Journal of Hydraulic Research Vol. 46 Extra Issue 1 (2008), pp. 129–137

Systematic exploration of pipeline network calibration using transients

Exploration systématique du calibrage d'un réseau de canalisations en utilisant des coupures

B.S. JUNG, Professional Engineer, *MWH Soft., 618 Michillinda Avenue Suite 200, Arcadia, CA 91007, USA*

B.W. KARNEY, Professor, *Department of Civil Engineering, University of Toronto, 35 St. George St., Toronto, ON, M5S 1A4, Canada*

ABSTRACT

Accurate information of pipeline and system properties is crucially important for precise computer simulation of pipe networks. Even though inverse transient analysis (ITA) techniques have been widely investigated for leak detection and friction factor calibration, many challenges still remain. One reason for these difficulties is that real water distribution systems invariably have many kinds of uncertainties including pipe diameter, wave speed and the water demand at the time of the tests. This paper investigates quantitatively how inaccuracies in such values invariably deteriorate the performance of calibration approaches. Thus, the paper argues that a systematic calibration should explicitly include these additional uncertainties during the ITA process. Two evolutionary approaches, namely Genetic Algorithms and Particle Swarm Optimization, are applied and are both compared and contrasted during the ITA iterations. The evolutionary algorithms help the search escape from poor local optima in multifaceted and complex problems and thus assist in locating global (or near-global) optima. However, most current approaches are shown to converge poorly as the full scale of the typical field problems is progressively reflected in the search space.

RÉSUMÉ

Une information précise sur les propriétés des canalisations et du système est d'une importance cruciale pour la précision de la simulation sur ordinateur de ces réseaux. Quoique les techniques d'analyse inverse de transitoire (ITA) aient été largement étudiées pour la détection de fuite et le calibrage du coefficient de frottement, beaucoup de défis demeurent encore. Une raison de ces difficultés est que, dans la réalité, les systèmes de distribution d'eau présentent invariablement de nombreuses incertitudes quant au diamètre des conduites, la célérité des ondes et la demande d'eau à l'heure des essais. Cet article étudie quantitativement comment les incertitudes sur ces valeurs détériorent invariablement la performance des approches de calibrage. Ainsi, le papier argue du fait qu'un calibrage systématique devrait explicitement inclure ces incertitudes additionnelles durant le processus d'ITA. Deux approches évolutives, à savoir des Algorithmes Génétiques et une Optimisation par Essaim de Particules, sont appliquées et sont toutes deux comparées et mises en contraste pendant les itérations d'ITA. Les algorithmes évolutifs aident la recherche à échapper à des optimums locaux insuffisants dans les problèmes complexes multi-facettes et aident ainsi à localiser des optimums globaux (ou quasi-globaux). Pourtant, la plupart des approches courantes ne paraissent converger que faiblement tandis que l'échelle complète des problèmes typiques concrets est progressivement reflétée dans l'espace de la recherche.

Keywords: Water distribution system, calibration, inverse transient analysis, sensitivity analysis, evolutionary optimization, particle swam optimization, genetic algorithm.

1 Introduction

Before any pipeline network model can be adequately used to analyze the performance of an existing water distribution system, the model must first be calibrated. Calibration involves adjusting network parameters until the model results closely approximate observed conditions, ideally obtained from measured field data. Since the confidence and value of the model is strongly dependent on the confidence associated with the parameter estimates, the calibration of water distribution system (WDS) is economically and strategically important.

Boulos and Wood (1990) present an explicit algorithm for directly determining a variety of design, operating, and calibration parameters. Their approach uses the steady-state network equations but requires a one-to-one relation between the selection of target parameters and the specification of pressure and flow constraints. Similarly, Lansey and Basnet (1991) present a nonlinear programming technique for estimating WDS parameters, showing that the model could match measured and computed values of pressure head, pipe flow, and tank levels. Pudar and Liggett (1992) calibrated a steady-state WDS model to detect leakage in a hypothetical network. Their objective function, the sum of squares of the differences between measured and calculated flows and heads, was minimized by the Levenberg–Marquardt method. However, Pudar and Liggett recognize that a shortage of data often renders the inverse problem underdetermined; in most realistic applications, inverse methods require a great deal of experimental data.

Revision received November 13, 2006/Open for discussion until August 31, 2008.

Considering that transient pressure waves are attenuated by leakage and water demand, several direct response monitoring techniques have been investigated for leakage detection. Over 80 years ago, Babbitt (1920) reviewed several leak detection techniques, almost all of which are still familiar in modern approaches including using a system's water hammer response. Brunone and Ferrante (2001) used the transient test presenting the influence on the pressure signal of size and shape of small leaks. Stoianov *et al.* (2002) used the direct response of a system to a pressure transient. A Wavelet Transform was applied for signal identification (de-noising) and combined with an Artificial Neural Network (ANN) approach for classification of the extracted features to determine leak location and quantification in a pipeline. However, this application of ANN requires a sufficient number of training data sets that are acquired by simulating a real leak case, which is problematic in field applications. Liou (1998) presents a method using the direct response of transient signals by impulse response extraction. The impulse response of the pipeline in the presence of noise is extracted by using cross-correlations between a low amplitude pseudo-random binary disturbance input and the system's output. This impulse response is applied for real-time pipeline leak detection; however, its application is limited to a single pipeline due to the linear system restriction.

