

## Research Article

# Blind Source Separation Model of Earth-Rock Junctions in Dike Engineering Based on Distributed Optical Fiber Sensing Technology

Huaizhi Su,<sup>1,2</sup> Meng Yang,<sup>1,2</sup> Kunpeng Zhao,<sup>3</sup> and Zhiping Wen<sup>4</sup>

<sup>1</sup>State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China

<sup>2</sup>College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, China

<sup>3</sup>National Engineering Research Center of Water Resources Efficient Utilization and Engineering Safety, Nanjing 210098, China

<sup>4</sup>Department of Computer Engineering, Nanjing Institute of Technology, Nanjing 211167, China

Correspondence should be addressed to Huaizhi Su; [su\\_huaizhi@hhu.edu.cn](mailto:su_huaizhi@hhu.edu.cn)

Received 17 October 2014; Accepted 2 March 2015

Academic Editor: Zhenhua Zhu

Copyright © 2015 Huaizhi Su et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Distributed temperature sensing (DTS) provides an important technology support for the earth-rock junctions of dike projects (ERJD), which are binding sites between culvert, gates, and pipes and dike body and dike foundation. In this study, a blind source separation model is used for the identification of leakages based on the temperature data of DTS in leakage monitoring of ERJD. First, a denoising method is established based on the temperature monitoring data of distributed optical fiber in ERJD by a wavelet packet signal decomposition technique. The temperature monitoring messages of fibers are combined response for leakages and other factors. Its character of unclear responding mechanism is very obvious. Thus, a blind source separation technology is finally selected. Then, the rule of temperature measurement data for optical fiber is analyzed and its temporal and spatial change process is also discussed. The realization method of the blind source separation model is explored by combining independent component analysis (ICA) with principal component analysis (PCA). The practical test result in an example shows that the method could efficiently locate and identify the leakage location of ERJD. This paper is expected to be useful for further scientific research and efficient applications of distributed optical fiber sensing technology.

## 1. Introduction

China is a country affected by monsoons significantly. Floods and other natural disasters are often caused for rainy season. A large number of dikes are built in order to reduce the loss caused by flooding and invasion of storm tide. However, the integrity and safety of dike projects are objectively undermined by crossing culverts and other hydraulic structures, especially earth-rock junctions [1]. The operating experience of hydraulic structures of dikes indicated that strengthening real-time location and identification of leakage hazards for the earth-rock junctions of dike projects (ERJD) had special significances to ensure the safety of entire culverts and dike projects [2]. Research and analysis of high-tech leakage monitoring in ERJD have gotten more and more attention in engineering and academia during improvement

and optimization of traditional monitoring techniques and methods [3]. The temperature tracing technology is a geophysical technique based on geophysics and it has broad application prospects in seepage monitoring [4]. Compared to the traditional leakage monitoring technology, such as piezometer tubes and osmometers, the early temperature tracing technology was more sensitive and effective and had lower unit acquisition cost. However, the early temperature tracing method was still a point-monitoring technology. The singular area of internal temperature in buildings would not be detected inevitably because of a large layout spacing of thermistor thermometers and other reasons [5].

DTS, as a new approach, was produced by combination of optical fiber sensing technology and leakage risk monitoring technique [6, 7]. The new leakage monitoring method, DTS, provides a possible way to the distributed, long-distance, and

large-scale monitoring [8, 9]. In general, the use of distributed fiber sensing technology in leakage monitoring of buildings is still in its infancy stage for theoretical research, experimental simulation, and engineering application research, especially for monitoring the combination between two mediums in ERJD. The consequent analysis of temperature monitoring data and processing aspects still need more studies. Therefore, an experimental model of ERJD was established and the distributed fiber leakage monitoring technology was developed according to the working features of ERJD. In-depth studies of models and methods for leakage identification of ERJD have important science and engineering applications based on features of monitoring data.

## 2. Denoising Research of DTS Monitoring Data with Wavelet Packet

Wavelet transform method was proposed by Morlet in the early 1980s, and then it changed rapidly to be a very powerful tool for digital signal processing, image processing, data compression, data denoising, and so forth.

Wavelet is a breakthrough to Fourier transform analysis. It is a time-frequency analysis method with fixed window size and changeable shape, time window and frequency window. Local features signal can be well characterized in both time domain and the frequency domain.

