

Inverse Optimization based Detection of Leaks from Simulated Pressure in Water Networks, Part 2: Analysis for Two Leaks

Peace K. Amoatey,¹ András Bárdossy² and Heidrun Steinmetz³

¹University of Ghana, Legon Accra, Ghana; ²University of Stuttgart, Stuttgart, Germany; ³Technische Universität Kaiserslautern, Kaiserslautern, Germany.

Abstract

Water networks lose significant volumes of water from the distribution network. This is because most water networks, especially in developing countries, do not have the techniques, equipment or monitoring systems required to enable the detection of leaks. An optimization-based approach is used to model leakage detection in water networks from simulated network pressures. This study explores the ability of the proposed model to simultaneously detect two leaks within the water distribution network and to determine the number of reference points required for the leaks to be detected. By changing the emitter property in the network hydraulic model, reference and simulated pressures are generated. Two nodes are injected with leaks of known magnitude and the model attempts to find the two reference nodes. Similarly, several simulated references are generated from stochastically simulated leaks. The reference and simulated pressures are compared with respect to selected observation or reference points within the network. The model detects leaks using the optimization functions for which the sum of squared errors (SSE) is equal to zero. For the range of leak sizes and the two scenarios considered, the model performs poorly (13%) in the *Same Vicinity* scenario and just below average (41%) in the leaks occurring in the *Far Apart* scenario. This may be due to the water network configuration affecting the sensitivity of the model to the leak sizes being considered. Relatively smaller leak sizes and various network configurations will be further investigated. Additionally, it was determined that a minimum of eight observation points is required for leaks to be detected.

1 Introduction

The occurrence of leaks in water networks disrupts the water supply as the pressure in the water network drops. This may significantly affect water users as part of the network is shut down or isolated during repairs (Palau et al. 2012). Water utilities require equipment or tools for the continuous measurement of flow and pressure for the easy detection of leaks. In the absence of such equipment, active leakage control (ALC) is practiced. It involves hunting for leaks in water networks using equipment which operates mainly on acoustic principles, such as listening sticks, noise correlators and data loggers (Hunaidi 2012).

Effective step testing as recommended by Boulos et al. (2008) is a practice where the test area with leakage is bracketed into a tight branched network with a flow meter installed on its input main. However, the use of such equipment and technology is much more effective and less time-consuming if the area where to look for a leak can be predicted.

Leak detection modeling in pipeline systems such as water, oil and gas lines has been widely studied. With respect to water networks, the most popular modeling approaches include: transients (Shucksmith 2011; Colombo et al. 2009; Vitkovsky et al. 2007;

Wu and Sage 2006; Wang et al. 2002; Liggett and Chen 1994); probabilistic (Poulakis et al. 2003; Vanik et al. 2000); multivariate statistical analysis, specifically principal component analysis, PCA (Palau et al. 2012; Gertler et al. 2010; Nowicki et al. 2012; Nowicki and Grochowski 2011); support vector machines (De Silva et al. 2011; Mashford et al. 2009); and simulation-optimization inverse modeling methods (Mirats-Tur et al. 2014; Jasper et al. 2013).

Pressure transients (pressure waves) are created by the sudden interruption of a section of the water network by shutting down or opening valves. The transients propagate signals which are reflected in the presence of a leak (Brunone and Ferrante 2001; Liggett and Chen 1994). The reflected signal and the arrival time are measured to find the position of the leak. Transient methods are not difficult but include the possibility of leak signals being confused with noise from the surroundings (Shucksmith 2011; Colombo et al. 2009; Vitkovsky et al. 2007).

The challenge of noise is predominantly in direct transient methods (DTM); the inverse transient approach for detecting leaks in water networks was investigated by Wu et al. (2011); Karney et al. (2009); Kapelan et al. (2003); Vitkovsky et al. (2001); and Covas et al. (2001). Hydraulic transients of known intensity

are injected at a given location to obtain pressure signals that are recorded at pre-determined locations within the network. The leak position is determined with the help of solvers that use optimization routines. Although the results were better than DTM, transient methods are not generally applicable to complex pipe networks with branches and loops with varying demands (Wu et al. 2011; Karney et al. 2009; Kapelan et al. 2003; Vitkovsky et al. 2001; Covas et al. 2001; Vitkovsky et al. 2007).

Poulakis et al. (2003) and Vanik et al. (2000) explored the use of flow test data and Bayesian modeling to find the magnitude and location of a possible leakage event. The authors additionally investigated the uncertainties in the estimated leakage. They found that the method required a statistical model formulation which was tedious and time consuming for an entire water distribution network.

Support vector machines (SVMs), which are multivariate statistical methods that identify patterns in the negative pressure wave and build a classification system of the wave structure to effectively distinguish between normal adjustment states and leak states, have been explored for leak detection in water networks (De Silva et al. 2011; Mashford et al. 2009). An SVM has the advantage of being applicable to an entire network. However, the presence of sensors in the network is key to successful leak detection and this is a challenge for most developing country water networks.

PCA performs the feature extraction from flow or pressure measurements in the water network over a period (Hsu et al. 2010). It works on the premise that both leaks and demand in a water distribution system occur at nodes (colinear). The magnitude of the leak is proportional to the projection of the total inflow, consumption, and leak displaced in each principal component (PC) direction (Palau et al. 2012; 2004; Gertler et al. 2010). The PCs are the eigenvectors of the covariance matrix (Nowicki et al. 2012; Nowicki and Grochowski 2011; Gertler et al. 2010).

Multivariate methods perform significantly well. SVM in the De Silva et al. (2011) and Mashford et al. (2009) studies detect leaks within 300 m of the actual leak while Nowicki et al. (2012) and Nowicki and Grochowski (2011) report that PCA can accurately detect leaks within a reasonable distance. However, that reliability depended on the optimal location of sensors and this is a challenge in developing countries.

