Contamination Warning in Water Networks: General Mixed-Integer Linear Models for Sensor Location Design

Marco Propato[1]

Abstract: A mixed-integer linear program is proposed to identify optimal sensor locations for early warning against accidental and intentional contaminations in drinking water distribution systems. The general model can be applied to unsteady hydraulic conditions. Furthermore, it may accommodate different design objectives whose problem formulations vary only by the cost function coefficients while decision variables and linear constraints remain identical. Such a feature is very important since several requirements may be factors for practical design of warning systems. Linear constraint matrix properties show that the solution may often be found at the root (no branching). If not, a procedure is proposed to identify a significant set of discrete decision variables whose integrality constraints can be always relaxed. This result is coupled with good data preprocessing to minimize auxiliary continuous variables and constraints, allowing for efficient computation and increasing model applicability to large problems. The methodology is illustrated on a small and a midsize network.

DOI: 10.1061/(ASCE)0733-9496(2006)132:4(225)

CE Database subject headings: Water distribution systems; Water quality; Sensors; Contamination; Security; Terrorism; Optimization models.

Introduction

Regulatory agencies and water utilities have long been concerned about detection of accidental contaminations in drinking water distribution systems (DWDS) and are increasingly concerned of deliberate attacks. Conventionally, the detection of water quality deterioration in DWDS is mainly accomplished through routine distribution system monitoring. However, such practice is often inadequate for two main reasons. First, while the regulatory body provides some directions on locations and frequency of water sampling, it is not known if such water quality samples are representative of the entire system water quality. Second, although contamination events may have a low probability of occurrence, their impact on the population can be severe. The current practice may be acceptable to detect chronic, low intensity contamination problems but it becomes unacceptable when sudden, possibly high intensity changes occur in the water quality.

Although technological limits still exist, recent progress in the development of water quality sensor devices and communication networks permit a continuous water quality monitoring system to detect contaminations occurring in DWDS and therefore to provide early warnings to prompt the appropriate response strategy to eliminate or mitigate consumer exposure. The solution of the sensor location problem is a preliminary and necessary step for the development of adequate water quality monitoring systems.

In general, design requirements are multiple: they should provide maximum public health protection, cover all possible threats, should be affordable, and be able to identify the contamination source. Several difficulties make the problem challenging. Source location, start time, duration, and strength of contamination events are in general unknown. Contaminant fate and transport depend on water quality characteristics and several time varying water pathways governed by consumer water usage patterns and system operation. Moreover, assessing the health impact on consumers for a given intrusion event is a difficult task. In particular, the human response to a particular pathogenic or chemical dose is hard to predict.

The most recently proposed approaches for sensor location design in DWDS have different mathematical formulations and require different computational efforts; both depending on design objectives, simplifying assumptions, and solution methods adopted. They include integer programming (IP) models that are solved for steady state flow conditions (Lee and Deininger 1992; Kumar et al. 1997; Watson et al. 2004; Berry et al. 2005), and more general models applicable to unsteady flows solved using genetic algorithms or IP models (Berry et al. 2004; Ostfeld and Salomons 2004; Uber et al. 2004). A variety of objective functions have been introduced as surrogate of public health impact minimization: time to detection, population exposed, contaminated water delivered to consumers, detection coverage, extent of contamination, and failed detections. For practical design problems, a set of possible solutions should be given considering multiple objectives and their tradeoffs. If not, since optimality is in general model dependent, a sensor layout identified by focusing only on a specific design objective risks may be suboptimal with respect to other important requirements.

[1]Research Scientist, Hydraulics and Civil Engineering Research Unit, CEMAGREF, 50 Ave. de Verdun, Gazinet 33612, Bordeaux, France. E-mail: marco.propato@bordeaux.cemagref.fr

Note. Discussion open until December 1, 2006. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on August 24, 2005; approved on December 30, 2005. This paper is part of the ***Journal of Water Resources Planning and Management***, Vol. 132, No. 4, July 1, 2006. ©ASCE, ISSN 0733-9496/2006/4-225–233/$25.00.

The main goal of this work is to illustrate a design model, applicable to steady or unsteady flow conditions, that may accommodate different design objectives in a unique coherent mathematical formulation. Initially, a 0-1 polynomial programming problem is proposed for identification of optimal sensor locations to detect an ensemble of spatial and temporal random contamination events. This formulation is then transformed to a mixed-integer linear program (MILP) with the introduction of auxiliary continuous variables and constraints. Various design objectives may be described as identical MILP problems whose formulation differences are only the cost function coefficients and not the decision variables nor the constraints.

The model proposed here may reproduce equal solutions of most recent approaches when design characteristics—in particular design objective, network dynamical properties, and contamination ensembles—are set equal. However, compared to existing models, this formulation does not require data aggregation and its computational complexity may be lower as a result of the "integer friendly" properties of the constraint matrix, as well as of the reduced total number of auxiliary continuous variables and constraints. Moreover, it provides the designer with a general tool to generate solutions according to different objectives and to study their tradeoffs without employing disparate mathematical models and assumptions whose interrelationships may be more difficult to understand.

The paper is organized as follows. The next section describes the general framework and the assumptions made. The optimization problem is then formulated for different design objectives followed by a detailed comparison with other proposed design methods. The description of the computational algorithms to optimize problem solution and problem coefficients calculation precedes the application of the model to two illustrative network examples. Finally, concluding remarks are presented.

Methodology and Assumptions

Before describing the sensor location model, a general simulation framework is presented to define the different computational tasks and assumptions made to assess consumer exposure to contaminants. Although the general framework may be simple conceptually, the selection and verification of assumptions and models is a complex activity. For practical applications, the sensor design model should be viewed as a tool for generating alternative solutions to assist the iterative design process. The designer can and should study the robustness of the solutions under different model assumptions before recommending to the water utility a particular sensor layout. A similar scheme is considered in Propato and Uber (2004b). Below, the main results are summarized with a description of the particular assumptions made.

Network Hydraulic Model

An extended period simulation network model is used to estimate the time variation of pipe flow rates and storage tank water levels driven by time varying consumer water demands and system operation (Walski et al. 2003). Such models are increasingly used by water utilities to reflect typical current or projected operation over a specified time frame.

Contamination Event Ensemble

Modeling network contaminations is a difficult task and an open research issue. In general, a contamination event is described by joint probabilities of random variables that include location, start time, strength, and duration.

Accidental contaminations may be triggered by a pressure loss in correspondence of a cross connection. Under normal demand conditions, such events occur when there is a power loss or a main break. Thus, models require data or assumptions of the relevant joint probabilities (Lindley and Buchberger 2002; LeChevallier et al. 2003).

