Journal of Cleaner Production 29-30 (2012) 28–37

Contents lists available at SciVerse ScienceDirect

Journal of Cleaner Production

journal homepage: www.elsevier.com/locate/jclepro

A rule-based decision support system for sensor deployment in small drinking water networks

Ni-Bin Chang [a,*], Natthaphon P. Pongsanone [a], Andrew Ernest [b]

[a] Civil, Environmental, and Construction Engineering Department, University of Central Florida, Orlando, FL 32816, USA
[b] Center for Water Resource Studies, Western Kentucky University, Bowling Green, KY 42101, USA

ARTICLE INFO

Article history:
Received 2 May 2011
Received in revised form
12 December 2011
Accepted 8 February 2012
Available online 23 February 2012

Keywords:
Decision support system
Drinking water
EPANET
Graph theory
Rule-based sensor deployment
Systems analysis

ABSTRACT

The physical layout of drinking water utilities makes them inherently vulnerable to contamination incidents caused by routine operations. These contaminations present environmental health concerns including but not limited to total trihalomethanes, lead, and chlorine residual issues. To achieve the goal of cleaner production, sensor placement in municipal drinking water networks in response to possible public health threats has become one of the most significant challenges currently facing drinking water utilities, especially in small-scale communities. Long-term monitoring is needed to develop modern concepts and approaches to risk management for these utilities. We developed a Rule-based Decision Support System (RBDSS), a methodology to generate near-optimal sensor deployment strategies with low computational burden, such as those we often encountered in large-scale optimization analyses. Three rules were derived to address the efficacy and efficiency characteristics of such a sensor deployment process: (1) intensity, (2) accessibility, and (3) complexity rules. Implementation potential of this RBDSS was assessed for a small-scale drinking water network in rural Kentucky, United States. Our case study showed that RBDSS is able to generate the near-optimal sensor deployment strategies for small-scale drinking water distribution networks relatively quickly. The RBDSS is transformative and transferable to drinking water distribution networks elsewhere with any scale.

1. Introduction

Drinking water distribution systems are inherently vulnerable to accidental or intentional water contamination incidents. Because those networks are large, spatially distributed, and complicated infrastructures, the possibility of human-related influences is significantly high (Buckle, 2000; Haestad et al., 2003; Karamouz et al., 2010). For example, in developing countries like Guatemala, inadequate clean water and waterborne bacterial infection among young children are the cause of disease and productivity losses equivalent to 2% of gross domestic product (Norstrom, 2007; Tune and Elmore, 2009); therefore, the numerous studies being conducted for vulnerability assessment, risk reduction, monitoring sensor network, and contamination warning system are rigorous. A recent case study of vulnerability assessment of water supply system components in a major city included five criteria: distribution, spread, visibility, exposure, and recovery. It found that the failure of water distribution networks and water treatment plants generated the highest human losses among other water supply failures (Karamouz et al., 2010). Because these incidents often have severe immediate and long-term human health consequences, drinking water distribution networks require intensive monitoring and security considerations using real-time early warning systems (EWS) (Clark and Deininger, 2001; National Research Council, 2002).

To build a functional EWS, a sensor location system should be designed to satisfy multiple criteria with or without optimization schemes (Berry et al., 2003), yet the optimization of sensor deployment locations is often necessary because of the high cost of monitoring devices and to achieve the highest degree of protection for a finite number of sensors (Thompson et al., 2007, 2009). Therefore, various methodologies for layout design of monitoring stations have been proposed in the past decade in water distribution systems to detect the migration of any contaminations that can cause adverse effects on consumer health (Kessler et al., 1998; Al-Zahrani and Moied, 2001; Woo et al., 2001; Haught et al., 2003; Ostfeld and Salomons, 2004; Berry et al., 2003, 2005, 2006; Propato, 2006; Ghimire and Barkdoll, 2006; Preis et al., 2007; Aral et al., 2010; Hart and Murray, 2010; Weickgenannt et al., 2010). Numerous technical approaches were developed for optimizing

* Corresponding author. Tel.: +1 407 823 1375.
E-mail address: nchang@mail.ucf.edu (N.-B. Chang).

0959-6526/$ – see front matter
doi:10.1016/j.jclepro.2012.02.010

sensor placement, including mixed-integer programming (MIP) models (Lee et al., 1991; Lee and Deininger, 1992; Watson et al., 2004; Berry et al., 2004, 2005; Propato et al., 2003), combinatorial heuristics (Kessler et al., 1998; Kumar et al., 1999; Ostfeld and Salomons, 2004), general-purpose metaheuristics (e.g., Ostfeld and Salomons, 2004), and lagrangian heuristics (Berry et al., 2008). Ideally a large number of sensors would increase the monitoring coverage of a network, but it would also increase cost accordingly; however, these methodologies are usually not applicable to small communities and developing countries due to the complication of methodologies and the lack of resources such as funding and technical computing software. More important, the lower capital cost of the projects will more likely be granted if performance is comparable to the higher-cost project, especially in small communities or developing countries. The remaining scientific question is how to generate sensor deployment strategies with low computational burden, such as we often encountered in large-scale optimization analyses.

This study developed a Rule-based Decision Support System (RBDSS), a new method for sensor deployment, to generate near-optimal sensor deployment strategies with low computational burden to improve consumers' health and safety by preventing civilians from consuming contaminated water. Three rules were derived to address the efficacy and efficiency characteristics: (1) intensity, (2) accessibility, and (3) complexity rules. Such an RBDSS is thus designed to minimize the total number of costly sensors and maximize the monitoring coverage to promote the cost-effectiveness of an EWS in small communities. Because a real-world case study is the most adequate research strategy for theoretical research to the real-world implementation, RBDSS was applied to the water distribution network in Hardin No. 1 County in Kentucky to validate the methodology. In this work we provide the formulation of the three rules for RBDSS, present a real-world application and results of an RBDSS, and apply these results to a rural community in Kentucky.

