A Water Leakage Risk Assessment Model for Shield Tunnel Based on Kalman Filter Data Fusion Method

Wang Jiaxu¹,², Xue Yadong¹,², Li Xuhui³, Chen Yuelong³, Fu Yanbin⁴

1. Key Laboratory of Geotechnical and Underground Engineering (Tongji University), Ministry of Education, Shanghai 200092, China;
2. Department of Geotechnical Engineering College of Civil Engineering, Tongji University, Shanghai 200092, China;
3. Guangdong Yuehai Pearl River Delta Water Supply Co., Ltd, Guangzhou 511455, China;
4. College of Civil Engineering and Transportation, Shenzhen University, Shenzhen 518060, China;

*Corresponding author: yadongxue@126.com

Abstract. Multi-source information monitoring and data fusion analysis system has high practical value for the risk management of tunnel structure, especially in complex ground. The analysis of structure health will be getting more comprehensive and real-time with multisource monitoring technologies. Risk warning will be more accurate and timelier. Based on one shield tunnel project, this research analysed the water leakage risk of shield subsea tunnel in unfavourable geological conditions by using the method of geological information and structural monitoring data fusion. Firstly, Kalman filter method was used to denoise the monitoring data, and the effect of different monitoring data was compared; Then, the risk factor of water leakage was analysed, and the relationship between risk factors, geological information and monitoring data was established; Finally, the risk factors based on multi-source monitoring information were fused and analysed by using the improved D-S evidence theory method, and the final analysis results of water leakage risk were obtained. A multisource dynamic risk analysis model of tunnel water leakage was established. The results showed that the Kalman filter method has better noise reduction effect on the high-frequency monitoring data than that of low-frequency monitoring data, and the multi-source dynamic risk analysis model has a good prediction effect on the water leakage risk of tunnel lining in unfavourable ground condition.

Keywords: Water leakage, Risk assessment, Kalman Filter, D-S evidence theory, Data fusion

1. Introduction

In recent years, with the development of wireless sensing, photogrammetry, acoustic emission sensor and other new monitoring technologies, tunnel risk monitoring has gradually developed from traditional manual monitoring to semi-automatic and automatic monitoring. Meanwhile, multisource sensing system has been widely used in tunnel safety monitoring. Through multisource information monitoring, major risks could be identified more quickly and accurately. In addition, management team could predict the evolution of the risk through the method of multi-source information fusion, and track the effect of risk disposal. Therefore, the information fusion method based on multisource sensing system has high application value for the safety and risk management of tunnel engineering.

The research needs to focus on two problems: First, real time noise reduction. Common noise reduction methods include equipment denoise and data filter, which have been used in many
But real time noise reduction is still being researched; Secondly, the method of multisource data fusion, including data, feature and decision fusion \[7,9\]. Zhou et al. \[10\] analysed the fusion of multisource data in data and feature aspects by using AHP and DS evidence theory, and established a set of risk quantitative analysis model. Xie et al. \[11\] analysed the multisource data fusion of the decision aspect of the collapse risk of the new Austrian tunnel by using the numerical simulation method and the fuzzy comprehensive evaluation method. Compared with data level and feature level fusion, it has better fault tolerance and anti-interference capability.

In this study, based on a shield tunnel project, the multisource data fusion method was adopted to evaluate the water leakage risk of tunnel structure. First, the Kalman filter method was used to pre-process monitoring data. Then, the factors of the leakage risk were analysed. The improved D-S evidence theory method was used to fuse multisource data. Finally, the two processes were combined, a dynamic risk assessment model of water leakage of shield tunnel structure was established.

2. Multisource perceptual system

2.1. Case study backgrounds

This study is based on section B3 of Pearl River Delta Water Allocation Project. Section B3 is 11.36km long and has four shield tunnel sections. The underwater shield section is 4.2km long, crossing Lianhuashan river and Shiziyang river. The diameter of the tunnel is 8.65m and depth is about 46m. Complex geological conditions is an important challenge for tunnel construction, including 9 major faults, unidentified uneven weathered strata and mud-bearing stratum. Mud-water shield method is used to excavate for underwater section. The roadmap of B3 section and geological section map for shield tunnel section 2 is shown in figure 1 and figure 2:

![Figure 1. Tunnel roadmap of B3 section.](image1)

![Figure 2. Geological section map of shield section 2.](image2)

Water leakage, segment damage and nonuniform settlement are typical structural risks of this project. Therefore, multisource perceptual system is used to monitor the state of structure, which could find potential risks of the structure in advance and take early measures.

