Vitkovsky, John; Simpson, Angus Ross; Lambert, Martin Francis
Leak detection and calibration of water distribution systems using transients and genetic algorithms 29th Annual Water Resources Planning and Management Conference: preparing for the 21st century, 6 June - 9 June 1999, Tempe, Arizona, United States/ Erin M. Wilson (ed.): 9 p

Leak Detection and Calibration of Water Distribution Systems using Transients and Genetic Algorithms

by

John P. Vítkovský, Angus R. Simpson and Martin F. Lambert

Department of Civil and Environmental Engineering
The University of Adelaide
Australia 5005

asimpson@civeng.adelaide.edu.au

Citation: Vítkovský, J.P., Simpson, A.R. and Lambert, M.F. (1999). "Leak detection and calibration of water distribution systems using transients and genetic algorithms." *Water Distribution Systems Conference,* Division of Water Resources Planning and Management, American Society of Civil Engineers, Tempe, Arizona, 7-9 June.

Leak Detection and Calibration of Water Distribution Systems using Transients and Genetic Algorithms

John P. Vítkovský, Angus R. Simpson and Martin F. Lambert

Abstract

The use of genetic algorithm optimisation applied to solving engineering problems has gained popularity over the last 10 years. Applications to the design of water distribution systems based on genetic algorithm optimisation first appeared in the early 1990s. This paper starts out with a brief review of the past use of genetic algorithms applied to aspects of water distribution systems. Leak detection and calibration of pipe internal roughnesses in a network are important issues for water authorities around the world. Computer simulation of water distribution systems has become a routine task of water authorities and consultants. One of the big unknowns in developing these models is the condition of the pipes, especially if they are old. It is very difficult to obtain reliable estimates of the roughness height for each pipe in the system using steady state calibration techniques. Liggett and Chen at Cornell University in 1994 developed an innovative technique called the inverse transient technique. The technique is able to determine, from unsteady pressure traces at a number of nodes in the network, the locations and magnitudes of any leaks that are occurring and the friction factor for each pipe in the network. An alternative approach to solving the minimization problem is presented in this paper. Genetic algorithm optimisation is used. A population of solutions is generated with each string representing values of the decision variables that are to be found. These include the magnitudes of leaks at nodes in the network and friction factors for each pipe. A forward transient analysis is performed for each string in the population that represents different combinations of leak magnitudes and friction factors. The sum of the absolute deviations between the measured transient pressures and the pressures predicted by the numerical model are determined and are used to determine the fitness of the string. The smaller the sum of the deviations then the larger the fitness that is assigned to the string. The genetic algorithm operators that are used include tournament selection, crossover and mutation. A new crossover operator is introduced. The genetic algorithm optimisation technique that has been developed in the research is applied to an example network. The results are encouraging and compare favorably with the inverse transient technique.

Introduction

The least accurately known parameters in water distribution system are the pipe internal roughnesses or Darcy-Weisbach friction factors (Task Committee on Water Supply Rehabilitation Systems, 1987). Over time, the friction factor of a pipe changes due to tuberculation (the build up of deposits on the pipe wall). This can cause friction factor values to increase with age. Different pipes in a network are subject to different conditions including variations in dissolved solids loadings, flow, pressure and temperature. Thus, in an aging network, reliable estimates of friction factors can be difficult to obtain. Previous methods used for network calibration have generally employed steady state analysis as a basis for network simulations (Rahal *et al.*, 1980; Walski, 1983; Savic and Walters, 1995). As a general observation results were poor using a single steady state case, and better when using extended period analysis. A lot of field tests made use of fire flows for steady state case to generate data.

A new direction is possible if an unsteady state or transient model is utilized. Pudar and Liggett (1992) introduced inverse steady state analysis that used sets of measured steady state pressure data at different nodal positions to both calibrate pipe roughnesses and locate leaks in pipe networks. A potential problem predicted by Pudar and Liggett was the applicability of their solution method to very large water distribution networks. Liggett and Chen (1994) posed a solution to this potential shortcoming that was to use analysis of transient events rather than steady state calibration. The minimization method used to fit the measured data to the model by Liggett and Chen (1994) was an adjoint Levenberg-Marquardt minimization method. The search space for the problem of minimization of deviations between measured and model predicted pressures is enormous. An alternative method based on the genetic algorithm (GA) technique is presented in this paper.

