

## Research Article

# A Fuzzy Clustering Logic Life Loss Risk Evaluation Model for Dam-Break Floods

Yantao Zhu,<sup>1,2,3</sup> Xinqiang Niu,<sup>4,5</sup> Chongshi Gu ,<sup>1,2,3</sup> Bo Dai ,<sup>1,2,3</sup> and Lixian Huang<sup>6</sup>

<sup>1</sup>State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing, China

<sup>2</sup>National Engineering Research Center of Water Resources Efficient Utilization and Engineering Safety, Hohai University, Nanjing, China

<sup>3</sup>College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing, China

<sup>4</sup>Changjiang Institute of Survey, Planning, Design and Research, Wuhan, China

<sup>5</sup>National Dam Safety Research Center, Wuhan, China

<sup>6</sup>Materials Science and Engineering, University of California-Los Angeles, Los Angeles, CA, USA

Correspondence should be addressed to Chongshi Gu; [csgu@hhu.edu.cn](mailto:csgu@hhu.edu.cn) and Bo Dai; [bo.dai@hhu.edu.cn](mailto:bo.dai@hhu.edu.cn)

Received 10 June 2020; Revised 20 August 2020; Accepted 15 February 2021; Published 25 February 2021

Academic Editor: Narayanan Kumarappan

Copyright © 2021 Yantao Zhu 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.

A dam is a complex and important water-retaining structure. Once the dam is broken, the flood will cause immeasurable damage to the lives and properties of the downstream people, so it is particularly important to have the dam risk management. Since the dam-break flood is a severe-consequence low-frequency event, the corresponding fatalities caused by it are difficult to estimate due to the lack of relevant data and poor data continuity. This paper analyzes the direct and indirect factors affecting the risk of life loss in dam failures and studies the characteristics, distribution rules, and membership functions of each factor. An adaptive differential evolution method is constructed through an optimization of the mutation factors and cross factors of the differential evolution method. This proposed evaluation method also combines with the fuzzy clustering iterative method that is capable of evaluating the similarity of life loss in dam accidents. The effectiveness of the proposed method is verified by 16 dam-break case studies.

## 1. Introduction

Dam-break floods can pose great threats to people's lives and properties. In 1852, the South Fork Dam in Pennsylvania was destroyed, killing 2,142 people. In 1928, the St. Francis Dam in the United States burst, killing 450 people. The collapse of the Malpasset Arch Dam in France in 1959 killed 421 people. The Vajont Dam in Italy collapsed in 1963, killing 1,994 people. In 1975, a chain of two dams in China, Banqiao and Shiwan, collapsed, killing 26,000 people. In 1979, the Machchhu Dam burst in India, killing 1,800 people [1–3]. In 2017, part of the Oroville Dam in USA was damaged, killing 34 people, and about 188,000 people were ordered to evacuate the area [4]. Consequently, dam safety risk analysis entered the vision of researchers by late 20<sup>th</sup> century. Beyond the safety of the dam structure and

construction itself, dam safety risk analysis also evaluates the impact of dam failure on downstream areas, which responds to people's increasing consciousness of self-protection and self-rescue [5–8].

A comprehensive consequence assessment is an indispensable part of dam risk analysis. As a complex assessment system, it involves multiple assessment factors such as loss of life, economic loss, and social and environmental impacts [9, 10].

Life loss due to dam failures refers to the fatalities caused by floods in the downstream areas of the dam after the dam failure. The two most direct indicators of LOL are the mortality rate of the population at risk and the death rate of the population at risk downstream the failed dam. The life loss evaluation model and outburst flood model [11, 12] are the main focuses in this field of dam failure life safety research.

Brown and Graham [13] used a mathematical statistical method to analyze the data of some dam-break life loss in the history of the United States and other countries in the world and established a simple empirical estimation formula for dam-break life loss (Brown & Graham method). Assal et al. [14] of BC Hydro, Canada, based on the empirical statistics and regression analysis of predecessors, applied the dam-break simulation technique and probability theory to estimate the life loss of dam breaks (Assaf method). Dekay from the University of Colorado and McClelland from the United States Bureau of Reclamation (USBR) [15] proposed an empirical estimation formula (Dekay & McClelland method) to describe the nonlinear relationship between the life loss and the population at risk, considering the difference in severity of dam-break floods. [16] Graham [17] suggested that the life loss of dam breaks should be estimated based on the severity of dam-break floods and built a mortality rate table of dam-break risk with regard to the release time of flood forecasts (Graham method). McClelland and Bowles [18] proposed a method to estimate the life loss based on the division of flooded areas and risk analysis in the dam-break downstream (McClelland & Bowles method). Reiter [19] in Finland proposed a simplified life loss estimation method (RESCDAM method) from Graham's method. The United States Army Corps of Engineers, USACE, USBR, and the Australian National Committee on Large Dams (ANCOLD) also supported the construction of LIFESim and other life loss calculation models [20]. Based on the above studies, Jonkman et al. [21, 22] constructed the life loss and economic loss models of low-probability high-loss accidents by analyzing the key parameters of accidents and the loss rate of accident damage areas. Fan and Jiang [23] proposed a quantitative calculation method for flood control risk rate according to the logical process of flood control project overtopping failure and introduced empirical formulas to predict fatalities. Based on the analysis of Craham and RESCDAM methods, Zhou et al. [24] analyzed the loss rates and correction coefficients under various working conditions based on the data of eight dam breaks and proposed a corresponding empirical model (Li-Zhou method). Song and He [25] studied the estimation methods of individual life loss and social life loss, analyzed the corresponding uncertainty factors, and improved the traditional estimation methods. Sun et al. [26] used Graham method to calculate the loss of life in dam-break flood simulation based on Monte Carlo and T hypercube sampling. Peng and Zhang [27, 28] built a life loss analysis model based on Bayesian networks by analyzing flood paths and other uncertain factors. Ge [29] proposed a fast evaluation method of potential consequences based on catastrophe theory.

In addition, many scholars have adopted the method of fuzzy mathematics. Wang et al. [30] established the life loss model of dam breaks based on fuzzy matter elements; Wang et al. [31] introduced indirect influencing factors of dam-break life loss and also built a prediction model of life loss. Wang et al. [32] introduced the fuzzy clustering iterative model to the dam-break life loss model. Li et al. [33] established the risk grade evaluation model of the dam-break life loss in China based on the theory of variable fuzzy sets

theory. The above three are independent of the simulation method based on fuzzy mathematics and consider the loss factors of dam failure life of the weighting matrix by using either the expert assigned AHP method or the adjacent fuzzy scale method [34, 35]. Li [36] built a dam-break comprehensive evaluation function which utilized the linear weighting method integrating the weight of life and economic, social, and environmental consequences of dam failures.

## 2. Influential Factors of Life Loss

According to the degree of influences, the influential factors of a dam-break life loss of the population at risk can be divided into direct and indirect categories as shown in Figure 1. Direct factors include warning time ( $WT$ ), flood severity ( $FS$ ), rescue capability  $RC$ , break time  $BT$ , and understanding degree of the flood severity by population at risk  $U D$ . Indirect factors include the temperature at the time of dam break, the distance between the risk population and the dam site  $L$ , and the downstream terrain conditions [37, 38], as shown in Figure 2. The influential factors are analyzed below.

### 2.1. Direct Factors

**2.1.1. Flood Severity  $FS$ .** The flood severity is a parameter that indicates the severity of damage to downstream lives and buildings. The severity of a dam-break flood is influenced by the dam type, water discharges, reservoir capacities, and downstream topographical and geomorphic conditions. In general, flood severity can be defined as the flood flow per unit section width,  $FS = (Q/W) = DV$ , where  $Q$  is the flood flow per section ( $m^3/s$ ),  $W$  is the width of the section ( $m$ ),  $D$  is the depth of the flood, and  $V$  is the flow rate of the flood. The severity of a flood and the life loss rate in a disaster have positive correlation, so the distribution is ascending half ridge type for the membership function [39], that is,

$$\mu(FS) = \begin{cases} 0, & FS \leq a_1 \\ \frac{1}{2} + \frac{1}{2} \sin \frac{\pi}{a_2 - a_1} \left( FS - \frac{a_1 + a_2}{2} \right), & a_1 < FS < a_2 \\ 1, & FS \geq a_2 \end{cases} \quad (1)$$

From literature studies, when  $FS < 0.1 (m^2/s)$ , the flood will not destroy the buildings, so  $\mu(FS) = 0$ ; when  $FS > 11.0 (m^2/s)$ , the flood will destroy all the buildings in downstream areas, so  $\mu(FS) = 1$ . Therefore,  $a_1 = 0.1$  and  $a_2 = 11.0$ .

