

## Analysis of Observation Data of Earth-Rockfill Dam Based on Cloud Probability Distribution Density Algorithm

Han Liwei<sup>1,\*</sup>, Yu Hongtao<sup>1</sup>, Zhang Hongyang<sup>1</sup>, Lee Kunghon<sup>2</sup> and Xu Cundong<sup>1</sup>

<sup>1</sup>School of Water Conservancy, North China University of Water Resources and Electric Power, 450045, Zhengzhou, Henan, China

<sup>2</sup> Electronics and Information Department, CDCM Institute of Data Storage, Singapore, 139956, Singapore

Received 19 January 2014; Accepted 27 July 2014

### Abstract

Monitoring data on an earth-rockfill dam constitutes a form of spatial data. Such data include much uncertainty owing to the limitation of measurement information, material parameters, load, geometry size, initial conditions, boundary conditions and the calculation model. So the cloud probability density of the monitoring data must be addressed. In this paper, the cloud theory model was used to address the uncertainty transition between the qualitative concept and the quantitative description. Then an improved algorithm of cloud probability distribution density based on a backward cloud generator was proposed. This was used to effectively convert certain parcels of accurate data into concepts which can be described by proper qualitative linguistic values. Such qualitative description was addressed as cloud numerical characteristics--  $\{Ex, En, He\}$ , which could represent the characteristics of all cloud drops. The algorithm was then applied to analyze the observation data of a piezometric tube in an earth-rockfill dam. And experiment results proved that the proposed algorithm was feasible, through which, we could reveal the changing regularity of piezometric tube's water level. And the damage of the seepage in the body was able to be found out.

*Keywords:* Cloud Theory; Cloud Probability Distribution Density; Earth rockfill dam; Observation Data

### 1. Introduction

There is inherent uncertainty in the monitoring data of an earth-rockfill dam. So the analysis is closely related to the probability density of the monitoring data. High possibility data in the domain of data make a greater contribution to knowledge than low possibility data due to the randomness of monitoring data. If the probability density of data is overlooked, the data may become a departure from the value of the real knowledge.

In light of the uncertainties and complexities in dam engineering problems, risk analysis is more essential now than ever before [1]. Many methods have been proposed to solve the problem of uncertainty. The stochastic theory and fuzzy mathematics have proved to be key solutions for handling the fuzziness or randomness in spatial data.

As for the stochastic theory, Wu [2] thought that the spatial variability of soil properties was mainly dominated by their geological provenance, so the constraint random field was proposed and implemented utilizing random field theory and regionalized variable theory of geo-statistics to ensure that the random field realizations exactly match the parameters at the sample locations. LI [3] proposed a probability method for infinite slope stability analysis considering the variation of soil shear strength parameters with depth. Since the gravity dam and its foundation bore great loads and the operating conditions were very complex, the random FEM was used by Erfeng Zhao to calculate the effects under environmental loads [4].

Mechanical time-varying characteristics parameters of gravity dam foundation were calculated dynamically based on a hierarchical diagonal neural network with monitoring data.

However, only random uncertainties were considered for damage identification in most monitoring approaches. And fuzzy constituted a kind of uncertainty too. So an optimal group of damage fuzzy sets was used to classify a set of observations at any unknown state of damage under the principles of fuzzy pattern recognition based on maximum approaching degree [5]. Fuzzy mathematics were adopted to deal with the dam monitoring data and the membership function of the fuzzy matrix was established under the principle of maximum membership and evaluation set given that the relationship between monitoring variables of dam safety monitoring was fuzzy [6]. The main character of the fuzzy support vector machine was introduced and its formal description in detail was given by Long [7]. Then, a new method of fuzzy support vector machine was subjected to training and testing processes.

From the analysis above, it is clear that only randomness or fuzziness was considered in these papers. So in order to reduce the influence of fuzziness and randomness, cloud theory was proposed to cope with fuzziness and randomness, and the advantage of transition between the qualitative concept and quantitative description. In cloud theory, the observation data were incomplete samples of parent populations. So the probability distribution density of the parent population that described the observation data was the mathematical expectation function of the cloud model that consisted of cloud drops corresponded with the data. This was derived from the observed data and it was used by the

\* E-mail address: hanliwei13@163.com

cloud model's sample characteristics to estimate the parent population characteristics. The estimation error between the cloud model's sample and the parent population would be greatly reduced if the cloud probability distribution density and radiation energy of the corresponding data were fully taken into consideration.

## 2. Cloud Probability Distribution Density Algorithm

### 2.1 Cloud model [8]

Definition of cloud model: Suppose  $U$  is the set  $U=\{u\}$  which is the universe of discourse, and  $T$  is a term associated with  $U$ . The membership degree of  $u$  is in  $U$  to the term  $T$ ,  $C_T(u) \in [0,1]$  is a random number with a stable tendency. The cloud of  $T$  is a mapping from the universe of discourse  $U$  to the unit interval  $[0, 1]$ .

There is:

$$C_T(u): U \rightarrow [0,1]; u \in U; u \rightarrow \forall C_T(u)$$

The cloud model (Li et al, 1995) is a model of uncertainty transition between a linguistic term of a qualitative concept and its numerical representation. The cloud model can be characterized by three digital parameters- $C(Ex, En, He)$ , the expected value- $Ex$  represents the center of gravity of a cloud, the entropy- $En$  is a measurement of the fuzziness of the concept over the universe of discourse, which shows how many elements could be accepted by the term- $C$ , and the hyper entropy- $He$  is a measurement of the uncertainty of entropy- $En$ . The larger the value of  $He$ , the more random the set of membership degrees is distributed. Figure 1 shows the three numerical characteristics of the cloud model.

