Research on pre-processing of geotechnical engineering safety monitoring data

Zhihao Fu[1]*, Yuting Fu[2] [a]

[1] China Water Resources Pearl River Planning, Surveying & Designing Co., Ltd, Guangzhou, Guangdong, China
[2] College of Industrial Design, Hubei Institute of Fine Arts, Wuhan, Hubei, China

*5753079@qq.com, [a]994234868@qq.com

Abstract. Feedback-based construction methods using safety monitoring data are widely used in geotechnical engineering. However, improving the quality of monitoring data and conducting accurate calculations and analysis remain challenging areas that require further research. This article introduces the principles of wavelet transform, the processing flow of wavelet packet transform, data anomaly detection, and monitoring data denoising methods. The effectiveness of the data processing is demonstrated through a case study in engineering. The results show that the quality of processed monitoring data is significantly improved, enabling a more direct and efficient reflection of data change patterns. Furthermore, this article infers the deformation state of the rock mass and the composition of the rock mass structure in reverse by starting from the changing pattern of monitoring data, taking examples of the multi-point displacement meter and anchor stress meter, combined with the general law of surrounding rock deformation during tunnel excavation and basic geological knowledge. This process provides basic information such as the depth of the loose failure zone and the range of rock mass parameters for subsequent efficient feedback analysis. This article's proposed data processing and analysis ideas have practical significance and are operable in the automation processing, intelligent analysis, and decision-making of monitoring data. They also provide experience and references for related research and application work in this field.

Keywords: geotechnical engineering, safety monitoring, wavelet analysis, wavelet packet transform, data pre-processing

1 Introduction

Intelligent water conservancy construction is a critical factor in promoting the high-quality development of water conservancies. Its main feature is the "four predictions" function, which includes prediction, warning, simulation, and contingency planning. The "four predictions" prediction component relies on analyzing and interpreting historical data to identify patterns and trends, allowing for early prediction of specific

D. Li et al. (eds.), *Proceedings of the 2023 9th International Conference on Architectural, Civil and Hydraulic Engineering (ICACHE 2023)*, Advances in Engineering Research 228,
https://doi.org/10.2991/978-94-6463-336-8_12

events. The accuracy, effectiveness, and regularity of historical data, particularly monitoring data, are critical to the function's success. Therefore, due to the influence of observation conditions, comprehensive research and application of monitoring data pre-processing techniques are essential before analysis. Monitoring data is the most fundamental and critical information element of the "four predictions" function, making it a significant concern for all parties involved in the project.

The application of feedback construction technology in underground engineering is increasingly widespread. The general method for analyzing the stability of underground caverns based on monitoring data is as follows: the rock parameters are obtained through the inversion analysis (or parameter inversion) of tracking data, and then numerical simulation analysis is performed based on the rock parameters, predicting the subsequent deformation and stability of the rock mass, and guiding construction design. Therefore, pre-processing monitoring data is the first and most crucial step of feedback construction. This article does not discuss the inversion analysis and related optimization algorithms and focuses on the pre-processing of monitoring data. Currently, there are three main ways to pre-process and analyze monitoring data: mechanism analysis, mathematical and physical analysis, and a hybrid analysis based on the two. Among them, mechanism analysis infers the rock mass's deformation status and structure based on various monitoring data and the general laws of rock deformation during excavation combined with basic geological knowledge. The mathematical and physical analysis mainly includes detecting and interpolating singular values in monitoring data, denoising, and pattern recognition of monitoring signals. This article studies and explores the pre-processing of monitoring data from the perspectives of mathematical and physical analysis and mechanism analysis, then summarizes the related processing flow. This article can provide experience and references for relevant research and application work.