Recently, a calibration method called "Inverse Transient Analysis" (ITA) has been widely investigated and advocated (Liggett and Chen, 1994; Nash and Karney, 1999; Vitkovsky *et al.*, 2000; Jung and Karney, 2004a). The approach here is to create a transient condition and to use high-frequency pressure transducers to record the system's response. Next, an optimization model is used to adjust (and thus evaluate) system parameters to bring the predicted and measured responses into reasonable agreement. Many different optimization models and approaches have been developed for ITA. Gradient methods like the Levenberg–Marquardt method were initially most widely used (Liggett and Chen, 1994; Nash and Karney, 1999). The gradient-based methods have the good advantage of computational efficiency; however, they may fail to converge, due to instability problems, or may converge to a local rather than global optimum (Back *et al.*, 1997; Jung and Karney, 2006; Jung *et al.*, 2006). Recently, Genetic Algorithms (GA) have been popular in water distribution optimization (Goldberg, 1989; Simpson *et al.*, 1994; Dandy *et al.*, 1996) to cope with using the traditional gradient-type methods. Vitkovsky *et al.* (2000) performed GA optimization in conjunction with the inverse transient method to detect leaks and friction factors in WDS. Kapelan and Savic (2003) introduce the hybrid genetic algorithm, called HGA, for ITA. Vitkovsky *et al.* (2003) found the performance of ITA is strongly dependent on the quantity and location of data measurement site so they presented an approach for determining the configuration of measurement sites that produces optimal results. Jung and Karney (2004b) used GA and Particle Swarm Optimization (PSO) for the numerical investigation of ITA. Recently, Kapelan *et al.* (2005) summarized the role of "prior information" on parameters incorporated in the calibration of WDS transient simulation models by adding the sum of weighed squared prior information residuals to the classic sum of weighted squared sample information residuals. The explicit type of prior information improves the conditioning of ill-posed inverse transient problem and thus improves the efficiency of algorithm. Kim (2005) applied an impulse response method as a leak detection algorithm. To consider the impact of unsteady friction, Kim employed frequency-dependent friction under laminar conditions and local and convective velocity accelerations for turbulent conditions. A GA was integrated into the impulse response method to calibrate the location and the quantity of leakage.

Despite ongoing research into ITA techniques for water distribution systems, many significant challenges remain for field applications. In fact, even though model-based methods are excellent for numerical calibration tests (Liggett and Chen, 1994; Vitkovsky *et al.*, 2000; Jung and Karney, 2004b), these methods have not been widely applied to field experiments. One of the main reasons is that the existing transient models are slow and awkward for accurately representing the full and complex behavior of a system. Many complications, like unsteady friction factors and the possible presence of air, remain as challenging issues. In addition, real water distribution systems often have many uncertainties in addition to the leakage rates and friction factors that are conventionally considered as unknowns. For example, even properties such as the pipe diameter, wave speed (i.e., wall material), the exact value of the water demand at the time of the tests, and uncertain measurement accuracy, all add to the intricacy and difficulty of obtaining a reliable calibration. Filion and Karney (2003) provided an overview of several important issues of uncertainty and error that arise in modeling water distribution systems.

This paper systematically explores inverse transient analysis for leak detection and system calibration in a small test network. Evolutionary algorithms, specifically both a Genetic Algorithms and a Particle Swarm approach are applied as optimization tools to solve an inverse problem in which system parameters (such as lumped leak coefficients and friction factors) are determined from measured or numerical transient pressure head data. This paper also investigates the difficulties and limitations of ITA as the problem grows in complexity, with obvious application to real WDSs. Specific uncertainties such as pipe diameter, wave speed, and water demand, in additional to the traditional issues of leakage and friction factor, are explicitly included to create a more systematic and comprehensive ITA approach.

2 Calibration uncertainties

The primary objective of a simulation is to reproduce the behavior of a real or proposed system and its spatial and dynamic characteristics in a realistic and useful way. To accomplish this goal, data are supplied that depict the physical characteristics of the system, the loads placed on the system, and the boundary conditions in effect (Walski *et al.*, 2001). However, even if all of the data gathered describing the model match the real system, the calculated pressures and flows will always have some discrepancies with the observed ones. One of the main reasons arises from mathematical assumptions associated with achieving computational tractability in the simulation model and associated technical difficulties in representing the real system. Filion and Karney (2003) provided the discussion of many common sources of errors and uncertainties in hydraulic solution. Classic water hammer theory based on the assumptions of linear elastic behaviour of pipe-walls and quasi-steady-state friction losses is used to predict the maximum and minimum pressure surges; this is relatively accurate for simulating hydraulic transients in metal pipe, but it is considerably less precise for plastic pipes, particularly when the surge is generated by rapid changes in flow conditions. This unsteady friction characterization is a challenging question and research is on-going (e.g., Brunone *et al.*, 1991; Vitkovsky, 2001; Covas *et al.*, 2004).

Another critical reason for the problems with ITA, especially for field applications, is the uncertain representation of the real system. Traditionally, ITA is accomplished by adjusting internal pipe friction values and nodal leakages until an agreement between observed and computed pressures and flows are obtained. This approach implicitly assumes that other system characteristics like pipe diameters, water demands, wave speeds, pressures, and flows are known or directly measured; by contrast, pipe friction values and nodal demands are typically estimated. Hence, there is room for adjustment in these values to improve the accuracy of the estimation. However, numerous factors can contribute to disagreement between model and field observation. All uncertain input data are candidates for adjustment during calibration and in the quest to obtain reasonable agreement between model-predicted behavior and actual-field behavior.

Even determining a parameter like pipe diameter is not as straightforward as it might seem. The internal diameter of pipe may differ from the nominal diameter that is all too commonly used in modeling. Further, the internal diameter often decreases over time as corrosion, tuberculation, and scaling occur. The reduced pipe diameter has a great influence on the head loss, velocity and residence time through a pipe. A 10% decrease in the pipe diameter, for example, will increase head loss by nearly 40% (Walski *et al.*, 2001). Steady-state hydraulic calibration models generally employ a resistance coefficient, instead of adjusting both the friction coefficient and diameter. By adjusting friction coefficients in calibration but leaving diameter at assumed known values, the number of variables is reduced by half and the calibration process is simplified. However, this mathematical simplification less accurately represents the physical system. Moreover, it would be even more problematic during transient calibration (or, for that matter, for water quality simulation) because pipe diameter is strongly linked through velocity to system dynamics.

