Genetic algorithms for leak detection in water supply networks

S. Mambretti & E. Orsi
DIIAR, Politecnico di Milano, Italy

Abstract

In many countries water losses can be much higher than 50%, with great economic losses because of the energy required for pumping and for the primary treatments that are very often required.

Very often losses can be detected only when water rises in the streets, or during an expensive campaign, which are very useful to punctually find existing losses, but with the disadvantage that the actual situation of the network remains unknown.

To this end, a methodology has been developed which requires the installation of a number of instruments on the network, in order to measure pressures and discharges; then, the demands at the nodes are changed by means of a genetic algorithm and the network is simulated with a computer program, in order to match the readings of the instruments; losses are higher where the demands have been most increased.

In this paper, the general methodology is presented, and the application on a case study, performed in a comparatively small town (Castegnato), in the north of Italy, with 8000 inhabitants and 9 km^2 of extension, and where the modelled network length is equal to 43.3 km and water losses exceed 60% of the total daily volume.

The preliminary simulations show that the methodology is able to detect the different hypothesized scenarios and has been of great support in the decision about the number, typology and position of the devices to be installed. Moreover, it is shown that this method allows the detection of the areas where the pipe conditions are most critical, therefore providing a new indicator for deciding, in case of pipe rupture, whether the conduit is to be repaired or replaced.
Keywords: water supply networks, water losses, genetic algorithms, hydraulic modelling.

WIT Transactions on The Built Environment, Vol 122,
www.witpress.com, ISSN 1743-3509 (on-line)
doi:10.2495/UW120061

1 Introduction

Water distribution networks, as a primary part of water supply systems, represent one of the main infrastructure assets of society, and therefore an effective, efficient and energy saving management is necessary in order to optimize their performances. Water losses from water distribution systems are often the main problem for water utility companies. In some countries, and among them Italy, there are regions where leakages (or, more precisely, Non Revenue Water) are up to 70% of the water volume supplied to the network. This volume is a serious economic damage for the companies, because of the costs of pumping and treatment, the need to invest in the system to increase its capacities and to find new sources, while they diminish because of the pollution and the increase in demand.

In particular, managers need to know where to operate, and how (repair or substitution of a pipe, for instance) [1]. This issue is usually dealt with as a multi-objective optimization problem, where the objective function is represented by the performance of the network and the costs of the rehabilitation [2, 3].

Common objective functions are the volumes of undelivered water or the numbers of customers affected by interruptions due to pipes bursts [4]. Such a condition inspired the authors to develop models able to generate pipe breaks [1] or have available a good database about previous breakages [5, 6].

Another objective to be pursued is the increase of the network efficiency through the reduction of water losses. The limited funds available constrain the invested annual budget and increase the importance of interventions time scheduling.

In this paper, a new methodology is developed to identify the areas where losses are most expected, by means of data collection (discharge and pressure) from instruments positioned on the water supply network, and successive comparison of the data collected with those simulated by means of a computer program. The results of the model should match the readings of the instruments. Under the hypothesis, the model is a good representation of the real network, the differences between simulated and recorded data are due to the different demands imposed at the nodes: in the model, the volume of water losses are not inserted, because they are not known; in the actual network, instead, they are present.

The algorithm is developed in order to find the best combination of the discharges outflowing from the network, and imposed as “demand” in each node, that allow the modelled values to be closest to those recorded. This best combination of demands is carried out by means of an improved genetic algorithm, as described as follows.

To the simulated water supply network, a number of scenarios are imposed; then, the genetic algorithm is run in order to identify the areas. It is found that the areas where the highest losses are imposed are properly found; this allowed us to confirm that the positions of the devices which have been identified within the network are correct and in a following phase of the research, the instruments

www.witpress.com, ISSN 1743-3509 (on-line)

are actually positioned; when real data are available, the procedure might be improved.