A number of studies have used optimization methods (Sousa et al. 2015; Mirats-Tur et al. 2014; Ishido and Takahashi 2014; Jasper et al. 2013). A simulation–optimization inverse modeling approach which characterizes and detects leak locations was studied by Jasper et al. (2013). In their study, the difference between real and simulated measurements for a known sensor configuration is minimized in order to find the leak locations. Mirats-Tur et al. (2014) varied the demand in any node in the water network, performed sensitivity analysis of pressure measurements and found that leaks were detected within 150 m of the actual leaks, a better performance than the SVM which detected leaks within 300 m of the actual leak (De Silva et al. 2011;

Mashford et al. 2009). Ishido and Takahashi (2014) introduced a new algorithm for real time leak detection and found that the method can be successfully applied to a wider class of water distribution networks. A burst leak is identified by the ratio of two differences of pressure measurements known as the head loss ratio (HLR).

Sousa et al. (2015) used an optimization model which minimized the differences between measured and estimated pressures at the monitoring points to identify leaky pipes. The authors solved the objective function with a simulated annealing algorithm. It was found that the method is reliable, fast and easy to implement. They obtained encouraging results for single or for two simultaneous leaks at different locations.

The method used in this study is similar to that of Sousa et al. (2015) in that single and two simultaneous leaks are investigated. This study is similar to Ishido and Takahashi (2014) in that it uses differences in pressure. It is similar to the Jasper et al. (2013) study in two ways. First, it is an inverse modeling approach and second, a leak is located by minimizing pressure differences between real and simulated measurements.

This study is also similar to De Silva et al (2011) and Mashford et al. (2009) in that pressures are simulated in the hydraulic model of the network by changing the emitter coefficient of network nodes. Moreover, the range of leak sizes tested (0.1 L/s/m–0.3 L/s/m) is the same in both studies. This study is similar to Wu et al. (2011), Karney et al. (2009), Kapelan et al. (2003) and Covas et al. (2001) in that they also used a (different) inverse modeling and optimization approach.

This study differs from the aforementioned studies in two ways: the simulated pressures are randomly (Monte Carlo) generated in the hydraulic model of the network by randomly varying the emitter coefficient; and the leak is detected by minimizing the differences between reference pressures (generated by injecting a leak of known magnitude assigned to a node) and the randomly simulated pressures within the network.

This study differs from Jasper et al. (2013) who used a known sensor configuration while we use observation points perceived to be potential sensor locations or measurement points. We compare real and simulated pressures instead of two sets of simulated pressures. This work also differs from Sousa et al. (2015) in that it compares real and estimated pressures instead of two sets of simulated pressures.

We differ from Ishido and Takahashi (2014) in that we use the ratio of pressure differences instead of two sets of simulated pressures. Our approach differs from Wu et al. (2011), Karney et al. (2009), Kapelan et al. (2003) and Covas et al. (2001) because we simulate pressure by changing nodal emitter coefficients within the network model instead of creating pressure transients by suddenly opening and shutting valves.

Finally, we differ from De Silva et al. (2011) and Mashford et al. (2009) by incrementing the emitter coefficient in steps of 0.001 instead of using a Monte Carlo generation of nodal pressures. De Silva et al. (2011) and Mashford et al. (2009) used SVMs to detect

leaks in the network while we use an optimization approach to detect leaks.

The optimization method used here is relatively fast and computationally less demanding. It is also applicable to various kinds of water networks even when sensors are not installed. The method was tested first for detecting a single leak in the water network. It was found that the model could better detect a leak in a dense network than in a less dense one. Additionally, the model detects a leak within 400 m of the actual simulated leak (Amoatey 2015).

In Paper 1 (Amoatey et al. 2018) the proposed leak detection modeling method was used to detect a single leak in the network. The performance of the model for leaks in a gridded network was better than in the branched network. In this paper the model is used to detect two leaks in a network. In addition, we explore the minimum size of a leak that is detectable and how many observation points are necessary for the leak to be detected. In this paper, leak sizes in the range of 2 m³/h to 6 m³/h are tested. Based on the results, relatively smaller leak sizes will be further tested.

2 Methodology: Inverse Modeling of Simulated Leaks

A minimization-related approach to the Monte Carlo method is applied to a water network. Simulated pressures p_s are compared with the reference pressures p_r generated from leaks of known magnitude assigned to selected reference (control) nodes. The sum of squares differences between the two pressures compared must be zero for a leak node to be detected. Mathematically, the model is:

$$SS = \text{diag} \left[(Y - J_n \tilde{x})(Y - J_n \tilde{x})' \right] = \text{vector of dimension } n \quad (1)$$

where:

- SS = sum of squares, which is equal to 0,
- $Y = n \times p$ matrix,
- \tilde{x} = vector of simulated pressures,
- $J_n = n \times 1$ vector of ones,
- n = number of model runs, and
- p = number of reference points.

It can be expressed in simple terms by the sum of squared errors (SSE) as:

$$SSE = \sum_{i=1}^n (P_r - P_s)^2 = 0 \quad (2)$$

where:

- SSE = sum of squared errors,
- P_r = reference pressures at the of known leak,
- P_s = simulated nodal pressures, and
- n = number of runs.

To simulate a leak on a selected node, the emitter coefficient, equivalent to a leak flow rate, is changed. The model is then run to generate pressures on all other nodes in the network (see Equations 3 and 4). Emitter coefficients of 0.13 L/s/m to 0.3 L/s/m, which are equivalent to the 2 m³/h to 6 m³/h inferred from Mashford et al. (2009) and Trifunović et al. (2009), are selected. These are derived from the equations for network pressure (Mashford et al. 2009; Trifunović et al. 2009; Tabesh et al. 2009):

$$EC = \frac{Q}{P^\alpha} \quad (3)$$

thus

$$Q = EC * P^\alpha \quad (4)$$

where:

- EC = emitter coefficient (L/s/m^α),
- P = nodal pressure (m),
- α = EPANET exponent for pressure, and
- Q = leak flow rate (L/s).

The model detects leaks by minimizing differences between simulated reference pressures from selected reference nodes and several randomly simulated leaks. The simulated reference pressures are obtained from two simulated reference leaks of known magnitude. The several randomly simulated pressures are given by modeled runs of Monte Carlo simulated leaks.

