WATER LOSS DETECTION VIA GENETIC ALGORITHM OPTIMIZATION-BASED MODEL CALIBRATION

Zheng Yi Wu
Haestad Methods Solution Center, Bentley Systems, Incorporated
27 Siemon Co Dr, Suite 200W,
Watertown, CT06795, USA
Email: zheng.wu@bentley.com

Paul Sage
United Utilities PLC, Dawson House, Great Sankey,
Warrington, WA5 3LW,
United Kingdom
Email: paul.sage@uuplc.co.uk

Abstract

Identifying how much water is being lost from water networks and where the losses are occurring is of great importance to water utilities both for operational and planning reasons as well as for reputation. In this paper, an optimization-based approach is presented for simultaneously quantifying and locating water losses via the process of hydraulic model calibration. The model calibration is formulated as a nonlinear optimization problem that is solved by using a genetic algorithm. The method is developed as an integrated framework of hydraulic simulation and optimization modeling. Case studies are presented to demonstrate how the integrated approach is applied to water loss detection. The results obtained show that the method is effective at detecting water loss as part of the hydraulic calibration of the network model. The accuracy of water loss detection is dependent on the quality of the field observed data and model granularity. However, it has been shown that the approach can be used for reducing the uncertainty of the water loss identification by locating water loss hotspots, which could lead to improved operating revenues at water utilities.

Key Words
Water distribution system, loss, leakage, hydraulic model, parameter estimation, model calibration, genetic algorithm, optimization.

INTRODUCTION

Water loss represents a major portion of non-revenue water (NRW) for a water system. According to the International Water Association (IWA)'s best practice recommendations for water balance studies more than 65% of NRW arises from unauthorized water consumption, meter inaccuracies and leakages from the water mains source-to-taps infrastructure. Conventional water loss studies focus on the calculation of supply balances for water systems or their subsystems such as district meter areas (DMAs). Over the last decade, the concepts and methods developed for system wide water balance calculations (IWA 2000 and AWWA 2003) have been based upon water assets' physical data and the statistics of pipe bursts, service connections and underground conditions. Performance measures and indicators are used to support the managerial approaches to minimize different components of water losses. These concepts and methods have been adopted by countries around the world (Thornton, 2002; Lambert & McKenzie 2002; Lambert & Fantozzi 2005). More recently, Almandoz et al. (2005) and Walski (2006) developed an alternative approach for assessing leakage through network hydraulic simulation. As far as the authors are aware, none of the previous methods have considered how network models can be used to locate the water loss or leakage in a system. Nevertheless, a means by which network models can be used to predict leakages quickly within DMAs is important as it assists a water utility to take the correct proactive measures to accurately minimize the leakage.

Leakage detection apparatus has been developed for detecting potential water loss or leakage from local water mains. However, it is time-consuming and costly for leak detection crew to apply the current detection procedures to every main. Leaks are also becoming harder to find. One reason for this is pressure management. Not only is leakage being routinely suppressed by planned pressure reduction but the same reduction leads to less acoustic disturbance from those leaks. At the same time, sounding for leaks has become more difficult as ferrous mains are replaced with less acoustically responsive plastic pipes replace

In the UK, leakage targets are a key part of water utilities measured performances, with severe financial penalties being incurred if targets are not met. A need has arisen to develop a systematic approach that identifies likely leakage hotspots so that detection crews can identify leaky mains more quickly in turn leading to quicker repairs.

Sage (2005) has applied a genetic algorithm (GA) optimization technique to all-main models and demonstrated some initial success at predicting leakage hotspots throughout a water distribution system. The method is based on water network modeling with leakage modeled as a demand. The optimization process is to determine whether a demand node is to supply a running burst or background leakage. One variable, taking a value between 0 and 1, is assigned to each demand node. A node is deemed to supply a running burst if the variable attributed to the node has a value that is greater than the pre-calculated average Internal Condition factor (ICF), otherwise the node is designated as one with background leakage only. Each GA solution is comprised of one set of values for all the demand nodes. The solution fitness is evaluated by comparing the field observed pressure and the simulated pressure. The leakage detection was undertaken by first analyzing the normalised field test data with head losses and associated area inflows, which were measured with a density of one data logger per 200 houses. The initial field data analysis revealed useful information about locations of likely leakage hotspots. The GA optimization process was performed to identify unknown flows out of a study area and also identified customers with large unrecorded consumptions. Leakage detection runs were generally limited to the low flow conditions at night as the more variable day time demands were found to mask leakage behaviour. Several optimisation runs were required for each study area to test if the likely leakage hotspot locations were predicted in the same localities.