2 Genetic algorithms

Complex and multi-objective optimization problems are often solved by means of Evolutionary Computation. The term Evolutionary Computation (EC) [7] represents a large spectrum of heuristic approaches to simulate evolution, which include: genetic algorithms (GA) [8, 9], Evolutionary Strategies [10, 11], Evolutionary Programming [12] and Genetic Programming [13]. GAs are one of the most known EC, because of their central use in water resources planning and management. They have been largely used in the last two decades in order to improve the efficiency of water distribution systems because traditional calculus-based and enumerative optimization methods were not able to cope with the geometrical complexity of these systems (a complex network of pipes, junctions and hydraulic control elements) and the request to respond to the water demand of customers with required quality and in a reliable cost-effective and sustainable way.

In this paper, the objective function is set as:

$$min\left(\sum_{i=1}^{Number\ of\ Control\ Nodes}\frac{\left|H_{meas}-H_{comp}\right|}{\left|H_{meas}\right|}\cdot W_H + \sum_{i=1}^{Number\ of\ Control\ Links}\frac{\left|Q_{meas}-Q_{comp}\right|}{\left|Q_{meas}\right|}\cdot W_Q + \frac{\left|Q_{Globally\ Expected}-Q_{Globally\ Computed}\right|}{\left|Q_{Globally\ Expected}\right|}\cdot W_{GE}\right) \tag{1}$$

where H is the head, Q is the discharge and W the weights to be used to assess the different importance to the parameters to be used when actual measurements are carried out and the expected precision of the instruments is different. Equation (1) is the general objective function, while in this paper the weights W are set equal to one because the scenarios are still being computed and have not measured yet, and therefore all the values have the same level of accuracy.

The algorithm is set as follows:

1. The discharge at nodes are turned into binary numbers with eleven figures in order to store discharges from 0.000 l/s (binary: 00000000000) to 2.047 l/s (binary: 11111111111);
2. The initial population is randomly generated;
3. The objective function is evaluated for each individual of the population;
4. (a) For ten generations, crossover is used to generate the next population, with a number of the best individuals being kept without changing (elitism);

(b) The eleventh population is generated with a Monte Carlo method from only one parent, accepting the new individuals only if their objective function values are better than those of their parents;

5. Points three and four are repeated until the imposed number of iteration is reached.

3 Case study

The case study is the water supply network of Castegnato, a small town in the north of Italy with around 7900 inhabitants and which network is divided into two disconnected parts. In one of the two sub-networks, water is pumped from two wells called "Franchi" and "Coronino" (and therefore the network will be called "Franchi-Coronino"); in the other sub-network, water is pumped from a well called "Cavour" (and therefore the network will be called "Cavour"). All pumps convey water directly in the network, with head variable within the range 30–50 meters.

With regards to the pumps, Cavour has a discharge of 20 l/s; Coronino is around 30 l/s and Franchi around 20 l/s.

The pipes in the network have a total length of 45 km and materials are steel and plastic.

In figure 1, the water supply network for Castegnato is shown and highlighted is the division in the two sub-networks "Cavour" and "Franchi-Coronino".

During the year 2005, the pumped volume was equal to 1˙800˙171 m^3; the invoiced volume was equal to 751˙866 m^3 and therefore the losses were equal to 58%. During the year 2006, 39 repairs have been performed on the network, and for that year the pumped volume was equal to 1˙625˙070 m^3; the invoiced volume was equal to 810˙000 m^3 with 50% of losses. Losses are decreased, but their value is still high.

Repairs are performed only when the losses are visible, i.e. when water flows on the streets. There is therefore the need to identify losses even when they are not visible, as these are the most frequent cases.

To this end, a number of devices to measure pressures or discharges are to be positioned in the network: 10 pressure transducers and 2 flow meters for the Cavour network, which is smaller, and 15 pressure transducer and 3 flow meters for the Franchi-Coronino network.

In order to understand whether their number and position are appropriate, different scenarios have been simulated to check whether they can be reconstructed by the improved genetic algorithm described in section 2.

