# Multi-source data fusion method for structural safety assessment of water diversion structures

Sherong Zhang, Ting Liu and Chao Wang

## ABSTRACT

Building safety assessment based on single sensor data has the problems of low reliability and high uncertainty. Therefore, this paper proposes a novel multi-source sensor data fusion method based on Improved Dempster–Shafer (D-S) evidence theory and Back Propagation Neural Network (BPNN). Before data fusion, the improved self-support function is adopted to preprocess the original data. The process of data fusion is divided into three steps: Firstly, the feature of the same kind of sensor data is extracted by the adaptive weighted average method as the input source of BPNN. Then, BPNN is trained and its output is used as the basic probability assignment (BPA) of D-S evidence theory. Finally, Bhattacharyya Distance (BD) is introduced to improve D-S evidence theory from two aspects of evidence distance and conflict factors, and multi-source data fusion is realized by D-S synthesis rules. In practical application, a three-level information fusion framework of the data level, the feature level, and the decision level is proposed, and the safety status of buildings is evaluated by using multi-source sensor data. The results show that compared with the fusion result of the traditional D-S evidence theory, the algorithm improves the accuracy of the overall safety state assessment of the building and reduces the MSE from 0.18 to 0.01%.

Key words | BP neural network, D-S evidence theory, multi-source data fusion, safety evaluation, structural safety, water diversion project

Sherong Zhang
Ting Liu (corresponding author)
Chao Wang
State Key Laboratory of Hydraulic Engineering
  Simulation and Safety,
Tianjin University,
135 Yaguan Road, Tianjin 300350,
China
E-mail: lting_0109@tju.edu.cn

## HIGHLIGHTS

- A new method is proposed to evaluate the safety status of water diversion structures by fusing multi-source heterogeneous sensor data.
- A multi-sensor hierarchical data fusion model suitable for the structural characteristics of the water diversion project is established.
- The classical D-S evidence theory is improved and combined with BPNN to reduce the uncertainty of sensor data.

## GRAPHICAL ABSTRACT
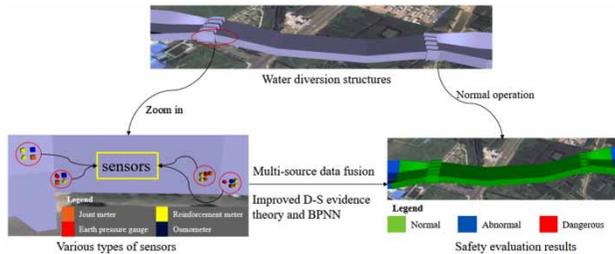


## INTRODUCTION

Inter-basin water diversion project is an effective measure to control the uneven distribution of water resources and solve the contradiction between supply and demand for water resources (Valipour 2016, 2017; Bazrkar *et al.* 2017). According to statistics, more than 350 water diversion projects have been built in more than 40 countries around the world (Jia 2016), which has made great contributions to the economy, public security, and ecological benefits. There are many types of water diversion structures, and many safety risk factors are inevitably exposed during operations (Mehta *et al.* 2020). Especially in the case of aging, such as building settlements, cracks, and leakage (Samadi *et al.* 2014; Mehta & Yadav 2020). The various structures in the water diversion project belong to the series structure, and the impact of any link problem is far-reaching and huge. Therefore, how to reasonably evaluate the overall safety of water diversion structures and discover abnormalities in time is particularly important.

The long-term real-time monitoring of water diversion project using sensors is an important means to ensure the safe operation of buildings. Although the data acquisition and analysis platform for main structures of water diversion project has been established at this stage, which can monitor the safety status of buildings in real-time (Jiang *et al.* 2013; Xiao *et al.* 2019). However, there are still some limitations in the overall safety evaluation of buildings. First of all, the analysis of the monitoring data of the diversion building is mainly to establish a single measuring point mathematical model to determine the local structural state of the diversion building, but there are obvious limitations in reflecting the overall structural state of the building. Second, due to the existence of

various noises and abnormal values in the monitoring data, it is usually necessary to rely on expert experience for identification and processing, which greatly reduces the efficiency of analysis. Third, in actual situations, the occurrence of safety accidents in water diversion projects is often the result of the joint action of various types of monitoring indicators (such as stress, strain, displacement, and seepage pressure). The current analysis method does not consider the relationship between the data, so it is difficult to obtain accurate information reflecting the overall safety status of the building.

Aiming at the problem of multi-source data security evaluation of hydropower projects, scholars have proposed the dam safety performance fusion evaluation model and carried out typical engineering applications. Su *et al.* (2018) combined the Dempster–Shafer (D-S) evidence theory with a set pair theory, integrated multi-source spatiotemporal information of dam safety, and identified and evaluated the structural behavior of dams. Jiang & He (2016) proposed a multi-point fusion evaluation method for the overall dam service status based on the joint distribution function. Liu *et al.* (2012) proposed a comprehensive analysis method of high dam prototype monitoring data based on multi-source information fusion. This method can effectively process a large amount of monitoring data from multiple monitoring points, and output the comprehensive evaluation results of dams at multiple points in real-time. However, there are few studies on the application of information fusion theory in the safety evaluation of diversion structures.

To address the above problems, this study uses information fusion technology to fuse multi-source heterogeneous sensor data and reveals the internal relationship between the

overall performance and local characteristics of water diversion structures. According to the characteristics of water diversion structures and based on the hierarchical theory of data fusion, a multi-sensor hierarchical data fusion model is proposed. Through the establishment of the data-level, feature-level and decision-level data fusion models, the safety status of the water diversion project is evaluated.

There are three innovations in this study. First, a new method is proposed to evaluate the safety status of water diversion structures by fusing multi-source heterogeneous sensor data. Second, a three-layer data fusion evaluation model is established according to the structural characteristics of diversion structures, including data-level fusion, feature-level fusion, and decision-level fusion. Third, the classical D-S evidence theory is improved and combined with Back Propagation Neural Network (BPNN) to reduce the uncertainty of sensor data and improve the accuracy and efficiency of safety evaluation. Besides, the evaluation model can also be extended to the pipeline, bridge, and other similar long-line structure buildings, and the novel method can be extended to other engineering fields for fusion diagnosis. In practical application, safety inspection is also an important means to ensure the operation of the project. A large amount of unstructured text data will be generated during the inspection process of the water diversion project. This part of the data has not been considered in this study.

## LITERATURE REVIEW

### Data fusion technology

Data fusion technology refers to the comprehensive processing of data from different information sources through various effective methods to obtain accurate and reliable reasoning decisions. Data fusion technology has been widely used in various disciplines, including but not limited to structural health monitoring (Liu *et al.* 2020; Wu & Jahanshahi 2020; Zhu *et al.* 2020), environmental monitoring (Long *et al.* 2020; Yang *et al.* 2020), mechanical fault diagnosis (Azamfar *et al.* 2020; Huang *et al.* 2020; Zhang & Deng 2020), and aerospace (Osegueda *et al.* 2003; Brierley *et al.* 2014).

Generally speaking, data fusion can achieve different fusion levels according to different fusion objectives

(Castanedo 2013). Data fusion is usually divided into three levels: data level, feature level, and decision level. The data level is used for the integration of similar sensor data, the feature level is used for the integration of heterogeneous sensor data, and the decision level l obtains the final evaluation result through multi-source data fusion. Existing data fusion algorithms can be divided into three categories: statistical methods include weighted average, Kalman filter, Bayesian estimation, D-S evidence theory, etc.; information theory methods include support function, cluster analysis, and entropy theory; artificial intelligence methods including artificial neural network, fuzzy set theory, and expert system. The choice of different fusion methods is the main problem of building safety evaluation using data fusion technology. How to choose the optimal fusion method for each safety assessment task is a challenging problem. For this reason, this paper designs a set of safety evaluation system and multi-source data fusion method suitable for water transfer project.

### D-S evidence theory and BPNN

D-S evidence theory, as a classic data fusion algorithm with the ability to deal with uncertain information, has strong engineering practicability (Yue *et al.* 2010; Zhao *et al.* 2020). However, in the application of classical D-S evidence theory, the basic probability assignment (BPA) of the key parameter is often obtained through empirical formulas or statistical methods, which are subjective and leads to low credibility of the results (Guan *et al.* 2008). Therefore, this paper adopts BPNN to obtain BPA. BPNN has strong nonlinear mapping ability, good fault tolerance, and robustness and is widely used in the fusion of multi-source heterogeneous data (Zhang *et al.* 2019; Wang 2020). This paper trains the characteristic data collected by various sensors in the water diversion structure to obtain the BPA value. To improve the dynamic applicability in the process of data fusion and eliminate the interference of effective abnormal monitoring data, the adaptive weighted average algorithm is introduced to calculate the input data of BPNN. Previous studies have shown that the adaptive weighted average method (AWAM) is superior to the traditional weighted average method in improving the fusion accuracy (Bin *et al.* 2011; Ren *et al.* 2012).