The model supported leak detection approach has been taken forward by UUNW as it is likely to provide economic and regulatory benefits for the Company and might be taken up by other water utilities as awareness of the approach grows. The solution method focuses on identifying demand uncertainty by using the observed flow and pressure field test data, collected for model calibration. The approach is similar to optimization-based model calibration (Wu et al. 2002), which allows users to identify system parameters in an integrated environment. In the following sections, UUNW's leak detection approach has been discussed in accordance with the optimization based model calibration methods.

GENERALIZED MODEL CALIBRATION

A generalized model calibration is formulated for optimizing the parameters for a water distribution system model, including:
1. Pipe roughness factor f_i for pipe group i, all the pipes within one group will be either set to the same roughness coefficient or applied the same multiplier to the original roughness coefficients;
2. Demand adjustment multiplier $m_{j,t}$ for junction group j at time step t, junctions within the same demand group are applied the same demand adjustment multiplier;
3. Link operation status $s_{k,t}$ for link k (pipes, valves and pumps) at time step t.

The fittest parameters are obtained by minimizing the discrepancy between the model predicted and the field observed values of junction pressures (hydraulic grades) and pipe flows for given boundary conditions including tank levels, control valve settings and pump speeds. The optimized calibration is then defined as a nonlinear optimization problem for a given time *t* step as follows.

Search for: $\vec{X} = (f_i, m_{j,t}, s_{k,t})$ $i = 1, ..., NI; j = 1, ..., NJ; k = 1, ..., NK.$ (1)

Minimize: $F(\vec{X})$ (2)

Subject to: $\overline{f_i} \le f_i \le \underline{f_i}$ (3)

$\overline{m_{j,t}} \le m_{j,t} \le \underline{m_{j,t}}$ (4)

$s_{k,t} \in \{0,1\}$ (5)

Where $\vec{X}$ represents a set of model parameters, $\overline{f_i}$ and $\underline{f_i}$ are the upper and lower limits of roughness factor for pipe group *i*, $\overline{m_{j,t}}$ and $\underline{m_{j,t}}$ are the upper and lower limits for demand adjustment multiplier for junction group *j* at time step *t* and $F(\vec{X})$ is the objective function that measuring the goodness-of-fit between the field observed values and the model simulated values. The calibration objective can be defined in three different distance functions of (1) the sum of difference square; (2) the sum of absolute differences and (3) the maximum difference among all the data points. The model parameter identification problem formulated by Eq(1) to Eq(5) is solved by using a genetic algorithm (Wu et al. 2002).

LEAK DETECTION

Leakage or water loss is modeled as demand in addition to the recorded demand that can be calculated and assigned by different loading methods using customer meter information (Wu et al. 2003). Actual demand conditions vary throughout a system and over time. Field observation data are measured and collected at different time of day. Each dataset collected at one time corresponds to a specific demand loading condition. In order for that the model simulation results more closely represent the observed data, the simulated results must be given for the same demand loading and boundary conditions as the observed data is collected.

To apply the calibration method formulated as Eq(1) to Eq(5) to water loss or leak detection, demand for each junction is decomposed into a representative demand and demand adjustment (water loss). Representative demand is defined as a baseline demand (fixed amount such as average day demand, peak hour etc.) and a demand pattern (diurnal variation coefficients over 24 hours), which are determined by using the metered consumption data and flow measurement. For junction *j* at time step *t*, the total demand is given as:

$$Q_j(t) = Q_j^b \times pat(t) \times m_{j,t}$$ (6)

$$Q_j^l(t) = Q_j(t) - Q_j^b(t) \times pat(t)$$ (7)

where $Q_j(t)$ designates the total demand at junction *j* at time step of *t*; Q_j^b denotes the baseline demand at junction *j*; $pat(t)$ represents the pattern coefficient at time step of *t*, $m_{j,t}$ is the demand multiplier for junction group *j* at time *t* and $Q_j^l(t)$ is the demand adjustment of junction *j*, which represents an additional demand namely water loss or leakage. Baseline demands and demand patterns are usually

constructed when building a hydraulic network model. To achieve a good estimate of water loss or leakage through this calibration approach, total system inflow, boundary conditions and the data collection time must be recorded for providing accurate demand calibration.

EXAMPLE I

This case study is undertaken using the same example system as originally designed by UUNW for applying the model for leak management (Sage 2005). The system layout is illustrated in Figure 1. It is a trivial system, consisted of 21 pipes and supplied by one reservoir. The model includes a customer base demand of 1.0 L/s at all the nodes, which is multiplied by a typical domestic demand profile to reproduce a representative diurnal profile. Leakage is considered as another demand but with a much flatter diurnal profile. The leakage detection test is then progressed as follows.

Test Design

It is assumed that the following field data can be obtained for leak detection, including:
- Total inflow into the system;
- The pressures at 5 nodes xx1, xx2, xx3, xx7, xx9 and 101.

