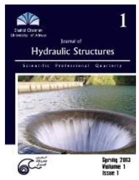
Journal of Hydraulic Structures
***J. Hydraul. Struct.*, 2016; 2(1):58-70**
DOI: 10.22055/jhs.2016.12651



Calculation of Leakage in Water Supply Network Based on Blind Source Separation Theory

JinliangGao[1a]
ShihuaQi[1]
Jun Nan[1]
Juanjuan Li[1]

Abstract

The economic and environmental losses due to serious leakage in the urban water supply network have increased the effort to control the water leakage. However, current methods for leakage estimation are inaccurate leading to the development of ineffective leakage controls. Therefore, this study proposes a method based on the blind source separation theory (BSS) to calculate the leakage of water supply network. The method uses fast independent component analysis (FastICA) algorithm to separate flow signal of laboratory and practical measuring area, adopts trend similarity to solve the uncertainty of separation sequence to get hourly change curve of user usage and physical leakage, and embeds the leakage model into amplitude optimization model to solve amplitude uncertainty to obtain physical leakage value. The study found that the estimation of leakage level using the blind source separation is reasonably accurate and facilitates the identification of the subsequent reduction in water leakage. This can provide scientific evidence for leakage reduction and the investment of pressure relief devices in the next stage.

Keywords: blind source separation theory (BSS); fast independent component analysis (FastICA); leakage simulation; water consumption; water supply network (WSN); water leakage.

Received: 24 February 2016; Accepted: 2 July 2016

1. Introduction

Leakage in water supply network is a general problem in water supply enterprises and different measures such as Minimum Night Flow (Tabesh 2009), "Water Balance" (AWWA 1999; Winarni 2009), Network Leakage Hydraulic Model (Almandoz, et al. 2005) and Regional

[1] Harbin Institute of Technology, No. 73 Huanghe Rd, Harbin, 150090, P.R. China
[a] Corresponding author; JinliangGao. Tel.: 0086-13936477968; fax: 0086-451-82311926.
E-mail address:gjl@hit.edu.cn

Leakage Estimation (Buchberger & Nadimpalli 2004)have been adopted to calculate the physical leakage. However, the methods are generally ineffective due to the use of inaccurate analysis of regional leakage quantity and sequential characteristics.

Minimum Night Flow requires data collection at night with no water consumption or during the hour with minimum water consumption (Tabesh 2009), and significant amount of manpower and materials. Consequently, the method requires significant disruption to the regional water supply and has high uncertainty in the actual regional water consumption estimation. Similarly, "Water Balance", proposed by the International Water Association (IWA), calculates the amount of leakage based on the difference between the production and consumption, which needs to install many flow meter or water meter in the branch of water network (AWWA 1999). It also has several drawbacks such as inaccurate customer metering, use of estimated values for unbilled or unmetered consumption, need for the installation of significant amount meters (Winarni 2009). Further, the Water Balance is effective for accumulated leakage and unsuitable to analyze the hourly variation leakage. Similarly, the Network Leakage Hydraulic Model produces subjective results since it primarily depends on the experience (Almandoz, et al. 2005), while, the Regional Leakage Estimation does not predict the real time leakage (Buchberger & Nadimpalli 2004). Therefore, it is necessary to develop alternative methods that can estimate the actual water leakage accurately. One such alternative method for water leakage estimation is the Blind Source Separation(BSS) (Comon & Jutten 2010), which is used for processing signals that have unknown source similar to water consumption and leakage amount. Overall, this study proposes an algorithm to obtain the curve of water leakage changes with time, which is the foundation for reducing water leakage through adjusting hourly pressure of the inlet of water supply network (WSN).

2. Methods

2.1. BBS Theory

BSS, which is a powerful signal processing technology developed in the 1980s, is a combination of artificial neural network, medical signal processing, statistical signal processing and information theory (Comon & Jutten 2010). BSS detects the sensor signal from a multi-input-and-output nonlinear dynamic system, and consequently finds an inverse system to estimate and reconstitute the original signal source (Comon & Jutten 2010). In this process, the original signal and the method that the source signal used to obtain the observation signal are unknown.

