# Signal Processing and Optimization Process for leakage detection and localization

2nd INTERNATIONAL CCWI / WDSA

**Daniel Barros[1], Pedro Resende[2], Bruno Brentan[3], Gustavo Meirelles[4],, Idel Montalvo[5], Edevar Luvizotto[6], Joaquín Izquierdo[8]**

[1,3,4] Hydraulic Engineering and Water Resources Department, School of Engineering – Federal University of Minas Gerais, Belo Horizonte, Brazil

[2,6] School of Civil Engineering, Architecture and Urban Studies, University of Campinas, Campinas, Brazil

[5]IngeniousWare GmbH, Karlsruhe, Germany

[7]FluIng, Institute for Multidisciplinary Mathematics, Universitat Politècnica de València, Valencia, Spain.

[3] *brentan@ehr.ufmg.br*

## *EXTENDED ABASTRACT*

Several classes of algorithms, such as clustering (WU et al, 2016), forecasting (MOUNCE et al, BIREK et al, 2014), filtering (YE et al, 2011) and optimization (HAGHIGHI and RAMOS, 2011) are widely applied for leakage detection and localization. Mainly in the last decades, when the monitoring systems have become popular and the processing capacity of computer have increased, data mining and optimization have changed the routine of water companies.

Observing the hydraulic system, measured pressure and flow are the result of a non-linear combination of consumed water and leakages, modified by control devices. In signal processing field, the leakages and consumers can be read as sources. These sources have different behaviors and their actions will result in a final pressure and flow signals, acquired by sensors. Bearing it in mind, a set of techniques for signal separation without previously knowledge of the sources, so called, blind source separation (BSS) can be applied on pressure and flow data for leakage detection. Independent Component Analysis (ICA), one algorithm from the class of BSS, have been used successfully in health-related research (WESSEL, 2018), seismic activities (CAPUANO *et al*., 2017), image analysis (AHMAD *et al*., 2017), discharge flow separation in water systems (ZHENG *et al*., 2016). ICA is able to differentiate a linear combination in a data set, extracting components of source signal without variation. (LANGLOIS *et al*., 2010).

For the Battle of Leakage Detection and Isolation (BattleDIM), ICA is applied on hydraulic data, for source separation helping to identify water consumption and noise sources. The results show that noise source is significantly affected by leakages. The second stage of leak detection methodology, a statistical controlling method, Abrupt Change Point Detection (CPD), is applied using noise data as input. CDP assumes that changes in data affect the statistical properties of the dataset, such as average and standard deviation. Because of this, in general, CPD techniques monitor changes in some of these parameters, identifying abrupt changes (LAVIELLE, 2005; BRENTAN *et al*., 2017).

Optimization processes also gained space in research on WDS, such as in the optimization of pump and valve control (Mala-Jetmarova, et al, 2017), network calibration (Savic, et al, 2009), etc. Since BattleDIM offer a non-calibrated model, and pressure/flow dataset, calibration process is applied in two different stages. The first one, a dataset of normal condition is used to calibrate roughness and base demand using Particle Swarm Optimization algorithm linked to Epanet hydraulic simulator. The results are used on a second stage, where pressure and flow during detected leakages by the ICA-CDP algorithm is used for calibration of a leakage event. A 24h dataset is used and leakage flow and nodal leaking position are calibrated. The main objective of this stage is to identify the region of leaking pipe.

## METHODOLOGY

The methodology used in this work is divided into two parts: Leakage Detection and Leakage Isolation. The first one is based on data process and signal analysis and has the main goal to identify when leakages start in the network. The second part, Leakage Isolation, is based on optimization process for calibration, and it's responsible to localize the leakage region.