The classical D-S evidence theory may produce unreasonable results when synthesizing high conflict evidence. To solve this problem, scholars have proposed various solutions (Deng *et al.* 2017; Deng & Wang 2020). Murphy (2000) proposed to perform arithmetic averaging on the initial evidence set, and then combine it with D-S evidence theory. This method can effectively fuse high conflict evidence but ignores the correlation between the evidence. The Bhattacharya distance (BD) is a random measurement that considers the probability distribution between two samples (Bi *et al.* 2019). This paper introduces BD to measure the distance between different evidences output by BPNN to resolve conflicts between evidences and improve the reliability of evaluation results. In complex situations, existing studies have shown that hybrid data fusion methods are often superior to single fusion methods (Xie & Guan 2008; Gong 2009; Wu *et al.* 2018).

Therefore, this paper proposes a hybrid data fusion algorithm based on Improved D-S theory and BPNN for multi-sensor data fusion of water diversion projects. First, the AWAM is used to extract the characteristic values of the same type of monitoring data from multiple measurement points, and the results are used as the input data of BPNN. Then, the output data trained by BPNN is used as the BPA of improved D-S. Finally, the improved D-S theory is applied to the overall safety evaluation of the building.

## METHODOLOGY

### Research framework

As shown in Figure 1, the research framework of this paper includes three major parts: data collection and evaluation index acquisition, the process of safety evaluation, and addressing a case study. First, the sensor data of reinforcement gauge, joint meter, osmometer, and earth pressure gauge are collected. According to the type of sensor data, the evaluation index of the building is determined, namely stress, displacement, seepage pressure, and earth pressure. Second, combining the Improved D-S method with BPNN, a novel method of building structure safety evaluation based on multi-source data fusion technology is proposed. Before that, the improved support function is introduced to preprocess
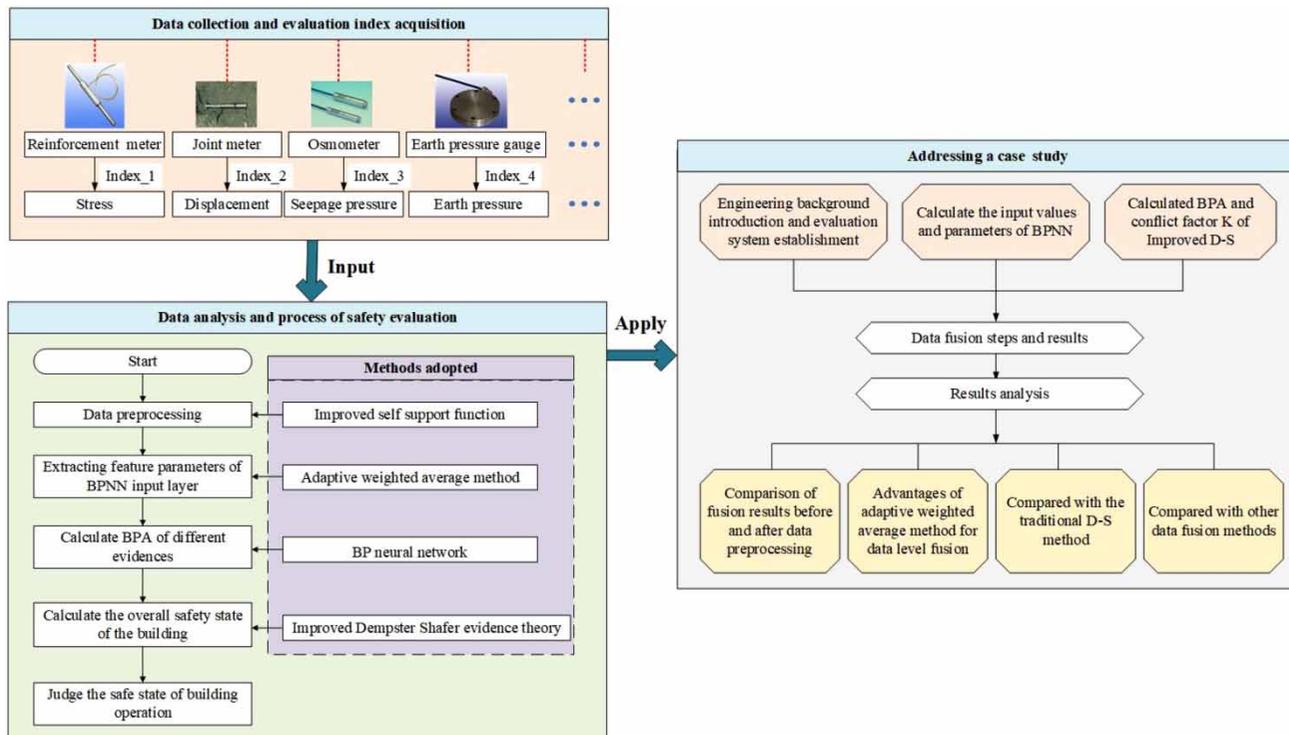


**Figure 1** │ Research framework.

the original data. Use AWAM to extract the feature values of preprocessed data as the input data of BPNN. Finally, the method is applied to the structural safety evaluation of an inverted siphon building in the Henan section of the middle route of the south to north water diversion project, and the evaluation results were analyzed.

## Data preprocessing

### Improved algorithm of support function with self-support

Due to the uncertainty of the environment and sensor failure, the collected data may be invalid or abnormal. The fusion results with invalid abnormal data cannot reflect the real safety status of buildings, so it is necessary to preprocess the invalid abnormal data. The data preprocessing method based on support function can identify abnormal data with large errors and improve the accuracy and reliability of data fusion (Yager 2001). Considering that the sensor collects data many times during the operation of the building, the reliability of the data can be evaluated by measuring the data consistency at each point in the collection interval. Therefore, this paper introduces a self-support degree function to improve the algorithm. The improved algorithm considers the credibility of different sensor data at the same time and the credibility of the data collected from the same measuring point in the entire observation interval. The improved algorithm is used to preprocess the original data, which can improve the accuracy of subsequent fusion evaluation.

Unknown factors such as geographical location and environmental conditions of water diversion projects will affect the collection of monitoring data, making the monitoring data mixed with uncertain or even wrong information. To improve the accuracy of subsequent building safety assessment, it is necessary to find the relationship between different data to determine the reliability of the data. The exponential decay support function has good reliability and high precision in data fusion. The support degree of calculation avoids the absoluteness of non-zero or non-one. Therefore, this paper introduces and improves the exponential decay support function to judge the validity of the data. Assuming that the data of the measuring points $i$, $j$ at time $t$ are $a_i(t)$ and $a_j(t)$, $i$, $j =$ 1, 2, 3, …, $m$. The calculation equation of the exponential decay support function is as follows (Shi *et al.* 2018):

$$d_{ij} = \exp\left(-\beta \times (a_i(t) - a_j(t))^2\right) \qquad (1)$$

where the parameter $\beta$ is the support attenuation factor, which is usually artificially set to 1. For a given $|a_i(t) - a_j(t)|$, the attenuation amplitude of support is changed by adjusting the size of $\beta$.

Considering the reliability of the data collected from the same measuring point for many times. To reduce the subjective error caused by setting the value of parameter $\beta$ artificially. In this study, the close degree of the data collected by the measuring point $i$ in this part is named self-support degree $d_i(t)$. The square root $\sqrt{d_i(t)d_j(t)}$ of the self-support degree of the measuring point $i$, $j$ is taken as the attenuation factor of the support degree. Then, the calculation equation of self-support is shown in Equation (2) and the improved support function is shown in Equation (3).

$$d_i(t) = \frac{1}{1 + |a_i(t) - \bar{a}_i|} \qquad (2)$$

where $\bar{a}_i$ represents the average value of $k$ data collected by measuring point $i$.

$$d_{ij}(t) = \exp\left(-\sqrt{d_i(t)d_j(t)} \times (a_i(t) - a_j(t))^2\right) \qquad (3)$$

The improved support function $d_{ij}(k)$ depends not only on the support degree of different measuring points at the same time but also depends on the self-support degree of measurement points. In this way, the influence of the monitoring data with large errors caused by the instrument itself and the environment on the fusion value is reduced, and the validity of the data to be fused is improved.