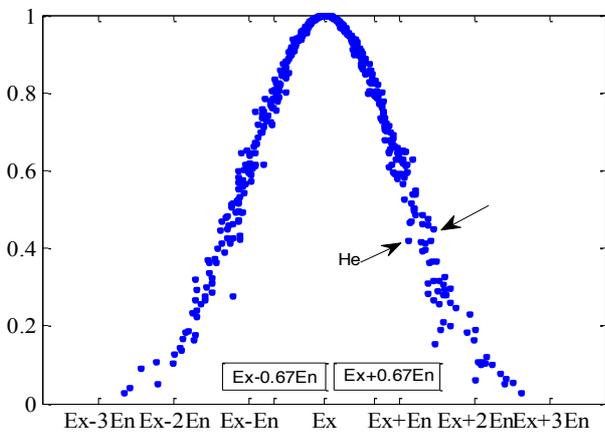


Fig. 1. Cloud model

Through statistical analysis, the cloud drops that contribute to qualitative concept- $C$  in universe- $U$ , are within interval  $[Ex-3En, Ex+3En]$ , and the contribution rate is 99.74%. The cloud drops within interval  $[Ex-0.67En, Ex+0.67En]$  account for 22.33% of the total quantitative numerical, and the contribution rate is 50%. These cloud drops are called the “backbone element”; the cloud drops within interval  $[Ex-En, Ex+En]$  account for 34.33% of the total quantitative numerical, and the contribution rate is 68.26%. These cloud drops are called the “basic element”; the cloud drops within interval  $[Ex-3En, Ex-2En]$  and  $[Ex+2En, Ex+3En]$  account for 34.33% of the total

quantitative numerical, and the contribution rate is 27.18%. These cloud drops are called the “peripheral element”; the cloud drops within interval  $[Ex-2En, Ex-En]$  and  $[Ex+En, Ex+2En]$  account for 34.33% of the total quantitative numerical, and the contribution rate is 4.3%. These cloud drops are called the “weak peripheral element”. The contribution of cloud drops in different regions to the qualitative concept is shown in Figure 1.

### 2.2 Cloud probability distribution density

Suppose  $L_n = \{l_1, l_2, \dots, l_n\}$  are the cloud drops of samples- $W$  to concept- $T$  in universe- $L$ . Samples- $W$  comes from the parent population- $\Omega$ , and  $n$  is the sample size,  $f(l)$  is the probability density function of parent population- $\Omega$ . Suppose  $x = \varphi(l-l_i)$ , ( $i=1,2, \dots, n$ ). There is a Borel measurable function within  $(-\infty, +\infty)$ , which represents stochastic membership conception  $T$  of  $x$ , as is denoted by  $C_T(x)$ . The membership conception- $T$ 's data energy of  $l_i$  can be radiated to  $l$ . Suppose  $t_i =$  “around of  $l_i$ ”, and the initial distribution of radiated cloud drops can be expressed by  $Q(x) = \sum C_T(x) = \sum C_T(g(l-l_i))$ .  $C_T(x)$  is a radiation brightness function for conception- $T$  in universe- $L$ .  $x = (l-l_i)/d, d > 0$  can be deduced from  $x = \varphi(l-l_i)$ , in which  $d$  is called radiation unit, and formula (1) is deduced to be [9]:

$$\hat{f}(l) = \frac{1}{n} \sum_{i=1}^n C_T(x) = \frac{1}{nd} \sum_{i=1}^n C_T\left(\frac{l-l_i}{d}\right) \quad (1)$$

From formula (1) we can see that the cloud probability distribution density is an estimation of the parent cloud model according to data radiation.  $C_T(x)$  and  $d$  are keys to radiation estimation. It is important to obtain an appropriate radiation brightness function and the radiation unit through sample- $W$ . Cloud drops radiation takes into consideration randomness and fuzziness at the same time, selects the radiation spot of the data energy as the kernel, and redistributes the total radiation brightness of the sample- $W$ . Thus, new membership sets  $Q(x) = \{Q_1/nd, Q_2/nd, \dots, Q_n/nd\}$  of concept- $T$  have taken shape. It is a random distribution in the basic universe- $L$  by the cloud drops radiation. When  $x = (l-l_i)/d$ , then  $Q_n/nd$  is the kernel estimation (Parzen, 1962). According to formula (1), the probability density of cloud drops can be estimated based on the data radiation energy of cloud drops. In contrast, the initial concept sets can be established from the probability density function of spatial data distribution through cloud transformation. Obviously, it is an inverse cloud process. Data radiation provides an inverse explanation for cloud transformation from the perspective of physics.

The derivative of the radiation brightness function can be expressed as:

$$C_T'(x_0) = \lim_{\Delta x \rightarrow 0} \frac{\Delta C_T(x)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{C_T(x_0 + \Delta x) - C_T(x_0)}{\Delta x} \quad (2)$$

Which is shown in Figure 2.

The radiation brightness function  $C_T(x)$  must match the constraint condition [10]:

$$0 \leq C_T(x) \leq \frac{1}{\sqrt{2\pi}\sigma} \quad (3)$$

$$\int_{-\infty}^{+\infty} C_T(x)dx = \int_R |C_T(x)|dx = 1 \quad (4)$$

$$\lim_{|x| \rightarrow \infty} xC_T(x) = 0 \quad (5)$$

$$\int_R xC_T(x)dx = 0 \quad (6)$$

$$\int_R x^2 C_T(x)dx = \sigma^2 \quad (7)$$

$$\int_R C_T^2(x)dx = \frac{1}{\sqrt{2\pi\sigma}} \quad (8)$$

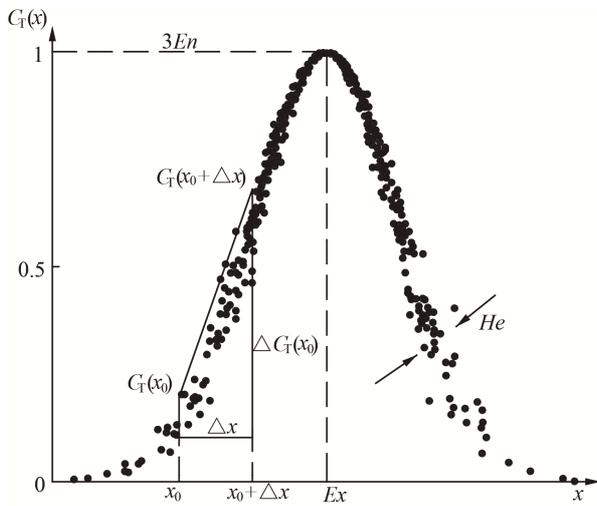


Fig. 2. The derivative of  $C_T(x)$

According to formulae (3) ~ (8), another characteristic of the radiation brightness function can be demonstrated.