2. Methodology of rule-based decision support system

The RBDSS developed in this study is a decision support system to optimize sensor deployment location based on three rules, the intensity, accessibility, and complexity rules, for applications in small communities and developing countries to maximize the protection of contaminant exposure to the population and minimize the cost for sensor deployment. Because the Maximum Contaminant Levels (MCLs) are regulated by the US Environmental Protection Agency (EPA), the intensity rule, which has primary focus on concentration of contaminants, was analyzed prior to the accessibility and complexity rules in this RBDSS. To retrieve the information of population exposure in the context of the intensity rule, the EPA's water quality network (EPANET) model was applied for the vulnerability assessment (Rossman et al., 1994). EPANET is software developed by EPA's Water Supply and Water Resources Division (EPA, 2011) that models water distribution piping systems and performs extended-period simulation of the hydraulic and water quality behavior within pressurized pipe networks. In principle, the accessibility rule addresses the flow fraction downstream at a node driven by the downstream water demand within the prescribed spatiotemporal pattern of a drinking water distribution network. Thus, the fraction of water flow can be assumed as a surrogate index to indicate the percentage of population that could be affected when an unexpected contaminant intrusion occurs. The complexity rule aims to deconstruct the configuration of the node structure and translate the configuration to account for a larger population that might possibly be affected by an event within a network, eliminating the need to deal with temporal variability. Each of these three rules can be analyzed independently based on the same set of the collected data, and all three rules may also be grouped together, with flexibly contributing to the final decision of sensor deployment by differing collective methods. To illustrate the robustness of these three rules, the rules were applied in three scenarios for residual chlorine, total trihalomethanes (TTHM), and lead simulated by EPANET.

The analytical process of RBDSS consists of four phases, including data collection, dynamic simulation, development, and evaluation (Fig. 1). The collected data consist of Geographic Information System (GIS), water quality, and water flow data. The RBDSS is designed to ease the burden of large-scale sensor location optimization to minimize cost and maximize coverage of protection in drinking water networks with the aid of a predetermined number of sensors. Within this context, EPANET, EXCEL®, and LINGO 10.0 were selected to support essential dynamic simulations, data analysis, and bubble sorting of data and selection of sensor locations, respectively, in which EXCEL® was used to handle data streams in support of EPANET simulation and LINGO 10.0 optimization modules.

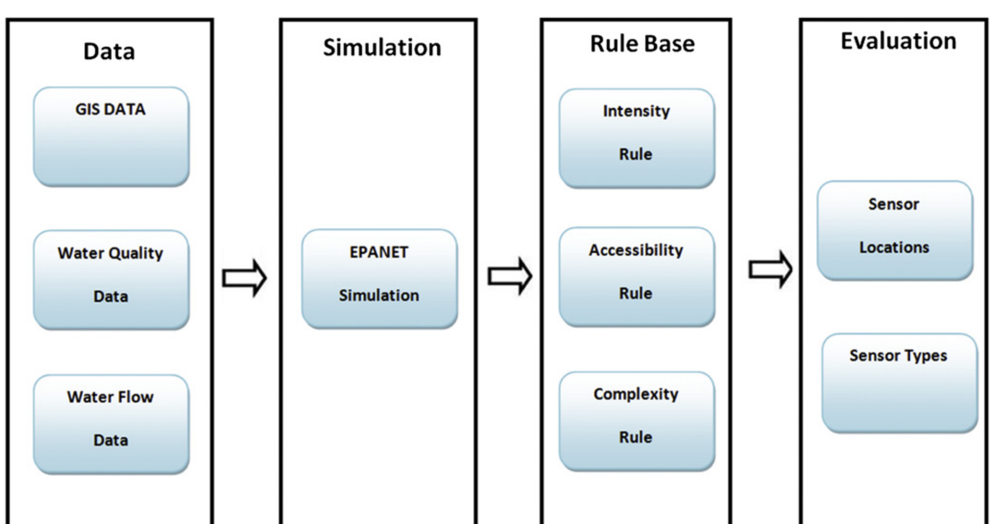


Fig. 1. Schematic of the RBDSS process.

3. Theory of the rule-based decision support system

3.1. Intensity rule

The intensity rule is designed with respect to population exposure to contamination incidents. The principle of this rule is to ensure that the concentration of targeted microorganisms, disinfection by-products, disinfectants, inorganic chemicals, organic chemicals, and/or radionuclides is under MCLs, except for the residual chlorine concentration, which must meet the minimum concentration requirement of 0.2 mg L^{-1} but not exceed the MCL of 4 mg L^{-1}, regulated by the EPA (EPA, 2006; EPA, 2009). The first-order decay was used in the chlorine residual analysis. Thus, the intensity rule may be versatile in association with several chemical species of concern in the drinking water distribution networks to prevent fatalities in any incidental or accidental events. Regardless of the effect of diurnal variation, the node shall not exceed the MCLs at any time during the day; nodes are ranked from highest to lowest exceedance, with nodes that exceed the MCLs ranked highest, and the top "k" nodes are selected for deploying sensors. However, some chemicals must meet the minimum concentration standard (i.e., 0.2 mg L^{-1} of residual chlorine concentration; EPA, 2006). In this case, the objective is to minimize the summation of total concentrations at these nodes that violate (i.e., deceed) the minimum concentration standard. To quickly sort out the nodes with higher priority, a bubble sort optimization scheme (Horowitz et al., 2007) was used in all three scenarios, including chlorine residual, TTHM, and lead. Because chlorine concentration is the parameter related to chlorine residual and TTHM scenarios simultaneously, the intensity rule must be implemented based on the upper and lower bounds of chlorine concentration to screen out these nodes with priority. The bubble sort optimization scheme can rank the data by swapping them in the correct order from highest to lowest based on the chlorine concentration. The nodes with chlorine concentration below the minimum requirement established by EPA will be collected as candidate nodes for testing the chlorine residual scenario, whereas the nodes with chlorine concentration over the maximum requirement established by the World Health Organization will be collected as candidate nodes for testing the TTHM scenario.

To determine the near-optimal solution with a quick screening tool when the total number of nodes involved is high, LINGO 10.0®, an optimization solver, may be used to optimize the selection process of sensor locations. Because the intensity rule can be applicable to any chemicals or microorganisms regulated by the EPA, the scenarios must specify the chemicals or microorganisms of interest. To detect exceedance–deceedance situations, simulation of the dynamic concentrations in a water distribution system using a well-developed simulation model, such as EPANET, may be performed. Using the outputs from EPANET, we can consider two objective functions concurrently in two separate small-scale optimization models. One objective function of this small-scale screening model is to maximize the detection limit of summation of exceedance concentrations of contaminant, such as TTHM and lead, in the network (note that TTHM is a byproduct of chlorinating water that contains natural organics). The other objective function is to minimize the summation of deceedance concentrations of chlorine residual as regulated by the EPA. These two small-scale optimization models can be formulated independently and applied collectively to finalize the implementation of the intensity rule. With this small-scale integrated simulation and optimization model, the sensor deployment can be carried out based on the assumptions that the budget is limited, the cost of the same type of sensor deployment is equal at every node location, and each monitoring event can generate the optimal solution independently with no mutually related effect.