2.2 Fast independent component analysis (FastICA)

Comon (1994) suggested the core algorithm, namely independent component analysis (ICA),for blind source separation. It synchronously observes overlapping signals in multi channels and separates the observation signals into independent parts to estimate the source signal. Fast independent component analysis (FastICA) is a high-speed algorithm based on ICA proposed by Hyvärinen (Hyvärinen1999; Hyvärinen & Oja 2000). FastICA has fast rate of convergence, and is appropriate for processing non-Gaussian data. It can extract an independent part from the observation signals every time.

2.3 Evaluation index for BSS

The performance of FastICA can be evaluated by assessing two criteria, namely similarity coefficient and performance index. The similarity coefficient was taken as the criterion to evaluate the similarity between separated and real water consumption leakage.

Similarity coefficient between separated consumption and leakage in different operating conditions $s'(t) = [s'_1(t), s'_2(t)]^T$ and the real consumption and leakage $s(t) = [s_1(t), s_2(t)]^T$ was used as an assessment criterion of leakage separation, as shown in equation (1):

$$\xi_{y,s} = \frac{\left|\sum_{t=1}^{T}[s(t) - E\{s(t)\}] \cdot [y(t) - E\{y(t)\}]\right|}{\sqrt{\sum_{t=1}^{T}[s(t) - E\{s(t)\}]^2 \cdot \sum_{t=1}^{T}[y(t) - E\{y(t)\}]^2}} \quad (1)$$

where set y and s 1 and 2, respectively. y=1 represents real water consumption, while y=2 represents real leakage. Similarly, s =1 represents separated water consumption, while s =2 represents separated leakage. When ξ_{11} is close to 1, the real water consumption approaches the separated water consumption value, whereas the real leakage approaches the separated value of leakage when ξ_{22} is approximately 1.

2.4 FastICA for BSS

The signal noise was present from the flow meter in water supply network due to the interference of the external environment. Therefore, preprocessing of the signals was used to guarantee the precision of the signal separation. In this study, FastICA was used to apply blind source separation and the preprocessing included the following:

- The physical leakage was obtained from the total flow of water supply network and the mathematical model was built according to the blind source separation theory, as shown in equation (2):

$$X = AS \quad (2)$$

Where *X* is the observation signal, *A* is the hybrid matrix of the source signals.

- Tabesh, et al. (2002) suggested that the node flow points are actual water consumption and the physical leakage. The model of regional water supply network BSS can be consequently subdivided as shown in equation (3):

$$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}, S = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} q_y \\ q_l \end{bmatrix}, A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad (3)$$

where Q_1 is regional total input volume and Q_2 is main branch volume, q_y and q_l are the actual user water consumption and physical leakage. The total water supply was obtained from the combination of the actual user water consumption and physical leakage. The physical leakage can be separated from the WSN using the FastICA algorithm.

- Whitening was used before separation, because they completely lost connection after whitening, as shown in equation (4):

$$\bar{X}(t) = Q_w X(t) \tag{4}$$

where Q_w is the whitening matrix.

FastICA algorithm, also known as fixed point algorithm, is a non-Gaussian maximization algorithm. It searches the non-Gaussian maximum using the fixed-point iteration theory. This algorithm adopts the Newton iteration method to batch the observational variable at a large number of sampling sites. It uses maximized negative entropy as objective function and processes an independent element from the observation signals every time. The flowchart for FastICA is illustrated in Fig. 1.

