



Weight analysis of dam break risk consequences influencing factors

Zongkun Li¹, Wei Li¹, Wei Ge¹

¹School of Water Conservancy and Environment Engineering, Zhengzhou University 450001, China

Correspondence to: Wei Li (gewei@zzu.edu.cn)

5 **Abstract.** There are numerous influencing factors of the risk consequences of dam break. Scientific and reasonable index system and its weight distribution is one of the key element for comprehensive evaluation of the dam-break risk. Taking into consideration of 20 factors, including hazards, exposure and vulnerability factors, the evaluation index system of the consequences of the dam break risk is constructed. Using the Statistical Cloud Model (SCM) to improve the entropy method, we establish the weight calculation model of the influencing factors of the dam break risk consequences. The results shows
10 that the top 5 factors with the highest weight are risk population, flood intensity, alert time, risk understanding and distance from the dam. Compared to the traditional algebraic weight calculation methods, the result is basically consistent with the algebraic weight distribution, and increases the range by 2.03 times, supporting a more scientific basis for recognizing and evaluating the dam break risk consequences.

Keywords: Weight, dam break, risk consequences, entropy method, statistical cloud model

15 1 Introduction

Comprehensive evaluation of the risk consequences of dam break is the overall description of the severity of the consequences of the dam failure (Ling et al., 2009). The factors that can influence the risk consequences of dam break are usually composed of 3 factors, namely, hazards, exposure and vulnerability (Okada, 2004; Smith, 2013). The vulnerability factors can be further divided into 4 aspects: loss of life, economic loss, social impact and environmental impact. From the
20 point of view of system science, dam break flood disaster system is a dynamic system with high dimension, complexity and uncertainty. It accords with the development trend of risk assessment research "from low dimensional linearity to complex high-dimensional nonlinearity", "from single scale to multi-dimensional space-time scale", "from single scenario to combined scenario", "from certainty to uncertainty" (Zou et al., 2013).

The previous research on the index system of risk consequences and its weight is not sufficient. The uncertainty of the
25 impact of dam failure is explored, and suggestions for research index system are given (Lee and Noh, 2003; Wagenaar et al., 2016). The relationship between the hazard influencing factors and relationship between exposure and vulnerability factors is very complicated and the different types of flood including dam break flood can cause different degree of life loss (Jonkman et al., 2018; Wisner and Uitto, 2009). The indirect loss index for nature disasters is introduced and their weight is calculated by the traditional algebraic method (Daniell et al., 2018). The DAMBREAK computer programme is utilized to analysis the



downstream environmental impact and present 21 influence receptors, but the weight distribution of them is too average (Colomer and Gallardo, 2008). The Statistical Cloud Model (SCM) is used on the qualitative and quantitative transformation to analyze the regional water safety systems, but is not combine with the weight calculation (Ren et al., 2017). In the quantitative evaluation of risk consequences, we need to consider the combined effects of various factors, in which weight is a key part of it. The function of weight is to coordinate and balance the difference between the indexes. It is a measure to unify each index without considering the dimension difference between the indexes. In order to evaluate the risk consequences more comprehensively and objectively, many influencing factors are needed. However, too many indicators, more than 9 for example, will bring such problems like the difficulty in expert scoring and consistency test, and too average weight distribution.

In the course of calculating the weights, different methods have their own emphasis. For example, entropy weight method as one of the important methods of weight calculation, does not adequately consider the subjective opinions of experts. The analytic hierarchy process (AHP) is faced with the difficulty of consistency checking when dealing with the conditions of multiple factors (>9)(Su et al., 2016). When previous studies used the data of Statistical Cloud Model (SCM) to calculate weights, they had neglected the entropy when applied the SCM to convert the subjective opinions, resulting in the imperfection of information utilization (Mithas et al., 2011; Wan et al., 2015). These mentioned defects all lead to lack of scientificity in the calculation of weight. This manuscript introduces the SCM, which can reflect the fuzziness and randomness, to improve the entropy method for analyzing the weight of influencing factors of dam break risk consequences. The scientific influencing factors' weight will provide an important basis for further research on the dam-break risk comprehensive evaluation and for the establishment and improvement of dam risk management theory.