2.2. System framework

With the development of sensor technology and big data analysis technology, multisource perceptual system has become an important means of monitoring infrastructure such as tunnels. The multisource perceptual system of this project is composed of sensing nodes, intelligent terminals, cloud server and visualization platform, as shown in figure 3:
There are two approaches for data collection and recording: automatic method and manual method. For automatic method, the sensing node is composed of one or more sensors. As shown in figure 4 and figure 5, the sensing node is responsible for data collection of a specific monitoring task. The sensing node in a region transmits the data to the intelligent terminal by wired or wireless way, and the intelligent terminal summarizes and sends the data to cloud server. In addition, the sensor node is responsible for the data collection of a specific monitoring task. The intelligent terminal is also responsible for monitoring the status of each sensor and passing the commands sent by cloud server to the corresponding sensors. Manual method mainly refers to the process of technicians measuring on site through special equipment, and recording the data into server. The cloud server is responsible for the update, storage, analysis and call of data, and feedback the response data or analysis results to the platform according to the command of the visual platform.

For the structural safety risk of the project, the monitoring items, monitoring equipment, monitoring mode, monitoring frequency and other information designed are shown in table 1:

| Object       | Item            | Equipment | Mode | Frequency     |
|--------------|-----------------|-----------|------|---------------|
| Structure    | Vertical        | Total station | Manual | 2 times / day |
|              | contraction     |           |      |               |
|              | Tunnel settlement| Total station | Manual | 2 times / day |
|              | Horizontal      | Total station | Manual | 2 times / day |
|              | displacement    |           |      |               |
|              | Structure joint opening | Displacement sensor | Automatic | 3 times / hour |
|              | Bolt stress     | Stress sensor | Automatic | 3 times / hour |
| Ground       | Water pressure  | Piezometer | Automatic | 3 times / hour |

3. Data analysis
3.1. Data preprocessing

Due to the complex environmental noise in the construction site, the original monitoring data often have large volatility, which is difficult to intuitively and accurately reflect the real state of the measured entity. In this section, according to the characteristics of noise and monitoring data, Kalman filtering method is used to reduce the noise of sensor monitoring data and artificial monitoring data, and the noise reduction effect is compared.

3.1.1. Noise reduction method

Kalman filtering method is simple and reliable, and can be used for real-time data denoising [12,13]. The core idea of the Kalman filter method is to use the process estimation at the previous moment and the measurement estimation at the current moment for data fusion to obtain the optimal estimation of real value of the data at current moment, and then to iteratively calculate the data at the next moment [13]. Process estimation and measurement estimation are shown in equations (1) and (2):

\[ X_k = A \cdot X_{k-1} + B \cdot u_{k-1} + \omega_{k-1} \]  
\[ Z_k = H \cdot X_k + \nu_k \]  

In the formula, \( X_k \) is the real value of the current moment, \( X_{k-1} \) is the real value of the previous moment, \( u_{k-1} \) is the control quantity of the previous moment, \( \omega_{k-1} \) is the process error of the previous moment, \( \nu_k \) is the measurement error of the current moment, and \( A, B, H \) are the calculation matrices. Due to the high frequency of real-time monitoring, the process variation at the current time and the previous time is negligible relative to the process error. Therefore, the optimal estimation of the monitoring data at the previous time can be used as a priori estimation of the current time monitoring data, as shown in formula (3):

\[ \hat{X}_k = \hat{X}_{k-1} \]

In the formula, \( \hat{X}_k \) is the prior estimate of the current moment and \( \hat{X}_{k-1} \) is the optimal estimate of the previous moment. Similarly, the prior estimation of the covariance of the process error at the current moment is shown in equation (4):

\[ P_\epsilon = P_{\epsilon-1} + Q \]

In the formula, \( P_\epsilon \) is the prior estimate of the current time process error covariance and \( P_{\epsilon-1} \) is the optimal estimate of the previous time process error covariance. The Kalman increment represents that the process estimation and measurement estimation are integrated into the optimal estimation of the current time parameters in a certain proportion. The ratio is determined by the covariance of the process error and the covariance of the measurement error, as shown in Equation (5):

\[ K_k = \frac{P_{\epsilon-1} \cdot H^T}{H \cdot P_{\epsilon-1} \cdot H^T + R} \]

In the formula, \( R \) is the covariance of the measurement error. The optimal estimates of the current time can be obtained by using the process prior estimates and the measurement estimates as Kalman increments, as shown in formula (6):

\[ \hat{X}_k = \hat{X}_{\epsilon-1} + K_k \cdot (Z_k - H \cdot \hat{X}_{\epsilon-1}) \]