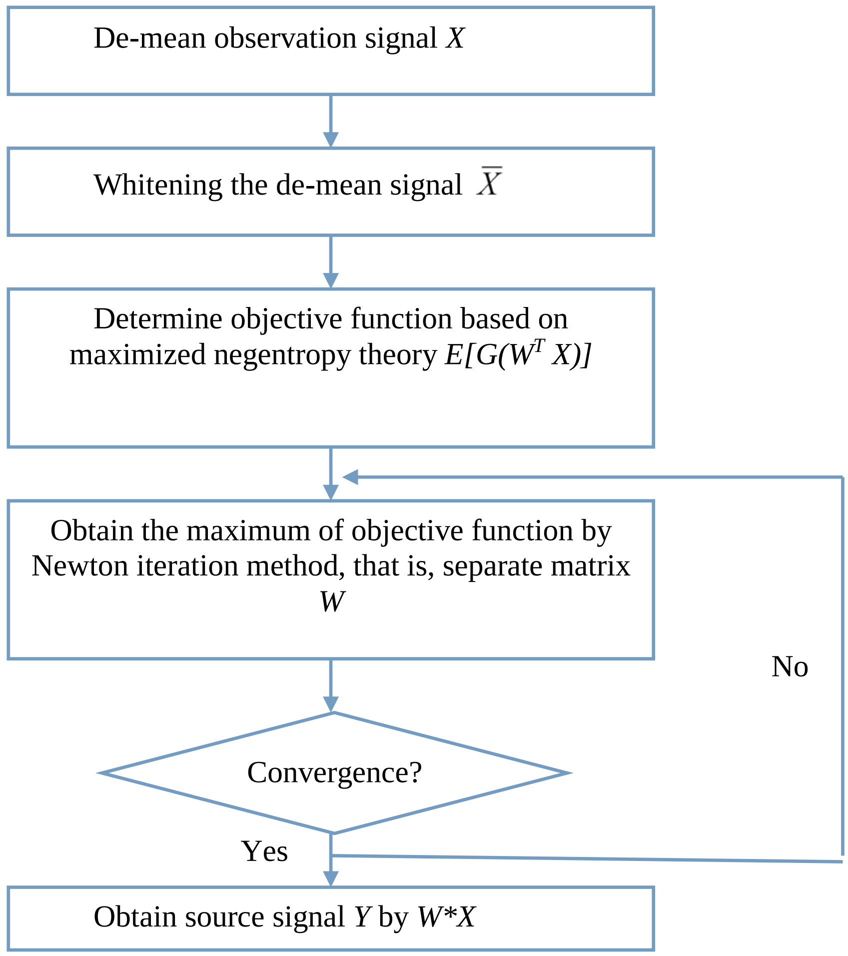


Fig. 1. Processing steps of FastICA

BSS Programming was established by using relevant MATLAB toolbox (Mathworks, 2014).

2.5 Amplitude solution

The uncertainty of the signal obtained by blind source separation contains sequential uncertainty and amplitude uncertainty (Tong 1991). Sequential uncertainty restricts the ability to separate the physical leakage from different components, while the amplitude uncertainty limits the ability to obtain the actual leakage and to analyze the dynamic change in the physical leakage.

In this study, trend similarity was adopted to solve uncertainty of separation sequence. Over determined systems were established to calculate source signal mean and variance based on the law of mass conservation, and solved by the genetic algorithm. The amplitude uncertainty was solved through converting the amplitude into significant physical leakage. The specific solutions of amplitude uncertainty include: embedding the leakage model into amplitude optimization model; setting average pressure for the laboratory pipe network; networking total flow and separated water consumption as given value; setting leakage coefficient and leakage index as unknown quantities of leakage model; setting amplitude translation amount and amplitude proportion coefficient as unknown quantities of the water model; establishing the water balance equation and using multiple iterative operation of genetic algorithm optimization to obtain leakage coefficient, leakage index, amplitude translation and amplitude ratio; and sequentially solving the problem of water consumption amplitude. In this study, the Genetic Algorithm consisted of 4 independent variables, population of 500, crossover rate of 0.8 and mutational rate of 0.1. Water balance equation is shown in equation (5):

$$\begin{cases} Q_i = Q_{y_i} + Q_{l_i} \\ Q_{y_i} = (a_i + x)y \quad i = 1,2,3\ldots 12 \\ Q_{l_i} = k_i P_i^{\beta} \end{cases} \tag{5}$$

where a_i is the mean of separated water consumption in each operating condition, x is the translation of separated water consumption, y is the amplitude proportionality coefficient of separated water consumption(L/s), P_i is the average water head of pipe network(mH$_2$O), Q_i is the total flow in pipe network(L/s), Q_{y_i} is the separated water consumption of pipe network(L/s) and Q_{l_i} is the actual leakage of pipe network(L/s).