Suppose  $l_i$  is an independent random observation in data space- $R$ , the expectation of formula (1) can be expressed as:

$$E\hat{f}_n(l) = \frac{1}{nd} \sum_i EC_T\left(\frac{l-l_i}{d}\right) = \frac{1}{nd} \sum_i \int_R C_T\left(\frac{l-x}{d}\right) f(x)dx \quad (9)$$

$$= \frac{1}{d} \int_R C_T\left(\frac{l-x}{d}\right) f(l-x)dx$$

$$\left| E\hat{f}(l) - f(l) \right| = \left| \int_R (f(l-x) - f(l)) \frac{1}{d} C_T\left(\frac{l-x}{d}\right) dx \right| \quad (10)$$

$$= \left| \int_R A(x)dx \right| \leq \left| \int_{|\eta|} A(x)dx \right| + \left| \int_{R-\eta} A(x)dx \right|$$

In formula (10),  $\eta$  is a neighborhood of 0. If  $\eta \rightarrow 0$ , then  $\left| \int_{|\eta|} A(x)dx \right| \rightarrow$  arbitrarily small, which can be addressed as  $O_1$ , then

$$\left| E\hat{f}(l) - f(l) \right| \leq O_1 + \left| \int_{R-\eta} f(l-x) \frac{1}{d} C_T\left(\frac{x}{d}\right) dx \right| \quad (11)$$

$$+ f(l) \left| \int_{R-\eta} \frac{1}{d} C_T\left(\frac{x}{d}\right) dx \right|$$

$$\leq O_1 + \sup \left| \frac{x}{d} C_T\left(\frac{x}{d}\right) \int_R f(x)dx \right| + f(l) \left| \int_{R-\eta} \frac{1}{d} C_T\left(\frac{x}{d}\right) dx \right| \leq O_1 + O_2 + O_3$$

If  $\eta \rightarrow \infty$ , then  $\left| \frac{\eta}{d} \right| \rightarrow \infty$ . After that, according to formula (5),  $O_2 \rightarrow 0$  can be obtained. Finally,  $O_3 = \int_{R-\frac{\eta}{d}} C_T(x)dx$ ,

$\lim_{n \rightarrow \infty} \frac{\eta}{d} \rightarrow R$ , and based on formula (4),  $O_3 \rightarrow 0$  can be obtained. So  $\lim_{n \rightarrow \infty} E\hat{f}(l) = f(l)$  can be demonstrated.

### 2.3 Radiation unit and best radiation unit

According to formula (1) and formula (10), the evaluation  $\hat{f}(l)$  of the parent population's probability density function depends on the radiation unit  $d$ , the monitoring value  $l_i$  and the number  $n$  of monitoring value (sample size). After the observation, the monitoring value of  $l_i$  and the number of  $n$  are known, but radiation unit  $d$  is unknown. So  $d$  must be decided firstly. Suppose  $E$  is the expectation, and  $D$  is the variance, according to formula (4) and formula (5), the variance of  $C_T(x)$  can be obtained:

$$\begin{aligned} D\hat{f}(l) &= \frac{1}{n^2 d^2} \sum_i DC_T\left(\frac{l-l_i}{d}\right) = \frac{1}{nd^2} DC_T\left(\frac{l-l_1}{d}\right) \\ &= \frac{1}{nd^2} E \left( C_T\left(\frac{l-l_1}{d}\right) - EC_T\left(\frac{l-l_1}{d}\right) \right)^2 \\ &= \frac{1}{nd^2} \left( EC_T^2\left(\frac{l-l_1}{d}\right) - (EC_T\left(\frac{l-l_1}{d}\right))^2 \right) \leq \frac{1}{nd^2} EC_T^2\left(\frac{l-l_1}{d}\right) \quad (12) \end{aligned}$$

$$= \frac{1}{nd^2} \int_R C_T^2\left(\frac{l-x}{d}\right) f(x)dx$$

$$= \frac{1}{nd^2} \int_R C_T^2(x/d) f(l-x)dx$$

$$\lim_{x \rightarrow \infty} D\hat{f}(l) = 0 \quad (13)$$

Suppose  $\int_R xC_T(x)dx = 0, k = \int_R x^2 C_T(x)dx < \infty$ , then

$$\begin{aligned} \frac{1}{d^2} (E\hat{f}(l) - f(l)) &= \frac{1}{d^2} \int_R C_T(x) (f(l-xd) - f(l)) dx \\ &+ \frac{1}{d^2} f'(l) \int_R xC_T(x) dx \quad (14) \end{aligned}$$

Suppose  $|\eta| \leq 1$ , and according to the mean value theorem,  $\frac{1}{d^2} (E\hat{f}(l) - f(l))$  can be expressed as:

$$\begin{aligned} \frac{1}{d^2} (E\hat{f}(l) - f(l)) &= \int_R C_T(x) \frac{f(l-xd) + xdf'(l) - f(l)}{d^2} \\ &= \frac{1}{2} \int_R C_T(x) x^2 f''(l - \eta xd) dx \quad (15) \end{aligned}$$

According to the dominated convergence theory, if  $f''(l) \neq 0$ ,  $O$  is the infinitesimal of higher order, then

$$E\hat{f}(l) - f(l) = \frac{1}{2} f''(l)kd^2 + O(d^2) \quad (16)$$

The mean square error is:

$$\begin{aligned} MSE(\hat{f}(l)) &= D\hat{f}(l) + (E\hat{f}(l) - f(l))^2 \\ &= \frac{1}{nd} f(l) \int_R C_T^2(x) dx + \frac{(f''(l))^2 k^2 d^2}{4} + O \\ &= A(d) + O \end{aligned} \quad (17)$$

From formula (17) we can see that the value of radiation unit is the key factor, and it is essential to obtain the appropriate radiation unit through sample  $W$ .