2 Pre-processing of Monitoring Data Based on Wavelet Packet Decomposition

2.1 Basic Principles of Wavelet Transform

The information obtained from engineering safety monitoring can be characterized as a signal that varies over time or space. This signal includes both useful information and errors or noise. To effectively remove errors and extract essential features, conducting thorough research on monitoring data analysis is necessary. Safety monitoring data can be represented as a digital signal sequence comprised of various frequency components. The significant low-frequency components typically represent the trend of the valuable signal, while high-frequency components often indicate sudden changes and anomalies. Therefore, signal processing techniques can be utilized for analyzing monitoring data.

Wavelet analysis overcomes the shortcomings of the short-time Fourier transform in single resolution and has the characteristic of multi-resolution study, with the ability to represent local information of signals in both time and frequency domains. Both the time window and the frequency window can be dynamically adjusted according to the specific form of the call. Generally, a lower time resolution can be used in the low-frequency part of the signal (where the signal is relatively stable). In contrast, a higher frequency resolution can be used in the high-frequency region (where the frequency changes are insignificant) in exchange for accurate time positioning. Because of these characteristics, wavelet analysis can detect transient components in standard signals and display their frequency components. It is known as a "mathematical microscope" and is widely used in various time-frequency analysis fields. The wavelet packet transform is based on the wavelet transform and can achieve uniform frequency band division with better time-frequency characteristics. Therefore, the wavelet packet has a higher application value in signal processing. Due to space limitations, this article does not elaborate on the relevant theory and formulas of the wavelet transform. In the case of multi-resolution wavelet transform, low-frequency (i.e., scale space) is decomposed, and high-frequency is retained in each step. In contrast, wavelet packet transform simultaneously decomposes low-frequency and high-frequency in each transformation, resulting in more precise frequency bands. The transformation processes of the two methods are illustrated in Figure 1.

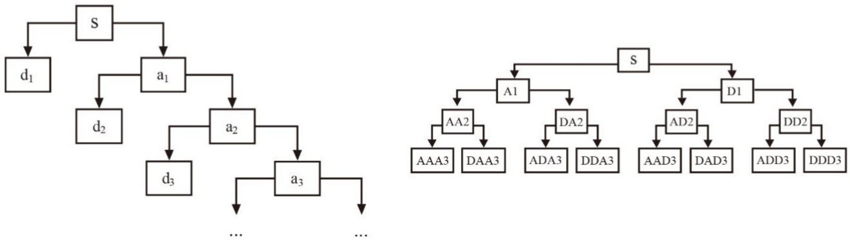

Fig. 1. Schematic of wavelet decomposition and wavelet packet decomposition

2.2 The Basic process of wavelet packet transform

Engineering safety monitoring data presents challenges due to its characteristics of multiple equipment types, large and varied data volume, and long duration. Considering the current demand for intelligent simulation and precise decision-making in engineering, there is a crucial practical need to quickly process large amounts of monitoring data, improve data quality, and conduct related research. This paper proposes a workflow for monitoring data processing using wavelet packet decomposition, as depicted in Figure 2. The workflow was implemented using the PyWavelets package based on the Python language. After thorough application verification, the analysis workflow was reasonable, and the program produced accurate results.