2.6 Laboratory based case study

An experimental platform of water supply network leakage (Fig. 2) was established to apply blind source separation for the water supply network leakage separation. This platform was able to simulate different pipe network scenarios by controlling the speed of the pumps and on-off control of the valves such as single water source loop pipe network leakage simulation, single water source ring combination of pipeline leakage simulation and multi- water sources loop pipe network leakage simulation. The water supply pipes were setup vertically to simulate the real pipes with different heights. A control cabinet was used to control the water pumps, and 3 horizontal multi-stage speed centrifugal pumps were used to supply water for the pipe network. Additionally, a water tank was used as the water supply source and backwater device. The pipes used in the water supply network were made up of Unplasticised Polyvinyl Chloride (UPVC) and an online pressure gauge was installed in the network to detect the pipe pressure.

Furthermore, a turbine flow meter was used at the outlet of the pump.

Fig. 2. Experimental platform of Leakage Separation of Water Supply Network

Water consuming nodes and leakage points are shown in Fig.3. There network consisted of 4 nodes in the water supply network and 3 water meters at every node to quantify the water consumption and the leakage. B, D, F and H represent users, while A, C, E and G represent leaking points. The leaking points were controlled by on-off controls at valves A, C, E and G. The leakage can be determined by controlling the level of valves A, C, E and G.

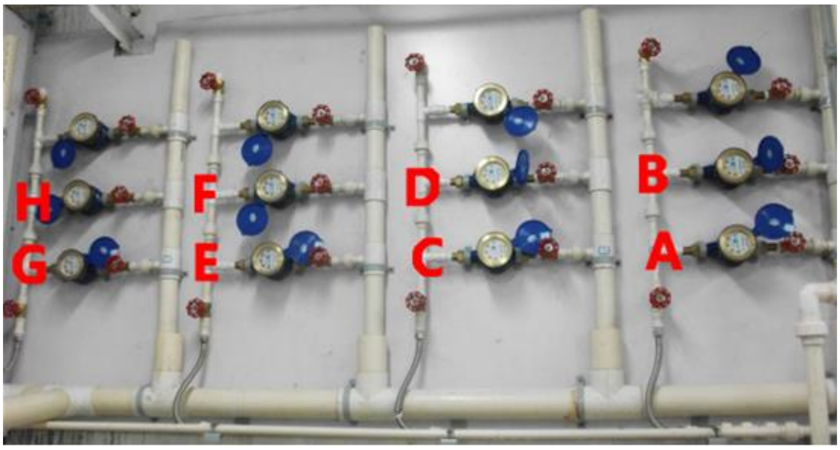


Fig. 3. Water Consuming Nodes and Leaking Points

Considering the features of a typical water supply network, the leakage separation of network was set up in laboratory. The Supervisory Control And Data Acquisition (SCADA) system data were considered as the observation signals, while the real water consumption and physical leakage were considered as the source signals. While the level of valves A, C, E and G were kept constant, the water consumption model was changed by altering the level of valves B, D, F and H. However, the frequencies of pumps and the flow rates were not changed. Additionally, the real water consumption and leakage were recorded, while the amount of leakage separated by applying BSS. The separation performance was evaluated by comparing the separated data with the real water consumption and leakage.

In a multi-water source loop pipe network leakage, the SCADA online data was used to record the flow observation signals of 9648 sets in a network under 14 operating conditions, as shown in Fig. 4.

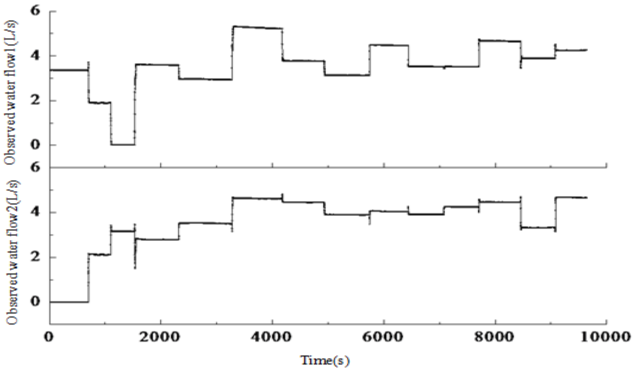


Fig. 4. Observation Signals (1. regional total input volume, 2. main branch volume) of Flow Meters

The separated signals of FastICA are shown in Fig. 5.