On the contrary, intentional contaminations are not in general associated with particular events that affect hydraulic dynamics. In principle, if a pump or a mobile-pressured tank is used, each network node can be considered as a potential intrusion site and only technical issues may limit contamination duration and strength.

Results below consider contamination events that may occur at any network locations and any start time, modeled as contaminant mass flow rate added to the downstream water of the node. These assumptions are consistent with intentional intrusions; a choice that is attractive for two main reasons. First, the solution of a design problem may dictate control of extreme behaviors. Second, intentional contaminations are simple to model without requiring system specific data allowing for model application to any network.

Contaminant Fate and Transport

Once in the DWDS, the contaminant mixes with the bulk water and if it does not react with water, its dynamics is determined by water flows, dilution, and mixing from the intrusion site to consumer nodes. In contrast, if the contaminant reacts with water, its fate may depend also on water quality properties. Some environmental variables are disinfectant residual, nutrients, pH, and temperature. For reactive contaminants a water quality model must be included to account for these types of reactions. For example, for pathogenic contaminations, several kinetic inactivation models have been proposed (Gyurek and Finch 1998) and multispecies water quality models have been employed to predict the outcome of disinfection kinetics (Shang et al. 2002; Propato and Uber 2004b).

Results here presented are for conservative contaminations. This model assumption may be again more consistent with the worst-case scenario of intentional attacks since the strength of the contaminant is not mitigated by reactions occurring in the water but depends only on water dilution and mixing. However, reactive contaminants that produce harmful byproducts also may be very important.

Contaminant Detection

The actual detection of a contamination event by the sensor network is a challenge itself. Although, the state of the art is rapidly changing, existing sensor devices have limited capabilities. Moreover, the information process algorithms needed to distinguish a contamination event from normal conditions can be complicated and may require the coordination of multiple sensor data.

J. Water Resour. Plann. Manage. 2006.132:225-233.

This paper assumes that sensors are able to instantaneously detect a contaminant as soon as the measured concentration exceeds a minimum accepted value and an alarm signal is raised.

Quantifying Contamination Damage

Assuming no consumer exposure after an alarm signal is raised, the desire is to identify a monitoring system minimizing health impact on consumers before detection. Unfortunately, the human response to a particular contaminant is hard to predict and is an open research area. For practical design problems, surrogate consumer exposure indicators have been considered by the research community (Watson et al. 2004). Here, design objectives to be minimized include:

- Time to detection: The time interval between the contamination start time and the detection time. Minimizing this objective allows for a more timely response to eliminate or limit consumer exposure.
- Population exposed: The total population exposed before detection. In the absence of population data, assuming that population is proportional to water use, it is possible to calculate the population fraction exposed as the ratio of the total daily water consumption at exposed households and the total daily water volume consumed in the DWDS. Population may be considered exposed in a variety of ways. For example, consumers are exposed if contaminant concentration, mass, or contaminated water volume ingested per capita exceed some accepted minimum value.
- Contaminated water consumed: This is the total volume of contaminated water delivered to consumers before detection. This quantity is similar to the previous one but it assumes that health effects may be proportional to the contaminated water quantity ingested.
- Contaminant mass consumed: This is the total contaminant mass that is consumed before detection. This quantity is simply given by the product of the water demand times the contaminant concentration. For this design objective, the effects on health are assumed proportional with the contaminant mass ingested.
- Number of failed detections: This is the total number of contamination events that go undetected by the monitoring system. Ideally, every contamination scenario should be detected, however for realistic designs with a limited number of sensors this goal may be difficult to achieve and a tradeoff solution has to be found.

Although not discussed here, there also exist other possible design requirements. The most significant one is to look for sensor locations whose monitoring data would allow the identification of the nodes or network areas that may be the source of contamination. This could significantly help water managers to take the appropriate steps to identify and eliminate the origin of the problem.

As previously discussed, practical design problems should study the set of possible solutions considering multiple objectives and their tradeoffs. However, here the goal is not to discuss multiobjective design but to develop a general optimization model for single-objective problems whose formulations and required computational efforts are independent of the design objective. Nonetheless, when multiobjective design is required, the identification of the solution Pareto-optimal front and the study of its properties is perhaps more easily carried out if each subproblem differs only by the cost coefficients and not by the particular mathematical form and simplifying assumptions.

Problem Formulation

This section initially presents the sensor location design formulation as a nonlinear 0-1 polynomial programming optimization problem for the general case of unsteady hydraulic conditions. Next, the problem is linearized and reformulated as a MILP problem.

0-1 Polynomial Programming Formulation

Each contamination scenario i, with $i=1,\ldots,q$, where q=total number of scenarios, is a contaminant injection that may vary in location(s), start time, duration, and strength.

An operational method is described to calculate objective function coefficients and constraints. Execute a water quality model for each contamination scenario i and do the following:

1. Record the set of all n_i potential sensor locations that are contaminated by scenario i and rank them according to the temporal order in which they are contaminated for the first time as the contaminant propagates through the network. If j indicates the rank position $(j=1,\ldots,n_i)$, then location with rank j is contaminated at time t_{ij} and $t_{i1}\leqslant t_{i2}\leqslant\ldots\leqslant t_{in_i}$;
2. Calculate the variation α_{ij} of the objective quantity to be minimized in the time interval $[t_{ij},t_{i,j+1}]$ between the contamination of two consecutively ranked locations.

For clarity, consider as an example the design objective to minimize time to detection, then $\alpha_{ij}=t_{i,j+1}-t_{ij}$. For other objective functions, calculate the quantities described in the section "Problem Objectives."

Let E_{ij} be the set of all potential sensor locations that have already been contaminated by scenario i at the time the contaminant reaches for the first time location with rank j. Recursively, E_{ij} includes all locations in $E_{i,j-1}$ and the additional node with rank j. By introducing the discrete variable δ_k to represent the absence $(\delta_k=1)$ or presence $(\delta_k=0)$ of a sensor at location k, consider the following function Z_i:

$$Z_i=\sum_{j=1}^{n_i}\alpha_{ij}\prod_{k\in E_{ij}}\delta_k=\alpha_{i1}\delta_{k_1^i}+\alpha_{i2}\delta_{k_1^i}\delta_{k_2^i}+\ \ldots\ +\alpha_{in_i}\delta_{k_1^i}\delta_{k_2^i}\cdot\ \ldots\ \cdot\delta_{k_{n_i}^i} \tag{1}$$

where the subscript k_j^i indicates the potential sensor location k having rank j for scenario i.