**2.1.2. Rescue Capability  $RC$ .** The rescue capability  $RC$  considers the evacuation condition of the disaster, the rescue ability of the emergency department, the self-rescue ability of the individuals affected by the disaster, and so forth. The rescue ability is the success rate obtained by people who use

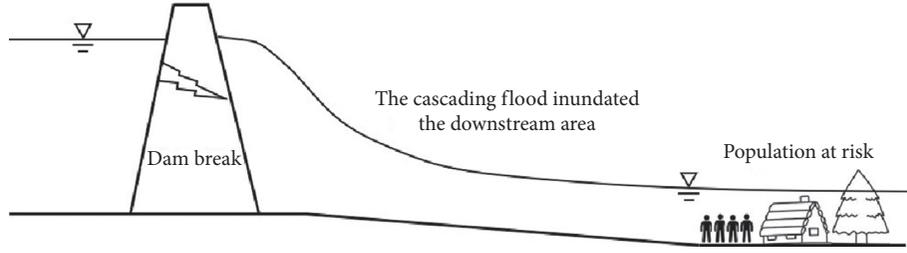


FIGURE 1: Life loss caused by dam-break flood sketch.

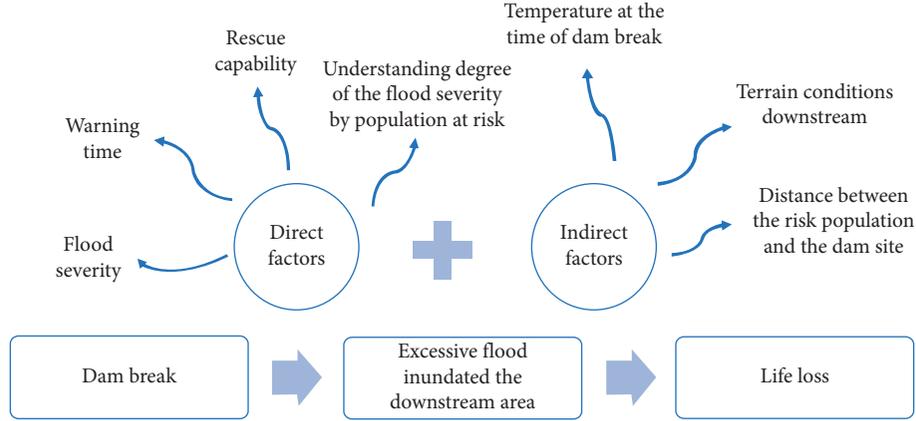


FIGURE 2: Influential factors of the life loss in a dam-break flood.

all possible means to escape. The rescue consequence is the survival rate of the population at risk after the disaster. And the membership function of the emergency rescue ability is expressed as

$$\mu(RC) = 1 - RC. \quad (2)$$

**2.1.3. Warning Time  $WT$ .** Warning time is a key factor related to the evacuation success of people in dam's downstream areas: not long enough  $WT$  may result in evacuation failure, so  $WT$  is an influential factor of the dam-break life loss.  $WT$  refers to the time interval from the point people at risk receive the flood warning to the point they evacuate. Due to the negative correlation between  $WT$  and life loss, the membership function of  $WT$  adopts the distribution of a descending half ridge:

$$\mu(WT) = \begin{cases} 1, & WT \leq a_1 \\ \frac{1}{2} - \frac{1}{2} \sin \frac{\pi}{a_2 - a_1} \left( WT - \frac{a_1 + a_2}{2} \right), & a_1 < WT < a_2 \\ 0, & WT \geq a_2 \end{cases}. \quad (3)$$

From the studies of BRUS [40], when  $WT$  is less than 0.25 h, the alarm can be considered invalid, since the people at risk cannot manage to evacuate in time; when  $WT$  is longer than 1 h, it can be considered as the state of "full warning," since it provides enough time for people to

evacuate; when  $WT$  is between 0.25 h–1 h, the state is considered "partial warning." Therefore,  $a_1$  is 0.25 and  $a_2$  is 1.

**2.1.4. Dam-Break Time  $BT$ .** The dam-break time has a great impact on the life loss. For example, in a case where  $BT$  is in the midnight, like 3 a.m., the probability that downstream residents receive the flood warning and evacuate in time is low. Consequently,  $BT$  influences the effectiveness of warning message transmission. In the same flooded area, different dam-break times will bring about fluctuations to the life loss distribution. Regarding the daily event segments,  $WT$  can be divided into working time (7:00–17:00), rest time (17:00–22:00), and sleep time (22:0–7:00). The membership function is accordingly as follows:

$$\mu(T_D) = \frac{0.4}{T_W} + \frac{0.6}{T_R} + \frac{0.9}{T_S}. \quad (4)$$

Regarding the seasons,  $WT$  can be divided into spring (March–May), summer (June–August), autumn (September–November), and winter (December–February), and its corresponding membership function is

$$\mu(T_Y) = \frac{0.5}{T_{SPR}} + \frac{0.3}{T_{SUM}} + \frac{0.6}{T_{FAL}} + \frac{0.9}{T_{WIN}}. \quad (5)$$

Regarding the fuzzy operation, the membership function of dam-break time  $BT$  is

$$\mu(BT) = \mu(T_D) \cdot \mu(T_Y). \quad (6)$$

**2.1.5. Understanding Degree of the Flood Severity by Population at Risk ( $U D$ ).** When a dam breaks, under the same set of objective influential factors ( $FS, BT, WT$ ),  $U D$  is the most important subjective factor to improve the life rescue rate in the process of rescue and evacuation.  $U D$  mainly includes three aspects of cognition: (1) people's subjective understanding of the flood disasters' severity and possible escape time after the flood occurrence, (2) the subjective escape consciousness, and (3) the understanding of escape routes and measures.  $U D$  can be divided into four levels, as shown in Table 1.

## 2.2. Indirect Factors

**2.2.1. Temperature at the Dam-Break Time.** The temperature at the dam-break time affects the life loss of the population at risk in a dam-break accident. Freezing cold conditions can hinder the evacuation process while burning hot conditions may lead to the outbreak of diseases after the evacuation. Temperatures can change with different geographical conditions, climates, seasons, and so forth.

**2.2.2. Distance between Population at Risk and the Dam Site ( $L$ ).**  $L$  refers to the distance between the area where population at risk is located and the damaged dam. The farther away from the dam, the longer time available for people to receive the warning and prepare to escape, so the life loss rate reduces with increasing  $L$ . Conversely, smaller  $L$  can result in higher life loss rate.

**2.2.3. Characteristics of the Dam.** Dam characteristics include the dam type, dam height, storage capacity, service life, dam materials, and other dam structural parameter properties.

**2.2.4. Terrain Conditions around the Downstream Areas.** When downstream areas have high terrain or some geographic/man-made barriers, these terrain conditions raise the chance of a successful evacuation and thus dramatically reduce the life loss rate.

## 3. Risk Evaluation Model of Life Loss Caused by Dam-Break Floods

To evaluate the dam-break life loss risk, various influential factors are selected to perform classification and similarity sorting to the target dam and known dam-break cases, using the method of adaptive differential evolution (ADE) optimization of fuzzy clustering iteration model: three known dams containing the most common parameters with the target dam are selected. The exponential smoothing method is used to perform risk evaluation of the target dam's life loss rate.

**3.1. Similarity Assessment of Life Loss.** In this paper, a fuzzy clustering iteration model optimized by the ADE method is used to evaluate the similarity of life loss in dam accidents in

China. The following part studies the application of the fuzzy clustering iterative model, the adaptive differential algorithm, and the fusion method for the similarity evaluation of life loss in dam accidents.

**3.1.1. Fuzzy Clustering Iterative Model.** Fuzzy clustering model (FCM) is a fuzzy clustering method based on the traditional clustering cognition and the concept of fuzzy set. In this method, each object has a membership degree about a certain class. The objects ( $u_i$ ) in the clustering object set ( $U$ ) are categorized to a certain class because of some characteristics in a certain membership degree. The core idea of fuzzy clustering iterative method is to cluster the samples in the sample space near several clustering centers by a clustering algorithm and to identify the category of each sample regarding the relative membership degree in the generated fuzzy clustering matrix. The fuzzy clustering iterative model can be used to rank the similarity of life loss samples from multiple dam accidents.

