



## Water Leakage Detection in Dikes by Fiber Optic

Jerome I. Mars, Amir Ali Khan, Valeriu Vrabie, Alexandre Girard, Guy d'Urso

### ► To cite this version:

Jerome I. Mars, Amir Ali Khan, Valeriu Vrabie, Alexandre Girard, Guy d'Urso. Water Leakage Detection in Dikes by Fiber Optic. EAGE 2010 - 72nd European Association of Geoscientists and Engineers Conference and Exhibition incorporating SPE EUROPEC 2010, Jun 2010, Barcelone, Spain. pp.P221. hal-00505226

**HAL Id: hal-00505226**

**<https://hal.science/hal-00505226>**

Submitted on 23 Jul 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

P221

## Water Leakage Detection in Dikes by Fiber Optic

J.I. Mars\* (Grenoble Institute of Technology), A.A. Khan (CIH Grenoble), V. Vrabie (University of Reims), A. Girard (EDF) & G. D'urso (EDF)

### SUMMARY

---

We propose a method for the identification of leakages in dikes using the temperature data obtained through fiber optic distributed temperature sensors. We showed how it is possible to treat leakage identification as a source separation problem. The sources were considered as defining the response of the ground, the known structures in the path of the fiber sensors (drains), the seasonal variations, the precipitations and, of course, the leakages, the last ones being our desired signals. We showed that with the help of techniques based on data decomposition and source separation (by PCA and ICA techniques), we can identify the leakages.

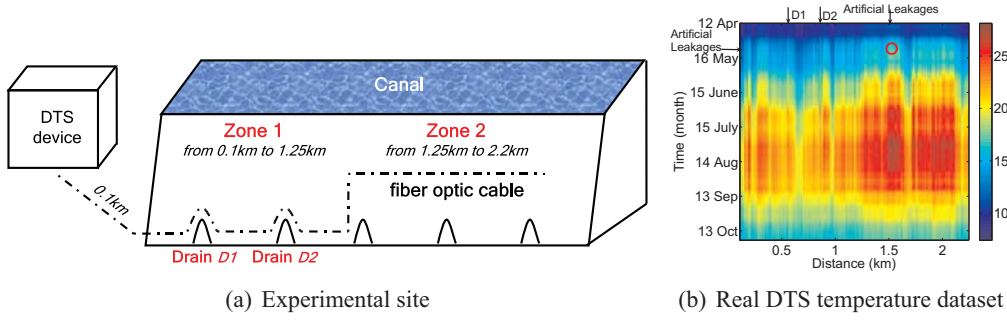
## Introduction

Fiber optic sensors are employed in diverse domains covering applications as engineering structures monitoring, fault detection in electrical circuits, fire detection systems, parameter sensing in oil and gas industry, etc [1],[2]. In subsurface engineering domain, an important issue is the detection of anomalies, such as flow of water, in the dikes to avoid disaster. In this regard, thermometric based method employing optical fiber Distributed Temperature Sensors (DTS) provides a very efficient method. The major advantage of this technic based on optical fiber is their commercial viability (low-cost), ability to multiplex large number of sensors along a single fiber and environmental robustness. DTS have the capability to provide a monitoring range that may go up to 30 km. Classical spatial and temperature resolutions are respectively around 1 m and  $0.01^{\circ}\text{C}$ . The goal of this paper is to propose an efficient and automated system based on the processing of DTS data to identify singularities such as water leakages in dike structures. The basic concept behind temperature acquisition is that a change of ground temperature is brought about by a significant flow of water through the structure due to leakages. Unfortunately, this change of temperature can equally be brought about by other natural factors (rainfall, day and seasonal effects, interference with existing structures like drains, etc). Additionally, as the fiber optic cable is buried in ground, the temperature signals acquired by DTS are also strongly influenced by the response of the near surface (ground) where the acquisitions are made. In first step, we propose a processing based on singular value decomposition to separate and isolate the ground response. Secondly, a detection processing based independent component analysis algorithm is proposed for leakage detection. We validate the proposed system on real temperature data sets under different scenarios.