Initially, reference pressures are simulated by assigning two emitter coefficients representing leaks of known magnitude to two nodes within the water network in EPANET Hydraulic Simulation software. The reference pressures (P_r) of selected reference point nodes are noted (Trifunović et al. 2009). Next, several random simulations are run by an automated process that randomly selects the node and randomly assigns a random emitter coefficient (leak) until the specified number of runs is reached. The simulated pressures (P_s) at the reference points are also noted. Last, the model finds or detects the two leaking nodes by comparing reference pressures P_r and P_s .

The probabilities of occurrence of each pair of node-emitter-coefficient combinations that involve either or both reference nodes that have minimal differences are evaluated.

3 Application to Case Study Network

The Baifikrom Water Network is one of eight separate distribution systems of the Ghana Water Company in the Central Region of Ghana. It serves a population of ~122 000 (estimate based on 2000 census) in 32 communities. The water treatment plant located at Baifikrom, near Mankessim, covers an area of ~250 km² (AVRL 2008; Ghana Statistical Service 2001). The design per capita consumption rates range from 30 L/d to 75 L/d and the total demand is 318 m³/h. From Baifikrom, water is pumped into a reservoir at Mankessim and then flows into three reservoirs located at the highest points within the network for distribution into the communities by gravity (AVRL 2008).

Pipes are made of asbestos cement, high density polyethylene and polyvinyl chloride, making the pipe length of the network ~150 km. The pipe sizes range from 75 mm to 500 mm (AVRL 2008). About 60% of the transmission and distribution mains in the Baifikrom water network lie along the N1 highway (Elubo–Accra–Aflao Road) and have been laid at depths of up to 1.5 m depending on the size of the pipe. The pipeline crosses the highway at 6 points (Mankessim, Abonku, Anokye, Saltpond, Abandze and Anomabo; see Figure 1). At these points, the pipe is laid 2 m deep and sleeved in a larger asbestos cement pipe.

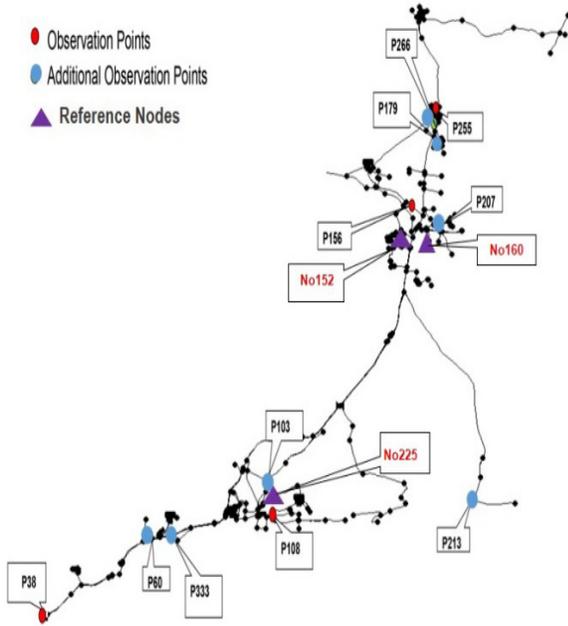


Figure 1 Map showing reference points (Amoatey 2015).

The range of leak sizes simulated was 0.1 L/s/m to 0.3 L/s/m, equivalent to 2.0 m³/h to 6.0 m³/h. Two reference nodes were selected as reference leaking nodes for the assignment of reference pressures. The capability of the model to detect two leaks at a time was investigated using two scenarios: leaks occurring in the *Same Vicinity*, for which nodes No160 and No152 were selected; and leaks occurring *Far Apart* for which No152 and No225 were selected (see Figure 1). To generate reference pressures for the leaks in the *Same Vicinity* scenario, leak coefficients 0.2637 and 0.2647 were assigned to No160 and No152. Similarly, leak coefficients 0.2052 and 0.2417 were assigned to No152 and No225 for the *Far Apart* scenario. The reference pressures for selected reference nodes were recorded.

For the minimum number of reference nodes required for a leaking node to be identified, the analysis began with 4 reference points and was later increased to 6, 8 and 10 points. P38, P108, P156 and P255 were first selected, followed by P103 and P333 (6 reference points), P207 and P266 (8) and P179 and P213 (10). The additional reference nodes selected (making 6,

8, 10 and 12) from different parts of the network are shown in Figure 1.

4 Results: Model Performance Evaluation

The solution parameter set was determined by minimizing an objective function between the reference and simulated pressures. Thus, SSE should be zero, implying that the model run yields pressures similar to the reference leak scenarios under consideration. The performance of the model is presented as probabilities of model runs. Subsequently, five classes were identified and labelled. These represented nodes which paired with either of the two reference nodes as the first or second of the paired nodes.

Thus model runs with SSE = 0 involving the reference nodes No160–No152 and No152–No160 which were designated as the first of the five categories. This was followed by model runs involving nodes that combined with No160 as the first of the paired nodes (No160–NoXA). The third category is model runs involving nodes which combined with No152 as the first of the paired nodes (No152_NoXB) then those that combined with No160 as the second of the paired leak as the fourth category (NoXD–No160). Finally, the fifth category consists of runs in which nodes combined with No152 as the second of the paired nodes (NoXC–No152). The probabilities of each of these 5 categories were computed.

In the same way, in analyzing the performance of the model for leaks occurring in *Far Apart* five categories were designated with No152–No225 and No225–No152 as the first category, No152–NoXE as the second, No225–NoXF as the third, NoXG–No152 as the fourth and NoXH–No225 as the fifth category.

For each scenario, two sets of simulation runs, of 50 000 iterations and 100 000 iterations, were performed. The modeling results and their implications are given. The discussion is organized based on the two scenarios, *Same Vicinity* and *Far Apart*, each for both a *Moderate* (50 000 simulation runs) and a *High* (100 000 simulation runs) number of model runs. For the *Same Vicinity* scenario, nodes No152 and No160 were studied while nodes No152 and No225 were studied in the *Far Apart* scenario. For each scenario, the likely leaking nodes identified were analyzed to see how many of them were paired with either of the control nodes in the respective scenarios.

4.1 Leaks in Same Vicinity: Moderate

The model performance is estimated as the probability of the nodes occurring in model runs which gave an SSE of zero. Nodes which are directly connected to either of the reference nodes in addition to the pairing of the reference nodes were further evaluated. The summary of the resulting objective functions is given in Table 1.