Suppose that  $n$  samples of life loss in a dam accident have a set  $X$ , where

$$X = \{x_1, x_2, \dots, x_n\}. \quad (7)$$

Consider one random sample  $x_{im}$ .

$$X = \{x_{i1}, x_{i2}, \dots, x_{im}\},$$

$$X = (x_{ij})_{n \times m} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix}. \quad (8)$$

In the formula,  $x_{ij}$  represents the eigenvalue of the  $j^{\text{th}}$  index of the  $i^{\text{th}}$  sample.  $i = 1, 2, \dots, n$  and  $j = 1, 2, \dots, m$ .  $n, m$  are the number of samples and the number of indicators, respectively. Due to different data dimensions of different categories, the eigenvalues in the matrix  $X$  need to be normalized. The normalization formula is as follows:

$$r_{ij} = \frac{(x_{ij} - x_{\min}(j))}{(x_{\max}(j) - x_{\min}(j))}. \quad (9)$$

In the formula,  $x_{\max}(j)$  and  $x_{\min}(j)$  are the maximum and minimum values of index  $j$ , respectively.  $r_{ij}$  is the normalized value of the  $j^{\text{th}}$  index of the  $i^{\text{th}}$  sample.

According to equation (9), the normalized matrix  $R$  can be obtained as follows:

$$R = (r_{ij})_{n \times m} = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1m} \\ r_{21} & r_{22} & \dots & r_{2m} \\ \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & \dots & r_{nm} \end{bmatrix}. \quad (10)$$

For  $n$  samples of life loss with  $m$  indexes, clustering analysis is conducted according to  $c$  samples, and the membership matrix of fuzzy clustering is obtained as follows:

TABLE 1: Degree of understanding.

Level	Degree	Description
I	Fuzzy	When people receive the flood warning, they cannot correctly understand the severity of this flood, fail to thoroughly understand and properly estimate the downstream flood propagation extent and its resulting damage to the downstream areas, are unable to estimate the effective escaping time before the flood arrives, fail to respond properly to escape successfully, and fail to determine possible escape routes and measures
II	Half-fuzzy	People have vague understanding of the flood severity by the time they receive the warning messages, have vague awareness of the damage to the downstream drainage systems by the flood but not accurately estimate the damage extent, and know they should take timely measures to escape and choose an escape route but cannot manage to choose the best path and the best measures
III	Half-definite	People receiving the flood alert have better understanding of the flood severity, have a more specific estimation on the downstream flood damage degree, and know they should choose escape routes and escape measures and make necessary preferential choices for this scope of drainage
IV	Definite	When people receive the risk warning, they have an accurate understanding of the flood and an accurate estimation of the flood damage extent to the downstream areas and can effectively select the most effective escape route and the escape measures to successfully evacuate

$$U = (u_{ij})_{n \times m} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1m} \\ u_{21} & u_{22} & \cdots & u_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ u_{n1} & u_{n2} & \cdots & u_{nm} \end{bmatrix}. \quad (11)$$

where  $\sum_{h=1}^c u_{hi} = 1; 0 \leq u_{hi} \leq 1; \sum_{i=1}^n u_{hi} \geq 1$ . In the formula,  $u_{hi}$  is the relative membership degree of the  $i^{\text{th}}$  life loss samples in a dam break to the  $h$  category:  $i = 1, 2, \dots, n$  and  $h = 1, 2, \dots, c$ .  $n$  is the number of samples, and  $c$  is the number of classifications.

Let the fuzzy clustering centers of  $c$  categories be

$$S = (s_{jh})_{m \times c} = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1c} \\ s_{21} & s_{22} & \cdots & s_{2c} \\ \cdots & \cdots & \cdots & \cdots \\ s_{m1} & s_{m2} & \cdots & s_{mc} \end{bmatrix}. \quad (12)$$

In the formula,  $s_{jh}$  is the relative membership degree of the  $j^{\text{th}}$  indicator for the  $h^{\text{th}}$  category. In this paper, generalized Euclidean distance is used to represent the difference between the  $i^{\text{th}}$  sample and the  $h^{\text{th}}$  category:

$$\|r_i - s_h\| = \left[ \sum_{j=1}^m (r_{ij} - s_{jh})^2 \right]^{(1/2)}. \quad (13)$$

In the formula,  $r_i = (r_{i1}, r_{i2}, \dots, r_{im})^T$  represents the eigenvalue vector of the  $i^{\text{th}}$  index and the  $m^{\text{th}}$  index;  $s_h = (s_{1h}, s_{2h}, \dots, s_{mh})^T$  represents the clustering center vector of category  $h$ . Since different evaluation indexes are of different importance in clustering, index weight is introduced. Equation (13) can be extended to

$$d(r_i - s_h) = \|\omega_j \cdot (r_i - s_h)\| = \left[ \sum_{j=1}^m [\omega_j \cdot (r_{ij} - s_{jh})]^2 \right]^{(1/2)}, \quad (14)$$

where  $\omega_j$  is the weight of the  $j^{\text{th}}$  evaluation index. From formula (14), the weighted generalized Euclidean weight distance  $D(r_i - s_h)$  is defined as follows:

$$D(r_i - s_h) = u_{hi} \cdot \omega_j \cdot (r_i - s_h) = u_{hi} \cdot \left[ \sum_{j=1}^m [\omega_j \cdot (r_{ij} - s_{jh})]^2 \right]^{(1/2)}. \quad (15)$$

The objective function formula (15) is established to solve the optimal fuzzy clustering matrix, fuzzy clustering central matrix, and weight. The objective is to minimize the sum of the squares of weighted generalized Euclidean weight distances of all life loss samples for all categories:

$$\min, \{F(\omega_j, s_{jh}, u_{hi})\} = \sum_{i=1}^n \sum_{h=1}^c [u_{hi} \cdot \omega_j \cdot (r_i - s_h)]^2, \quad (16)$$

$$\text{s.t., } \sum_{h=1}^c u_{hi} = 1, \quad 0 \leq u_{hi} \leq 1,$$

According to the Lagrangian function method, the Lagrangian function is constructed, and the fuzzy clustering iterative model can be obtained:

$$\omega_j = \left[ \frac{\sum_{i=1}^n \sum_{h=1}^c u_{hi}^2 \cdot (r_{ij} - s_{jh})^2}{\sum_{k=1}^m \sum_{i=1}^n \sum_{h=1}^c u_{hi}^2 \cdot (r_{ik} - s_{kh})^2} \right]^{-1},$$

$$u_{hi} = \frac{\sum_{j=1}^m [\omega_j \cdot (r_{ij} - s_{jh})]^2}{\sum_{k=1}^m \sum_{j=1}^m [\omega_j \cdot (r_{ij} - s_{jk})]^2}, \quad (17)$$

$$s_{jh} = \frac{\sum_{i=1}^n u_{hi}^2 \omega_j^2 r_{ij}}{\sum_{i=1}^n u_{hi}^2 \omega_j^2}.$$

The above three formulas are the iterative formulas of fuzzy clustering cycle. After iterative calculation, the membership matrix of the optimal fuzzy classes, the optimal fuzzy clustering matrix center, and the optimal weight vector can be obtained.

### 3.1.2. Adaptive Differential Evolution (ADE).

#### (1) Differential Evolution Algorithm

The differential evolution (DE) algorithm is an optimization algorithm based on the greedy idea. It

uses floating point coding instead of binary coding in the traditional genetic algorithm, which inherits the advantages of the genetic algorithm in population search and greatly improves its search efficiency. Another important characteristic of DE algorithm is its mutation process. In the DE algorithm, the mutation values are not generated randomly; instead, they are evolved from the differences among randomly selected individuals in the population in the feasible solution space. The global search ability of this algorithm is ensured by the search strategy of exchanging information among individuals. The DE algorithm, with simple structure, few control parameters, and small population size, is suitable for computer program implementation. Therefore, it was proved to be the fastest evolutionary algorithm among all the participating algorithms in the first IEEE program algorithm competition, with affirmed convergence speed and stability was also affirmed.