Decision variables are a set of binary variables, x_i, defined as

$$x_i = \begin{cases} 1, & \text{if node } i \text{ is selected} \\ 0, & \text{otherwise} \end{cases}.$$

Submodel 1: prevents the contaminant from exceeding the MCLs:

$$\text{Maximize} \quad \sum_{i=1}^{k} C_{i,\max} x_i, \tag{1}$$

subject to:

$$S \le ks \quad \forall k \in N, \tag{2}$$

$$\sum_{i=1}^{k} x_i \le k, \text{ and} \tag{3}$$

$$C_{i,\max} \ge C_{MCL} \quad \forall i. \tag{4}$$

Submodel 2: performs the quality control of minimum concentration standard:

$$\text{Minimize} \quad \sum_{i=1}^{k} C_{i,\min} x_i, \tag{5}$$

subject to:

$$S \le ks \quad \forall k \in N, \tag{6}$$

$$\sum_{i=1}^{k} x_i \le k, \text{ and} \tag{7}$$

$$C_{i,\min} \le C_{MS} \quad \forall i, \tag{8}$$

where S is the total budget for sensor deployment (US$); s is the cost of each deploying sensor per node ($); N is the number of junctions in the water distribution (dimensionless); k is the number of total sensor available to be deployed (dimensionless); i is the subscript representing the I sensor to be deployed up to k locations, $1 = 1, 2, \ldots, k$; $C_{i,max}$ is the concentration of contaminant of interest at a node that exceeds the MCLs at i location (mg/L); $C_{i,min}$ is the concentration of chlorine residual at node that deceeds the minimum concentration standard at i location (mg/L); C_{MCL} is the MCLs regulated by the EPA (mg/L); and C_{MS} is the minimum concentration standard regulated by the EPA (mg/L).

The objective function in equation (1) represents the maximum summation of concentration of contaminant that exceeded MCLs in the drinking water distribution network, from which the candidate nodes for sensor deployment are determined. The objective function in equation (5) represents the minimum summation of concentration of residual chlorine concentration that violated the minimum standard of the EPA in the drinking water network, from which the candidate nodes for sensor deployment are determined. Equations (2) and (6) represent cost constraint of sensors, which is determined by dividing the total budget (S) by the cost per sensor deployment (s) to ensure the number of sensors (k) do not exceed the upper bound as defined as the right-hand-side values in the constraints. Equations (3) and (7) represent the constraint of the number of the selected nodes, not to exceed the number of sensors (k), which is determined by the total budget for sensor deployment. Equation (4) represents the constraint of MCLs associated with the objective function represented by equation (1). Equation (8) represents the constraint of minimum concentration standard associated with objective function represented by equation (5).

3.2. Accessibility rule

Population exposure to potential contaminants is a specific concern related to the flow pattern in the network. The accessibility rule can be defined as the flow fraction from the main pipeline to secondary pipes in the remaining part of a network. For residential area, because the water flow in a particular pipeline at a given time step is driven by the downstream water demand within a spatio-temporal pattern, the fraction of water flow can be assumed as a surrogate index to indicate the percentage of population that could be affected when an unexpected contaminant intrusion occurs. This approach does not have to specify a certain node or pipeline at which the contaminant intrusion happens. Rather, the goal is to propose an optimal design to ensure maximum protection to the portion of the population residing in that part of the network. This implies that the higher the flow fraction at a certain node located at a cross or a T pipe fitting, the larger the population that could be affected by contaminant intrusion and could be protected by the sensor deployment strategies. From an economic perspective, placing sensors in a high-based demand area may exhibit greater efficacy than deploying sensors in a low-based demand area.

Because a higher flow fraction leads to greater population protection, the design objective of the accessibility rule is to maximize flow fractions associated with the predetermined number of sensors for deployment:

$$R = \sum_{j=1}^{k} \frac{q_j}{Q_j} = \sum_{j=1}^{k} r_j \quad \forall k \in N, \tag{9}$$

where Q_j is flow rate from the main pipe at j location; q_j is flow rate from the subroutine at j location; r_j is the flow fraction ($r_j = q_j/Q_j$) at j location; R is the maximum summation of the flow fractions for k sensors; k is the predetermined total number of sensors for deployment; and N is the number of junctions in the drinking water distribution networks.

The objective function can be achieved by calculating the flow fraction of every node with at least one or more secondary pipes connected to the main pipe. The flow fractions are then ranked from highest to lowest, and the top "k" nodes are selected based on the ranking system for possible sensor deployment. Such an analysis may be deemed a supplemental step in addition to the intensity rule or may be performed independently for a small-community, should the community have no resources to carry out the essential calculation involved in the intensity rule.

3.3. Complexity rule

The complexity rule originates from a branch of the graph theory of computation in computer science that classifies problems according to their inherent difficulties. In this case, the advantage of applying the complexity rule is its ability to solve sensor placement issues in a more explicit way for small-scale drinking water distribution networks that contain fewer intersections or loops among pipelines (Deuerlein et al., 2009). Path Element Reduction, in which the nodes of the block are distinguished according to their nodal degree (number of connected links), was used in the application of graph decomposition in RBDSS (Deuerlein et al., 2009). To translate the complexity rule into a programming algorithm, graph theory should be applied to develop the complexity formulas:

$$X = \sum_{i=1}^{k} x_i \quad \forall k \in N, \tag{10}$$

where X is the maximum summation of inner nodes within k path nodes; x_i is the number of inner nodes within impact zone, r_i, of the path node at i location; N is the number of junctions in the drinking water distribution network; and k is the predetermined total number of sensors for deployment; and

$$r_i = \frac{\sum_{m=1}^{l} d_{i,m}}{l_i}, \tag{11}$$

where r_i is the impact zone of the path node at i location; $d_{i,m}$ is the distance from the path node at i location to the inner node at m location; and l_i is the number of inner nodes within the impact zone of the path node at i location.

