OPTIMAL ALLOCATION OF MONITORING STATIONS AIMING AT AN EARLY DETECTION OF INTENTIONAL CONTAMINATION OF WATER-SUPPLY SYSTEMS

Luca Cozzolino, Carmela Mucherino, Domenico Pianese, Francesco Pirozzi
University of Naples Federico II, Department of Hydraulic and Environmental Engineering Girolamo Ippolito
Via Claudio, 21 – 80125 - Napoli (Italy)
lcozzoli@unina.it; mucherin@unina.it; pianese@unina.it; francesco.pirozzi@unina.it

Abstract

A stochastic approach is proposed, capable of specifying the optimal allocation of different sets of monitoring stations aiming at an early detection of the intentional contamination of a water distribution network. The approach is based on the use of the Monte Carlo technique for the generation of different Users' Water Demand Scenarios (U'WDS), which are variable both within the day and from one day to another. Once the generation of scenarios is performed, a direct analysis of the hydraulic behaviour of the distribution network by a capable hydraulic model allows, for each U'WDS, the evaluation of the hydraulic flow characteristics, such as the mean flow velocities through each link of the network. Given an intentionally contaminated input node, the hydraulic characteristics of the flow calculated for each U'WDS allow the evaluation, by means of a lagrangian transport model, of the arrival time of substance transported at each node of the network. Then, by varying the node chosen as the possible contamination source, it is possible to evaluate, for each node of the network, the arrival times from all potential intrusion nodes. For each operating condition scenario, the locations of monitoring stations are chosen at nodes which maximize the number of upstream nodes characterized by arrival times smaller than a pre-assigned value (Early Warning Time, EWT). Finally, the optimal set of monitoring stations is chosen by a statistical analysis of the results obtained with reference to the generated scenarios.

Keywords
Monitoring, monitoring stations, contamination, intentional contamination, water supply systems, water distribution networks, early warning.

1. INTRODUCTION

Since September 11th, 2001, the attention towards threats caused by possible terrorist attacks has been increasing exponentially, putting a lot of pressure on technicians and policy makers in order to implement policies and strategies aiming at the protection of water supply infrastructures: these attacks, whose effects may be potentially catastrophic, include cyber or physical disruption, biological, chemical or radioactive contamination. General disruption (for instance, concerning municipal tanks or pumping stations) or contamination of water supply/distribution systems, causing long periods of service interruption, may have important consequences on public health and morale, and on economic and industrial activities. But also the case of contamination concerning moderately extended portions of a water distribution system can cause generalized panic, and also, direct effects consisting of a number of illnesses or life losses among the water consumers.

Even if security experts believe that the probability (*hazard*) of threatening water supply/distribution systems by introducing sufficient quantities of chemical, radioactive or microbial contaminants is substantially low, the possibility that a determined terrorist can access a water distribution system, endangering the life of thousands people, makes the *risk* very high. It is common knowledge that water distribution networks are the most vulnerable part of the whole water supply/distribution system, since they provide numerous points of unauthorized potential access. In the past, research about the contamination of municipal drinking water distribution systems has mainly considered the threats due to accidental events [1], [2]. Recently, a number of papers have been proposed, concerning both methodologies for optimal allocation of monitoring stations against terrorist attack [3], [4], [5], [6], and the design of contaminant detection systems aiming at early warning in water distribution networks [7], [8]. In this paper, with the aid of a case study, a methodology aiming at an optimal allocation of monitoring stations for the early detection of intentional contamination in water distribution systems is proposed, taking in account the variability of users' water demand. The proposed approach, though based on the generation of different water demand patterns, is quite different from the one proposed by other authors [9], and consists of the following steps:

a) a number of equally-probable Users' Water Demand Scenarios (*U'WDS*) in a water distribution network is generated via Monte Carlo simulation; in this approach, seasonal, daily and stochastic components of demand patterns variability could be taken in account;
b) a proper level of service is fixed, consisting of the maximum time elapsed before contamination detection (Early Warning Time, *EWT*);
c) given the water demand patterns, the hydraulic behaviour of the water distribution system during a sufficiently long time interval is simulated, taking into account the changes of the users' water demand;
d) each node of the water distribution system can be considered a potential contamination source because of a threat event: taking into account different instants at which the attack could begin, the spread of a non reacting tracer in the water distribution system is simulated, making use of a convection-dispersion model; every node reached by the non reacting-tracer during the spreading simulation can be considered definitively contaminated;
e) for each source node, and for each instant at which the attack could begin, the nodes reached by the tracer before chosen *EWT* is exceeded are considered; these nodes are then considered potential locations for contaminant monitoring stations;
f) for each potential starting point of the contamination attack, sets of increasing numbers of monitoring stations are considered, capable of covering the network partially or totally;
g) finally, statistics of the sets of detection stations are considered, determining the optimal layout of the monitoring system.

2. GENERATION OF A SET OF USERS' DEMAND PATTERNS BY MEANS OF THE MONTECARLO METHOD

The discharges flowing in municipal water distribution networks are variable, in time and space, due to the variability of users' demand. Since these variations are partially random, the application of a deterministic approach doesn't allow for the consideration of all the possible water demand scenarios. As a consequence, a probabilistic approach has to be adopted, based on the hydraulic analysis of a number of randomly generated operating conditions, accordingly to the variations of one or more of the components that characterise the water demand: i) variations through the years; ii) oscillations in the year; iii) variations in the week and/or in the day; iv) random variations, connected to the random behaviour of the users. Since the first two components (respectively, annual and seasonal) are closely correlated to the specific characteristics of the users, in this paper, according to [10] and [11], only the daily and random components of the users' water demand have been considered, assuming that in each node the mean daily variation of the water demand consists of N_{ti} = 48 different operating conditions (each of them 30 minutes long), and follows a trend characterised by an am peak and two less important pm peaks, one in the afternoon and the other one in the night. Particularly, in each node of the distribution system (whose number is equal to *S*) a peak coefficient has been adopted, depending on the number of users served by the node itself, considering the random component of the demand as a variable that is statistically independent from the others. To this aim, if *k* indicates the generic time interval in which the day is subdivided (k = 1, 2, …., N_{ti}), the Demand Coefficient ($DC_{j,k}$) referring to the node *j* and to the interval *k* is defined as:

$$DC_{j,k} = \frac{V_{j,k}}{E\left[\sum_{K=1}^{48} V_{j,k} \middle/ 48 \right]} \quad (1)$$

where the numerator and the denominator represent, respectively, the water volume supplied at the node *j* during the time interval *k*, and the mean water volume delivered to the users located at the same node during a 30 minute interval. If the Demand Coefficients $DC_{j,k}$ could be considered as random variables, then they are log-normally distributed with mean $E[DC_{j,k}]$ and coefficient of variation $CV[DC_{j,k}]$. If we consider $CV[DC_{j,k}]$ independent of *j* and *k*, i.e. $CV[DC_{j,k}] = CV = constant$, the values of the Demand Coefficients have to be evaluated by means of the following equation:

$$(DC_{j,k}) = 10^{\left\{ \log E[DC_{j,k}] - \frac{1}{2}\log\left(1+CV^2[DC_{j,k}]\right) + u_f \left[\frac{\log\left(1+CV^2[DC_{j,k}]\right)}{\ln 10} \right]^{1/2} \right\}} \quad (2)$$

where u_f is the standard normal random deviate. Finally, in order to generate $N_{Sim}=N_d \times N_{ti}$ water demand patterns by means of the Monte Carlo technique (being N_d the number of days considered in the analysis), the mean values of the water demand coefficients (*E[DCj,k])* and the corresponding coefficient of variation (*CV*) are needed [10] [11]. Sufficiently long periods of "unsteady" hydraulics scenarios are taken in account by means of sequences of steady operating conditions, since the hydraulics could be regarded as steady in each of the N_{sim} operating conditions. If the hydraulics could be considered periodic, with period N_d, then it is possible to consider N_{sim} different hydraulics scenarios, each of them starting from a different users' water demand condition