The DE algorithm includes random initialization, mutation operation, cross, and selection. Its optimization process is mainly realized through a population containing several individuals. This population should be initialized to determine the corresponding individuals within the population. Each individual is a feasible solution, and the population of feasible solutions is the feasible solution space of the optimization problem. The main analysis steps of the differential evolution method are as follows.

#### (1) Population Initialization

Suppose that the population size of differential evolution is  $N_p$ , and the initial population is obtained by random method, which is expressed as  $X^0 = [x_1^0, x_2^0, \dots, x_{N_p}^0]$ , where  $x_2^0 \in R^n$ . The individual in the population can be expressed as  $x_i^G = [x_{i1}^G, x_{i2}^G, \dots, x_{iD}^G]$ , where  $G$  represents the current evolutionary algebra,  $i = 1, 2, \dots, N_p$ , and  $i$  represents the population number of the individual;  $D$  represents the number of genes contained in each individual. The initial population component  $x_{ij}^0$  is firstly initialized:

$$x_{ij}^0 = L(j) + \text{rand}(0, 1) * (U(j) - L(j)), \quad (18)$$

where  $U(j) - L(j)$ , respectively, represent the maximum and minimum values of the component, so as to ensure that the initial component does not exceed the component limit.

#### (2) Mutation

The mutation operation of differential evolution method is different from that of other genetic algorithms. It is manifested in the superposition of an external random disturbance based on individual components, and this disturbance is based on two randomly selected individuals in the set of two noninferior solutions. The difference between the two random individuals is

calculated, and then the mutation vector  $v_i^G$  is obtained.

$$v_i^G = x_{r1}^G + F(x_{r2}^G - x_{r3}^G). \quad (19)$$

In the formula,  $v_i^G$  is the mutation vector,  $x_{r1}^G$  represents the individual in the current population, and  $r2, r3$  represents any integer between  $[1, N_p]$ , that is, the noncleaved set.  $G$  represents the evolutionary algebra, and  $F$  is the mutation factor between  $[0, 2]$ . The mutation factor  $F$  controlled the degree of mutation of random parameter vector. Larger mutation factor  $F$  means stronger global search ability, while smaller  $F$  indicates stronger local search capability.

#### (3) Cross

The cross operation simulates the process of propagating mutation during evolution to generate regular random disturbance in the population. In order to enrich the diversity of the population, the individual information within the population is exchanged, and the mutant is operated according to the following formula to generate experimental individuals  $u_{ij}^G$ :

$$u_{ij}^G = \begin{cases} v_{ij}^G & \text{if Rand}(j) \leq CR \text{ or } j = Rnb(i), \\ x_{ij}^G & \text{if Rand}(j) > CR \text{ or } j \neq Rnb(i). \end{cases} \quad (20)$$

In the formula,  $i$  represents the population number,  $j$  represents the individual gene,  $\text{Rand}(j) \in [0, 1]$  represents the random quantity distributed uniformly,  $Rnb(i) \in [1, D]$  represents the random integer,  $CR$  is the cross factor, and the value is  $[0, 1]$ .

The advantage of the cross operation according to formula (20) is that the variability of the experimental individuals  $u_{ij}^G$  is guaranteed. No matter how the random quantity is taken, the existence of at least one mutation gene  $v_{ij}^G$  can be guaranteed and the effectiveness of evolution can be maintained.

#### (4) Selection

The selection simulates the survival of the fittest in the evolution process: parent individuals and experimental individuals are compared and the better ones are selected to enter the next generation. This selection process can also be referred to as "greedy" selection strategy. To implement a "greedy" strategy, every individual in the population will be filtered by the fitness function. When the experiment individual  $u_i^G$  in the fitness function  $f(\cdot)$  has superior results to the current individual  $x_i^G$ , the current results of the mutation cross  $u_i^G$  will replace those of  $x_i^G$  into the offspring; otherwise, keep the current individuals into the offspring. The fitness function  $f(\cdot)$  represents the target problem that needs to be optimized. If the goal of the

algorithm optimization is to take the minimum value, the selection operation can be performed as shown in equation (21):

$$x_i^{G+1} = \begin{cases} u_i^G, & \text{if } f(u_i^G) < f(x_i^G), \\ x_i^G, & \text{else.} \end{cases} \quad (21)$$

## (2) Construction Process of the ADE Algorithm

The standard DE method has excellent global optimization ability and fast convergence. Mutations and cross operations are the main sources of offspring generation. However, although the analysis of the parameters from the DE algorithm is found, its sensitivity to the mutation factor  $F$  and cross factor  $CR$  is very high, much higher than the sensitivity to the mathematical model of the optimization. In this case, the DE algorithm efficiency, effectiveness, and robustness of the evolutionary parameters, to a great extent, are dependent on the following main research the structure of the adaptive differential evolution algorithm and the algorithm flow

### (1) Construction of the ADE Algorithm

In the standard DE method, the difference vector is related to the speed of the population convergence, and the mutation factor  $F$  controls the modulus of the difference vector. At the early stage of the population evolution, the individual distribution is relatively dispersed through random initialization. In this case, the step size of the difference vector is relatively large and the evolution speed is fast. By running the ADE algorithm, each individual in the population constantly evolves towards the direction with the higher fitness function value, and individuals in a population are relatively gathering together. In this case, the difference vector step size from the mutation process is small, the degree of population mutation is greatly reduced, and the population begins to homogenize. Consequently, the fitness function values of the offspring cannot override those of the parent generation. So, the rate of population evolution is greatly reduced, and it may be trapped into the optimal solution in a local space. In other words, when the mutation factor  $F$  is too large, the population evolution process will oscillate heavily, and the optimal solution cannot be found for a long time. When the mutation factor  $F$  is too small, the search step size will decrease, which will easily lead to the homogenization of individuals in the population. Obviously, constant mutation factor is not a positive strategy for effective optimization algorithm. Therefore, for the mutation factor  $F$ , this paper adopts an adaptive mutation factor operator that changes proportionally with the evolution algebra, which can be expressed as

$$F = F_0 \cdot 2^{\exp[(1-G_{\max})/(G_{\max}+1-G)]}. \quad (22)$$

In formula (22),  $F_0$  is the initial mutation factor,  $G_{\max}$  is the maximum evolution algebra, and  $G$  represents the current evolution algebra.

It can be seen from formula (22) that the mutation factors  $F$  are negatively correlated with the evolutionary algebra. In the early stage of operation, the global search ability is high, and the diversity of the population is guaranteed. In the later stage of operation, the local search ability is improved, and the accuracy of search is improved. At the same time, the speed of operation is accelerated.

In addition, the cross factor  $CR$  is also an essential parameter that can severely affect the optimization effect of the population, because  $CR$  plays an important role in coordinating the local search speed and the global optimization ability. When  $CR$  is larger, the contribution of the mutation vector to the test individuals is higher, which helps to improve the population diversity. When  $CR$  is smaller, more test individuals are provided by the parent individuals, which helps to narrow the search range and improve the local optimization accuracy.  $CR$  determines the degree of mutation in the cross operation, so the adaptive search strategy optimization of  $CR$  is as follows:

$$CR = CR_0 \cdot 2^{\exp[(1-G_{\max})/(1+G)]}, \quad (23)$$

where  $CR_0$  is the initial cross factor and  $G$  is evolutionary algebra. The cross factor  $CR$  is positively correlated with evolutionary algebra. That is, with the gradual accumulation of samples, the diversity gradually decreased, but the participation degree of the mutant gradually increased. So, the rate of the mutation is not going to slow down. When the maximum evolution rate is reached, the convergence rate also reaches its peak.

### (2) Adapt the Differential Evolution Algorithm Steps

*Step 1. Initialization:* initialization was carried out according to equation (18). Spatial dimension  $D$ , population size  $N_p$ , initial cross probability  $CR_0$ , initial mutation factor  $F_0$ , and initial evolutionary algebra  $G$  are set, and random initialization is carried out within the value range.

*Step 2. Generate the operation factor:* equation (22) is applied to calculate the parameters.

*Step 3. Evolution operation:* use equations (19), (20), and (21) to carry out the mutation operation, cross operation, and selection, respectively, to complete the evolution of the present generation.

*Step 4. Individual evaluation:* the fitness is calculated for each individual in the population of

the current generation. The best fitness means the best value of this generation.