Genetic Algorithm Optimisation

Genetic algorithms mimic the way populations of species genetically evolve to suit their environment over many generations. Using this analogy a process can be used to evolve a population of potential solutions representing engineering design problems towards improved solutions. These solutions will satisfy the specified constraints while minimizing or maximizing one or more objective functions. There are six main steps to the implementation of a genetic algorithm for application to the network calibration problem. These steps are based upon the steps described by Simpson and Goldberg (1994). The steps are as follows:

1. Encoding of a chromosome to represent the decision variables to be optimised.
2. Generation of an initial population.
3. Analysis of each chromosome to assess the performance of each member of the population.
4. Computation of the fitness of a chromosome.
5. Generation of a new population using genetic operators.

6. Production of successive generations.

The problem of fitting a numerically modelled pressure or hydraulic grade line (HGL) trace to a measured HGL trace may be solved by minimizing the sum of the absolute differences. The objective of the GA is usually formulated to maximize the fitness of a string. This may be achieved in Eq. 1 by maximizing the negative of the sum of absolute differences , where M is the total number of data points and H_i^* and H_i are the measured and modelled HGLs respectively.

$$fitness = -\sum_{i=1}^{M} \left| H_i^* - H_i \right| \qquad (1)$$

The variables to be solved for are the Darcy-Weisbach pipe friction factors (f) and lumped leak coefficients. Both of these types of decision variables may take on any value within a range of continuous values. Representation of decision variables within a GA string is usually by discrete values the continuous range.

Leakage from a pipe in a water distribution system is simulated using an orifice equation as shown in Eq. 2. This equation may be applied by accounting for continuity at nodal positions.

$$Q_L = C_d A_0 \sqrt{2gH_L} \qquad (2)$$

where Q_L = leak discharge, C_d = orifice discharge coefficient, A_o = area of leak orifice, g = gravitational acceleration and H_L = head difference across the leak.

To exploit the continuous nature of the variables in the string, a new crossover operator is introduced. This operator is based upon a similar crossover operator first used by Savic and Walters (1995) named one child average crossover. This Savic and Walters operator exhibited quick convergence that produced fast stagnation of the population. Traditional crossover operators produce two child strings rather than the one produced by Savic and Walters average crossover. A new operator has been created called two child staggered average crossover (Vítkovský and Simpson, 1997). The schematic of this operator with three crossover points is shown in Figure 1.

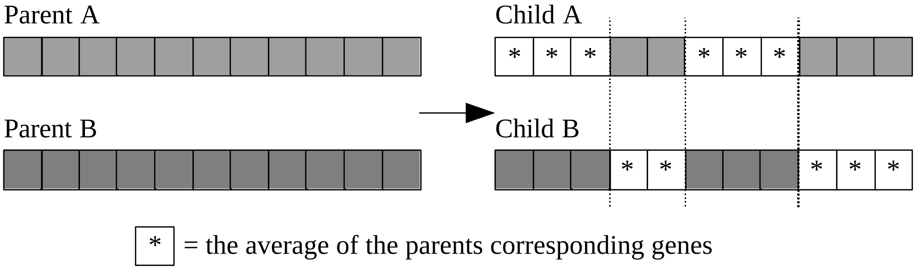

Figure 1. Two Child Staggered Average Crossover Example

In addition to the modification of the crossover operator, the mutation and initial population creation have also been adapted for continuous variables. Eq. 3

demonstrates how a gene in a string can be mutated. In this case the operator is a continuous adjacency mutation where *RND* is a uniformly distributed random number between 0.0 and 1.0, and *Step_Size* is the maximum increment allowed in a gene.

$$gene = gene + Step\_Size(2 \cdot RND - 1) \quad \ldots\ldots\ldots\ldots (3)$$

To demonstrate how continuous adjacency mutation is applied an example is shown in Figure 2. The possible range for a new gene is one *Step_Size* located either side of the original gene value. Given the operation of this type of mutation, different *Step_Size* values must be given for different types of parameters. For example the *Step_Size* for a friction factor is orders of magnitude larger than that for a lumped leak coefficient.