3. HYDRAULIC ANALYSIS OF THE WATER DISTRIBUTION NETWORK

The hydraulic model applied is based [12] on the solution of a set of S non linear equations in S unknowns (the potential heads at the nodes of the water distribution system, h_j). Since the model has been diffusely described elsewhere [12] and also within a companion paper [13], it is only very briefly illustrated here. Generally speaking, when the flow is oriented from the node n to the node j (with n hydraulically located upstream to j), the hydraulic head loss along the pipe connecting the nodes n and j, $h_{n,j}$, could be evaluated as:

$$h_{n,j} = h_n - h_j = SF_{n,j} \cdot \beta_{n,j} \cdot \frac{Q_{n,j}^{\alpha_{n,j}}}{D_{n,j}^{\omega_{n,j}}} \cdot l_{n,j} = r_{n,j} \cdot Q_{n,j}^{\alpha_{n,j}} \tag{3}$$

whereas, when the flow is oriented from the node j to the node z, hydraulically located downstream, the hydraulic head loss along the pipe connecting the nodes j and z, $h_{j,z}$, could be evaluated as:

$$h_{j,z} = h_j - h_z = SF_{j,z} \cdot \beta_{j,z} \cdot \frac{Q_{j,z}^{\alpha_{j,z}}}{D_{j,z}^{\omega_{j,z}}} \cdot l_{j,z} = r_{j,z} \cdot Q_{j,z}^{\alpha_{j,z}} \tag{4}$$

In the equations (3) and (4), D and l represent the diameter and the length of the pipe, respectively; Q is the discharge flowing in the pipe; the exponents α and ω are parameters depending on the type of flow (laminar or turbulent); β is a coefficient that could be itself a function of the discharge, the pipe diameter and the roughness of walls, SF is a safety factor, to be considered because of the uncertainness involved in the correct estimation of head losses; and $r = SF \cdot (\beta \cdot l) / D^{\omega}$. The continuity equation at node j is:

$$\sum_{n=1}^{N_{j1}} Q_{n,j} - \sum_{z=1}^{N_{j2}} Q_{j,z} \pm Q_j = 0 \tag{5}$$

where: N_{j1} and N_{j2} = number of pipes respectively inflowing and outflowing from the node j; and Q_j is the discharge directly inflowing or outflowing from the node j. Substituting in the equation (5) the expressions of $Q_{n,j}$ and $Q_{j,z}$ obtained, respectively, from the equations (3) and (4), it is possible to obtain

$$\sum_{n=1}^{N_{j1}} \left[\left| \frac{h_n - h_j}{r_{n,j}} \right|^{1/\alpha_{n,j}} \right] - \sum_{z=1}^{N_{j2}} \left[\left| \frac{h_j - h_z}{r_{j,z}} \right|^{1/\alpha_{j,z}} \right] \pm Q_j = 0 \tag{6}$$

The set of non linear equations (6) in the unknowns h_j is solved by using a Newton-Raphson procedure, applied in the way proposed in [14], and generalised in [13].

4. EVALUATION OF THE ARRIVAL TIMES, AT VARIOUS NODES, OF CONTAMINANTS INTRODUCED AT GIVEN POINTS OF THE NETWORK

The water quality model used for the analysis ([11] and [15]) is able to take into account the variations of contaminants' concentration due to the following phenomena: i) water mixing at the nodes of the system (tanks and pipes junctions); ii) advection; iii) dispersion; iv) molecular diffusion; v) reaction; vi) volatilisation. However, in the present application, the model is just used to trace the contaminant from a node of the system to another, without dispersing, reacting or decaying. The model consists of a set of mass balance equations able to explain the variations of the contaminant present:

i) in the water contained in the tanks existing in the system

$$C_t\big|_v = \frac{V_v \cdot C_{t-\Delta t}\big|_v + \Delta t \cdot \sum_{n=1}^{N_{j1}} Q_{n,j} \cdot C\big|_{s=L} + \Delta t \cdot M_v - \Delta t \cdot K_v \cdot C_{t-\Delta t}\big|_v}{V_v + \Delta t \cdot \sum_{z=1}^{N_{j2}} Q_{j,z} + \Delta t \cdot Q_{wd}} \tag{7}$$

ii) in the water flowing through the nodes

$$C_j = \frac{\sum_{n=1}^{N_{j1}} Q_{n,j} \cdot C_{n,j} + M_j}{\sum_{z=1}^{N_{j2}} Q_{j,z} + Q_j} \tag{8}$$

iii) in each pipe of the network

$$\frac{\partial C}{\partial t} = -U \cdot \frac{\partial C}{\partial s} + D_T \cdot \frac{\partial^2 C}{\partial s^2} + D_M \cdot \frac{\partial^2 C}{\partial s^2} - K \cdot C \tag{9}$$

In these equations, the meaning of the symbols is the following: C = contaminant concentration at instant t and abscissa s; $C_{n,j}$ = contaminant concentration at the terminal cross-section of the pipe connecting nodes n and j; U = flow velocity; D_T and D_M = eddy and molecular diffusion coefficients; K_v = contaminant decay constant in tanks, related to bulk reaction phenomena, such as volatilisation and reaction with other substances contained in the water; K = contaminant decay constant in the pipes, related to both bulk and wall reactions; M_j and M_v = mass of contaminant directly introduced into node j and the tank V; Δt = time step; V_v = mean water volume in the tank V during Δt; $C_t|_V$ and $C_{t-\Delta t}|_V$ = contaminant concentration in the tank at t and t-Δt; $C|_{s=L}$ = contaminant concentration in the last node of the pipe whose flow enters the tank. In the present application, the parameters D_T , D_M, and K were set equal to zero. K_v is obviously set equal to zero too, because in the water distribution network used as case-study there are no reservoirs.
The solution technique used in order to take into account the advection phenomena without introducing numerical dispersion is based on a lagrangian approach, not far from that proposed in [16]: the volume of water contained in each pipe, whose length is l_y (alternatively equal to $l_{n,j}$ or $l_{j,z}$), is subdivided into EC_y elementary cells, which are traced as they move downstream through the pipe. The number of cells EC_y is variable from one pipe to another. With the chosen values for the coefficients, equation (9) can then be discretised, in the framework of the lagrangian approach, as

$$C_{i,y}^{t} = C_{i-1,y}^{t-\Delta t} \qquad (10)$$

where $C_{i,y}^{t}$ is the concentration in the cell i of the pipe y at time t. If the flow is not steady but slowly variable in time, as it was supposed in the present paper, the number of cells varies from one operating condition to another, and concentration has to be properly redistributed when passing from one operating condition to the following one. The water quality module has been applied, for each users' demand scenario, to evaluate the *arrival time* of the contaminant at each node of the network, given a pre-assigned node for contaminant intrusion.