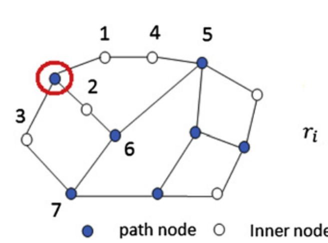
For simplification of network analysis, the nodes of the block are first distinguished according to their nodal degree (number of connected links). All nodes can be categorized into two groups: a path node has one or more pipes connected to the main pipe; an inner node is located between two path nodes (Fig. 2) by which the impact zone can be calculated. In the impact zone, the number of path nodes to receive deployed sensors is equal to the predetermined total number of sensors for deployment. The higher the number of nodes is within a determined circular radius, the greater the population in this targeted area. Thus the objective of the complexity rule is to determine the number of path nodes with the maximum combined number of inner nodes based on path nodes' individual impact zones.

The impact zone of a particular path node is determined by averaging the distance from all the inner nodes with a hydraulic connection to the path node. The number of inner nodes located within the determined circular radius of impact zone is then counted. Next, all the path nodes are ranked from highest to lowest based on the number of inner nodes. Finally, the top "k" path nodes are selected for sensor deployment. Again, such an analysis may be deemed a supplemental step to the intensity and accessibility rules or may be performed independently for small communities with no resources to carry out the essential calculation involved in the intensity rule.

4. Application of the rule-based decision support system

Statistics compiled by the Kentucky Division of Water (DOW) from public water system data in 1995, and subsequently reported in 1996, indicate that the greatest violators of federal drinking water regulations are small systems serving 3300 people or fewer (CERS, 2007). Analysis of these data reveals that 78% (942) of violations occurred in public water systems that serve fewer than 500 people (CERS, 2007). Of the 1207 total violations cited by Kentucky DOW, 93% were monitoring and reporting infractions (CERS, 2007). Yet small communities can rarely afford to integrate effective monitoring system into their networks; large cities typically have abundant resources to establish EWS to monitor water supplies and distribution network. Hence, cost-effective EWS for small-scale drinking water networks are desperately needed to monitor small drinking water networks and improve public safety.

To test the practicality of employing the RBDSS, the three rules were applied on the water distribution network in Hardin County Water District No. 1, a part of Elizabethtown in Kentucky (Fig. 3). The population estimate in 2009 was 99,770 (U.S. Census Bureau, 2009). The county relies solely on the Pirtle Spring water treatment plant, located on the west side of the water distribution



$$r_i = \frac{d_1 + d_2 + d_3 + d_4 + d_5 + d_6 + d_7}{7}$$

Fig. 2. Calculation of the impact zone for circled node (Deuerlein et al., 2009).

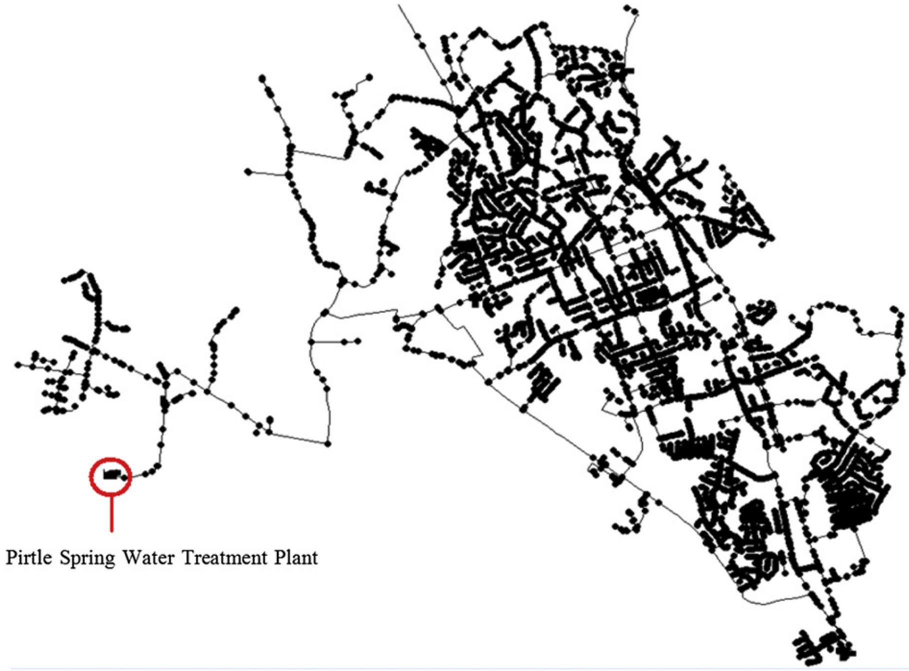


Fig. 3. Hardin No. 1 network.

network (Fig. 3), as its primary water source. The capacity of the plant is 2 million gallons per day to supply residential areas. The chlorine dosage of the treatment plant is 1.70 mg L^{-1}, and no rechlorination stations are used to maintain the chlorine residual. The majority of the population is located at the Fort Knox army military base north of Elizabethtown.

Although Hardin is not a big city, the assessment for optimal sensor deployment was based on five different sections of the network (Fig. 4) to ease the presentation. Three scenarios based on EPANET simulation were prepared for delineating residual chlorine, TTHM, and lead, respectively, with the assumption that the available budget can be distributed to deploy 10 sensors in each scenario.

The RBDSS was applied to run the intensity, accessibility, and complexity rules in series to prioritize the location of the sensors. Although the three rules are independent from one another, they may be applied in series to discern each node's potential for sensor deployment. In other words, the collected data were analyzed first by the intensity rule to pinpoint more than 10 candidate nodes. The node contenders for sensor deployment obtained from the intensity rule were then evaluated by the accessibility rule to narrow down

Fig. 4. Hardin No. 1 network divided into 5 sections. WTP represents the water treatment plant.

Fig. 5. Section 2 of the network. Green and blue circles represent the nodes selected for chlorine residual and lead scenarios, respectively. The red dot represents Tank 26653 in which lead was injected to simulate the third scenario. The nodes selected for chlorine residual and lead scenarios, are represented in green and blue circles respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the candidate list. Finally, the final selected nodes were generated by the complexity rule based on the candidate nodes obtained from the accessibility rule to finalize the 10 nodes for sensor deployment.