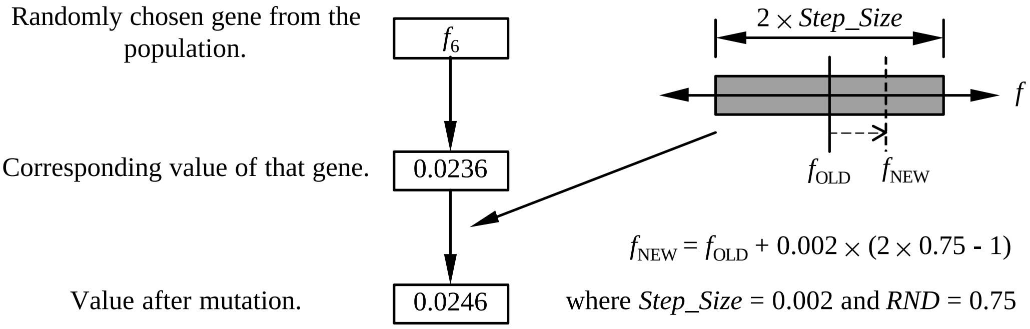

Figure 2. Continuous Adjacency Mutation Example

Existing code written at the University of Adelaide (Anderson and Simpson, 1996) has been modified to accommodate these new operators.

Results on the Test Network

The Test Network (Figure 3) considered in this paper consists of 11 pipes and 7 nodes and is fed from a reservoir and a constant inflow at node 6 (based on a network introduced by Pudar and Liggett, 1992). There is a variable demand at node 4. This network allows a number of different events to be analysed and multiple leakage points to be chosen. The transient is initiated at node 7 by reducing the flow demand over a small amount of time (much like a valve would).

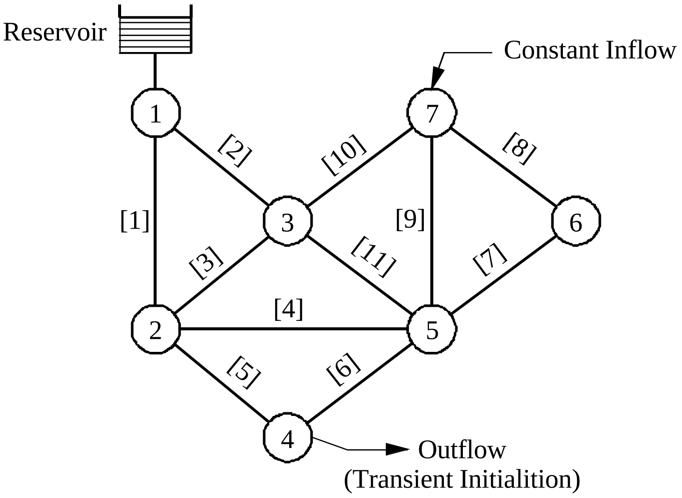

Figure 3. Test Network

The forward transient pressure variation data used in this example includes the hydraulic grade line at nodes 2, 3, and 6 over a 40 second interval (corresponding to 72 data points per node). There is a simulated leak at node 2 which has a lumped leak coefficient, C_dA_0, of 1.0×10^{-4}, that corresponds to an approximate leak hole diameter of 13mm (the pipe being 254mm). The assumed Darcy-Weisbach friction factors for the network are shown in the second column of Table 1. Chromosomes have been encoded for this Test Network such that the first set of genes in the chromosome correspond to the friction factors and the second set to the leak candidates (Figure 4).