5. OPTIMAL ALLOCATION OF MONITORING STATIONS

For each of the N_{Sim} users' demand scenarios, the quality model has to be repeatedly used, considering each node of the network, in turn, as a potential contaminant source. As a consequence, the quality module has to be applied N_t times ($N_t = N_d \times N_{ti} \times S_c$). By means of the quality module, for a given scenario, it's possible to evaluate the arrival times, at each node, for all the potential input nodes. Thus, for a given users' demand scenario and contaminant source node, all the nodes which are reached by the contaminant in a time smaller than a pre-assigned Early Warning Time *ETW* are potential locations for a monitoring station: in turn, these nodes "cover" upstream nodes characterised by the arrival time smaller than the one fixed for warning. Of course, if the source node changes, for a given scenario, the covered nodes change too.
In the approach considered here, it has been hypothesized that the monitoring stations are allocated one at time at different nodes, and the analysis has been conducted by defining proper non-dimensional indexes, named *Presence Indexes* PI_i. In more detail, for each scenario, the node which covers the greatest number of potential contamination source nodes is chosen as suitable for the first monitoring station to be allocated. For each of these maximum coverage-nodes, the *First Presence Index* PI^1 is defined as the probability that a monitoring station is suitable for the allocation in that node, and is estimated as the ratio between the number of *U'WDS* in which the given node is a maximum-coverage node, and the total number of *U'WDS*. Then, the first monitoring station is allocated in the node n^1 which exhibits the greatest PI^1. For each users' demand scenario, a node is suitable for the allocation of the second monitoring station if it is a second-order maximum-coverage node, that is, if it covers the maximum number of nodes not covered by the first station. The *Second Presence Index* PI^2 measures the conditional probability that such a node is suitable for the allocation of a monitoring station, given the occurrence of n^1, and is estimated as the ratio between the number of scenarios in which the given node is a second-order maximum-coverage node, and the total number of occurrences of n^1. Then, the second monitoring station is allocated in the node n^2 which exhibits the greatest PI^2.
A similar definition stands for all the cases with more than 2 monitoring stations, where *Presence Indexes* PI^m, that are successive to PI^2, can be introduced.

6. APPLICATION OF THE PROCEDURE TO A CASE-STUDY

For the sake of simplicity, the procedure proposed in this paper was applied to the same case-study presented in a previous paper [11], considering a supply period N_d = 100 days long. For each day, N_{ti} = 48 different conditions of the water demand are considered. Figure 1 shows the water distribution network considered in the case-study.

Using the software described in the section 3, the hydraulic behaviour of the network was evaluated for each of the $N_{Sim} = 100 \times 48 = 4800$ generated conditions, obtaining, for each pipe, the value of the water discharge and flow velocity, and, for each node, the value of the potential head.

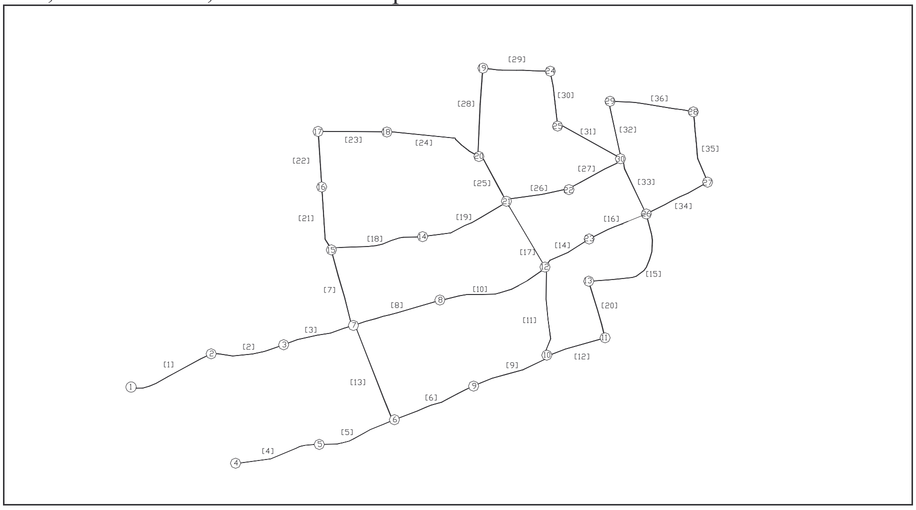

Fig. 1. Water distribution network considered in the case-study

For each working condition scenario, the quality module was applied $S_c = 30$ times, since all nodes of the network were considered to be, in turn, a potential contaminant source. Considering a tabular procedure very similar to the one proposed in [17], it was possible to evaluate, for each node where a monitoring station could be located, and for three *EWT* values (1, 3, and 6 hours): i) the nodes characterised by arrival times smaller than the chosen *EWT*, and thus the potential location for a monitoring station; ii) the node characterised by the maximum number of potentially monitored nodes (which is assumed as the location of the first monitoring station); iii) the node characterised by the maximum number of potentially monitored nodes not already monitored by the first station (which is assumed as the location of the second monitoring station); iv) the node characterised by the maximum number of potentially monitored nodes not already monitored by the first two stations (which is assumed as the location of the third monitoring station); etc. The monitoring stations obtained for the three different early warning times are shown in Table 1: when *EWT*=6 h, 3 monitoring stations are enough to cover all nodes; reducing *EWT*, the number of monitoring stations increases to 5.