The function Z_i is actually the cost to be minimized. In fact, for a given sensor layout, assume that the first sensor detecting the contamination is placed at location k' having rank j'. By continuing with the previous design objective example, the time to detection is given by $(t_{ij'}-t_{i1})$ where t_{i1}=start time of contamination i. Now, all sets E_{ij} with $j<j'$ include potential sensor locations that do not have a sensor $(\delta_k=1)$, forcing all corresponding polynomials to be 1. In contrast, all sets E_{ij} with $j\geqslant j'$ include location k' $(\delta_{k'}=0)$, forcing all corresponding polynomials to be 0. Therefore, as desired, $Z_i=\Sigma_{j=1}^{j'-1}\alpha_{ij}=\Sigma_{j=1}^{j'-1}(t_{i,j+1}-t_{ij})=t_{ij'}-t_{i1}$. The meaning of coefficients α_{ij} becomes more clear: α_{ij}=objective function increase ΔZ_i between two sensor location solutions that raise an alarm signal for contamination i at locations having rank j and $j+1$ (detection times t_{ij} and $t_{i,j+1}$), respectively.

Consistent with the above definition, the last coefficient α_{in_i} is the weight given if scenario i goes completely undetected before time t_{is} when the first symptoms of exposure are observed on the population (e.g., hospitalization). When the design objective is time to detection minimization, its value is simply $\alpha_{in_i}=t_{is}-t_{in_i}$.

The ideal goal is to determine optimal sensor locations that minimize Z_i for every scenario (at optimality, $Z_i=0$ for all i) given a resource constraint that only a maximum number n_s of sensors is available. In practice, unless n_s is on the order of the total number of contamination potential sources, such a requirement is very restrictive: for realistic contamination ensembles and sensor location solutions there will always be some scenarios with non-null impact ($Z_i>0$). Therefore, a tradeoff must be found.

Although other approaches may be possible, a solution is found here by seeking the minimization of a weighted average cost function Z

$$\min_{\boldsymbol{\Delta}} Z(\boldsymbol{\Delta}) = \min_{\boldsymbol{\Delta}}\sum_{i=1}^{q}\omega_i Z_i(\boldsymbol{\Delta}) = \min_{\boldsymbol{\Delta}}\sum_{i=1}^{q}\omega_i\sum_{j=1}^{n_i}\alpha_{ij}\prod_{k\in E_{ij}}\delta_k \tag{2}$$

Subject to

$$\sum_{k\in S}\delta_k \geqslant N-n_s \tag{3}$$

$$\delta_k \in \{0,1\} \quad \forall \; k\in S \tag{4}$$

where $\boldsymbol{\Delta}$=vector of δ_k; the positive coefficient ω_i represents the weight assigned to contamination scenario i ($\Sigma_i\omega_i=1$); and S=network node set of all N potential sensor locations. In practice, coefficient ω_i can be estimated as the probability of contamination occurrence by analyzing historical data of unintentional contamination events or by considering the different physical measures that are in place to protect network nodes against intentional attacks.

Eq. (3) ensures that a maximum n_s sensors are built; and Eq. (4) defines the variables δ_k to be binary. If there is the additional requirement for each contamination i to be detected, a set of linear constraints can be added to problems (2)–(4) of the form $\Sigma_{k\in E_{in_i}}\delta_k\leqslant n_i-1$ (at least one $\delta_k=0$) for $i=1,\ldots,q$.

The objective function Z is not linear and thus problem (2)–(4) is difficult to solve using standard solution methods. However, by introducing auxiliary variables and linear constraints, Z can be rewritten in a linear form with respect to the decision variables.

MILP Problem Formulation

The objective function Z is expressed by the summation of positive terms, each including a 0-1 polynomial of the form $x_{ij}=\Pi_{k\in E_{ij}}\delta_k$. By noticing that set E_{ij} has j elements, variable x_{ij} can be transformed into a set of auxiliary linear constraints (Chang and Chang 2000)

$$x_{ij}\geqslant\sum_{k\in E_{ij}}\delta_k-j+1$$

$$x_{ij}\geqslant 0 \tag{5}$$

1. If $\Pi_k\delta_k=1$ then $\Sigma_k\delta_k=j$, thus Eqs. (5) become $x_{ij}\geqslant 1$ and $x_{ij}\geqslant 0$. Since $\alpha_{ij}\geqslant 0$, optimal x_{ij} is forced to be one.
2. If $\Pi_k\delta_k=0$ then $\Sigma_k\delta_k<j$, thus Eqs. (5) become $x_{ij}\geqslant e(e\leqslant 0)$ and $x_{ij}\geqslant 0$, forcing optimal x_{ij} to be zero.

Moreover, it is not necessary to require variables x_{ij} to be integral (although optimal values are binary), thus saving in computational effort to determine a solution.

If each 0-1 polynomial present in Z is substituted by the new continuous auxiliary variables x_{ij}, problems (2)–(4) becomes

$$\min_{\boldsymbol{\Delta},\boldsymbol{X}} Z=\min_{\boldsymbol{\Delta},\boldsymbol{X}}\sum_{i=1}^{q}\sum_{j=1}^{n_i}\omega_i\alpha_{ij}x_{ij} \tag{6}$$

Subject to

$$x_{ij}\geqslant\sum_{k\in E_{ij}}\delta_k-j+1 \quad \text{for } j=1,\ldots,n_i \text{ and } i=1,\ldots,q$$

$$\sum_{k\in S}\delta_k\geqslant N-n_s$$

$$x_{ij}\geqslant 0 \quad \text{for } j=1,\ldots,n_i \text{ and } i=1,\ldots,q$$

$$\delta_k\in\{0,1\} \quad \text{for } k=1,\ldots,N$$

where $\boldsymbol{X}$=vector of x_{ij}.