4 Scenarios

Four scenarios have been set for each of the two sub-networks, imposing the losses in different areas of the town. In figure 2 one of those scenarios is reported and in figure 3 the reconstruction carried out with the genetic algorithm is shown.

www.witpress.com, ISSN 1743-3509 (on-line)

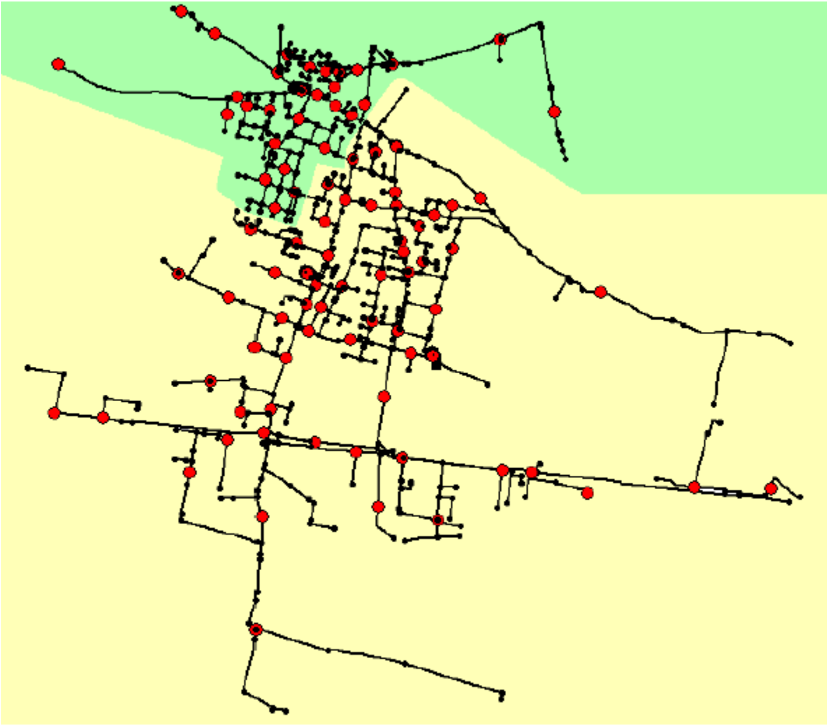

Figure 1: Water supply network in Castegnato. The northern network is “Cavour”; southern is “Franchi-Coronino”.

As can be seen (figures 2 to 4) the areas where losses are hypothesized have been identified. However, there is the need for a numeric indicator of the quality of the result. To this end, the difference between the target and the computed losses to each node is not a good indicator of the performance, as this method does not allow a “perfect” identification of the position of the losses.

Therefore, it has been decided to divide the town into subareas and to compute the differences of discharges within them.

Dividing the area into sub-areas with dimensions approximately equal to 150 x 150 m^2, the losses computed for the scenario shown in figure 2 are reported in table 1. In the table, only the subareas where losses are expected or detected are reported. As can be seen, the computer code found losses also in the subareas 4,7 (loss: 5.36 l/s) and 7.4 (loss: 2.48 l/s) which are quite far from the actual losses areas and, therefore, which identification is an error. Subareas numbers are imposed as shown in figure 4.

Figure 2: Franchi-Coronino sub-network in Castegnato: losses set up as target.

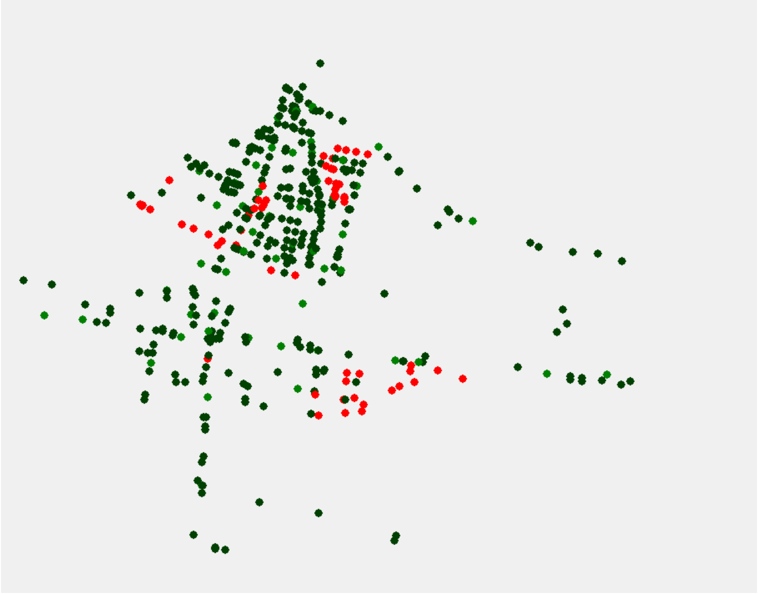

Figure 3: Franchi-Coronino sub-network in Castegnato: losses reconstructed with the model.