Its main advantage is the ability to do local refinement and analysis for signal, which overcomes the difficulties of not processing the nonstationary signals for Fourier transform technique and short-time Fourier transform method.

During practical applications, wavelet transform will disperse the stretching factor  $a$  and translation factor  $\tau$  in order to adapt to the computer calculation.  $a$  takes  $a_0^0, a_0^1, a_0^2, \dots, a_0^j$ . The sampling interval of displacement amount  $\tau$  is  $\Delta\tau = a_0^j\tau_0$ . Thus, wavelet function is

$$\phi_{a,\tau}(t) = a_0^{-j/2} \phi(a_0^{-j}(t - ka_0^j\tau_0)), \quad (1)$$

where  $j$  and  $k$  are positive integers.

Taking  $a_0 = 2$  and  $\tau_0 = 1$  in practical operation, wavelet basis function,  $\phi_{a,\tau}(t)$ , can be expressed as follows:

$$\phi_{j,k}(t) = 2^{-j/2} \phi(2^{-j}t - k). \quad (2)$$

Corresponding discrete wavelet transform is

$$\text{WT}_x(j, k) = \langle x(t), \phi_{j,k}(t) \rangle = \int_R x(t) \overline{\phi_{j,k}(t)} dt. \quad (3)$$

The signal can be decomposed into a high frequency portion named  $D$  and a low frequency part marked  $A$  by wavelet transformation. Compared to the wavelet transform, wavelet packet decomposition analysis is a more sophisticated approach, which well disassembles the high and low frequency parts of each layer simultaneously. It greatly improves the resolution of signal and overcomes the insufficient of wavelet decomposition in "low-resolution of high-frequency." The wavelet packet decomposition has more broad application prospects and the 3-layer wavelet packet decomposition process was expressed in Figure 1.

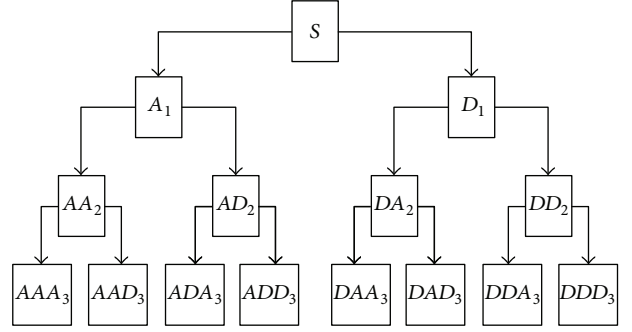


FIGURE 1: Transform decomposition process of wavelet packet.

As shown in Figure 1, the decomposition relationship of 3-layer wavelet packet is as follows:

$$S = AAA_3 + AAD_3 + ADA_3 + ADD_3 + DAA_3 + DAD_3 + DDA_3 + DDD_3. \quad (4)$$

The denoising procedure of wavelet packet for DTS leakage measurement data is summarized as below.

- (1) Wavelet packet decomposition. Select wavelets and determine the decomposition level  $N$  then decompose  $N$ -layer wavelet packet for original signal  $S$ . During data processing, the wavelet packet decomposition level is determined by the wavelet packet noise reduction effect.
- (2) Thresholds quantization. Select the appropriate threshold quantization method to quantify treatment for each wavelet packet coefficients decomposed.
- (3) Reconstruction of wavelet packet. Reconstruct the low-frequency coefficients of wavelet packet  $N$ -layer and its corresponding high-frequency coefficients.

Among the previous steps, the keys are selection of the wavelet packet decomposition level and approach to the threshold quantization. Wavelet packet decomposition level is generally not more than 5-layer. The threshold quantization can be processed by denoising with the soft threshold and hard threshold method.

Coefficient of wavelet packet of hard threshold estimated is

$$\bar{w}_{j,k} = \begin{cases} w_{j,k}, & |w_{j,k}| \geq \lambda \\ 0, & |w_{j,k}| < \lambda. \end{cases} \quad (5)$$

Coefficient of wavelet packet built for soft threshold is given:

$$\bar{w}_{j,k} = \begin{cases} \text{sign}(w_{j,k}) (|w_{j,k}| - \lambda), & |w_{j,k}| \geq \lambda \\ 0, & |w_{j,k}| < \lambda. \end{cases} \quad (6)$$

For (5) and (6),  $w_{j,k}$  represents wavelet packet coefficient and  $\lambda$  is preset threshold or threshold value, which is calculated in the form

$$\lambda = \frac{\sigma \sqrt{2 \lg N}}{\lg(j+1)}, \quad (7)$$

where  $\sigma$  is noise variance,  $j$  is decomposed scale, and  $N$  is signal length.