Another uncertainty in a water distribution system is water usage or demand, which continuously varies for many reasons. In order for a transient simulation to accurately reflect system dynamics, these demand fluctuations should be incorporated into the model; but, at least over short intervals, demands are typically modeled as constant, independent of both pressure and time. Indeed, little data is usually available about exact values of instantaneous demand in complex real pipe systems. Also, there is a spatial problem that the water distribution modeling equations are based on the simplifying assumption that water is withdrawn at a junction node. In reality, however, water usage tends to occur along the entire length of pipe. Spatially redistributing water usages that occur along a length of pipe to the junction nodes in the model is widely used and known as "demand allocation". The demand allocation process is a possible source of error that should be considered when calibrating a model. It is conceivable that a model could incorporate all of the locations where water is withdrawn from the system by placing junction nodes where the service lines are connected to the water main. This approach, however, would significantly increase the number of pipes and nodes required in the model, thereby increasing its complexity and the data requirements, particularly for a transient model. McInnis and Karney (1995) presented a formulation to permit system demand as a distributed pipe flux in transient model to provide a more uniform distribution of water consumption. This approach was compared with an aggregated discrete constant withdrawal model and an aggregated withdrawal using an equivalent orifice that accounted for the pressure-dependent nature of demand.

Wave speed is another challenging uncertainty in a pipeline. Wave speed is a function of many fluid and pipe properties (pipe diameter, thickness and material; water density, elasticity, temperature, air and solids content; pipe restraint conditions; etc.). Some of these conditions can be accurately assessed, but many are poorly defined and uncertain. For example, accurate measurement of the air content dispersed in fluid is difficult; however, even a tiny amount of gas throughout a liquid greatly reduces the propagation velocity of a pressure wave in a pipeline (Wylie and Streeter, 1993).

In this paper, three uncertainties (pipe diameter, water demand, and wave speed) are considered to examine systemically how these uncertainties can deteriorate system calibration. In addition, an ITA approach including these uncertainties, in addition to leak and friction factor, is explored. One of the goals is to assess the limits of the ITA approach when facing a significant uncertainty; that is, to determine when ITA is likely to fail, or at least to produce misleading and inaccurate results.

A more systematic and comprehensive approach to ITA can be partly assessed through a sensitivity study. The sensitivity of the ITA to wrong information certainly will strongly depend on system characteristics (including system topography, pipe size, material, and thickness) and the transient characteristics (such as severity and source of initiating event). In addition, the sampling rate, the range of inaccurate information, the quantity, and location of measurement sites (Vitkovsky *et al.*, 2003) would also crucially affect sensitivity results. Although important, such a detailed analysis is beyond the scope of this paper.

3 Evolutionary algorithms

Evolutionary Algorithms (EA) is an optimization system inspired by biological evolution and adaptation. It aims at developing more robust and efficient algorithms for solving complex real-world optimization problems. The specific problems dealt with by such complex systems are usually highly nonlinear and often contain inaccurate and noisy data (Back *et al.*, 1997). Probably, the most well known and widely used EA paradigm is the GA, created by John Holland and made popular at least for engineers by Goldberg (1989). In the GA approach, a population of possible solution is biologically encoded and genetic operators, designed to mimic genetic evolution, cause the population to evolve, and thus to explore the search space efficiently. In this paper, the probabilities of mutation and (uniform) crossover are obtained empirically as 0.02 and 0.5, respectively. Tournament selection and elitism (in which the best individual is copied to the next generation) are selected.

Recently, a new EA field has arisen, based on another population search procedure, called Swarm Intelligence (Kennedy and Eberhart, 1995). Swarm Intelligence argues that intelligent cognition derives from the interaction of individuals in a social environment and that the main ideas of sociocognition can be effectively applied to develop stable and efficient algorithms for optimization tasks. One of SI techniques, called Particle Swarm Optimization (PSO), has been developed by Kennedy and Eberhart (1995) to simulate the movement of a flock of birds searching for food. PSO has been used mainly for continuous optimization in which the population of potential solutions is called a "swarm". In the computer version of this search, the global exchange of information among all individuals, which are called "particles", takes place and each particle can profit from the discoveries of the rest of the swarm. Many variants of the PSO technique have been developed (Kennedy and Eberhart, 2001).

This paper uses a version of the algorithm derived by adding an inertial weight to the original PSO dynamics (Shi and Eberhart, 1998; Jung and Karney, 2004a; Jung and Karney, 2004b). Assuming that the search space is D-dimensional, we denote the current position by $X_i = (x_{i1}, x_{i2}, \ldots, x_{iD})$ of the ith particle of the swarm and by $P_i = (p_{i1}, p_{i2}, \ldots, p_{iD})$ the best position it ever had within the search space. Let g be the index of the best particle in the swarm and $V_i = (v_{i1}, v_{i2}, \ldots, v_{iD})$ the velocity (position change) of the ith particle. The swarm is manipulated according to the following equations:

$$v_{\mathrm{id}} = w v_{\mathrm{id}} + c_1 r_1 (p_{\mathrm{id}} - x_{\mathrm{id}}) + c_2 r_2 (p_{\mathrm{gd}} - x_{\mathrm{id}}) \quad (1)$$

$$x_{\mathrm{id}} = x_{\mathrm{id}} + v_{\mathrm{id}} \quad (2)$$

where $d = 1, 2, \ldots, D$; $i = 1, 2, \ldots, N$ and N is the size of the population; w is the inertial weight; c_1 and c_2 are positive constants; and r_1 and r_2 are two random values that are uniformly distributed in the range [0, 1]. Equation (1) is used to calculate the ith particle's new velocity, a determination that takes into consideration three main terms: (i) the particle's previous velocity, (ii) the distance of the particle's current position from its own best position, and (iii) the distance of the particle's current position from the swarm's best experience (position of the best particle). Thus, each particle or potential solution moves to a new position according to (2). The performance of each particle is measured using a predefined fitness function.

The initialization of the swarm is done using a uniform distribution over the search space. The inertial weight w plays an important role for the convergence behavior of the technique. It is used to control the impact of the previous history of velocities on its current value for each particle, regulating in this way the tradeoff between the global and local exploration abilities of the swarm, since large values of w facilitate global exploration of the search space (visiting new regions) while small values facilitate local exploration (fine-tuning the current search area). By gradually decreasing the inertial weight from a relatively large value to a small value through the course of the PSO run, the PSO tends to exhibit a more global search ability at the beginning of a run and a more local search ability near the end. Following an empirical study of PSO (Shi and Eberhard, 1999), a linearly decreasing inertia weight is used which starts at 0.9 and ends at 0.4, with $c_1 = 2$ and $c_2 = 2$.

