Civil Engineering and Environmental Systems
Vol. 23, No. 3, September 2006, 161–174

Positioning, within water distribution networks, of monitoring stations aiming at an early detection of intentional contamination

L. COZZOLINO*, C. MUCHERINO, D. PIANESE and F. PIROZZI

Dipartimento di Ingegneria Idraulica e Ambientale 'Girolamo Ippolito', Università degli Studi di Napoli Federico II, via Claudio 21, I-80125, Napoli, Italy

(Received 9 December 2005; in final form 11 April 2006)

A stochastic approach is proposed, aiming at the optimal allocation of increasing sets of monitoring stations for the early detection of the intentional contamination of water distribution networks. The approach is based on the use of the Monte Carlo technique for the generation of a number of time-varying hydraulic scenarios, each consisting of a succession of steady conditions related to different users' water demands. Given a time-varying hydraulic scenario, and chosing an injection node, the spreading of the contaminant through the network is evaluated by means of a Lagrangian advection model, and the arrival times to all the potential monitoring stations are calculated. If these operations are accomplished for all the source nodes, and for each of the time-varying hydraulic scenarios, a statistical analysis allows for the allocation of the monitoring stations which maximise the number of upstream nodes characterised by arrival times less than a pre-assigned value (*early warning time*).

Keywords: Monitoring stations; Intentional contamination; Water distribution networks; Early warning

1. Introduction

Since 11 September 2001, there has been an increasing interest towards threats caused by possible terrorist attacks on the water distribution systems, and towards the policies and strategies aiming at the protection of these infrastructures. Terrorist attacks on water distribution systems, whose effects may be potentially catastrophic, include cyber or physical disruption and 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 activity. Contamination of moderate sections of a water distribution system can cause generalized panic as well as a number of illnesses and loss of life among the water consumers.

*Corresponding author. Tel.: +39 (0)81 7683430; Fax: +39 (0)81 5938936; Email: lcozzoli@unina.it

Civil Engineering and Environmental Systems
ISSN 1028-6608 print/ISSN 1029-0249 online © 2006 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/10286600600789359

Even if security experts believe that the probability (hazard) of threatening water supply/distribution systems by introduction of sufficient quantities of chemical, radioactive or microbial contaminants is substantially low, the possibility that a determined terrorist can access a water distribution system, and endanger the lives of thousands of 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, as they provide numerous points of unauthorised potential access.

In the past, research about the contamination of municipal drinking water distribution systems has mainly considered the threats due to accidental events (Kessler *et al.* 1998, Kumar *et al.* 1999). Recently, a number of papers have been proposed concerning (i) the potential exposure of consumers to contaminants deliberately injected into a water distribution system, under deterministic and stochastic conditions (Nilsson *et al.* 2005) (ii) the methodologies for the optimal allocation of monitoring stations against terrorist attacks or external pollution (Kessler *et al.* 1998, Berry *et al.* 2005), and (iii) the design of contaminant detection systems aiming at early warning in water distribution networks, considering the variability of demand and random contamination events (Ostfeld and Salomon 2004).

Especially relevant to our discussion are the articles by Nilsson *et al.* (2005) and Ostfeld and Salomon (2004). In the first article (Nilsson *et al.* 2005) a Monte Carlo approach was applied in order to study the spreading, through an urban water distribution system, of a biochemical contaminant injected in a single point, located at the centre of the network. In particular, a thousand different runs were considered, taking into account the random variability of the water demand through a 55 h lapse. It was found that the stochastic demand could greatly affect the variability, from one run to another, of the total contaminant delivered at the nodes, demonstrating the usefulness of the Monte Carlo approach in problems related to the spreading of a contaminant through a water distribution network. In the article by Ostfeld and Salomon (2004), a deterministic demand pattern was instead considered, but the injection position and the time of the attack were pseudo-randomly generated, also considering the case of two or more contemporary injections: then, a genetic algorithm was used to determine, for a given number of monitoring stations, the position of a set of monitoring stations which maximised the detection likelihood.

In this article, a methodology aiming at an optimal allocation of monitoring stations for the early detection of intentional contamination in water distribution systems is proposed, taking contemporarily into account both the random variability of users' water demand and the uncertainties about the injection point of the contaminant. The proposed approach consists of the following steps.

(1) Given a water distribution network, a great number of equally-probable time-varying users' water demand scenarios (TDSs), each consisting of a sufficiently long sequence of constant demand steps, are generated by means of the Monte Carlo method, considering the random changes of the users' water demand: here, the daily and stochastic components of the demand patterns variability are taken into account.
(2) Using a steady hydraulic model, the hydraulic characteristics (potential head at the nodes, discharge through the pipes; flow velocities through the pipes, and so on) are evaluated for each step of the TDSs: in this way, a number of equally-probable time-varying hydraulic scenarios are obtained, each consisting of a long sequence of steady hydraulic conditions.
(3) A proper level of service is fixed, consisting of the maximum time elapsed before contaminant detection (early warning time EWT).
(4) Each node of the water distribution system is considered as a potential contamination source: for a given TDS, and a fixed node (potential source of contamination), an event consisting of the injection of a non-reacting tracer with constant mass rate is considered.

The spreading of the contaminant in the water distribution system is simulated, by making use of a proper quality model; every node reached by the non-reacting tracer during the spreading simulation is considered definitively contaminated: then, the nodes that are contaminated before the chosen EWT is exceeded are considered potential locations for the contaminant monitoring stations.