### Leakage Detection

Flow and pressure data are used to detect leaks, being separeted in two sources by ICA algorithm. Langlois et al. (2010) present the fundamentals of the ICA method, which can be denoted through a random vector observed $X = [X_1, X_2, \dots, X_m]^T$ where $m$ elements are mixtures of $m$ independent elements of a random vector $S = [S_1, S_2, \dots, S_m]^T$ given by:

$X = AS$ (eq. 1)
where $A$ is a mix matrix $m \times m$, the sample value of $X_j$ is indicated by $x_j$ and $j = 1,2, \dots, m$. ICA aims to find an unmixing matrix $W$, or the inverse of A, which provides $Y$, what is the best possible way to find $S$:

$Y = WX \cong S$ (eq. 2)

ICA is applied for pairs of flow and pressure, generating a large dataset of sources. These sources, usually, can represent with better precision, anomalies, than mixed signal, mainly when the noise source is analyzed. A post-processing methodology is used on the new dataset to identify the start time of leakages, the CPD method presented by Lavielle (2005). The algorithm simultaneously determines all change points, given a statistical control parameter, minimizing a penalized contrast function. The contrast function allows to accurately determine the points where statistical parameter control changes, and the penalty function seeks to reduce the number of points of changes (BRENTAN et al., 2017). In this study, average and standard deviation values are used as statistical control parameter.

### Leakage Location

The leak location process consists of two steps, the first one calibrates the pipe roughness and nodal demands for normal scenario, using the PSO algorithm. The second step uses the results of the calibration process to identify the leakage magnitude and region in the network. The sum squared error of dimensionless pressure, flow and tank level are used as objective function. A single-objective problem F is stated, as presented in equation 3.

$F = \sum_{t=1}^{T} \left[ \sum_{i=1}^{Nn} \left( \frac{P_{t,i}^m - P_{t,i}^o}{\overline{P^o}} \right)^2 + \sum_{j=1}^{Np} \left( \frac{Q_{t,j}^m - Q_{t,j}^o}{\overline{Q^o}} \right)^2 + \sum_{k=1}^{NT} \left( \frac{L_{t,k}^m - L_{t,k}^o}{\overline{T^o}} \right)^2 \right]$ (eq. 3)
where $P_{t,i}^m$ is the modelled pressure in the node $i$ in a network with $Nn$ monitored nodes, at time step $t$ in a simulation during $T$. $P_{t,i}^o$ is the observed pressure for the same node $i$ at time step $t$. $\overline{P^o}$ is the average pressure of observed dataset. $Q_{t,j}^m$ is the modelled flow in the pipe $j$, in a network with $Np$ monitored pipes, $Q_{t,j}^o$ is the observed flow in the same pipe and $\overline{Q^o}$ is the average flow of observed flow. Finally, $L_{t,k}^m$ is the modelled tank level at tank $k$ at time step $t$ in a network with $NT$ monitored tanks, and $L_{t,k}^o$ is the observed tank level at the same tank.

The fist calibration process has as variables all nodes and all pipes of the water system. PSO algorithm (EBERHART and KENNEDY, 1995) is applied to minimize the single-objective problem. A swarm or population is compound of possible solutions of the problem. In the first iteration of the method the particles' velocity and position are randomly distributed within a range of interest. The objective function is evaluated and the velocity and position are updated. The velocity update is calculated as a linear combination of three components, as presented in the equation (4).

$$v_{id}^{n+1} = \left[w.v_{id}^{n} + \frac{c_1.r_1(p_{id}^{n}-x_{id}^{n})}{\Delta t} + \frac{c_2.r_2(g_{id}^{n}-x_{id}^{n})}{\Delta t}\right] \text{ (eq. 4)}$$

$$x_{id}^{n+1} = x_{id}^{n} + v_{id}^{n+1}.\Delta t \text{ (eq. 5)}$$

where $d = 1,2,\ldots,D$, is the dimension of decision variable vector, $n = 1,2,\ldots,N$ and $N$ is the iteration number. $r_1$ and $r_2$ are random numbers within the range $[0,1]$. $P_i = (p_{i1}, p_{i2}, \ldots, p_{iD})$ is the best position occupied by a particle and $G_i = (g_{i1}, g_{i2}, \ldots, g_{iD})$ is the best position occupied by the swarm. Part of the update is weighted by the cognitive coefficient $(c_1)$ social coefficient $(c_2)$, Finally, the velocity $v_{id}^{n}$ is weighted by the coefficient of inertia (w).