The 720-h simulation was performed using EPANET to implement the intensity rule on the Hardin No. 1 drinking water distribution network. Chlorine residual and TTHM scenarios were simulated to select the nodes that could not meet the minimum standard or the MCLs, respectively. In contrast, the lead scenario was simulated where the sensors should be deployed to provide optimal level of protection for these residents of Handin No. 1 system. At present, chlorine residual is regulated by EPA to meet the minimum standard of 0.2 mg L^{-1} (EPA, 2006). During the simulation, the actual chlorine dosage (1.70 mg L^{-1}) was injected at Pirtle Spring water treatment plant located at the lower west location of the network (Fig. 4). After the scenario was simulated, the intensity rule was used to analyze the sensor deployment location by using equations (5)–(8) to select the nodes with simulated chlorine residual concentrations below the standard. The accessibility rule was then applied using equation (9), and the complexity rule was applied using equations (10) and (11) to determine the final sensor deployment nodes.

The second scenario evaluated TTHM. After the scenario was simulated, the intensity rule was used to analyze the sensor

Fig. 6. Section 5 of the network.

Table 1
Top 10 nodes selected for residual chlorine scenario by using the intensity, accessibility, and complexity rules in sequence.

No	Node ID		
	Intensity rule	Accessibility rule	Complexity rule
1	1267	1267	1267
2	13779	6731	1573
3	6731	6643	5035
4	6643	1573	6794
5	1573	25358	3309
6	25358	5035	4363
7	5035	769	1986
8	769	6794	24519
9	251	3309	2008
10	6794	4363	2151
11	3309	1986	
12	4363	22285	
13	1986	24519	
14	1224	2008	
15	22285	2151	
16	1008		
17	24521		
18	24519		
19	2008		
20	2151		

deployment location by using equations (1)–(4) to select the nodes with simulated chlorine residual concentrations below the standard. The accessibility rule and complexity rule were applied as in the first scenario. The MCLs states that TTHM must remain below 0.08 mg L^{-1}. According to the simulation for this scenario, none of the nodes in the network has chlorine concentration exceeding 4.0 mg L^{-1}, and because TTHM is a disinfection by-product of chlorine in the network, a chlorine concentration that remains below the MCLs for chlorine (4.0 mg L^{-1}) indicates that TTHM formation in the network is negligible. Thus, we combined this scenario testing with previous one.

Finally, lead, which is regulated by EPA, has an MCL of 0.015 mg L^{-1}; however, this simulation was intended to evaluate a possible accidental leakage or an intentional attack targeting the water tanks in Hardin No. 1 network. As expected, all 10 sensors to be deployed were located in pipe section 2 (Fig. 5) because the concentrations of lead decreases as it migrates farther from the source location (i.e., Tank 26653). In the simulation, although the network consists of four water tanks, a lead concentration of

Table 2
Top 10 nodes selected for lead scenario by using the intensity, accessibility, and complexity rules in sequence.

No	Node ID		
	Intensity rule	Accessibility rule	Complexity rule
1	24813	24813	24837
2	212	209	25845
3	209	25837	1470
4	25837	25845	25880
5	25845	25869	25852
6	25869	1470	180
7	1470	25880	188
8	25880	171	196
9	171	25852	25898
10	25852	180	213
11	180	188	
12	178	196	
13	170	25898	
14	188	161	
15	196	213	
16	25910		
17	156		
18	25898		
19	161		
20	213		

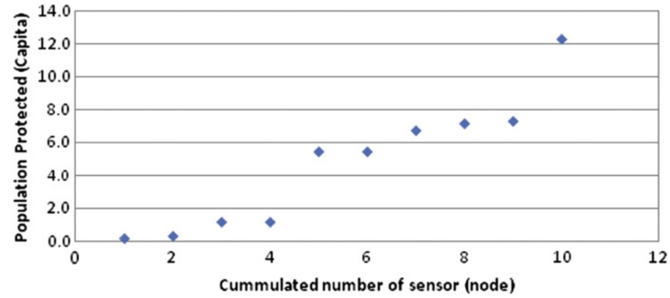


Fig. 7. Sensitivity analysis of sensor deployment for chlorine residual based on the size of the population protected.

15 mg L^{-1} was released at tank ID 26653 located in the area with the highest population density in the network (Fig. 6). Then, equations (1), (3)–(5) from the intensity rule, equation (7) from the accessibility rule, and equations (8) and (9) from the complexity rule were applied to indicate the nodes that could exceed the MCLs.

To test robustness of the RBDSS, sensitivity analysis was performed. Because the simulation showed that the TTHM scenario did not require any sensor deployment, residual chlorine and lead scenarios were the only cases considered for sensitivity analysis. Two indices, the size of the population protected and exposure level, were used for sensitivity analysis as the number of deployed sensors increased. To determine the size of the population protected at the sensor location, the baseline demand at the selected node was divided by water consumption per capita to determine the size of the population protected at that node. In our case, the average water consumption rate was 100 gal/day/capita. Likewise, the exposure assessment, defined as the amount of substance consumed by a person at a given exposure level of a specified chemical or organism, was calculated by multiplying the substance concentration at the selected node with the same water consumption rate per capita. These two indices may be collectively used for final robustness assessment of the optimal sensor deployment strategies.

5. Results

Ranking these selected nodes may further reveal the cost-effectiveness in sensor deployment should financial constraint be emphasized. In other words, the node with higher rank receives higher priority, implying a greater number of residents may be protected if the corresponding sensor can be deployed at that node. The rankings of sensor locations associated with different scenarios were summarized for the residual chlorine scenario (Table 1) and the lead scenario (Table 2). For the TTHM scenario, none of the nodes in the network exceeded the MCL of 4.0 mg L^{-1}; therefore, sensor deployment was not necessary. Because the network consists of 25,964 nodes and 15,600 pipes, only partial results were presented in this paper.

Table 3
Ranking of the selected sensors associated with two sensitivity analyses and peak concentration of chlorine residual at each node.