(5) Given a TDS, the operations at point (4) are repeated for each node of the network; finally, sets of increasing numbers of monitoring stations are considered, capable of covering the network partially or totally during the given time-varying hydraulic scenario.
(6) The operations at points (4) and (5) can be repeated for a great number of different TDSs, each starting at a different instant.
(7) Finally, the statistics of the sets of monitoring stations are considered, determining the optimal layout of the monitoring system.

In the following sections the model used to generate the water demand patterns will be illustrated, together with the hydraulic and quality models; moreover, a procedure aiming at the optimal allocation of the monitoring stations, which takes into account the stochastic variability of the hydraulic conditions in the water distribution system, is illustrated and applied to a real-world case.

2. Generation of a set of users' demand patterns by means of the Monte Carlo method

The discharges flowing in municipal water distribution networks are variable, in time and space, because of the variability of users' demand. As these variations are partially random, the application of a deterministic approach does not allow for the consideration of all the possible water demand scenarios. As a consequence, a probabilistic approach must be adopted, which has to be based on the hydraulic analysis of a number of pseudo-randomly generated operating conditions, according to the variations of one or more of the components that characterize the water demand: (a) variations from one year to another; (b) seasonal oscillations; (c) variations in the week and/or in the day; and (d) random variations, connected to the random behaviour of the users. As the first two components (respectively, annual and seasonal) are closely correlated to the specific characteristics of the users, in this article only the daily and random components of the users' water demand have been considered, without losing in generality.

It has been assumed that, in each node of the distribution system (whose number is equal to S), the mean daily variation of the water demand can be approximated by $N_{\mathrm{ti}} = 48$ different operating conditions (each of them 30 min long), following a trend characterized by an a.m. peak and two less important p.m. peaks, one in the afternoon and the other one in the night (figure 1).

If k indicates the generic time step in which the day is subdivided ($k = 1, 2, \ldots, N_{\mathrm{ti}}$), the *demand coefficient* ($\mathrm{DC}_{j,k}$) referring to the node j and to the interval k is defined as:

$$\mathrm{DC}_{j,k} = \frac{V_{j,k}}{\mathrm{E}\left[\sum_{k=1}^{N_{\mathrm{ti}}} V_{j,k}/N_{\mathrm{ti}}\right]} \tag{1}$$

where the numerator and the denominator represent, respectively, the water volume actually supplied at the node j during the time interval k, and the daily average water volume delivered to the users located at the same node during a time interval of 30 min. The demand coefficients $\mathrm{DC}_{j,k}$ at the nodes are random variables, which are assumed to be log-normally distributed with mean $\mathrm{E}[\mathrm{DC}_{j,k}]$ and coefficient of variation $\mathrm{CV}[\mathrm{DC}_{j,k}]$, and which can be evaluated by

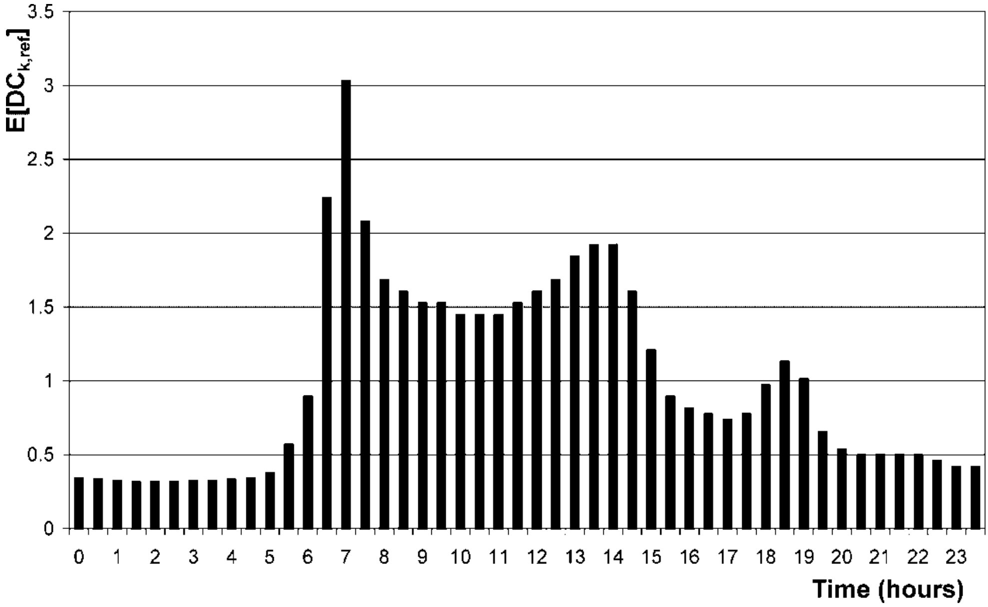

Figure 1. The average daily water demand coefficients of the network considered in the case-study.

means of the following equation:

$$\mathrm{DC}_{j,k} = 10^m \tag{2}$$

where

$$m = \log \mathrm{E}[\mathrm{DC}_{k,j}] - \frac{1}{2}\log(1 + \mathrm{CV}^2[\mathrm{DC}_{k,j}]) + u_\mathrm{f}\left\{\frac{\log(1 + \mathrm{CV}^2[\mathrm{DC}_{k,j}])}{\ln 10}\right\}^{1/2} \tag{3}$$

and u_f is the standard normal random deviate.

