_i = P(q_i = s_i), 1 \leq i \leq N \]
\[ \pi_i \geq 0, \sum_{i=1}^{N} \pi_i = 1 \]

(2)

Let \( \lambda = (A, B, \pi) \) be the parameter of the given HMM and \( O = o_1, o_2, \ldots, o_T \) be the sequence of observations, then the dynamic prediction of the security risk state of the small reservoir can be solved by the decoding problem in the hidden Markov model. That is, for a given model parameter and observation sequence, the most likely state sequence \( \max Q(P(O|\lambda)) \) is obtained.

3.3. HMM model algorithm for small reservoir safety risks

Knowing the observation sequence \( O= (o_1, o_2, \ldots, o_T) \), estimate the parameters of the model \( \lambda = (A, B, \pi) \), so that the observed sequence probability \( P(O|\lambda) \) is the largest under the model.

3.3.1. Supervised learning methods. If the training data contains \( S \) observation sequences of the same length and corresponding state sequences \( \{ (O_1, I_1), (O_2, I_2), \ldots, (O_S, I_S) \} \), then the maximum likelihood estimation method can be used to estimate the parameters of the hidden Markov model \[^3\]. The specific method is as follows:

- Estimation of the transition probability \( a_{ij} \). Let the time \( t \) in the sample be in the state \( I \), and the frequency at which the time \( t+1 \) shifts to the state \( j \) is \( A_{ij} \), then the estimation of the state transition probability \( a_{ij} \) is:
  \[ a_{ij} = \frac{A_{ij}}{\sum_{j=1}^{N} A_{ij}} \]  
  (3)

- Estimation of the observed probability \( b_j(k) \). Let the frequency in the sample be \( j \) and the frequency of observation \( k \) is \( B_{jk} \), then the probability \( b_j(k) \) of observing the state as \( j \) as \( k \) is:
  \[ b_{ij} = \frac{B_{jk}}{\sum_{k=1}^{M} B_{jk}}, j=1, 2, \ldots, N; k=1, 2, \ldots, M \]  
  (4)

The initial state probability \( \pi_i \) is estimated as the frequency in which the initial state is \( q_i \) in the \( S \) samples. Because supervised learning requires the use of training data, and manual labeling of training data is often costly, sometimes unsupervised learning is used.

3.3.2. Unsupervised learning algorithm. Suppose that given training data contains only \( S \) observation sequences \( \{ O_1, O_2, \ldots, O_S \} \) of length \( T \), and there is no corresponding state sequence. Our goal is to learn the parameters of Hidden Markov Model \( \hat{\lambda} = (A, B, \pi) \). We regard the observed sequence data as observation data \( O \), and the state sequence data as unobservable hidden data \( I \). It can be seen that the hidden Markov model is a probability model with hidden variables \[^4\]. See the following formula for details.

\[ P(O|\lambda) = \sum_{I} P(O|I, \lambda)P(I|\lambda) \]  
(5)

Its parameter learning can be implemented by the EM algorithm. The prediction problem is also one of the three major problems of the hidden Markov model. The model \( \hat{\lambda} = (A, B, \pi) \) and the observation sequence \( O = o_1, o_2, \ldots, o_T \) are known, and the state sequence \( I= (i_1, i_2, \ldots, i_T) \) having the largest conditional probability \( P(I|O) \) for a given observation sequence is obtained. This problem can be solved with the Viterbi algorithm. For the safety risk prediction of small reservoirs, after observing the observation sequence of the explicit variable, according to the Viterbi algorithm, the state sequence
with the highest probability of the observation sequence can be found. Furthermore, the safety risk of small reservoirs at a certain time T is predicted.

3.4. Calculation of HMM model parameters for small reservoir safety risks

According to the theoretical analysis of the small reservoir safety status and its transfer relationship, combined with the national reservoir accident registration data and Guangxi small reservoir operation related data, the transfer relationships between various types of safety status are obtained through statistics and analysis.

The theoretical analysis and case analysis are combined to construct the output relationship between various security risk states and observation layer indicators. The final results were obtained through expert investigation and discussion.

4. Case study

A reservoir in Zhejiang is located halfway up the mountain. Taking nine days as the research cycle, it is assumed that the reservoir has cracks on the eighth day and has been in a crack state since then.

According to the hidden Markov model constructed above, the Viterbi algorithm is used to solve the state with the highest probability. The first three results of the probability of the resulting state sequence are:

- \[ P_{S1} = P(\text{normal, normal, normal, normal, normal, normal, leaking, dam break}) = 5.65 \times e^{-6} \]
- \[ P_{S2} = P(\text{normal, normal, normal, normal, normal, normal, normal, normal}) = 2.37 \times e^{-7} \]
- \[ P_{S3} = P(\text{normal, normal, normal, normal, normal, normal, normal, leak}) = 8.56 \times e^{-8} \]

From the calculation results, the most likely state of the reservoir in the first seven days is the normal state. On the first day of the crack, the most likely state of the reservoir is leakage. The most likely state on the second day of the crack is the dam break.

5. Conclusion

- Comprehensively use multi-source data to transform the risk of mountain flood disasters of small reservoirs Risk assessment for disaster risk under conditions of easy-to-observe indicators.
- Based on the survey data of mountain torrents, reservoir operation management data and expert consultation, the disaster risk state transition matrix and the disaster risk output matrix are constructed to construct a hidden Markov model for mountain flood disaster risk prediction.
- Through the easy-to-observe explicit variables such as reservoir instability and reservoir geological disasters, the probability of various types of disaster risk is proactively predicted, providing necessary information for downstream mountain flood warning. The case study shows that the reservoir risk prediction method based on hidden Markov model is feasible.

References

[1] Jiang Shuhai. Reliability Analysis in Hydraulic Design of High-head Outet Works without Cavitation[J]. Journal of Hydraulic Engineering, 1990(05): 12-21.

[2] Li Hang. Statistical learning methods [M]. Tsinghua University Press, 2012:352.

[3] Cai Xun, Wang Zhaosheng, Zhu Siyu, Ding Jialin. A Real-Time Risk Assessment and Synthetical Early Warning Model of Earth and Rockfill Dams [J]. Advances in Science and Technology of Water Resources, 2018,38 (04): 57-63.

[4] Zhang Dongqing, Han Yubing, Ning xuanxi, Liu Xueni. Time Series Analysis Based on Wavelet-Domain HMM-Smoothing, Interpolation and Prediction [J]. Chinese Journal of Management Science, 2008 (02): 122-127.