20 **2 Methods**

2.1 Risk index system

The establishment of evaluation index system is a systematic process. Scientific and reasonable evaluation index system is the guarantee for accurate risk assessment of dam failure, and the evaluation result is helpful for later research. Influencing factors of dam-break risk consequences are many and complicated in both quality and quantity, direct and indirect contribution, natural and social ways (Zhou et al., 2014). We choose the representative indicators as much as possible to reduce the mutual influence and derivative of the indicators. For example, the risk population is the most direct factor of life loss, we only set it in the life bearing bodies, even though it is influencing the economic and social aspects, but in indirect and less crucial way (Dutta et al., 2003). In the selection of the economic impact factors, the selection of GDP(Gross Domestic Product) per capita can better reflect the economic situation of the dam area. Compared with the GDP of the area, it is more accurate. Similarly, the comprehensive ability of water environment, soil environment and social carrying capacity is also selected. Whether the established index system is scientific and reasonable is directly related to whether it can objectively reflect the nature of the vulnerability itself. On the basis of aforementioned factors and characteristic of dam-



break flood system, we establish the risk influencing factor index systems scientifically and reasonably as shown in Fig. 1.

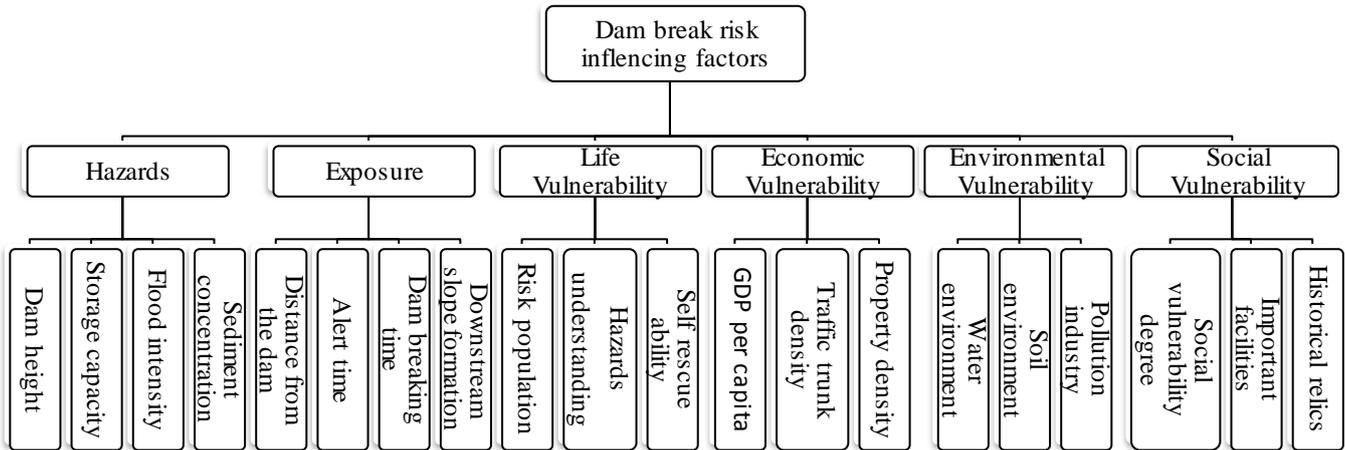


Figure 1. Index of dam-break risk consequences influencing factors .

2.2 Weight calculating model based on SCM-improved entropy method

Uncertainty is an intrinsic property of the objective world. The most important and most common uncertainties include fuzziness and randomness (Regas et al., 2010). The influencing factor system of the dam failure risk consequence is a multi-level and multi-index system with uncertainties (Li et al., 1995). In determining the importance of each risk factor to the comprehensive evaluation of the consequence, it needs a "quantitative conversion" of the uncertainty of the indicator. In the process of conversion, the expert's judgment makes a choice between many different factors that mutually affect each other and will absolutely lead to the ambiguity of boundaries, which is the fuzziness. On the other hand, the risk factors of dam break involve many aspects of life, economic loss, environmental and social impacts. In order to avoid the impact of expert's personal experience and subjective factors on the evaluation results, the risk factors of dam break need to adopt the method of group decision-making process. When an expert judged diverse risk factors, other experts must have some different opinions, reflecting in the randomness of judgments. Therefore, dam-break risk assessment system is a complex system integrating fuzziness and randomness. The SCM is invented under this context of the random and fuzzy feature of the dam break risk system. It describes the notions by the concept of clouds, reflects the randomness and fuzziness of concepts in natural language, and realizes the conversion between qualitative and quantitative information (Wang et al., 2016; Liu et al., 2018). In the process of group decision-making, the traditional method is only a simple algebraic operation of expert's ratings, which could not reflect the disagreements of different experts and the concentration of opinions. In fact, the experts' opinion is actually a rounded value that focuses on a certain degree of swing, which is using a stable tendency of the random number instead of the exact value, basically consistent with the central idea of SCM and the concept of entropy (Yari and Chaji,