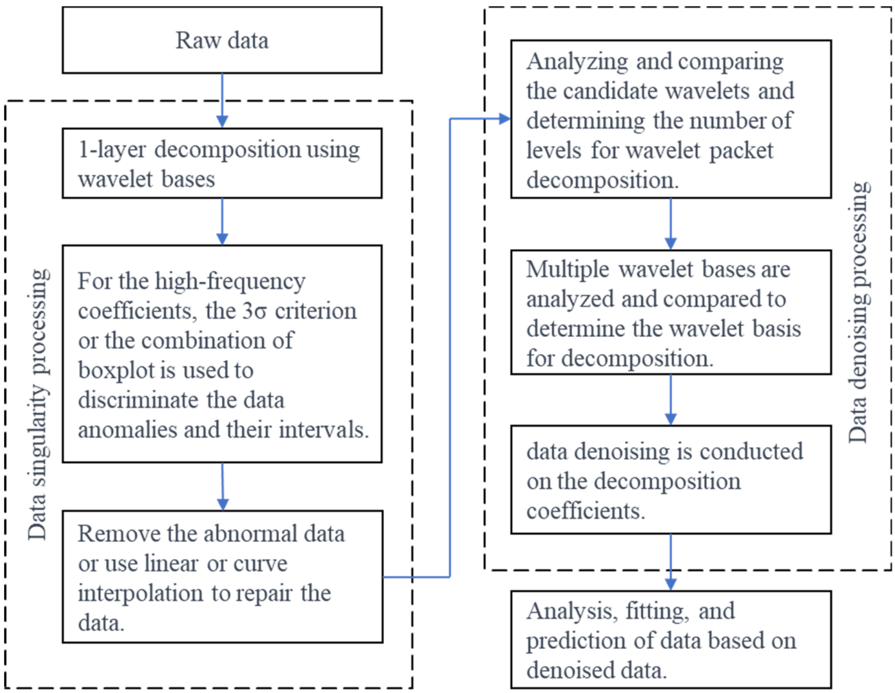

Fig. 2. Wavelet packet decomposition workflow diagram of monitoring data

Abnormal data detection.

In the information obtained after wavelet decomposition, the information about the trend of changes is distributed in the low-frequency coefficients. The information about sudden changes is mainly reflected in the modulus maximum value of high-frequency coefficients. Since wavelet analysis can identify the frequency changes of data and locate the position of its anomalies, it can diagnose the time, type, and magnitude of abnormal information by detecting the points of modulus maximum value. For abnormal data processing, the high-frequency noise is generally assumed to follow a Gaussian normal distribution so that the 3σ rule can be used for elimination. The box plot method can also be used to obtain the distribution law of the data statistically, combined with the 3σ control. For engineering safety monitoring data, the appearance of abnormal data often contains essential information, such as the anchor rod stress meter of underground chambers and the multi-point displacement meter around the room. During layered blasting excavation operations, they will be affected by blasting vibrations and the redistribution of surrounding rock stress, resulting in specific sudden changes in the measured values. Therefore, when detecting and processing data anomalies, objective analysis of data changes is also required to avoid the loss of important information. After determining the above information, data deletion or interpolation can be used to repair abnormal data.

Denoising of Monitoring Data.

Actual measurement data inevitably contains noise (error), and noise in the measured values will distort the measured values and the measured results. The company of noise significantly impacts the analysis of applied research data. It may affect parameter estimation, cause large residuals, and affect the safety monitoring model's fitting effect and extrapolation performance. Therefore, when applying engineering safety monitoring data, the noise in the data should be judged and processed to distinguish between true and false information so that the analysis work is based on reliable data. The usual wavelet decomposition method for denoising research first selects commonly used wavelet bases for time series of measured values based on experience and then sets a threshold to choose an appropriate threshold function to denoise the low-frequency or high-frequency coefficients. Finally, the processed coefficients are reconstructed into the original signal, then error judgment and subsequent analysis work are carried out. For the denoising of monitoring data, the wavelet bases are usually selected from the db or sym families of wavelets. There are many rules for determining the threshold, which means extracting weak proper signals from high-frequency information without eliminating appropriate high-frequency characteristic signals as noise signals during the denoising process. The threshold can generally be set to:

$$\lambda = \sigma\sqrt{2\,log(n)}. \tag{1}$$

In the Equation, n is the number of high-frequency coefficients at the corresponding decomposition level. Since the standard deviation σ of actual noise coefficients is generally unknown, the absolute standard deviation of high-frequency coefficients at the first level (i.e., the finest scale) of wavelet decomposition can be used to estimate σ. The threshold function is selected for the high-frequency coefficients of each layer of wavelet decomposition and is generally quantized using a soft threshold function: coefficients below the threshold λ are set to 0, while coefficients greater than or equal to λ are reduced by λ, which can remove the noise component concentrated in the high-frequency coefficients, while the hard threshold function preserves high-frequency coefficients bigger than the threshold. Other types of threshold functions can also be chosen according to the specific needs.