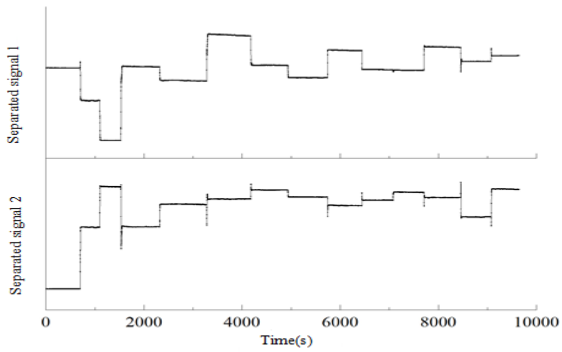


Fig. 5. Variation Trend of Separated Signals

As can be seen from Fig.5, the variation trend that separated signal 1, which is regarded as separation value for water consumption, is similar with the real water consumption. Multiple iterative operation of GA optimization was used to obtain the amplitude translation and proportionality coefficient in order to investigate the amplitude uncertainty. Consequently, the simulated water consumption of multi-water source loop network was obtained using the genetic algorithm optimization, while the simulated physical leakage was calculated from the difference between the production and consumption. The relative error between the simulated and real water consumption and leakage is shown in Table 1.

Table 1. Simulated Water Consumption and Leakage with BSS

Operating Condition	Real Leakage (L/s)	Simulated Leakage (L/s)	Relative Error (%)	Real Water Consumption (L/s)	Simulated Water Consumption (L/s)	Relative Error (%)
1	0.880	0.742	15.71	0.000	0.246	—
2	1.189	1.085	8.69	0.000	0.103	—
3	0.844	0.779	7.70	0.000	0.165	—
4	0.914	0.864	5.47	0.954	1.004	5.25
5	0.922	0.856	7.08	0.991	1.057	6.58
6	0.573	0.626	9.13	2.442	2.289	6.24
7	0.494	0.526	6.61	1.966	1.903	3.19
8	0.819	0.771	5.89	1.264	1.313	3.82
9	0.980	1.015	3.54	1.557	1.522	2.23
10	0.911	0.911	0.01	1.308	1.308	0.01
11	0.733	0.797	8.67	1.566	1.503	4.06
12	0.883	0.983	11.35	1.797	1.696	5.58
13	0.619	0.639	3.27	1.509	1.489	1.34
14	0.715	0.791	10.52	1.947	1.872	3.87
Total Flow(L/s)	11.476	11.385	0.80	17.301	17.471	0.98

From Table 1, it can be seen that the simulated water consumption fluctuates around the real water consumption, with a relative error ranging from0.01% to 6.58%. The simulated leakage fluctuates around the real leakage, with a relative error ranging from0.01% to 15.71%. The simulated total leakage error is 0.98%, while the simulated water consumption error is 0.8%.

The similarity coefficient for the real and simulated leakage is 0.876, while the similarity coefficient for real and simulated water consumption is 0.929. This suggests that the separation performance is fairly satisfied.

3. Case study

A secondary metering area in the urban water supply network of a southern-China coastal city was used as the case study. The urban water supply system in the area had SCADA and Geographic Information System (GIS) systems and had about 10 km of total length of the water supply network. There were inlet flow meters (A and B), a neighborhood inlet flow meter (C) and three pressure monitoring points (P1, P2 and P3) as shown in Fig. 6.

