

# A method for dam health diagnosis based on the improved attribute reduction

SHI Yu-qun, YANG Juan-juan & HAN Zi-long

(School of Water Resources and Hydropower Engineering, Wuhan University, 430072, Wuhan, Hubei)

**Keywords:** dam; health diagnosis; fuzzy rough set; attribute significance; golden section method

**Abstract.** Since the traditional rough set theory can easily result in information loss during attribute discretization and its attribute reduction is too complex, the fuzzy rough set theory and the golden section method are introduced for dam health diagnosis. With attribute fuzzification replacing attribute discretization, and attribute significance as a condition of attribute reduction, the dam health rough set diagnosis model is improved. Next, the improved dam health rough set diagnosis model is applied to a practical project. Results show that the improved attribute reduction put forward in this paper can more fully demonstrate factors influencing uncertainty of the dam health status. The diagnosis results, while more reasonably reflecting the dam's practical health status, can provide a new research path for dam health diagnosis.

## Introduction

Dam health diagnosis is an uncertain, complex issue involving multiple indexes and multiple layers of characteristics. In dam health diagnosis, the monitoring statistical model, the deterministic model and the hybrid model are the most widely-used. Targeting at the single measurement point, these models regard the relationship between the monitoring effect size and the environmental variable as a certain cause-and-effect relationship [1, 2]. However, modeling efforts combining multiple monitoring measurement points with multiple monitoring effect sizes are insufficient [3, 4]. Besides, few scholars take into account uncertainty of health diagnosis, which serves as a basis for dam health diagnosis. Uncertainty of dam monitoring information can result in an uncertain cognition of the dam health status. An in-depth research of uncertainty in dam health diagnosis is believed to be conducive to realization of intelligent dam health diagnosis. Thus, it has become a general research interest and a major difficulty in the field of dam health diagnosis [5].

The current dam monitoring models and the classical mathematical theories can hardly resolve uncertainty in dam health diagnosis. The rough set theory is a modern intelligent, informatized mathematical method devoted to studying the issue of uncertainty [6]. It can provide the theoretical basis for study of uncertainty in dam health diagnosis [7, 8].

Developed on the classification mechanism, the rough set theory deems classification to divide fuzzy subsets equivalent to each other but mutually disjoint. Under the prerequisite of maintaining the classification ability, the upper approximation set, lower approximation set and roughness are constructed to depict the issue of boundary uncertainty. The principle of attribute reduction is employed for index screening and attribute dimensionality reduction, which can effectively realize the mapping between the decision objective and the uncertainty.

The traditional rough set theory is oriented towards the discrete data. Nevertheless, discretization of continuous attributes might easily result in information loss. Meanwhile, the traditional attribute reduction algorithm can only eliminate redundant attributes one at a time. The computing process is especially complex. This paper, while applying the rough set theory to dam health diagnosis, introduces the fuzzy rough set theory and the golden section method to improve confirmation of the attribute decision table and the attribute reduction algorithm. In this way, during the diagnosis process, the original information can be more fully utilized for more effective description of factors

influencing uncertainty of the dam health status. Finally, the diagnosis results can more reasonably reflect the actual health status of the dam.

## Modeling

### Basic steps of dam health diagnosis based on the rough set theory

Below are basic steps of dam health diagnosis based on the rough set theory:

(1) Build the diagnosis index set,  $C$ : Select several proper monitoring effect sizes to build the diagnosis index set reflecting the dam health status— $C = \{c_1, c_2, \dots, c_m\}$ .

(2) Build the attribute decision table,  $S$ : Use the five-grade evaluation method,  $V = \{V_1, V_2, V_3, V_4, V_5\}$ , to build the attribute decision table,  $S = (U, C, A, D)$ , based on analysis results of the dam safety monitoring data. Where,  $U$  is the domain of discourse, namely the dam health status;  $C$  is the diagnosis index set;  $A$  is the condition attribute set of diagnosis indexes,  $A = \{a_1, a_2, \dots, a_m\}$ . Where, the diagnosis index,  $a_i$ , is corresponding to the degree of membership,  $\mu_{ij}$ , of the evaluation grade,  $V_j$ ;  $D$  is the decision attribute set,  $D = \{d\}$ .

(3) Attribute reduction: In the diagnosis index set,  $C$ , not all diagnosis indexes are necessities. Besides, significance of different diagnosis indexes to dam health diagnosis also varies. Attribute reduction is to eliminate indexes unnecessary to dam health diagnosis from the diagnosis index set based on the prerequisite of not changing the diagnosis capacity and based on the principle of equivalence in the rough set. Doing so can simplify complex issues, and alleviate negative influence of redundant information and uncertainty on dam health status diagnosis.