### Identify abnormal data

To identify abnormal data, the support degree of each monitoring data is calculated by improved support function, and

the support matrix is constructed as follows:

$$\mathbf{Z} = \begin{bmatrix} d_{11}(t) & d_{12}(t) & \cdots & d_{1m}(t) \\ d_{21}(t) & d_{22}(t) & \cdots & d_{2m}(t) \\ \cdots & \cdots & \cdots & \cdots \\ d_{m1}(t) & d_{m2}(t) & \cdots & d_{mm}(t) \end{bmatrix} \tag{4}$$

At time $t$, the consistency measurement $\gamma_i(t)$ of support degree between the data of measuring point $i$ and the data of other measuring points are shown as follows:

$$\gamma_i(t) = \frac{\sum\limits_{j=1, j \neq i}^{m} d_{ij}(t)}{m - 1} \tag{5}$$

where, $0 < \gamma_i(t) \leq 1$, the larger the value of $\gamma_i(t)$ is, the closer the monitoring data of measuring point $i$ to the monitoring data of other measuring points at time $t$. On the contrary, the monitoring data is likely to be abnormal data and should be eliminated. Based on this, the monitoring data $S_i(t)$ of all kinds of sensors in this part after preprocessing can be obtained at time $t$.

## BPA calculation based on AWAM and BPNN

### AWAM calculates the input data of BPNN

After eliminating the invalid abnormal data, the AWAM is used to fuse the similar monitoring data of various parts of the building to provide input data for BPNN calculation. Different monitoring data have different weights in the safety evaluation of building components. The AWAM based on the minimum mean square error theory is used to solve the weight of each sensor. Multiply the data received by each sensor by the corresponding weight, and add the result, which is the input value of BPNN.

Assuming that there are $m$ stress measurement points in a certain part of the building, and the variance are $\sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2$. Since the sensors are installed in different locations and have a certain distance, it is approximately considered that the monitoring data are independent of each other. Accordingly, the mean square error $\sigma$ satisfies Equation (6) (Haq 2020). The fusion of $\hat{X}$ and weight should meet Equation (7), where the effective stress monitoring data is $x_i$

and the corresponding weight is $w_i$. When the mean square error is minimum, the corresponding weight of $m$ stress sensors is shown in Equation (8). In the same way, the fusion values of other monitoring indicators in this part after eliminating abnormal values are obtained.

$$\sigma^2 = E[(x - \overline{X})^2] = E\left[\sum_{i=1}^{m} w_i^2 (x - x_i)^2\right] = \sum_{i=1}^{m} w_i^2 \sigma_i^2 \tag{6}$$

$$\hat{X} = \sum_{i=1}^{m} w_i x_i, \quad \sum_{i=1}^{m} w_i = 1 \tag{7}$$

$$w_i = \frac{1}{\sigma_i^2 \sum\limits_{i=1}^{m} \frac{1}{\sigma_i^2}} \tag{8}$$

### BPNN-enabled BPA

The neural network is a typical model to construct nonlinear complex relationships (Samadi *et al.* 2015). BPNN is currently the most widely used neural network model, and its learning process consists of two parts: forward propagation and back propagation. In the forward propagation process, the input pattern is passed from the input layer to the output layer through the processing of hidden layer neurons. If the desired output cannot be obtained in the output layer, error back propagation is performed. At this time, the error signal propagates from the output layer to the input layer, and the connection weights and thresholds of each layer are adjusted along the way, so that the error is continuously reduced until the accuracy requirements are met. The algorithm uses the gradient descent method to make the weight converge to the minimum point at the fastest speed through repeated training of multiple samples, and find the minimum value of the error function.

There is a highly nonlinear relationship between the operational safety state of buildings and multi-sensor data. According to the requirements of building safety evaluation, the BP network structure designed in this paper is shown in Figure 2. It is a three-layer BPNN with a hidden layer.

Where the input vector $\mathbf{I} = [I_1, I_2, \ldots, I_n]$, the number of input layers $n$ is determined by the type number of building safety monitoring indicators. The expected output vector $\mathbf{U} = [U_1, U_2, \ldots, U_j]$, the actual output vector
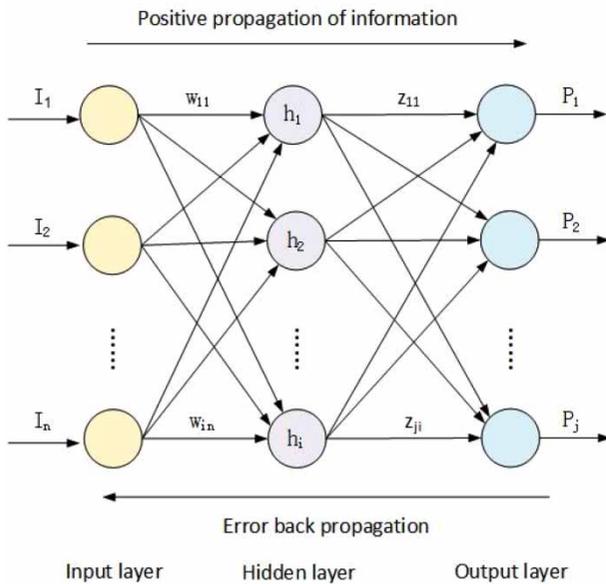
**Figure 2** │ Structure of BPNN.

$\mathbf{P} = [P_1, P_2, \ldots, P_j]$, the number of output layers $j$ is determined by the number of building safety evaluation level. The connection weight between the input layer and hidden layer is $w_{in}$, and that between the hidden layer and the output layer is $z_{ji}$. The threshold of each neuron in the hidden layer is $b_n$, and that of each unit in the output layer is $b_i{}'$. The number of neurons in the hidden layer $i$ is determined by the empirical Equation (9).

$$i = \sqrt{n + j} + a \qquad (9)$$

where $a$ is the natural number in [1, 10].

In this study, the BPNN is used to locally integrate the heterogeneous monitoring data of various parts of the building, and to initially judge the safety status of each part of the building. The fusion value of each monitoring index obtained by adaptive weighted average fusion is used as the feature parameter input of the BPNN. The unipolar sigmoid function is used in the output layer. The steps of neural network fusion are as follows:

1. BPNN initialization. Initial values are given to the initialization variables, including the initial random values $w_{in}$ and $z_{ji}$, the thresholds $b_i$, and $b_j{}'$.
2. The result of AWAM is used as the sample set to train the neural network. The activation function uses the Sigmoid

function, and the calculation equation is as follows (11), which maps the input range of neurons from $(-\infty, +\infty)$ to (0,1).

$$f(x) = \frac{1}{1 + e^{-x}} \qquad (10)$$

3. Calculate the input and output of each neuron in the hidden layer and the output layer.
4. Reverse calculation of the unit errors of the output layer and the hidden layer. The calculation equation for the neuron error of the $k$th output layer is shown in Equation (16).

$$\delta_k = (c_k - h_k)h_k(1 - h_k) \qquad (11)$$

where $c_k$ is the expected value of the sample.

5. Update the connection weight $z_{ji}$ (Equation (17)) and the output threshold $b_i{}'$ (Equation (18)) between the hidden layer and the output layer according to the weight error $\delta_k$.

$$z_{ji}(N + 1) = w_{ji}(N) + \alpha\delta_k h_i \qquad (12)$$

$$b_i'(N + 1) = b_i'(N) + \beta\delta_k \qquad (13)$$

6. Update the learning input mode and input times, repeat steps (2)–(8) until the error and learning times meet the specified requirements.

The mean square error *MSE* (Equation (14)) of the test sample is used to measure the quality of the network performance. The smaller the error value, the better the BP network fusion result.

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}\left(y_r^{(i)} - y_p^{(i)}\right)^2 \qquad (14)$$

where $N$ is the number of training samples; $y_r^{(i)}$ is the actual output value of the $i$th sample in the test set; $y_p^{(i)}$ is the output value of the BPNN after the $i$th sample in the test set has passed the simulation.

## Introducing BD to improve conflict factor

In D-S theory, the hypothesis recognition framework consists of $n$ independent and complete propositions. Satisfying

$2^\Theta \rightarrow [0, 1]$, $m(\varnothing) = 0$, and $\sum_{i=1}^n m(A_i) = 1$, $m(A)$ reflects the trust degree of evidence to proposition $A$, which is called the BPA. If $m(A) \neq 0$, $A$ is called a focal element. The BPA of $n$ evidence in the framework of recognition is $m_1, m_2, \ldots, m_n$, then the synthesis rule is as follows (Li *et al.* 2020):

$$(m_1 \oplus m_2 \oplus \cdots \oplus m_i)(A) = \frac{\sum_{\bigcap_{i=1}^n A_i = A} \prod_{i=1}^n m_i(A_i)}{1 - k} \quad (15)$$

$$k = \sum_{\bigcap_{i=1}^n A_i = \varnothing} \prod_{i=1}^n m_i(A_i) \quad (16)$$

where $k$ is the conflict factor, and $i$ is the number of focal elements in the recognition framework $\Theta$.

If the value of $k$ is large, it indicates that the conflict between the evidence is large. This may cause the fusion result to be inconsistent with the actual situation, leading to decision-making errors. In this study, the source of evidence is improved by introducing BD, and the high conflict evidence is corrected and then combined iteratively using synthesis rules to improve the accuracy of the fusion results.