Even though EAs are usually quite efficient and typically converge to near global optimum solutions, they can guarantee convergence only when given unlimited computer resources; for finite runs, convergence occurs only in a probabilistic sense. The nature of the probabilistic convergence strongly depends on the population size and number of generations, but practical selection must also take into account computational time. Thus, an optimal selection of number of trials, population size, and number of generations continues to be a challenging multi-objective question. Certainly increasing the population, the number of generations, or the number of random seeds will all tend to decrease the risk of premature convergence, but at significant computational cost. Jung *et al.* (2006), in a benchmark study of quite a broad class of EAs for a range of optimization problems, show that the performance of EAs strongly depends on the optimization problem characteristics. In Jung's (2005) approach, GA and PSO are tested with benchmark functions using multiple independent runs. Jung found that although multiple independent runs do help to improve convergence, single runs usually achieve excellent approximations. Considering the high computational cost of transient analysis, single runs with different optimization approaches (e.g., GA and PSO) are used here to help control and assess convergence.

4 Case studies

A case study, based on a small network introduced by Pudar and Liggett (1992) and applied to several other approaches (Liggett and Chen, 1994; Vitkovsky *et al.*, 2000), is extended to illustrate a more challenging calibration application. The network system shown in Fig. 1 comprises one reservoir, 7 nodes and 11 pipes. Water is drawn from the higher reservoir (50 m) to the downstream network and supplied to two external demands (at nodes 4 and 7 with flows of 20 and 40 L/s, respectively). The pipe diameter, length, and Darcy-Weisbach friction factors for all pipes are initially given as 0.2, 500 and 0.02 m, respectively. The pipe material is set as steel with a wave speed of 1200 m/s. One computational reach is selected for each pipe so the time step is 5/12 s. All nodal elevations except the reservoir are at datum.

To initiate transient conditions for the ITA procedure, a simple transient event caused by an instantaneous valve closure at node

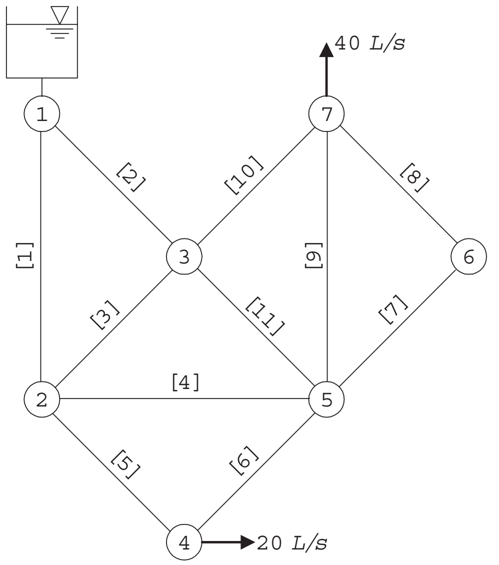

Figure 1 Network system of 7 nodes and 11 pipes.

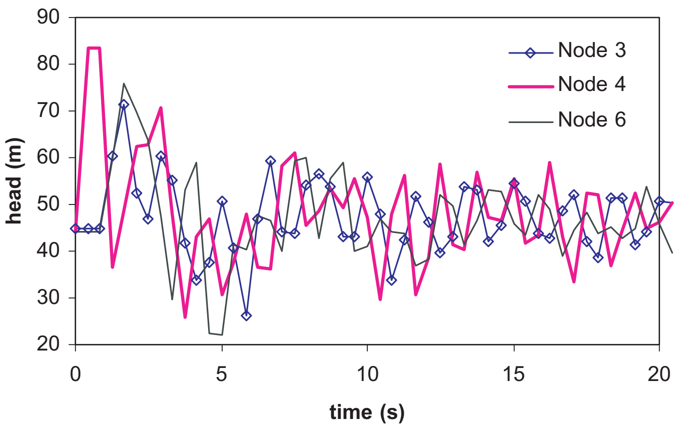

Figure 2 Head traces: measurement data for inverse analysis.

4 is assumed. The leak point is located at node 2 and the leak area is $0.001\ m^2$. Figure 2 shows the transient head traces of nodes 3, 4 and 6 due to the rapid valve closure at node 4. The possible leakage nodes during the ITA are selected as all 4 "internal" nodes (i.e., all nodes excluding the reservoir and demand locations). The permitted range of leak area is 0 to $0.005\ m^2$. Pipe friction factors (11 unknowns) are chosen for all pipes within the range of 0.012 to 0.04. The number and location of measurement sites are selected for each case study differently; the length of record for the measurement sites is fixed at 40 s, so each numerical measurement site provides 97 $[40/(5/12)+1]$ values of the pressure head. Following an empirical study of PSO by Shi and Eberhard (1999), the maximum velocities, meaning the maximum change the particles of PSO can move in a single iteration, for leakage and friction factor are set as 0.0005 and 0.003, which are approximately 10% of search ranges. The population size and generation (iteration) number for both optimizations are selected depending on specific case studies.

4.1 *ITA for leakage and friction factor*

Traditional ITA approaches that only consider leakage and friction factor are first explored. The population size and number of generations (iterations) for both optimizations are set at 200. For the inverse calculations, three different measurement combinations—one site (node 6), three sites (nodes 3, 4 and 6) and six sites (nodes 2–7)—are selected to evaluate how this data influences the ITA results. The objective function of GA and PSO is to minimize the square sum of the difference between measured head trace and the predicted values during the inverse calculation.