Signal to noise ratio (SNR) and root mean square error (RMSE) are usually selected as a signal denoising performance evaluation, which are calculated by following equations:

$$\text{SNR} = 10 \lg \left( \frac{\sum_{n=1}^N s^2(n)}{\sum_{n=1}^N [s(n) - \bar{s}(n)]^2} \right), \quad (8)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} [s(n) - \bar{s}(n)]^2},$$

where  $s(n)$  is original signal,  $\bar{s}(n)$  represents signal after denoising by wavelet packet, and  $N$  is signal length.

### 3. Blind Source Separation Models and Methods of DTS Temperature Monitoring Data

Seepage will lead to a change in the stability temperature field in the media and a partial volatile region can be generated [10, 11]. However, the temperature field distribution of structure body is affected not only by leakage field, but also soil characteristics, natural phenomena, optical fiber laying elevation, cross artificial cavities, and other factors. Further, the measured value of the optical fiber temperature sensing system is affected. Therefore, the temperature field of structure body is a result caused by multiple-factor under the combined effects. The positioning of leakage points inside ERJD under national condition can be attributed to the blind source separation problem for the influence of multiple-factor [12].

Leakages, soil characteristics or other factors causing changes in temperature monitoring data unit have a corresponding source component which is a function of the distance between the sampling point and the starting point. Soil characteristics, natural phenomena, and other factors should be possibly separated out of the temperature data except for the leakage factor. The separated data can be a result of temperature field change only affected by the leakage factor as much as possible. And the leakage factor is just the final concern.

The effects of leakage factor, soil characteristics, or other factors on the measured temperature data can also be considered to be independent units. The temperature data in most regions along fiber are not affected by leakages, gutters, and other factors, which are small probability events on time and space. Based on non-Gaussian distribution and statistical independent of the source signal, ICA is selected in blind source separation technique. In addition, PCA is determined as a pretreatment method for ICA technique in order to reduce the computational complexity and effective multiple-variable correlation.

The key of PCA is the certain degree of relevance and redundancy of information between the variables. At the same time, adjacent sampling points have similar work environment and strong correlation, which provide a possibility to the use of PCA. The mathematical model of PCA can be obtained by the following process. First, a  $p$  dimensional

random vector is consisted by  $p$  scalars and it is expressed as  $X = (X_1, \dots, X_p)'$ .  $p$  is defined as an arbitrary number.  $X$  is used for orthogonal transformation, ordering  $Y = T'X$ , and  $T$  is orthogonal matrix. Each component of  $Y$  is not relevant, and the variance of the first variable of  $Y$  is the largest, the second variable coming after. This process can be represented by the following matrix:

$$\begin{aligned} Y_1 &= t_{11}X_1 + t_{12}X_2 + \dots + t_{1p}X_p = T_1'X, \\ Y_2 &= t_{21}X_1 + t_{22}X_2 + \dots + t_{2p}X_p = T_2'X, \\ &\vdots \\ Y_p &= t_{p1}X_1 + t_{p2}X_2 + \dots + t_{pp}X_p = T_p'X. \end{aligned} \quad (9)$$

The definition of singular value decomposition (SVD) of PCA can be described:  $M$  represents a matrix with  $m$  rows and  $n$  columns, whose elements belong to the field of real numbers or complex field. There is the following decomposition:

$$M = USV^T, \quad (10)$$

where  $U$  represents a unitary matrix with  $m \times m$  order and is called left singular matrix of  $M$ ,  $V$  is a unitary matrix with  $n \times n$  order and is named right singular matrix of  $M$ ,  $V^T$  is transpose of matrix  $V$ , the same bellow,  $S$  represents a positive semidefinite diagonal matrix with  $m \times n$  order, and elements on the diagonal matrix are singular values of  $M$ , which are equal to the square root of eigenvalues of  $M^T * M$ . It is orthogonal matrix as the elements of unitary matrix are real numbers. Relationship between the main component of  $Y$  and monitoring data matrix  $X$  can be obtained by

$$Y = XT. \quad (11)$$

Thus,

$$X = YT', \quad (12)$$

where matrix of  $T$  is constituted by eigenvectors of diagonal matrix and  $X^T X$  is orthogonal matrix.

