Normalization processing and research of multi-field data in reservoir seepage investigation

KeXun Zheng\textsuperscript{1,2}, Yu Bo\textsuperscript{3}, Xianggang Liu\textsuperscript{1,2}, Daru Hu\textsuperscript{1,2}, Xiaohua Han\textsuperscript{1,2}, Senlin Wang\textsuperscript{1,2}, and Sheng Liu\textsuperscript{1,2}

1 HydroChina Guiyang Engineering Corporation Geotechnical Engineering Co., Ltd, Guizhou, 550081; 2. PowerChina Guiyang Engineering Corporation, Ltd, Guizhou, 550081

Abstract: Field analysis is one of the important methods to investigate the leakage of reservoir. Groundwater seepage field, groundwater temperature field, groundwater chemical field (conductivity field) and isotope field are widely used in leakage of reservoir. In this paper, the indexes of a single field are normalized according to some factors, including the relationship between field indexes and leakage sources, the relationship between field indexes and natural field background values of geological media, the variation of indexes in adjacent space, and the changes of indexes at different times. After processing, the characteristic values of the single field are obtained, including the tracer index characteristic value, the background index characteristic value, the gradient index characteristic value and the time series index characteristic value. Based on the reliability of the data and the correlation with the leakage, different weighting factors can be selected for each characteristic value to obtain a normalized field comprehensive index characteristic value. The indexes of different fields can also be superimposed according to the normalized characteristic value to realize the comparison and verification of each field. It makes full use of the limited data thus to reducing the interference information and amplifying the leakage signal. And it makes the field analysis method of groundwater leakage quantitative and systematic, which can effectively process the field data and more significantly delineate the possible groundwater leakage location.

Keywords: reservoir leakage investigation; field analysis method; multi field data; normalization processing; comprehensive index eigenvalue.

1. Introduction
By 2018, 98822 reservoirs of various types have been built in China, including 23841 dams with a height of more than 15m and a total reservoir storage capacity of 895.3 billion m\textsuperscript{3}. Part of the reservoirs in China were built in the 1950s and 1970s, mostly in the process of "survey, design and construction at the same time". On the one hand, due to a low-level requirement of anti-seepage design specifications and poor quality of site construction, repeated treatments have been conducted for the prominent leakage problem. On the other hand, after decades of operation of the reservoir, some of the grouting curtains are gradually aging, which results in an increased leakage and making
the dam in an unfavorable state.

Field analysis is one of the common methods to study groundwater leakage. Field refers to the distribution of objects in space, which is a special form of material existence characterized by spatial position function. The information of groundwater related variables has experienced special geological processes in a certain period in the geological medium space. The characteristic parameters and their changes of groundwater seepage, temperature, conductivity, chemical composition and isotope reflect some characteristics and laws of the objective existence of the groundwater system. Groundwater seepage field analysis method, groundwater temperature field analysis method, groundwater chemical field analysis method, conductivity (or resistivity) field analysis method and isotope field analysis method are widely used in engineering.

At present, the general application of field analysis method is to measure the representative indexes of each field of borehole, cavern or groundwater exposed point, and then make the corresponding index hole depth distribution curve, index profile (axis) distribution curve, or the cloud map of the index on the survey section. Then, according to the shapes of the curve and the cloud map, the abnormal area of the index in space is delineated by experience. Combined with the building structure and geological characteristics, corresponding assessments are made on the leakage. This kind of data processing, analysis and judgment process can not fully reflect the time and space change information of a single index, which will leave out more information and make full use of the data. The judgment of leakage information is still based on a qualitative analysis, which is greatly influenced by experience. Different field indexes have great differences, which can not achieve the comparative verification among these indexes.

2. Normalization of field data

The indexes of a single field are normalized according to some factors, including the relationship between field indexes and leakage sources, the relationship between field indexes and natural field background values of geological media, the variation of indexes in adjacent space, and the changes of indexes at different times. After processing, the characteristic values of the single field are obtained, including the tracer index characteristic value, the background index characteristic value, the gradient index characteristic value and the time series index characteristic value. According to the reliability of the data and the correlation with the leakage, different weighting factors can be selected for each characteristic value to obtain a normalized field comprehensive index characteristic value. Based on the aforementioned, the quantitative research and judgment is realized by making full use of the limited data, thus to reducing the interference information, amplifying the leakage signal, and effectively delineating the location of the leakage channel. The indexes of different fields can also be superimposed according to the normalized characteristic value to realize the comparison and verification of each field.