The best radiation unit is defined as a unit with the minimum mean square error. According to formula (17), the main part  $A(d)$  of mean square error is a function of  $d$ . Suppose the derivative of  $A(d)$  is 0, then:

$$-\frac{f(l)}{nd^2} \int_R C_T^2(x) dx + (f''(x))^2 k^2 d^3 dx = 0 \quad (18)$$

And there is:

$$d^5 = -\frac{f(l) \int_R C_T^2(x) dx}{n(f''(x))^2 k^2} \quad (19)$$

Substitute formula (19) into formula (17), formula (20) is expressed by:

$$MSE(\hat{f}(l)) = \frac{5f^2(l)}{4n^{0.8}} \left( \int_R C_T^2(x) dx \right)^{0.8} [kf''(l)]^{0.4} + O\left(\frac{1}{n^{0.8}}\right) \quad (20)$$

Suppose  $f(l) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{l^2}{2\sigma^2}}$ , then:

$$d = \left( \frac{\sigma^9}{2n(l^2 - \sigma^2)^2 e^{-\frac{l^2}{2\sigma^2}}} \right)^{0.2} \quad (l \neq \pm\sigma) \quad (21)$$

However,  $f(l)$  is unknown in the real situation. So the best radiation unit cannot be worked out from formula (15).

According to the properties of radiation brightness function-  $C_T(x)$ , when  $n \rightarrow$  maximum, there is  $f(l_i) \approx \hat{f}(l_i)$ .

As  $f'(l) = \lim_{h \rightarrow 0} \frac{f(l+d) - f(l)}{d}$ ,

$f''(l) = \lim_{h \rightarrow 0} \frac{f(l+d) - 2f(l) + f(l-d)}{d^2}$ ,  $\lim_{n \rightarrow \infty} d_n = 0$ , Suppose

radiation unit is  $d$ , and get rid of the sign of limits, there is:

$$f''(l_i) = \frac{\hat{f}(l_i + d) - 2\hat{f}(l_i) + \hat{f}(l_i - d)}{d^2} \quad (22)$$

Substitute  $f(l_i) \approx \hat{f}(l_i)$  and formula (20) into formula (19), and the iteration computational equations of radiation unit can be worked out.

$$\left. \begin{aligned} d^0 &= (\max(l) - \min(l)) / (n - 1) \\ d_i^{j+1} &= \left\{ \frac{0.26833 \hat{f}^j(l_i)(h^j)^4}{n[\hat{f}^j(l_i + d^j) - 2\hat{f}^j(l_i) + \hat{f}^j(l_i - d^j)]^2} \right\}^{0.2} \\ d^{j+1} &= \text{median}(d_i^{j+1}) \\ \hat{f}^{j+1}(l) &= \frac{1}{nd^j} \sum_T C_T \left( \frac{l-l_i}{d^j} \right) \end{aligned} \right\} \quad (23)$$

#### 2.4 The best radiation brightness function

The best radiation brightness function can be determined by the best radiation unit.

Suppose  $k \left( \int_R C_T^2(x) dx \right)^2 \rightarrow$  minimum, there is:

$$C_T(x) = \begin{cases} \frac{3}{20\sqrt{5}}(5-x^2) & |x| < \sqrt{5} \\ 0 & |x| \geq \sqrt{5} \end{cases} \quad (24)$$

#### 2.5 The cloud numerical characteristics based on data radiation

On the basis of a normal cloud generator and a backward cloud generator, the mapping relationship between qualitative and quantitative can be established through three numerical characteristics: Expectation ( $Ex$ ), Entropy ( $En$ ) and Hyper Entropy ( $He$ ). In the process of qualitative and quantitative transformation through the cloud model, it is important to consider the premises of using cloud probability density that is in line with the data. Each cloud drop radiates the data energy for a concept in the number field space. Thus, all the cloud drops distribution obeys a certain probability density distribution on the whole. From the perspective of radiation data, cloud numerical characteristics must be closely related to the probability density of cloud drops, and the probability density of it should not be overlooked. Otherwise, we cannot get the parent population cloud. The algorithm of cloud numerical characteristics is described as follows.

First, estimate the probability density of cloud drops based on data radiation. Second, use the probability density of cloud drops to measure the weight of each sample of cloud drops to the parent population. Finally, according to the weighted average algorithm of spatial distribution, calculate the three numerical characteristics (Expectation ( $Ex$ ), Entropy ( $En$ ) and Hyper Entropy ( $He$ )).

$$\hat{Ex}(l) = \frac{\sum_{i=1}^n \hat{f}(l_i) l_i}{\sum_{i=1}^n \hat{f}(l_i)} \quad (25)$$

$$\hat{En}(l) = \sqrt{\frac{\sum_{i=1}^n \hat{f}(l_i)(l_i - \hat{Ex})}{\sum_{i=1}^n \hat{f}(l_i)}} \quad (26)$$

$$\hat{He}(l) = \sqrt{\frac{\sum_{i=1}^n [\hat{f}(En(l_i))(En(l_i) - \hat{En}(En(l_i)))^2]}{\sum_{i=1}^n \hat{f}(En(l_i))}} \quad (27)$$

From formula (25) ~ formula (27) we know that the algorithm fits the spatial distribution characteristics of cloud drops. The algorithm is easy to understand and is superior to the arithmetic average method.