Finally, in order to generate $N_\mathrm{Sim} = N_\mathrm{d} \times N_\mathrm{ti}$ water demand conditions by means of the Monte Carlo technique (N_d being the number of days considered in the analysis), the mean values of the water demand coefficients ($\mathrm{E}[\mathrm{DC}_{j,k}]$) and the corresponding coefficient of variation $\mathrm{CV}[\mathrm{DC}_{j,k}]$ are needed. The mean value $\mathrm{E}[\mathrm{DC}_{j,k}]$ can be correlated (Gargano and Pianese 2000) to the population served by each node and to the average demand coefficient $\mathrm{E}[\mathrm{DC}_{k,ref}]$ of the network taken as a whole (figure 1). The values of the nodal coefficient of variation $\mathrm{CV}[\mathrm{DC}_{j,k}]$ which are commonly used in literature lie in the range of 0.1–0.2 (Bao and Mays 1990, Babayan *et al.* 2005), tending to an asymptotic value for increasing population served: here, a uniform value $\mathrm{CV} = 0.2$ has been taken for demonstrative purposes.

After the generation of the water demand conditions is accomplished, the corresponding steady hydraulics are solved. Sufficiently long periods of time-varying hydraulic scenarios are taken into account by means of successions of steady operating conditions: more in particular, N_Sim different time-varying hydraulic scenarios have been considered, each starting from a different water demand condition.

3. Hydraulic analysis of the water distribution network

The hydraulic model applied is based (Pianese and Masini 1994) 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). As the model has been widely described elsewhere, 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 through the pipe connecting the nodes n and j, $h_{n,j}$, can be evaluated as:

$$h_{n,j} = h_n - h_j = \beta_{n,j} \frac{Q_{n,j}^{\alpha_{n,j}}}{D_{n,j}^{\omega_{n,j}}} l_{n,j} = r_{n,j} Q_{n,j}^{\alpha_{n,j}} \tag{4}$$

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, $h_{j,z}$, can be evaluated as:

$$h_{j,z} = h_j - h_z = \beta_{j,z} \frac{Q_{j,z}^{\alpha_{j,z}}}{D_{j,z}^{\omega_{j,z}}} l_{j,z} = r_{j,z} Q_{j,z}^{\alpha_{j,z}}. \tag{5}$$

In equations (3) and (4), D and l represent the diameter and the length of the pipe, respectively; Q, the discharge flowing in the pipe; the exponents α and ω are parameters depending on the type of flow (laminar or turbulent) and on the hydraulic behaviour (smooth, transitional, or very rough) of the pipe; β is a coefficient that could be itself a function of the discharge, the pipe diameter and the roughness of walls, and $r = (\beta l)/D^{\omega}$. The continuity equation at node j is:

$$\sum_{n=1}^{N_{j1}} Q_{n,j} - \sum_{z=1}^{N_{j2}} Q_{j,z} \pm Q_j = 0 \tag{6}$$

where N_{j1} and N_{j2} are the number of pipes, respectively, inflowing and outflowing from the node j and Q_j the discharge directly inflowing or outflowing from the node j. Substituting in equation (6) the expressions of $Q_{n,j}$ and $Q_{j,z}$ obtained, respectively, from equations (4) and (5), one has

$$\sum_{n=1}^{N_{j1}} \left[\left| \frac{h_n - h_j}{r_{n,j}} \right|^{1/\alpha_{n,j}} \right] - \sum_{z=1}^{N_{j2}} \left[\left| \frac{h_j - h_z}{r_{j,z}} \right|^{1/\alpha_{j,z}} \right] \pm Q_j = 0. \tag{7}$$

The set of non-linear equations (7) in the unknowns h_j, together with the corresponding equations which are valid for pumps, valves, reservoirs and so on, is solved using a Newton–Raphson descent procedure, with a constant backtracking coefficient (Mignosa 1987).

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 (Pirozzi *et al.* 2002, Cozzolino *et al.* 2005) is able to take into account the variations of contaminant concentration due to the following phenomena: (a) water mixing at the nodes of the system (tank and pipe junctions); (b) advection; (c) dispersion, molecular and turbulent diffusion; (d) reactions; and (e) volatilisation. However, in the present application, the model is just used to trace the contaminant from one node of the system to another, without dispersing, reacting, or decaying.

Generally speaking, the model consists of a set of mass balance equations able to explain the variations of the contaminant present.

(1) In the water contained in the tanks existing in the system

$$C_t|_{\mathrm{v}} = \frac{V_{\mathrm{v}} C_{t-\Delta t}|_{\mathrm{v}} + \Delta t \sum_{n=1}^{N_{j1}} Q_{n,j} C|_{s=L} + \Delta t M_{\mathrm{v}} - \Delta t K_{\mathrm{v}} C_{t-\Delta t}|_{\mathrm{v}}}{V_{\mathrm{v}} + \Delta t \sum_{z=1}^{N_{j2}} Q_{j,z} + \Delta t Q_{\mathrm{wd}}}. \tag{8}$$

(2) In the water flowing through the nodes

$$C_j = \frac{\sum_{n=1}^{N_{j1}} Q_{n,j} C_{n,j} + M_j}{\sum_{z=1}^{N_{j2}} Q_{j,z} + Q_j}. \tag{9}$$

(3) In each pipe of the network

$$\frac{\partial C}{\partial t} = -U \frac{\partial C}{\partial s} + D \frac{\partial^2 C}{\partial s^2} - KC. \tag{10}$$