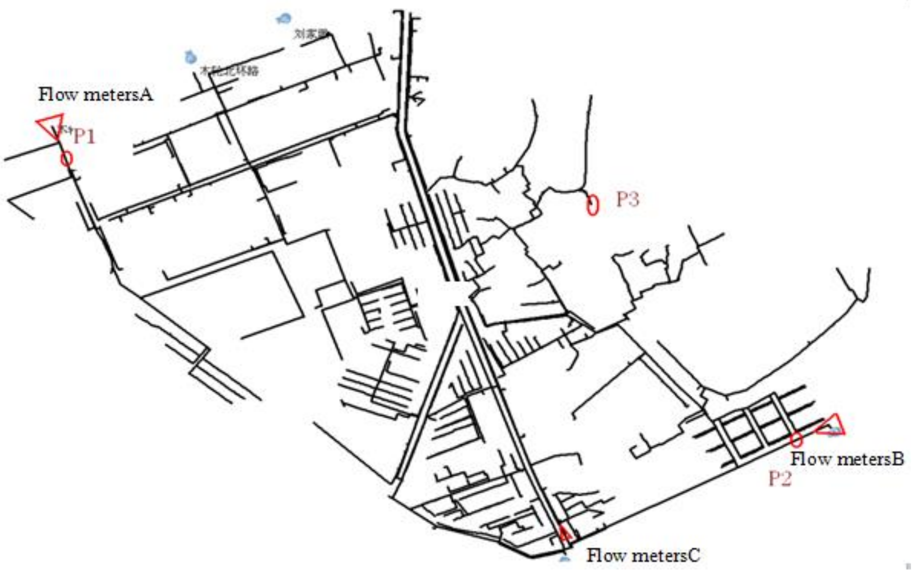


Fig. 6. Network Topology of Metering Area

There are multi-water source and branch-loop combined water supply networks in this area. The data for three flow meters on date X can be obtained and shown in Fig.7(a), while the pressure data is plotted in Fig. 7(b).

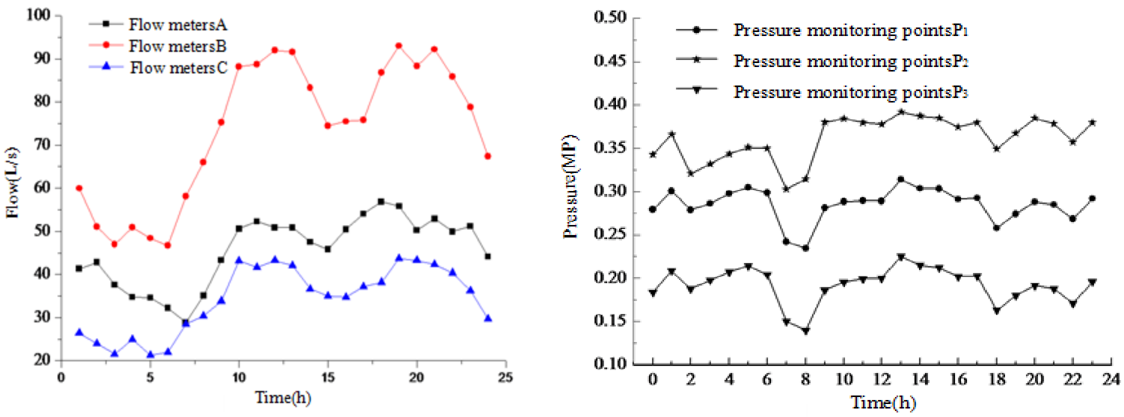


Fig. 7. a. Flow meter measured data, b. Pressure data

According to the laboratory scale leakage separation method study, the two flow observation signals were the inputs, where the sum of flow through meters A and B was one input, while the amount of flow through meter C was another. By using FastICA, the variation trend shown in Fig. 8 can be obtained.

The water balance equation was developed based on the energy conservation principles in order to eliminate the uncertainty in the amplitude of separated water consumption. Genetic optimization leads to the simulated water consumption and simulated leakage, as shown in Table 2.