Figure 3 shows the search results for finding the leakage for four possible leakage nodes and friction factors for all pipes. Both optimizations show, regardless of the number of measurement sites, fast initial convergence during the first 50 generations, although PSO exhibits a faster and more precise convergence than does GA. Table 1 presents the calibration results using GA and PSO with different combinations of the measurement sites. Both optimizations using fewer measurement sites should be carefully accessed since less information may increase the risk of the problem becoming underdetermined. In an underdetermined system, apparent convergence may not provide a correct parameter set, although it would provide the decent model fit. Figure 3 and Table 1 show that using fewer measurement sites leads to a better fitness value and faster convergence, but also to less accurate calibration results. Also, Table 1 shows both optimizations provide accurate calibration results for leakages, but show less precision for the friction factors. Jung and Karney (2004a) also found that the inverse calibration of friction factor is less accurate than that of leakage in their numerical experiments. Chen (1995) and following Vitkovsky *et al.* (2003) propose a convincing reason: the sensitivities with respect to the leaks are typically three orders of magnitude greater than those with respect to the friction factors. The mismatch in sensitivities, therefore, causes the optimization to occur for the leak parameter rather than for both parameter types.

4.2 *ITA with wrong system information*

The same ITA calculations for leakages and friction factors are now performed again, but with three new sources of uncertainty

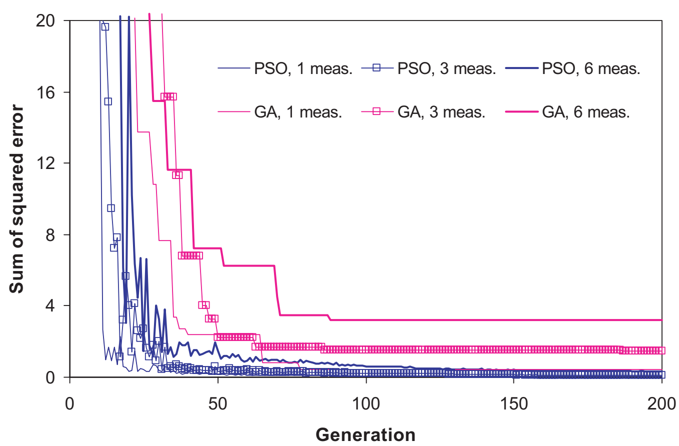

Figure 3 Evolution procedure for leakage and friction factor.

Table 1 Leakage and friction factor calibration with GA and PSO

	Node/ pipe	Real	Number of measurement sites					
			1		3		6	
			GA	PSO	GA	PSO	GA	PSO
Possible	2	1	1.02	1.00	1.02	1.00	1.02	1.00
leakage	3	0	0.00	0.00	0.00	0.00	0.00	0.00
node	5	0	0.00	0.00	0.00	0.00	0.00	0.00
(10^{-3} m^2)	6	0	0.00	0.00	0.00	0.00	0.00	0.00
Unknown	1	2	1.77	2.28	1.88	2.06	2.13	1.95
friction	2	2	2.06	1.77	2.04	1.94	1.80	2.04
factor	3	2	1.64	3.24	1.47	2.79	2.39	1.93
(10^{-2})	4	2	2.70	3.00	1.42	2.26	3.58	1.52
	5	2	3.38	4.00	3.45	4.00	1.33	4.00
	6	2	2.32	2.97	3.76	4.00	2.90	1.20
	7	2	3.27	4.00	3.89	3.46	2.68	2.73
	8	2	1.55	3.11	3.25	2.50	2.19	2.93
	9	2	1.53	1.98	1.46	2.19	1.88	1.85
	10	2	1.84	1.20	1.46	1.20	1.53	1.77
	11	2	1.86	2.77	3.29	3.22	2.21	2.61
Sum of square error		—	0.453	0.0324	1.46	0.0341	3.23	0.346

(in fact, sources of inaccuracy): pipe diameter, water demand, and wave speed. These terms provide a greater degree of realism but complicate the numerical convergence. The parameters for both optimizations are same as those in Section 4.1 but measurement sites are fixed with three nodes (nodes 3, 4 and 6).

Different pipe diameters and wave speeds in the system create uneven travel times in constructing characteristic grids. The smallest Courant number in the system can be adjusted to unity by dividing the pipe into smaller computational units. The process of discretization is repeated until the smallest Courant number exceeds 0.75. After the discretization, a linear timeline interpolation (Jung and Karney, 2004b) is used to obtain head and flow at a grid point in the characteristic mesh for the uneven computational units. This rediscretization problem arises whenever wave speed or pipe lengths are uneven.

For convenience, the uncertainties are introduced with wrongly specified parameter values in the case study. First, a 20% error in the diameter of pipe 6 (0.2–0.16 m) was introduced and its impact on convergence is explored. Specifically, what was done was that the forward simulation was run with a diameter of 0.2 m and the inverse simulation used a diameter of 0.16 m, thus distorting the nature of the response. Second, incorrect information for the demand at node 7 was explored. For simplicity, the water demand was assumed constant with time but a 20% error (40 to 48 L/s) was introduced. Third, the wrong information for the pipe material was used. A plastic pipe, instead of steel, at pipe 6 was assumed, so the corresponding wave speed was taken as 300 m/s not 1200 m/s as it was in the forward calculation. The squared sum of the difference between measured and calculated data, with determined leakages and friction factors using wrong information of diameter, demand and wave speed are 10600, 230, and 20200, respectively. Hence, at least in this case, the wave speed is most influential parameter among the three deteriorating ITA in this case studies, although clearly its assumed uncertainty was bigger than the others.

Figure 4 shows the average relative error between the corrected and calibrated leakages and frictions using incorrect base information. For comparison, each error is divided by its search range to show its relative error. The fitness values of GA are similar with those of PSO and the fitness results using the wrong information of diameter, demand, and wave speed are 5300, 3.1, and 17000, respectively. As expected, the wrong information for system parameters contaminates the system calibration so the calibrated results are much poorer than those of the case studies in Section 4.1. Due to the varying influence of each term, demand errors are less influential than pipe diameter and wave speed. As in the previous case studies (Section 4.1), using the wrong demand shows the leak calibration is relatively accurate but the friction factor calibration is less so. However, the case using the wrong pipe diameter provides poor calibration results for leakages as well as friction factors. Because pipe diameter has a more direct influence on the system dynamics, a small diameter discrepancy can greatly contaminate the calibration results using transient signals.