Table 1. Allocation of Monitoring Stations for different early warning times

Early Warning Time	First Station Node	Second Station Node	Third Station Node	Fourth Station Node	Fifth Station Node
1 hour	30	7	18	11	28
3 hours	28	18	21	23	11
6 hours	28	17	9	-	-

To understand the approach considered for the evaluation of the nodes where monitoring stations have to be positioned, it could be useful to see the trends of the PI_j values evaluated for the case of an early warning time equal to 6 hours (Fig. 2). In this figure, one can see that the first monitoring station is allocated at node 28, since the corresponding PI_j value (PI_{28}) attains its maximum value just at this node ($PI_{28} = PI^1 = 0.676$). In turn, the second and third monitoring stations have to be allocated, respectively, at node 17 and 9, since the conditioned mass probability function values attain their maximum values just at these nodes ($PI_{17} = PI^2 = 0.244$ and $PI_9 = PI^3 = 0.282$).

7. CONCLUSIONS AND FUTURE DEVELOPMENTS

In this paper, a stochastic approach has been proposed aiming at an optimal allocation, within a water distribution network, of monitoring stations finalized to the early detection of intentional contamination of water. In a certain working condition, a node is suitable for the allocation of a monitoring station if it covers the greatest number of potential source nodes. The *First Presence Index* PI^1 has been defined, which takes into account the probability that a node is suitable for the allocation of a monitoring station, having considered the set of all the possible working conditions. When the monitoring stations are allocated one at a time, the first position coincides with the node which exhibits the greatest *First Presence Index*. Successive *Presence Indexes* have been defined, which aim to help in defining the allocation positions of successive monitoring stations. An example has been considered, in order to demonstrate the feasibility of the proposed approach for the application to the real world. The above defined *Presence Index* could be usefully re-defined to take into account the actual volume of contaminated water supplied by the distribution network: this approach will be the subject of future communications.

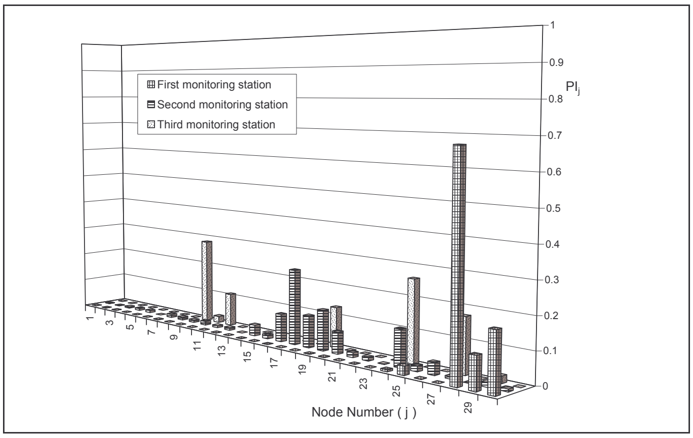

Fig. 2. Conditioned mass probability functions of nodes suitable for the allocation of monitoring stations

References

[1] Kessler A., Ostfeld A., Sinai, G., "Detecting Accidental Contaminations in Municipal Water Networks", *Journal of Water Res. Plan. and Man.*, ASCE, Vol. 124, 1998, pp. 192–198.

[2] Kumar, A., Kansal, M.L., Arora, G., Ostfeld, A., Kessler, A., Discussion on the paper "Detecting Accidental Contaminations in Municipal Water Networks" by Kessler A., Ostfeld A. and Sinai, G., *Journal of Water Res. Plan. and Man.*, ASCE, Vol. 125, 1999, pp. 308–309.