In these equations, the symbols represent the following: C is the contaminant concentration at time t and abscissa s; $C_{n,j}$, the contaminant concentration at the last cross-section of the pipe connecting nodes n and j; U, the mean flow velocity; D, the dispersion coefficient (which comprises also molecular and turbulent diffusion); K_{v} the contaminant decay constant in the tanks related to bulk reaction phenomena, such as volatilisation and reaction with other substances contained in the water; K, the contaminant decay constant in the pipes related to both bulk and wall reactions; M_j and M_{v}, the mass rate of contaminant directly introduced into node j and the tank V; Δt, the time step used for the calculation of changes in water quality characteristics; V_{v}, the mean water volume in the tank V during Δt; $C_t|_{\mathrm{V}}$ and $C_{t-\Delta t}|_{\mathrm{V}}$ the contaminant concentration in the tank at t and $t - \Delta t$; and $C|_{s=L}$, the contaminant concentration in the last node of the pipe whose flow enters the tank. In the present application, the parameters D and K were set equal to zero. K_{v} was obviously set equal to zero too, because in the water distribution network used as the case-study there are no reservoirs.

The solution technique applied in order to take into account the advection phenomena without introducing numerical dispersion is based on a Lagrangian approach (Pianese *et al.* 1997a,b), similar to that proposed by Rossman *et al.* (1994): 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 (10) can then be discretized, in the framework of the Lagrangian approach, as:

$$C_{i,y}^{t} = C_{i-1,y}^{t-\Delta t} \tag{11}$$

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 article, 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 the contaminant injection.

5. Optimal allocation of monitoring stations

Given the difficulty of simulating an urban water distribution system for sufficiently long time intervals, the time-varying hydraulic scenarios can be usefully approximated by successions of steady hydraulic conditions, each related to a given water demand condition. In particular, for each of the N_{ti} time intervals into which a day is subdivided, the set of the S nodal values of DC_{jk} is generated by means of the procedure proposed by Gargano and Pianese (2000): then, multiplying these values by the respective nodal daily average discharges, a set of discharges delivered at the nodes is obtained. The generation of N_{ti} demand conditions can be repeated for each of the N_{d} days of the supply period under consideration, obtaining a water demand pattern, N_{d} days long, consisting of $N_{\text{Sim}} = N_{\text{d}} \times N_{\text{ti}}$ water demand conditions: the steady hydraulic conditions related to each of the generated water demand conditions can be calculated by means of the above described hydraulic model. Once a little more than N_{Sim} steady hydraulic conditions are available (obtained considering $N'_{\text{Sim}} = [(N_{\text{d}} + 1) \times N_{\text{ti}}]$ conditions instead of the $N_{\text{Sim}} = N_{\text{d}} \times N_{\text{ti}}$ strictly needed), it is possible to consider up to N_{Sim} time-varying hydraulic scenarios, each starting from a different steady hydraulic condition.

If a node of the water distribution system is chosen as a source of the contamination, it is possible to evaluate the spreading of the contaminant through the network from the instant of injection, which is assumed to be coincident to the instant of the beginning of the time-varying hydraulic scenario. All the nodes that are reached by the contaminant in a time smaller than a pre-assigned EWT are potential locations for a monitoring station: following the definition by Ostfeld and Salomon (2004), these nodes, which 'cover' the injection point, form the *Domain of Pollution Event* (DOPE) related to the source node. If a different source node is considered, for the given time-varying hydraulic scenario, the related DOPE changes accordingly. Conversely, all the source nodes that are covered by a given potential monitoring station form its *Domain of Detection* (DODN). For each time-varying hydraulic scenario, the nodes n_1 which exhibit the largest DODN are said to be *first order*, the nodes n_2 that cover the greatest number of potential source nodes not covered by the first order nodes are *second-order nodes*, and so on.

If N_{Sim} equally probable time-varying hydraulic scenarios are considered, it is possible to estimate the probability PI_1 (*first presence index*) that a node x is of the first-order type

$$\text{PI}_1(x) = P[n_1 = x] \tag{12}$$

as the ratio between the number of unsteady hydraulics scenarios in which the given node x is a first-order node, and the total number of time-varying hydraulic scenarios. The node M_1 which exhibits the greatest PI_1 is chosen for the allocation of the first monitoring station:

$$M_1\colon \text{PI}_1(M_1) = \max_{1 \le x \le S} \{\text{PI}_1(x)\}. \tag{13}$$

From the definitions above, it is clear that the water quality module has to be applied N_t times ($N_t = N_{\text{d}} \times N_{\text{ti}} \times S$), to evaluate the *presence indexes*. If it is possible to implement a second monitoring station, then the probability PI_2 (*second presence index*) that the node y is of the second order, conditioned to the event $M_1 = x$

$$\text{PI}_2(y) = P[n_2 = y | M_1 = x] \tag{14}$$

can be evaluated as the ratio between the number of scenarios in which the given node y is a second-order node, and the total number of occurrences of x as first-order node. The node M_2 which exhibits the greatest PI_2 is chosen for the allocation of the second monitoring station:

$$M_2\colon \text{PI}_2(M_2) = \max_{1 \le y \le s} \{\text{PI}_2(y)\}. \tag{15}$$

Similar definitions stand for all the cases with more than two monitoring stations, where presence indexes PI_m, that are successive to PI_2, can be introduced.

The optimisation procedure outlined above fits to the case that the monitoring stations are implemented one at a time (which is frequently the case when financial resources are scarce), and the position of the monitoring stations previously implemented influences the position of the stations implemented after: in fact, the proposed methodology is intended to maximise the detection likelihood, also in the case that not all of the planned monitoring stations have been implemented at a given time. If two or more stations are implemented simultaneously, the procedure will lead to a local optimum, and different optimisation techniques will be more adequate (see, for example Ostfeld and Salomons 2004).