Objective function Z in Eq. (2) is initially described by the sum of $\Sigma_{i=1}^{q}n_i$ terms, each with a 0-1 polynomial. Since an auxiliary continuous variable and an inequality constraint (and a positivity constraint) is introduced for each polynomial, the total number of cost function terms can be significantly lowered by consolidating Z. In particular, one can: (1) sum over all terms with equal polynomial; and (2) eliminate all terms that have a null cost function coefficient, $\alpha_{ij}=0$. For example, consider two distinct contamination scenarios whose contaminant propagation, due to different hydraulics regimes, is described by node ID sequences $1\to2\to3\to4$ and $1\to2\to3\to5$, respectively. For this design problem, only five polynomials—and thus auxiliary variables and constraints—are necessary to set up the model (δ_1, $\delta_1\delta_2$, $\delta_1\delta_2\delta_3$, $\delta_1\delta_2\delta_3\delta_4$, and $\delta_1\delta_2\delta_3\delta_5$). In addition, if nodes 2 and 3 are contaminated simultaneously (within a water quality time step) in both scenarios, the term associated to polynomial $\delta_1\delta_2$ can be eliminated from Z. As it will become clear through illustrative examples, such variable and constraint reduction can be particularly significant to ease the computational effort required to solve practical design problems.

Problem Objectives

Mathematically identical problem formulations may be obtained if different design objectives are considered, whose differences are only the objective function coefficients. For each contamination scenario i, as contaminated locations are recorded and ranked temporally, α_{ij} are calculated as ΔZ_i in the time intervals $[t_{ij},t_{i,j+1}]$:

1. Time to detection (TD): $\alpha_{ij}=t_{i,j+1}-t_{ij}$.
2. Population exposed (PE): α_{ij}=new population exposed in $[t_{ij},t_{i,j+1}]$.
3. Contaminated water consumed (CWC): $\alpha_{ij}=\int_{t_{ij}}^{t_{i,j+1}}\Sigma_k d_k(t)\cdot H[c_k(t)-c_m]dt$ where $d_k(t)$=water demand at time t, the Heaviside function $H(x)$ is 1 if $x>0$, and 0 otherwise. The summation is over all consumer nodes.
4. Contaminant mass consumed (CMC): $\alpha_{ij}=\int_{t_{ij}}^{t_{i,j+1}}\Sigma_k c_k(t)\cdot d_k(t)dt$.
5. Number of failed detections: To maintain formulation uniformity, for every contamination scenario i let $\alpha_{ij}=1$ for $j=n_i$ and $\alpha_{ij}=0$ otherwise.

The objective function Z in Eq. (2) becomes $Z=\Sigma_i\omega_i\Pi_{k\in E_{in_i}}\delta_k$ where each polynomial includes all n_i potential locations that may be contaminated by scenario i. The contamination is not detected when none of these locations have a sensor. Then cost function Z increases since the polynomial is one ($\delta_k=1$ for all $k\in E_{in_i}$),

J. Water Resour. Plann. Manage. 2006.132:225-233.

otherwise Z remains unchanged. Problem linearization is simpler since only one continuous variable has to be introduced for each scenario.

For all objective designs, to maintain consistency with the description given when considering time to detection minimization, coefficients α_{in_i} are calculated as ΔZ_i in the time intervals $[t_{in_i}, t_{is}]$.

Comparison with Alternative Approaches

This section discusses and compares the proposed design method and the most recent and significant approaches.

The key feature of problem (6) with respect to existing models is that not only it records the set of all contaminated locations by a contamination event, but it also records the temporal order in which they are contaminated. This allows the effects of a contamination to be described as a summation of positive contributions as the contaminant propagates through the network.

The method proposed by Ostfeld and Salomons (2004) defines a maximum level of service (LOS), the contaminated water volume consumed prior to detection. The optimal solution is the one that maximizes the number of detected contaminations within the selected LOS, assigning equal cost to all detected scenarios regardless of the actual contaminated water volume consumed at the time of detection. In principle, different solutions may be obtained for different LOS. Thus, for practical problems, additional criteria have to be used to identify the tradeoff solution. Moreover, the assignment of the prespecified LOS requires some a priori information that may not be always available. In contrast, the method described here can identify the optimal sensor layout by minimizing the actual contaminated water volume consumed before detection. Thus, contaminations with different contaminated water consumption prior to detection contribute differently to the cost function and the optimal sensor solution is unique. If genetic algorithms were to be used to solve problems (2)–(4) using cost Z as fitness function, identification of a sensor solution would not be more computationally difficult than in Ostfeld and Salomons; except that also the temporal order of location contamination has to be recorded when constructing the domains of detection for each scenario. However, as discussed in the next section, the linear relaxation in Eq. (6) may bring significant computational benefits for practical problems.

Based on the model introduced by Berry et al. (2005), Watson et al. (2004) have been the first to propose alternative design objectives for steady flow conditions. The model proposed here cannot only describe multiple objectives in the general case of unsteady flow conditions, but its mathematical structure is simpler, and uniform for different design objectives.

The MILP problem proposed by Berry et al. (2004) is philosophically similar to the one illustrated here and may also accommodate various objectives. For equal designs, the two models have the same sensor solutions and costs; and without any reduction of the auxiliary continuous variable set, both problems are parameterized with a similar number of decision variables and constraints [although formulation (6) does not require a dummy variable and related set of constraints for each contamination scenario]. However, as discussed above, taking into account node contamination order allows for a significant variable and constraint reduction that is not foreseen in the Berry et al. formulation. For example, the two contamination scenarios described previously require ten continuous variables and constraints for Berry et al. compared to five of problem (6). More importantly, the different mathematical structure of the linear constraints, derived from the different meaning of the decision variables and of the cost function coefficients, facilitates the investigation of the problem scalability properties.

Total Unimodularity and Problem Scalability

Problem (6) can be solved using standard branch and bound techniques. The number of discrete variables depends on the size of the network and it is very well known that IP problems are often $\mathcal{NP}$ hard. Sensor placement design scalability in DWDS has also been recently discussed in Berry et al. (2005) and Watson et al. (2005).

Branch and bound algorithms always initially solve the relaxed problem and if the solution is already integral they stop. For this specific integer problem, as also noticed in other sensor location design models (Berry et al. 2004), integral solutions are often already found at the root (no branching). This somehow striking result is due to the properties of the feasible space polyhedron defined by the linear constraints. It turns out that most of the polyhedron vertices are integral. Thus, it is very likely that the solution of the relaxed problem, which has to be a vertex, is integral. However, since in general not all polyhedron extremes are integral, a nonintegral optimal solution may be possible. For this situation, before starting the branching procedure, the computational effort could be reduced if some of the integrality constraints could be relaxed by exploiting the integer "friendly" properties of the feasible space.

Consider the following definition: a matrix whose elements are 0, 1 or -1 is called totally unimodular (TU) if the determinant of every submatrix is 0, 1, or -1. If the constraint matrix of an IP problem is TU and the right hand side constraint vector is integral, then all polyhedron vertices are integral. It follows that an IP problem with such properties can be efficiently solved with a standard simplex algorithm (Papadimitriou and Steiglitz 1998).