### 2.6 Cloud expectation function based on data radiation

The cloud expectation function is the probability density function of the spatial entities' parent population. It is also the mathematical expectation function of cloud drops that are in line with the sample of spatial data. So the cloud expectation function can be acquired by the probability density of cloud drops. The cloud expectation function is:

$$C_T(l) = \frac{1}{\hat{f}(l)} e^{-\frac{[\hat{f}(l)(l - \hat{Ex})]^2}{2[\hat{En}(l)]^2}} \quad (28)$$

If the probability density of cloud drops is equal anywhere in the number field space, then  $\hat{f}(l) = 1$ , and the cloud expectation function is the same as the function

$$C_T(l) = e^{-\frac{(l - \hat{Ex})^2}{2(\hat{En})^2}}$$

of the normal cloud model.

According to the cloud expectation function, the three numerical characteristics of the cloud model are acquired from the given data by using the cloud generator. Based on this, it is possible to describe the morphological characters of the cloud model. So the cloud radiation expectation function is in line with the backward cloud generator.

### 2.7 cloud radiation algorithm

The cloud radiation algorithm is based on the backward cloud generator, which is an uncertainty transformation model realizing the transformation between a numeric value and its linguistic value. It effectively converts a certain number of accurate data to the concept indicated by appropriate qualitative linguistic values  $\{Ex, En, He\}$  which are the characteristics of the whole drops. Based on the data radiation, a new improved algorithm of backward cloud generator is described as follows [11]:

**Input:** Coordinate value  $l_i$  of each cloud drops and its degrees of certainty- $C_T(l_i)$ ;

**Output:**  $Ex, En$  and  $He$ , cloud drops' number  $N$ ;

- Fit the cloud expectation curve

$$C_T(l) = \frac{1}{\hat{f}(l)} e^{-\frac{[\hat{f}(l)(l - \hat{Ex})]^2}{2[\hat{En}(l)]^2}}$$

of cloud drops, and get

$$\hat{Ex}(l) = \frac{\sum_{i=1}^n \hat{f}(l_i)l_i}{\sum_{i=1}^n \hat{f}(l_i)}$$

which is the estimation value of

$Ex$ ;

- Eliminate the dots that fit the condition of  $C_T(l) > 0.999$ , and there are  $m$  cloud drops left;

- Calculate  $\hat{En}(l_i) = \frac{|f(l_i)(l_i - \hat{Ex})|}{\sqrt{-2 \ln(f(l_i)C_T(l_i))}}$ ;

- Calculate the estimation value of  $En$ :

$$\hat{En}(l) = \sqrt{\frac{\sum_{i=1}^m [\hat{f}(En(l_i))En(l_i)]}{\sum_{i=1}^m \hat{f}(En(l_i))}}$$

- Calculate the estimation value of  $He$ :

$$\hat{He}(l) = \sqrt{\frac{\sum_{i=1}^n [\hat{f}(En(l_i))(En(l_i) - \hat{En}(En(l_i)))^2]}{\sum_{i=1}^n \hat{f}(En(l_i))}}$$

### 3. Analysis of observation data based on cloud probability distribution density algorithm

Luhun reservoir is an earth-rockfill dam in Luoyang, China. 13 piezometric tubes of Luhun reservoir were selected for the purposes of calculation. Piezometric tubes S1~S9 were embedded in the body of the dam and piezometric tubes D2~D5 were embedded in the fault zone of the dam's foundation. Piezometric tubes S1, S4 and S7 were embedded in section 1 (0+813); piezometric tubes S2, S5 and S8 were embedded in section 2 (0+660); piezometric tubes S3, S6 and S9 were embedded in section 3 (0+540); piezometric tubes D2 and D4 were embedded in section 5 (0+640); piezometric tubes D3 and D5 were embedded in section 6 (0+680). See Figure 3.

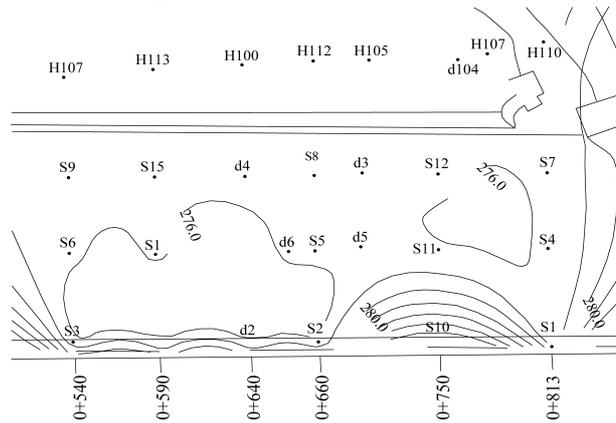


Fig.3. The location of each piezometric tube

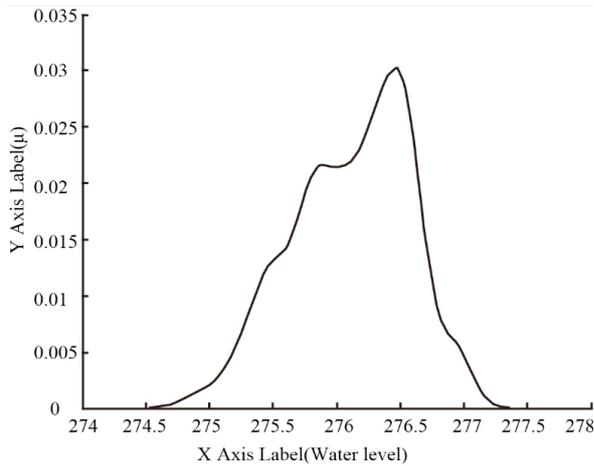
The probability distribution density of the piezometric tubes' water levels between 1979~1999 were acquired, and the results are shown in figure 4~ figure 16, in which the vertical axis represents probability density estimation and the horizontal axis represents water level value. The qualitative linguistic values (three numerical characteristics  $\{Ex, En, He\}$ ) are used to describe the monitoring results, which can be obtained through the improved backward

cloud generator algorithm. The three numerical characteristics  $\{Ex, En, He\}$  are shown in Table I.