## Acquisition Principle and Data Description

The Distributed Temperature Sensors (DTS) were successfully used for the acquisition of temperature data over long distances with high spatial and temperature resolutions [2]. In most commercial distributed temperature sensors, the acquisition principle of temperature profiles is based on Raman scattering using Optical Time Domain Reflectometry (OTDR) techniques [3]. Measuring the travel time of probe pulse and the intensity ratio, at the fiber input, gives the temperature profile along the entire length of the fiber using one-to-one relationship between spatial resolution and traveling time. Based on this physical principle, a DTS has been used as thermometric data monitoring setup by the EDF Group (a leader energy player in France) to study real and artificial leakages on one experimental test site. The aim of study is to extract the information pertaining to leakages (both natural and controlled) in the dike of canal. A schematic representation of this site is given in Fig. 1(a). A fiber optic cable was installed in the abutment at the toe end of the canal (at a depth of 1 m) so as to intercept the water leakages from the canal. Two distinct elevation levels (Zone1, from approximately 0.1 km to 1.25 km and Zone2, from approximately 1.25 km to 2.2 km) will be exposed with varying intensities to direct sunlight. The cable also circumvents two drains, D1 and D2, situated at 0.561 km and 0.859 km, respectively. The temperature data is acquired with a sampling interval of 1 m along the entire cable length. Continuous temperature monitoring is important in order to observe the temporal evolution of the leakages. The recorded data set is thus a two-dimensional temperature signal and can be written as:  $\mathbf{Y} = \{y(t, x) \mid 1 \leq t \leq N_t, 1 \leq x \leq N_x\}$ , where  $N_t = 168$  and  $N_x = 2200$  represents the total acquisition time and the number of observation points, respectively.

During 6 months in 2005, we record dataset everyday. On May 10th and 12th, three artificial leakages, L1, L2 and L3, were introduced on the site with different flow rates of 5, 1 and 1 lit/min and at different positions, 1.562, 1.547 and 1.569 km, respectively (in the red circle on Fig. 1(b)). A hot point (HP) was also introduced on day 10th at 0.674 km. On Fig. 1(b), we can identify spring and summer time temperature where the different colors represent the scale of the recorded temperature with dark blue showing lowest temperature and dark red the highest. On this figure, we can also note that the behavior of rainfall is expressed as line in horizontal axis (with temperature offset).



**Figure 1** Site and dataset recorded on 6 months (red circle expresses days with artificial leakages)

### Processing to detect leakages

Firstly, a preprocessing step is applied to remove seasonal effect. For this step, the dataset is filtered by a low pass filter and normalized (zero mean and unitary variance for each profiles, i.e. of each column of the recorded dataset  $\mathbf{Y}$ ). This preprocessing leads to attenuation of the temporal variability of the data (daily and seasonal variations). We can also identify and remove precipitation period using a criteria based on higher order statistics developed by authors but that does not fit into the scope of this paper and would be presented elsewhere [4]. We will see now how we can exploit this dataset to extract useful information linked to the leakage.

A recorded dataset can be considered as a linear mixture of sources. Likewise, since the sources originate from non-correlated phenomena, they are supposed to be independent of each other. The problem can thus be formulated as:  $\mathbf{Y}^T = \mathbf{M}\mathbf{F}^T$ , where  $\mathbf{Y} \in \mathbb{R}^{N_x \times N_t}$  represents the acquired data,  $\mathbf{M} \in \mathbb{R}^{N_t \times p}$  is the mixing matrix,  $\mathbf{F} \in \mathbb{R}^{N_x \times p}$  designates the matrix made up of the  $p$  independent sources, i.e. a source represents a column of the matrix  $\mathbf{F}$ , and  $^T$  denotes the matrix transposition. The identification of each source can thus be treated as a source separation problem. The classical techniques commonly used are Principal Component Analysis (PCA) and Independent Component Analysis (ICA). When the principal components are calculated from the covariance matrix of  $\mathbf{Y}$ , an efficient method for their calculation is given by the Singular Value Decomposition (SVD) [5] written as:

$$\mathbf{Y}^T = \mathbf{U}_N \Sigma_N \mathbf{V}_N^T = \sum_{j=1}^N \sigma_j \mathbf{u}_j \mathbf{v}_j^T = \mathbf{Y}_{sig}^T + \mathbf{Y}_{noise}^T = \sum_{j=1}^m \sigma_j \mathbf{u}_j \mathbf{v}_j^T + \sum_{j=m+1}^N \sigma_j \mathbf{u}_j \mathbf{v}_j^T \quad (1)$$

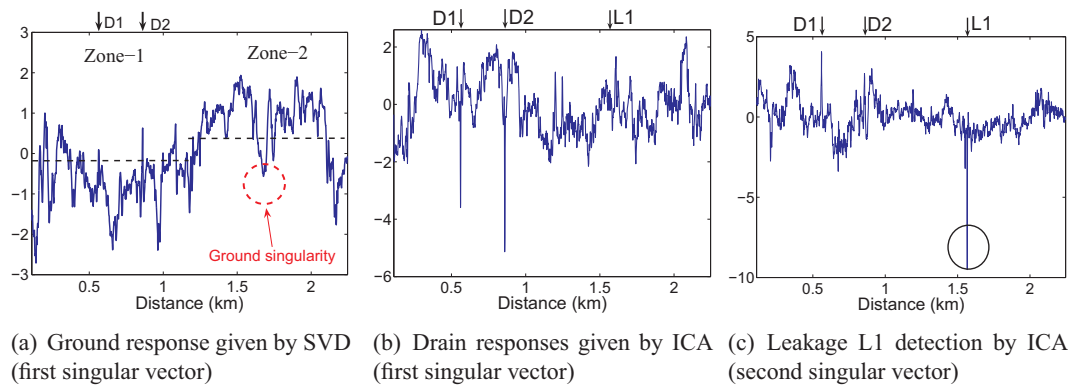
where  $N = \min(N_x, N_t)$ ,  $\Sigma_N \in \mathbb{R}^{N \times N}$  is a matrix containing on its diagonal the singular values  $\sigma_j \geq 0$  arranged in a descending order and  $\mathbf{U}_N \in \mathbb{R}^{N_t \times N}$  and  $\mathbf{V}_N \in \mathbb{R}^{N_x \times N}$  are orthogonal matrices, containing  $N$  left and right singular vectors  $\mathbf{u}_j \in \mathbb{R}^{N_t}$  and  $\mathbf{v}_j \in \mathbb{R}^{N_x}$ , respectively. As these vectors are orthonormal by construction, the estimated sources are decorrelated and normalized. The decorrelation allows to extract Gaussian sources, which is not sufficient in our case because sources associated with some factors such as the drains or leakages can be modeled by sparse and non-Gaussian sources. By SVD, it is also possible to decompose the initial data into two complementary subspaces, namely, the signal subspace and the noise subspace (Eq. (1)). For our dataset, the first one subspace being contained in a space of dimension  $m = 1$  has been identified as the ground response (Fig. 2(a)). The second subspace (dimension  $N - 1$ ) contains all other sources (drains and leakages) mixed together, and at this stage, PCA is not usefull to separate them (due to the orthogonality given by SVD).