*Step 5.* Terminate the judgment: judge the optimal value of this generation. If the value of fitness function meets the accuracy requirement or the number of iterations reaches the upper limit, the operation will terminate. Otherwise, the evolution times will increase by one and continue to iterate from Step 2.

**3.1.3. Similarity Evaluation Process.** In the following, the ADE method is introduced to the traditional fuzzy clustering iterative model (FCM) to analyze, identify, and rank multiple life losses caused by dam-break flood samples, so ADE can optimize the clustering efficiency of the traditional FCM model and construct an efficient and concise ADE-FCM dam-failure life loss similarity evaluation method.

(1) Search Variable Encoding

To fuse the ADE algorithm with the fuzzy clustering algorithm, the primary task is to determine the encoding mode of the optimization object and the search variables. As the ADE is actually an argument evolutionary algorithm, the search variables must be coded with real numbers. The purpose of fuzzy clustering iteration is to find the optimal clustering center, the fuzzy membership matrix, and the optimal weight matrix. Therefore, the selected optimization object cannot only optimize but also classify the clustering center. Therefore, the clustering center  $(s_{jh})_{m \times c}$  is selected as the optimized object and coded as a population individual, which means that there are  $m \times c$  populations of individuals (optimization variables) that need to be coded. The ADE vector can be expressed as

$$s_j = [s_{j,1}, s_{j,2}, \dots, s_{j,c(m-1)+1}, s_{j,c(m-1)+2}, \dots, s_{j,cm}]. \quad (24)$$

In other words,  $m \times c$  variables are divided into  $c$  groups; each group represents a clustering center and has  $m$  indicators. This method numbers the vectors and starts the calculation.

(2) Determine the Fitness Function

The objective function of FCM iteration is a generalized Euclidean function, so the following equation is used as the fitness function:

$$f = \{F(\omega_j, s_{jh}, u_{hi})\} = \sum_{i=1}^n \sum_{h=1}^c [u_{hi} \cdot \omega_j \cdot (r_i - s_h)]^2. \quad (25)$$

The goal is to minimize the fitness function value (Euclidean distance), which is the optimal partition of the data set.

The process of risk evaluation of life loss caused by a dam-break flood with ADE-FCM method is as follows:

*Step 1.* Initialize each parameter of the FCM: to initialize the cluster center  $s_{jh}$ , set the stop operation precision  $\epsilon$ , set a large constant  $\theta$ , and make the target function  $\{F(\omega_j, s_{jh}, u_{hi})\} = \theta$

*Step 2.* Initialize ADE parameters: let the initial evolution algebra  $G = 1$ , and set the population size  $N_p$ , the maximum evolution algebra  $G_{\max}$ , the initial cross factor  $CR_0$ , and the initial mutation factor  $F_0$

*Step 3.* Optimize the clustering center: use the adaptive differential optimization method to carry out mutation, cross, and selection operations to select the optimal clustering center of this generation

*Step 4.* Calculate the fuzzy membership matrix of this generation  $u_{hi}^G$

*Step 5.* Calculate the weight of this generation  $\omega_j^G$

*Step 6.* Calculate the fitness function of this generation  $f^G$

*Step 7.* Compare the fitness function value of the current generation with the fitness function  $f^G$  of the previous generation  $f^{G-1}$ . If  $|f^G - f^{G-1}| < \epsilon$  or  $G > G_{\max}$ , the operation will stop; otherwise, go to Step 3 to continue the iterative operation

**3.2. Estimation of Life Loss Rate.** After the similarity evaluation of the life loss, the exponential smoothing method was used to estimate the estimated samples.

The principle of the exponential smoothing method is to select the known samples that are similar to the samples to be estimated as well as giving more weight to those with higher similarity and giving less weight to those with lower similarity. The calculation results reflect the comprehensive influence of similar samples, and the results are more in line with the actual situation.

The formula for the exponential smoothing method to fit the sample to be evaluated is as follows:

$$N = \alpha N_1 + \alpha N_2 (1 - \alpha) + \alpha N_3 (1 - \alpha)^2 + (1 - \alpha)^3 \frac{N_1 + N_2 + N_3}{3}. \quad (26)$$

In the formula,  $N$  is the estimated value of the sample to be estimated,  $N_1, N_2, N_3$  are the three known data samples that are most similar to the sample to be estimated,  $\alpha$  is the smoothing coefficient,  $\alpha \in [0, 1]$ , and  $\alpha$  is related to the degree of stationarity of the known value. The data of the three dams which are most similar to the risk of life loss due

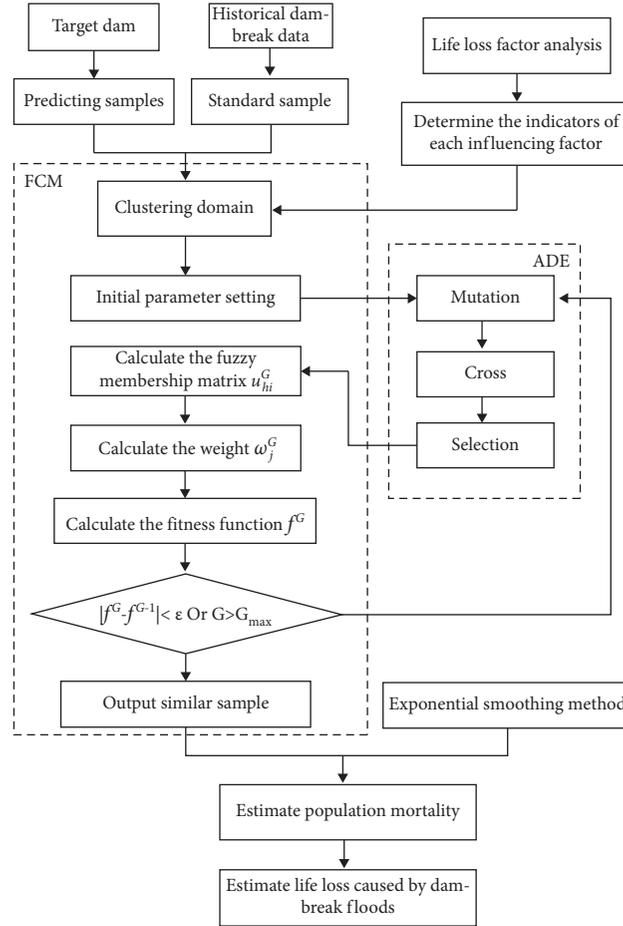


FIGURE 3: Calculation flow chart.

to dam-break flood to be estimated are taken into equation (26) to estimate the risk of life loss due to dam-break flood. The calculation process is shown in Figure 3.

#### 4. Case Study

Aiming at investigating the situation of dam breaches in China [41], the case studies have selected eight dams' (a total of 16 sets of data) life loss history data as the research objects of dam breaks; see Table 2. The locations of the eight dams are shown in Figure 4. Six influential factors  $BT, FS, RC, WT, U D,$  and  $L$  of the population mortality risk  $f$  are set as the characteristic values. A set of samples are selected to be estimated using the method of mutual interaction validation, while the rest of the samples are set as known references. The calculation results from the estimation are compared with those from the field data. Then, the feasibility of the estimation method is verified.

4.1. Membership Degree Calculation and Normalization Processing. According to the analysis of life loss factors in Section 2, the membership degree of each index is calculated, and the membership degree matrix is shown in the following equation:

$$x_{ij} = \left\{ \begin{array}{cccccc} & J_1 & J_2 & J_3 & J_4 & J_5 & J_6 \\ I_1 & 0.27 & 1.00 & 0.70 & 0.00 & II & 4 \\ I_2 & 0.27 & 1.00 & 0.70 & 0.70 & II & 11 \\ I_3 & 0.27 & 0.48 & 0.70 & 1.00 & II & 22.5 \\ I_4 & 0.20 & 0.18 & 0.30 & 0.87 & III & 4.5 \\ I_5 & 0.20 & 0.66 & 0.30 & 0.87 & I & 10.5 \\ I_6 & 0.27 & 0.20 & 0.60 & 1.00 & I & 0.75 \\ I_7 & 0.27 & 0.24 & 0.30 & 1.00 & I & 1.5 \\ I_8 & 0.54 & 1.00 & 0.95 & 1.00 & I & 0.25 \\ I_9 & 0.27 & 1.00 & 0.70 & 0.80 & I & 0.4 \\ I_{10} & 0.27 & 1.10 & 0.70 & 1.00 & IV & 9 \\ I_{11} & 0.27 & 0.64 & 0.70 & 1.00 & I & 28.5 \\ I_{12} & 0.27 & 0.24 & 0.70 & 1.00 & I & 52.5 \\ I_{13} & 0.27 & 0.72 & 0.80 & 1.00 & III & 2.5 \\ I_{14} & 0.27 & 0.12 & 0.80 & 1.00 & II & 7.5 \\ I_{15} & 0.27 & 0.15 & 0.80 & 1.00 & II & 12.5 \\ I_{16} & 0.18 & 0.22 & 0.95 & 1.00 & I & 12.5 \end{array} \right\}. \quad (27)$$

TABLE 2: Sample data of life loss caused by dam-break flood.