According to the qualitative explanatory rules,  $Ex$  refers to the water level of a piezometric tube;  $En$  refers to the dispersion degree between water level and expected water level of a piezometric tube;  $He$  refers to the influence of monitoring instruments and monitoring environment on monitoring level. From figure 4~ figure 16 and Table 1 we can see that the saturation line of Section 1 is the highest, the saturation line of Section 2 is the second, the saturation line of Section 3 is the lowest, the saturation line of each section is not high and in a stable state; all the piezometric tubes' dispersion degree of each section are in a low state and the maximum value ( $En$ ) is less than 1. The maximum value ( $He$ ) of piezometric tube in each section is 0.13, which indicates that the monitoring level is stable.

**Tab.1.** Numerical characteristics of piezometric tubes' water level

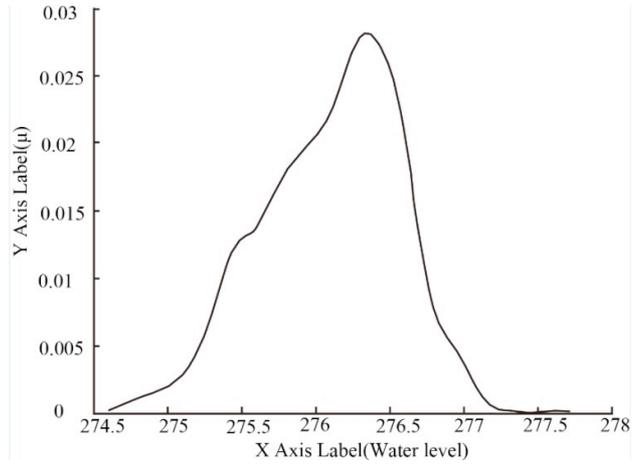
Location	Section	Piezometric tube	$Ex$	$En$	$He$	
In the body of the dam	Section 1	S1	276.12	0.49	0.13	
		S4	276.11	0.47	0.11	
		S7	276.07	0.5	0.13	
	Section 2	S2	276.11	0.42	0.07	
		S5	276.03	0.47	0.11	
		S8	276.02	0.46	0.07	
		S3	276.01	0.44	0.11	
		Section 3	S6	276.01	0.45	0.08
			S9	276.01	0.43	0.08
In the fault zone	Section 4	D2	275.77	0.83	8.1	
		D4	275.96	0.44	0.09	
	Section 5	D3	276.00	0.44	0.09	
		D5	276.05	0.46	0.13	



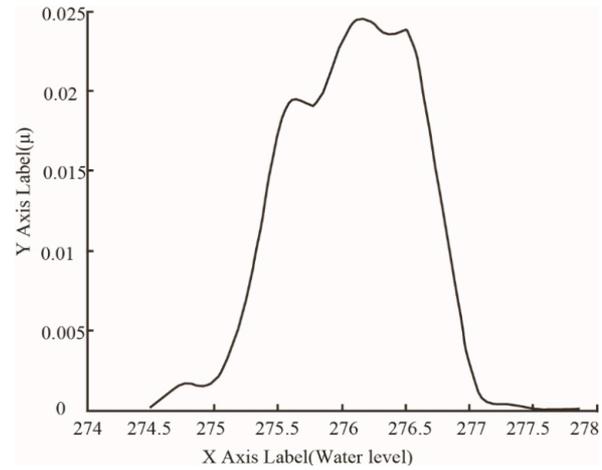
**Fig. 4.** Probability density of S1

The water level of the piezometric tubes in section 4 and section 5 is lower than that in the body of the dam. The water levels are in a stable state except in piezometric tube D2. The entropy ( $En$ ) of piezometric tube D2 is 0.83, and the hyper entropy ( $He$ ) of piezometric tube D2 is 8.1. This means that there are some problems about the dispersion degree and the monitoring level is affected by monitoring

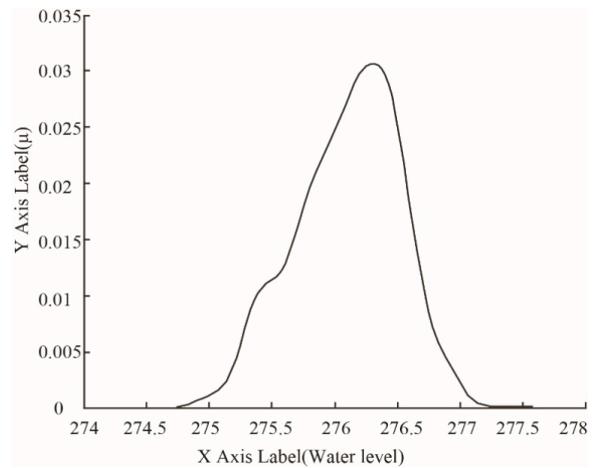
instruments and the monitoring environment. So piezometric tube D2 needs to be studied further and treated in the fault zone.