In statistics, the Bhattacharyya Distance (BD) is used to measure two discrete probability distributions. For two discrete probability distributions $p$ and $q$ in the $X$-number field, the BD is defined as follows:

$$BD(p, q) = -\ln(BC(p, q)) \quad (17)$$

where $BC(p, q) = \sum \sqrt{p(x)q(x)}$, $0 \leq BC \leq 1$ and $0 \leq BD \leq \infty$, BC is called the Bhattacharyya coefficient. According to the BD, the distance equation of $m_i$ and $m_j$ and the distance matrix can be derived as follows:

$$d_{BPA}(m_i, m_j) = d_{ij} = -\ln(BC(m_i, m_j)) \quad (18)$$

$$\mathbf{DM} = \begin{bmatrix} 0 & d_{12} & \cdots & d_{1n} \\ d_{21} & 0 & \cdots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \cdots & 0 \end{bmatrix} \quad (19)$$

It can be seen from Equation (24) that the matrix is a symmetric matrix, which is $d_{ij} = d_{ji}$, and the diagonal elements are zero.

The monitoring data of different parts of the building have a certain impact on the overall safety status of the building, so there is a certain connection between the evidence of each part layer. Although the conflict factor $k$ reflects the magnitude of conflict between focal elements, it does not consider the relationship between evidence. BD reflects the similarity of the probability distribution of each focal element among pieces of evidence. This study makes full use of the complementarity of the above two; a new expression of conflict factor is combined, as shown in Equation (25). The new conflict factor $k'$ is the result of the combination of conflict factor $k$ and evidence distance $d_{BPA}(m_i, m_j)$. When and only if both of them are zero, it means that there is no conflict between the pieces of evidence, which overcomes the error problem of judging evidence conflict by a single condition.

$$k' = \frac{d_{BPA}(m_i, m_j) + k}{2} \quad (20)$$

## CASE STUDY

### Engineering background and evaluation system

The middle route of the south to north water diversion project is a long-distance and super large water conservancy project in China, with a total length of 1,431.98 km and a total head difference of about 100 m. There are many kinds of buildings along the project, including cross buildings (such as inverted siphon, aqueduct, and highway bridge) and control buildings (such as sluice, tunnel, and pump station) and a total of 1,796. The water supply was officially opened on 12 December 2014, benefiting more than 60 million people. To ensure the safe operation of the project, it is necessary to set up appropriate monitoring items and many measuring points according to the type and structure of buildings. The safety monitoring items of water conveyance structures can be divided into deformation, stress, strain, seepage, and temperature. The corresponding sensors installed on various buildings include displacement gauge, earth pressure gauge, strain gauge, osmometer, thermometer, joint gauge, steel bar gauge, water level meter, etc. The evaluation grade of building safety state is divided

into three grades: normal (A), abnormal (B), and dangerous (C). Then, the evaluation is set as $\phi = \{A, B, C\}$.

In this study, the long-term safety monitoring data of an inverted siphon located in the Henan section of the middle route of the south to north water diversion project is selected to evaluate the safety status of diversion buildings. The safety evaluation system and fusion model established for the inverted siphon are shown in Figure 3. The building is divided into five parts, and $a$, $b$, $c$, and $d$ in the figure represent the number of monitoring points of the same kind of monitoring indicators on the pipe body. The monitoring indicators include four types of displacement, earth pressure, stress, and seepage pressure. The types and quantities of monitoring indicators for different parts of buildings are different. The specific information of the sensor layout is shown in Table 1.

This research proposes a three-tier information fusion framework for the safety evaluation of diversion buildings, including data-level fusion, feature-level fusion, and decision-level fusion. The data layer obtains the feature information
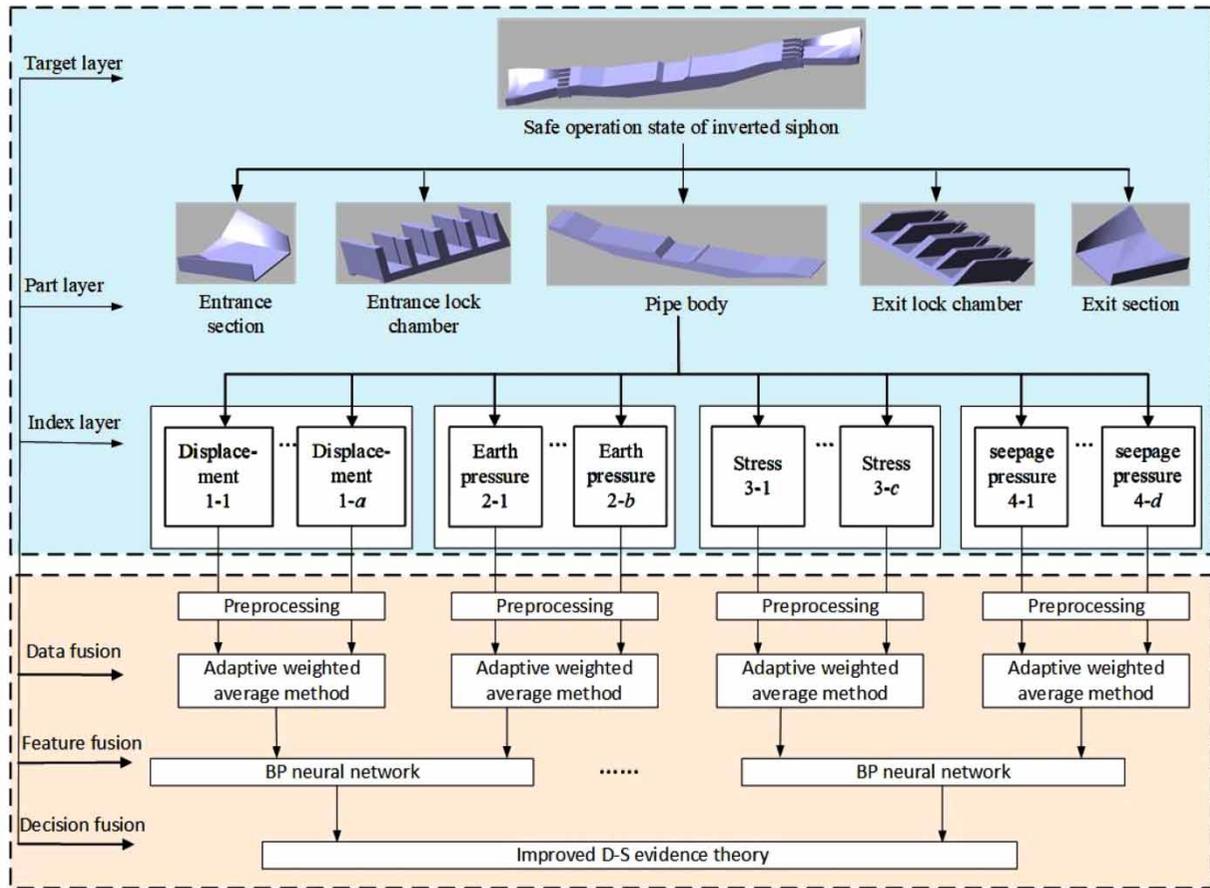


**Figure 3** │ Evaluation system and fusion model of inverted siphon.

**Table 1** │ Type and number of sensors

| Part | Entrance part | | Entrance lock chamber | | | | Pipe body | | | | Exit lock chamber | | | | Exit part | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Index | C2 | C4 | C1 | C2 | C3 | C4 | C1 | C2 | C3 | C4 | C1 | C2 | C3 | C4 | C2 | C4 |
| Type | S2 | S4 | S1 | S2 | S3 | S4 | S1 | S2 | S3 | S4 | S1 | S2 | S3 | S4 | S2 | S4 |
| Number | 1 | 4 | 4 | 5 | 7 | 7 | 10 | 14 | 48 | 8 | 4 | 5 | 7 | 8 | 1 | 3 |

by directly fusing the original data monitored by multiple sensors. The feature layer fuses the feature information to obtain the local judgment result. Finally, the overall safety status of the building operation is judged by the decision-making layer. The fusion information and results between different levels are transferred to other levels through the database, and the data fusion is realized in this interactive way. Data-level fusion has the advantages of fusing a large amount of original data and providing detailed information for feature-level and decision-level fusion. Feature-level fusion provides association information for decision-level fusion by analyzing feature information comprehensively. Decision-level fusion has the advantages of small sensor dependence, strong anti-interference ability, and good flexibility.

Accordingly, the multi-source data fusion method proposed in this paper is divided into three steps. The first step is to use the AWAM for data-level fusion of similar monitoring data in building parts. The second step is to use the BPNN for the feature-level fusion of heterogeneous sensors in the same part of the building. The third step is to use the D-S theory to fuse the fusion data of different parts of the building. Through data fusion of multi-type sensor data, the overall safe operation state of the building is judged.