No.	Name of dam	Break date	$BT$	$FS$ (m <sup>2</sup> /s)	$RC$ (%)	$WT$ (h)	$U D$	$L$ (km)	$f$ (%)
1				19.7	30	2	II	1-7	18.9
2	Liujiatai	1963/8/8	3:55	12.4	30	0.6	II	7-15	10.4
3				5.4	30	0	II	15-30	0.5
4				3.3	70	0.25	III	3-6	0.8
5	Hengjiang	1970/9/15	8:00	6.7	70	0.25	I	6-15	1.9
6				3.41	40	0	I	0.5-1	4.4
7	Dongkoumiao	1971/6/2	5:50	3.72	70	0	I	1-2	2.7
8				20.22	5	0	I	0-0.5	49.9
9	Lijiazui	1973/4/29	23:30	16.6	30	0.4	I	0-0.8	27
10				9.75	30	0	IV	6-12	12.7
11	Banqiao	1975/8/8	1:00	6.5	30	0	I	12-45	8.8
12				3.7	30	0	I	45-60	1.34
13				7.1	20	0	III	0-5	2.1
14	Shimantan	1975/8/8	1:00	2.86	20	0	II	5-10	0.81
15				3.08	20	0	II	10-15	2.07
16	Gouhou	1993/8/27	22:00	3.6	5	0	I	12-13	10.5

In the above equation,  $I_1, I_2 \dots I_{16}$  are 16 samples, one of which was set as the sample to be evaluated, and the rest were 15 known samples;  $J_1, J_2 \dots J_6$  were the 6 factors affecting the severity of life loss caused by a dam-break flood, which were the time of break  $BT$ , flood severity  $FS$ , rescue capability  $RC$ , warning time  $WT$ , understanding degree of the flood severity by population at risk  $U D$ , and the distance between the risk population and the dam site  $L$ , respectively. Among these six factors, only  $U D$  and  $L$  did not experience membership degree calculation, and others had membership degree between 0 and 1. Therefore, values of  $U D$  and  $L$  needed to be normalized. Matrix (30) was normalized according to equation (9), and its normalized matrix was

$$r_{ij} = \begin{pmatrix} & J_1 & J_2 & J_3 & J_4 & J_5 & J_6 \\ I_1 & 0.27 & 1.00 & 0.70 & 0.00 & 0.67 & 0.08 \\ I_2 & 0.27 & 1.00 & 0.70 & 0.70 & 0.67 & 0.21 \\ I_3 & 0.27 & 0.48 & 0.70 & 1.00 & 0.67 & 0.43 \\ I_4 & 0.20 & 0.18 & 0.30 & 0.87 & 0.33 & 0.09 \\ I_5 & 0.20 & 0.66 & 0.30 & 0.87 & 1.00 & 0.20 \\ I_6 & 0.27 & 0.20 & 0.60 & 1.00 & 1.00 & 0.01 \\ I_7 & 0.27 & 0.24 & 0.30 & 1.00 & 1.00 & 0.03 \\ I_8 & 0.54 & 1.00 & 0.95 & 1.00 & 1.00 & 0.00 \\ I_9 & 0.27 & 1.00 & 0.70 & 0.80 & 1.00 & 0.01 \\ I_{10} & 0.27 & 1.10 & 0.70 & 1.00 & 0.00 & 0.17 \\ I_{11} & 0.27 & 0.64 & 0.70 & 1.00 & 1.00 & 0.55 \\ I_{12} & 0.27 & 0.24 & 0.70 & 1.00 & 1.00 & 1.01 \\ I_{13} & 0.27 & 0.72 & 0.80 & 1.00 & 0.33 & 0.05 \\ I_{14} & 0.27 & 0.12 & 0.80 & 1.00 & 0.67 & 0.14 \\ I_{15} & 0.27 & 0.15 & 0.80 & 1.00 & 0.67 & 0.24 \\ I_{16} & 0.18 & 0.22 & 0.95 & 1.00 & 1.00 & 0.24 \end{pmatrix}. \quad (28)$$

4.2. *Parameter Settings.* Firstly, the parameters of the fuzzy clustering algorithm were set. The number of samples was 16, the index dimension was  $j = 6$ , and the number of categories to be classified was  $c = 3$ . Due to the low degree of data dimensions and functionality in the clustering problem, the parameters of the ADE part were set as follows: population size ( $N_p$ ) was 10, maximum evolutionary algebra ( $G_{\max}$ ) was 10000, initial cross factor ( $CR_0$ ) was 0.45, and initial mutation factor ( $F_0$ ) was 0.45.

4.3. *Similarity Evaluation.* After setting the parameters, the optimal fuzzy clustering center ( $s_{jh}$ ), the optimal fuzzy membership matrix ( $u_{hi}$ ), and the membership degree ( $\omega_j$ ) can be obtained. In this paper, the MATLAB program was used to achieve the following calculation results:

$$s_{jh} = \begin{pmatrix} & i = 1 & i = 2 & i = 3 \\ j = 1 & 0.27 & 0.26 & 0.30 \\ j = 2 & 0.53 & 0.24 & 0.94 \\ j = 3 & 0.65 & 0.68 & 0.72 \\ j = 4 & 0.95 & 0.97 & 0.74 \\ j = 5 & 0.82 & 0.77 & 0.69 \\ j = 6 & 0.40 & 0.17 & 0.13 \end{pmatrix},$$

$$\omega_j = \begin{pmatrix} \omega & j = 1 & j = 2 & j = 3 & j = 4 & j = 5 & j = 6 \\ h = 1 & 0.25 & 0.46 & 0.57 & 0.79 & 0.19 & 0.28 \\ h = 2 & 0.17 & 0.25 & 0.99 & 0.63 & 0.46 & 0.16 \\ h = 3 & 0.39 & 0.68 & 0.75 & 0.50 & 0.37 & 0.07 \end{pmatrix}. \quad (29)$$

According to the feature consensus of variable fuzzy set in variable fuzzy set theory, the classification of the sample was calculated as follows:



FIGURE 4: Distribution of dams in China.

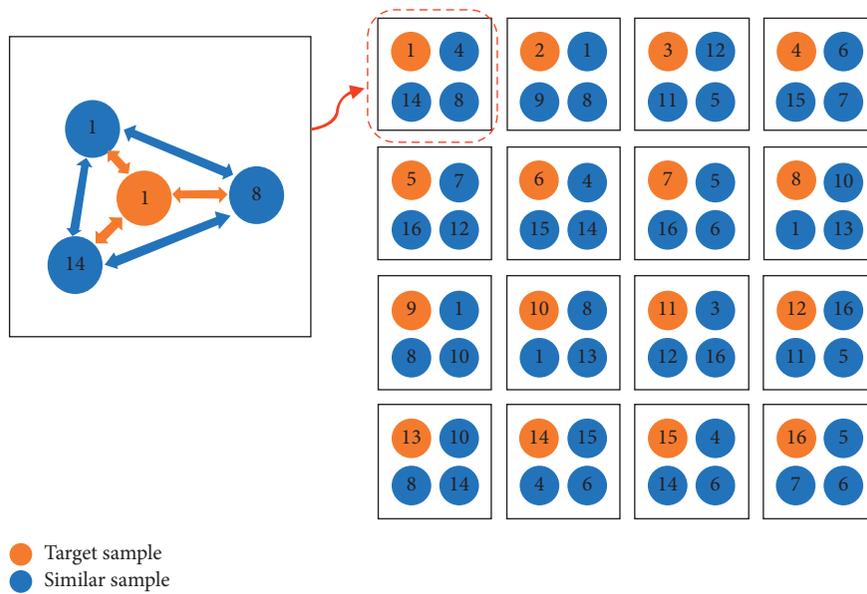


FIGURE 5: Results of similarity evaluation.