Of the three cases studied, inaccurate information for the wave speed gives rise to the worst calibration result. Figure 4 clearly demonstrates that, in this case, neither optimization approach can find accurate friction factors, nor even current leakages. The reason for these poor calibration results is that a fourfold change in wave speed at pipe 6 strongly affects not only the magnitude of the transient response, but also the phase of the dynamic behavior. These distortions cause a significant mismatch in the ITA results. A simple but effective approach to improve this problem is to use mild transient instead of rapid one, thus minimizing the

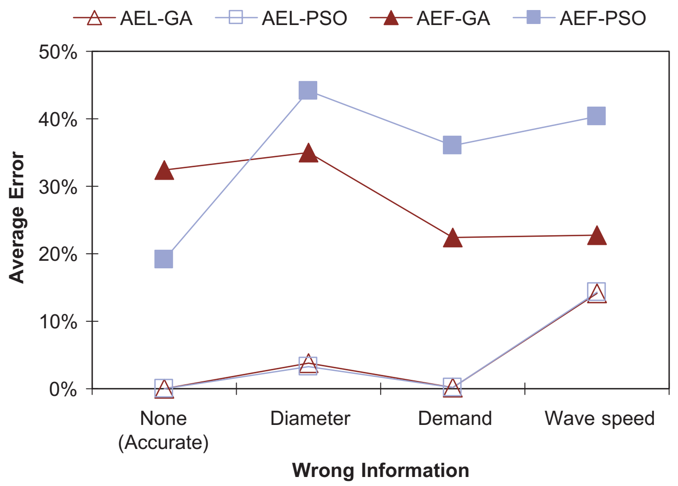

Figure 4 Average errors of leakage and friction factor. AEF, average error friction factor; AEL, average error of leakage.

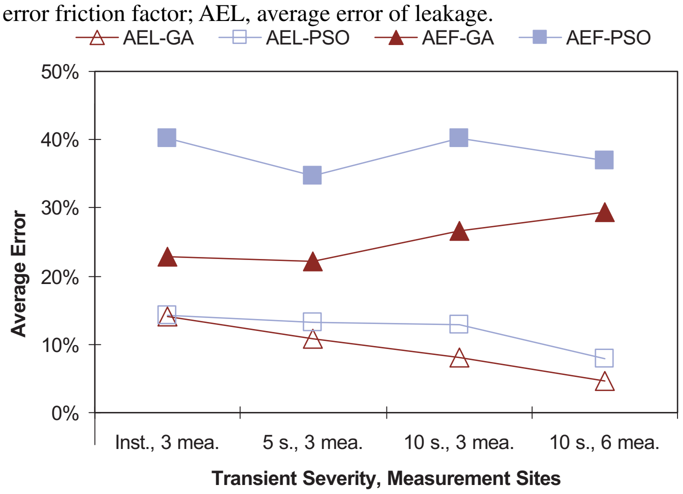

Figure 5 Calibration with mild transient and increased measurement sites.

effect of the wave speed. However, Jung and Karney (2004a) indicated mild transients often lead to a false convergence; too mild a transient simply does not uniquely characterize the system's response. Hence, the increased number of measurement sites would be helpful to improve the "non-sensitivity" of mild transient for ITA. Figure 5 shows two less sudden transients having a 5 and 10 s valve closure at node 4 and using an increased number of measurement sites (6 nodes). As expected, the fitness values with the less severe transients of 5 s valve closure (1260 in GA; 1200 in PSO) and 10 s valve closure (201 in GA; 187 in PSO) are much smaller than that with rapid transient (16700 in GA; 16700 in PSO). The leakage calibration with GA and PSO becomes more accurate as transient becomes milder in Fig. 5. Also, the addition of measurement sites increases the calibration accuracy for leakage detection. However, both a milder transient and increased measurement sites have little benefit for producing a more accurate calibration of friction factors.

4.3 ITA with all uncertainties

Since the case studies in Section 4.2 show that only a small amount of wrong information about the system parameters causes a high contamination of the ITA results, an obvious remedy is to treat all parameters that are uncertain as unknowns. Hence, calibration with all parameters is potentially desirable for ITA, especially for field applications.

The current case study considers all parameters as calibration candidates: four possible leakages at nodes 2, 3, 5, and 6; unknown friction factors for all 11 pipes; unknown pipe diameters for all 11 pipes; an unknown water demand at node 7; and unknown wave speeds for all 11 pipes. However, an increased number of unknowns and different sensitivities of system parameters are crucially challenging problem in both PSO and GA optimizations. The results of the case studies with inaccurate system information indicate that large changes in wave speed strongly affects the magnitude and phase of the transient response. Taking this into account, preliminary calibration tests are explored using PSO with 100,000 evaluations. Table 2 shows ITA results for four unknowns for all parameters using one measurement site at node 6. All parameters are first tested; then the test is repeated with known wave speed. The ITA results shown in Table 2 indicate the wave speeds, at least in this case study, affect more strongly the inverse calibration procedure than the other parameters, so this uncertainty tends to spoil the calibration of the others. Therefore, a more simplified binary range of wave speeds, 300 or 1200 m/s are considered in the case studies. The two values of wave speed roughly correspond with two different pipe materials, such as plastic and steel.