## BATTLEDIM – CASE STUDY

The methodology is applied to the L-Town network provided in BattLeDIM. The dataset is compound by 1 tank water level sensor, 3 flow sensors, 33 pressure sensors and 82 Automated Metered Readings. These points are monitored in 2018 and 2019 every 5 minutes. For 2018, 10 leakages are fixed, providing for participants the pipe fixed and the time.

The application of the ICA is done with different combination of pressure and flow. The choice of pairs (pressure and flow data) follows the division of the network into areas proposed by BattLeDIM, (e.g areas A, B and C). Area A and B are supplied by two reservoirs, while area C is supplied by a tank.

The ICA method resulted, for each application, two sets of components in which they are analyzed by the CPD technique and several changes points in the signals are indicated. Figure 1 shows the results of these applications in one pair of pressre/flow data. The first component follows the demand pattern. The second component has variations in the data directly associated with the leaks.

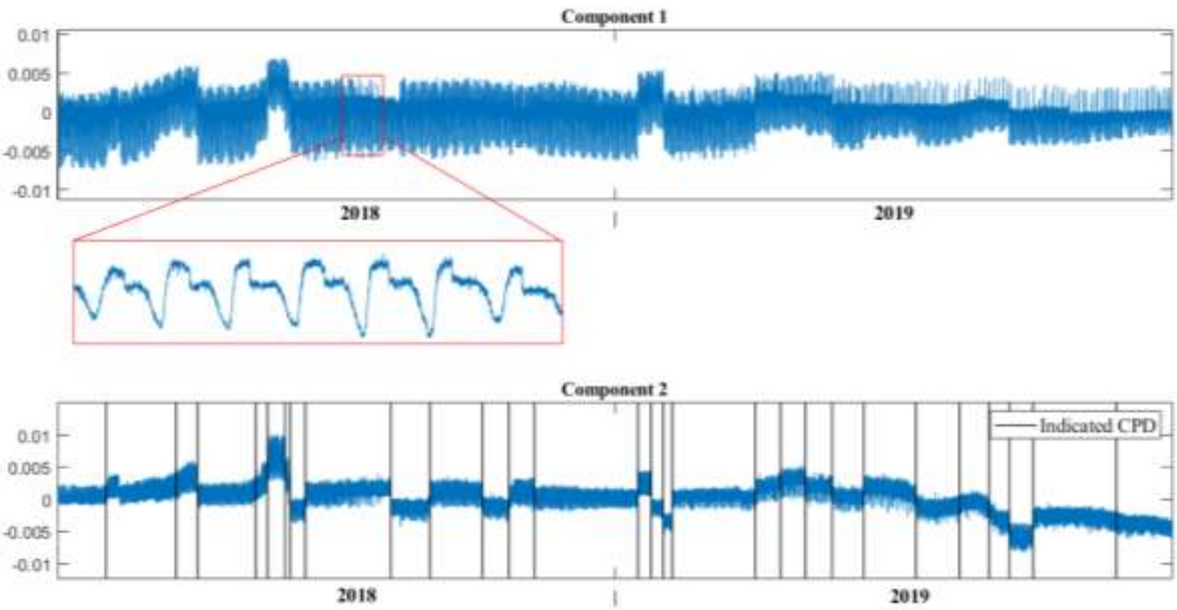


**Figure 1** – Leak detection - application of ICA and CPD

The CPD indications are used to built the report file, with the start time for possible leakages in 2019.