3. Field data normalization method

The field analysis methods widely used in engineering include groundwater seepage field analysis, groundwater temperature field analysis, groundwater chemical field analysis, conductivity (or resistivity) field analysis and isotope field analysis. The normalization method of field data is as follows.
3.1. Measured data
First of all, the water level, temperature, hydrochemistry, conductivity (or resistivity) and isotope field data $V$ of the water source in the study area are measured, and the measured values of each field of borehole, cavern or outcrop of groundwater are measured. According to the index change law of the water source, multiple repeated measurements can be carried out at different times.

3.2. Data interpolation
The field space can be divided into one-dimensional line (borehole or section line), two-dimensional surface (section) and three-dimensional geological space. For one-dimensional data such as borehole or section line, necessary data interpolation is conducted according to the measured values to make the field data evenly distributed in the elevation or horizontal direction. The representative data series of measured data are $(Z, V_Z)$ or $(X, V_X)$. Based on a certain field data of groundwater seepage point in different boreholes and underground caverns, plane interpolation is performed to obtain the field distribution data of the study section. The representative data series of the measured data are $(X, Z, V_{XZ})$.

3.3. Normalization of single field data
According to the relationship between the index of a single field and the leakage source, the relationship between the index and the background value of the natural field of geological media, the index changes in adjacent space, and the index changes at different times, the normalized processing is made to obtain the characteristic values of the tracer index, the background index, the gradient index, and the time series index of a single index.

3.3.1. Normalization of single borehole data.
A single borehole is a typical representative of one-dimensional field, which reflects the spatial variation characteristics of the field along the vertical direction of the borehole. Single borehole data normalization processing is to process the data series $(Z, V_Z)$.

3.3.1.1. Tracer index.
The tracer index reflects the close degree between the field representative value measured at a certain depth of the borehole and the corresponding field representative value of the leakage water source.

$$F_{Z1} = 1 - \frac{|V_Z - R|}{|V_Z - R|_{\text{max}}}$$

(1)

Where:
$F_{Z1}$ is the characteristic value of the tracer index of a field at the elevation of $Z$ in the borehole, and the calculated result is $0 \sim 1$, dimensionless. $0$ represents the biggest difference between the measured field index of the borehole and the leakage water source, and the possibility of leakage is small; $1$ represents that the measured field index of the representative borehole is the same as that of the leakage water source, so the possibility of leakage is high; $V_Z$ is the representative value of a field index at the elevation of $Z$ in the borehole, which is obtained by further interpolation of the measured values in different depth of the borehole;
$R$ is the representative value of a field index of the leakage water source, which is measured at the leakage water source;

3.3.1.2. Background index.

Background index reflects the abnormal degree of the measured representative value of a depth field in the borehole and the background value of the corresponding field index under general conditions.

$$F_{Z2} = \frac{|V_Z - B_Z|}{|V_Z - B_Z|_{\text{max}}}$$

Where:

$F_{Z2}$ is the characteristic value of the background index of a field at the elevation of $Z$ in the borehole, and the calculated result is $0 \sim 1$, dimensionless. 0 represents that the measured field index of the borehole is the same as the background value under normal conditions, i.e., there is no abnormality and the possibility of leakage is small; 1 represents that the difference between the measured field index and the background value of the representative borehole is the largest, there are anomalies, and the possibility of water leakage is high;

$B_Z$ is the background value of a field index at the elevation of $Z$ in the borehole, which can be obtained through theoretical analysis or empirical value;

3.3.1.3. Gradient index.

The gradient index reflects the degree of spatial variation between the measured value of a certain depth field and the measured value of the corresponding field index near the borehole.

$$F_{Z3} = \frac{|V_Z - V_{Z-\Delta Z}|}{|V_Z - V_{Z-\Delta Z}|_{\text{max}}}$$

Where:

$F_{Z3}$ is the gradient index eigenvalue of a field at the elevation of $Z$ in the borehole, and the calculated result is $0 \sim 1$, dimensionless. 0 represents that the measured field index of the borehole has no change at the adjacent elevation, i.e., there is no abnormality and the possibility of leakage is small; 1 represents that the measured field index of the borehole changes the most near the elevation, there are anomalies, and the possibility of water leakage is high;

$V_{Z-\Delta Z}$ is the representative value of a field index at the elevation of $Z-\Delta Z$ in the borehole;

3.3.1.4. Time series index.

The time series index reflects the time series change degree between the measured representative value of a certain depth field and the measured value of corresponding field index in other time periods.

$$F_{Z4} = \frac{|V_{Zt} - V_{Z(t-\Delta t)}|}{|V_{Zt} - V_{Z(t-\Delta t)}|_{\text{max}}}$$

Where:

$F_{Z4}$ is the time series index eigenvalue of a field at the elevation of $Z$ in the borehole, the calculated result is $0 \sim 1$, dimensionless. 0 means that the measured field index of the borehole has no change compared to a certain period in the past, i.e., there is no abnormality and the possibility of leakage
small; Compared with a certain period of time in the past, the measured field index of 1 represents that the change is the largest, there are anomalies, and the possibility of water leakage is high; 
\( V_{XZ} \) is the representative value of a field index at the time \( t \) at the elevation \( Z \) in the borehole; 
\( V_{Z,t\sim t+}\Delta t \) is the representative value of a field index at the time \( t - \Delta t \) at the elevation \( Z \) in the borehole; 
\( \Delta t \) is the time difference between the two measurements, generally considering the time when the leakage water source field index has changed significantly.

3.3.2. Normalization of profile data.
Data normalization of profile data is to process data series \((X, Z, V_{XZ})\).

3.3.2.1. Tracer index.
The tracer index reflects the close degree between the field representative value of a certain position in the research section and the corresponding field representative value of the leakage water source.

\[
F_{XZ1} = 1 - \frac{|V_{XZ} - R|}{|V_{XZ} - R|_{\text{max}}} 
\]

Where:
\( F_{XZ1} \) is the characteristic value of the tracer index of a field at the coordinates of \((X, Z)\) in the profile, the calculated result is \(0 \sim 1\), dimensionless. 0 represents the biggest difference between the field index of the profile and the leakage water source, and the possibility of leakage is small; 1 represents that the field index of the representative section is the same as that of the leakage water source, so the possibility of leakage is high;
\( V_{XZ} \) is the representative value of a field index at the coordinates of \((X, Z)\) in the profile, which is obtained by interpolating the measured field data of groundwater seepage points in different boreholes or underground caverns in the profile;
\( R \) is the representative value of a field index of the leakage water source, which is measured at the water source;

3.3.2.2. Background index.
Background index reflects the abnormal degree of the measured representative value of a certain position field in the profile and the background value of the corresponding field index under general conditions.

\[
F_{XZ2} = \frac{|V_{XZ} - B_{XZ}|}{|V_{XZ} - B_{XZ}|_{\text{max}}} 
\]

Where:
\( F_{XZ2} \) is the characteristic value of the background index of a field at the coordinates of \((X, Z)\) in the section, the calculated result is \(0 \sim 1\), dimensionless. 0 represents that the location field index in the section is the same as the background value in normal conditions, i.e., there is no abnormality and the possibility of leakage is small; 1 means that the difference between the location field index and the background value is the largest, there is an anomaly, and the possibility of water leakage is high;
\( B_{XZ} \) is the background value of a field index at the coordinates of \((X, Z)\) in the profile, which can be obtained by theoretical analysis or empirical value;
3.3.2.3. Gradient index.

Gradient index reflects the spatial variation of the representative value of a position field in the profile and the measured value of the corresponding field index in the adjacent area.

\[
F_{XZ3} = \frac{|V_{XZ} - V_{X-AZ} - \Delta Z|}{V_{XZ} - V_{X-AZ} - \Delta Z_{\text{max}}} \tag{7}
\]

Where:

- \(F_{XZ3}\) is the gradient index eigenvalue of a field at the coordinates of \((X, Z)\) in the profile, and the calculated result is \(0 \sim 1\), dimensionless. 0 means that there is no change in the field index near the area, that is, there is no anomaly and the possibility of leakage is small; 1 means that the field index changes most in the adjacent area, there are anomalies, and the possibility of water leakage is high;
- \(V_{X-AZ, Z-AZ}\) is the representative value of a field index of the adjacent area along a certain direction at the coordinates \((X, Z)\) in the profile;