Table 1 Reference nodes combined with possible leak nodes (modified from Amoatey 2015).

Reference Leaking Node	Reference Points	Likely Leaking Nodes	Out of 50 000 Runs (%)	Nodes Combined with Reference Node	Probability of Connected Nodes
No160 and No152 (0.2637 and 0.2647)	4	457	0.914	38	0.0832
	6	452	0.904	38	0.0840
	8	451	0.902	38	0.0843
	10	123	0.246	17	0.1382

About 450 runs gave SSE of zero for 4, 6 and 8 reference points. In all, 47 different nodes including the reference nodes were identified in the resulting model runs. These node pairs with their assigned emitter coefficients generate nodal pressures within the network that are similar to those generated by the reference nodes (No152 and No160) and their assigned emitter coefficients (0.2637 and 0.2647). However, 38 (out of the ~450) runs were paired with one of the two reference nodes including the reference nodes themselves. When the reference points were increased to 10, the number of runs with SSE equal to zero was reduced to 123 involving 30 nodes. Of these, 17 were paired with either No160 or No152. The performance of the model in detecting leaks in Same Vicinity is relatively poor as the probability for nodes connected with either of the reference leaking nodes was just 13% with 10 reference points and much lower (8%) for 4, 6 and 8 reference points (see Table 1).

The relatively high number of runs that resulted in SSE = 0 is possible because both reference nodes No160 and No152 are individually connected to other nodes in the network. There is some form of ripple effect on the surrounding nodes due to the presence of the leak. The other nodes detected along with No160 and No152 for 10 reference points are shown in Figure 2.

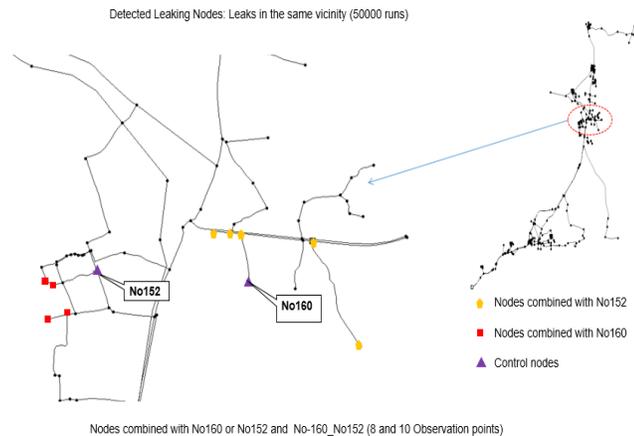


Figure 2 Nodes combined with No160 or No152 (Moderate); Amoatey (2015)

A close observation of Figure 2 reveals that nodes which paired with No152 are situated in the vicinity of No160 while nodes combined with No160 are located in the vicinity of No152. This behavior could be due to an effect of the random nature of the simulation process although this needs to be further investigated. Additionally, all nodes are ~400 m to 800 m away from

reference leaking nodes (see Figure 2). The probabilities of the occurrence of either of the reference nodes according to the five classes are shown in Table 2.

Table 2 Summary of runs; two leaks in Same Vicinity (modified from Amoatey 2015).

Reference points	Occurrence of Reference Nodes in Possible Leaking Nodes ¹	Category	Counts	Probabilities
4, 6 and 8 ¹	No160–No152 and No152–No160	1	3	0.079
	No160–NoXA	2	3	0.079
	No152–NoXB	3	12	0.316
	NoXC–No160	4	11	0.289
	NoXD–No152	5	9	0.237
	Total			38
10	No160–No152 and No152–No160	1	4	0.333
	No160–NoXA*	2	0	0.000
	No152–NoXB*	3	7	0.583
	NoXC*–No160	4	3	0.250
	NoXD*–No152	5	2	0.167
	Total			16

¹NoXA* represents all nodes which combine with reference node No160 as second leaking node; NoXB* represents all nodes which combine with reference node No152 as second leaking node; NoXC* represents all nodes which combine with reference node No160 as first leaking node; and NoXD* represents all nodes which combine with reference node No152 as first leaking node.

The performance of the model, given as probabilities in the tables, shows that with 4, 6 and 8 reference points No160–No152 and No152–No160 have 8% of the count of zero SSEs. The count of nodes paired with No160 in the category No160–NoXA was about 8% while those in the third category (No152–NoXB) performed better (32%). NoXC–No160 node pairs performed better than No160–NoXA (29%) while the fifth category (NoXD–No152) was about 24%. Increasing the reference points to 10 led to an improvement in the performance of the model. Counts of No160–No152 and No152–No160 of the resulting zero SSEs were ~33%. Node pairs in No152–NoXB and NoXD–No152 were 58% and 17%. There was no count for the node pairs No160–NoXA but their counterpart pair (NoXC–No160) showed only 25%.

Nodes which were paired with the reference node No152 are situated in the neighborhood of No160. This is most likely due to the random selection of nodes and random assignment of emitter coefficients in the leak simulation process. Reference node No160 paired only with No152 and can therefore be easily detected.

4.2 Leaks in Same Vicinity: High

As mentioned earlier, in order to further understand the performance of the model in detecting leaks within the range considered, the number of model runs was increased. Again, several model runs showed similar pressures within the network to No160 and No152 as determined by their assigned emitter coefficients. With 4 and 6 reference points, ~900 runs gave an SSE of zero involving 38 nodes (see Table 3). Among these nodes, 72 were paired with either No152 and No160. With 8 and 10 reference points, the number of runs with SSE = 0 reduced to 229 involving 29 nodes. In this case 36 of them were paired with one of the reference

nodes. Again, the model performed poorly yielding about 15% with 8 and 10 reference points. The results are summarized in Table 3. Other nodes which combined with either No160 or No152 with 8 and 10 reference points are shown in Figure 3. The counts and probabilities of the occurrence of No160 and No152 in the five categories for the leaks in the Same Vicinity scenario are presented in Table 4.

Table 3 Summary of results for 100 000 runs, leaks in Same Vicinity (modified from Amoatey 2015).