Sensor rank	Sensor ID	Population protected (capita)	Exposure level (mg/capita/d)	Peak concentration of chlorine residual (mg/L)
1	1267	0.1	0.0	0.00
2	1573	0.3	0.0	0.00
3	5035	1.1	0.0	0.00
4	6794	1.1	0.0	0.00
5	3309	5.4	0.0	0.00
6	4363	5.4	0.0	0.00
7	1986	6.7	0.0	0.00
8	24519	7.1	0.0	0.00
9	2008	7.3	0.0	0.00
10	2151	12.3	0.3	0.01

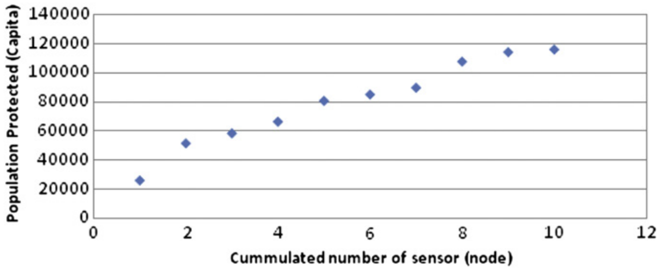


Fig. 8. Sensitivity analysis of sensor deployment for lead detection based on the size of the population protected.

6. Discussion

The RBDSS outputs show that the selected nodes for sensor deployment to detect residual chlorine are located throughout the water network, except section 3 where the Pirtle Spring water treatment plant is located. This section should have the highest residual chlorine concentration, and there are no rechlorination stations in other sections of the drinking water distribution network. Residual chlorine simulation results indicate that the summation of the residual chlorine concentration of the selected nodes equals zero. In other words, these selected nodes are highly unlikely to be effectively disinfected by the chlorine dosed at the plant. Because none of these nodes are located along the main pipelines, the lack of disinfection at the selected nodes would not cause a significant negative impact on the majority of the population.

Several sensors could be applied in this case study, and other commercial sensors are available depending on the contaminants considered (Appendix A). The sensors for residual chlorine detection can be AccuChlor 2 Residual Chlorine Measurement System, CL17 Free Residual Chlorine Analyzer, or Series B20 Residual Chlorine Recorder with type B sensor (Appendix A). For sensitivity analysis, the population protected, the number of people who benefit from the sensor being deployed based on the base demand of water, is greater when a larger number of sensors can be deployed (Fig. 7). The total number of protected residents is significantly small (only 12 people can be protected when 10 sensors are deployed) because the chlorine residual is continuously decayed while the water is aged in the pipeline, especially when no rechlorination station is installed in the water distribution, as in the Hardin No. 1 network (Fig. 7). The node farthest hydraulically from the water treatment plant or the source of chlorination will have the lowest chlorine residual concentration. As a result, the selected nodes are those located far from the water treatment plant and the highly populated area. In other words, these nodes are located at in low population density areas, and as a result, the deployed sensors can only protect a small number of people. The lower exposure levels, determined by how much contaminant is consumed by each person, indicate that the levels of residual chlorine that effectively disinfect

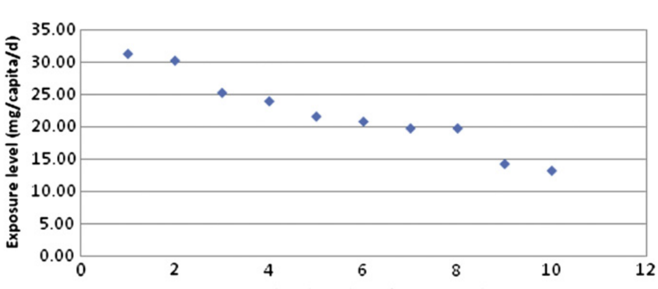


Fig. 9. Sensitivity analysis of sensor deployment for lead detection based on the exposure assessment.

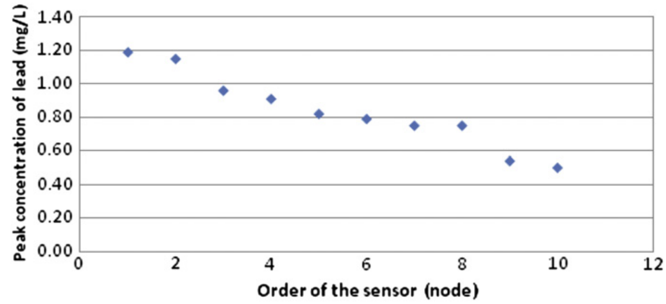


Fig. 10. Peak concentration of lead at the selected nodes.

at those nodes are below the minimum standard. Similarly, the selected nodes ranked first through ninth have peak residual chlorine concentrations of 0.0 mg L^{-1} (Table 3), indicating that the water flows at these nodes have no disinfection, except 10th; therefore, they were selected to deploy sensors to detect such violations.

Finally, lead release due to either a terrorist attack or a pipe corrosion scenario was explored. Based on the observations of the EPANET simulation outputs (Fig. 5), these nodes are located along the first pipe section, which receives most of the outflow from the water tank and has a significantly higher lead concentrations than the other pipe sections. Therefore, these selected nodes would require the installation of type A sensor of SMART 2 Colorimeter with the 3660-SC Reagent System Portable Cyanide Analyzer or Deltatox® instrument (see Appendix A). In sensitivity analysis, the size of the population protected can be significantly increased as the total number of sensor to be deployed is increased (Fig. 8). When 10 sensors are deployed in the lead scenario, the maximum possible number of population protected is 116,362 based on the average water consumption per capita. This population is much higher than the number in the residual chlorine scenario because all the deployed sensors are located in the highly populated area of the county; thus, more of the optimal locations for sensor deployment were in pipe section 2 to maximize the protection for largest population residing in this region. In addition, the sensitivity analysis for exposure assessment (Fig. 9) indicates that the higher the number of sensor to be deployed, the lower the exposure level of the substance to the population. The level of exposure is decreased instantly as more sensors are deployed. The higher the number of sensors deployed, the higher the probability of contaminant being detected prior to the end users' consumption. Therefore, the amount of water being consumed per capita per day will be reduced because the public warning must be announced as soon as the contaminant is detected. For example, when one sensor is deployed, the level of exposure is 31.42 mg/capita/day; but when 10 sensors are deployed, the level of exposure decreases dramatically to 13.20 mg/capita/day.

Table 4
Ranking of the selected sensors associated with two sensitivity analysis parameters and peak concentration of lead at each node.