3.3.2.4. Time series index.

The time series index reflects the time series change degree between the field representative value of a certain position in the section and the corresponding field representative value of other time periods in that position.

\[
F_{XZ4} = \frac{|V_{XZ} - V_{XZ(t-\Delta t)}|}{V_{XZ} - V_{XZ(t-\Delta t)}_{\text{max}}} \tag{8}
\]

Where:

- \(F_{XZ4}\) is the characteristic value of time series index of a field in the section with coordinates \((X, Z)\). The calculated results are \(0 \sim 1\), and there is no dimension. 0 represents that the field index of the section has no change compared to that of a certain period in the past, that is, there is no abnormality and the leakage probability is small; 1 represents that the field index of the section changes the most, there is abnormality and the possibility of leakage water is large;
- \(V_{XZ}\) is the measured representative value of a field index at the time \(t\) at the coordinate \((X, Z)\) in the profile;
- \(V_{XZ(t-\Delta t)}\) is the measured representative value of a field index at the time \(t-\Delta t\) at the coordinate \((X, Z)\) in the profile;
- \(\Delta t\) is the time difference of two measurements, generally considering the time when the index of leakage water source field has changed obviously;

3.4. Superposition of eigenvalues of single field index

According to the reliability of the data and the correlation with the leakage, different weights are selected to superimpose the characteristic values of the tracer index, the background index, the gradient index and the time series index. And the unified eigenvalue of a field's comprehensive index is obtained.

For single borehole data:

\[
F_{z} = \frac{\sum_{i=1}^{n} \alpha_i F_{z_i}}{n} \tag{9}
\]

Where:
$F_Z$ is the characteristic value of a field with elevation $Z$ in the borehole. The calculated results are $0 \sim 1$, and there is no dimension. 0 represents that the field index of drilling is normal, and the leakage probability is small; 1 represents that the field index of the drilling hole is abnormal, and the possibility of leakage water is high;

$F_Z$ is the characteristic value of tracer index, background index, gradient index and time series index of a field at $Z$ elevation in the borehole;

$\alpha_i$ is the weight of the characteristic value of tracer index, background index, gradient index and time series index respectively. According to the reliability of data and the correlation with leakage, the value of $\alpha_i$ is selected between $0 \sim 1$, $\sum_{i=1}^{n} \alpha_i = 1$. If the reliability of each characteristic value is similar to the leakage, $1/n$ can be taken;

$n$ is the number of calculated eigenvalue categories, and the characteristic value of tracer index, background characteristic value, gradient eigenvalue and time series eigenvalue are all calculated as 4. For profile data:

$$F_{XZ} = \frac{\sum_{i=1}^{n} \alpha_i F_{XZi}}{n} \quad (10)$$

Where:

$F_{XZ}$ is the field eigenvalue at the position $(X, Z)$ in the profile, the calculated results are $0 \sim 1$, dimensionless. 0 represents that the field index of the profile is normal and the possibility of leakage is small; 1 represents that the field index of the representative section is abnormal, and the possibility of water leakage is high;

$F_{XZi}$ is the characteristic value of tracer index, background index, gradient index and time series index of a field located at $(X, Z)$ in the profile.

### 3.5. Analysis of field index eigenvalue

For a single borehole data, draw the relationship curve between the field comprehensive characteristic value $F_Z$ and elevation $Z$ of different elevations in the borehole. And according to the comprehensive characteristic value of the field, the possible large, medium and small leakage sections are quantitatively divided.

For the profile data, the $F_{XZ}$ cloud maps of field integrated eigenvalues at different positions of the profile are drawn, and the possible large, medium and small leakage areas are quantitatively divided according to the field integrated eigenvalues.

### 3.6. Comprehensive analysis of each field data

According to the groundwater seepage field, temperature field, chemical field, conductivity (or resistivity) field and isotope field, repeat the steps from section 3.2 to 3.5, respectively delineate the leakage abnormal hole section or area of each field, compare and calibrate the abnormal hole section or area of each field, and finally find out the leakage location in combination with other geological survey and exploration methods.