(4) Health diagnosis: The simplified diagnosis indexes in the attribute decision table,  $S$ , are adopted as the lower layer diagnosis indexes. Then, a proper recursion algorithm is employed to integrate the degree of membership,  $\mu_{ij}$ , representing the lower diagnosis index attributes, into the decision value representing the diagnosis object (the upper diagnosis index or the final diagnosis objective) attributes. In accordance with certain identification rule, the decision value and the evaluation grading set are put together for mapping comparison. Then, the diagnosis results of diagnosis objects can be obtained.

### Basic steps of attribute reduction improvement

**Introduction of the fuzzy rough set.** As mentioned above, the traditional rough set is oriented towards discrete data, so condition attributes and decision attributes generally adopt the clear, limited symbols (such as “0”, “1” and “2”) [8]. As to continuous attributes, discretization is necessary, but might lead to serious information loss because of insufficient utilization of information. Concerning the defect, some scholars have put forward the fuzzy rough set theory by combining the fuzzy set with the rough set [9]. According to the fuzzy rough set theory, attribute discretization is replaced by attribute fuzzification, and the degree of fuzzy membership is replaced by the clear limited symbols. Under the condition, condition attributes and decision attributes are all fuzzy, but the attribute value of objects in the fuzzy equivalence class is still continuous. This not only avoids information loss caused by attribute discretization, but also more accurately reflect the attribute status of diagnosis indexes.

This paper introduces the fuzzy rough set to dam health diagnosis, and uses the concept of attribute significance in the fuzzy rough set as a condition for attribute reduction. Assume that  $\mu_{pos}(x)$  is the degree of membership of the research object,  $x$ , in the domain of discourse, to the fuzzy positive domain, then significance of the decision attribute can be defined as:

$$Sig(A, d) = \frac{\sum_{x \in U} \mu_{pos}(x)}{|U|} \quad (1)$$

In Eq. (2),  $|U|$  denotes the number of research objects in the domain of discourse.

**Introduction of the golden section method.** The traditional attribute reduction algorithm can delete one redundant attribute at a time. Assume that there are  $m$  condition attributes. Under the extreme conditions, it is necessary to calculate the dependency of  $m(m+1)/2$  attribute combinations. Especially when the dependency of the decision attribute on multiple condition attribute combinations is to be calculated, it is even more complex, because the compound attribute dependency should be worked out. Besides, the more the compound attributes are, the higher the degree of computing complexity is. Thus, this paper introduces the golden section method [10], through which condition attributes with a higher degree of significance are selected out to delete many redundant attributes at a time. This can significantly improve the reduction efficiency and information utilization rate.

According to the idea of the golden section method, the left point of golden section,  $x_1=a+0.382(b-a)$ , and the right point of golden section,  $x_2=a+0.618(b-a)$ , are defined in accordance with the golden section proportion within the interval of current bracket,  $[a, b]$ . After comparison, if results of the left point,  $x_1$ , are proved to be more favorable, then set  $a=x_1$ , and delete the left region of the left point,  $x_1$ . Similarly, if results of the right point,  $x_2$ , are better, then set  $b=x_2$ , and delete the right region of the right point,  $x_2$ . The deletion process is repeated in the new region obtained. In this way, every computing can narrow down the search area by 0.382 or 0.618 until the optimal area is found. The algorithm features fast convergence and is suitable for the heuristic algorithm based on attribute significance.

## Case study

A concrete double-curvature arch dam with the highest dam height of 160m is selected as case study. In order to monitor the dam's safety status, facilities monitoring horizontal displacement, horizontal displacement, joint displacement, dam foundation uplift pressure, seepage flow, dam body stress, and local stress are set up to obtain adequate monitoring data. Based on the improved attribute reduction put forward in this paper, the health status of the dam is comprehensively diagnosed.

### Establishment of the index attribute decision-making table

According to the dam's monitoring layout, the index set for dam health diagnosis is built:

$$C = \{c_1, c_2, c_3, c_4, c_5, c_6, c_7\}$$

= {horizontal displacement, vertical displacement, joint displacement, dam foundation uplift pressure, seepage flow, dam body stress, local stress }

According to the five-grade evaluation method frequently used to evaluate the dam health status [5], the dam health status grade remark set is built:

$$V = \{V_1, V_2, V_3, V_4, V_5\}$$

= {normal, basic normal, slightly abnormal, seriously abnormal, malignantly abnormal}

$$= \{[1, 0.85], (0.85, 0.6], (0.6, 0.4], (0.4, 0.2], (0.2, 0]\} \quad (2)$$

Take the health status of the dam from 2008~2016 as the evaluation object,  $x_i$  ( $i=1,2,\dots,9$ ). Through preliminary analysis of the monitoring data of various effect sizes (diagnosis index set,  $C$ ), the degree of membership,  $\mu(c_i)$ , of the diagnosis index,  $c_i$ , to the health status is confirmed from the perspective of numerical value and tendency. After the degree of membership of various diagnosis indexes is preliminarily confirmed, the decision attribute,  $d$ , can also be confirmed. The decision attribute table (Table 1) is built.