The leakage scenarios are designed as:
1. Uniformly distributed leak, 1.0 L/s of leakage at all demand node;
2. Big downstream leak, 5.75 L/s of leakage at downstream node 117 and 0.75 L/s for the rest of demand nodes;
3. Big upstream leak, 5.75 L/s of leakage at upstream node 101 and 0.75 L/s for the rest of demand nodes;
4. Big downstream leak and wrong pipe diameter for pipe 'ad', pipe size is deliberately set to a wrong size of 400 mm.
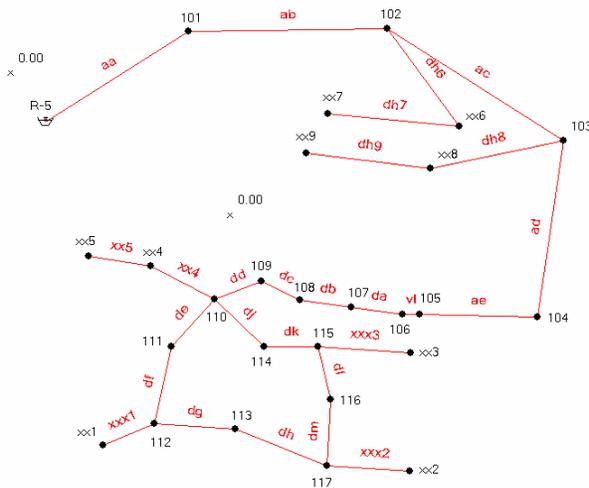
Figure 1 Layout of example system for leak detection study

Each of four scenarios is built into the model as an additional set of demand on top of the base demand, and simulation results are extracted to represent model calibration field test data that will subsequently be used in the GA for leak detection. The calibration setup is based upon the base demand, namely just the customer metered consumption of 1.0 L/s at nodes. Optimization runs are performed to identify the

correct demand multipliers that result in the total nodal demand given as Eq(6), consequently, the leak is identified by Eq(7).

The simulated results above are based on known base demands and base demand profiles. The next step is to investigate the performance of the GA-based calibration tool (Wu et al. 2004) applied to the deliberately designed leakage scenarios by using WaterGEMS (Bentley 2006).

Results

Due to the stochastic mechanism of GA optimization and the leakage uncertainty, the solution produced by one single run may not represent statistic significance for leakage detection. Ten optimization runs with different random seeds were therefore performed and the expected mean of leakage demand found for each leakage scenario. The leak detection optimization for all four scenarios was undertaken at the 5:00 AM demand condition, a low flow time for the system. The results of the ten GA optimization runs with different random seeds ranging from 0.1 to 1.0 with increment of 0.1,. are given below.

1. Uniformly distributed leak

For this testing scenario, the node (metered customer) demand is 0.5 L/s at 5:00 AM, while the leakage is assumed 1.0 L/s at every node, the total demand is 1.5 L/s at each node, this results in a demand multiplier of 3 (1.5/0.5) corresponding to the ratio of the calibrated demand to the original demand. The GA-based calibration runs are initiated to identify the correct demand factor for the demand nodes such that the differences between the observed and simulated values are minimized.

Table 1 Optimization results for detection of uniformly distributed leak

Node ID	Optimized demand factors of GA-based calibration runs										Average Factor	Correct factors
	1	2	3	4	5	6	7	8	9	10		
101	3.25	2.05	4.50	1.75	3.70	0.95	1.95	3.05	3.05	2.90	2.72	3.00
102	1.20	4.05	0.75	4.55	2.00	5.00	3.70	2.05	5.00	4.95	3.33	3.00
103	3.15	2.15	3.15	4.00	2.35	2.40	2.30	3.35	1.65	0.95	2.55	3.00
104	5.00	5.00	2.40	0.90	3.65	2.80	2.20	0.70	3.20	5.00	3.09	3.00
105	3.50	5.00	1.15	0.55	4.15	2.90	5.00	5.00	1.30	3.65	3.22	3.00
106	4.35	3.45	4.05	2.35	3.20	0.95	2.45	3.55	1.65	5.00	3.10	3.00
107	2.35	1.25	5.00	1.10	2.40	0.60	1.90	2.00	5.00	3.40	2.50	3.00
108	1.80	5.00	3.95	1.85	2.90	1.45	0.55	5.00	0.55	1.85	2.49	3.00
109	5.00	2.00	2.50	2.60	4.75	2.15	5.00	3.00	3.30	4.25	3.46	3.00
110	1.40	5.00	3.30	0.60	0.60	2.30	3.25	2.35	3.90	1.50	2.42	3.00
111	2.00	0.75	5.00	1.15	5.00	5.00	3.75	1.35	1.85	1.90	2.78	3.00
112	4.15	2.05	1.50	5.00	2.45	3.45	3.70	3.80	5.00	4.90	3.60	3.00
113	5.00	5.00	3.10	1.35	2.30	2.65	3.10	4.60	0.75	0.50	2.84	3.00
114	4.15	2.10	2.95	5.00	0.50	3.05	0.85	2.05	2.10	3.40	2.62	3.00
115	2.10	2.10	2.15	5.00	4.25	5.00	5.00	3.90	4.70	1.85	3.61	3.00
116	0.90	2.50	3.65	5.00	5.00	5.00	1.45	1.25	2.90	3.25	3.09	3.00
117	1.50	2.40	2.00	5.00	1.45	1.95	3.20	3.05	5.00	3.45	2.90	3.00
xx7	4.85	1.95	5.00	3.70	5.00	5.00	4.65	4.55	2.45	1.40	3.86	3.00