2.3 Wavelet Packet Transform Case Study Verification

Based on the monitoring data obtained from a multi-point displacement sensor in an underground factory building, a 1-level wavelet decomposition using the sym4 wavelet was performed. The resulting data includes the original data and the high-frequency, low-frequency, and noise components after decomposition, as shown in Figure 3. The selection of the sym4 wavelet was based on prior experience.

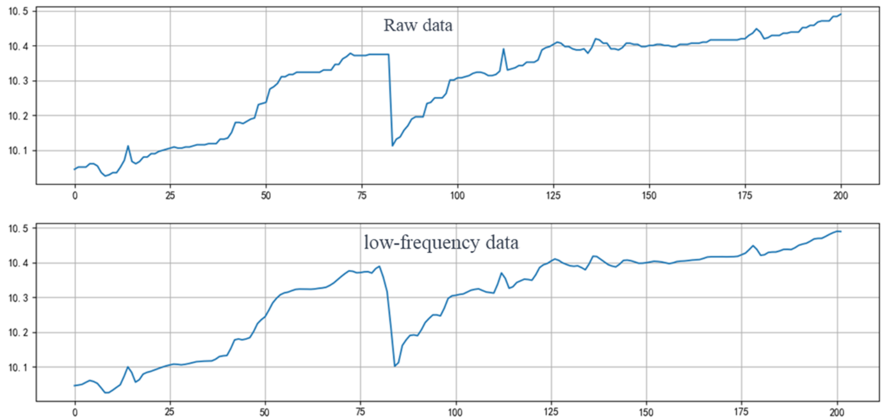

Fig. 3. Raw data and low-frequency data (trend of change)

After wavelet packet decomposition, it compares the original monitoring data sequence with the reconstructed low-frequency data. It shows that the wavelet packet decomposition has better preserved the original data trend. Boxplot and statistical analysis of the reconstructed high-frequency data yield key parameters shown in Figure 4, where the high-frequency part mainly reflects data noise. By examining the kernel density plot, it can be intuitively observed that the noise exhibits Gaussian typical distribution characteristics.

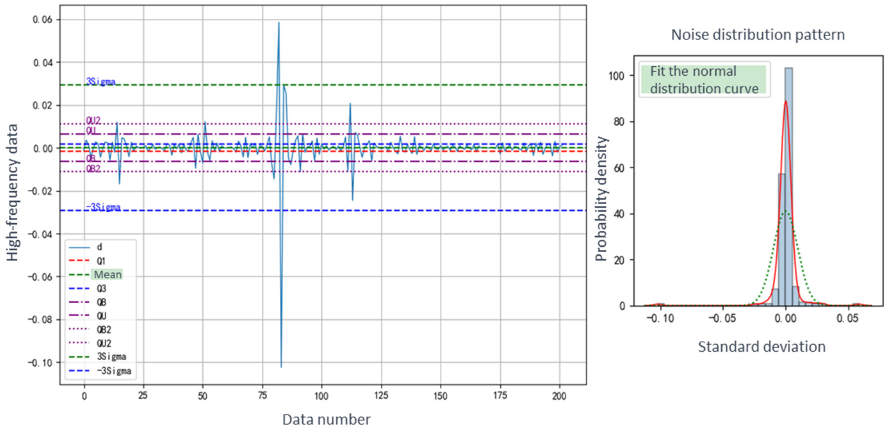

Fig. 4. Distribution of high-frequency data and noise after single-layer wavelet packet decomposition for multi-point displacement monitoring

The commonly used 3σ criterion based on the normal distribution assumption is to calculate the mean and standard deviation of the data, which has certain limitations due to the small robustness of the mean and standard deviation and the significant influence of outliers. The boxplot method for handling outlier data can be introduced in practical applications. Compared with the 3σ criterion, the boxplot is more intuitive in showing the distribution shape of the data and is based on quartiles and interquartile ranges, which has certain robustness and provides a more objective approach to identifying data anomalies. In practice, the two methods can be used together. By comparing the original data and noise distribution, the range of abnormal data in the data sequence can be visually identified, and methods such as interpolation based on intrinsic physical connections or mathematical techniques such as linear interpolation, Lagrange interpolation, and polynomial curve fitting can be used to repair the data sequence.