6. Application of the procedure to a case-study

The procedure outlined above was applied to the case study presented in a previous article (Cozzolino *et al.* 2005), considering a supply period $N_\mathrm{d} = 100$ days long and, for each day, $N_\mathrm{ti} = 48$ different spatial distributions of the water demand (hydraulic time steps): in figure 1, the average daily demand coefficients of the network are plotted, whereas in table 1 the daily average values of the discharge are reported, with the population served at each node.

Figure 2 shows the layout of the case study water distribution network, which consists of 30 nodes and 36 steel pipes, whose diameters and lengths are reported in table 2. The feeding points of the system are the nodes 1 and 4, that are directly linked to a regional water supply system, which is served by a number of tanks and pumping stations: there are no such devices throughout the local network considered. Due to the presence of pressure regulation

Table 1. Nodal daily average discharges.

Node	Population served	Discharge (l/s)
2	300	0.910
3	200	0.606
5	350	1.061
6	1000	3.032
7	1150	3.487
8	1050	3.184
9	700	2.123
10	950	2.881
11	1150	3.487
12	1200	3.639
13	450	1.365
14	600	1.819
15	600	1.819
16	250	0.758
17	300	0.910
18	300	0.910
19	350	1.061
20	850	2.577
21	700	2.123
22	850	2.577
23	400	1.213
24	325	0.985
25	450	1.365
26	1150	3.487
27	650	1.971
28	425	1.289
29	550	1.668
30	750	2.274
Total	18,000	54.580

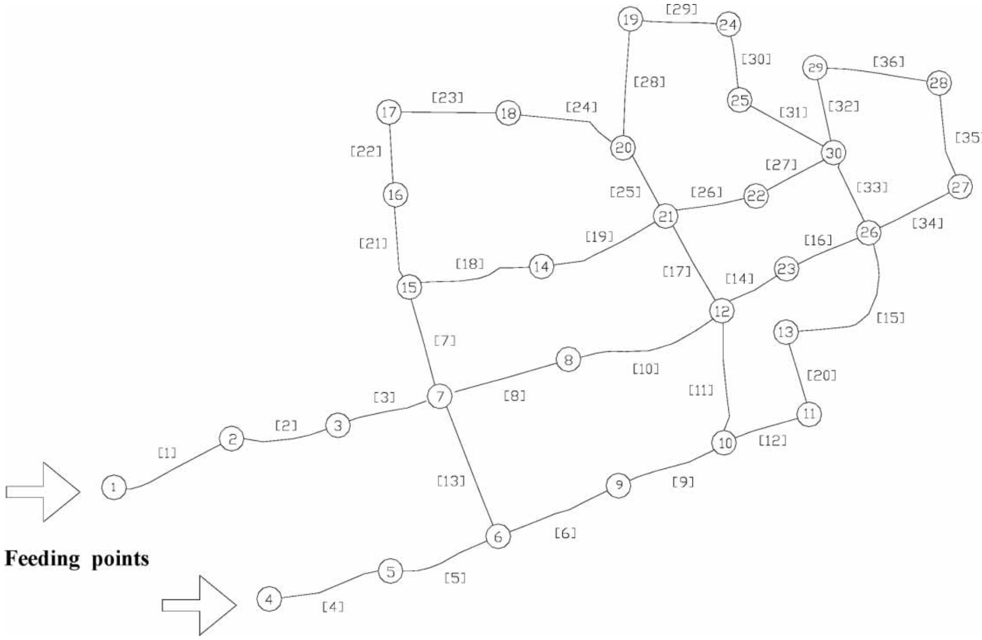

Figure 2. The layout of the water distribution network considered in the case-study.

Table 2. Pipeline characteristics of the water distribution system.

Link	Diameter (mm)	Length (m)	Link	Diameter (mm)	Length (m)
1	400	336	19	250	352
2	400	283	20	250	226
3	400	281	21	200	247
4	400	333	22	200	215
5	400	415	23	200	267
6	250	310	24	200	370
7	300	305	25	200	203
8	250	350	26	150	245
9	250	350	26	150	245
10	250	427	28	200	343
11	150	342	29	200	292
12	200	234	30	200	213
13	300	230	31	100	276
14	200	202	32	150	229
15	200	343	33	150	233
16	200	242	34	150	267
17	150	295	35	100	278
18	250	358	36	100	325

devices at the feeding points, the head at these nodes can be approximately taken as constant ($h_1 = h_4 = 110.0$ m a.m.s.l). The terrain elevation ranges between 51.0 and 57.0 m a.m.s.l., whereas the height of the buildings ranges between 4.0 and 15.0 m.

Prior to considering the effect of the choice of different values of the level of service EWT, it is interesting to evaluate the arrival times at node 28 of the contaminant injected, respectively, at the feeding points 1 and 4, considering 48 daily average and cyclically repeated demand conditions (pattern of figure 1). Considering, for the two feeding points, 48 contamination events, each starting at the beginning of a different constant demand step, the plot of figure 3 is obtained.