TABLE 3: Results estimated by different models.

Sample	Actual value	ADE-FCM	FME	Error of ADE-FCM	Error of FME
1	0.189	0.1816	0.428	0.0074	0.239
2	0.104	0.0891	0.152	0.0149	0.048
3	0.005	0.0233	0.09	0.0183	0.085
4	0.008	0.0238	0.01	0.0158	0.002
5	0.019	0.0840	0.02	0.0650	0.001
6	0.044	0.0139	0.113	0.0301	0.069
7	0.027	0.0617	0.06	0.0347	0.033
8	0.499	0.1359	1.497	0.3631	0.998
9	0.270	0.2665	0.668	0.0035	0.398
10	0.127	0.3083	0.269	0.0187	0.142
11	0.088	0.0399	0.175	0.0481	0.087
12	0.0134	0.0576	0.0162	0.0442	0.0028
13	0.021	0.2477	0.028	0.0038	0.007
14	0.0081	0.0262	0.0119	0.0181	0.0038
15	0.0207	0.0204	0.0431	0.0003	0.0224
16	0.105	0.0325	0.1056	0.0725	0.0006

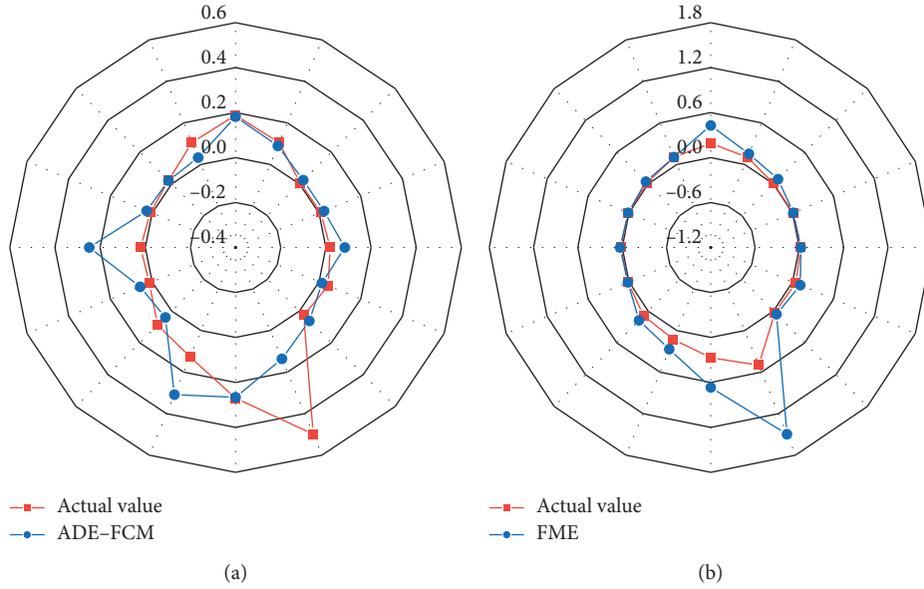


FIGURE 6: Comparison of estimation results.

$$H_i = \sum_{h=1}^c u_{hi} h \begin{cases} 1 \leq H_i \leq 1.5, & \text{belong to level 1,} \\ \vdots & \vdots \\ c-1 < H_i \leq c-0.5, & \text{belong to level } c-1, \\ c-0.5 < H_i \leq c, & \text{belong to level } c. \end{cases} \quad (30)$$

where  $i = 1, 2, \dots, n$ ,  $h = 1, 2, \dots, c$ , and  $c = 3$ .

$$u_{hi} = \left\{ \begin{array}{cccccc} & i=1 & i=2 & i=3 & \dots & i=14 & i=15 & i=16 \\ h=1 & 0.25 & 0.04 & 0.77 & \dots & 0.12 & 0.15 & 0.31 \\ h=2 & 0.20 & 0.02 & 0.17 & \dots & 0.83 & 0.81 & 0.58 \\ h=3 & 0.56 & 0.95 & 0.06 & \dots & 0.05 & 0.04 & 0.10 \\ H & 2 & 3 & 1 & \dots & 2 & 2 & 2 \end{array} \right\}. \quad (31)$$

According to the fuzzy membership matrix, the similarity of each sample to be evaluated was sorted, and Figure 5 was obtained.

**4.4. Estimation of Life Loss Rate.** According to equation (26), 16 known samples were estimated according to the smoothing index method, and the difference values between estimated results and the actual field data were calculated. The calculation results of this paper were evaluated by comparison with the Fuzzy Matter-Element Analysis [11], as shown in Table 3.

From the comparisons shown in Figures 6 and 7, the calculation accuracy of the method presented in this paper is obviously better. The relative errors between the estimated values and the actual values in this paper are between  $[0, 0.37]$ , and most of the errors are between  $[0,$

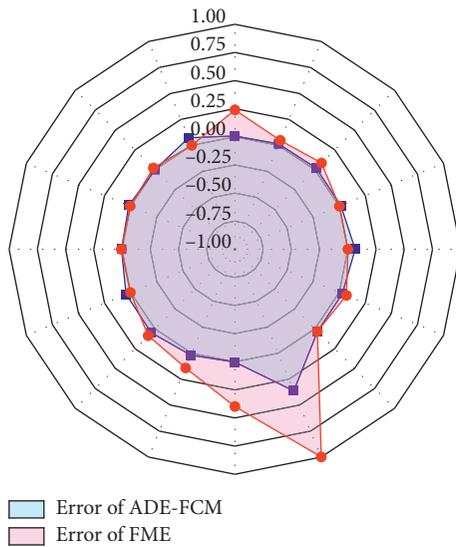


FIGURE 7: Comparison of estimation errors.

0.01]. Therefore, the life loss estimation method after similarity evaluation by ADE-FCM method has certain feasibility and accuracy in estimating the life loss rate caused by dam-break floods. In theory, it can result from the following: (1) Through ADE method to optimize the fuzzy clustering center, a better fuzzy clustering center is obtained, so the similarity evaluation is more scientific and reasonable, and the influence of the known samples and the sample to be estimated can be more accurately grasped. (2) Through ADE-FCM model, a weight based on objective facts is obtained, and the weight analysis shows that the weight of  $j = 2, 3, 4$  (flood severity, rescue capability, and warning time) is larger, which is consistent with the weight results calculated by adjacent fuzzy scale method in literature [28]. So, the life loss estimation method proposed in this paper is scientific and feasible.

## 5. Summary and Conclusion

This study analyzes the influential factors of dam-break flood life loss through the ADE-FCM method. The similarity of eight dam-break cases and their life loss data are analyzed, the life loss at the target dam is estimated with a comprehensive exponential smoothing method, and the results are compared with the actual values. The main research contents are as follows:

- (1) The direct and indirect factors affecting the risk of life loss due to dam failure are analyzed, and the characteristics, distribution rules, and membership functions of the factors are studied. Based on the research of fuzzy clustering iterative method, the FCM method is introduced into the similarity evaluation of life loss caused by dam-break floods.
- (2) An adaptive differential evolution method was constructed by improving the mutation factor and cross factors of the differential evolution method.

Based on the fusion of ADE and FCM, a fuzzy clustering iterative method based on adaptive differential evolution method (ADE-FCM) was established, and a similarity evaluation method for life loss caused by dam-break floods was established.

- (3) The principle of exponential smoothing method was studied. On the basis of the similarity evaluation, the method was used to estimate the target sample and compare the results with the actual data. Based on case studies, ADE-FCM was used to rank the similarity of life loss in dam accidents, and the risk population mortality of the target was estimated based on the exponential smoothing method.

## Data Availability

The raw data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant nos. 51739003 and 51909173), Free Exploration Project of Hohai University (Grant no. B200201058), Basic Research Project Funded of National Key Laboratory (Grant no. 20195025912), and Open Foundation of Changjiang Survey, Planning, Design and Research Co., Ltd. (Grant no. CX2019K01).