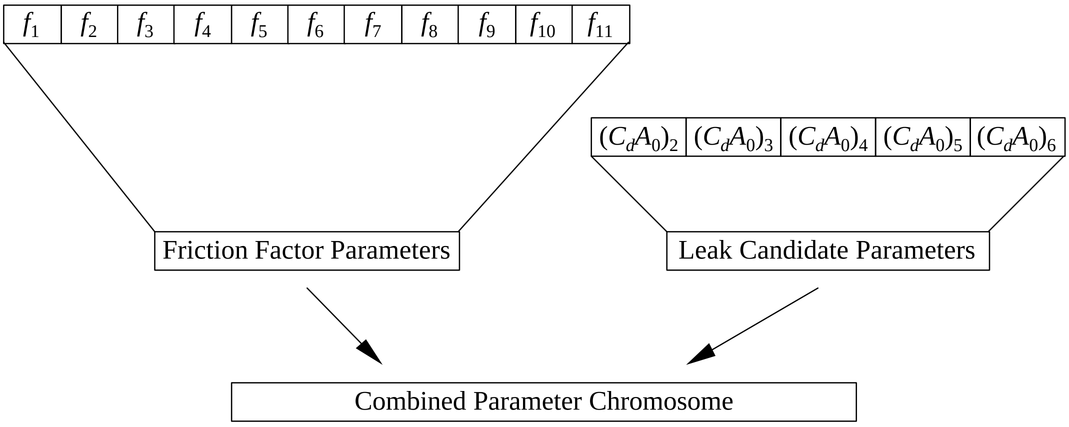

Figure 4. Chromosome Representation for Test Network

A typical result from an application of the GA calibration and leak detection is shown in Tables 1 and 2 that correspond to friction factors and leaks respectively. An average error 3.43% is found in the friction factors and an error of 0.50% is found for the leak at node 2. The size of the calibrated leaks in the nodes where the were no real leaks were approximately three orders of magnitude smaller than that of the real leak.

Their effect on the transient model were negligible and correspond to a hole in the pipe of about the size of a pinprick. These can then be discounted as being too small.

Table 1. GA Friction Factor Calibration Results

Pipe No.	Darcy-Weisbach Friction Factor, f		
	Correct Results	Typical GA Run	Percentage Error
1	0.040	0.0410	2.50%
2	0.040	0.0393	-1.84%
3	0.040	0.0388	-2.89%
4	0.025	0.0253	1.20%
5	0.025	0.0259	3.60%
6	0.025	0.0247	1.34%
7	0.030	0.0283	-5.52%
8	0.030	0.0268	-10.6%
9	0.030	0.0300	0.00%
10	0.020	0.0215	7.61%
11	0.020	0.0199	0.67%
Overall Average Absolute Error			3.43%

Table 2. GA Leak Detection Results

Leak Candidate	Lumped Leak Coefficient, C_dA_0		
	Correct Results	Typical GA Run	Percent Error
Node 2	1.0×10^{-4}	9.95×10^{-5}	-0.50%
Node 3	0.0	1.19×10^{-7}	/
Node 4	0.0	1.33×10^{-7}	/
Node 5	0.0	1.89×10^{-7}	/
Node 6	0.0	1.83×10^{-7}	/

There is large variation in the friction factors observed. This can be explained by checking the sensitivity of the fitness function with respect to the parameters. The friction factors exhibited a very low sensitivity compared to those of the lumped leak coefficients. This suggests that leaks may be found more easily than the friction factors. The sensitivity for each parameter can be calculated by finding the partial derivative of the fitness with respect to that parameter. The magnitudes (rather than the signs) of these values can then be used to show the relative ease with which a parameter can be found compared to another parameter. A large sensitivity in a parameter corresponds to a greater confidence in a calibrated solution for that parameter. The sensitivities for the friction factors and the lumped leak parameters are shown in Figures 5 and 6 respectively. The sensitivity for the lumped leak parameters are approximately 3 orders of magnitude larger than the friction factor sensitivities. This reflects the original observation that the lumped leak coefficients had less error than the friction factors.