Now consider if, or to what extent, such a property applies to the problem examined here. From Eq. (6), linear constraints (excluding positivity and integrality) are rewritten in the matrix form $[\boldsymbol{A_1}\,\boldsymbol{A_2}][\boldsymbol{X}\ \boldsymbol{\Delta}]^T \leqslant \boldsymbol{B}$, where $\boldsymbol{A_1}$ is a $(n_x+1)\times n_x$ matrix (n_x the size of $\boldsymbol{X}$), $\boldsymbol{A_2}$ a $(n_x+1)\times N$ matrix, and $\boldsymbol{B}$ a n_x+1 vector of integer numbers. As an example, linear constraints (5) for a contamination scenario described by node ID sequence $11 \rightarrow 7 \rightarrow 23$ are written as

$$\left(\begin{array}{ccc|ccc} \multicolumn{3}{c|}{\overbrace{\hphantom{-1\ \ 0\ \ 0}}^{\boldsymbol{A_1}}} & \multicolumn{3}{c}{\overbrace{\hphantom{1\ \ 0\ \ 0}}^{\boldsymbol{A_2}}} \\ -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & 1 & 1 & 1 \end{array}\right) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \delta_{11} \\ \delta_7 \\ \delta_{23} \end{pmatrix} \leq \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$$

In general, matrix $\boldsymbol{A_1}$ is a negative identity matrix, except the row describing the resource constraint that contains all zeros. Matrix $\boldsymbol{A_2}$ is a (0,1) matrix, except the resource constraint row that contains all -1. Its non-null row entries are 1 for locations in E_{ij}, the set of all contaminated potential sensor locations when scenario i reaches location with rank j. An important property TU matrices is that if matrix $\boldsymbol{A}$ is TU $[\boldsymbol{I}\,\boldsymbol{A}]$ and $-\boldsymbol{A}$ are also TU, where

J. Water Resour. Plann. Manage. 2006.132:225-233.

- Let $\mathcal{A} = \{$set of columns of a generic matrix $\boldsymbol{A}\}$;

1. ◇ Sort columns of $\boldsymbol{A_2}$ according to the ranking order from the contamination scenario that affects the highest number of potential sensor locations,

   ◇ Initialize sets $\mathcal{A}_2^c = \mathcal{A}_2$ and $\mathcal{A}_2^d = \emptyset$;

2. **For** row $i = n_x, \ldots, 1$ of $\boldsymbol{A_2}$:

   ◇ Remove the minimum set $\mathcal{D}$ of columns to obtain a $C1P$ row,

   ◇ Update sets $\mathcal{A}_2^c = \mathcal{A}_2^c - \mathcal{D}$ and $\mathcal{A}_2^d = \mathcal{A}_2^d \cup \mathcal{D}$;

3. Sort n_d columns of $\boldsymbol{A_2^d}$ according to their initial position in $\boldsymbol{A_2}$;

4. **For** column $\boldsymbol{c_j}$ of $\boldsymbol{A_2^d}$, $j = 1, \ldots, n_d$:

   ◇ Check if it can be placed back as first column in $\boldsymbol{A_2^c}$ without destroying the $C1P$, **if yes**:

   • Update sets $\mathcal{A}_2^c = \mathcal{A}_2^c \cup \boldsymbol{c_j}$ and $\mathcal{A}_2^d = \mathcal{A}_2^d - \boldsymbol{c_j}$;

5. $\boldsymbol{A_2} = [\boldsymbol{A_2^c}\ \boldsymbol{A_2^d}]$

Fig. 1. Heuristic procedure to reduce integrality constraints

$\boldsymbol{I}$=identity matrix. Therefore, with formulation (6), one may focus only on the TU properties of matrix $\boldsymbol{A_2}$.

Suppose you can identify a submatrix $\boldsymbol{A_2^c}$ of columns of $\boldsymbol{A_2}$ that is TU. Then, problem (6) can be solved by relaxing the integrality constraints on the decision variables $\boldsymbol{\Delta^c}$ associated with constraint coefficients $\boldsymbol{A_2^c}$. In fact, problem linear constraints can be also written as $[\boldsymbol{A_1}\ \boldsymbol{A_2^c}][X\ \Delta^c]^T \leqslant (\boldsymbol{B} - \boldsymbol{A_2^d}\Delta^d)$ where $\boldsymbol{A_2^d}$ and Δ^d=remainder of $\boldsymbol{A_2}$ and Δ, respectively. The right-hand side of the inequality is ensured to be integral if decision variables Δ^d are constrained to be integral. It follows that since $[\boldsymbol{A_1}\ \boldsymbol{A_2^c}]$ is TU, for any integral Δ^d all vertices of the polyhedron defined by the inequality constraints are integral.

Unfortunately, the problem to identify the biggest TU submatrix $\boldsymbol{A_2^c}$ of columns of $\boldsymbol{A_2}$ is known to be $\mathcal{NP}$ complete (Garey and Johnson 1979). However, some peculiar properties for this problem can be exploited. In what follows, the resource constraint row of all -1 is eliminated. Later, it will be shown that the TU properties of $\boldsymbol{A_2}$ also hold when this row is included.

If the network has no loops, each contamination event occurs at a single point and flow velocities are constant, then network nodes are always contaminated in the same temporal order. For this simple case, there always exists a permutation of columns of $\boldsymbol{A_2}$ such that the manipulated matrix has the consecutive one property ($C1P$): each row has only one block of consecutive ones. Such a matrix is TU. For the general case, $\boldsymbol{A_2}$ may not be TU and its structure depends on the network hydraulic dynamical properties and modeled contamination events. From a system dynamics point of view, the TU property may be lost all the times there is a time inversion of contaminated locations, for instance location k is contaminated before location k' for a contamination event and vice versa for another. However, for realistic networks and contamination ensembles, one can expect that a significant part of $\boldsymbol{A_2}$ still fulfills the $C1P$. A relatively simple and quick heuristic procedure to identify a $C1P$ submatrix of $\boldsymbol{A_2}$ is described in Fig. 1.