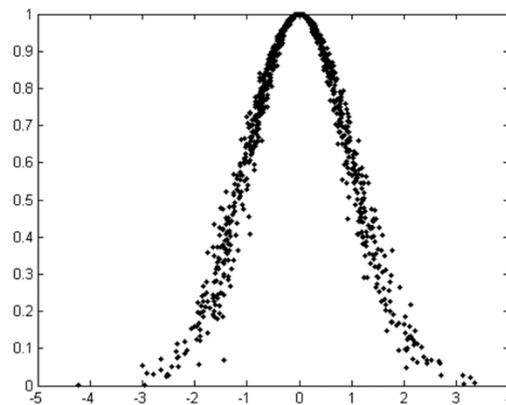


2012; Wang et al., 2016).

2.2.1 SCM Theory

The SCM, which was proposed by Li Deyi, is a model of uncertainty transformation between a qualitative concept and quantitative numerical representation (Li et al., 1995; Li et al., 2014). It mainly reflects the fuzziness and randomness of the concept of things or human knowledge in the objective world and integrates these two together. Constituting the mutual mapping between qualitative and quantitative, cloud generator is the key to its practical application.

Membership cloud: Suppose a universe $U=\{x\}$, L is the language value of the link in U . The membership degree $R_L(x)$ of the element x in U to the qualitative concept expressed by L is a stable random number. The membership degree distributed in the universe of discourse is called the membership cloud as shown in Fig. 2.



10

Figure 2. Sketch map of a Membership Cloud.

The x and y axes are for the expectation number and probability of distribution respectively. $R_L(x)$ takes a value between 0 and 1, whereas the cloud represents the mapping from the universe U to the interval $[0,1]$, that is: $R_L(x): U \rightarrow [0,1], \forall x \in U, x \rightarrow R_L(x)$

15 It can be seen that the qualitative concept to the quantitative value on the universe U is a one-to-many mapping relation, rather than a one-to-one relationship on the traditional fuzzy function. The degree of membership of x to L is a probability distribution, not a fixed value. SCM uses the expectation (Ex), entropy (En) and hyper entropy (He) as a whole to characterize an uncertain concept.

Expectation(Ex): The mathematical expectation of cloud drop distribution in the universe of discourse, that is, the domain value corresponding to the centric of the area under the coverage of the membership cloud, which is the domain value x of the degree of membership. Generally, it is the point most capable of characterizing the qualitative concept, reflecting the information centre value of the corresponding fuzzy concept.

Entropy (En): En is a measure of the ambiguity of a qualitative concept, reflecting the range of values that can be accepted



by the concept in the universe U . In the SCM, entropy is mainly used to measure the ambiguity and probability of qualitative concepts, reflecting the uncertainty of qualitative concepts. The larger the En is, the larger the range of values can be accepted by the concept and the more obscure the concept is. It embodies the flexibility of qualitative language.

Hyper-entropy (He): The measure of En uncertainty, entropy of entropy, reflects the discreteness of cloud drops. When the He is larger, the dispersion of cloud droplets is greater, that is, the greater the randomness of the membership value is, and the greater the "thickness" of the cloud can be. When it is closer to the concept centre or away from the centre, the randomness is relatively small, which is similar to a person's subjective feelings.

Cloud Generator: Generator is the most basic cloud algorithm, which can achieve quantitative range and distribution rules from the qualitative information expressed in language value. Cloud generators are mainly divided into the forward cloud generator and the backward cloud generator. The conversion process from qualitative concept to quantitative representation is conducted the forward cloud generator, the conversion process from quantitative representation to qualitative concept is produced by the backward cloud generator.

2.2.2 Entropy Method

The subjective weight analysis method is more dependent on the experts' opinions, and the consistency test under many factors is very difficult (Yari and Chaji, 2012). Therefore, this manuscript introduces entropy weight method as an objective weight calculation method. Entropy is a measure of uncertainty or randomness in information theory (Ouyang and Shi, 2013). In general, the more uncertain or random the event is, the more information it will contain, so the bigger entropy is. Therefore, the most important part of the entropy method is to obtain the differences in information, which is the degree of variation (Wang and Chen, 2016). According to the degree of variation of each index, we can calculate the entropy of each factor, and then use the entropy to adjust the weight of it, and finally, the objective weight value of the factors in the system is obtained.