For data denoising, the high-frequency coefficients of each wavelet decomposition level are processed using the global threshold calculated above. An appropriate threshold function (soft threshold function, hard threshold function, or other custom threshold function) is selected for processing. After processing, regular wavelet packet reconstruction can be used to achieve signal denoising. In practical applications, another feasible denoising method is to analyze the energy characteristics of the signal after wavelet packet decomposition, that is, to directly remove wavelet packet coefficients with relatively small energy characteristics and then perform the reconstruction. Since wavelet packet decomposition preserves the original information completely, this method only retains the essential part of the data information, which is more direct, convenient, and physically meaningful for identifying and analyzing the change law of engineering monitoring data.

3 Analysis of Mechanism for Monitoring Data Changes

The monitoring data undergo mathematical analysis, data anomaly identification, and noise reduction processing, resulting in a significant improvement in data quality and an increase in regularity compared to the original data. However, further mechanism analysis is required to determine and select the numerical model calculation parameters based on the monitoring data change pattern, which involves combining the general principles of rock deformation during tunnel excavation and basic geological knowledge to infer the deformation state of the rock mass and the composition of the rock structure inside the rock mass.

3.1 Multi-point displacement measurement and analysis

Multi-point displacement measurement instruments increase the density of measurement points within a specific range, allowing for the accurate measurement of relative displacement changes between multiple points. As a result, they can provide relatively accurate information on the composition of rock structures. This article does not introduce the construction and measurement principles of multi-point displacement meters but only provides a theoretical analysis of measurement results. The measurement points of a multi-point displacement meter are generally 4-5. The research in this article uses a 4-point multi-point displacement meter as an example.

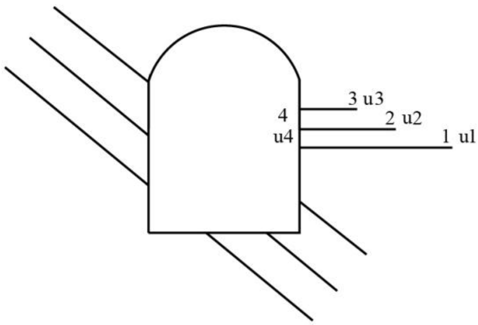

Fig. 5. Schematic diagram of multi-point displacement meter

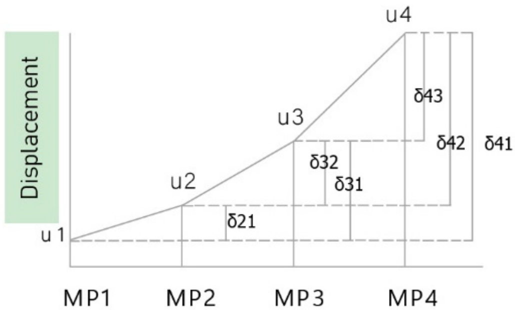

Fig. 6. The relationship between absolute displacement and relative displacement of monitoring points

Assume that the multi-point displacement meter is arranged as shown in Figure 5. Points 1, 2, and 3 are fixed points on the measuring rod, and point 4 is a measuring point on the surface of the excavation. During the excavation process, stress release will cause deformation in the direction towards the cavity, resulting in absolute displacement values u1, u2, u3, and u4 along the measuring rod direction for each point. In general, the total displacement values between the measurement points satisfy Equation (2), where the displacement at the cavity surface is the largest and gradually decreases with increasing depth.