Table 1: Target and computed losses for the scenario shown in figure 2 of the Franchi-Coronino sub-network.

Subarea	Target [l/s]	Computed [l/s]
3,8	3.80	0.0
4,8	9.40	1.46
5,8	0.00	1.47
6,8	5.17	6.44
7,8	1.88	0.00
3,7	0.00	2.48
4,7	0.00	5.36
6,7	0.00	0.48
5,6	0.00	0.45
5,5	2.82	0.00
4,4	0.00	0.47
5,4	2.82	0.49
6,4	3.86	3.48
7,4	0.00	2.48
8,4	0.00	0.43

5 Conclusions and future development of the research

Water distribution networks are a primary part of water supply systems and represent a main infrastructure asset of the society, which need effective, efficient and energy saving management.

In some countries Non Revenue Water is up to 70% of the water volume supplied, which is a serious economic damage.

In this paper, a new methodology is developed to identify the areas where losses are most expected, by means of data collection (discharge and pressure) from instruments positioned on the water supply network, and successive comparison of the data collected with those simulated by means of a computer program.

To test the methodology, a case study has been identified, where devices are to be installed and data collected.

The position of the devices and the validity of the developed methodology are checked by means of a number of scenarios before their installation.

It is found that the areas where the highest losses are imposed are properly found; this allowed us to confirm that the positions of the devices which have been identified within the network are correct and in a following phase of the research the instruments are actually positioned; when real data are available, the procedure might be improved.