**Fig. 5.** Probability density of S4



**Fig. 6.** Probability density of S7



**Fig. 7.** Probability density of S2

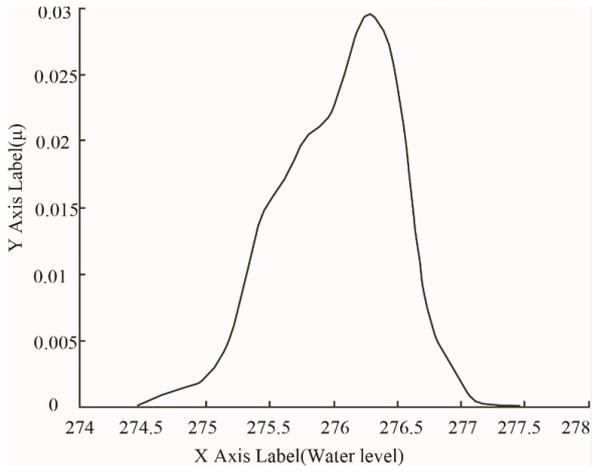


Fig. 8. Probability density of S5

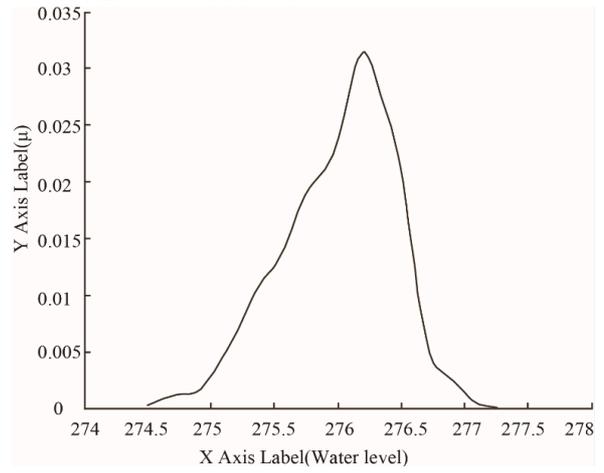


Fig. 11. Probability density of S6

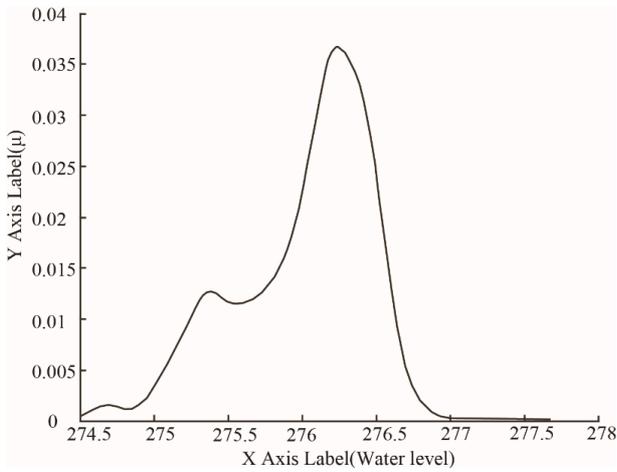


Fig. 9. Probability density of S8

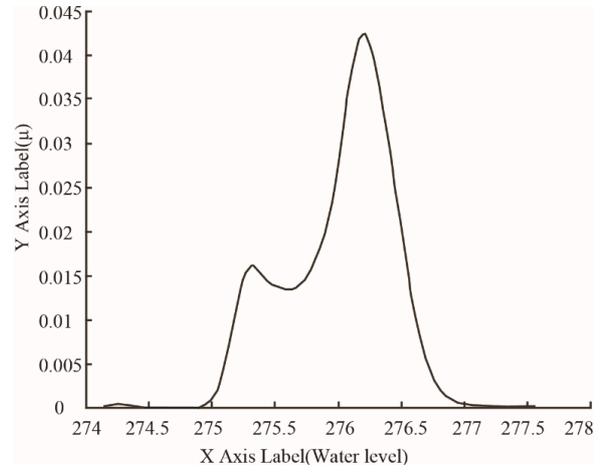


Fig. 12. Probability density of S9

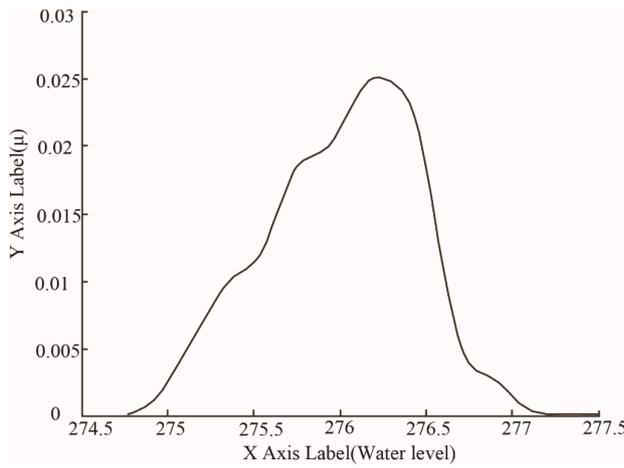


Fig. 10. Probability density of S3

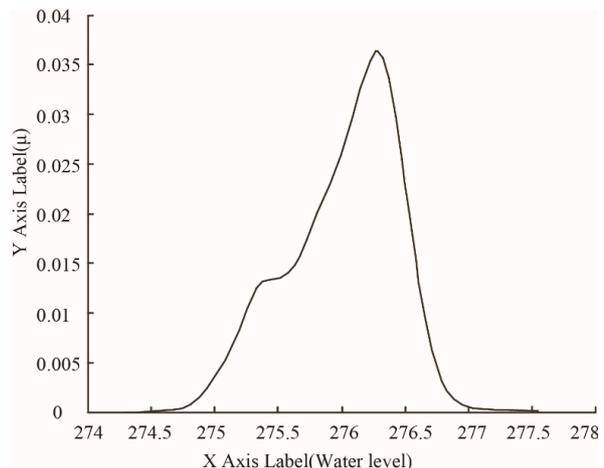


Fig. 13. Probability density of D2

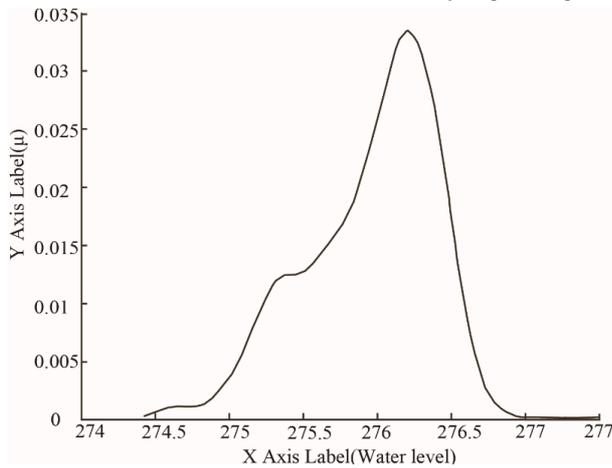


Fig. 14. Probability density of D4

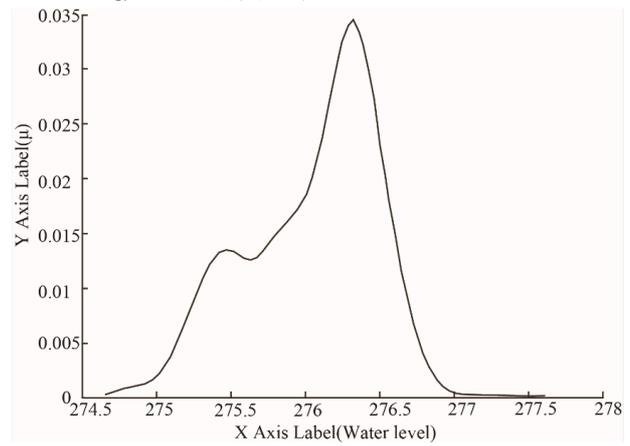


Fig. 16. Probability density of S1

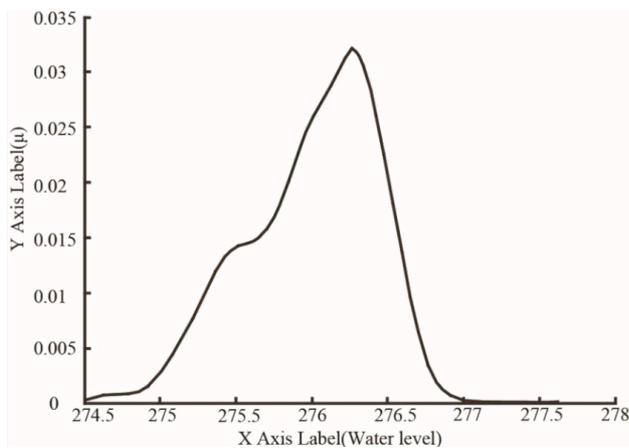


Fig. 15. Probability density of D3

#### 4. Conclusions

This paper proposes an improved algorithm of cloud probability density based on a backward cloud generator and applies it to the analysis of monitoring data of piezometric tube water levels in an earth-rockfill dam. The corresponding cloud probability density distribution map and cloud numerical characteristics are demonstrated and serve to monitor the seepage of the dam. The results show that most of the tubes are in a stable state. Only D2 presents excessive entropy and hyper entropy, which may be caused by a change in seepage or outdated equipment, but which needs further inspection.

However, the algorithm still has the following loopholes. First, due to limited space, this paper fails to consider the special distribution of data. So the influence of location uncertainty on the monitoring of the parent population has been overlooked. Second, this paper assumes that the monitoring data are in line with Gaussian distribution, which requires further proof and research.

#### Acknowledgments

The authors wish to thank Natural Science Fund Committee (51279064 and U1304511) and the editor of the journal.