With step 1 the columns of $\boldsymbol{A_2}$ are sorted in a way that the submatrix corresponding to the constraints of the chosen contamination (typically one at the water source that would affect all potential sensor locations) is triangular and thus $C1P$. When considering the constraints (rows) of other contamination scenarios, any column permutation that is not cyclic to the left (append a column as the first matrix column) would create at least one non-$C1P$ row in the triangular part. Thus, step 2, going from one row to the next, can only remove columns to maintain the $C1P$ for all rows already $C1P$. The algorithm moves upward because to reduce the block number in each contamination scenario, it is always better to start from the row that contains the highest number of ones. The $C1P$ by cyclic permutations is evaluated at steps 3 and 4. Finally, in step 5 matrix $\boldsymbol{A_2}$ is reassembled and problem (6) can be solved by imposing integrality constraints only on the variables associated with $\boldsymbol{A_2^d}$.

Now consider the addition of the resource constraint row of all -1. The determinant of *any* submatrix $\boldsymbol{M}$ of $\boldsymbol{A_2^c}$ containing elements of this row is equal to the negative determinant of a matrix $\boldsymbol{M'}$ with the same elements of $\boldsymbol{M}$ except for all consecutive +1 at the resource constraint row. Matrix $\boldsymbol{M'}$ is TU since it has the $C1P$. Thus, any $\boldsymbol{M}$ is also TU $[|\det(\boldsymbol{M})| = |\det(\boldsymbol{M'})| = 0, 1]$. It follows that $\boldsymbol{A_2^c}$ is also TU with the resource constraint row.

Unfortunately, such a heuristic procedure does not guarantee that $\boldsymbol{A_2^c}$ is the biggest possible TU part of $\boldsymbol{A_2}$, which would require increased computation. However, as it will be more evident through the illustrative examples, the dimension n_d of $\boldsymbol{\Delta^d}$ can already be significantly lower than the initial number N of discrete decision variables, facilitating the branching step. If an optimal solution is not already found at the root, this proposed procedure may increase problem (6) computational limits.

Data Collection Algorithms

One immediate way to collect problem coefficients is to run a water quality simulation for each contamination scenario. Considering that these events may vary in network location, start time, duration, time profiles, and strength, the collection of data for a meaningful ensemble may require a prohibitive computational effort. Moreover, in the design process of practical problems, the robustness of proposed solutions should be studied under different model assumptions; for example different water demand dynamics or contamination characteristics. For a more efficient analysis, it would then be advisable to decouple the different computational tasks previously described.

Several water quality models calculate water quality dynamics in a very efficient way. For example, with *EPANET* programmer's toolkit (Rossman 2000) it is possible to execute the hydraulic model first and independently evaluate water quality dynamics a posteriori. When dealing with many contamination scenarios this is a significant computational advantage since there is no need to run a full hydraulic-water quality simulation for every modeled contamination event.

If the contaminant is conservative, or its reaction rate is first order, it is possible to relate the contaminant concentration time series $c_i(t)$ at node i to the contaminant mass or concentration injections $u_j(k)$ done at node j at previous time steps k through a general linear input–output (I/O) model of water quality in DWDS (Propato and Uber 2004a)

$$c_i(t) = \sum_{j=1}^{N} \sum_{k=1}^{t} \theta_{ij}^k(t) u_j(k) \quad (7)$$

where the summations are extended to all possible N contamination nodes and time steps. The operational definition of impact coefficient $\theta_{ij}^k(t)$ is the concentration at time t and consumer node i, from a constant unit contamination done at node j between times $t=k$ and $t=k+1$ (all other contamination sources equal to zero). For a conservative contaminant, the impact coefficients account for water mixing while for a first order reactive species such terms are mitigated/amplified by the decay/growth [Propato and Uber (2004a) provide an example calculation].

Contamination Event Ensemble

For a conservative (or first order reactive) contaminant and detection limit $c_m=0$, I/O relations (7) suggest that the contamination event ensemble could be made of single point contaminations with unit time step duration and different start times that scan the entire possible contamination time window for each potential contamination source location. A sensor layout minimizing the cost function for all these "unit" contamination events will likely minimize the cost with respect to contaminations of different characteristics, since the effects on consumers from these more complicated profiles can always be written as a linear combination of effects from unit contamination events.

The considerations above are not always valid if $c_m>0$. However, to improve data collection, Eq. (7) can be used for these cases. Once impact coefficients are calculated for all consumer nodes and possible contamination sources, contaminant concentration profiles at any node i due to a modeled contamination scenario are obtained by multiplying contamination source injection rates with the impact coefficients without the need to execute a water quality simulation for every scenario. Among other computational factors, such an approach is convenient when the contamination ensemble includes a number of scenarios greater than the number of water quality simulations necessary to calculate the impact coefficient time series. This number depends on time granularity of contaminations profiles, the time window when contaminations can occur, and potential contamination locations. However, it does not depend on contamination start time, time profile, duration, and strength.

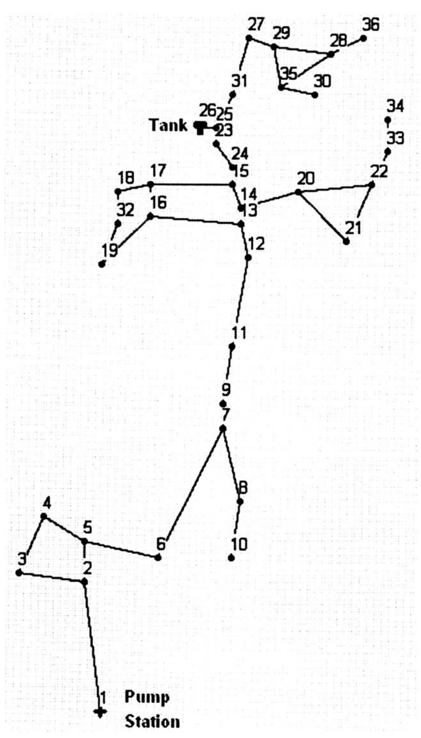

Fig. 2. Diagram of network example 1

Table 1. Example of Cost Function Coefficients Calculation

Rank	Node ID	TD (min)	PE (%)	CWC (1,000 L)	CMC (kg)
1	13	0	0.00	0.00	0.0
2	14	10	1.24	0.1	0.07
3	15	15	1.86	0.2	0.11
4	24	20	5.27	0.6	0.25
5	23	30	7.74	1.3	0.47
—	—	—	—	—	—
29	34	2,025	86.37	1,224.0	5.25
30	36	2,125	86.68	1,406.5	5.47
—	—	2,880	86.68	2,001.8	6.01

Model Application

A small and a midsize network are presented as illustrative example of the design methodology proposed.

Example 1

The network is a modified version of EPANET's example 2 network [for more details, see Boccelli et al. (1998)]. It comprises 34 consumer nodes whose daily periodic demands are fulfilled by a pump station that operates on a 24 h cycle and a storage tank. This network has been chosen because it is relatively small but its hydraulic dynamical behavior is complex enough for the presence of pipe loops and abrupt changes of flow direction occurring in the main trunk when the pump station is off and water demand is supplied only by the reservoir tank (See Fig. 2).