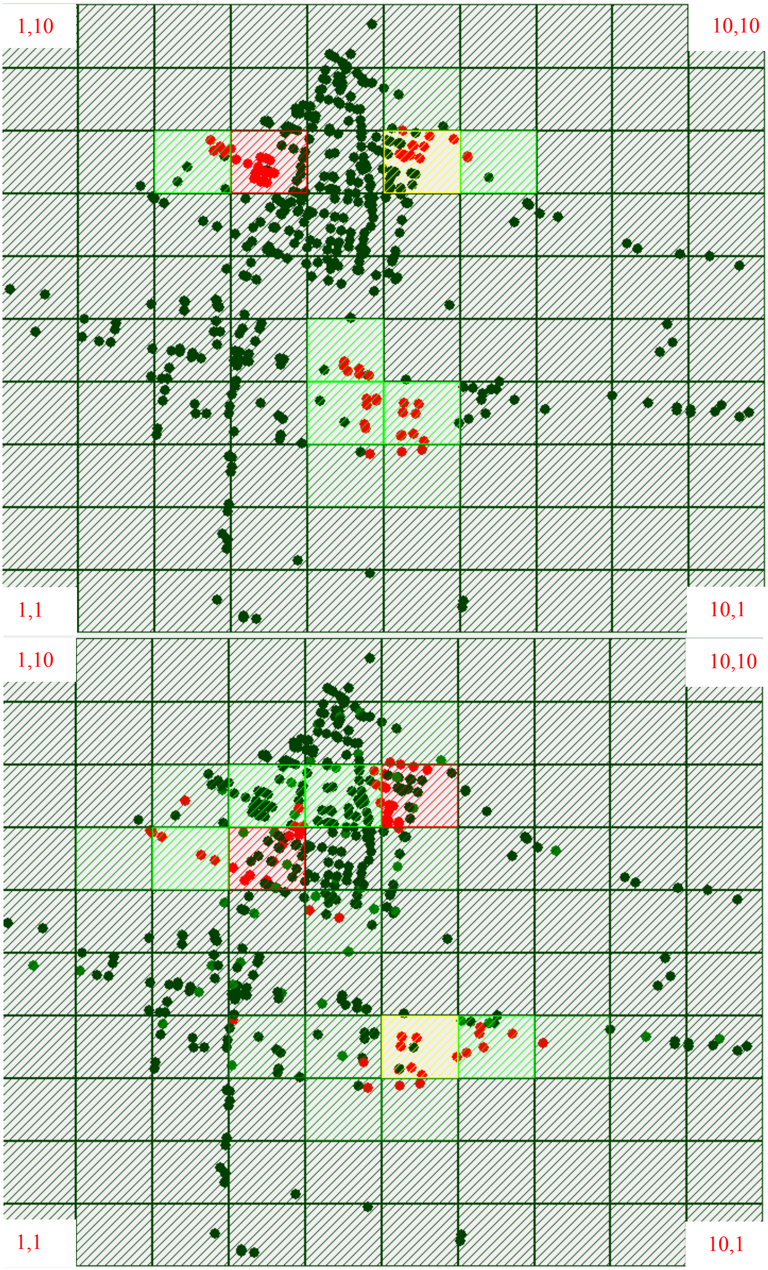

Figure 4: Matrix definition of the subareas in the network; comparison between target (above) and modeled (below).

The results carried out show that the scenarios are properly reconstructed, even if errors are obviously present; however, the goal to identify the areas where losses are concentrated seems to be reached.

Future developments will comprise the analysis of real data collected on the network, and the improvement of the computer program. Moreover, data will also be collected by means of a portable flow meter.

The final goal of this research is the development of a new methodology able not only to find the areas where losses are most expected, but also the improvement of the existing indicators of water supply management.

References

[1] Alvisi S., Grata S., Franchini M. (2006) "Leakage detection planning in water distribution systems", in "Management of Water Networks" Proceedings of the Conference "Efficient Management of Water Networks. Design and Rehabilitation Techniques." Bertola and Franchini (Eds), Ferrara 2006, De Angeli Editore, Milano, Italy.

[2] Halhal, D., Walters, G. A., Savic, D. A., Ouazar, D. (1999) "Scheduling of water distribution system rehabilitation using structured messy genetic algorithms" *Evolutionary Computation*, Volume 7, No. 3; MIT Press.

[3] Giustolisi, O., Laucelli, D., Savic, D.A., (2006) "Development of rehabilitation plans for water mains replacement considering risk and cost-benefit assessment", *J. of Civil Engineering and Environmental Systems*, Taylor & Francis, UK, No.3, Vo.23, pp. 175-190.

[4] Engelhardt M.O., Skipworth P.J., Savic D.A., Saul A.J., Walters G.A. (2000) "Rehabilitation strategies for water distribution networks: a literature review with a UK perspective", *Urban Water*, 2, pp. 153-170.

[5] Male, J. W., Walski, T. M., and Slutsky, A. H. (1990) "Analyzing water main replacement policies." *J. Water Resour. Plan. Manage.*, 116(3), 362–374.

[6] Sundahl A., (1996) "Using break data on water pipe systems for renewal planning" COST Action C3 workshop, 18 and 19 June 1996, Brussels.

[7] Back T., Fogel D., Michalewicz Z. (1997) *Handbook of evolutionary computation* IOP Publishing Ltd. and Oxford University Press, New York and Oxford.

[8] Holland J. H. (1962) "Outline for a logical theory of adaptive systems" *Journal of the ACM*, Volume 9, Issue 3; ACM.

[9] Holland J. H. (1975) "Adaptation in natural and artificial systems" University of Michigan Press.

[10] Rechenberg I. (1973) *Evolutionsstrategie: Optimierung Technischer Systeme Nach Prinzipien der Biologischen Evolution* Fromann-Holzboog

[11] Schewefel H.P. (1981) *Numerical Optimization of Computer Models*, John Wiley and Sons, New York.

[12] Fogel L., Owens A., Walsh M. (1966) *Artificial Intelligence Through Simulated Evolution*, John Wiley & Sons, Inc, New York.

[13] Koza J. R. (1992) *Genetic Programming* MIT Press.