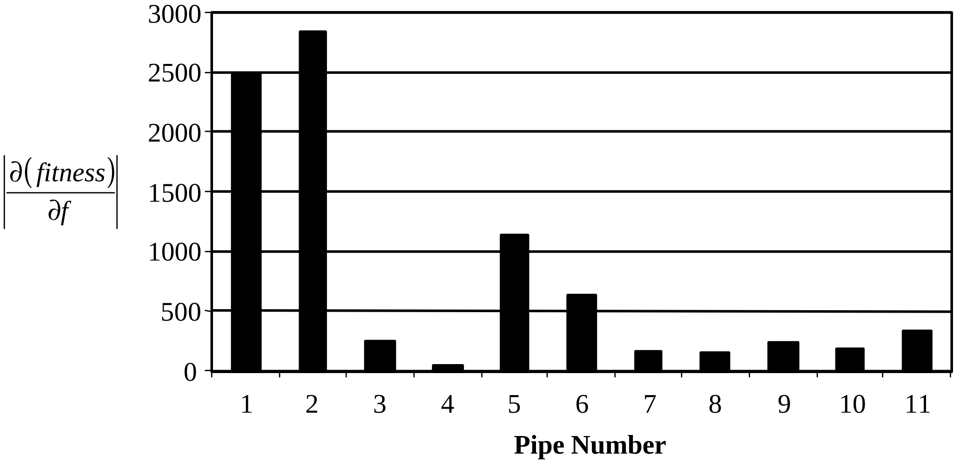

Figure 5. Sensitivity of Fitness with Respect to Friction Factors

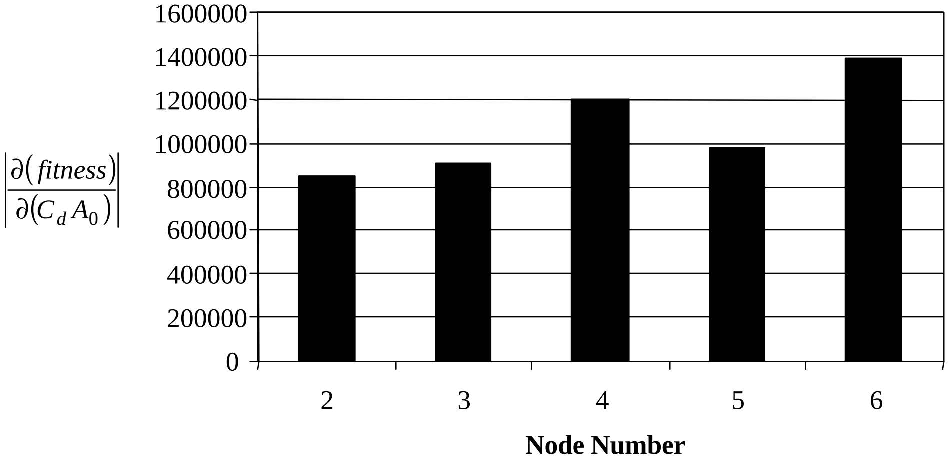

Figure 6. Sensitivity of Fitness with Respect to Lumped Leak Coefficients

A plot of the convergence of fitness for this run is shown in Figure 7. The convergence is understandably fast at the beginning of the GA run but is slower towards the end of the run. A reason for this is the low sensitivity of friction factors such as those for pipes 4, 7, 8 and 10. Any changes in these parameters do not tend to have a great affect on the fitness of a string. Thus many strings may have similar fitnesses while having differing friction factors for the low sensitivity pipes. This means that some parameters are less important when modeling and should have a lower associated confidence.

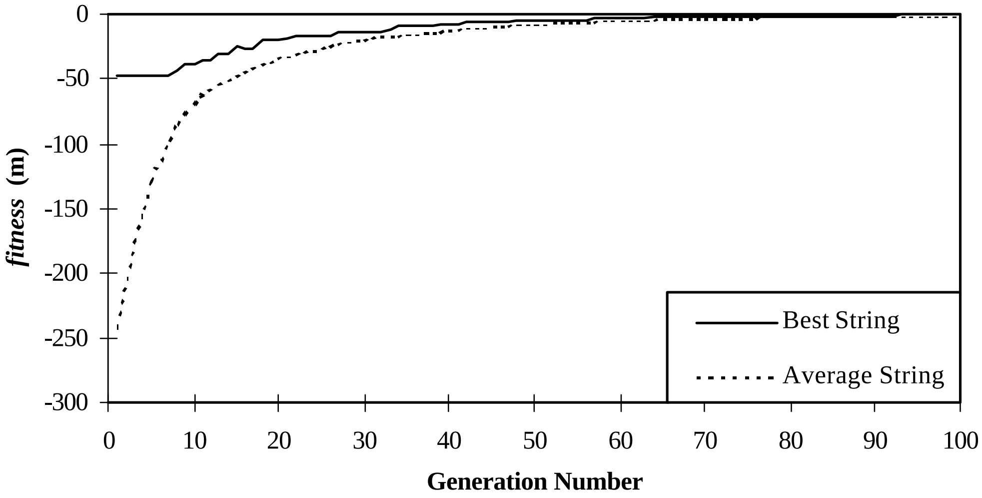

Figure 7. Convergence of Genetic Algorithm

Conclusions

Satisfactory results have been gained for both leak detection and friction factor system calibration for the Test Network based on the genetic algorithm technique. These results highlight the usefulness and potential of this technique for application on larger networks. A new crossover operator has been introduced in this paper that takes advantage of the continuous nature of the variables being represented in the genetic algorithm string.