Reference Leaking Node	Reference Points	Likely Leaking Nodes	Out of 100 000 runs (%)	Nodes Combined with Reference node	Probability of Connected Nodes
No160 and No152 (0.2637 and 0.2647)	4	897	0.90	73	0.0814
	6	886	0.89	73	0.0824
	8	229	0.23	36	0.1572
	10	229	0.23	36	0.1572

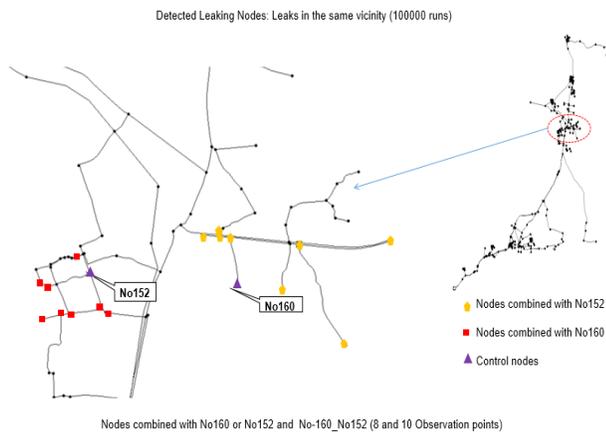


Figure 3 Nodes combined with No160 or No152 (Same Vicinity, High); Amoatey (2015).

Table 4 100 000 runs, Same Vicinity (modified from Amoatey 2015).

Reference Points	Occurrence of Reference Nodes in Possible Leaking Nodes ¹	Category	Counts	Probabilities
4 and 6	No160–No152 and No152–No160	1	4	0.056
	No160–NoXA*	2	13	0.181
	No152–NoXB*	3	21	0.292
	NoXC*–No160	4	20	0.278
	NoXD*–No152	5	15	0.194
	Total			73
8 and 10	No160–No152 and No152–No160	1	4	0.111
	No160–NoXA*	2	7	0.194
	No152–NoXB*	3	11	0.306
	NoXC*–No160	4	10	0.278
	NoXD*–No152	5	4	0.111
	Total			36

¹NoXA* represents all nodes which combine with reference node No160 as second leaking node; NoXB* represents all nodes which combine with reference node No152 as second leaking node; NoXC* represents all nodes which combine with reference node No160 as first leaking node; and NoXD* represents all nodes which combine with reference node No152 as first leaking node.

The results show that with 4 and 6 reference points, the No160–No152 and No152–No160 counts were 6% of the zero SSEs. Nodes paired with No160, No160–NoXA were ~18% while those in the category No152–NoXB performed much better with 29%. NoXC–No160 pairs performed better than No160–NoXA with ~28% while NoXD–No152 was ~19%.

There was a slight improvement with 8 and 10 reference points. The performance of the model for two leaks in the Same Vicinity scenario using this inverse modeling approach is poor. It attained only 15% with 8 and 10 reference points. When the resulting node pairs were further analyzed, the counts for No160–No152 and No152–No160 of the zero SSEs increased to 11%. Node pairs in categories 2 through 5 were 19%, 31%, 28 and 11%. Again, nodes which were paired with the reference node No152 are situated in the environs of No160 and vice versa. Also again, this could be attributed to the random selection of nodes and random assignment of emitter coefficients in the leak simulation process and will be further studied. The nodes identified along with the reference nodes are shown in Figure 4.

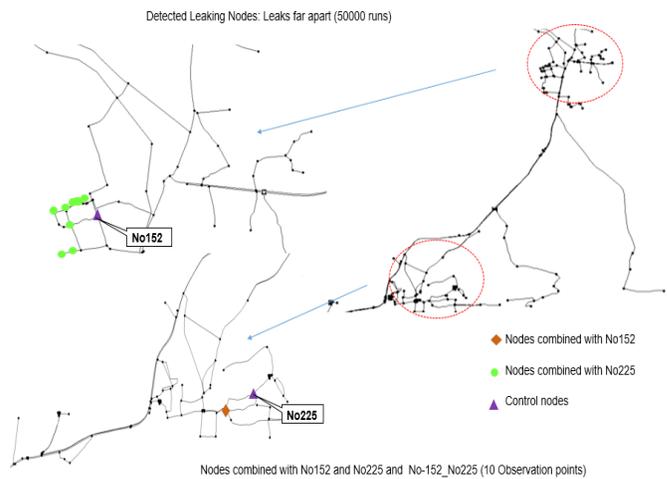


Figure 4 Nodes combined with No152 and No225 (Moderate); Amoatey (2015).

4.3 Leaks in Far Apart Scenario: Moderate

Similarly, for leaks occurring at a distance from each other, the Far Apart scenario, the generated pressures were compared using 4, 6, 8 and 10 reference points as in the previous scenario. With No152 and No225 as reference (leaking) nodes, the resulting likely leaking nodes are summarized in Table 5. With 4, 6 and 8 reference points, 41 model runs involving 18 different nodes were obtained (Table 5). Of the 41 model runs, 27 of them were paired with one of the reference nodes. When the reference points were increased to 10, the model runs that gave SSE = 0 reduced to 27, involving fifteen nodes. Seventeen nodes were paired with the reference nodes (Table 5). These likely leaking nodes are shown in Figure 5.

Table 5 Results for Moderate; leaks occurring in Far Apart (modified from Amoatey 2015).