## References

- [1] X. Zhang, "Collection of typical cases of dam failures and accidents at hydropower stations," *Dam and Safety*, vol. 1, no. 1, pp. 13–16, 2015.
- [2] J. Zhang, Z. Yang, and J. Jiang, "An analysis on laws of reservoir dam defects and breaches in China," *Scientia Sinica Technologica*, vol. 47, no. 12, pp. 1313–1320, 2017.
- [3] J. Sun, "Achievement of reservoir dam safety management in China and challenges," *China Water Resources*, vol. 20, pp. 1–6, 2018.
- [4] A. Koskinas, A. Tegos, P. Tsira et al., "Insights into the Oroville dam 2017 spillway incident," *Geosciences*, vol. 9, no. 37, 2019.
- [5] J. Zhang and Y. Xiang, "Analysis on the impact of climate change on the water conservancy project safety," *Scientia Sinica Technologica*, vol. 48, no. 10, pp. 1031–1039, 2018.
- [6] Z. Li, W. Ge, J. Wang et al., "Strategic consideration of dam safety management and risk management in China," *Advances in Water Science*, vol. 26, no. 4, pp. 589–595, 2015.
- [7] L. Yan, *Study on Operation Safety Risk Analysis Method of Dam*, Tianjin University, Tianjin, China, 2011.
- [8] B. Richard and R. E. Waite, "Evaluation for a series of utah power and light hydropower dams, including risk assessment: the owner perspective," in *Proceedings of the 6th ASDSO Annual Conference*, pp. 1–5, Association of State Dam Safety Official, Albuquerque, New Mexico, October 1989.

- [9] ANCOLD, *Guidelines on Assessment of the Consequence of Dam Failure*, Tatura: Australian National Committee on Large Dams, Hobart, Australia, 2000.
- [10] M. L. DeKay and G. H. McClelland, "Predicting loss of life in cases of dam failure and flash flood," *Risk Analysis*, vol. 13, no. 2, pp. 193–205, 1993.
- [11] P. Dimitriadis, A. Tegos, A. Oikonomou et al., "Comparative evaluation of 1D and quasi-2D hydraulic models based on benchmark and real-world applications for uncertainty assessment in flood mapping," *Journal of Hydrology*, vol. 534, pp. 478–492, 2016.
- [12] A. D. F. Tschiedel and R. C. D. De Paiva, "Uncertainty assessment in hydrodynamic modeling of floods generated by dam break," *RBRH*, vol. 23, no. 30, pp. 1–17, 2018.
- [13] C. A. Brown and W. J. Graham, "Assessing the threat to life from dam failure," *Journal of the American Water Resources Association*, vol. 24, no. 6, pp. 1303–1309, 1988.
- [14] H. Assaf, D. N. D. Hartford, and J. D. Cattanaach, "Estimating dam breach flood survival probabilities," *Ancold Bulletin*, vol. 107, pp. 23–42, 1997.
- [15] M. L. Dekay and G. H. McClelland, "Predicting loss of life in cases of dam failure and flash flood," *Risk Analysis*, vol. 13, no. 2, pp. 193–205, 1993.
- [16] J. Sheng, D. Li, Q. Cai, and X. Peng, "Research progress and its practice of key techniques for dam risk assessment and management," *Scientia Sinica Technologica*, vol. 48, no. 10, pp. 1057–1067, 2018.
- [17] W. J. Graham, "A procedure for estimating loss of life caused by dam failure," *Sedimentation & River Hydraulics*, vol. 6, no. 5, pp. 1–43, 1999.
- [18] M. McClelland and D. S. Bowles, "Towards improved life loss estimation methods: lessons from case histories," in *Proceedings of the Rescdam Seminar*, Seinäjoki, Finland, October 2000.
- [19] P. Reiter, *Loss of Life Caused by Dam Failure: the RESCDAM LOL Method and its Application to Kyrkosjarvi Dam in Seinäjoki*, PR Water Consulting Ltd, Helsinki, Finland, 2001.
- [20] D. S. Bowles and M. Aboelata, "Evacuation and life-loss estimation model for natural and dam break floods," *Nato Science*, vol. 78, pp. 363–383, 2006.
- [21] S. N. Jonkman, J. K. Vrijling, and A. C. W. M. Vrouwenvelder, "Methods for the estimation of loss of life due to floods: a literature review and a proposal for a new method," *Natural Hazards*, vol. 46, no. 3, pp. 353–389, 2008.
- [22] S. N. Jonkman, A. Lentz, and J. K. Vrijling, "A general approach for the estimation of loss of life due to natural and technological disasters," *Reliability Engineering & System Safety*, vol. 95, no. 11, pp. 1123–1133, 2010.
- [23] K. Zhou, L. Li, and J. Sheng, "Evaluation model of loss of life due to dam breach in China," *Journal of Safety and Environment*, vol. 7, no. 3, pp. 145–149, 2007.
- [24] Z. Fan and S. Jiang, "Application of tolerance risk analysis method in decision-making of flood prevention," *Journal of Hydraulic Engineering*, vol. 36, no. 5, pp. 618–623, 2005.
- [25] J. Song and X. He, "Discussion on analysis method for risk of life loss caused by dam failure in China," *Journal of Hohai University (Natural Sciences)*, vol. 36, no. 5, pp. 628–633, 2008.
- [26] Y. Sun, D. Zhong, M. Li, and Y. Li, "Theory and application of loss of life risk analysis for dam break," *Transactions of Tianjin University*, vol. 16, no. 5, pp. 383–387, 2010.
- [27] M. Peng and L. M. Zhang, "Analysis of human risks due to dam-break floods-part 1: a new model based on Bayesian networks," *Natural Hazards*, vol. 64, no. 1, pp. 903–933, 2012.
- [28] M. Peng and L. M. Zhang, "Analysis of human risks due to dam break floods-part 2: application to Tangjiashan landslide dam failure," *Natural Hazards*, vol. 64, no. 2, pp. 1899–1923, 2012.
- [29] W. Ge, Y. Jiao, H. Sun et al., "A method for fast evaluation of potential consequences of dam breach," *Water*, vol. 11, no. 11, p. 2224, 2019.
- [30] Z. Wang, C. Gu, and H. Liu, "Life loss estimation based on fuzzy matter-element and exponential smoothing method," *Journal of Yangtze River Academy of Sciences*, vol. 26, pp. 25–28, 2009.
- [31] S. Wang, T. Bao, and L. Chen, "Fuzzy prediction model of life loss in dam break based on fuzzy matter-element analysis," *Hydropower Energy Science*, vol. 29, pp. 46–49, 2011.
- [32] J. Wang and Y. Yuan, "Prediction of dam break life loss based on variable fuzzy clustering iterative model," *Hydropower Energy Science*, vol. 215, no. 6, pp. 82–85, 2012.
- [33] W. Li, Z. Li, W. Ge, and S. Wu, "Risk evaluation model of life loss caused by dam-break flood and its application," *Water*, vol. 11, no. 7, p. 1359, 2019.
- [34] S. Chen, J. Yuan, and Y. Li, "Fuzzy clustering generation selection model based on variable fuzzy sets and its application," *Journal of Dalian University of Technology*, vol. 48, pp. 881–886, 2008.
- [35] S. Chen, "Variable fuzzy set theory--double discussion on mathematical and logic errors in extenics," *Journal of Dalian University of Technology*, vol. 47, pp. 618–624, 2007.
- [36] L. Li, R. Wang, and J. Sheng, "Study on evaluation model of dam break consequence severity," *Journal of Safety and Environment*, vol. 6, pp. 1–4, 2006.
- [37] D. Koutsoyiannis, P. Dimitriadis, F. Lombardo, and S. Stevens, "From fractals to stochastics: seeking theoretical consistency in analysis of geophysical data," in *Advances in Nonlinear Geosciences*, A. A. Tsonis, Ed., Springer, Berlin, Germany, pp. 237–278, 2018.
- [38] T. Wang, "Study on the risk analysis of dam safety emergency management region and the methods of losses assessment," Master thesis, Changjiang River Scientific Research Institute, Wuhan, China, 2010.
- [39] N. D. Singpurwalla and J. M. Booker, "Membership functions and probability measures of fuzzy sets," *Journal of the American Statistical Association*, vol. 99, no. 467, pp. 867–877, 2004.
- [40] T. L. Wahl, "Prediction of Embankment dam breach parameters: a literature review and needs assessment," DSO-98-004. Dam Safety Research Report, USBR, Washington, DC, USA, 1998.
- [41] K. Zhou and L. Li, "Investigation of broken dams and the law of loss of life in China," *Dams and Safety*, vol. 5, pp. 14–18, 2006.