Future research will compare for large networks the effectiveness of the GA calibration solution technique versus the inverse transient technique of Liggett and Chen (1994). Analytical searching methods like the Newton-Raphson and Levenberg-Marquardt methods used in the inverse transient method can fail to converge or converge to a local minimum rather than the global minimum. This has been observed with the Test Network when searching for a large number of unknowns. There is a trade-off between the inverse transient method (based on analytical techniques) fast speed (and possibility of non-convergence and finding a local minima) and the GAs comprehensive search (guaranteed convergence to a solution, although not necessarily the global minimum, but slow speed).

The focus of on-going research at the University of Adelaide is to further develop both the inverse transient technique and the genetic algorithm technique that uses transient data to calibrate and detect leaks in a network. Experimental investigations are also being carried out. The aim is for these techniques to be eventually employed with a continuous data acquisition system in water distribution systems in the field. The technique will operate giving a regular update of the state of health of a network. The deterioration of pipes (indicated by high friction cases) can be detected and also the water authority then deal with existence of any leaks.

References

Anderson, A., and Simpson, A.R. (1996). "Genetic Algorithm Optimisation Software in Fortran." *Research Report No. R136*, March, Department of Civil and Environmental Engineering, The University of Adelaide, Australia.

Liggett, J.A., and Chen L. (1994). "Inverse Transient Analysis in Pipe Networks." *Journal of Hydraulic Engineering*, ASCE, 120(8), 934-955.

Pudar, R.S., and Liggett, J.A. (1992). "Leaks in Pipe Networks." *Journal of Hydraulic Engineering*, ASCE, 118(7), 1031-1046.

Rahal, C.M., Sterling, M.J.H., and Coulbeck, B. (1980). "Parameter Tuning for Simulation Models of Water Distribution Networks." *Institution of Civil Engineers*, Proceedings, Part 2, Vol. 69, 751-762.

Savic, D.A., and Walters, G.A. (1995). "Genetic Algorithm Techniques for Calibrating Network Models." Report No. 95/12, Centre for Systems and Control Engineering, University of Exeter, Devon, UK.

Simpson, A.R., and Goldberg, D.E. (1994). "Pipeline Optimisation via Genetic Algorithms: from Theory to Practice." *2nd International Conference on Water Pipeline Systems*, Proceedings, Edinburgh, Scotland, May, 309-320.

Task Committee on Water Supply Rehabilitation Systems (1987). *Water Supply System Rehabilitation*. T.M. Walski (Editor), ASCE.

Vítkovský, J.P., and Simpson, A.R. (1997). "Calibration and Leak Detection in Pipe Networks Using Inverse Transient Analysis and Genetic Algorithms." *Research Report No. R157*, August, Department of Civil and Environmental Engineering, The University of Adelaide, Australia.

Walski, T.M. (1983). "Technique for Calibrating Network Models." *Journal of Water Resources Planning and Management*, ASCE, 109(4), 360-372.

About the Authors

John P. Vítkovský, BE (Hons), Postgraduate Student, Department of Civil & Environmental Engineering, The University of Adelaide, Adelaide, SA 5005, Australia, Phone: +61-8-8303-4323, Fax: +61-8-8303-4359, email: jvitkovs@civeng.adelaide.edu.au

Angus R. Simpson, M. ASCE, Senior Lecturer, Department of Civil & Environmental Engineering, The University of Adelaide, Adelaide, SA 5005, Australia, Phone: +61-8-8303-5874, Fax: +61-8-8303-4359, email: asimpson@civeng.adelaide.edu.au

Martin F. Lambert, Senior Lecturer, Department of Civil & Environmental Engineering, The University of Adelaide, Adelaide, SA 5005, Australia, Phone: +61-8-8303-5838, Fax: +61-8-8303-4359, email: mlambert@civeng.adelaide.edu.au