Observing figure 3, it is possible to verify that: (i) the arrival times at node 28 are always a little bit higher when node 4 is considered as the injection node; as a consequence, hydraulically speaking, node 28 could be considered farther from node 4 than from node 1; (ii) depending on the time at which the injection of the contaminant starts, the arrival times at node 28 can differ by a size order; in particular, the arrival times range between 5874.42 and 58,120 s if node 1 is considered as the injection node, and between 5357.48 and 57,730 s if node 4 is considered as the injection node; and (iii) because of the different values of the discharges requested from users flowing through the water distribution network, the arrival times at node 28 are higher if the contaminant is injected during the evening or the night, and smaller if the injection occurs at early morning; as a consequence, for the network examined in this article, the EWT has to be chosen not higher than 1 h if the contaminant injected is dangerous when small amounts are ingested, whereas it can be higher (for instance, 3 or 6 h) if the contaminant injected is dangerous only after the ingestion of medium or high amounts.

Using the software described in section 3, the hydraulic behaviour of the network was evaluated for each of the $N_{\mathrm{Sim}} = 100 \times 48 = 4800$ generated demand conditions, obtaining the value of the water discharges and flow velocities in the pipes, and the value of the potential head at the nodes. Finally, N_{Sim} different time-varying hydraulic scenarios were considered, consisting of sufficiently long successions of steady hydraulics conditions.

For each time-varying hydraulic scenario, the quality module was applied $S = 30$ times, as all the nodes of the network were considered in turn as a potential contaminant source: each event consisting of a constant injection of conservative substance, starting at the beginning of the scenario. The mass rate chosen for each injection was 1 g s^{-1}, whereas its duration was 50,000 s.

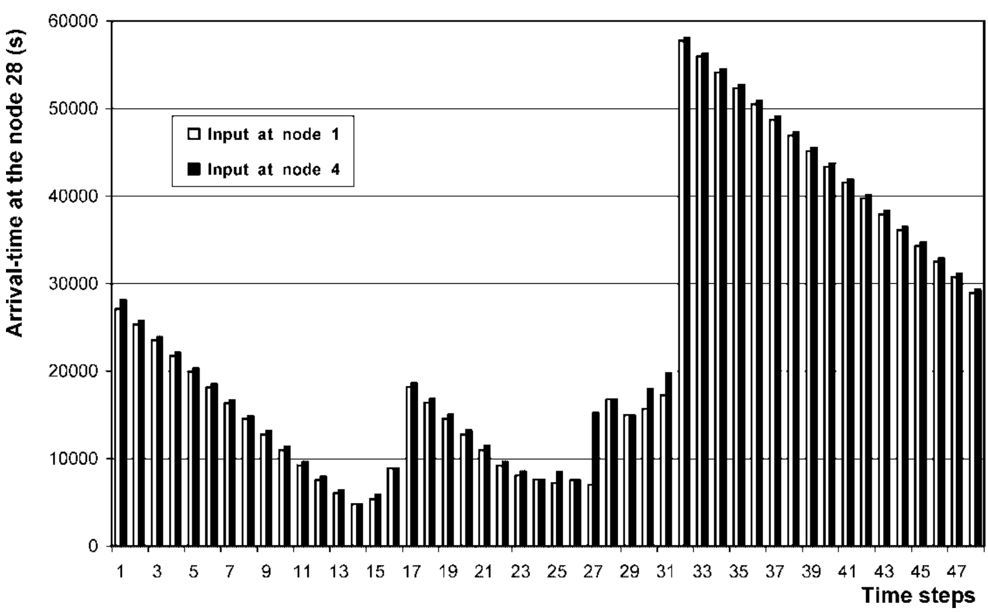

Figure 3. Arrival times at the node 28 of the contaminant injected at nodes 1 and 4, considering daily average demand conditions.

Applying a tabular procedure very similar to the one proposed in Arora (2003), it was possible to evaluate, for each scenario, and for three EWT values (1, 3, and 6 h):

(a) the nodes characterized by arrival times smaller than the chosen EWT, and therefore potential locations for a monitoring station;
(b) the nodes characterized by the maximum number of monitored nodes (first-order nodes);
(c) the nodes characterized by the maximum number of monitored nodes not already monitored by the first-order nodes (second-order nodes);
(d) the nodes characterized by the maximum number of monitored nodes not already monitored by the first- and second-order nodes (third-order nodes); etc.

Finally, the PI_1 was evaluated for each of the $S = 30$ network nodes, and the node exhibiting the greatest PI_1 was chosen for the allocation of the first monitoring station. Given the position M_1 of the first monitoring station, the PI_2 of the remaining network nodes were evaluated, allowing for the allocation of the second monitoring station, and so on.

To lighten the approach proposed for the evaluation of the nodes suitable for the allocation of a monitoring station, it could be useful to observe the plot of the PI_1 values calculated for the case EWT $= 6$ h.

In figure 4, one can see that the first monitoring station has to be allocated at node 28, as the first presence index PI_1 attains its maximum value at this node ($PI_1(28) = 0.676$). In turn, the second and third monitoring stations (figures 5 and 6) have to be allocated, respectively, at nodes 17 and 9, as the conditioned probability mass functions attain their maximum values at these nodes ($PI_2(17) = 0.244$ and $PI_3(9) = 0.282$).

The monitoring stations obtained for the three different EWTs are shown in table 3: when EWT $= 6$ h, three monitoring stations suffice to cover all the nodes, and the first monitoring station should be allocated at node 28; when considering EWT $= 1$ h and EWT $= 3$ h, the number of monitoring stations needed to cover the network increases to 5, and the first station should be allocated at nodes 30 and 28, respectively. Interestingly, the sets of monitoring stations supplied by the procedure differ in the two last cases: this is not surprising, because the method implies that the monitoring stations are implemented one at a time, and that the position of the first monitoring station influences the allocation of the remaining stations.