Table 1. Dam safety evaluation index attribute decision-making table

Year	$x_1(2008)$	$x_2(2009)$	$x_3(2010)$	$x_4(2011)$	$x_5(2012)$	$x_6(2013)$	$x_7(2014)$	$x_8(2015)$	$x_9(2016)$
$\mu(c1)$	0.64	0.63	0.62	0.74	0.78	0.80	0.84	0.86	0.85
$\mu(c2)$	0.73	0.70	0.74	0.80	0.76	0.82	0.79	0.83	0.85
$\mu(c3)$	0.65	0.63	0.69	0.73	0.72	0.75	0.80	0.82	0.85
$\mu(c4)$	0.70	0.72	0.70	0.76	0.81	0.80	0.82	0.80	0.77
$\mu(c5)$	0.64	0.60	0.70	0.69	0.74	0.77	0.79	0.83	0.82
$\mu(c6)$	0.73	0.75	0.78	0.79	0.83	0.85	0.81	0.76	0.77
$\mu(c7)$	0.60	0.65	0.67	0.76	0.81	0.76	0.74	0.71	0.68
$d$	BN	BN	BN	BN	BN	BN	BN	BN	BN

Note: BN=Basic Normal.

### Attribute reduction

#### (1) Attribute fuzzification

Based on Eq. (2), the triangular membership function in fuzzy mathematics is adopted for fuzzification of the continuous attributes in the decision table. Table 2 shows the attribute fuzzification results of the diagnosis index,  $c_1$ .

Table 2. Dam health diagnosis index ( $c_1$ ) attribute fuzzification decision table

	Year	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$x_9$
$c_1$	$V_{11}$	0	0	0	0	0	0	0.20	0.30	0.25
	$V_{12}$	0.10	0.15	0.20	0.70	0.90	1	0.80	0.70	0.75
	$V_{13}$	0.90	0.85	0.80	0.30	0.10	0	0	0	0
	$V_{14}$	0	0	0	0	0	0	0	0	0
	$V_{15}$	0	0	0	0	0	0	0	0	0

#### (2) Attribute reduction based on attribute significance

Table 2 represents fuzzification results of various diagnosis index attributes. According to the rough set theory, the fuzzy positive domain membership degree of the research object to the condition attribute is  $\mu_{pos}(x)$ . Then, significance of decision attribute,  $d$ , to a single condition attribute can be given by Eq. (1),  $Sig(a_i, d)$  ( $i=1, 2, \dots, 7$ ), which is 0.267, 0.217, 0.261, 0.211, 0.300, 0.172 and 0.328, respectively.

Then, the diagnosis index of attribute significance is ranked in a descending order in the light of attribute significance, and the set,  $P = \{c_7, c_5, c_1, c_3, c_2, c_4, c_6\}$ , can be obtained. After that, attribute reduction based on the golden section method is adopted to maintain the 0.618m condition attributes, namely  $c_7, c_5, c_1$  and  $c_3$ . 0.382m redundant condition attributes, namely  $c_2, c_4$  and  $c_6$ , are deleted to obtain the set of critical diagnosis indexes,  $C' = \{c_7, c_5, c_1, c_3\} = \{\text{local stress, seepage flow, horizontal displacement, joint displacement}\}$ .  $C'$  is actually a subset of  $C$ . Horizontal level is a comprehensive reflection of the dam health status. The joint displacement reflects wholeness of the dam structure. The seepage flow indicates the working status of the dam foundation vadose. Local stress is an implication of the dam's structural stress state. After attribute reduction, the diagnosis indexes coincide with the major factors influencing health diagnosis of the dam in terms of physical significance. This also suggests feasibility of the improved attribute reduction to be applied to dam health diagnosis.