From the theory of SVD, it can be obtained that

$$X = USV'. \quad (13)$$

Comparing formula (12) with (13), it can be known that  $T = V$ ; then,  $Y = US = XV$ .

Therefore, the main component of matrix is the product of left singular matrix and its corresponding singular values and is also equal to the product of monitoring data matrix and the right singular matrix. The number of principal components can be determined through the cumulative variance contribution rate. Variance contribution rate and cumulative variance contribution rate are described follows:

$$\begin{aligned} \varphi_k &= \frac{\lambda_k}{\sum_{k=1}^p \lambda_k}, \\ \psi_m &= \frac{\sum_{k=1}^m \lambda_k}{\sum_{k=1}^p \lambda_k}, \end{aligned} \quad (14)$$

where  $\lambda_k$  represents the eigenvalues of covariance matrix and  $\varphi_k$  and  $\psi_m$  are, respectively, called variance contribution rate and cumulative variance contribution rate.

In order to avoid inconsistencies between dimensions and units of data, the data is typically normalized:

$$X_i^* = \frac{X_i - E(X_i)}{\sqrt{D(X_i)}}. \quad (15)$$

The mathematical model of ICA can be expressed by

$$\begin{aligned} x_1 &= a_{11}s_1 + a_{12}s_2 + \cdots + a_{1n}s_n, \\ x_2 &= a_{21}s_1 + a_{22}s_2 + \cdots + a_{2n}s_n, \\ &\vdots \\ x_n &= a_{n1}s_1 + a_{n2}s_2 + \cdots + a_{nn}s_n, \end{aligned} \quad (16)$$

where  $x_1, x_2, \dots, x_n$  represent  $n$  random variables, named mixed signals;  $s_1, s_2, \dots, s_n$  are  $n$  independent components, called source signals;  $a_{ij}$  ( $i, j = 1, 2, \dots, n$ ) is mixing coefficient.

Parallel algorithm calculation steps of several independent components are described as follows:

- (1) standardize the data so that the mean and variance of each column are, respectively, 0 and 1;
- (2) data whitened of  $z$  is gotten from the data obtained from Step (1);
- (3)  $m$  is selected as the number of independent component to be estimated;
- (4) the mixing coefficient of  $w_i$  is initialized,  $i = 1, 2, \dots, m$ , and every  $w_i$  is changed to be a unit 2-norm; then, matrix  $W$  is orthogonalised by the method of Step (6);
- (5) update each  $w_i$ ,  $w_i \leftarrow E\{zg(w_i^T z)\} - E\{g'(w_i^T z)\}w_i$ ; the function of  $g$  can be selected from formula (17) to formula (19);
- (6) matrix  $W = (w_1, \dots, w_m)'$  is orthogonalised,  $W \leftarrow (WW')^{-1/2}W$ ;
- (7) return to Step (5) if the result is not converged.

The function of  $g$  can be selected from the following functions:

$$g_1(y) = \tanh(a_1 y), \quad (17)$$

$$g_2(y) = y \exp\left(-\frac{y^2}{2}\right), \quad (18)$$

$$g_3(y) = y^3. \quad (19)$$



FIGURE 2: Model experiment of ERJD.

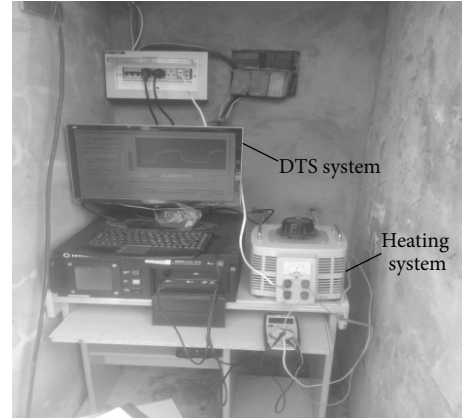


FIGURE 3: DTS monitoring and heating system.