As shown in Table 1, although the optimized demand factors by one individual optimization run do not seem to provide much indication as to the leakage level and its location, the average demand factors over

ten runs do illustrate the allocation of the deliberately allocated leak demand that is uniformly distributed across the system. Table 2 illustrates the good goodness-of-fit between the observed values and the simulated values, including hydraulic grade for node xx1, xx2, xx3, xx7, xx9 and 101, and also the flow of pipe 'aa'.

Table 2 Comparison of the observed and simulated values for optimizing uniformly distributed leak detection

Elements	Obs- ered	Simulated values for different GA optimization runs									
		1	2	3	4	5	6	7	8	9	10
xx1 (m)	285.7	285.6	285.7	285.7	285.7	285.7	285.7	285.6	285.6	285.7	285.6
xx2 (m)	285.5	285.5	285.5	285.5	285.4	285.5	285.5	285.5	285.5	285.5	285.5
xx3 (m)	286.2	286.2	286.2	286.2	286.2	286.2	286.2	286.2	286.2	286.2	286.2
xx7 (m)	295.6	295.6	295.6	295.6	295.5	295.7	295.4	295.5	295.6	295.6	295.6
xx9 (m)	294.4	294.4	294.4	294.4	294.4	294.6	294.4	294.4	294.4	294.6	294.6
101 (m)	298.5	298.5	298.5	298.5	298.5	298.5	298.5	298.5	298.5	298.5	298.5
aa (L/s)	30.0	30.0	30.0	30.0	30.0	30.0	30.0	30.0	30.0	30.0	30.0

2. Big leak upstream

This test case was designed to deliberately allocate a big leak of 5.75 L/s at node 101 an upstream node near water source R-1, with a leak demand of 0.75 L/s being assigned to the rest of the demand nodes. The corresponding demand multipliers to be identified are 12.5 for node 101 and 2.5 for the rest nodes

Table 3 Optimization results for detection of big upstream leak

Node ID	Optimized demand factors of GA-based calibration runs										Average Factor	Correct factors
	1	2	3	4	5	6	7	8	9	10		
101	5.40	12.80	8.90	7.20	12.10	6.50	7.30	14.90	4.10	5.70	8.49	12.50
102	3.30	1.80	4.40	8.60	2.60	6.00	3.30	1.00	4.00	2.50	3.75	2.50
103	3.80	3.00	1.20	1.20	4.20	1.00	1.00	2.20	1.20	8.30	2.71	2.50
104	6.60	4.90	1.30	1.00	1.30	1.10	1.20	1.20	1.10	2.10	2.18	2.50
105	2.00	5.00	3.40	1.20	1.50	1.00	2.60	7.80	2.80	2.80	3.01	2.50
106	1.40	1.30	1.40	3.00	1.30	4.20	1.90	1.00	3.00	1.10	1.96	2.50
107	5.00	1.00	1.20	1.50	1.50	1.80	2.60	4.20	4.50	1.10	2.44	2.50
108	1.20	1.70	1.20	1.20	1.60	1.00	1.10	1.70	4.50	1.30	1.65	2.50
109	1.20	2.50	5.20	1.90	3.30	4.30	4.00	1.00	5.00	3.10	3.15	2.50
110	2.40	1.10	1.80	1.60	1.40	1.20	1.00	2.20	1.20	3.00	1.69	2.50
111	1.90	1.20	2.60	3.40	1.60	1.70	1.00	1.10	1.30	3.50	1.93	2.50
112	1.00	4.30	1.80	1.10	4.20	2.40	4.20	4.20	2.90	2.60	2.87	2.50
113	4.40	4.20	5.20	1.70	1.00	2.50	2.60	1.90	1.30	1.00	2.58	2.50
114	1.90	4.90	4.20	3.90	3.20	2.80	1.70	1.30	1.10	1.90	2.69	2.50
115	1.60	1.30	1.80	2.00	2.50	2.40	2.50	3.20	1.10	2.10	2.05	2.50
116	3.00	1.20	1.60	6.60	7.10	5.10	5.10	3.70	2.10	4.30	3.98	2.50
117	1.00	1.00	1.10	2.10	1.00	1.40	1.10	1.30	2.50	1.50	1.40	2.50
xx7	5.80	2.00	6.10	1.00	1.80	6.60	11.50	2.90	11.10	1.30	5.01	2.50