In this table, C1, C2, C3, and C4 represent displacement, earth pressure, stress, and seepage pressure, respectively. S1, S2, S3, and S4 represent sensor joint meter, earth pressure gauge, reinforcement meter, and osmometer, respectively.

## Fusion steps and results

According to the three-level data fusion algorithm proposed in this paper, the specific steps of the safety state evaluation of the diversion building are as follows.

### Step 1: data-level fusion

First, the improved support function is introduced to preprocess the data collected by sensors of five parts of the building to eliminate the invalid outliers with a large deviation. Second, the AWAM is used to fuse the data of the same kind of sensors in each part according to Equations (6)–(8).

Figures 4 and 5, respectively, represent the monitoring data of displacement and stress from 2013 to 2019 after data preprocessing and data-level fusion. According to Table 1, these two monitoring indicators only belong to the entrance lock chamber, pipe body, and exit lock chamber. Figures 6 and 7, respectively, represent the monitoring data of earth pressure and seepage pressure from 2013 to 2019 after data preprocessing and data-level fusion. These two monitoring indicators are installed in five parts of an inverted siphon. It can be seen from these figures that the fusion data of displacement, earth pressure, and stress changes greatly at the initial stage of building operation. After 2015, they all change periodically and stably with the seasonal temperature range, and the variation range is within the normal range. Before the water supply, the osmometer value is mainly affected by the groundwater level. Due to the pumping and drainage work in the construction site in the early stage, the osmometer of the building floor is basically in the state of no water, and the change of seepage pressure is very small. After the water was officially filled on 12 December 2014, the seepage pressure at each measuring point of the inverted siphon showed an increasing trend. Among them, the seepage pressure of the pipe body section is larger, reaching the maximum value in the flood season of 2015. It can be seen that the inverted siphon is affected by groundwater in
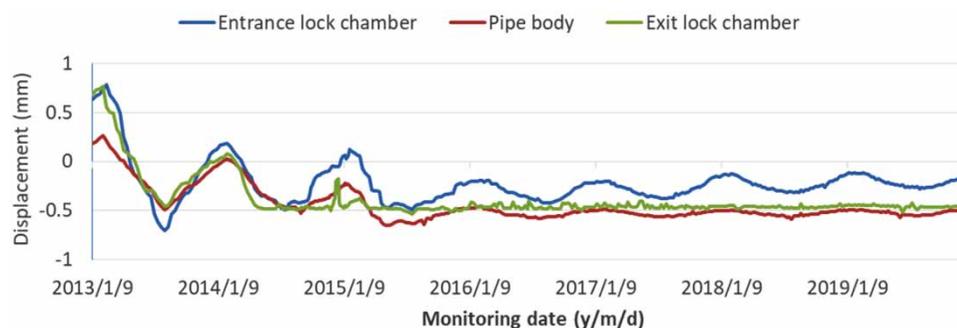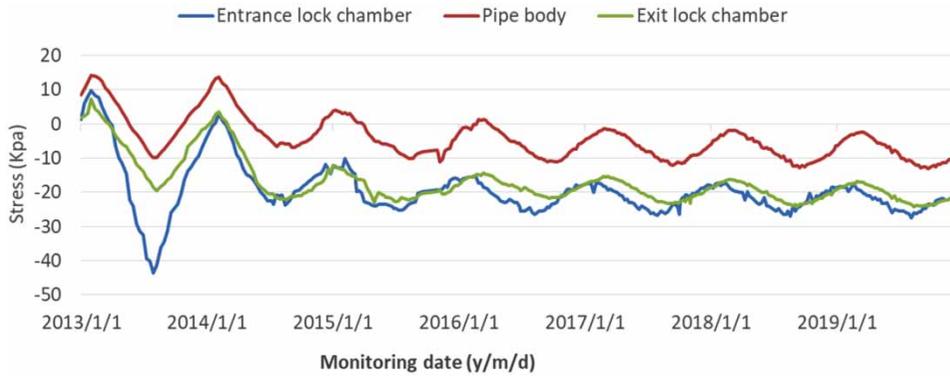


**Figure 4** | Data-level fusion results of displacement.

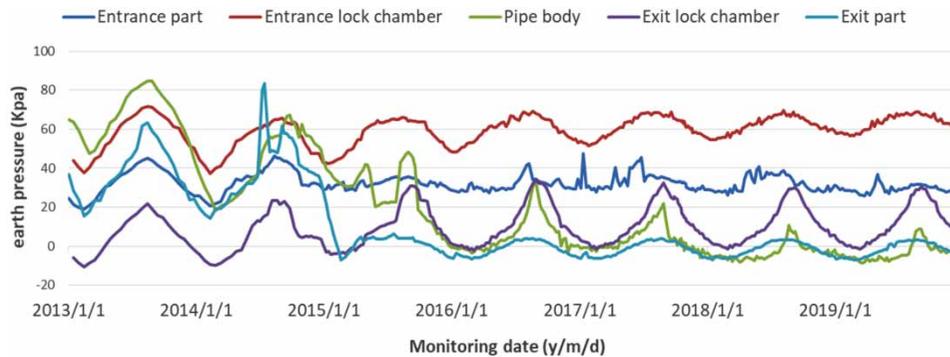**Figure 5** │ Data-level fusion results of stress data.

**Figure 6** │ Data-level fusion results of earth pressure.
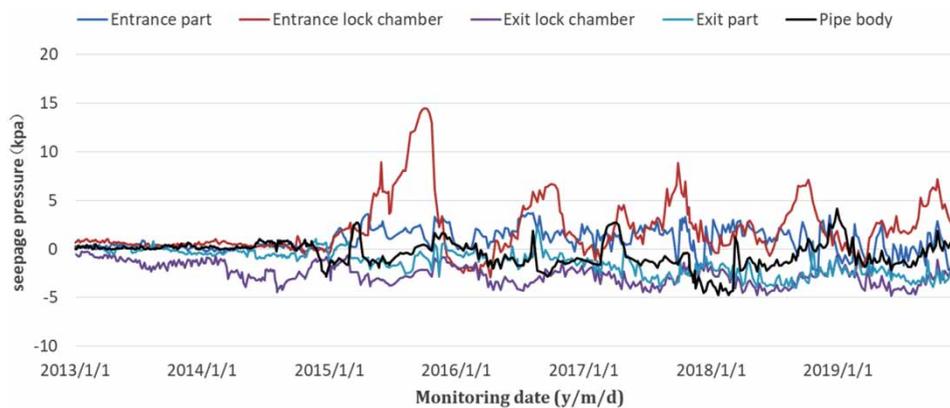
**Figure 7** │ Data-level fusion results of seepage pressure.

the flood season. From the changing trend of measured values, most of the measured values are normal, so it can be judged that the project is in a safe operating state.

Therefore, the long-term operation trend of each building can be preliminarily judged according to the data-level fusion results.

## Step 2: feature-level fusion

Before constructing the BPNN, the number of neurons in the input layer must be determined. Each part of the building partition in this study is equivalent to a sub-neural network. The data-level fusion results of different sensors in different parts are used as the input of the neural network. The number of neurons in the input layer is equal to the number of sensor types contained in each part. Taking the entrance part as an example, as can be seen from Table 1 that there are two kinds of sensors in this part: earth pressure gauge and osmometer, so the number of input layers is two. Similarly, the number of neurons in the input layer of other buildings can be obtained as shown in Table 3. Since the different data collected by different sensors, the larger input data will play a greater role in simultaneous interpreting and lead to longer training time. Therefore, it is necessary to normalize the training data to the interval [0, 1].

The number of neurons in the hidden layer is calculated by the empirical Equation (9) listed in the 'Identify abnormal data' section. The number of neurons with the minimum error is selected as the number of hidden layer neurons corresponding to each sub-neural network. Figure 8 shows the change of *MSE* with the number of hidden layers. *MSE* takes the average value of 10 training times of each sub-neural network. In the case of the minimum error, the number of hidden layer neurons corresponding to the entrance part, entrance lock chamber, pipe body, exit lock chamber, and exit part is shown in Table 3.

The number of neurons in the output layer is determined according to the number of building safety evaluation level. According to the evaluation set as $\phi = \{A, B, C\}$, the number of output layer neurons of the BPNN is set to 3. In this study,

the output of the BPNN is defined by binary. The definition of the output is shown in Table 2.

In this study, the BPNN is used to fuse heterogeneous sensor data in different regions. MATLAB education version 2019b software was used to build and train the BPNN model. A three-layer BPNN is designed as the feature-level data fusion algorithm, that is, the number of hidden layers is 1. The number of each layer of the neural network is determined according to Table 3.