Finally, the covariance of process errors is updated, as shown in formula (7):

\[ P_\epsilon = (I - K_k \cdot H) \cdot P_{\epsilon-1} \]

The above is a complete set of iterative process of Kalman filtering method. In general, the process error and measurement error are in line with the normal distribution with the expectation of 0. According to the accuracy of monitoring equipment and expert experience, the appropriate initial value of process covariance \( P_0 \) and monitoring error covariance \( R \) are selected. The initial value of monitoring data is taken as \( \hat{X}_0 \), and the data can be continuously predicted and corrected by repeating
the process from Formula (3) to (7), so as to smooth the data and reduce the influence of measurement noise on data analysis.

3.1.2. Preprocessing results
Taking the monitoring data of f111 fault area in the second tunnelling section as an example, the opening amount of segment joint, bolt stress, longitudinal convergence and tunnel settlement of 360 ring are denoised. Among them, joint opening and bolt stress are high frequency monitoring, which were recorded twice an hour. And vertical contraction and tunnel settlement are low frequency monitoring, which were recorded twice a day. So four days of high frequency monitoring data (about 200 records) and twenty-five days of low frequency monitoring data (about 50 records) were processed, which could reflect overall trends of monitoring. The processing results are shown in Figures 6-9:

![Joint opening data preprocessing](image1)

**Figure 6.** Joint opening data preprocessing

![Bolt stress data preprocessing](image2)

**Figure 7.** Bolt stress data preprocessing

![Vertical contraction data preprocessing](image3)

**Figure 8.** Vertical contraction data preprocessing

![Tunnel settlement data preprocessing](image4)

**Figure 9.** Tunnel settlement data preprocessing

The results shows that Kalman filtering method has better denoising effect for high frequency monitoring data than that for low frequency monitoring data. The processing of abrupt data points has strong lag, and some denoised data points have shifted for low frequency data. The reason may be related to the process estimation error. When the monitoring frequency is low, the process variation of the two monitoring cannot be ignored compared with the estimation error, so the deviation is produced.

3.2. Data fusion
Multi-source perception system provides comprehensive and diverse monitoring data for risk management. Effectively using multi-source data to reflect the service state of the structure, and combining with the risk evolution mechanism for multi-factor risk assessment, data fusion method is needed. In this section, the risk of tunnel lining leakage is analysed. According to the multi-source monitoring data, the D-S evidence theory method is used for data fusion, and a multi-source data fusion analysis model for the probability of leakage risk is established.

3.2.1. Risk analysis
Lining leakage is one of the important common risks of shield tunnels. Small-scale water leakage will accelerate the corrosion rate of bolts and bring hidden dangers of tunnel electricity. Large-scale water
leakage may lead to failure of tunnel structure, casualties and other serious consequences. Leakage risk mainly comes from geological and structural aspects, as shown in figure 10:

Figure 10. Risk analysis of water leakage of tunnel lining.

Geological aspects include two risk factors: First, the strong weathered strata in the fault fracture zone area, resulting in poor grouting quality behind the lining wall, prone to structural deformation. Secondly, some strata in the water area tunnelling section are rich in water, water pressure is large, and water gushing is easy to occur. The geological risk factors can be directly obtained by preliminary survey and design data. The structure also includes two risk factors: First, the lining joint is opened, resulting the failure of the waterproof belt and the leakage of water. This factor can be quantitatively characterized by the monitoring data of the opening of the circumferential joint of the lining and the longitudinal convergence monitoring data. Second, longitudinal uneven settlement, resulting the damage of adjacent ring joints, which was prone to water leakage. This factor can be quantitatively characterized by tunnel settlement monitoring data and bolt stress monitoring data.

3.2.2. Data fusion method
D-S evidence theory has an adaptive and flexible algorithm, which can carry out uncertainty reasoning, fuse heterogeneous data from multiple sources full of randomness, fuzziness and uncertainty, and can intuitively express the degree of uncertainty \(^{[14-16]}\). The method framework is shown in figure 11:

Figure 11. Framework of D-S evidence theory.