[3] Ostfeld, A., Salomon, E., "Optimal Layout of Early Warning Detection Stations for Water Distribution Systems Security", *Journal of Water Res. Plan. and Man.*, ASCE, Vol. 130, 2004, pp. 377–385.

[4] Uber, J., Janke, R., Murray, R., Meyer, P., "Greedy Heuristic Methods for Locating Water Quality Sensors in Distribution Systems", *World Water Environmental Resources Congress 2004*, G. Sehlke, D.F. Hayes, D. K. Stevens - Editors, June 27 – July 1, 2004, Salt Lake City, Utah, USA, pp. 1–9.

[5] Berry, J.W., Fleischer, L., Hart, W.E., Phillips, C.A., Watson, J-P., "Sensor Placement in Municipal Water Networks", *Journal of Water Res. Plan. and Man.*, ASCE, Vol. 131, 2005, pp. 237–243.

[6] Shastri, Y., Diwekar, U., "Optimal sensor placement for water distribution network security", *AIChE 2005 Annual Meeting*, Cincinnati ,OH, USA,

[7] Ostfeld, A., Salomon, E., "An Early Warning Detection System (EWDS) for Drinking Water Distribution Systems Security", *World Water and Environmental Resources Congress 2003*, Paul Bizier, Paul DeBarry - Editors, June 23–26, 2003, Philadelphia, Pennsylvania, USA, Vol. 118, 2004, pp. 34.

[8] Ostfeld A., Kessler A., Goldberg, I., "A contaminant detection system for early warning in water distribution networks", *Engineering Optimization*, Taylor & Francis, Vol.36, 2004, pp. 525–538.

[9] Ostfeld, A., Salomon, E., "A Stochastic Early Warning Detection System Model for Drinking Water Distribution Systems Security", *World Water Environmental Resources Congress 2004*, G. Sehlke, D.F. Hayes, D. K. Stevens - Editors, June 27 – July 1, 2004, Salt Lake City, Utah, USA, pp. 454–459.

[10] Gargano, R., Pianese, D., "Reliability as a tool for hydraulic networks planning", *Journal of Hydraulic Engineering,* ASCE, Vol.5, May, 2000, pp. 354-364.

[11] Cozzolino, L., Pianese, D., Pirozzi, F., "DBPs control in water distribution systems through optimal chlorine dosage and disinfection stations allocation", *Desalinisation*, Elsevier B.V., 2005, 176, pp. 113-125.

[12] Pianese, D., Masini, P., "Sui transitori che si sviluppano nelle reti idrauliche in pressione munite di luci di efflusso regolabili e di dispositivi di attenuazione dei fenomeni di moto vario". *Atti del Seminario "Moto Vario nei Sistemi Acquedottistici.* Bari (Italy), 1994, pp. 159-183 (in italian).

[13] Cozzolino, L., Covelli, C., Mucherino, C., Pianese, D., "Hydraulic Reliability of Pressurized Water Distribution Networks for On-Demand Irrigation". This issue, Essex (En), 2005.

[14] Lam, C.F., Wolla, M.L., "Computer Analysis of Water Distribution Systems. Part. II – Numerical Solution", *Journal of Hydraulic Division,* ASCE, Vol. 98, 1972, pp. 447-460.

[15] Pirozzi, F., Pianese, D., d'Antonio, G., "Water quality decay modelling in hydraulic pressure systems", Water Science and Technology: Water Supply, Vol. 2, No.4, IWA Publishing, 2002, pp 111-118.

[16] Rossman, L.A., Clark, R.M., Grayman W.M., "Modeling Chlorine Residuals in Drinking-Water Distribution Systems". *Journal of Environmental Engineering*, ASCE,Vol. 120, 1994, pp. 803-820.

[17] Arora, G., "Water quality modelling and monitoring in distribution systems". Thesis for the partial fulfillment PhD in Civil Engineering, University of Delhi (India), 2003.