Due to the complexity of the optimization, the calibration starts with unknown value of all parameters for one pipe and one node only and progressively increases the number of pipes and nodes being calculated to two, four, and, finally, the value of all parameters. Due to the increased number of unknowns, the population size and generation (iteration) number for both

Table 2 Preliminary calibration tests using PSO

	Node pipe	Real	All	All except wave speed
Leakage (10^{-3} m^2)	2	1	4.94	1.01
	3	0	3.80	0.00
	5	0	10.0	0.00
	6	0	7.86	0.00
Friction factor (10^{-2})	1	2	2.29	1.58
	2	2	3.38	2.42
	3	2	2.32	1.20
	4	2	1.63	1.20
Pipe diameter (10^{-1} m)	1	2	1.00	2.00
	2	2	1.51	2.00
	3	2	1.39	2.00
	4	2	1.44	2.00
Water demand (10^{-2} m^3/s)	7	4	6.00	4.12
Wave speed (10^3 m/s)	1	1	1.06	—
	2	1	1.03	—
	3	1	1.25	—
	4	1	0.893	—
Sum of square error		—	3610	0.0622

optimizations are set as 500 and 200, respectively. For the inverse calculation, three measurement sites at nodes 3, 4, and 6 are selected. The ranges of leak area and friction factor are same as those in Section 4.1. The ranges of pipe diameter and water demand are 100–200 mm and 20–60 L/s, respectively. For the PSO approach, the maximum velocities of pipe diameter and water demand are set as 10 mm and 4 L/s, respectively.

Figure 6 shows the evolutionary history of both GA and PSO for all parameters. Similarly to the case studies in Section 4.1, PSO provides faster convergence and produces more accurate results than GA except for the case of the all 11 unknowns. Figure 7 shows the average relative errors between real and calibrated five parameters. Both optimizations in Fig. 7 show accurate calibration for first three cases (all 1, 2, and 4); however, they do not perform well for the case of all 11 pipes with unknown parameters. The reason for the poor convergence is likely the different sensitivities of the parameters, as shown in the case studies of Section 4.1, as well as the increased number of unknowns. Certainly, the friction factor is always determined the least accurately due to its lower sensitivity.

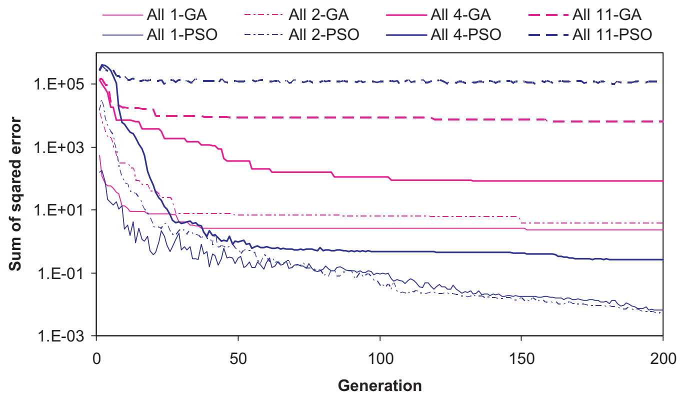

Figure 6 Evolutionary procedure of all parameters.

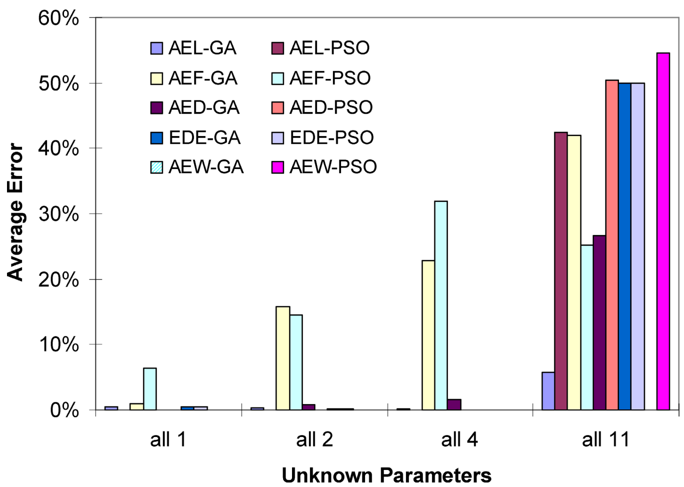

Figure 7 Average errors of different number of parameters. AED, average error of diameter; EDE, error of demand; AEW, average error of wave speed.

5 Conclusion

Although ITA is a mathematically elegant method of finding system parameters using a system's responses, many challenging problems remain as the full complexity of a field application is approached. Complexity and uncertainty in water distribution systems substantially deteriorates the conventional ITA of leakage and friction factor. These difficulties in truly complex problems, at least partly limit ITA to numerical and laboratorial experiments or to well-defined field system. This paper investigates uncertainties in water distribution system and quantitatively discovers how the uncertainties influence system calibration. Pipe diameter, water demand, and wave speed are specifically considered here and a systematic calibration approach for including these additional uncertainties during the ITA process is presented.

As optimization methods, GA and PSO are applied to minimize the difference between the measured and the predicted heads during the inverse calculation. In the case studies with some incorrect information, both optimizations tend to provide relatively accurate calibration for leakages but poorer values for friction factors. The mismatch in sensitivities creates a difficult inverse problem for low sensitive parameters, especially in the systematic calibration that includes many uncertainties. Hence, these varying sensitivities should receive special attention in the future along with the many other practical challenges of ITA.

References

1. Babbitt, H.E. (1920). "The Detection of Leaks in Underground Pipes". *J. AWWA* 7, 589–595.
2. Back, T., Fogel, D.B., Michalewicz, Z. and Baeck, T. (1997). *Handbook of Evolutionary Computation*. Institute of Physics Publishing, Bristol, Philadelphia.
3. Boulos, P.F. and Wood, D.J. (1990). "Explicit Calculation of Pipe-Network Parameters". *J. Hydraul. Eng.* 116(11), 1329–1344.
4. Brunone, B. and Ferrante, M. (2001). "Detecting Leaks in Pressurised Pipes by Means of Transients". *J. Hydraul. Res.* 39(5), 539–548.
5. Brunone, B., Golia, U.M. and Greco, M. (1991). "Some Remarks on the Momentum Equations for Fast Transients". *International Meeting on Hydraulic Transients with Column Separation*, 9th Round Table, IAHR, Valencia, Spain, pp. 201–209.
6. Chen, L.C. (1995). "Pipe Network Transient Analysis — The Forward and Inverse Problems". PhD Thesis, Cornell University, Ithaca, NY, USA.
7. Covas, D., Stoianov, I., Ramos, H., Graham, N., Maksimovic, C. and Butler, D. (2004). "Water Hammer in Pressurized Polyethylene Pipes: Conceptual Model and Experimental Analysis". *Urb. Wat. J.* 1(2), 177–197.
8. Dandy, G.C., Simpson, A.R. and Murphy, L.J. (1996). "An Improved Genetic Algorithm for Pipe Network Optimization". *Wat. Resour. Res.* 32(2), 449–458.