A more realistic technique not driven by the orthogonality condition is based on the independence of the sources (Independent Component Analysis). ICA is a blind decomposition of a multichannel data set made up of unknown linear mixtures of unknown source signals based on the assumption that the sources are mutually statistically independent. The goal of ICA is to estimate the mixing matrix  $\mathbf{M}$  and/or the

source matrix  $\mathbf{F}$  from the noise subspace  $\mathbf{Y}_{noise}^T$  given by first SVD with the only hypothesis that the sources containing in are independent. ICA is usually resolved by a two step algorithm, consisting of a prewhitening step and a higher-order step. The first step is directly carried out by SVD on  $\mathbf{Y}_{noise}^T$  to obtain the whitened (decorrelated and normalized) vectors  $\mathbf{v}_j$ . The second step then comprises of finding a rotation matrix  $\mathbf{B} \in \mathbb{R}^{i \times i}$ , which diagonalizes the tensor of 4<sup>th</sup> order cross cumulants constructed with the columns of  $\mathbf{V}_i$ . One of the popular algorithms for finding this rotation matrix is the joint approximate diagonalization of eigenmatrices (JADE) algorithm [6]. This second step provides independent vectors  $\tilde{\mathbf{v}}_j$  from the decorrelated ones  $\mathbf{v}_j$ . These independent vectors are the columns of the matrix  $\tilde{\mathbf{V}}_i = \mathbf{V}_i \mathbf{B}$ . From ICA, it is now possible to find each source by decomposition.

## Results

The methodology adopted for the identification of leakages will now be resumed. As a first step of data decomposition in subspaces, PCA is applied. Signal subspace built by the first singular vector is identified as ground response (Fig. 2(a)) where zone1 and zone2 can be characterized. Additionally a ground singularity has been detected. On the site, this singularity has been identify as a specific soil. The precipitations and the leakages are ephemeral phenomena in time and time/space domains respectively and are considered as “noise space” (the residue). This in turn means that they are not coherent to all the acquired data and are thus not going to be revealed in the first vector obtained by PCA. This residue is of dimension  $N - 1$  and ICA has been then applied on but considering only “ $i$ ” sources to be estimated, which means that only the decorrelated sources  $\mathbf{v}_2$  to  $\mathbf{v}_{2+i}$  are considered in the second step of the ICA algorithm. The first ICA source allows to identify clearly drains  $D1$  and  $D2$  (Fig. 2(b)) and the second ICA source brings mainly information on leakage  $L1$  (Fig. 2(c)).



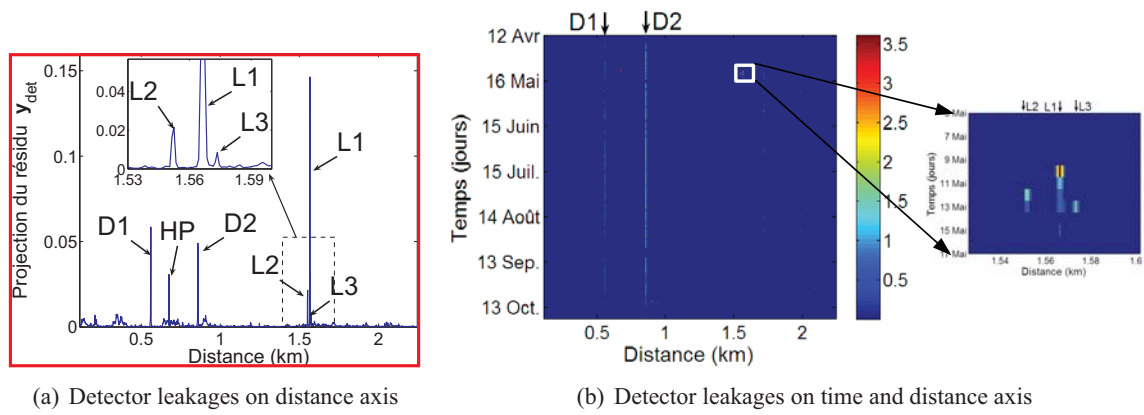
**Figure 2** Source separation on DTS dataset

In Fig. 3(a), we present the final result of the detector plotted on distance axis. It is possible to identify on this plot, each element as drain 1 and 2, the leakages 1, 2 and 3 and the hot spot. On the 6 months duration, by using this efficient processing, we can identify each phenomena on their respective days and positions ((Fig. 3(b)).