The contribution of the numerical value in high-frequency or common consensus factor to the qualitative concept is greater than that of the numerical value in low-frequency (Yang and Nataliani, 2017). The En in the SCM could coincide with the idea of the entropy method in essence (Sheng et al., 2016; Dong et al., 2010). This manuscript makes use of the similar connotations of the SCM and entropy method. The objective advantages of entropy method need to be based on large amounts of score samples, which can be produced by the SCM cloud generator to get enough samples from limited experts' opinions. This manuscript attempts to use the SCM of qualitative-quantitative conversion model to improve the entropy method and make a scientific and objective response to the weight of risk factor.

2.2.3 Improved entropy method based on SCM

Suppose there are n indicators (column vectors) and m experts (row vectors). Each indicator computes the expectation and variance according to the cloud model. Statistical equation for calculating the j th indicator is as follow (Li et al., 1995):



$$Ex_j = \bar{x}_j = \frac{1}{m} \sum_{i=1}^m x_{ij} \quad (i = 1, \dots, m; j = 1, \dots, n) \quad (1)$$

$$En_j = \sqrt{\frac{\pi}{2} \frac{1}{m} \sum_{i=1}^m |x_{ij} - Ex_j|} \quad (i = 1, \dots, m; j = 1, \dots, n) \quad (2)$$

$$He_j = \sqrt{\frac{1}{m-1} \sum_{i=1}^m (x_{ij} - Ex_j)^2 - En_j^2} \quad (i = 1, \dots, m; j = 1, \dots, n) \quad (3)$$

The weight equation for the indicator calculated with the use of the conventional algebraic method is as follow:

$$5 \quad \omega_j = \frac{Ex_j}{\sum_{j=1}^n Ex_j} \quad (j = 1, \dots, n) \quad (4)$$

This algebraic method is easy to use, but it does not make any use of the changes of En in the SCM and may be misleading. For example, when the average scores of all indicators are the same, the weight of each indicator will calculate the same result. However, En_j and He_j could change greatly, but will not make enough reflection of the change in original equation, so an improved model is needed to replace this equation, as follow:

$$10 \quad \hat{\omega}_j = \begin{cases} \frac{Ex_j}{\ln(1+En_j)+1} \cdot \frac{1}{\sum_{j=1}^n \frac{Ex_j}{\ln(1+En_j)+1}} & (En_j \neq 0) \\ \frac{Ex_j}{\sum_{j=1}^n Ex_j} & (En_j = 0) \end{cases} \quad (5)$$

If the En_j is not equal to 0, the equation of the weight is revised and the cloud entropy is involved in the calculation. The larger the cloud entropy, the more divergence of opinions the expert has on the index, so the weight of the index should be reduced. The smaller the entropy is, the smaller the expert's disagreement on the indicator, so the weight of the indicator should be increased. When the minimum entropy En_j is equal to 0, indicating that the indicators of the experts have the same score, and then the weight of the equation remained unchanged.

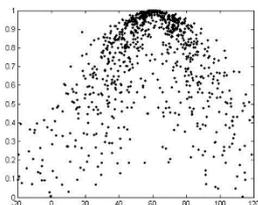
3 Results and discussion

3.1 Experts' scoring

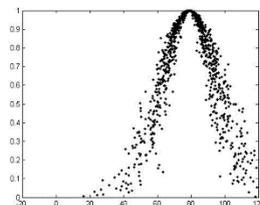
According to the requirement of data volume based on entropy method, we invite 20 experts to score the index system. Each index scoring adopts 100 integral points system, according to the importance without any comparison between each other. Scoring points should be scored from the perspective of comprehensive assessment of the risk consequences of the dam break. This scoring way can reveal the experts' opinion properly without imply any preference of the factors, and makes the scoring process easier. In accordance with the result of the score obtained by the backward cloud generator based on Eq. (1) to Eq. (3), its Ex_j , En_j , He_j are obtained. In order to reflect the model characteristics of expert scoring more intuitively, the outstanding advantage of SCM, we present the sketch map of these 20 factors' membership cloud as in the Fig. 3:



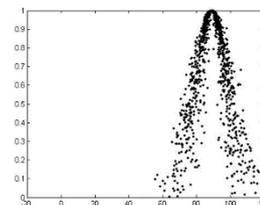
Hazards (H)



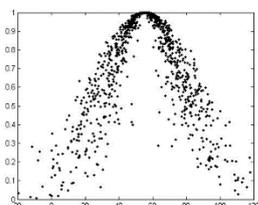
H1: Dam height
 (Ex:60.5 En:29.5 He:15.3)



H2: Storage capacity
 (Ex:79 En:16.8 He:3.6)