Reference Leaking Node	Reference Points	Likely Leaking Nodes	Out of 50 000 Runs (%)	Nodes Combined with Reference Nodes	Probability of Connected Nodes
No152 and No225 (0.2052 and 0.2417)	4	41	0.082	27	0.634
	6	41	0.082	27	0.634
	8	41	0.082	26	0.634
	10	27	0.054	17	0.630

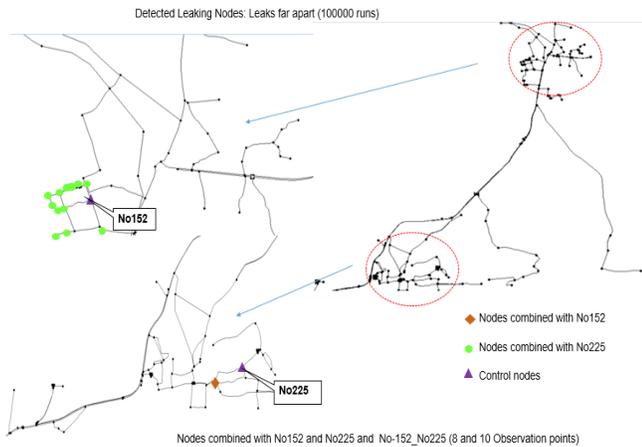


Figure 5 Nodes combined with No152 or No225 (High); Amoatey (2015)

The model performs above average (63%) in the Far Apart scenario, better than in Same Vicinity. For the leaks occurring in Far Apart, as shown in Table 5, out of the resulting 27 model runs, only 1 node (No91) paired with No152 and it is close to No225. No91 is directly connected to No225 (see Figure 5). All other model runs that involved No152 paired it with No225 or No225 was paired with other nodes. The reason for this is that No152 is connected to several nodes while No225 is connected to relatively few nodes as it occurs in a branch of the network. Thus, few nodes close to the reference node No225 are paired with No152, making it easy to identify the leaking node.

In modeling the detection of two leaks in a network, the identified likely leaking nodes are paired with the reference nodes in an interesting way. If other nodes take the first position in the pair with reference nodes, the likely leaking nodes are located in the vicinity of the second reference node and vice versa (Amoatey 2015).

The performance of the model, represented as counts and probabilities in respect of the five categories, is shown in Table 6. Since No152 did not pair with other nodes, the model runs yield higher probabilities involving No225 (that is far away from No152), unlike in Same Vicinity (see Figure 4). The No152–NoXE categories yielded a performance of 29% while

the No225–NoXF combinations yielded 41%. The categories involving No152–NoXG and NoXH–No225 yielded respectively 0% and 6%.

Table 6 Possible leak nodes combined with reference nodes (modified from Amoatey 2015).

Reference Points	Occurrence of Reference Nodes in Possible Leaking Nodes ¹	Category	Counts	Probabilities
4, 6 and 8	No152–No225 and No225–No152	1	4	0.154
	No152–NoXE*	2	0	0.000
	No225–NoXF*	3	8	0.308
	NoXG*–No152	4	1	0.038
	NoXH*–No225	5	13	0.500
	Total			26
10	No152–No225 and No225–No152	1	4	0.235
	No152–NoXE*	2	0	0.000
	No225–NoXF*	3	5	0.294
	NoXG*–No152	4	1	0.059
	NoXH*–No225	5	7	0.412
	Total			17

¹NoXE* represents all nodes which combine with reference node No160 as second leaking node; NoXF* represents all nodes which combine with reference node No152 as second leaking node; NoXG* represents all nodes which combine with reference node No160 as first leaking node; and NoXH* represents all nodes which combine with reference node No152 as first leaking node.

It can be inferred that, when two leaks occur in a network, as the leaks become further apart very few nodes show similar pressures to the reference leaking nodes. This is consistent with theory since the influence of each individual leak is not experienced by the other. This again implies that leaks occurring far apart could be easily found by this model compared to leaks occurring in the Same Vicinity. Nevertheless, if the expanse of the area is not large, as in the case of No152–No160, then locating the two leaks will also not be so tedious. Previous studies using optimization approaches (Sousa et al. 2015; Ishido and Takahashi 2014; Jasper et al. 2013 and Mashford et al. 2009) did not perform simulations involving two leaks and therefore the observed dynamics cannot be fully explained without further studies. Several instances of the Far Apart scenario will establish this.

4.4 Leaks in Far Apart Scenario: High

For leaks occurring Far Apart when the model runs were increased to 100 000, fewer node pairs with their assigned emitter coefficients showed similar pressures as the No152 and No225 compared to leaks in the Same Vicinity. The number of model runs with SSE equal to zero reduced from 75 involving 22 nodes for 4 and 6 reference points to 46 involving 20 nodes with 8 and 10 reference points. The model runs involving nodes that combined with either No152 or No225 are listed in Table 7. A summary of the model performance in the high model runs category is given in Table 7.

Table 7 Results for 100 000 runs, Far Apart (modified from Amoatey 2015).

Reference Leaking Node	Reference Points	Likely Leaking Nodes	Out of 100 000 Runs (%)	Nodes Combined with Reference Node	Probability of Connected Nodes
No152 and No225 (0.2173 and 0.2749)	4	75	0.075	44	0.587
	6	75	0.075	44	0.587
	8	46	0.046	27	0.413
	10	46	0.046	27	0.413

The model performance in detecting leaking nodes is just above average (58%) for 4 and 6 reference points but below average (41%) for 8 and 10 reference points. It is noteworthy that the model performance decreases with an increased number of runs (50 000 to 100 000 runs). The optimal point of reduced performance needs to be further investigated and established. Looking again at Table 7 above, it is easier to identify two leaks occurring in Far Apart within the network because relatively fewer model runs yield an SSE of zero in comparison with the leaks in Same Vicinity.

The reason for this is, as discussed earlier, due to the interconnectivity of No160 and No152 compared to No152 and No225. The resulting leaking nodes are located 400 m to 800 m away from the modeled reference leaking nodes. The counts and probabilities of the occurrence of No152 and No225 in the five categories are shown in Table 8. With 10 reference points, the counts for No152–No225 as well as No225–No152 were 15%. The No152–NoXE categories were 0% while the No225–NoXF combinations yielded 44%. The categories involving NoXG–No152 and NoXH–No225 were respectively 7% and 33%.

Table 8 Occurrence of reference nodes for leaks in Far Apart (modified from Amoatey 2015).

Reference Points	Occurrence of Reference Nodes in Possible Leaking Nodes ¹		Counts	Probabilities
	Category			
4, 6 and 8	No152–No225 and No225–No152	1	4	0.091
	No152–NoXE*	2	0	0.000
	No225–NoXF*	3	18	0.409
	NoXG*–No152	4	2	0.045
	NoXH*–No225	5	20	0.455
	Total		44	1.000
10	No152–No225 and No225–No152	1	4	0.148
	No152–NoXE*	2	0	0.000
	No225–NoXF*	3	12	0.444
	NoXG*–No152	4	2	0.074
	NoXH*–No225	5	9	0.333
	Total		27	1.000

¹NoXE* represents all nodes which combine with reference node No160 as second leaking node; NoXF* represents all nodes which combine with reference node No152 as second leaking node; NoXG* represents all nodes which combine with reference node No160 as first leaking node; and NoXH* represents all nodes which combine with reference node No152 as first leaking node.