Contamination source and potential sensor locations are all network nodes. The contamination event ensemble is made of conservative contaminations of unit time step duration of 15 min at a constant rate of 1 kg/min, with equal probability of occurrence (ω_i=cost $\forall$ i) at locations and start times during a normal day of operation. Detection limit is set to $c_m=0$ and water quality simulation and sampling time steps are set to 5 min. For big networks, a statistically significant ensemble can be generated. However, a complete numeration has been performed here. The maximum detection time t_{is} is set to 2 days from each contamination start time.

Table 1 shows an example calculation of cost function coefficients and polynomials of a contamination i occurring at node 13. For example, if time to detection has to be minimized, $\alpha_{i1}=t_{i2}-t_{i1}=10-0=10$ min, $\alpha_{i2}=15-10=5$ min, $\ldots, \alpha_{i30}=2,880-2,125=755$ min (similarly for other design objectives), the cost function becomes $Z_i=\alpha_{i1}x_{i1}+\alpha_{i2}x_{i2}+\ldots+\alpha_{i30}x_{i30}$ and $x_{i1}\geqslant\delta_{13}-1+1$(this variable could actually be eliminated by noting that $x_{i1}=\delta_{13}$ at optimality), $x_{i2}\geqslant\delta_{13}+\delta_{14}-2+1,\ldots,$ and $x_{i30}\geqslant\delta_{13}+\delta_{14}+\ldots+\delta_{34}+\delta_{36}-30+1$.

J. Water Resour. Plann. Manage. 2006.132:225-233.

Table 2. Optimal Sensor Locations for Network 1

n_s	DT	PE	CWC	CMC
1	32	13	14	13
2	32-35	7-14	9-15	12-30
3	32-35-34	7-13-23	6-12-23	12-30-32
4	7-17-13-23	5-11-17-23	12-30-32-34	10-26-32-35

By proceeding as above for all 3,456 $(36 \cdot 24 \cdot 60/15)$ contamination events, the initial number of polynomials is 52,740. Of these, only 1,599 are different. The total number of continuous variables and constraints to be introduced can be reduced to 1,507 by recognizing that 92 polynomials have a null cost function coefficient (potential sensor locations contaminated simultaneously).

The initial number of discrete variables is 36, the entire potential sensor location set. Although the branch and bound algorithm could always find a naturally integral solution at the root for any resource constraint n_s and objective design, the procedure described in Fig. 1 shows that at least 20 discrete variables can always be assumed continuous.

As summarized in Table 2 for different n_s, optimal sensor locations depend on the design objective. Fig. 3 shows the decrease of the normalized cost function $Z/Z_{\max}$, where $Z_{\max}$=maximum cost function obtained when no sensor is present, as a function of n_s. The rate of decrease of Z is also dependent on the design objective. In particular, the dramatic gain obtained as soon as a sensor is put in the network when the CWC is to be minimized, is due to the presence of the reservoir tank and $c_m=0$. If there are no sensors, once the contaminant goes in the reservoir tank it contaminates a very large water volume that is later consumed, thus amplifying the impact on consumers. As soon as a sensor is placed in the system, such a large contaminated water volume is no longer consumed because an alarm signal is raised first. For a different design objective, for example CMC, the contaminant mass consumed coming out from the reservoir tank is not that significant compared to other regimes. Thus, the reduction of CMC is relatively less important when the first sensor is in place. Although there is ample room to argue about assumptions and results, the main point here is to show that practical sensor location design needs to take into account different objectives and should study their tradeoffs.

Example 2

This example is presented to study model applicability to practical bigger problems. A midsize real DWDS is considered (Propato and Uber 2004b). It contains 474 consumer nodes, a treatment plant, six storage tanks, and seven locations where water is withdrawn by neighboring DWDS. The nonstationary hydraulic model has been calibrated using real historical data. Water flows in main trunk lines are unusually high due to the water sold to other utilities, which accounts for two thirds of the total water volume entering the system from the reservoir (approximately $17 \cdot 10^6$ L/day).

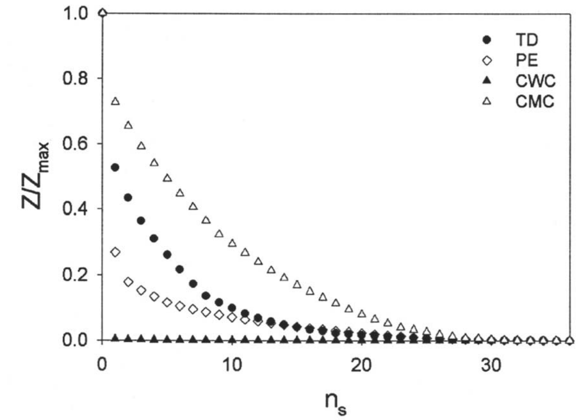

Fig. 3. Optimal normalized cost functions versus n_s

Contamination source locations and potential sensor locations are all 488 network nodes. The contamination event ensemble is made up of a complete numeration of equally likely conservative contaminations with unit time step duration of 30 min at a constant rate of 1 kg/min. Water quality sampling time step is 15 min while other simulation parameters equal those of previous examples.

The initial cost function for 23,424 $(488 \cdot 24 \cdot 60/30)$ contamination scenarios is described by the summation of 1,256,739 terms. However, only 165,645 continuous variables and constraints are necessary, almost ten times smaller than the initial number. The procedure to identify the $C1P$ constraint submatrix shows that of the initial 488 discrete variables, at least 215 can always be assumed continuous, improving the branch and bound algorithm efficiency. Although not extensively tested as for the previous example network, an optimal solution is also very often found at the root for different n_s and design objectives, while the optimal sensor layout varies depending of the design objective considered.

Table 3 summarizes some of the more significant design parameters for the two illustrative examples to demonstrate the validity of the method with respect to computational issues.

Conclusions

Several sensor location models for water quality monitoring have recently been proposed. The main contribution of this work is to provide a consistent simulation framework to systematically and quantitatively assist the iterative design process. A general formulation for optimal sensor placement expresses different design objectives in structurally identical mixed-integer linear programs. The proposed facility location model includes previous related work in a coherent unique formulation without adding, if not reducing, mathematical and computational complexity. With respect to previous approaches, the key feature of such a general scheme is to record not only the set of contaminated locations but

Table 3. Design Parameters Summary

Network example	N	q	$\Sigma_i n_i$	n_x	Total linear construction[a]	Relax. var.	Integr. var.
1	36	3,456	52,740	1,507	$1,507+1$	20	16
2	488	23,424	1,256,739	165,645	$165,645+1$	215	273

[a]Excluding positivity constraints.