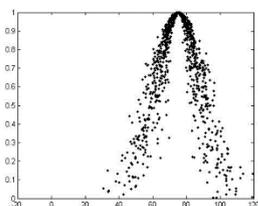


H3: Flood intensity
 (Ex:89 En:9.5 He:3)

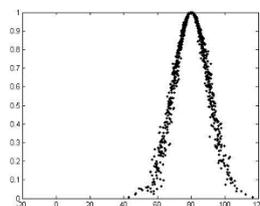


H4: Sediment concentration
 (Ex:55.5 En:24.4 He:6.7)

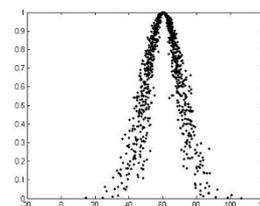
Exposure (E)



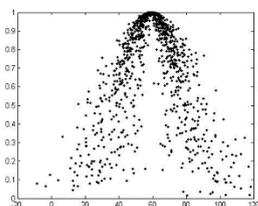
E1: Distance from the dam
 (Ex:75.2 En:13.5 He:3.8)



E2: Alert time
 (Ex:80.2 En:10.6 He:1.2)



E3: Dam breaking time
 (Ex:60.5 En:10.5 He:2.3)



E4: Downstream slope
 formation
 (Ex:59.3 En:18.4 He:7.6)