For the leakage position of horizontal section line and three-dimensional space, it can be calculated by similar method according to the above research ideas.
4. Implementation cases
Taking the reservoir leakage study of Maotiaohe cascade IV Zhai Xiang Kou hydropower station in Guizhou Province as an example, the feasibility of field data normalization method for reservoir leakage investigation was demonstrated. In summer, the temperature field data was obtained by measuring the water temperature of each borehole arranged along the grouting gallery. The water temperature of the boreholes was measured with an interval of 1m. The temperature of reservoir water was significantly higher than that of natural groundwater, which had a good tracer, and the leakage water was quite different from the background value. ZK43 borehole was selected for data normalization processing according to this method. The orifice elevation of ZK43 borehole is EL.1067.51m and the depth is 198m.

For the tracer index, because the reservoir storage capacity is small and the difference between different positions in the reservoir is small, the average value of the measured reservoir temperature on that very day is taken as the leakage source temperature of reservoir water, i.e., \( R = 20.2 ^\circ C \). According to equation (1), the characteristic value \( F_{Z1} \) of tracer index at different depths of borehole can be calculated.

For the background value of borehole water temperature, as the grouting gallery is more than 30m beneath the ground, the rock mass underlying the grouting gallery will maintain normal temperature throughout the year provided that it is not affected by the leakage of reservoir water. The lowest water temperature 13.1\(^\circ\)C measured in the borehole ZK37 (located at CH.0+282m within EL.1050 ~ EL.1059m) can be regarded as the natural groundwater temperature at EL.1059m which is not affected by reservoir water. The elevation and temperature at the bottom of each borehole are different. The minimum geothermal gradient from 1059m elevation to the bottom of each borehole is 3.4\(^\circ\)C / 100m, which is regarded as the natural geothermal gradient \( k_0 \) of the survey area, thus the natural temperature field without the influence of reservoir water can be restored, and the natural water temperature background values \( B_z \) of different elevations can be obtained. The maximum temperature difference between the measured water temperature and the background value is 5.7\(^\circ\)C. According to equation (2), the characteristic value \( F_{Z2} \) of the background index at different depths of the borehole can be calculated.

For the gradient index, the distance of 1m is taken as the index \( AZ \). Calculate \( |V_z - V_{z-\Delta z}| \), in turn. The maximum value is 1.4\(^\circ\)C. The characteristic value \( F_{Z3} \) of the borehole gradient index can be calculated according to equation (3).

For the time series index, the water temperature measured in winter of this borehole in the 1980s was collected. The reservoir water is significantly characterized by a low temperature, which is quite different from the summer high temperature anomaly measured this time with test interval of 5m. According to the measured data, the winter water temperature corresponding to each elevation of the measured water level is obtained by piecewise linear interpolation. The maximum temperature difference of corresponding elevation in winter and summer is 10.1\(^\circ\)C. According to equation (4), the characteristic value \( F_{Z4} \) of time series index in different depth of borehole can be calculated.

According to the natural background temperature gradient, the natural background value of the water temperature at the bottom of the borehole is basically the same as that of the reservoir water in
summer. It is difficult to distinguish whether it is caused by the leakage of reservoir water or by the natural groundwater, which may interfere in the analysis and determination. When calculating the comprehensive index eigenvalue of the temperature field of the borehole, the corresponding weight of the tracer index eigenvalue $\alpha_i$ is taken as 0.1, the weights of background index eigenvalue, gradient index eigenvalue and time series index eigenvalue are all taken as 0.3, and the comprehensive index eigenvalue $F_Z$ of the temperature field of the borehole is calculated according to equation (9).

The relationship curve between the characteristic value of each index and the characteristic value of comprehensive index $F_Z$ and elevation is made respectively. $F_Z \geq 0.5$ is the most likely area of leakage, which is EL.1030 ~ EL.1045m and EL.982 ~ EL.989m respectively; $0.5 > F_Z \geq 0.3$ is the area with medium leakage possibility, which is EL.997 ~ EL.1045m, EL.989 ~ EL.994m and EL.978 ~ EL.982m respectively. Later, it is confirmed that those are the peripheral fracture area or fissure leakage area of pipeline leakage; $F_Z < 0.3$ is the elevation of EL.994 ~ EL.997m and EL.978m below, the rock mass is complete, so it is unnecessary to carry out leakage investigation and treatment.