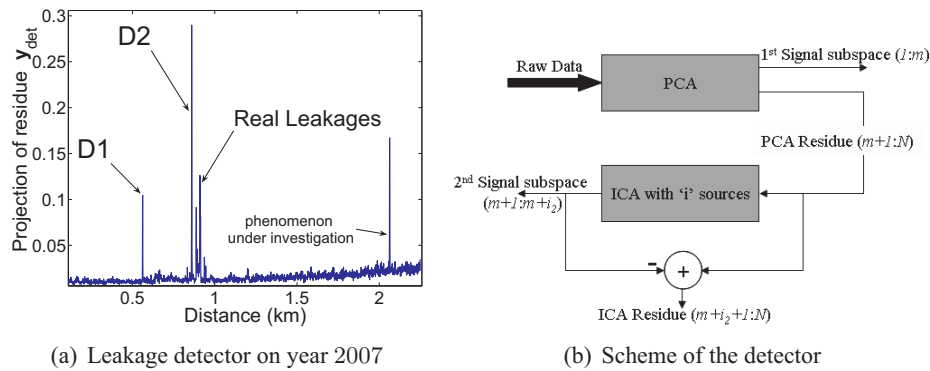
By applying the same detection processing (Fig. 4(a)) two years latter (2007), we find, on this site, the same external structures (Drains 1 and 2) with some real leakages (concentrated around drain 2) (Fig. 4(b)). After detection these real leakages have been investigated on the field. The proposed system was validated under different scenarios and the repeatability of the system was also verified by periodic analysis. We also prove the efficiency of this processing on complete simulated dataset.

## Conclusions

In the present work, we have proposed a method for the identification of leakages in dikes using the temperature data obtained through fiber optic distributed temperature sensors. We showed how it is



**Figure 3** Results of the detection processing on simulated leakages .



**Figure 4** Proposed scheme for detection of the leakages and results

possible to treat leakage identification as a source separation problem. The sources were considered as defining the response of the ground, the known structures in the path of the fiber sensors (drains), the seasonal variations, the precipitations and, of course, the leakages, the last ones being our desired signals. We have shown that with the help of techniques based on data decomposition and source separation (by PCA and ICA techniques), we can identify the leakages. It was found that for the best results, a single source should be utilized for constructing the signal subspace with PCA. The corresponding residue should then be treated using ICA on two sources. The future work would focus on the testing of this system on other data sets acquired at different sites as well as the development of other measures to quantify the shapes and variability of the leakages.

## References

- [1] Hartog, A. H. [2001] Progress in distributed fiber-optic temperature sensing. *Proc. SPIE, Fiber Optic Sensor Technology and Applications*, **4578**, 43-52.
- [2] Grosswig, St., Graupner, A. and Hurtig, E. [2001] Distributed fiber optical temperature sensing technique - a variable tool for monitoring tasks. *Proc. 8th Intl. Symposium on Temperature and Thermal Measurements in Industry and Science*, **19-21** June, 9-17.
- [3] Rogers, A. [2002] Distributed optical fiber sensing, in *Handbook of Optical Fiber Sensing Technology*, J. M. Lopez-Higuera, Ed. Wiley, 271-308.
- [4] Khan A.A, Vrabie V., Mars J.I. and Girard A. [2008] A least square approach for bidimensional source separation using higher order statistics criteria, *16th European Signal Processing Conference (EUSIPCO 2008)*, Lausanne, Aug. 2008 pp L6-6
- [5] Scharf L.L [1991] Statistical Signal Processing: Detection Estimation and Time series analysis, New-york Addison-Wesley
- [6] Khan A.A, Vrabie V., Mars J.I., Girard A., and Durso G. [2009] A Source Separation Technique for Processing of Thermometric Data from Fiber Optic DTS Measurements for Water Leakage Identification in Dikes *IEEE Transactions of Instrumentation and Measurements*, accepted.