Table 3 indicates that the average demand factor for node 101 is 8.49, greater than the other nodes. It does indicate that a large leak may exist at the upstream node. However, as the big leak was assigned to an upstream node, the flow into the rest of system, where most of the field observed pressures were taken, does not seem to be very sensitive to the node demand adjustment. Hence, the identified demand factors for the rest of nodes were not consistently matched with their expected demand factors. To improve the accuracy, more flow data may be required apart from the total inflow observation for pipe 'aa'.

Table 4 Comparison of the observed and simulated values for optimizing big upstream leak detection

Elements	Obs-ered	Simulated values for different GA optimization runs									
		1	2	3	4	5	6	7	8	9	10
xx1 (m)	289.5	289.5	289.5	289.5	289.6	289.5	289.6	289.5	289.5	289.5	289.5
xx2 (m)	289.4	289.4	289.4	289.4	289.4	289.4	289.4	289.4	289.4	289.5	289.4
xx3 (m)	289.9	289.9	289.9	289.9	289.9	289.9	289.9	290.1	289.9	289.9	289.9
xx7 (m)	295.8	295.8	296.5	296.1	296.0	296.5	295.8	295.8	296.7	295.5	295.9
xx9 (m)	294.9	294.9	295.6	295.5	295.4	295.6	295.4	295.4	295.9	295.1	294.7
101 (m)	298.5	298.5	298.5	298.5	298.5	298.5	298.5	298.5	298.5	298.5	298.5
aa (L/s)	30.0	30.0	30.0	30.0	30.0	30.0	30.0	30.0	30.0	30.0	30.0

3. Big leak downstream

This case is contrasts with case 2. A big leak of 5.75 L/s is assigned to downstream node 117 and 0.75 L/s leakage demand is assigned to the rest of nodes. GA optimization calibration runs were performed to identify the correct demand factors of 13.34 for node 117 and 1.67 for the other nodes. The results for the ten optimization runs have been presented in Table 5. The average demand factors indicate that the downstream node 117 and immediate adjacent nodes (116,115,113 and 112) are significantly greater than the other nodes. The goodness-of-fit between the observed pressures and flows are shown as in Table 6.

Table 5 Optimization results for detection of big downstream leak

Node ID	Optimized demand factors of GA-based calibration runs										Average factors	Correct factors
	1	2	3	4	5	6	7	8	9	10		
101	1.00	1.00	1.00	1.00	1.05	1.00	1.00	1.00	1.00	1.00	1.01	1.67
102	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.05	1.01	1.67
103	1.00	1.00	1.05	1.00	1.00	1.00	1.00	1.00	1.00	1.15	1.02	1.67
104	1.05	1.10	1.10	1.00	1.05	1.10	1.10	2.65	1.00	2.10	1.33	1.67
105	1.10	1.10	1.00	2.60	3.05	1.00	1.00	1.65	1.05	1.50	1.51	1.67
106	1.15	1.00	1.40	1.05	1.45	1.60	1.60	1.00	1.20	1.15	1.26	1.67
107	1.20	1.00	1.05	1.10	1.15	1.15	1.15	1.30	1.60	1.10	1.18	1.67
108	1.05	1.10	1.40	1.00	1.05	1.25	1.25	1.00	1.05	1.40	1.16	1.67
109	1.25	1.40	2.25	1.05	1.10	1.25	1.25	1.10	1.10	1.55	1.33	1.67
110	1.90	1.10	1.55	1.00	2.00	4.40	4.40	1.10	1.25	1.05	1.98	1.67
111	1.10	15.00	1.00	1.40	1.20	2.50	2.50	2.00	1.25	1.60	2.96	1.67
112	7.55	3.35	4.05	13.75	6.20	3.25	3.25	8.05	9.65	7.00	6.61	1.67
113	3.20	2.55	5.40	10.40	15.00	15.00	15.00	5.40	7.05	15.00	9.40	1.67
114	2.45	2.55	5.60	1.00	2.50	2.15	2.15	4.95	4.80	3.35	3.15	1.67
115	13.70	3.95	7.25	12.10	7.30	7.35	7.35	6.55	8.95	5.75	8.03	1.67
116	2.30	15.00	5.85	4.00	6.50	4.65	4.65	2.25	6.70	9.10	6.10	1.67

| 117 | 15.00 | 3.80 | 15.00 | 2.55 | 4.40 | 7.35 | 7.35 | 15.00 | 7.35 | 1.85 | 7.97 | 13.34 |
| xx7 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.10 | 1.01 | 1.67 |