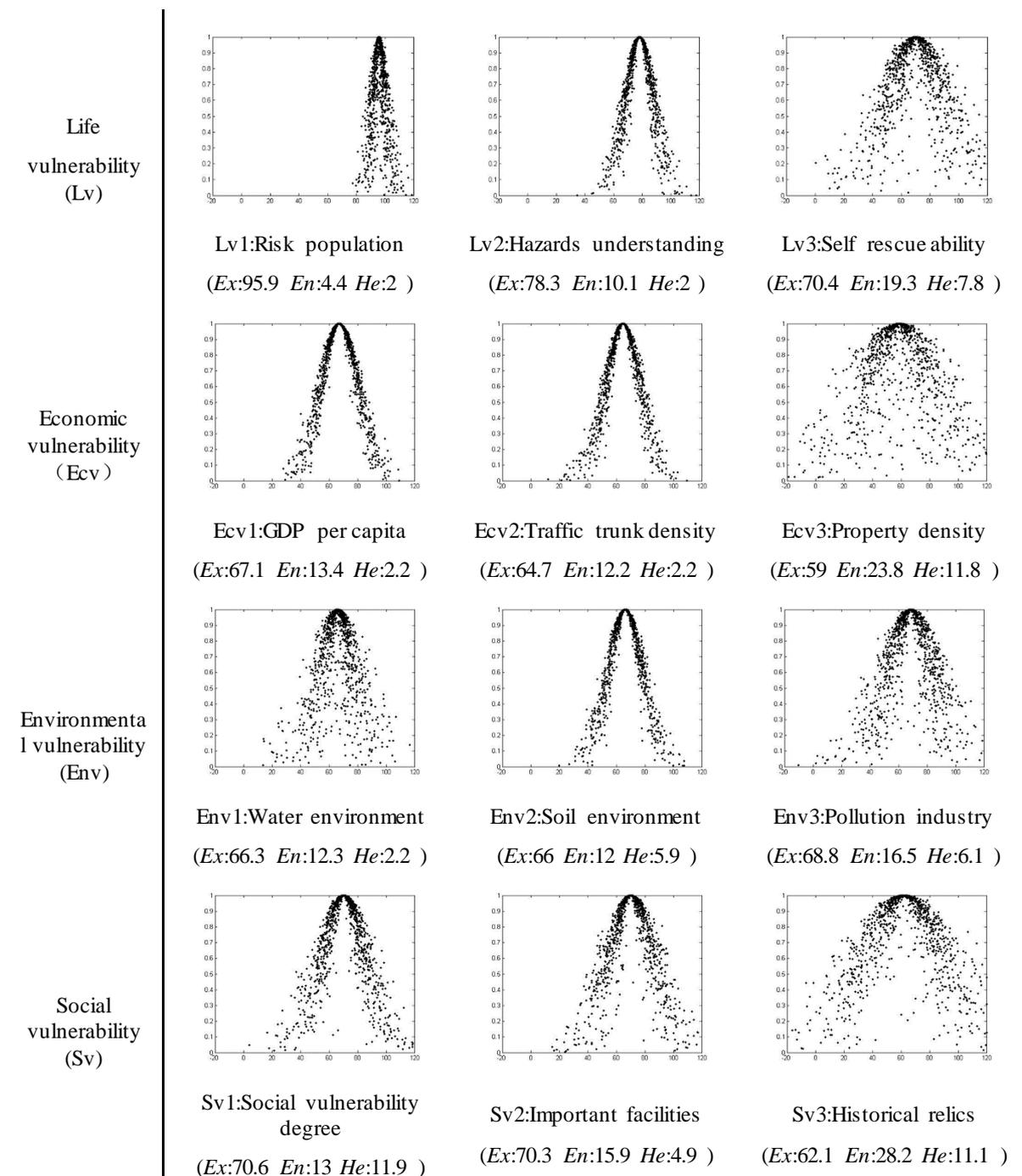


Figure 3. Sketch map of 20 indexes' membership cloud.

As shown in the Fig. 3, the centre vertex of the cloud is *Ex*, *En* represents the width of the cloud, and *He* represents the degree of dispersion of cloud distribution, that is, the thickness of cloud lines. For instance, the closer the *Ex* to the right side



of the axis, the higher the expert's score. The En of Lv3 is larger than that of Lv1, we can find the cloud is wider, and the He of E4 is larger than that of E2, so the distribution of the cloud is obviously “thicker” than E2. In a word, the membership cloud can obviously reflect the degree of divergence and randomness of expert opinions.

3.2 Weight calculation

5 After the result of the scoring is processed by the backward cloud generator according from the Eq. (1) to (3) and (5), the improved weight distribution result and result comparing the algebraic is shown in Table 1:

Table 1. Results comparison of the weight distribution.

	Indexes	ω - (SCM-improved model)	ω - (Algebraic method)
Weight	Risk population	0.093	0.069
	Flood intensity	0.069	0.064
	Alert time	0.061	0.057
	Hazards understanding	0.060	0.056
	Distance from the dam	0.053	0.054
	Storage capacity	0.053	0.056
	Social vulnerability degree	0.051	0.050
	Water environment	0.048	0.047
	Soil environment	0.048	0.047
	Important facilities	0.048	0.050
	GDP per capita	0.048	0.048
	Traffic trunk density	0.047	0.046
	Pollution industry	0.046	0.049
	Dam breaking time	0.046	0.043
	Self rescue ability	0.046	0.050
	Downstream slope formation	0.039	0.042
	Historical relics	0.037	0.044
Property density	0.037	0.042	
Dam height	0.036	0.043	
Sediment concentration	0.034	0.040	
Range		0.059	0.029
Multiple			2.044

3.3 Discussion

In order to verify the validity of the method, the results of the distribution contrast of the original and improved ones are
 10 drawn as Fig. 4.