#### References

1. Rohaninejad, Mohsen, Bagherpour, Morteza, "Application of risk analysis within value management: A case study in dam engineering". *Journal of Civil Engineering and Management*, 19(3), 2013, pp. 364-374.
2. Zhenjun Wu, Wang Shuilin, Ge Xiurun, "Slope reliability analysis by random FEM under constraint random field". *Rock and Soil Mechanics*, 30(10), 2009, pp. 3086-3092.
3. Li Dianqing, Qi Xiaohui, Zhou Chuangbing, Zhang Limin. "Reliability analysis of infinite soil slopes considering spatial variability of soil parameters". *Chinese Journal of Geotechnical Engineering*, 35(10), 2013, pp. 1799-1806.
4. Erfeng Zhao, Kai Zhu, Jiangyang Liu. "Random dynamical model for mechanical time-varying characteristics of gravity dam foundation". *Engineering Journal of Wuhan University*, 46(1), 2013, pp. 89-94.
5. Altunok Erdogan, Taha Mahmoud M. Reda, Epp David S, Mayes Randy L, Baca Thomas J, "Damage pattern recognition for structural health monitoring using fuzzy similarity prescription". *Computer-Aided Civil and Infrastructure Engineering*, 21(8), 2006, pp. 549-560.
6. Cui Shaoying, Bao Tengfei, Pei Yaoyao, Chen Shanying, "Data processing of dam safety monitoring based on fuzzy mathematical approach", *Water Resources and Power*, 30(10), 2012, pp. 45-48.
7. Yanjun Long, Jianquan Ouyang, Xinwen Sun, "Network Intrusion Detection Model based on Fuzzy Support Vector Machine", *Journal of Networks*, 8(6), 2013, pp. 1387-1394.
8. Li Deyi, Du Yi, *Artificial intelligence with uncertainty*, Beijing: National defense industry press, 2005, pp.137-182.
9. Wang. Shuliang, "Data Field and Cloud Model Based Spatial Data Mining and Knowledge", *Thesis of Wuhan University*, 2002, pp.83-117.
10. Shannon, C., Weaver, W., *Mathematical theory of communication*, Illinois: University of Illinois Press, 1949, pp.158-178.
11. Li Deren, Wang Shuliang, Li Deyi, *Spatial data mining theories and applications*, Beijing: Science press, 2006, pp.425-470.