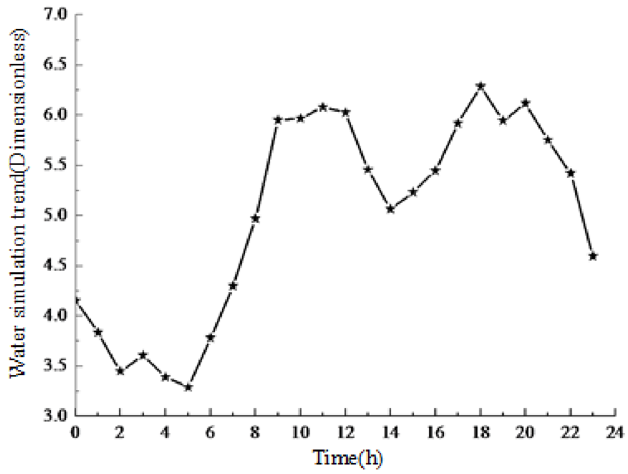


Fig. 8. Variation trend of separated water consumption

Table 2. Simulated Water Consumption and Leakage in 24 hours

Time(h)	1	2	3	4	5	6
Simulated Water Consumption(l/s)	83.214	75.908	67.053	70.669	65.700	63.366
Simulated Leakage(l/s)	16.033	16.814	15.900	16.228	16.578	16.808
Time(h)	7	8	9	10	11	12
Simulated Water Consumption (l/s)	74.689	86.536	102.011	124.556	124.958	127.578
Simulated Leakage (l/s)	16.589	14.753	14.600	16.400	16.647	16.681
Time(h)	13	14	15	16	17	18
Simulated Water Consumption (l/s)	126.317	113.261	104.172	108.069	112.978	123.864
Simulated Leakage (l/s)	16.669	17.389	17.106	17.039	16.692	16.758
Time(h)	19	20	21	22	23	24
Simulated Water Consumption (l/s)	132.267	124.472	128.453	119.994	112.350	93.383
Simulated Leakage (l/s)	15.531	16.131	16.575	16.433	15.828	16.642

The actual total volume of water supply on date X in this metering area was 10292.29 m^3, and the simulated leakage volume after blind source separation is 1414.16 m^3. Thus, the simulated leakage percentage was 13.74% of the actual water supply volume. The water audit analysis showed that the daily production and marketing difference on date X was 15.68%. Since 80% of the production and marketing difference is leakage, the leakage percentage was 12.54%. The error between the leakage percentage from FastICA and the one shown by water audit was only 1.2%. Therefore, it can be concluded that calculating the leakage percentage using separating leakage with FastICA algorithm is reliable.

4. Results

This study applied BSS for both laboratory water supply network and regional water supply network. In the laboratory leakage separation experiment, the relative error range between simulated and real leakage is only 0.01%~15.71%, while the error in simulated total leakage is 0.98%.Furthermore, the similarity coefficient between simulated and real leakage is 0.876. In the regional water supply network, the error between calculated leakage rate and the rate given by the water audit is only 1.2%. A significant portion of the data was obtained by laboratory leakage separation experiment, which was limited to topology of WSN, node flow and pressure control. The control mode of these experiments was achieved by adjusting water valves and pump speeds. Therefore, the reliability of the results needs further investigation using different experimental water network and different conditions. In the actual field test with traditional methods, it is difficult to detect hourly amount of water leakage resulting in unreliable and inaccurate test results. The leakage amount in the main branch of WSN needs to be measured in Minimum Night Flow. However, such measurement is difficult since it requires the installation of a significant number of water meters, especially in complex loop WSN in china. Consequently, the measured flow data were all off-line with low accuracy. In contrast, real-time monitoring data is reliable and accurate. However, establishing the remote communication is a major issue. In addition, the installation of excessive water meters would increase the experimental error.

Furthermore, the water leakage obtained from the pressure reducing experiment is unreliable, since the changing pressure would have affected water consumption. Similarly, node leakage model only considers single node leakage rather than total leakage of each node of regional WSN. However, in this study, the estimation of total leakage of regional WSN was calculated based on BSS with the measurement of two water volume (usually is regional total input volume and main branch volume) and one network average pressure.