Similar observations can be made with respect to the probabilities associated with the occurrence of the reference nodes as in the example of 50 000 runs. It is easier to identify two leaks occurring in Far Apart within the network because fewer

node–emitter coefficient combinations give similar pressures to the reference leaking nodes in comparison with two leaks in Same Vicinity. This is, as discussed earlier, because of the interconnectivity of No160 and No152 compared to No152 and No225. Also, due to the distance between the two nodes, the pressure differences caused by the leak has no influence on the other reference leaking node.

Table 9 gives a summary of the performance of the model using 10 reference points for the two scenarios with each set of simulation runs and shows that, for the range of leak sizes considered in this study, leaks occurring far apart are easier to detect than those occurring in the same vicinity. This has implications for further research and practice as both scenarios are likely occurrences in water distribution networks

Table 9 Summary of model performance for two leaks (modified from Amoatey 2015).

Scenarios	No. of runs	Performance	Remark on Performance
Same Vicinity	50 000	13%	Poor
	100 000	15%	Poor
Far Apart	50 000	63%	Above Average
	100 000	41%	Fair (Lower number of runs)

5 Conclusion

The optimization-based approach simulates pressures within the hydraulic model of the water distribution network. In the two scenarios considered, the detection of two simultaneous leaks in the Same Vicinity scenario was poor while the performance was just above average for the Far Apart scenario. The choice of reference nodes needs to be better optimized to improve results for Same Vicinity. A minimum of 10 reference points is required to detect leaks efficiently in the case study water distribution network. The performance of the model requires further study to improve it. The stochastic nature of the pairing of the nodes in the leak simulation process shows some characteristics with respect to the physical location of the other nodes detected which will be further investigated.

It must be emphasized that regular calibration of the water distribution network is mandatory for the model to better predict leaking nodes. Despite the average performance of the optimization based model, it is relatively easy to implement, computationally undemanding and fast. However, the model needs to be improved to make it applicable to water distribution networks without sensors.

For leaks in the range of the emitter coefficients we considered, the model performs better for the leaks occurring in Far Apart than for leaks in Same Vicinity. This will be further studied to better understand the model. The capability of the model to detect relatively smaller leaks than the sizes considered in this paper has been explored and findings will be published soon. The required degree of accuracy in calibrating the model for the

model to work well will be studied later. The performance of the model when a tolerance is set ($SSE < 0.05$) will also be investigated.

Acknowledgments

The authors wish to acknowledge the UG–Carnegie Next Generation of African Academics (NGAA) program for sponsoring an off-campus writing workshop that resulted in this paper. I am thankful that Dr. Iddi of the University of Ghana assisted in the mathematical notation of the model. I am highly indebted to Professors Steinmetz and Bárdossy for their guidance during my PhD studies. I am grateful to the University of Ghana and GETFUND for their financial support and to Ghana Water Company Limited for support during the data collection period.

References

- Amoatey, P. K. 2015. *Leakage Management in the Urban Water System in Ghana: Estimation and Detection Modelling*. Stuttgart: University of Stuttgart. PhD dissertation. ISBN 978-3-8356-7291-8.
[https://stg.ibs-bw.de/aDISWeb/app?service=direct/0/Home/\\$DirectLink&sp=S127.0.0.1:23022&sp=SAKSWB-Id-Nr470460091](https://stg.ibs-bw.de/aDISWeb/app?service=direct/0/Home/$DirectLink&sp=S127.0.0.1:23022&sp=SAKSWB-Id-Nr470460091)
- Amoatey, P. K., A. Bárdossy and H. Steinmetz. 2018. "Inverse Optimization based Detection of Leaks from Simulated Pressure in Water Networks, Part 1: Analysis for a Single Leak." *Journal of Water Management Modeling* 26: C460. doi: 10.14796/JWMM.C460.
- AVRL (Aqua Vitens Rand Limited). 2008. *Network Design Report*. AVRL Report 1–44.
- Boulos, P., T. Schade and C. W. Baxter. 2008. "Locating Leaks in Water Distribution Systems Using Network Modeling." *Journal of Water Management Modeling* 16: R228–21. doi: 10.14796/JWMM.R228-21.
- Brunone, B. and M. Ferrante. 2001. "Detecting Leaks in Pressurized Pipes by Means of Transients." *Journal of Hydraulic Research* 39 (5): 539–47.
- Colombo, A. F., P. Lee and B. W. Karney. 2009. "A Selective Literature Review of Transient-Based Leak Detection Methods." *Journal of Hydro-environment Research* 2:212–27.
- Covas, D., I. Stoianov, D. Butler, C. Maksimovic, N. Graham and H. Ramos. 2001. "Leakage Detection in Pipeline Systems by Inverse Transient Analysis: From Theory to Practice." In *Proceedings, Sixth International Conference on Computing and Control in the Water Industry (CCWI 2001), Leicester, England*, vol 2, 3–16. Baldock, UK: Research Studies.
- De Silva, D., J. Mashford and S. Burn. 2011. *Computer Aided Leak Location and Sizing in Pipe Networks: Urban Water Security Research Alliance Technical Report No. 17*, 2nd ed. Canberra, ACT: (CSIRO) Commonwealth Scientific and Industrial Research Organisation.
<http://www.urbanwateralliance.org.au/publications/UWS-RA-tr17.pdf>
- Gertler, J., J. Romera, V. Puig and J. Quevedo. 2010. "Leak Detection and Isolation in Water Distribution Networks using Principal Component Analysis and Structured Residuals." In *2010 Conference on Control and Fault Tolerant Systems*, 191–196. New York: IEEE.
<https://doi.org/10.1109/SYSTOL.2010.5676043>
- Ghana Statistical Services. 2001. *Population Census 2000*. Accra: Government of Ghana.
- Hsu, C. C., M. C. Che. and L. S. Chen. 2010. "Integrating Independent Component Analysis and Support Vector Machine for Multivariate Process Monitoring." *Computers and Industrial Engineering* 59 (1): 145–56.
- Hunaidi, O. 2012. "Detecting Leaks in Water-Distribution Pipes." *Construction Technology Update* 79:1–5.
- Ishido, Y. and S. Takahashi. 2014. "A New Indicator for Real-Time Leak Detection in Water Distribution Networks: Design and Simulation Validation." *Procedia Engineering* 89:411–7.
- Jasper, M. N., G. Mahinthakumar, S. Ranjithan and E. D. Brill. 2013. "A Sensitivity Analysis of Data Measurement Types for Leak Detection in Water Distribution Systems." In *World Environmental and Water Resources Congress, 2013, Cincinnati, Ohio*, 611–6. Reston, VA: American Society of Civil Engineers.
<https://doi.org/10.1061/9780784412947.059>
- Kapelan, Z. S., D. A. Savic and G. A. Walters. 2003. "A Hybrid Inverse Transient Model for Leakage Detection and Roughness Calibration in Pipe Networks." *Journal of Hydraulic Research* 41 (5): 481–92.
- Karney, B., D. Khani, M. R. Halfawy and O. Hunaidi. 2009. "A Simulation Study on Using Inverse Transient Analysis for Leak Detection in Water Distribution Networks." *Journal of Water Management Modeling* 17: R235–23. doi: 10.14796/JWMM.R235-23.
- Liggett, J. A. and L. C. Chen. 1994. "Inverse Transient Analysis in Pipe Networks." *Journal of Hydraulic Engineering* 120 (8): 934–55.
- Mashford, J., D. De Silva, D. Marney and S. Burn. 2009. "An Approach to Leak Detection in Pipe Networks using Analysis of Monitored Pressure Values by Support Vector Machine." In *Proceedings of the 2009 Third International Conference on Network and System Security*, 534–9. New York: IEEE Explore.
<https://doi.org/10.1109/NSS.2009.38>
- Mirats-Tur, J. M., P. A. Jarrige, J. Meseguer and G. Cembrano. 2014. "Leak Detection and Localization using Models: Field Results." *Procedia Engineering* 70:1157–65.