In BattleDIM' network, the calibration process demand and roughness are composed by 1687 variables (782 nodal demands and 905 roughness). One week of pressure, flow and tank level data are used for the process. Figure 02 shows the comparison of modelled and observed flow in reservoir R1 and R2.

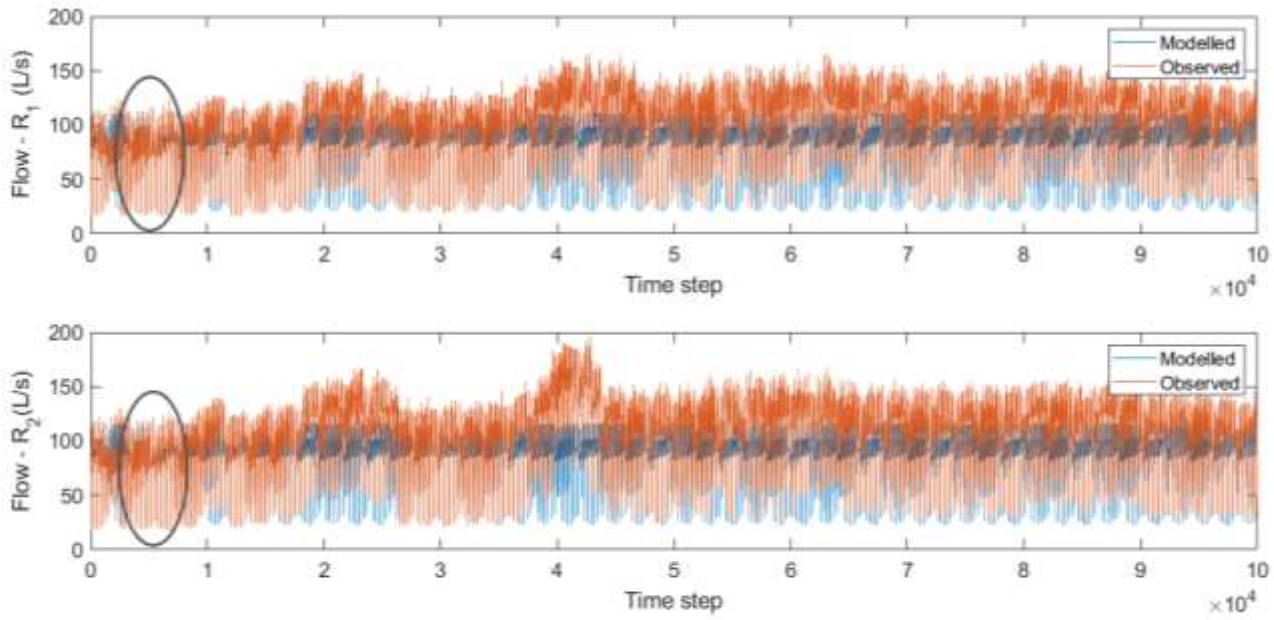


**Figure 2**– Leak Isolation - application of PSO for calibration of nodal demand and pipe roughness.

It's possible to verify that leakages increase the outlet flow of reservoirs R1 and R2, making the difference between modelled and observed data increase. This fact can be used to estimate the magnitude and the start time of the leakages in the network. The black ellipsis highlights the period used for calibration.
With the results of the first calibration, the region of leaks is determined, in a second calibration process. A unique and constant leakage flow and node are calibrated. While the simplification of constant flow makes the hydraulic simulation softer, this also can affect the precision of the results. For the 10 leakages of 2018, the marginal error, calculated as the average Euclidian distance between the leaking nodes and the real leaking pipes is around 300m.

**Keywords:** Leakage Detection and Localization, Blind Source Separation, Change Point Detection, Calibration, Particle Swarm Optimization.