5. Discussion

It is worthy to note that the results obtained in this study were relatively more accurate compared to the results obtained using the traditional methods. Water consumption and node water leakage has different characteristics of pressure, trend, requirement and output. The measurement value of two water volume and one network average pressure were interacted results of water consumption, node water leakage, pipeline head loss and node pressure. The leakage coefficient and index can be obtained based on the node leakage model using measured volume of regional water leakage, then the coefficient and index can be used to adjust pressure and control leakage.

This study proposed that system input volume consisted of two variation of water amount, namely water consumption and water leakage. Therefore, the water leakage curves were obtained each hour using FastICA. Kalman Filter Algorithm, Wavelet Analysis Algorithm or any other algorithm are recommended to separate water leakage from system input volume for further study, and compare the efficiency and accuracy of each algorithm.

6. Conclusion

In water supply network, linear relationship is difficult to obtain. This study investigated a method that uses BSS as a substitute for nonlinear relationship. BSS can only provide the trend of time, while the obtaining the trend of amplitude was not possible. Calculation of leakage in WSN based on BSS is simulated in both laboratory and field. Hourly water leakage of WSN can be simply obtained based on BSS using only two water volume (usually is regional input volume and main branch volume) and one network average pressure. The results concluded that the BSS is a relatively accurate and effective method to determine the leakage and to control the network leakage. BSS also can distribute all leakage volume to each working condition and can provide accurate data support of leakage characteristics analysis and leakage control hydraulic model. This can provide solid support for leakage reduction purposes and for the investment in pressure relief devices in the next stage.

7. Acknowledgement

This research was supported by the National Water Pollution Control and Treatment Science and Technology Major Project (2014ZX07405002C), National Natural Science Foundation of China (51278148) and The Union Project of Industry-study-research of Guangdong Provincial Department of Education (2011A090200040). The partial paper are from WDSA2014 "Study on Leakage Rate in Water Distribution Network Using Fast Independent Component Analysis".

Reference

1. Almandoz, J., Cabrera, E., Arregui, F., et al. 2005 Leakage assessment through water distribution network simulation. *Journal of water resources planning and management*. 131(6), 458-466.
2. American Water Works Association , 1999 2nd edition, Manual M36, *Water audits and leak detection*, AWWA.
3. Buchberger, S. & Nadimpali, G. 2004 Leak estimation in water distribution system by statistical analysis of flow readings. *J. Water Resour. Plann. and Manage*. 130(4), 321-329.
4. Comon, P. 1994 Independent component analysis, a new concept? *J. Signal Processing*. 36(3), 287-314.
5. Comon, P. & Jutten, C. (Eds.). 2010 Handbook of blind source separation: Independent com ponent analysis and applications. Academic press.
6. Gao, J., Qi, S. & Wu, W. 2014 Study on leakage rate in water distribution network using fast independent component analysis. *WDSA 2014, Bari, Italy: Procedia Engineering.* 89(2014), 934-941.
7. Hyvärinen, A. 1999 Fast and robust fixed-point algorithms for independent component analysis. *J. Neural Networks.* 10(3), 626-634.
8. Hyvärinen, A. & Oja, E. 2000 Independent component analysis: algorithms and applications. *J. Neural Networks.* 13(4-5), 411-430.
9. Mathworks. 2014 < http://www.mathworks.com/matlabcentral/fileexchange/47069-convolutive-bss > (Accessed 04 May 2015)
10. Rossman, L. 2000 *EPANET User's Manual*. United States Environmental Protection Agency, Cincinnati, U.S.A.
11. Tabesh, M., Tanyimboh, T.T. & Burrows, R. 2002 Head driven simulation of water supply networks. *IJE Transactions A: Basics.* 15(1), 11-22.
12. Tabesh, M., Asadiyani, A.H. & Burrows, R. 2009 An Integrated Model to Evaluate Losses in Water Distribution Systems. *Water Resources Management.* 23(3), 477-492

13. Tong, L., Liu, R. & Soon, V.C. 1991 Inderminacy and identifiability of blind identification. *IEEE Transactions on Circuits and Systems*. 38(5), 499-509.

14. Winarni, W. 2009 Infrastructure Leakage Index (ILI) as water losses indicator. *Civil Engineeri ng Dimension.* 11(2), 126-134.