- Nowicki, A. and M. Grochowski. 2011. "Kernel PCA in Application to Leakage Detection in Drinking Water Distribution System." In *Proceedings of the Third International Conference on Computational Collective Intelligence: Technologies and Applications*, 497-506. Berlin: Springer.
- Nowicki, A., M. Grochowski and K. Duzinkiewicz. 2012. "Data-Driven Models for Fault Detection using Kernel PCA: A Water Distribution System Case Study." *International Journal of Applied Mathematics and Computer Science* 22 (4): 939–49.
- Palau, C. V., F. J. Arregui and M. Carlos. 2012. "Burst Detection in Water Networks Using Principal Component Analysis." *Water Resources Planning and Management* 138 (1): 47–54.
- Palau, C. V., F. J. Arregui and A. Ferrer. 2004. "Using Multivariate Principal Component Analysis of Injected Water Flows to Detect Anomalous Behaviors in a Water Supply system—A Case Study." *Water Science and Technology: Water Supply* 4 (3): 169–81.
<https://doi.org/10.2166/ws.2004.0052>
- Poulakis, Z., D. Valougeorgis and C. Papadimitriou. 2003. "Leakage Detection in Water Pipe Networks using a Bayesian Probabilistic Framework." *Probabilistic Engineering Mechanics* 18:315–27.
- Shucksmith, J. 2011. "Leak Detection using Pressure Transients." Presentation to the Water Industry Forum, Pennine Water Group, University of Sheffield.
- Sousa, J., L. Ribeiro, J. Muranho and A. Sá Marques. 2015. "Locating Leaks in Water Distribution Networks with Simulated Annealing and Graph Theory." *Procedia Engineering* 119:63–71.
- Tabesh, M., A. H. Asadiani Yekta and R. Burrows. 2009. "An Integrated Model to Evaluate Losses in Water Distribution Systems." *Water Resources Management* 23:477–92.
- Trifunović, N., S. Sharma and A. Pathirana. 2009. "Modelling Leakage in Distribution System Using EPANET." Paper presented at the IWA Water Loss Conference, 2009, Cape Town, South Africa.
- Vanik, M. W., J. L. Beck and S. K. Au. 2000. "Bayesian Probabilistic Approach to Structural Health Monitoring." *Journal of Engineering Mechanics* 126 (7): 738–45.
- Vitkovsky, J. P., A. R. Simpson, M. F. Lambert and J. A. Liggett. 2007. "Experimental Observation and Analysis of Inverse Transients for Pipeline Leak Detection." *Water Resources Planning and Management* 133 (6): 519–30.
- Vitkovsky, J. P., A. R. Simpson, M. F. Lambert and X. J. Wang. 2001. "An Experimental Verification of the Inverse Transient Technique for Leak Detection." In *Proceedings of the Sixth Conference on Hydraulics in Civil Engineering: The State of Hydraulics, Hobart, Tasmania*, 373–80. Barton, ACT: Institution of Engineers, Australia.
- Wang, X. J., M. F. Lambert, A. R. Simpson, J. A. Liggett and J. P. Vitkovsky. 2002. "Leak Detection in Pipelines using the Damping of Fluid Transients." *Journal of Hydraulic Engineering* 128 (7): 697–711.
- Wu, Z. Y., M. Farley, D. Turtle, Z. Kapelan, J. Boxall, S. Mounce, S. Dahasahasra, M. Mulay and Y. Kleiner. 2011. *Water Loss Reduction*. Exton, PA: Bentley Institute Press.
- Wu, Z. Y. and P. Sage. 2006. "Water Loss Detection via Genetic Algorithm Optimization-Based Model Calibration." In *Eighth Annual International Symposium on Water Distribution Systems Analysis. Cincinnati, Ohio, August 27-30, 2006*. Reston, VA: American Society of Civil Engineers.
[https://doi.org/10.1061/40941\(247\)180](https://doi.org/10.1061/40941(247)180)