### Comprehensive diagnosis

After the critical diagnosis indexes are confirmed using the attribute reduction method, the weight recursion algorithm can be employed to obtain the membership degree,  $\mu_V(x_i)$ , of the dam

health status of different years to the grade remark set based on the attribute membership degree of various critical diagnosis indexes shown in Table 1. The diagnosis results are shown in Table 3. According to significance of diagnosis indexes, including horizontal displacement ( $c_1$ ), joint displacement ( $c_3$ ), seepage flow ( $c_5$ ) and local stress ( $c_7$ ), to dam health diagnosis, the weight-- $\omega(c_1, c_3, c_5, c_7) = \{0.30, 0.20, 0.30, 0.20\}$ —can be determined.

Table 3. Health diagnosis results of the dam in different years

Year	$x_1(2008)$	$x_2(2009)$	$x_3(2010)$	$x_4(2011)$	$x_5(2012)$	$x_6(2013)$	$x_7(2014)$	$x_8(2015)$	$x_9(2016)$
$\mu_V(x_i)$	0.634	0.625	0.668	0.727	0.762	0.773	0.797	0.813	0.807
Diagnosis result	BN	BN	BN	BN	BN	BN	BN	BN	BN

Note: BN=Basic Normal.

Based on Table 3 and Eq. (2), it can be observed that the membership degree,  $\mu_V(x_i)$ , of the dam is in the section of (0.85,0.60). The health status is mostly “BN” However, in 2008~2009, the membership degree,  $\mu_V(x_i)$ , is close to 0.60, and its health status belongs to “BN” or rather “slightly abnormal”. In 2015~2016, the membership degree,  $\mu_V(x_i)$ , is close to 0.85, and the health status belongs to “BN” or rather “normal”. This is because the dam was further reinforced concerning safety defects existing in the initial operation period of the dam. Since then, the dam health status has been greatly improved. This indicates not only favorable effects of dam reinforcement, but also feasibility of the health diagnosis method proposed in this paper.

## Conclusions

Dam health diagnosis is full of uncertainty, while the rough set theory can provide theoretical basis to resolve uncertainty of the dam health diagnosis. However, the traditional rough set theory can easily result in information loss during the discretization process of continuous attribute, and is complex in attribute reduction. Thus, this paper introduces the fuzzy rough set theory to dam health diagnosis. By substituting attribute fuzzification with attribute discretization, fuzzy membership degree with clear limited symbols, the confirmation method of the attribute decision table is thus improved. Meanwhile, the golden section method is introduced. With membership significance as a condition for attribute reduction, condition attributes with a high attribute significance are selected to realize the purpose of deleting multiple redundant attributes at a time. This helps improve the original attribute reduction. During the diagnosis process, the improved attribute reduction method put forward by this paper can make more full use of the original information and more effectively describe factors influencing uncertainty of the dam health status. In terms of diagnosis results, the actual health status of the dam is more reasonably reflected. To sum up, this paper offers a new research approach for dam health diagnosis.

## Acknowledgments

This research is supported by the National Natural Science Foundation of China (51379162).

## References

- [1] Tonini D. Observed behavior of several leakier arch dams. Proc. ASCE, Journal of the Power Division, Vol.82, December, 1956, 135-139.
- [2] ICOLD. Methods of analysis for the prediction and the verification of dam behaviour. Swiss Committee on Dams, 2003.

- [3] Ching-Yun Kao, Chin-Hsiung Loh. Monitoring of long-term static deformation data of Fei-Tsui arch dam using artificial neural network-based approaches. *Structural Control and Health Monitoring*. 2013, 20:282-303.
- [4] Jiang Zhenxiang, He Jinping. Method of fusion diagnosis for dam service status based on joint distribution function of multiple points. *Mathematical Problems in Engineering*. 2016. DOI: 10.1155/2016/9049260.
- [5] He Jinping. Advances in comprehensive diagnostic method of dam healthy status. *Hydropower and New Energy*, 2016, (1):1-6.
- [6] Pawlak Z. Rough sets [J].*International Journal of Computer and Information Science*, 1982, 11(5):341-356.
- [7] Lei Peng, Wu Zhongru, Gu Chongshi. A study on rough set model of dam safety monitoring. *Engineering Journal of Wuhan University*, 2005, 38 (3):45-49.
- [8] Xu Kunzhen, He Jinping, XIA Wanqiu. A fusion evaluation model of dam safety based on rough set and conditional information entropy. *Engineering Journal of Wuhan University*, 2014, 47 (1): 46-49.
- [9] Dubois D, Prade H. Rough sets and fuzzy rough sets. *International Journal of General Systems*. 1990, 17: 191-209.
- [10] J. Alikhani Koupaei, S.M.M. Hosseini, F.M. Maalek Ghaini. A new optimization Algorithm Based on Chaotic Maps and Golden Section Search Method. *Engineering Applications of Artificial Intelligence*, 2016, (4):201-214.