$$u4 > u3 > u2 > u1 \quad (2)$$

In practical monitoring, the values measured at each measuring point are the relative displacement values δ41, δ42, and δ43 between the fourth point in the figure and the fixed points on other measuring rods in the direction of the measuring rod. Equation (3) and Figure 6 represent the relative and absolute displacement relationship.

$$\delta_{41} = u4 - u1;\ \delta_{42} = u4 - u2;\ \delta_{43} = u4 - u3 \quad (3)$$

From the figure, if the distance between point 1 and the tunnel wall is far enough or u1 is very small, it can be assumed that u1≈0. Using the relationship between total and relative displacement, it can be analyzed that the deformation of the surrounding tunnel rock and the rock mass's structure. Since the actual monitoring may involve complex situations, here only compare, and explore several common relationships (assuming that the absolute displacement of the tunnel wall point is the largest) in Table 1.

Table 1. Typical Analysis of Relative Displacement Relationships

NO.	Relative displacement relationship	Absolute displacement relationship	Description
1	$\delta 41 \approx \delta 42 > \delta 43$	$u1 \approx u2 \approx 0 < u3 < u4$	The loosening range of the surrounding rock mass of the tunnel after excavation will not exceed the depth of Point 2.
2	$\delta 41 \approx \delta 42 > \delta 43$	$u1 \approx u2 \neq 0 < u3 < u4$	The rock mass between points 1 and 2 is relatively intact and can be considered to deform synchronously. The rock mass beyond points 2 to the tunnel periphery has poor properties.
3	$\delta 41 \approx \delta 42 < \delta 43$	$u3 < u1 \approx u2 < u4$	There may exist joint fractures or fractured zones between point 2 and point 3. The rock mass between points 1 and 2 and between points 3 and 4 is relatively intact.
4	$\delta 41 < \delta 42 < \delta 43$	$u3 < u2 < u1 < u4$	There may be fractured zones, weak layers, or joints between points 1 and 2, and between points 2 and 3, with larger displacements. Still, their effects did not propagate, indicating that the surrounding rock properties between point 3 and the tunnel are relatively good.
5	$\delta 41 < \delta 43 < \delta 42$	$u2 < u3 < u1 < u4$	There may be a fractured zone between points 1-2 and 2-3, and point 2 is subject to relatively small compressive and tensile forces from points 1 and 3, respectively. Therefore, the displacement of point 2 is minimal, indicating that it is likely located in a fractured zone with poor connectivity to the surrounding rock mass.
6	$\delta 41 \approx \delta 42 \approx \delta 43$	$u1 \approx u2 \approx u3 < u4$	The rock mass between points 1 and 3 has a relatively intact property, while the rock mass between points 3 and 4 has poor properties, leading to synchronous deformation of measuring points 1, 2, and 3.

Table 1 only briefly analyses several possible situations in the monitoring process of a 4-point multi-point displacement meter. In the actual monitoring process, other more complex cases may also occur. Still, compared with the engineering reality, most of them can obtain conclusions that are more consistent with fact. In addition to comparing the displacement of each measurement point, the convergence of rock deformation around the tunnel and the stability of the tunnel can also be judged based on factors such as the displacement rate of each measurement point. By analyzing the displacement changes, it can be obtained that information on the distribution of the rock types around the tunnel and the initial depth of the surrounding rock loosening zone.

An underground power station project excavated the main hall chamber in nine phases. Two multiple-point displacement meters on the upstream side wall of the monitoring section at 0+139.0 of the main halls were analyzed. Among them, C6B-CF-IV-M-06 is located at an elevation of 1008.7 m and is at the upper part of the wall, while C6B-CF-IV-M-08 is situated at a height of 998.78 m and is at the middle part of the wall. The displacement value changes of the two devices are shown in Figure 7.