J. Water Resour. Plann. Manage. 2006.132:225-233.

also the time order in which they are contaminated.

The problem solution exhibits nice integer friendly properties in the sense that an integral solution is often already found at the root before the branching step. However, when branching is needed, unimodularity properties of the constraint matrix may allow us to always keep relaxed several integrality constraints, thus improving computational efficiency and pushing model applicability limits.

These properties were demonstrated on a small and a midsize network. In general, different optimal sensor layouts are obtained if different design objectives are considered. These results suggest that for practical design problems a set of solutions should be given considering multiple objectives and their tradeoffs.

An attractive research topic is to consider the contamination source identification problem. Clearly, valuable water quality sampled data to identify contamination sources are only available with a good selection of sensor locations. However, the methodology proposed here cannot be easily extended to this design objective.

Acknowledgment

This research was supported by Marie Curie International Reintegration Grants within the 6th European Community Framework programme (Contract Number: MIRG-CT-6–2005–014825.

Notation

The following symbols are used in this paper:

$\boldsymbol{A}_2^{c(d)}$ = inequality constraints matrix for continuous (discrete) variables;

$\boldsymbol{B}$ = inequality constraint vector;

$c_k(t)$ = contaminant concentration at node k and time t;

c_m = minimum accepted contaminant concentration;

$d_k(t)$ = water demand at node k and time t;

E_{ij} = set of all potential sensor locations already contaminated at time t_{ij};

N = total number of potential sensor locations;

n_i = total number of potential sensor locations contaminated by scenario i;

n_s = maximum number of sensors available;

n_x = total number of x_{ij} variables;

q = total number of contamination scenarios;

S = set of all potential sensor locations;

t_{ij} = first time contamination i hits potential sensor location with rank j;

t_{is} = time first contamination symptoms are observed;

$u_j(k)$ = contaminant injection at node j and time k;

$\boldsymbol{X}$ = vector of x_{ij};

x_{ij} = auxiliary continuous variable;

Z = design objective function to be minimized;

Z_i = design objective function for contamination scenario i;

α_{ij} = cost function coefficient;

$\boldsymbol{\Delta}^{c(d)}$ = vector of continuous (discrete) δ_k;

δ_k = discrete variable indicating absence/presence of sensor at location k;

$\theta_{ij}^k(t)$ = impact coefficient time series; and

ω_i = probability given to contamination scenario i.

References

Berry, J., Hart, W., Phillips, C., and Uber, J. (2004). "A general integer-programming-based framework for sensor placement in municipal water networks." *Proc., World Water and Environmental Research Conf.*, Salt Lake City.

Berry, J., Hart, W., Phillips, C., and Watson, J. (2005). "Scalability of integer programming computations for sensor placement in water networks." *Proc., World Water and Environmental Research Congress*, Anchorage, Alaska.

Boccelli, D., Tryby, M., Uber, J., Rossman, L., Zierolf, M., and Polycarpou, M. (1998). "Optimal scheduling of booster disinfection in water distribution systems." *J. Water Resour. Plan. Manage.*, 124(2), 99–111.

Chang, C.-T., and Chang, C.-C. (2000). "A linearization method for mixed 0-1 polynomial programs." *Comput. Oper. Res.*, 27(10), 1005–1016.

Garey, M., and Johnson, D. (1979). *Computers and intractability, A guide to the theory of NP completeness*, Freeman, New York.

Gyurek, L., and Finch, G. (1998). "Modeling water treatment chemical disinfection kinetics." *J. Environ. Eng.*, 124(9), 783–793.

Kumar, A., Kansal, M., and Arora, G. (1997). "Identification of monitoring stations in water distribution system." *J. Environ. Eng.*, 123(8), 746–752.

LeChevallier, M., Gullick, R., Karim, M., Friedman, M., and Funk, J. (2003). "The potential for health risks from intrusion of contaminants into the distribution system from pressure transients." *J. Water Health*, 1(1), 3–14.

Lee, B., and Deininger, R. (1992). "Optimal locations of monitoring stations in water distribution system." *J. Environ. Eng.*, 118(1), 4–16.

Lindley, T., and Buchberger, S. (2002). "Assessing intrusion suscept ibility in distribution systems." *J. Am. Water Works Assoc.*, 94(6), 66–79.

Ostfeld, A., and Salomons, E. (2004). "Optimal layout of early warning detection stations for water distribution systems security." *J. Water Resour. Plan. Manage.* 130(5), 377–385.

Papadimitriou, C., and Steiglitz, K. (1998). *Combinatorial optimization: Algorithms and complexity*, Chap. 13, Dover, Mineola, N.Y.

Propato, M., and Uber, J. (2004a). "Linear least-squares formulation for operation of booster disinfection systems." *J. Water Resour. Plan. Manage.*, 130(1), 53–62.

Propato, M., and Uber, J. (2004b). "Vulnerability of water distribution systems to pathogen intrusion: How effective is a disinfectant residual?" *Environ. Sci. Technol.*, 38(13), 3713–3722.

Rossman, L. (2000). *EPANET user's manual*, Risk Reduction Engineering Laboratory, U.S. Environmental Protection Agency, Cincinnati.

Shang, F., Propato, M., and Uber, J. (2002). "Extension of EPANET for multispecies modeling in water distribution system." *Proc., ASCE Environmental and Water Resources Institute Conf.*, ASCE, Roanoke, Va.

Uber, J., Janke, R., Murray, R., and Meyer, P. (2004). "Set covering formulation for locating water quality sensors in distribution systems." *Proc., World Water and Environmental Research Conf.*, Salt Lake City.

Walski, T., Chase, D., Savic, D., Grayman, W., Beckwith, S., and Koelle, E. (2003). *Advanced water distribution modeling and management*, Haestead, Waterbury, Conn.

Watson, J., Greenberg, H., and Hart, W. (2004). "A multiple-objective analysis of sensor placement optimization in water networks." *Proc., World Water and Environmental Research Conf.*, Salt Lake City.

Watson, J., Hart, W., and Berry, J. (2005). "Scalable high-performance heuristics for sensor placement in water distribution networks." *Proc., World Water and Environmetal Research Congress*, Anchorage, Alaska.