The corresponding derivative functions are described as follows:

$$\begin{aligned} g_1'(y) &= a_1 (1 - \tanh^2(a_1 y)), \\ g_2'(y) &= (1 - y^2) \exp\left(-\frac{y^2}{2}\right), \\ g_3'(y) &= 3y^2, \end{aligned} \quad (20)$$

where  $a_1$  is a constant, valuing in  $[1, 2]$ , and the value is usually taken as 1;  $\tanh(x)$  represents hyperbolic tangent function.

#### 4. Application Case Analyses

A model experiment platform was formed and shown in Figure 2 to simulate the existence of leakages for ERJD. The DTS monitoring system of Sentinel DTS-LR produced by Sensornet was applied. The DTS monitoring and heating system was shown in Figure 3. A schematic diagram of the model was given in Figure 4. Multimode four core armored cable of 50/125  $\mu\text{m}$  was used as the monitoring optical cable with a type of ZTT-GYXTW-4A1a.

The ZTT-GYXTW-4A1a was a special optical cable which could be used to be heated. And its structure section was drawn in Figure 5.

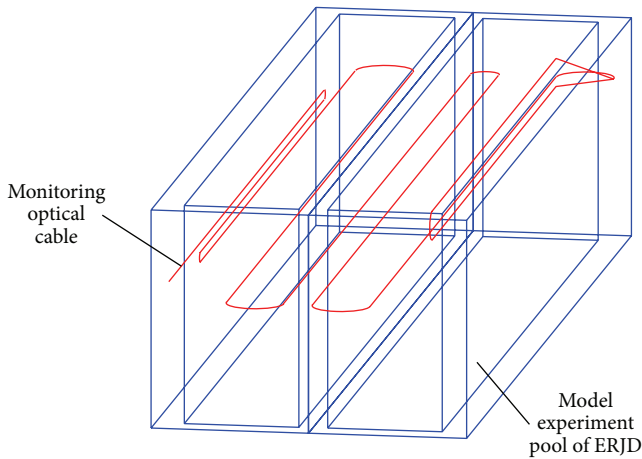


FIGURE 4: Schematic diagram of the model.

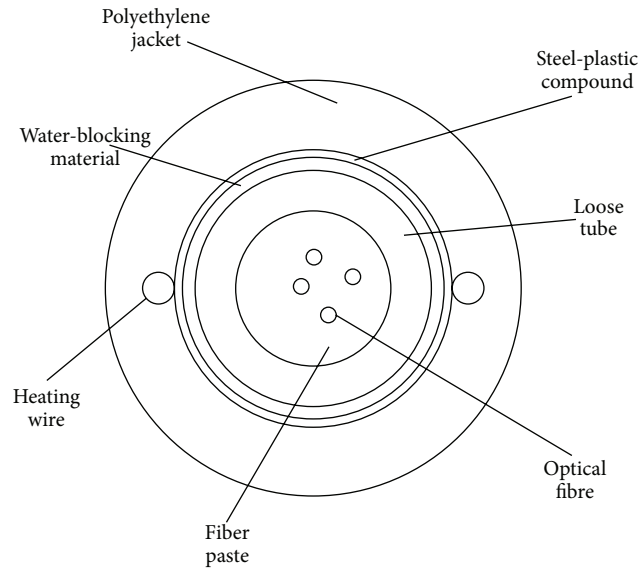


FIGURE 5: Structure section of ZTT-GYXTW-4A1a.

The leakage location was washed by a water pipe and the velocity of seepage was controlled. The monitoring data was noted as the stable velocity was formed.

For simplicity, just one leakage point was introduced in the model, and the sampling distance was 1 m and the sampling interval was 2 hours for 5 days. Thus, each measuring point obtained 60 groups of data. First, the original temperature data should be denoised and a typical point was selected as an example analyzed. After calculation, the decomposition levels of the wavelet packet were 3-layer. Figure 6 was shown to compare the raw signal with the wavelet denoising signal.

The first principal component was calculated and listed in Figure 7.