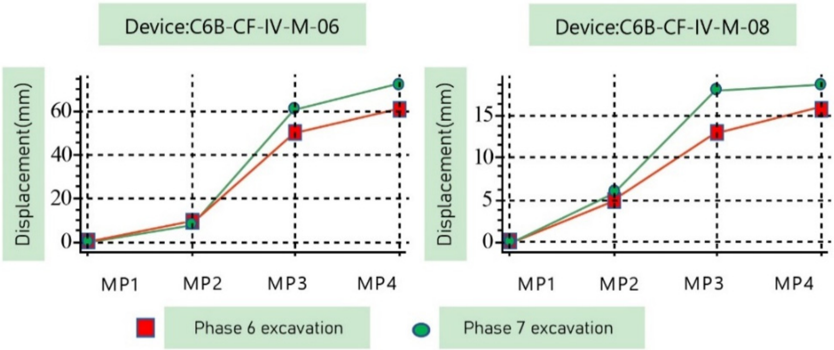

Fig. 7. Displacement change curve of the upstream wall of an underground factory

As the figure shows, the side wall's displacement change conforms to the first case in Table 1. The rock mass deformation between the third measuring point of the side wall and the surrounding rock is relatively large, indicating that the shallow rock mass of the side wall has significantly been disturbed during construction and excavation. In terms of displacement value, the cumulative displacement value of C6B-CF-IV-M-06 is much higher than that of C6B-CF-IV-M-08, indicating that on the one hand, the upper rock mass of the side wall is still disturbed as the excavation goes downward. On the other hand, the displacement of measuring point 2 of device M-06 is more significant, indicating that there may be deep deformation. By comparing with the excavation situation on site, a fault was exposed at station 0+139.0, so in the subsequent construction and excavation, the impact of construction blasting on the surrounding rock should be controlled, and specific support measures should be added to ensure the stability of the chamber at this location.

3.2 Monitoring and Analysis of Anchor Stress Meters

The focus of anchor monitoring is on two issues: one is the magnitude of anchor stress values, and the other is the relationship between the depth of the failure zone around the hole and the length of the anchor. The anchor stress meter monitoring data analysis is consistent with the multi-point displacement meter data analysis. By comparing the values of anchor stress meters at different depths, the distribution of structural surfaces around the hole and the loose failure zone of the surrounding rock can be judged. Suppose the anchor stress in the monitoring data is close to or exceeds the tensile strength of the anchor. In that case, it can be determined that the surrounding rock deformation or disturbance is significant, and support should be increased promptly.

On the other hand, if the depth of the loose failure zone around the hole is equal to or greater than the length of the existing anchor, the surrounding rock may undergo deep deformation. The current anchor support cannot meet the requirements of surrounding rock stability. It is necessary to consider increasing the anchor support appropriately to limit the further development of the loose failure zone.

For example, during a specific period in the construction process of an underground workshop, the number of practical points of the anchor stress meter was 410, of which 18% of the measurement points had monitoring values greater than 200 MPa. Among these points, 75% of the measurement points had a depth of ≤5 m and 67% had a depth of ≤3 m. The points with monitoring values greater than 300 MPa accounted for 10% of the effective total measurement points, and among these points, 84% of the measurement points had a depth of ≤5 m and 74% of the measurement points had a depth of ≤3 m. Therefore, it can be seen from the comparison that most of the anchor stress during the current period is below 200 MPa, and the depth of the points with relatively large anchor stress measurements is less than 3 m, indicating that the deformation of the surrounding rock of the workshop belongs to shallow deformation. The depth of surrounding rock disturbance is basically within the range of 3 m.

Using the above method to analyze the monitoring data of the multi-point displacement meter and the anchor stress meter, the parameter range of the current engineering rock mass and the depth of rock mass disturbance and damage can be roughly determined. In the feedback analysis of information, not only should the calculated value and the monitoring value be close in numerical value, but more importantly, the deformation law of the analysis results should conform to reality. Only in this way can the prediction be in line with reality and indeed used for engineering. Due to the large amount and variety of engineering monitoring data, we usually need to establish a dedicated monitoring database for statistical data analysis in the application process. The wavelet packet decomposition and reconstruction mentioned earlier and the identification of the numerical variation laws of each measurement point can be implemented using programming packages such as PyWavelets, Scipy, or other programming languages to achieve batch and rapid monitoring data processing.