5. Conclusions

The method described in this paper can normalize the index of a single field according to its relationship with the leakage source, the relationship with the background value of the natural field of geological media, the index changes in the adjacent space, and the index changes at different times. And then the tracer characteristic value, background characteristic value, gradient characteristic value...
and time series characteristic value of the single index can be obtained. According to the reliability of the data and the correlation with the leakage, different weighting factors can be selected for each characteristic value to obtain a unified field index system. Based on this, the quantitative research and judgment is realized, which can make full use of the limited data, thus to reducing the interference information, amplifying the leakage signal, and effectively delineating the location of the leakage channel. The indexes of different fields can be superimposed according to the same standard to realize the comparison and verification of each field. It makes the field analysis method of groundwater leakage quantitative and systematic, which can effectively process the field analysis data and obtain the possible groundwater leakage location information more significantly.

Acknowledgments
Special thanks are given for the invaluable support received from the science and technology project "Study on comprehensive evaluation of operation reliability of important depression in hydropower project" of POWER CHINA, and the major scientific research project "Research on key technologies for reservoir construction, development and utilization of karst depression" of POWER CHINA Guiyang Engineering Corporation Limited.

References
[1] Huang Qiang, Liu Dong, Wei Xiaoting, Yang Yuanyuan. Analysis on the reason why the number of dams built in China is the highest in the world [J]. Journal of Hydroelectric Engineering, https://kns.cnki.net/kcms/detail/11.2241.TV.20210507.1657.009.html.
[2] Fan Zhechao. Theory Study and Engineering Application of Integrated Tracer Methods for Leakage from Embankment [D]. Nanjing: Hohai University. 2006.
[3] Zou Chengjie. A study of geothermal field and karstic leakage in karstic area [J]. Carsologica Sinica, 1989, 8(2):143-150.
[4] Zheng Kexun, Zhang Guojun. Research on karst leakage pipeline detection of left bank seepage line of Zhaixiangkou Hydropower Station [J]. Guizhou Water Power, 2012, 26(2):10-21.
[5] Yuan Daoxian, Jiang Yongjun, Shen Licheng, etc., 2016. Modern karst [M]. Beijing: Science Press.
[6] Zou Chengjie, 1994. Karst engineering geology of water conservancy and hydropower[M]. Beijing: Water resources and electric power press.
[7] Peng Tubiao, Yuan Jianxin, Wang Huiming, 2011. Geological Manual of Hydropower Engineering [M]. Beijing: China water resources and Hydropower Press.
[8] Ouyang Xiaozhong, 2013. karst geology [M]. Beijing: China water resources and HydropowerPress.
[9] Shen Chunyong, Yo Bo, etc., 2015. Karst investigation and treatment of water conservancy and hydropower projects [M]. Beijing: China water resources and Hydropower Press.
[10] Li Shucai, etc., 2007. Advance prediction of karst fissure water and unfavorable geological conditions [J]. Journal of rock mechanics and Engineering, 26(2): 218-225.
[11] Zhao Rui, etc., 2011. Review of research on karst leakage and seepage control of reservoir [J]. Groundwater, 33(2): 20-22.
[12] Zou Chengjie, 1999. Preliminary study on geological model and mathematical model of reservoir karst seepage [J]. Carsologica Sinica, 9(3): 231-240.
[13] S. Altuğ Leakage study of the west side of the oymapinar reservoir, turkey, *Bulletin of the International Association of Engineering Geology - Bulletin de l'Association Internationale de Géologie de l'Ingénieur* volume 13, pages147–152(1976)

[14] S. Gangopadhyay, An approach to estimation of leakage from a karstic limestone reservoir, *Bulletin of the International Association of Engineering Geology - Bulletin de l'Association Internationale de Géologie de l'Ingénieur* volume 20, pages189–191(1979)

[15] Mengling Wang, Junrui Chai, Zengguang Xu & Yuan Qin, Leakage safety analysis of anti-seepage measures in reservoir basins: a case study of the Okinawa seawater pumped storage system in Japan, *Arabian Journal of Geosciences* volume 14, Article number: 345(2021)