## SUMMARY

Leakage detection and localization is one of the most important challenges of water companies. This work presents a multi-step methodology for contributing with discussions in this field. For the Battle of Leakage Detection and Isolation (BattleDIM), Independent Component Analysis, a blind source separation algorithm, is applied on hydraulic data, for source separation helping to identify water consumption and noise sources. The results show that noise source is significantly affected by leakages. In a second stage of leak detection methodology, a statistical controlling method, Abrupt Change Point Detection (CPD), is applied using noise data as input. CDP assumes that changes in data affect the statistical properties of the dataset, such as average and standard deviation. Calibration process is applied in two different stages in order to localize leakages. The first one, a dataset of normal condition is used to calibrate roughness and base demand using Particle Swarm Optimization algorithm linked to Epanet hydraulic simulator. The results are used on a second stage, where pressure

and flow during detected leakages by the ICA-CDP algorithm is used for calibration of a leakage event. A 24h dataset is used and leakage flow and nodal leaking position are calibrated. The main objective of this stage is to identify the region of leaking pipe.

## REFERENCES


AHMAD, Fawad et al. Illumination normalization using independent component analysis and filtering. The Imaging Science Journal, v. 65, n. 5, p. 308-313, 2017.

BIREK, Lech; PETROVIC, Dobrila; BOYLAN, John. Water leakage forecasting: the application of a modified fuzzy evolving algorithm. Applied Soft Computing, v. 14, p. 305-315, 2014.

BRENTAN, Bruno M. et al. On-line cyberattack detection in water networks through state forecasting and control by pattern recognition. In: World Environmental and Water Resources Congress 2017. 2017. p. 583-592.

CAPUANO, P. et al. Convolutive independent component analysis for processing massive datasets: a case study at Campi Flegrei (Italy). Natural Hazards, v. 86, n. 2, p. 417-429, 2017.

EBERHART, R., & Kennedy, J. (1995, November). Particle swarm optimization. In Proceedings of the IEEE international conference on neural networks (Vol. 4, pp. 1942-1948). Citeseer.

HAGHIGHI, Ali; RAMOS, Helena M. Detection of leakage freshwater and friction factor calibration in drinking networks using central force optimization. Water resources management, v. 26, n. 8, p. 2347-2363, 2012.

LANGLOIS, Dominic, et al. Dominique. An introduction to independent component analysis: InfoMax and FastICA algorithms. Tutorials in Quantitative Methods for Psychology, v. 6, n. 1, p. 31-38, 2010.

LAVIELLE, Marc. Using penalized contrasts for the change-point problem. Signal processing, v. 85, n. 8, p. 1501-1510, 2005.

MALA-JETMAROVA, H., Sultanova, N., & Savic, D. (2017). Lost in optimisation of water distribution systems? A literature review of system operation. Environmental modelling & software, 93, 209-254.

MOUNCE, Stephen R.; MACHELL, John. Burst detection using hydraulic data from water distribution systems with artificial neural networks. Urban Water Journal, v. 3, n. 1, p. 21-31, 2006.

SAVIC, D. A., Kapelan, Z. S., & Jonkergouw, P. M. (2009). Quo vadis water distribution model calibration?. Urban Water Journal, 6(1), 3-22.

WESSEL, Jan R. Testing multiple psychological processes for common neural mechanisms using EEG and independent component analysis. Brain topography, v. 31, n. 1, p. 90-100, 2018.

WU, Yipeng et al. Burst detection in district metering areas using a data driven clustering algorithm. Water research, v. 100, p. 28-37, 2016.

YE, Guoliang; FENNER, Richard Andrew. Kalman filtering of hydraulic measurements for burst detection in water distribution systems. Journal of pipeline systems engineering and practice, v. 2, n. 1, p. 14-22, 2011.

ZHENG, Chengzhi; GAO, Jinliang; HE, Wenjie. Leakage discharge separation from the total water supply flow in water distribution system by constrained independent component analysis (CICA). Water Science and Technology: Water Supply, v. 16, n. 3, p. 611-619, 2016.