Main problem and all its possible results should be determined firstly, which are mutually exclusive and discrete, as shown in formula (8). The set of all subsets of the result set is called the power set of the identification framework, as shown in equation (9):

\[
\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\} \tag{8}
\]

\[
2^\Theta = \{\phi, \{\theta_1\}, \{\theta_2\}, \{\theta_1, \theta_2\}, \{\theta_1, \theta_3\}, \ldots, \{\theta_1, \theta_2, \ldots, \theta_n\}\} \tag{9}
\]

Evidence is the basis for judging the main problem, which can quantitatively give the probability distribution of each subset in the power set of the identification framework, namely BPA. BPA can be seen as a mapping from each element of \(2^\Theta\) to a real number from 0 to 1, denoted by \(m\), which satisfying Equation (10):

\[
\sum_{A \in 2^\Theta} m(A) = 1 \tag{10}
\]

BPA can be calculated by statistical survey method, fuzzy membership method, etc \(^{[15]}\). Fuzzy membership method is to use membership function in fuzzy mathematics to calculate the probability
distribution of each level of evidence, in which the membership function in the form of exponential is the most commonly used. According to the upper and lower limits of each interval, the calculation process is as follows:

\[ f_i(x_i) = \exp\left[-\frac{(x_i - a_{ij})^2}{b_{ij}}\right] \]  

(11)

\[ a_{ij} = \frac{x_{ij}(L) + x_{ij}(R)}{2} \]  

(12)

\[ b_{ij} = \frac{x_{ij}(R) - x_{ij}(L)}{2} \]  

(13)

In the formula, \( x_{ij}(L) \) is the lower limit of the interval, \( x_{ij}(R) \) is the upper limit of the interval and \( f_i(x_i) \) is the membership function. In order to make the judgment of the problem to be identified more accurate, multiple evidences are often introduced in the judgment process, and the judgment results of each evidence are fused. In DS evidence theory, the most commonly used method is orthogonal evidence fusion, whose rules are defined as follows (14) and (15):

\[ K = \sum_{r \cap A_i \neq \emptyset} \prod_{i=1}^{n} m_i(A_r) \]  

(14)

\[ m(A) = \frac{\sum_{r \cap A_i = A} \prod_{i=1}^{n} m_i(A_i)}{1 - K} \]  

(15)

K is the conflict coefficient, which quantitatively characterizes the contradiction between the two fused evidence judgment results. The greater the K, the greater the contradiction between the evidence, and the worse the reliability of the positive blending law. Therefore, it is easy to use the weighted average method \(^{[16]}\) when encountering evidence fusion with greater contradiction, and its rule definition is as follows (16) to (18):

\[ d_r = \sum_{i=1}^{n} \sqrt{\sum_{j=1}^{m} (m_i(A_j) - m_i(A_j))^2} \]  

(16)

\[ \omega_r = \frac{1/d_r}{\sum_{r=1}^{m} 1/d_r} \]  

(17)

\[ m(A) = \sum_{i=1}^{N} (\omega_i \times m_i(A)) \]  

(18)

In the evidence fusion, the appropriate threshold is selected for the conflict coefficient K, the weighted average method is used for the evidence fusion with high conflict, and the positive blending method is used for the evidence fusion with low conflict, and a more accurate final judgment result is obtained.

3.2.3. Risk factor fusion

The probability of leakage risk in this project is the problem to be identified. According to the relevant norms of tunnel and underground engineering risk management, the risk probability is divided into four levels. Each risk factor in the risk analysis module can be used as the basis for judgment. Since the risk factors can be quantitatively evaluated by geological exploration information and monitoring data, based on relevant research and expert experience \(^{[17,19]}\), each risk factor and monitoring data also be divided into four levels, as shown in table 2:

Table 2. Risk factor classification of leakage risk.

| Risk factor / Monitoring data | Level I | Level II | Level III | Level IV |
|------------------------------|---------|----------|-----------|----------|
|                              |         |          |           |          |
Based on the geological information and monitoring data of the 545th ring in the f27 fault area of the second tunnelling section, the fuzzy membership method is used to obtain the risk factors BPA as shown in table 3:

**Table 3. Monitoring data and BPA of each risk factor of ring 545.**

| Risk factor / Monitoring data | Data       | Level I | Level II | Level III | Level IV |
|------------------------------|------------|---------|----------|-----------|----------|
| Vertical contraction         | 27.44mm    | 0       | 0.5751   | 0.3862    | 0.0387   |
| Tunnel settlement            | 13.28mm    | 0       | 0.6488   | 0.3397    | 0.0115   |
| Circumferential joint opening| 2.34mm     | 0       | 0.3101   | 0.5922    | 0.0977   |
| Bolt stress                  | 474.33MPa  | 0       | 0.1305   | 0.6021    | 0.2674   |
| Water pressure               | 0.45MPa    | 0       | 0.0243   | 0.6633    | 0.3124   |
| Weathering degree            | strongly weathered | 0 | 0.0013 | 0.7477 | 0.2510 |

The threshold of conflict coefficient K is 0.4, and the evidence fusion of risk factors under two categories is carried out by positive blending method and weighted fusion method. The fusion results are shown in table 4. According to relevant research and expert experience, the weighted average method is used to fuse the characteristics of the two types of evidence of geology and structure, and the fusion coefficient is \{0.6, 0.4\}. Finally, the probability distribution results of leakage risk in 545 rings are shown in table 5.