The transfer function of the hidden layer neuron of the neural network uses tansig(). Since the output is limited to (0, 1), the transfer function of output layer neurons uses logsig(). The Levenberg Marquardt algorithm with the fastest convergence speed, namely the traimlm() function is used to train the neural network. The sample sizes of the training set and test set for each sub-neural network are determined by the sampling frequency, data type, and pre-processing results in 2013–2019. The number of samples

**Table 2** │ BPNN output layer definition

| Building safety level | Normal | Abnormal | Dangerous |
|---|---|---|---|
| Output results | (0,0,1) | (0,1,0) | (1,0,0) |

**Table 3** │ BPNN structure design of building part layer

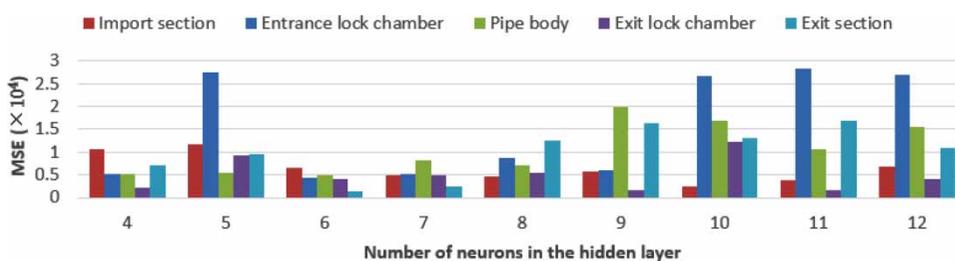| Part | Entrance part | Entrance lock chamber | Pipe body | Exit lock chamber | Exit part |
|---|---|---|---|---|---|
| Number of input layers | 2 | 4 | 4 | 4 | 2 |
| Number of hidden layers | 10 | 6 | 6 | 9 | 6 |
| Number of output layers | 3 | 3 | 3 | 3 | 3 |



**Figure 8** │ Different numbers of neurons and errors in the hidden layer.

and feature-level fusion results contained in different parts are shown in Table 4.

## Step 3: decision-level fusion

The recognition framework is composed of a building safety state evaluation set as $\phi = \{A, B, C\}$. The fusion results of the BPNN in step 2 are normalized to provide an initial $m$ value for D-S evidence theory. Equations (15)–(20) are used to evaluate the overall safety status of buildings. The evidence BPA and fusion results corresponding to the five parts are shown in Table 5. After the fusion of decision level, the BPA of the whole building in the normal condition is far greater than that of abnormal BPA, which is consistent with the actual operation state. It can be seen that data fusion based on D-S theory can eliminate the uncertainty

of building evaluation and improve the accuracy of evaluation results.

## Results analysis

Data preprocessing is an indispensable step before the safety evaluation of the whole building structure using data fusion technology. Figure 9 shows the stress monitoring data process line of three earth pressure gauges (Ej-1, Ej-2, Ej-3) from 2013 to 2019 in the entrance lock chamber. It can be seen that due to the interference of external factors, the monitoring data of the three soil pressure sensors will have varying degrees of fluctuations and sharp angles within a certain period of time. If these data are directly fused without preprocessing, the result is shown as the blue line in Figure 10. The data change process fluctuates

**Table 4** | Feature-level fusion results of building BPNN

| Part | Entrance part | Entrance lock chamber | Pipe body | Exit lock chamber | Exit part |
|---|---|---|---|---|---|
| Number of training sets | 210 | 280 | 230 | 230 | 228 |
| Number of test sets | 90 | 98 | 100 | 99 | 98 |
| Fusion results | (0.1138, 0.2437, 0.6495) | (0.0629, 0.3510, 0.5860) | (0.0460, 0.5390, 0.4150) | (0.2740, 0.1290, 0.5970) | (0.1710, 0.2730, 0.5560) |

**Table 5** | Evaluation results of each part and the whole building

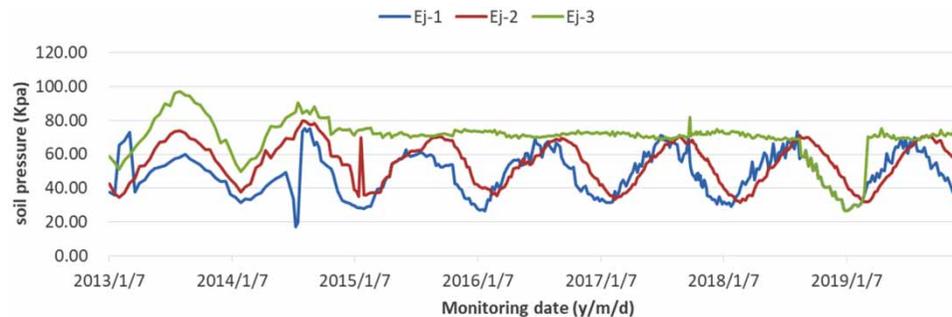| Evaluation level | Entrance part | Entrance lock chamber | Pipe body | Exit lock chamber | Exit part | Building safety |
|---|---|---|---|---|---|---|
| Dangerous | 0.113 | 0.063 | 0.046 | 0.274 | 0.171 | 0.00 |
| Abnormal | 0.242 | 0.351 | 0.539 | 0.129 | 0.273 | 0.01 |
| Normal | 0.645 | 0.586 | 0.415 | 0.597 | 0.556 | 0.99 |



**Figure 9** | Monitoring result process line of the earth pressure gauge in the entrance lock chamber.
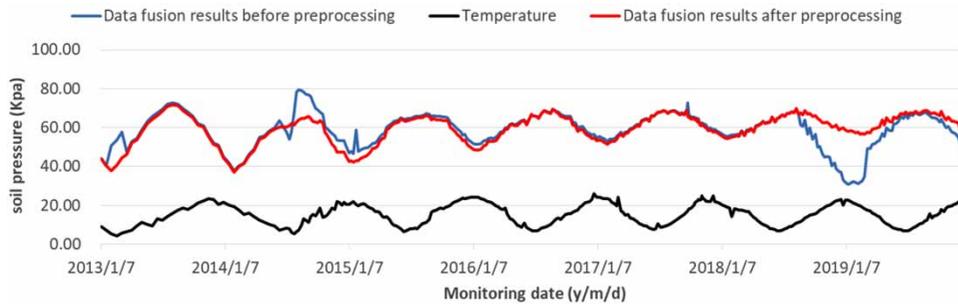
**Figure 10** │ Comparison of fusion results before and after preprocessing. Please refer to the online version of this paper to see this figure in color: https://doi.org/10.2166/hydro.2021.154.

greatly, which does not conform to the actual change law. After eliminating invalid outliers, the data-level fusion results are shown in the red line in Figure 10, and the fusion results tend to be periodic and stable. Combined with the temperature change line in Figure 10, it can be seen that the current stress changes periodically with temperature. As the temperature increases, the stress changes in the compression direction, and as the temperature decreases, the stress changes in the tensile direction. Therefore, the improved self-support algorithm introduced in this paper can eliminate uncertain data and improve the accuracy of data fusion.

To verify the effectiveness of the hybrid fusion algorithm proposed in this paper, taking the monitoring data of the

entrance part of the inverted siphon as an example, the AWAM is used to fuse the monitoring data within the same monitoring period. Figure 11 shows the trend of seepage pressure monitoring data and fusion data at the entrance section of buildings. Figure 12 shows the changing trend of monitoring data and fusion data after eliminating large fluctuation data. It can be seen from the figure that although the individual data collected at the measuring point Pa-1 fluctuates greatly due to external factors, the fluctuating data has little effect on the fusion result. The results show that the use of the AWAM for data fusion can effectively improve the accuracy of building safety evaluation monitoring data.

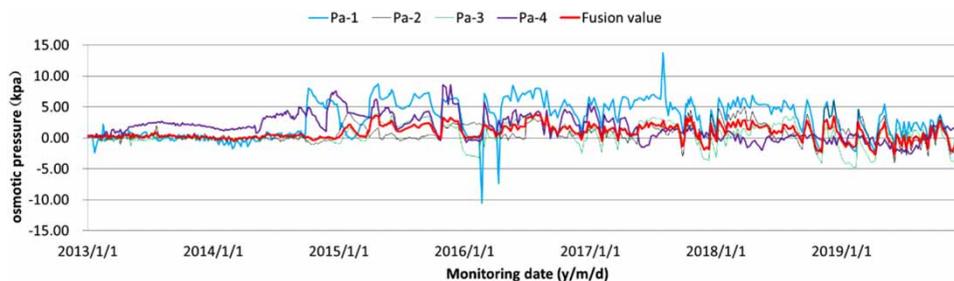Then, according to the five evidence sources in Table 5, the traditional D-S evidence theory and the algorithm



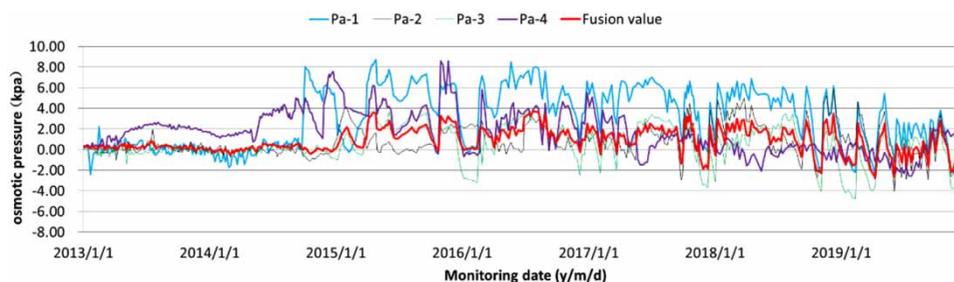**Figure 11** │ Monitoring data and fusion data of osmometer at the entrance part.