Table 6 Comparison of the observed and simulated values for optimizing big downstream leak detection

Elements	Obs-ered	Simulated values for different GA optimization runs									
		1	2	3	4	5	6	7	8	9	10
xx1 (m)	273.0	273.0	272.8	273.2	272.1	272.6	272.8	272.8	272.6	272.5	272.6
xx2 (m)	271.8	271.8	272.1	271.6	271.8	271.6	271.6	271.6	271.5	271.7	271.7
xx3 (m)	274.4	274.4	276.8	275.7	275.1	276.3	276.4	276.4	276.7	275.2	276.8
xx7 (m)	295.5	295.5	295.5	295.5	295.5	295.5	295.5	295.5	295.5	295.5	295.5
xx9 (m)	293.8	293.8	293.8	293.8	293.8	293.8	293.8	293.8	293.8	293.8	293.8
101 (m)	298.5	298.5	298.5	298.5	298.5	298.5	298.5	298.5	298.5	298.5	298.5
aa (L/s)	30.0	30.0	30.0	30.0	30.0	30.0	30.0	30.0	30.0	30.0	30.0

4. Big leak downstream and wrong pipe size

This testing scenario is similar to leakage scenario 3, except that a wrong pipe diameter of 400 mm is assumed for pipe 'ad', instead of 250 mm. The optimization calibration runs are launched to identify the leak under the wrong pipe size. As indicated by Table 7, the identified demand factors for node 117 and the adjacent nodes are not as significantly different from the rest of nodes as for case 3. But the impact of the wrong pipe size raises further interesting questions for the network modeler with more leakage being predicted further back upstream – this emphases that correct physical attributes are essential for this excise (or that the GA might even need further freedom to change pipe diameters).

Table 7 Optimization results for detection of big downstream leak and wrong pipe size

Node ID	Optimized demand factors of GA-based calibration runs										Average Factor	Correct factors
	1	2	3	4	5	6	7	8	9	10		
101	1.05	2.05	1.00	1.40	1.60	2.60	1.40	1.15	2.60	1.00	1.59	1.67
102	1.05	1.10	2.20	2.70	1.30	1.75	1.75	4.30	1.90	1.40	1.95	1.67
103	4.15	2.30	2.95	2.15	2.55	2.15	2.05	1.15	2.15	4.60	2.62	1.67
104	1.30	3.40	2.20	2.20	1.10	1.65	6.35	1.85	6.45	2.90	2.94	1.67
105	3.00	1.90	1.15	1.65	1.45	4.25	1.25	2.45	4.70	3.50	2.53	1.67
106	1.05	2.00	1.05	4.60	3.95	1.35	1.00	4.20	1.25	2.10	2.26	1.67
107	1.45	3.30	3.15	2.65	3.50	3.50	3.50	3.85	1.85	3.75	3.05	1.67
108	3.45	2.35	1.40	3.50	2.10	8.40	1.20	3.50	1.45	4.05	3.14	1.67
109	2.75	6.60	3.10	3.70	1.00	1.20	1.85	2.10	1.15	1.40	2.49	1.67
110	2.70	2.70	1.40	4.60	3.40	1.55	1.85	2.90	4.40	1.90	2.74	1.67
111	3.35	2.30	1.80	3.05	2.40	2.65	1.00	1.30	4.80	1.95	2.46	1.67
112	2.55	3.50	2.00	3.30	15.00	2.60	1.25	2.25	2.15	3.40	3.80	1.67
113	13.30	13.00	7.70	3.75	2.00	1.55	15.00	5.25	5.10	3.25	6.99	1.67
114	3.10	1.00	1.05	1.00	1.80	1.00	3.25	1.10	1.05	1.50	1.59	1.67
115	1.00	1.10	1.00	1.00	2.60	1.10	1.20	1.10	1.00	1.15	1.23	1.67
116	1.55	1.90	15.00	1.20	2.50	1.25	3.75	8.05	2.05	2.10	3.94	1.67
117	4.15	1.95	1.55	11.30	2.15	15.00	3.60	6.85	10.55	12.95	7.01	13.34
xx7	3.85	4.15	5.15	1.40	5.75	2.55	4.95	1.20	1.35	1.10	3.15	1.67

Table 8 Comparison of the observed and simulated values for optimizing big downstream leak detection and wrong pipe size