**Table 4. BPA of geological and structural factors.**

| Risk factor | Level I | Level II | Level III | Level IV |
|-------------|---------|----------|-----------|----------|
| Geographical | 0       | 0.0001   | 0.8634    | 0.1365   |
| Structural  | 0       | 0.2637   | 0.7266    | 0.0087   |

**Table 5. BPA of leakage risk of 545 ring.**

| Risk       | Level I | Level II | Level III | Level IV |
|------------|---------|----------|-----------|----------|
| Water leakage | 0       | 0.1055   | 0.8087    | 0.0858   |

It can be seen from the analysis that the occurrence probability of water leakage in the 545 ring is likely to be at the third level. Compared with the field practice, small-scale water leakage occurs at the circumferential and longitudinal segment joints of the side wall of the ring at 3 o’clock. It can be seen that the analysis model is in good agreement with the actual situation.

4. Conclusions
The main conclusions of this study are as follows:
Aiming at the fluctuation problem of monitoring data, Kalman filtering method is used for noise reduction, and good denoising effect is achieved. Therefore, Kalman filtering method has good application value in sensor data pre-processing.

- The noise reduction effects of monitoring data with different monitoring frequencies are compared, and it is found that the noise reduction effect of high-frequency monitoring data is better than that of low-frequency monitoring data. The reason may be that the neglect of process error by prior estimation method has a great influence on low-frequency monitoring.
- The D-S evidence theory method is used to conduct multisource data fusion analysis on the leakage risk of shield tunnel. The orthogonal fusion and weighted fusion are used for the comprehensive evidence fusion. The quantitative fusion of the risk factors in the data layer and the feature layer is realized, and a multisource analysis model of the leakage probability is formed. The analysis results are verified in practical engineering.

Acknowledgments

The work presented in this article was supported by Water Allocation Project of Pearl River Delta (CD88-GC02-2020-0038) Guangdong Yuehai Pearl River Delta Water Supply Co., Ltd and the National Natural-Science Foundation of China (Grant No. 52078377) and Shenzhen University.

References

[1] Bennett P, Kobayashi Y and Soga K 2010 J. Geotechnical Engineering. 163 147
[2] Mohamad H, Soga K and Bennett P 2012 J. Geotechnical and Geoenvironmental Engineering. 138 957
[3] Wang Z, Wang J and Sui Q 2016 J. Sensors and Actuators a-Physical. 250 145
[4] Khaleghi B, Khamis A and Karray F 2013 J. Information Fusion. 14 28
[5] Kumar M, Garg D and Zachery R 2007 J. IEEE Sensors Journal. 7 723
[6] Nguyen L and Nestorovic T 2016 J. Computing in Civil Engineering. 30 04015025
[7] Paneiro G and Rafael M 2020 J. Underground Space. 3 134
[8] Panou K and Sofianos A 2002 J. unnnelling and Underground Space Technology. 17 195
[9] Peng M and Zhang L 2012 J. Natural Hazards. 64 903
[10] Picandet V, Khelidj A and Bellegou H 2009 J. Cement & Concrete Research. 39 537
[11] Yin J and Huang H 2015 J. Aerospace Engineering. B 401
[12] Lin Y 1987 J. China Civil Engineering Journal. 16 714
[13] Wang L, Li Y and Liu W 2006 J. Journal of Xi’an University of Science and Technology. 26 353
[14] Zhou H, Zhao Y, Shen Q and Yang L 2020 J. Automation in Construction. 111 236
[15] Shahanshti S and Razavi S 2011 J. Journal of Construction Engineering & Management. 137 863
[16] Zhang L and Ding X. 2017 J. Journal of Aerospace Engineering. B 4015002
[17] Zhang Z Huang M and Wang W 2013 J. Tunnelling and Underground Space Technology. 38 244
[18] Santos J Cremona C and Orcesi A 2017 J. Journal of Structural Engineering. 143 04016162
[19] Zhang L, Wu X and Qin Y 2016 J. Risk Analysis. 36 278