**Figure 12** │ Monitoring data and fusion data after removing abnormal values.

proposed in this paper are used to evaluate the overall safety status of buildings. The evaluation results are shown in Table 6. The results show that the fusion results of the two algorithms all point to A level, that is, the inverted siphon building is in a normal operation state. Therefore, the BPA value of A-grade evaluated by the algorithm in this paper is obviously better than the traditional D-S evidence theory in the case of different fusion parameters, which proves that the Improved D-S fusion algorithm is more accurate. After improving the conflict factor, the probability accumulation of evidence is more obvious than the traditional D-S theory, and it is feasible to use the BPNN to solve BPA.

Finally, the AWAM, BPNN, D-S evidence theory, and the algorithm in this paper are used for data fusion. The comparison of fusion results and MSE is shown in Table 7. It can be concluded that the MSE using the three algorithms alone is much larger than the algorithm in this paper. Compared with the traditional D-S theory, this algorithm reduces the MSE from 0.18 to 0.01%. Therefore, it is proved that the three-level fusion model designed in this paper improves the accuracy of the system. The results show that the result of the multi-sensor hybrid fusion algorithm is more in line with the actual situation.

In multi-objective decision analysis, decision makers need to know the impact of the change of decision-making information on the final decision-making results. Therefore, this study analyzes the sensitivity of the BPA value of different parts of the building on the overall safety state of the building. In practice, the building safety management personnel will focus on the parts with evaluation level C, that is, the dangerous parts. So this study focuses on the analysis of the impact on the overall safety state of the building when the BPA of grade C in each part of the building changes slightly. The change range of BPA of grade C in each part of the building is $\delta A$. To ensure the sum of the BPA value of each part is 1, the changed value is added to the BPA of level A of the building, which is opposite to level C. In this study, the sensitivity coefficient $E$ was used to evaluate the sensitivity of each factor to the fusion results. The higher the sensitivity coefficient, the higher the sensitivity. The calculation formula is shown in the following formula. The data before the change are shown in Table 5, and the results of sensitivity analysis are shown in Table 8.

$$E = \frac{\Delta A}{\Delta F} \tag{21}$$

where $\Delta F$ is the change rate of each factor, and the value in this study is $\Delta F$. $\Delta A$ represents the change rate of the corresponding evaluation grade when factor $F$ changes $\Delta F$.

The results of the sensitivity analysis show that the entrance part of the inverted siphon has the highest

**Table 6** | Comparison of fusion results of different parameters

| Algorithm | Evaluation level | $\oplus_{i=1}^{2} m_{i'}$ | $\oplus_{i=1}^{3} m_{i'}$ | $\oplus_{i=1}^{4} m_{i'}$ | $\oplus_{i=1}^{5} m_{i'}$ |
|---|---|---|---|---|---|
| D-S evidence theory | C | 0.015 | 0.002 | 0.001 | 0.000 |
| | B | 0.181 | 0.226 | 0.059 | 0.030 |
| | A | 0.804 | 0.773 | 0.940 | 0.970 |
| The algorithm proposed in this paper | C | 0.026 | 0.002 | 0.001 | 0.000 |
| | B | 0.116 | 0.089 | 0.017 | 0.006 |
| | A | 0.857 | 0.910 | 0.982 | 0.994 |

**Table 7** | Fusion results and error comparison table

| Algorithm | Adaptive weighted averaging | BPNN | D-S evidence theory | The algorithm proposed in this paper |
|---|---|---|---|---|
| Results | (0.066, 0.252, 0.682) | (0.027, 0.221, 0.752) | (0.000, 0.030, 0.970) | (0.000, 0.006, 0.994) |
| MSE | 0.1690 | 0.1111 | 0.0018 | 0.0001 |

**Table 8** | Sensitivity analysis results

| Sensitivity coefficient | $E_A$ | $E_B$ | $E_C$ |
|---|---|---|---|
| P1 + 10% | − 0.0030 | 0.4516 | 1.3016 |
| P2 + 10% | − 0.0004 | 0.0449 | 0.9800 |
| P3 + 10% | − 0.0002 | 0.0067 | 1.0074 |
| P4 + 10% | − 0.0016 | 0.2369 | 1.2606 |
| P5 + 10% | − 0.0008 | 0.0953 | 1.1048 |
| P1 − 10% | − 0.0015 | 0.2163 | 1.1948 |
| P2 − 10% | − 0.0004 | 0.0415 | 0.9858 |
| P3 − 10% | − 0.0002 | 0.0029 | 1.0026 |
| P4 − 10% | − 0.0006 | 0.0789 | 0.9118 |
| P5 − 10% | − 0.0007 | 0.0890 | 1.0801 |

sensitivity to the overall safety evaluation of buildings. Therefore, the monitoring of the entrance section should be strengthened. The influence degree of the overall safety state of buildings in other parts is similar, which indicates that the overall safety state of inverted siphon depends on the comprehensive results of multiple indexes of multiple parts. Therefore, the multi-source data fusion method proposed in this paper is sensible and reliable.

In this table, P1, P2, P3, P4, and P5 represent the entrance part, entrance lock chamber, pipe body, exit lock, and exit part, respectively. $P1 + 10\%$ means that the BPA value of level C increases by 10% and that of level A decreases by 10%. $P1 - 10\%$ indicates that the BPA value of grade C decreases by 10% and the BPA value of level A increases by 10%. The meaning of other indicators is the same. $E_A$, $E_B$, and $E_C$ represent the sensitivity coefficients of A, B, and C in the final fusion results.

## CONCLUSIONS

With the development of information technology, a large number of multi-source heterogeneous data is generated every day during the operation of the project. Building safety evaluation based on single sensor data has low reliability and high uncertainty. Therefore, how to use these real-time data to evaluate the safety status of building structures is a challenging problem. This paper proposes a multi-source data fusion method based on Improved D-S and BPNN. In practical application, a three-level data fusion model based on the data level, feature level, and decision level is established according to the structural characteristics of water diversion structures to comprehensively evaluate the structural safety status. The results show that: (1) By introducing the improved support function to preprocess the original data, the invalid abnormal data collected due to sensor fault can be eliminated, and the accuracy of subsequent data fusion can be improved. (2) Data-level fusion based on the AWAM can effectively reduce the impact of large fluctuation data on evaluation results, improve the accuracy of monitoring data, and provide the input value for the neural network. (3) Taking the output of neural network as the evidence input of evidence theory can solve the problem that evidence is difficult to

obtain in the evidence theory fusion method and realize the local safety state evaluation of buildings. (4) By introducing BD to measure the distance between different evidence, the Improved D-S solves the conflict between evidence. Compared with the results of traditional D-S evidence theory, this method improves the accuracy of building overall safety state evaluation.

Compared with the single fusion method, the MSE of the hybrid data fusion method proposed in this paper is smaller, which proves that this method is an effective method for building safety evaluation. Therefore, it is feasible to evaluate the safety status of buildings through multi-level data fusion to ensure the safe operation of the project under the condition of a large building range and a large amount of monitoring data. In the follow-up research, we will consider how to combine sensor monitoring data and inspection text data to evaluate the safety status of buildings more accurately. In addition, with the development of Internet-of-things technology and the continuous operation of the project, the type and quantity of sensor data will continue to increase, forming a large amount of multi-source big data. Therefore, how to combine cloud computing and big data analysis and other advanced technologies to process these massive multi-source data and make a more scientific and accurate assessment of the building safety situation is also the research focus in the next stage.

## ACKNOWLEDGEMENTS

## DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

## REFERENCES

Azamfar, M., Singh, J., Bravo-Imaz, I. & Lee, J. 2020 Multisensor data fusion for gearbox fault diagnosis using 2-D convolutional

neural network and motor current signature analysis. *Mechanical Systems and Signal Processing* **144**, 106861.