4 Conclusion

The data quality can be improved by conducting mathematical analysis on the deformation patterns of safety monitoring data and processing them through anomaly recognition, correction, and noise reduction, laying a good foundation for subsequent analysis of data patterns.

The relationship between displacement monitoring data and rock mass shape was theoretically analyzed, and by analyzing the data, the deformation pattern of the rock mass can be objectively understood, and basic information such as the depth of the loose damage zone and the range of rock mass parameters can be obtained. The stress monitoring data of the anchor rods were statistically analyzed based on stress distribution and distribution depth. Through the analysis of the anchor rod stress, the deformation pattern of the surrounding rock mass can be perceived, and the deformation pattern can be verified with the displacement pattern to gain a comprehensive understanding of the engineering safety status, which has practical significance for improving the quality and efficiency of information feedback in construction.

This paper only focuses on the research and discussion of monitoring data from the perspectives of mathematics and mechanism, using the example of multi-point displacement and anchor rod stress meters. Relevant processing procedures were summarized, and case studies were provided. The data processing and analysis ideas proposed in this paper have practical significance and operability in the automation of monitoring data processing and intelligent analysis and decision-making, providing experience and reference for related research and application work.

References

1. Qiu Minghua. Application of data mining technology in water conservancy project safety monitoring management[J]. Water Conservancy Science and Technology and Economy, 2021,27(11):127-130.
2. Gu Chunfeng, Li Ling, Liu Chongqing. A brief description of the deformation monitoring of the South-to-North Water Diversion middle project[J]. Beijing Surveying and Mapping, 2022,36(10):1297-1301.
3. Han Shuai, Sun Leping, Yang Yiyun, et al. A Data Cleaning Method Based on Improved K-Means Clustering and Error Feedback[J]. Power System and Clean Energy, 2020,36(07):9-15.
4. Long Xiutang, Luo Ningning, Shen Yusheng. Construction Monitoring Design and Information Feedback of Long Span Tunnel with Variable Cross Section[J]. Journal of Henan Science and Technology, 2020, No.702(04):93-96.
5. Un Chol HAN, Chung Song CHOE, Kun Ui HONG, Hyon Il HAN. Intelligent back analysis of geotechnical parameters for time-dependent rock mass surrounding mine openings using the grey Verhulst model[J]. Journal of Central South University, 2021,28(10):3099-3116.
6. Zhang Chenming, Zhu Hehua, Zhao Haibin. Application of incremental displacement back analysis to hydropower station underground caverns[J]. Rock and Soil Mechanics, 2004, 25(2): =149-153.
7. Nie Xuejun, Hou Yucheng, Lu Zhaohui. Study on the Application of Wavelet Analysis in Data Processing for Dam Safety Monitoring[J]. Hongshui River, 2004, 23(2): 106-109.
8. Arfaoui, S., Ben Mabrouk, A., & Cattani, C. (2021). Wavelet Analysis: Basic Concepts and Applications (1st ed.). Chapman and Hall/CRC.
9. Wang Shiji. Processing the foundation pit monitoring data using Wavelet Analysis based on Matlab[J]. Geology of Anhui, 2022(S1):33-37.
10. Wei Yuming, Zhang Yongzhi, Wang Tao, et al. Damage identification based on structural dynamic responses[J]. China Earthquake Engineering Journal, 2017, 39(6): 1156-1160.

11. Zhang Yahui, Yang Kai, Yang Fan. Rotor fault diagnosis of induction motor based on wavelet packet energy analysis and signal fusion[J/OL]. Electrical Measurement & Instrumentation, 2021, 11: 1-9.
12. Fu Zhihao, Xiao Ming. Study on feedback analysis method for stability of surrounding rock in underground engineering[J]. Rock and Soil Mechanics, 2006, 27(S): 443-448.