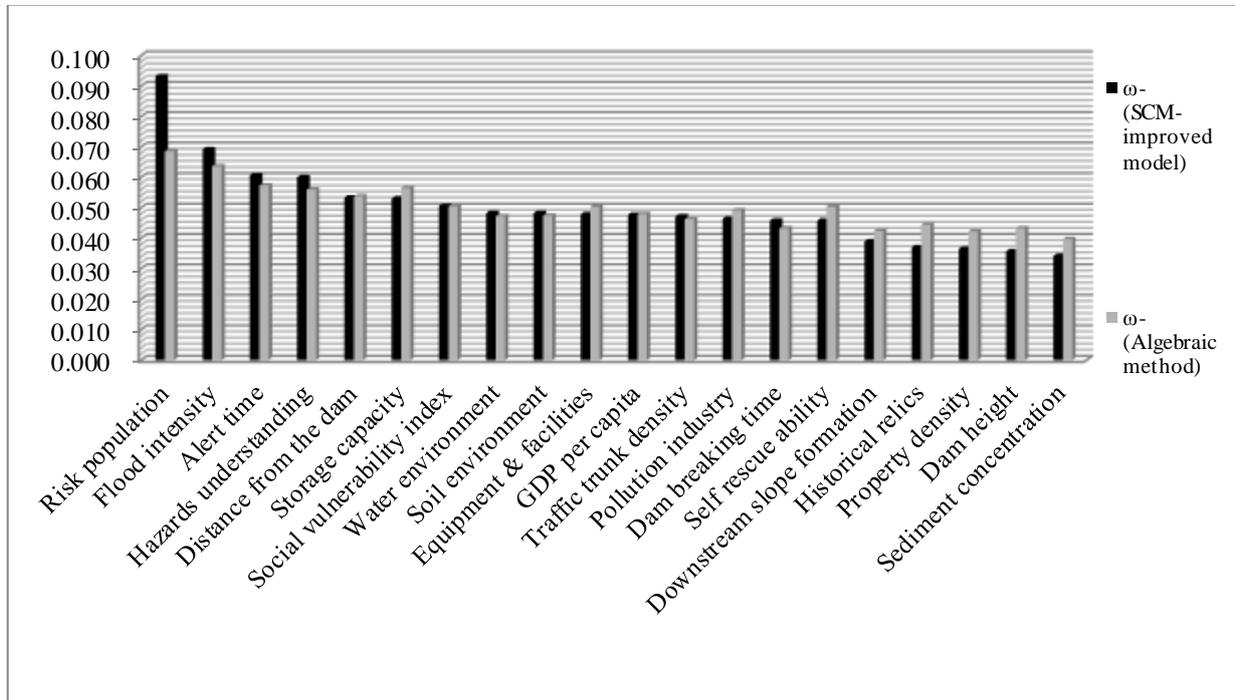


Figure 4. Weight comparing of 20 influencing factors.

According to the Fig. 3, Fig. 4 and Table 1, the analysis of figures shows following:

- (1) The top rankings have not changed after the adjustment, and still maintain the consistency of ranking. All of these top ranking factors are scored higher and the opinions are concentrated, which is in line with the objective situation. At the same time, the range increased by 2.04 times.
- (2) The distribution of weights is basically corresponding to the numerical value of *Ex*, reflecting the opinions of experts. At the same time, according to the adjustment of *En*, which reflects the difference of expert opinion, the weight of opinion unified index is further enlarged. Several factors are reduced the weight due to the large differences in opinions and the further reduction in adjusted weights. It reflects the validity of the entropy method in handling the weight distribution through the differences in opinions.

Thus, it can be seen that the SCM-improved entropy weight model is more in keeping with the general cognition of the people while ensuring the objective and fair data.

4 Conclusions

- 15 Dam break is a kind of low probability and high loss risk event with uncertainties. In this manuscript, risk factors are divided into hazards, exposure and vulnerability factors, and 20 factors are be selected as the main influencing factors of dam break risk consequences. We used SCM to improve the entropy method, based on the same ideas that these two methods are dealing with the divergence. The fuzziness index of the information is generated by the backward cloud generator, and then



applied to the improved formula of entropy weight calculation model. We establish the weight calculation model of influencing factors of dam break risk. The results indicate that (1) The result of weight calculation conforms to expert cognition, main factors' weight ranking is basically consistent with the one calculated by the traditional algebraic method. (2) Under the condition of 20 factors, the average problem of weight distribution is overcome, the difference between the maximum and the minimum is 2 times larger. (3) This model is easy to apply and can be improved by the accumulation of expert opinions. In a word, it is reasonable and feasible to apply this improved model to the weight analysis of dam break risk factors, providing a solid foundation for risk assessment and risk management theory.

Author contribution. Zongkun Li and Wei Ge provided the funding acquisition and supervision. Wei Li developed the model and wrote the original draft. Wei Li prepared the manuscript with contribution from all co-authors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgments. The work was supported by National Natural Science Foundation of China (Grant No. 51379192, 51679222, 51709239) and Key Project of Science and Technology Research of Education Department of Henan Province of China (Grant No. 18A570007).

References

- Colomer Mendoza, F. J. and Gallardo, I. A.: Design of a model to assess the environmental risk of leachate dams, *Waste Management*, 28, 2122-2133, 2008.
- Dong, Q., Ai, X., Cao, G., Zhang, Y., and Wang, X.: Study on risk assessment of water security of drought periods based on entropy weight methods, *Kybernetes*, 39, 864-870, 2010.
- Dutta, D., Herath, S., and Musiaka, K.: A mathematical model for flood loss estimation, *J. Hydrol.*, 277, 24-49, 2003.
- Daniell, J. E., Khazai, B., and Wenzel, F.: Chapter 6 – Indirect Loss Potential Index for Natural Disasters for National and Subnational Analysis, *Risk Modeling for Hazards & Disasters*, 139-173, 2018.
- Jonkman, S. N., Godfroy, M., Sebastian, A., and Kolen, B.: Brief communication: Post-event analysis of loss of life due to hurricane Harvey, 18, 1073-1078, 2018.
- Li, D. Y. and Du, Y.: *Artificial intelligence with uncertainty*, National Defence Industry Press, 15, 1583-1594, 2014.
- Li, D., Meng, H., and Shi, X.: Membership clouds and membership cloud generators, *Journal of Computer Research & Development*, 32, 16-21, 1995.
- Lee, J. S. and Noh, J. W.: The impacts of uncertainty in the predicted dam breach floods on economic damage estimation, *Ksce J. Civ. Eng.*, 7, 343-350, 2003.
- Liu, Y., Qin, Z. M., Hu, B. D., and Feng, S.: State fusion entropy for continuous and site-specific analysis of landslide