Bazrkar, M. H., Adamowski, J. & Eslamian, S. 2017 *Water System Modeling, in Mathematical Advances Towards Sustainable Environmental Systems*. Springer International Publishing, Switzerland, pp. 61–88.

Bi, S. F., Broggi, M. & Beer, M. 2019 The role of the Bhattacharyya distance in stochastic model updating. *Mechanical Systems and Signal Processing* **117**, 437–452.

Bin, G. F., Jiang, Z. N., Li, X. J. & Dhillon, B. S. 2011 Weighted multi-sensor data level fusion method of vibration signal based on correlation function. *Chinese Journal of Mechanical Engineering* **24** (5), 899–904.

Brierley, N., Tippetts, T. & Cawley, P. 2014 Proceedings of the Royal Society a-Mathematical Physical and Engineering Sciences **470** (2167), 20140167.

Castanedo, F. 2013 A review of data fusion techniques. *Scientific World Journal* 704504.

Deng, Z. & Wang, J. Y. 2020 A novel evidence conflict measurement for multi-sensor data fusion based on the evidence distance and evidence angle. *Sensors* **20** (2), 381.

Deng, X. Y., Xiao, F. Y. & Deng, Y. 2017 An improved distance-based total uncertainty measure in belief function theory. *Applied Intelligence* **46** (4), 898–915.

Gong, B. 2009 An algorithm of data fusion using artificial neural network and Dempster-Shafer evidence theory. Iita International Conference on Control, Automation and Systems Engineering, Proceedings, 11–12 July 2009, Zhangjiajie, China, pp. 407–410.

Guan, X., Yi, X. & He, Y. 2008 Study on algorithms of determining basic probability assignment function in Dempster-Shafer evidence theory. Proceedings of 2008 International Conference on Machine Learning and Cybernetics, 12–15 July 2008, Kunming, China, pp. 1–7, 121–126.

Haq, A. 2020 A maximum adaptive exponentially weighted moving average control chart for monitoring process mean and variability. *Quality Technology and Quantitative Management* **17** (1), 16–31.

Huang, M., Liu, Z. & Tao, Y. 2020 Mechanical fault diagnosis and prediction in IoT based on multi-source sensing data fusion. *Simulation Modelling Practice and Theory* **102**, 101981.

Jia, J. S. 2016 A technical review of hydro-project development in China. *Engineering* **2** (3), 302–312.

Jiang, Z. X. & He, J. P. 2016 Method of fusion diagnosis for dam service status based on joint distribution function of multiple points. *Mathematical Problems in Engineering* **2016** (pt. 9), 1–10.

Jiang, Y. H., Li, L. B., Zhang, Q. & Lu, Z. C. 2013 A study on wireless private network technology in an integrated automation system for safety monitoring of the south-to-north water diversion project. *Sustainable Development and Environment II, Pts 1 and 2* **409–410**, 1638.

Li, Y. Z., Yao, S. J., Zhang, R. Y. & Yang, C. 2020 Analyzing host security using D-S evidence theory and multisource information fusion. *International Journal of Intelligent Systems* **36** (2), 1053–1068.

Liu, X., Wu, Z. R., Yang, Y., Hu, J. & Xu, B. 2012 Information fusion diagnosis and early-warning method for monitoring the long-term service safety of high dams. *Journal of Zhejiang University – Science A* **13** (9), 687–699.

Liu, Y., Xu, C., Huang, B., Ren, X. W., Liu, C. Q., Hu, B. D. & Chen, Z. 2020 Landslide displacement prediction based on multi-source data fusion and sensitivity states. *Engineering Geology* **271**, 105608.

Long, D., Yan, L., Bai, L. L., Zhang, C. J., Li, X. Y., Lei, H. M., Yang, H. B., Tian, F. Q., Zeng, C., Meng, X. Y. & Shi, C. X. 2020 Generation of MODIS-like land surface temperatures under all-weather conditions based on a data fusion approach. *Remote Sensing of Environment* **246**, 111863.

Mehta, D. J. & Yadav, S. M. 2020 Analysis of scour depth in case of parallel bridges using HEC-RAS. *Water Supply* **20** (8), 3419–3432.

Mehta, D., Yadav, S. M., Waikhom, S. & Prajapati, K. 2020 Stable channel design of Tapi River using HEC-RAS for Surat region. *Environmental Processes and Management* **91**, 25–36.

Murphy, C. K. 2000 Combining belief functions when evidence conflicts. *Decision Support Systems* **29** (1), 1–9.

Osegueda, R. A., Seelam, S. R., Mulupuru, B. & Kreinovich, V. 2003 Statistical and Dempster-Shafer techniques in testing structural integrity of aerospace structures. *Smart Nondestructive Evaluation and Health Monitoring of Structural and Biological Systems II* **5047**, 140–151.

Ren, Y. F., Ge, Y. W. & Bai, X. C. 2012 Research on optimal weight choice of multi-MEMS gyroscope data fusion. *Advanced Mechanical Engineering II* **192**, 351–355.

Samadi, M., Jabbari, E. & Azamathulla, H. 2014 Assessment of M5' model tree and classification and regression trees for prediction of scour depth below free overfall spillways. *Neural Computing & Applications* **24** (2), 357–366.

Samadi, M., Jabbari, E., Azamathulla, H. M. & Mojallal, M. 2015 Estimation of scour depth below free overfall spillways using multivariate adaptive regression splines and artificial neural networks. *Engineering Applications of Computational Fluid Mechanics* **9** (1), 291–300.

Shi, P., Li, G. H., Yuan, Y. M. & Kuang, L. 2018 Data fusion using improved support degree function in aquaculture wireless sensor networks. *Sensors* **18** (11), 1–16.

Su, H. Z., Wen, Z. P., Sun, X. R. & Yan, X. Q. 2018 Multisource information fusion-based approach diagnosing structural behavior of dam engineering. *Structural Control & Health Monitoring* **25** (2), e2073.

Valipour, M. 2016 How do different factors impact agricultural water management? *Open Agriculture* **1** (1), 89–111.

Valipour, M. 2017 Global experience on irrigation management under different scenarios. *Journal of Water and Land Development* **32** (1), 95–102.

Wang, S. X. 2020 Multisensor data fusion of motion monitoring system based on BP neural network. *Journal of Supercomputing* **76** (3), 1642–1656.

Wu, R. T. & Jahanshahi, M. R. 2020 Data fusion approaches for structural health monitoring and system identification: past, present, and future. *Structural Health Monitoring – An International Journal* **19** (2), 552–586.

Wu, X. G., Duan, J. C., Zhang, L. M. & AbouRizk, S. M. 2018 A hybrid information fusion approach to safety risk perception using sensor data under uncertainty. *Stochastic Environmental Research and Risk Assessment* **32** (1), 105–122.

Xiao, R. Y., Shi, H. K., He, X. F., Li, Z. H., Jia, D. Z. & Yang, Z. X. 2019 Deformation monitoring of reservoir dams using GNSS: an application to south-to-north water diversion project, China. *IEEE Access* **7**, 54981–54992.

Xie, C. L. & Guan, Q. 2008 Compare research of data fusion and neural network diagnosis method. International Conference on Multimedia and Information Technology, Proceedings, 30–31 December 2008, Three Gorges, China, pp. 209.

Yager, R. R. 2001 The power average operator. *IEEE Transactions on Systems Man and Cybernetics Part A – Systems and Humans* **31** (6), 724–731.

Yang, J. C., Wen, J. B., Wang, Y. H., Jiang, B., Wang, H. H. & Song, H. B. 2020 Fog-based marine environmental information monitoring toward ocean of things. *IEEE Internet of Things Journal* **7** (5), 4238–4247.

Yue, S. H., Wu, T., Wang, Y. M., Zhang, K. & Liu, W. X. 2010 Dampster-Shafer evidence theory based multi-characteristics fusion for clustering evaluation. *Rough Set and Knowledge Technology (RSKT)* **6401**, 499–519.

Zhang, H. P. & Deng, Y. 2020 Weighted belief function of sensor data fusion in engine fault diagnosis. *Soft Computing* **24** (3), 2329–2339.

Zhang, F., Xue, H. F. & Zhang, Y. H. 2019 A new BP neural network fusion algorithm for multi-source remote sensing data on groundwater. *Applied Ecology and Environmental Research* **17** (4), 9083–9095.

Zhao, G. Z., Chen, A. G., Lu, G. X. & Liu, W. 2020 Data fusion algorithm based on fuzzy sets and D-S theory of evidence. *Tsinghua Science and Technology* **25** (1), 12–19.

Zhu, H. P., Gao, K., Xia, Y., Gao, F., Weng, S., Sun, Y. & Hu, Q. 2020 Multi-rate data fusion for dynamic displacement measurement of beam-like supertall structures using acceleration and strain sensors. *Structural Health Monitoring – An International Journal* **19** (2), 520–536.