Sensor rank	Sensor ID	Population protected (capita)	Exposure level (mg/capita/d)	Peak concentration of lead (mg/L)
1	24837	25,807	31.42	1.19
2	25845	51,484	30.36	1.15
3	1470	58,398	25.34	0.96
4	25880	66,435	24.02	0.91
5	25852	80,814	21.65	0.82
6	180	85,238	20.86	0.79
7	188	89,933	19.80	0.75
8	196	108,021	19.80	0.75
9	25898	114,537	14.26	0.54
10	213	116,362	13.20	0.5

The marginal sensitivity of sensors for lead detection based on the exposure assessment (Fig. 10) confirms the diminishing rate of return. The more sensors deployed, the smaller the marginal effect of sensor deployment. The cost-effectiveness of the RBDSS (Table 4) is collectively based on three indices, including population protected, exposure levels, and peak concentrations. The concentration of lead at a node can be as high as 1.19 mg L^{-1}, yet the concentration becomes 0.5 mg L^{-1} at the 10th selected node.

7. Conclusions

The RBDSS associated with the three rules described in this paper proved effective at simplifying and solving the sensor placement problem in a small-scale community, the Hardin County Water District No. 1 in Kentucky. Overall, correlation among the three rules can be drawn so that, based on the intensity rule, the location with the highest population density is proposed for more sensor deployment than others because higher exposure levels might occur along the main pipeline and water tanks. This vision is consistent with the fact that the flow factions of these areas picked up by the intensity rule should also be higher based on the accessibility rule, and the number of the inner nodes should be picked up more often based on the complexity rule.

Other findings of the RBDSS can be summarized as: (1) Temporal effects: the RBDSS is a well-suited tool for application that can eliminate the impact of changing water flow direction (e.g., the temporal effect) in the networks. (2) Sensor selections: our model integrates different commercially available sensor types into the decision-making process, making the proposed RBDSS more practical and user-friendly for solving real-world sensor placement issues. (3) Symmetry versus asymmetry of rules: although accessibility and complexity rules have different algorithms, the purpose of both rules is to maximize the protection of population exposure to contaminant, yet the complexity rule can decompose the complexity within a network and maximize the coverage for water flows coming through every pipeline. (4) Full-scale applications: this RBDSS is transformative and transferable to drinking water distribution networks elsewhere with any scale. (5) Given that further knowledge discovery is required to address extended issues, more rules might be required in the future.

Acknowledgment

The authors are grateful for the financial support from the Kentucky Science and Technology Foundation and the data provided by the Water Management District, Hardin, Kentucky.

Appendix A. Interim voluntary guidelines for designing and online contaminant monitoring system (ASCE, AWWA, and WEF, 2004):

Instrument/testing kit	Manufacturer	Parameters observed	Sensor type
Series B20 Residual Chlorine Recorder	Analytical Technology, Inc.	Free chlorine, chloramines	B
VVR Water Analysis System	Chemetrics	Ammonia, bromine, chlorine, chlorine dioxide, chromate, copper, cyanide, DEHA, formaldehyde, glycol, hydrazine, hydrogen peroxide, iron, molybdate, nitrate, nitrite, oxygen (dissolved) ozone, peracetic acid, phenols, phosphate, silica, sulfide, zinc	B
Six-CENSETM	Dascore	Chlorine (no reagents required), monochloramine or dissolved oxygen, pH, temperature, conductivity, ORP/REDOX	B
Ocean Seven 316 Water Probe	General Oceanics, Inc.	Pressure, temperature, conductivity, salinity, oxygen, pH, oxidation reduction potential.	A
Water Distribution Monitoring Panel (WDMP)	Hach	Chlorine, conductivity, pH, turbidity, pressure, temperature	A/B
AccuChlor 2 Residual Chlorine Measurement System	Hach	Chlorine	B
CL17 Free Residual Chlorine Analyzer	Hach	Chlorine	B
PolyToxTM Rapid Toxicity Test	InterLab Supply, Ltd.	pH, dissolved oxygen (ppm), temperature (°C), toxic metals (ppm)	A
Deltatox®	Strategic Diagnostics Inc./Azur Environmental	Phenol, lead, arsenic, mercury, sodium cyanide, selenium, potassium cyanide, chromium, PR-toxin, copper, aflatoxin, ochratoxin, rubratoxin, chloroform, ammonia, sodium lauryl sulfate, benzoyl cyanide, lindane, DDT, cresol, formaldehyde, malathion, carbaryl, flouroacetate, trinitrotoluene (TNT), parathion, 4-phehnyl toluene, carbofuran, pentachlorophenol, patulin, paraquat, diazinon, cyclohexamide, cadmium, quinine, dieldrin	B/C
SMART 2 Colorimeter with the 3660-SC Reagent System Portable Cyanide Analyzer	LaMotte Company	Alkalinity UDV, Aluminum, Ammonia, Nitrogen-LR (Fresh Water), Ammonia, Nitrogen-LR (Salt Water), Ammonia Nitrogen, Boron, Bromine LR, Bromine UDV, Cadmium, Carbohydrazide, Chloride, Chromium, Hexavalent, Chromium TesTab, Chromium (Total, Hex & Trivalent), Cobalt, COD COD SR 0-1500 without Mercury, COD HR 0-15,000 with Mercury, COD HR 0–15,000 without Mercury, Color, Copper BCA – LR, Copper Cuprizone, Copper DDC, Copper UDV, Cyanide, Cyanuric Acid, Cyanuric Acid UDV, DEHA, with Mercury, COD HR 0-15,000 without Mercury, Color, Copper BCA – LR, Copper Cuprizone, Copper DDC, Copper UDV, Cyanide, Cyanuric Acid, Cyanuric Acid UDV, DEHA, Dissolved Oxygen (DO), Erythorbic Acid, Fluoride, Hydrazine, Hydrogen Peroxide, Hydroquinone, Iodine, Iron, Iron UDV, Iron Phenanthroline, Lead, Manganese LR, Manganese HR, Mercury, Methylethylketoxime, Molybdenum HR, Nickel, Nitrate Nitrogen-LR, Nitrate TesTab, Nitrite Nitrogen-LR, Nitrite TesTab, Ozone LR, Ozone HR, pH CPR (Chlorphenol Red), pH PR (Phenol Red), pH TB (Thymol Blue), Phenol, Phosphate LR, Phosphate HR, Potassium, Silica LR, Silica HR, Sulfate HR, Sulfide LR, Surfactants, Tannin, Turbidity, Zinc LR	B

References

Al-Zahrani, M., Moied, K., 2001. Locating optimum water quality monitoring stations in water distribution system. In: Bridging the Gap: Meeting the World's Water and Environmental Resources Challenges, Proceedings of the ASCE Annual Conference on Water Resources Planning and Management. 2001, published on CD.