Elements	Obs-ered	Simulated values for different GA optimization runs									
		1	2	3	4	5	6	7	8	9	10
xx1 (m)	281.4	281.1	281.4	281.5	281.2	281.4	281.5	281.3	281.5	281.2	281.3
xx2 (m)	281.0	280.5	281.0	280.6	280.6	281.4	280.6	280.6	280.7	280.6	280.6
xx3 (m)	284.0	284.0	284.0	284.1	283.7	284.1	284.1	284.0	283.8	283.9	284.1
xx7 (m)	295.5	295.4	295.5	295.4	295.5	295.5	295.6	295.5	295.5	295.6	295.5
xx9 (m)	294.2	294.0	294.2	294.2	294.1	294.2	294.2	294.2	294.2	294.2	293.9
101 (m)	298.5	298.5	298.5	298.5	298.5	298.5	298.5	298.5	298.5	298.5	298.5
aa (L/s)	30.0	30.0	30.0	30.0	30.0	30.0	30.0	30.0	30.0	30.0	30.0

EXAMPLE II

An optimization-based calibration approach has been applied to a high leakage district meter area (DMA) in the UK (Sage 2005), where the field tests were carried out by UUNW with a higher density of pressure loggers than usual as part of the work to test out the methods being discussed here. However, UUNW has successfully identified leakage hotspots based on field test data used for normal all-mains model calibrations, typically deploying one logger per 200 houses.

The observed data contains the time series of flows into the DMA and pressures at a number of junctions. They are first processed and imported into the GA-based calibration tool as different field datasets with each set representing a complete set of data for one time step, For example, Field Date Set – Hour 6.5 contains the observed flow and the pressures at hour 6:30 AM. A total of 48 filed datasets over 24 hours have been imported into the GA-based calibration tool. One dataset at a particular time step can be used for model calibration at that time step.

To demonstrate the effectiveness of the GA-based calibration tool for the task, Four typical runs, two at peak demand (low pressure) hours of 6:30 AM and 6:00 PM, and two at low demand (high pressure) hours of 2:00 AM and 3:00 AM, have been summarized in Table 9. It shows that the observed inflow is consistently greater than the originally estimated demand. The difference between the observed inflow and originally estimated is deemed as a likely 'leakage' in the system. The genetic algorithm-based model calibration method is applied to determine how much the leakage is and where the leakages are while the pipe roughness is optimized as defined by Eq.(1) to Eq.(5).

Table 9 comparison of originally estimated, calibrated and observed demand at different hours

Hour	Originally Estimated Demand (L/s)	Observed Inflow (L/s)	Water Loss/Likely Leakage (L/s)	Calibrated Demand/Inflow (L/s)
2:00 AM	5.62	20.89	15.27	18.34
3:00 AM	4.92	18.00	12.08	17.30
6:30 AM	35.38	50.67	15.29	46.73
6:00 PM	35.41	47.00	11.59	46.97

The detected leakage at peak demand hours seem to be slightly lower than the low demand hours. This is because the pressure is greater at the low demand hour than high demand hour throughout the system. Hence, the GA process emulates the widely reported relationship between increased pressure and higher

leakage. The allocation of the leakage spots is presented in geographic information system (GIS). The detected leakage hotspots obtained for hour 2:00 AM is illustrated in Figure 3. The similar leakage maps can be readily produced for all the calibration snapshots.
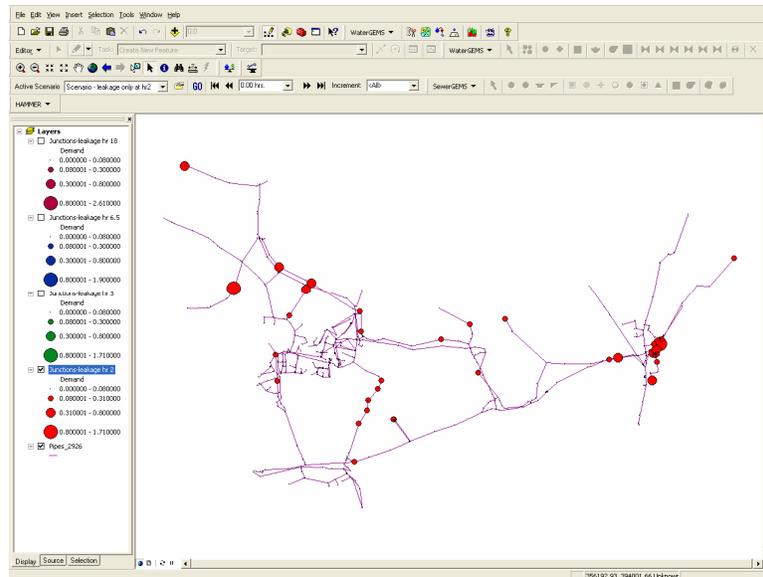

Figure 3 Leakage Map Detected by GA-based Model Calibration for DMA129-01 at Hour 2:00 AM

DISCUSSIONS

The optimization via model calibration has proved a viable approach for leak detection. For practical application, it requires more investigation of field data processes and field data analysis. In particular, UUNW has learnt the following lessons from the leakage detection studies carried out to date:

- An analysis of normalised field test data with head losses and associated area inflows can reveal useful information about locations of likely leakage hotspots;
- The density of field test equipment needs to be at least the same as that for a United Utility North West (UUNW) full field test for an all mains model. At UUNW this works out at about one logger for each 200 houses.
- Leakage hotspot prediction is an exercise that tries to achieve a better match between recorded trends and predicted ones;
- Assuming that high leakage might be more closely associated with some pipe materials and pipe ages than others doesn't necessarily lead to more reliable leakage hotspot predictions;
- several optimisation runs are required to test if likely leakage hotspot locations are repeatedly predicted in the same areas;
- Optimisation runs can be limited to the low flow conditions that prevail at night as day time demands are likely to be much more variable masking leakage behaviour;
- The GA process can identify unknown flows out of a study area and customers with large unrecorded consumptions;
- For trunk mains studies, the GA process can also identify upstream bulk meters that are over recording flows;
- Predicted leakage hotspots can only lie in the domain governed by the areas recorded incoming flow and pressure loggers that are on the periphery of the test area region.
- The method doesn't necessarily solve for system leaks but for any unusually high demand that is not metered or considered as base demand in the model.

10

- It is also essential that there is a significant head loss from the flow regulating element to the pressure monitoring point so that the pressure is sensitive to the model parameter adjustment of nodal demand and pipe roughness.

CONCLUSIONS

Reducing and managing leakage is of great importance to water utilities. This paper demonstrates that a method based on genetic algorithm optimization of a network model and its field test data offers a good support tool for network analysts to predict leakage hotspots thereby assisting utilities leakage reduction initiatives. The identification may not precisely point to the leak location, but significantly reduce the uncertainty and thus enable field crew to achieve better detection rates, more quickly within the areas where the leakage has been predicted. As a step forward, the optimization modelling tool needs improving so that modellers are more readily able to better deal with practical applications. This may include:

- Optimization for pressure dependent demand simulation;
- Rapid evaluation of field test data to suggest leaky locations;
- User interface enhancements for leakage prediction and leakage prediction outputs;
- Allowance for more model parameters for calibration such as pipe changing pipe diameters besides pipe roughness;
- Batch processing of GA optimisation runs.

REFERENCES

Almandoz J.; Cabrera, E.; Arregui, F.; Cabrera Jr., E. & Cobacho R. (2005). "Leakage Assessment Through Water Distribution Network Simulation." ASCE *J. of Water Resour. Plan. Manage*. 131(6).

American Water Works Association Water Loss Control Committee (AWWA). (2003). "Applying Worldwide BMPs in Water Loss Control." *J. of AWWA*, Jun., 65-79.

Bentley Systems, Incorporated (2006) "*WaterGEMS v8Users Manual*", Haestad Methods Solution Center, 27 Siemon Co Dr. Suite200W, Watertown, CT06795, USA.

International Water Association (IWA). (2000). *Losses from Water Supply System: Standard Terminology and Recommended Performance Measure*. IWA Task Force on Water Loss, London.

Lambert, A. O. & McKenzie, R. D. (2002). "Practical Experience in Using the Infrastructure Leakage Index.", *Proc. of IWA Conference in Leakage Management*, Lemesos, Cyprus, Nov., 2002.

Lambert, A. O. & Frantozzi, M. (2005) "Recent Advances in Calculating Economic Intervention Frequency for Active Leakage Control and Implications for Calculation of Economic Leakage Levels" *IWA International Conference on Water Economics, Statistics and Finance*, Rethymno, Greece, 8-10, July 2005.

Sage, P (2004). "Developments in Use of Network Models for Leakage Management at United Utilities North West." CIWEM North West and North Wales Branch Water Treatment and Distribution Conference, Warrington, UK, Nov 2005

Thornton, J., 2002, *Water Loss Control Manual,* McGraw-Hill, New York.

Walski, T.M., Bezts, W., Posluszny, E.T., Weir, M. and Whitman, B.E., 2006, "Modeling Leakage Reduction through Pressure Control," *J AWWA, 98:4, p. 147-155*.

Wu, Z. Y., Elio F. A. and Ernesto G. (2004) "Darwin Calibrator — Improving Project Productivity and Model Quality for Large Water Systems" *Journal of AWWA*, Vol. 96, No.10, pp27-34.

Wu, Z. Y. Wang, R. H., Diezo, D., Walski, T. (2003) "Mining Water Consumption and GIS-based Data for Loading Water Distribution Models", in Proceeding of 2003 World Water and Environmental Resource Congress, June 22-26, Philadelphia, Pennsylvania.

Wu, Z. Y, Walski, T., Mankowski, R., Cook, J. Tryby, M. and Herrin G. (2002) "Calibrating Water Distribution Model Via Genetic Algorithms", in *Proceedings of the AWWA IMTech Conference*, April 16-19, Kansas City, MI.