The corresponding variance contribution rate of the first principal component,  $\sigma_1 / \sum \sigma$ , was 87.21%, and it could be concluded that the contribution of the first principal component to information was larger. The difference of the principal component value between the leakage point presented and

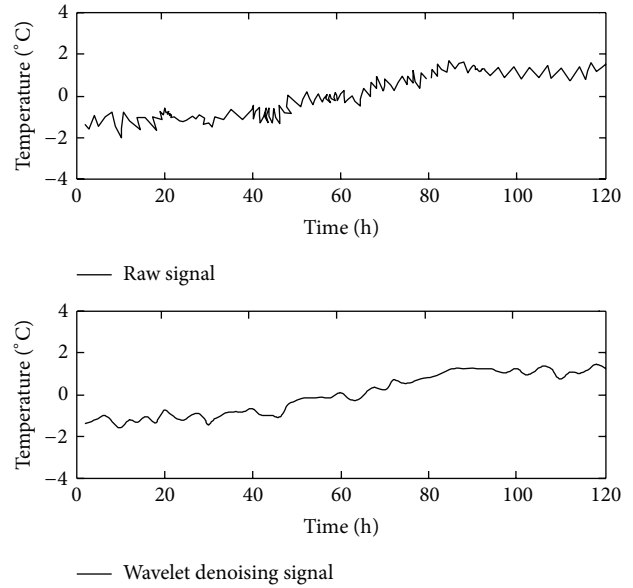


FIGURE 6: Comparative diagram between raw signal and the wavelet denoising signal.

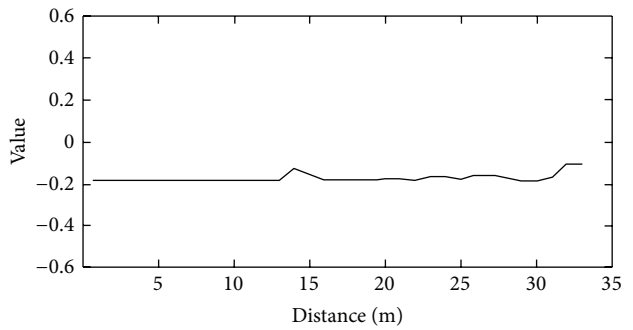


FIGURE 7: Result diagram after PCA method.

other points was much larger than the one of the first principal component. Therefore, the second and third principal components were selected to the next ICA, which meant  $i = 2$ . One of the two principal components was used to construct a separate component of the signal space, and the other one was made to build a separate part of the salvage space. The difference of the separate component 1 in the leakage point was more evident than the one of other points whose distribution was more stable. Thereby, the separate component 1 was considered to establish a salvage space. The separate component 2 constructed a separate component signal space. The corresponding grey-scale map of residual space was shown in Figure 8.

From the result of data processing, it can be known that the problem of the optical fiber temperature monitoring data could be analyzed by the blind source separation and ICA was used to process data. In Figure 8, the leakage area of 14 m was larger than the area of 32 m. And the test distance of the model was less than 32 m. Therefore, the area of 32 m

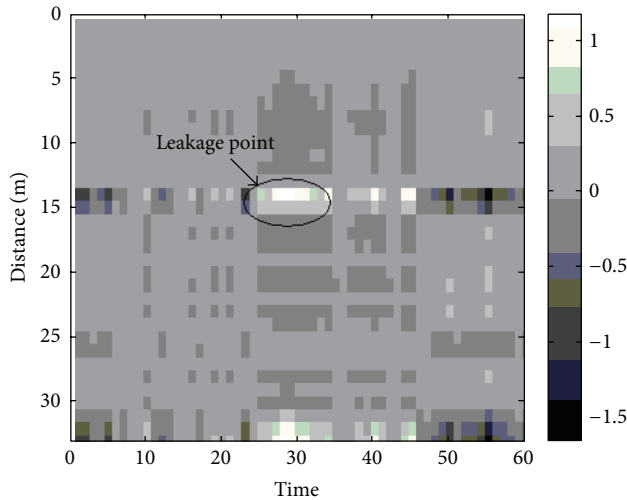


FIGURE 8: Grey-scale map of residual space.

was not considered. Finally the leakage location could be well implemented.