- stability changing regularities, *Nat. Hazards Earth Syst. Sci.*, 18, 1187-1199, 2018.
- Ling, H., Sun, Y. F., Wang, X. L., and An, J.: Environmental risk evaluation based on dam-break flood numerical simulation, Paper presented at the International Conference on Bioinformatics and Biomedical Engineering, 1-4, 2009.
- Mithas, S., Ramasubbu, N., and Sambamurthy, V.: How information management capability influences firm performance
5 , *Mis. Quart.*, 35, 237-256, 2011.
- Okada, N.: Urban Diagnosis and Integrated Disaster Risk Management, *Journal of Natural Disaster Science*, 26, 49-54, 2004
- Ouyang, S. and Shi, Y. L.: A new improved entropy method and its application in power quality evaluation, *Automation of Electric Power Systems*, 37, 156-159, 2013.
- Ragas, A. M. J., Huijbregts, M. A. J., Jong, H. D., and Leuven, R. S. E. W.: Uncertainty in environmental risk assessment:
10 implications for risk-based management of river basins, *Integr. Environ. Asses.*, 5, 27-37, 2010.
- Ren, Y., Yao, J., Xu, D., and Wang, J.: A comprehensive evaluation of regional water safety systems based on a similarity cloud model, *Water Sci. Technol.*, 76, 594, 2017.
- Smith, K. : *Environmental Hazards: Assessing Risk and Reducing Disaster* (New York: Routledge), 2013.
- Su, H. Z., Yang, M., and Wen, Z. P.: An approach using multi-factor combination to evaluate high rocky slope safety, *Nat.*
15 *Hazards Earth Syst. Sci.*, 16, 1449-1463, 2016.
- Wang, J. and Chen, X.: Application of Set Pair Analysis Method Based on Entropy Weight in the Air Quality Assessment, *Environ. Sci. Technol.*, 39, 177-180, 2016.
- Wagenaar, D. J., De Bruijn, K. M., Bouwer, L. M., and De Moel, H.: Uncertainty in flood damage estimates and its potential effect on investment decisions, *Nat. Hazards Earth Syst. Sci.*, 16, 1-14, 2016.
- 20 Wang, D., Liu, D., Ding, H., Singh, V. P., Wang, Y., Zeng, X., Wu, J., and Wang, L.: A cloud model-based approach for water quality assessment, *Environ. Res.*, 148, 24-35, 2016.
- Wu, S. W., Qin, P., Gao, J., and Yan, X. Q.: Application of Entropy Weight and Set Pair Analysis Theory to Risk Assessment of Dam Operation, *Journal of Yangtze River Scientific Research Institute*, 33, 36-40, 2016.
- Wang, M., Tang, D., Bai, Y., and Xia, Z.: A compound cloud model for harmoniousness assessment of water allocation,
25 *Environ. Earth Sci.*, 75, 977, 2016.
- Wisner, B. and Uitto, J.: *Life on the Edge: Urban social vulnerability and decentralized, Citizen-Based Disaster Risk Reduction in Four Large Cities of the Pacific Rim*: Springer Berlin Heidelberg, 2009.
- Wan, S. P., Xu, G. L., Wang, F., and Dong, J. Y.: A new method for Atanassov's interval-valued intuitionistic fuzzy MAGDM with incomplete attribute weight information, *Information Science*, 316, 329-347, 2015.
- 30 Yari, G. and Chaji, A. R.: Maximum Bayesian entropy method for determining ordered weighted averaging operator weights, *Comput. Ind. Eng.*, 63, 338-342, 2012.
- Yang, M. S. and Nataliani, Y.: A Feature-Reduction Fuzzy Clustering Algorithm Based on Feature-Weighted Entropy, *IEEE T. Fuzzy Syst.*, PP, 1, 2017.
- Zhou, Z., Wang, X., Sun, R., Ao, X., Sun, X., and Song, M.: Study of the comprehensive risk analysis of dam-break flooding



based on the numerical simulation of flood routing, Part II: Model application and results, *Nat. Hazards*, 72, 675-700, 2014.
Zou, Q., Zhou, J., Zhou, C., Song, L., and Guo, J.: Comprehensive flood risk assessment based on set pair analysis -variable fuzzy sets model and fuzzy AHP, *Stoch. Env. Res. Risk A.*, 27, 525-546, 2013.