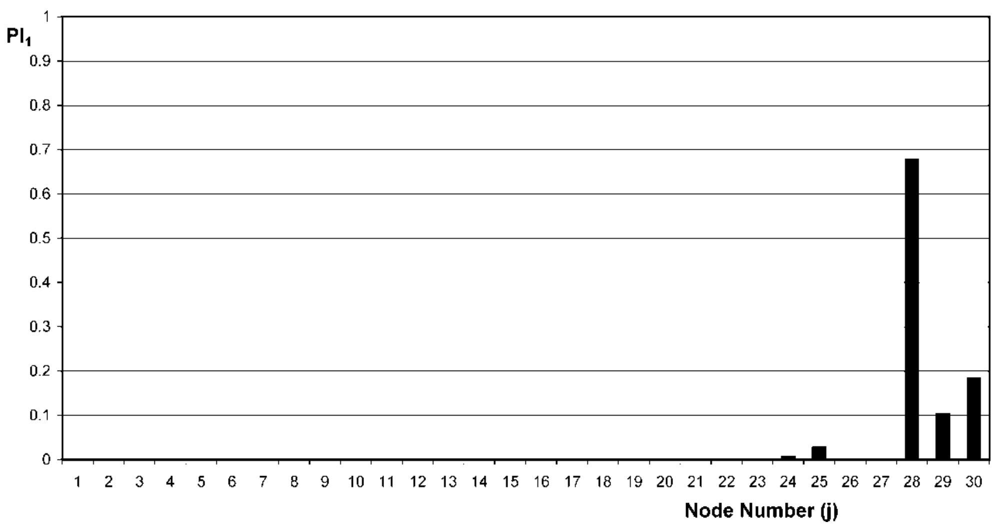

Figure 4. Probability mass functions of the nodes suitable for the allocation of the first monitoring station (EWT $= 6$ h).

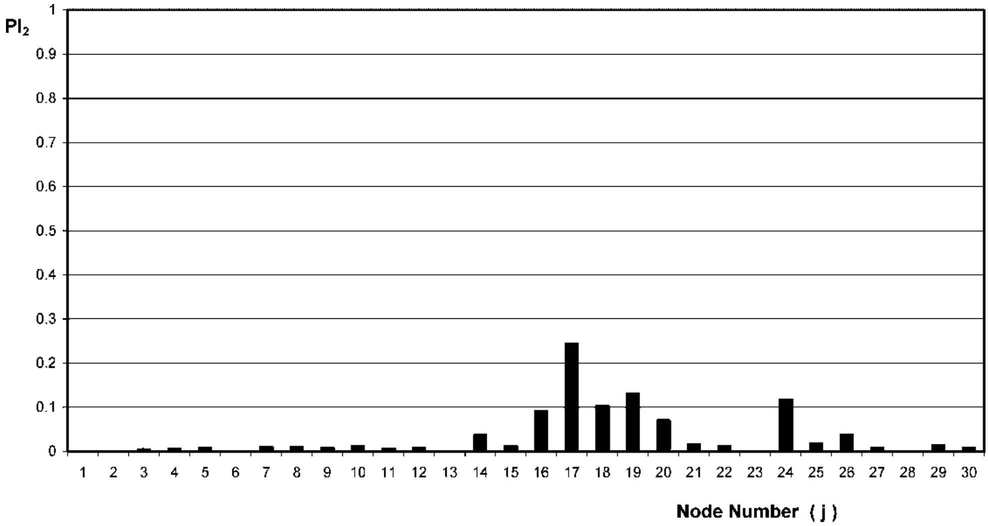

Figure 5. Conditioned probability mass functions of the nodes suitable for the allocation of the second monitoring station (EWT = 6 h).

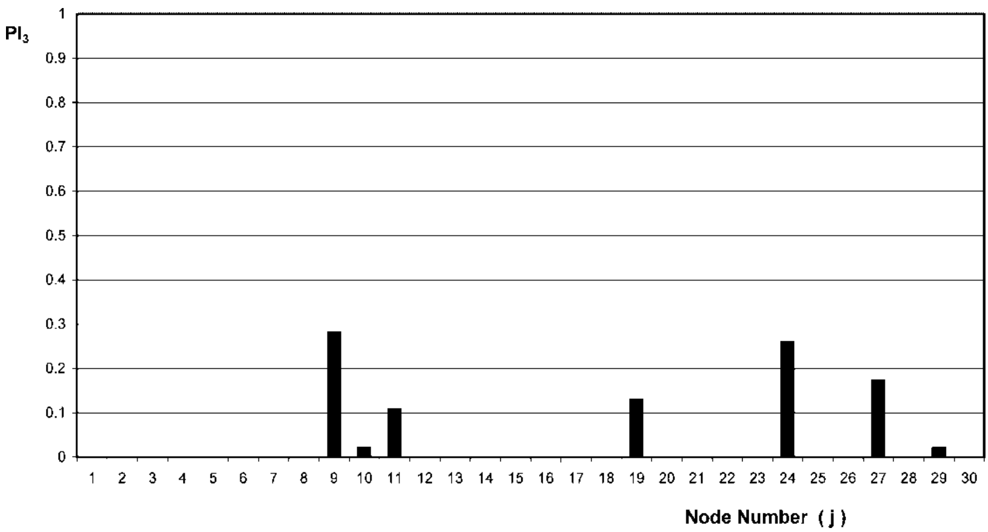

Figure 6. Conditioned probability mass functions of the nodes suitable for the allocation of the third monitoring station (EWT = 6 h).