## 5. Conclusions

In this paper, ERJD was the studied object and its leakage identification was for the research emphasis. Based on leakage characteristics of monitoring ERJD by DTS, the affecting factors of optical fiber temperature monitoring data were analyzed. The wavelet packet denoising method and threshold determining method were also studied. A blind source separation model of optical fiber temperature data was finally built and the corresponding temperature change process of the leakage source was extracted. The implementation method of blind source separation technique was discussed. In-depth research on characteristics of optical fiber temperature data was carried out and optimal applicable conditions for realization method of the blind source separation techniques were also analyzed. The blind source separation technique implementation combining ICA with PCA was finally proposed. With leakage fiber-optic model experiment of ERJD, the reliability of optical fiber temperature measuring data based on the blind source separation method is validated.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This research has been partially supported by Jiangsu Natural Science Foundation (SN: BK2012036), the National Natural Science Foundation of China (SN: 51179066, 51139001, 41323001, and 51409167), the Specialized Research Fund for the Doctoral Program of Higher Education of China (SN: 20130094110010), the Nonprofit Industry Financial Program

of MWR (SN: 201301061 and 201201038), the Open Foundation of State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering (SN: 20145027612), the Research Program on Natural Science for Colleges and Universities in Jiangsu Province (SN: 14KJB520016), the Priority Academic Program Development of Jiangsu Higher Education Institutions (SN: YS11001), and the Scientific Innovation Research Scheme for Jiangsu University Graduate (SN: 2013B25514).

## References

- [1] H.-H. Zhu, A. N. L. Ho, J.-H. Yin, H. W. Sun, H.-F. Pei, and C.-Y. Hong, "An optical fibre monitoring system for evaluating the performance of a soil nailed slope," *Smart Structures and Systems*, vol. 9, no. 5, pp. 393–410, 2012.
- [2] B. Shi, C.-S. Tang, L. Gao, C. Liu, and B.-J. Wang, "Observation and analysis of the urban heat island effect on soil in Nanjing, China," *Environmental Earth Sciences*, vol. 67, no. 1, pp. 215–229, 2012.
- [3] J. F. Yan, B. Shi, D. F. Cao, H. H. Zhu, and G. Q. Wei, "Experiment study on carbon coated heating optical fiber for sand leakage monitoring based on distributed temperature system," in *Proceedings of the 6th International Conference on Structural Health Monitoring of Intelligent Infrastructure*, Paper no. MS03-06, Hong Kong, 2013.
- [4] S. Johansson, "Localization and quantification of water leakage in ageing embankment dams by regular temperature measurements," in *Proceedings of the 17th International Congress on Large Dams (ICOLD '91)*, pp. 991–1005, Vienna, Austria, 1991.
- [5] O. Kappelmeyer, "The use of near surface temperature measurements for discovering anomalies due to causes at depths," *Geophysical Prospecting*, vol. 5, no. 3, pp. 239–258, 1957.
- [6] A. H. Hartog, "Distributed fiber-optic temperature sensors: principles and applications," in *Optical Fiber Sensor Technology*, K. T. Grattan and B. T. Meggitt, Eds., pp. 241–301, Kluwer Academic Publishers, New York, NY, USA, 2000.
- [7] H. Z. Su, J. Y. Li, J. Hu, and Z. Wen, "Analysis and back-analysis for temperature field of concrete arch dam during construction period based on temperature data measured by DTS," *IEEE Sensors Journal*, vol. 13, no. 5, pp. 1403–1412, 2013.
- [8] S. Yin, "Distributed fiber optic sensors," in *Fiber Optic Sensors*, F. T. Yu and S. Yin, Eds., pp. 201–229, Marcel Dekker, New York, NY, USA, 2000.
- [9] A. D. Kersey, "Optical fiber sensors for downwell monitoring applications in the oil and gas industry," in *Proceedings of the 13th International Conference on Optical Fiber Sensors (OFS-13 '99)*, pp. 326–331, 1999.
- [10] S. Grosswig, A. Graupner, and E. Hurtig, "Distributed fiber optical temperature sensing technique—a variable tool for monitoring tasks," in *Proceedings of the 8th International Symposium on Temperature and Thermal Measurements in Industry and Science*, pp. 9–17, 2001.
- [11] S. A. Wade, K. T. V. Grattan, B. McKinley, L. F. Boswell, and C. D'Mello, "Incorporation of fiber optic sensors in concrete specimens: testing and evaluation," *IEEE Sensors Journal*, vol. 4, no. 1, pp. 127–134, 2004.
- [12] H.-H. Zhu, L.-C. Liu, H.-F. Pei, and B. Shi, "Settlement analysis of viscoelastic foundation under vertical line load using a fractional Kelvin-Voigt model," *Geomechanics and Engineering*, vol. 4, no. 1, pp. 67–78, 2012.



**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