9. FILION, Y.R. and KARNEY, B.W. (2003). "Sources of Error in Network Modeling: A Question of Perspective". *J. AWWA* 95(2), 119–130.
10. GOLDBERG, D.E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Reading, Mass.
11. JUNG, B.S. (2005). "Optimization and Calibration of Water Distribution Systems: Exploring Evolutionary Approaches While Accounting for Fluid Transient". PhD Thesis, University of Toronto, Toronto, ON, Canada.
12. JUNG, B.S. and KARNEY, B.W. (2004a). "*Particle Swarm Optimization Compared to a Genetic Algorithm for Calibration of Water Distribution System*". *Proceeding of 9th International Conference on Pressure Surges*, Chester, UK, March.
13. JUNG, B.S. and KARNEY, B.W. (2004b). "Fluid Transients and Pipeline Optimization Using GA and PSO: The Diameter Connection". *Urb. Wat. J.* 1(2), 167–176.
14. JUNG, B.S. and KARNEY, B.W. (2006). "Optimization of Transient Protection Devices Using GA and PSO Approaches". *J. Wat. Resour. Plan. Man. ASCE* 132(1), 44–52.
15. JUNG, B.S., KARNEY, B.W. and LAMBERT, M.F. (2006). "Benchmark Tests of Evolutionary Algorithms: Mathematic Evaluation and Application to Water Distribution Systems". *J. Env. Info.* 7 (1), 26–37.
16. KAPELAN, Z. and SAVIC, D.A. (2003). "A Hybrid Inverse Transient Model for Leakage Detection and Roughness Calibration in Pipe Networks". *J. Hydraul. Res.* 41(5), 481–492.
17. KAPELAN, Z., SAVIC, D.A. and WALTER, G.A. (2005). "Incorporation of Prior Information on Parameters in Inverse Transient Analysis for Leak Detection and Roughness Calibration". *Urb. Wat. J.* 1(2), 129–143.
18. KENNEDY, J. and EBERHART, R.C. (1995). "*Particle Swarm Optimization*". Proc. IEEE Int. Conf. Neur. Net. IV, Piscataway, New Jersey, 1942–1948.
19. KENNEDY, J. and EBERHART, R.C. (2001). *Swarm Intelligence*. Morgan Kaufmann Publishers, San Francisco, California.
20. KIM, S.H. (2005). "Extensive Development of Leak Detection Algorithm by Impulse Response Method". *J. Hydraul. Eng.* 131(3), 201–208.
21. LANSEY, K.E. and BASNET, C. (1991). "Parameter Estimation for Water Distribution Networks". *J. Wat. Resour. Plan. Man.* 117(1), 126–114.
22. LIGGETT, J.A. and CHEN, L.C. (1994). "Inverse Transient Analysis in Pipe Networks". *J. Hydraul. Eng.* 120(8), 934–955.
23. LIOU, C.P. (1998). "Pipeline Leak Detection by Impulse Response Extraction". *J. Fluid. Eng. ASME* 120, 833–838.
24. MCINNIS, D. and KARNEY, B.W. (1995). "Transients in Distribution Networks: Field Tests and Demand Models". *J. Hydraul. Eng.* 121(3), 218–231.
25. NASH, G.A. and KARNEY, B.W. (1999). "Efficient Inverse Transient Analysis in Series Pipe Systems". *J. Hydraul. Eng.* 125(7), 761–764.
26. PUDAR, R.S. and LIGGETT, J.A. (1992). "Leaks in Pipe Networks". *J. Hydraul. Eng.* 118(7), 1031–1046.
27. SHI, Y.H. and EBERHART, R.C. (1998). "A Modified Particle Swarm Optimizer". *IEEE International Conference on Evolutionary Computation*, Piscataway, New York, pp. 69–73.
28. SHI, Y.H. and EBERHART, R.C. (1999). "*Empirical Study of Particle Swarm Optimization*". Proceedings of the 1999 Congress on Evolutionary Computation, Piscataway, New Jersey, pp. 1945–1950.
29. SIMPSON, A.R., DANDY, G.C. and MURPHY, L.J. (1994). "Genetic Algorithms Compared to Other Techniques for Pipe Optimization". *J. Wat. Resour. Plan. Man.* 120(4), 423–443.
30. STOIANOV, I., KARNEY, B., COVAS, D., MAKSIMOVIC, C. and GRAHAM, N. (2002). "*Wavelet Processing of Transient Signals for Pipeline Leak Location and Quantification*". Proceedings of the 1st Annual Environmental and Water Resources Systems Analysis (EWRSA) Symposium, Roanoke, Virginia, USA.
31. VITKOVSKY, J.P. (2001). "Inverse Analysis and Modeling of Unsteady Pipe Flow: Theory, Applications and Experimental Verification". PhD Thesis, University of Adelaide, Adelaide, SA, Australia.
32. VITKOVSKY, J.P., SIMPSON, A.R. and LAMBERT, M.F. (2000). "Leak Detection and Calibration Using Transients and Genetic Algorithms". *J. Wat. Resour. Plan. Man.* 126(4), 262–265.
33. VITKOVSKY, J.P., LIGGETT, J.A., SIMPSON, A.R. and LAMBERT, M.F. (2003). "Optimal Measurement Site Locations for Inverse Transient Analysis in Pipe Networks". *J. Wat. Resour. Plan. Man.* 129(6), 480–492.
34. WALSKI, T.M., CHASE, D.V. and SAVIC, D.A. (2001). *Water Distribution Modeling*. Haestad Press, Waterbury, CT, USA.
35. WYLIE, E.B. and STREETER, V.L. (1993). *Fluid Transients in Systems*. Englewood Cliffs, New Jersey.