Table 3. Allocation of the monitoring stations for different EWT.

EWT (h)	First station node	Second station node	Third station node	Fourth station node	Fifth station node
1	30	7	18	11	28
3	28	18	21	23	11
6	28	17	9	–	–

7. Conclusions

In this article, a stochastic approach has been proposed which aims at optimal allocations, within a water distribution network, of monitoring stations for the early detection of the intentional contamination of water. This approach, which is completely general, consists of the generation of equally-probable time-varying hydraulic scenarios, the application of a capable water quality model in order to evaluate the spreading of the contaminant from each of the potential source nodes and, finally, the allocation of the monitoring stations.

In the case that the monitoring stations are implemented one at a time, the allocation of the monitoring stations is accomplished by means of properly defined presence indexes. During a given time-varying hydraulic period, a node is suitable for the allocation of a monitoring station if it covers the greatest number of potential source nodes which release contaminant at the beginning of the period itself. 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 time-varying hydraulic periods. When the monitoring stations are allocated one at a time, which is the case in times of scarce financial resources, the first position coincides with the node which exhibits the greatest PI_1. Successive presence indexes have been defined, which aim to help in the allocation of the successive monitoring stations. An example has been considered in order to demonstrate the feasibility of the proposed approach for application in 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 (Cozzolino *et al.* 2006).

It has to be remarked that when two or more monitoring stations can be implemented simultaneously, the use of the presence indexes is expected to lead to a sub-optimal solution of the problem: in this case, a different optimisation method should be used, capable of taking into account the users' demand and uncertain hydraulic conditions.

References

Arora, G., Water quality modelling and monitoring in distribution systems. PhD thesis, University of Delhi, India, 2003.

Babayan, A., Kapelan, Z., Savic, D. and Walters, G., Least-cost design of water distribution networks under demand uncertainty. *ASCE J. Water Res. Plan. Manag.*, 2005, **131**(5), 375–382.

Bao, Y. and Mays, L.W., Model for water distribution system reliability. *ASCE J. Hydraul. Eng.*, 1990, **116**(9), 1119–1137.

Berry, J.W., Fleischer, L., Hart, W.E., Phillips, C.A. and Watson, J-P., Sensor placement in municipal water networks. *ASCE J. Water Res. Plan. Manag.*, 2005, **131**(3), 237–243.

Cozzolino, L., Mucherino, C., Pianese, D. and Pirozzi, F., Criteri per l'allocazione nelle reti idriche di stazioni di monitoraggio della qualità delle acque volte alla protezione da contaminazioni intenzionali, in *Proceedings of the Secondo seminario su "La ricerca delle perdite e la gestione delle reti di acquedotto"*, Perugia, Italy, 22 September 2005; Internal Report 1135, *University of Naples Federico II, Department of Hydraulic and Environmental Engineering* 2006 (in Italian).

Cozzolino, L., Pianese, D. and Pirozzi, F., DBPs control in water distribution systems through optimal chlorine dosage and disinfection stations allocation. *Desalination*, 2005, **176**(1–3), 113–125.

Gargano, R. and Pianese, D., Reliability as a tool for hydraulic networks planning. *ASCE J. Hydraul. Eng.*, 2000, **126**(5), 354–364.

Kessler, A., Ostfeld, A. and Sinai, G., Detecting accidental contaminations in municipal water networks. *ASCE J. Water Res. Plan. Manag.*, 1998, **124**(4), 192–198.

Kumar, A., Kansal, M.L. and Arora, G., Discussion of 'Detecting accidental contaminations in municipal water networks' by Kessler A., Ostfeld A. and Sinai, G. *ASCE J. Water Res. Plan. Manag.*, 1999, **125**(5), 308–309.

Mignosa, P., Sui problemi di verifica di reti di distribuzione idrica complesse. *Idrotecnica*, 1987, (6), 257–273 (in Italian).

Nilsson, K.A., Buchberger, S.G. and Clark, R.M., Simulating exposures to deliberate intrusions into water distribution systems. *ASCE J. Water Res. Plan. Manag.*, 2005, **131**(3), 228–236.

Ostfeld, A. and Salomon, E., Optimal layout of early warning detection stations for water distribution systems security. *ASCE J. Water Res. Plan. Manag.*, 2004, **130**(5), 377–385.

Pianese, D. and 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, in *Atti del Seminario sul Moto Vario nei Sistemi Acquedottistici*, Bari, Italy, 1994, 159–183 (in Italian).

Pianese, D., Pirozzi, F. and Taglialatela, L., Mathematical modeling of water quality in distribution systems, in *Proceedings of XXVII IAHR Congress, Theme A: 'Managing Water: Coping with Scarcity and Abundance'-San Francisco, USA, 10–15 August*, 1997a, pp. 875–880.

Pianese, D., Pirozzi, F. and Taglialatela, L., Influence of distinct processes on the quality of water in distribution systems, in *Proceedings of XXVII IAHR Congress, Theme A: 'Managing Water: Coping with Scarcity and Abundance'-San Francisco, USA, 10–15 August*, 1997b, pp. 881–886.

Pirozzi, F., Pianese, D. and d'Antonio, G., Water quality decay modelling in hydraulic pressure systems. *IWA Water Supply*, 2002, **2**(4), 111–118.

Rossman, L.A., Clark, R.M. and Grayman, W.M., Modeling chlorine residuals in drinking-water distribution systems. *ASCE J. Environ. Eng.*, 1994, **120**(4), 803–820.