Aral, M.M., Guan, J., Marslia, M.L., 2010. Optimal design of sensor placement in water distribution networks. Journal of Water Resources Planning and Management 136 (1). doi:10.1061/_ASCE_WR.1943-5452.0000001.

Berry, J.W., Fleischer, L., Hart, W.E., Phillips, C.A., Watson, J.P., 2005. Sensor placement in municipal water networks. Journal of Water Resources Planning and Management Division, ASCE 131 (3), 237–243.

Berry, J.W., Hart, W.E., Phillips, C.A., Uber, J.G., Watson, J.P., 2006. Sensor placement in municipal water networks with temporal integer programming models. Journal of Water Resources Planning and Management Division, ASCE 132 (4), 218–224.

Berry, J.W., Boman, E., Phillips, C.A., Riesen, L.A., 2008. Low-memory lagrangian relaxation methods for sensor placement in municipal water networks. In: World Environmental and Water Resources Congress 2008. Ahupua'a© 2008 ASCE, Published on CD.

Berry, J.W., Fleischer, L., Hart, W., Phillips, C., 2003. Sensor placement in municipal water networks. In: Proceedings CD of World Water and Environmental Resources Congress, Philadelphia, PA, June 23–26, pp. 3383–3392.

Berry, J., Hart, W.E., Phillips, C.A., Uber, J., 2004. A general integer-programming-based framework for sensor placement in municipal water networks. In: Proc. CD of World Water and Environment Resources Conference.

Buckle, P., 2000. Assessing Resilience and Vulnerability in the Context of Emergencies: Guidelines. Victorian Government Publishing Service. Retrieved October 20, 2002 from: www.anglia.ac.uk/geography/radix/resources/buckle-guidelines.pdf.

Center for Water Resource Studies (CERS), 2007. Bowling Green. KY 270-745-8895 accessed in April, 2007.

Clark, R.M., Deininger, R.A., 2001. Minimizing the vulnerability of water supplies to natural and terrorist treats. In: Proceedings of the American Water Works Association's IMTech Conference, held in Atlanta, GA, pp. 1–20. April 8–11, 2001.

Deuerlein, J., Wolters, A., Roetsch, D., 2009. Reliability analysis of water distribution system using graph decomposition. World Environmental and Water Resource Congress, 272–282.

EPA, 2006. ICR Treatment Study Fact Sheet: The Simulated Distribution System Test. US Environmental Protection Agency. EPA 814-B-96-002. http://www.epa.gov/ogwdw000/icr/icr_sds.html (accessed 03.05.10.).

EPA, 2009. National Primary Drinking Water Regulations. US Environmental Protection Agency. EPA 816-F-09–004. May 2009.

EPA, 2011. EPANET: Software that Models the Hydraulic and Water Quality Behavior of Water Distribution Piping Systems. http://www.epa.gov/nrmrl/wswrd/dw/epanet.html (accessed 15.11.11.).

Ghimire, S.R., Barkdoll, B.D., 2006. Heuristic method for the battle of the water network sensors: demand-based approach. In: Proceedings CD of the 8th Annual Water Distribution System Analysis Symp., Cincinnati, OH, USA.

Haestad, M., Walski, T.M., Chase, D.V., Savic, D.A., Grayman, W., Backwith, S., Koelle, E., 2003. Advanced Water Distribution Modeling and Management. Haestad Press, Waterbury, CT USA.

Hart, W.E., Murray, R., 2010. Review of sensor placement strategies for contamination warning system in drinking water distribution systems. Journal of Water Resources Planning and Management 136 (6). doi:10.1061/_ASCE_WR.1943-5452.0000081.

Haught, R.C., Goodrich, J.A., Herrmann, J., 2003. Evaluation of water monitoring instrumentation at EPA's water awareness technology evaluation research security center. In: Proceedings CD of Water Quality Technology Conference. American Water Works Association, Philadelphia, Pennsylvania.

Horowitz, Ellis, Sartaj, Sahni, Susan, Anderson-Freed, 2007. Fundamentals of Data Structures in C. Silicon Pr.

Karamouz, F.M., Saadati, S., Ahmadi, A., 2010. Vulnerability assessment and risk reduction of water supply systems. In: Proceedings CD of World Environmental and Water Resources Congress 2010: Challenges of Change. ASCE.

Kessler, A., Ostfeld, A., Sinai, G., 1998. Detecting accidental contaminations in municipal water networks. Journal of Water Resources Planning and Management Division, ASCE 124 (4), 192–198.

Kumar, A., Kansal, M.L., Arora, G., 1999. Discussion of `detecting accidental contaminations in municipal water networks'. Journal of Water Resources Planning and Management 125 (4), 308–310.

Lee, B.H., Deininger, R.A., 1992. Optimal locations of monitoring stations in water distribution system. Journal of Environmental Engineering 118 (1), 4–16.

Lee, B.H., Deininger, R.A., Clark, R.M., 1991. Locating monitoring stations in water distribution systems. Journal of American Water Works Association, 60–66.

Norstrom, A., 2007. Planning for Drinking Water and Sanitation in Peri-Urban Areas. Swedish Water House Report 21. SIWI, 2007.

National Research Council, 2002. Making the Nation Safer: The Role of Science and Technology in Countering Terrorism. The National Academies Press, Washington, D.C.

Ostfeld, A., Salomons, E., 2004. Optimal layout of early warning detection stations for water distribution systems security. Journal of Water Resources Planning and Management 130, 377–385.

Preis, A., Mayochik, Y., Ostfeld, A., 2007. Multiobjective contaminant detection response model. In: Proceedings CD of World Environmental and Water Resources Congress. ASCE.

Propato, M., Piller, O., Uber, J., 2003. A sensor location model to detect contaminations in water distribution networks. In: Proceedings CD of World Water and Environmental Resources Congress. ASCE.

Propato, M., 2006. Contamination warning in water networks: general mixed-integer linear models for sensor location design. Journal of Water Resources Planning and Management Division, ASCE 132 (4), 225–233.

Rossman, L.A., Clark, R.M., Grayman, W.M., 1994. Modeling chlorine residuals in drinking-water distribution systems. Journal of Environmental Engineering 120 (4), 803–820.

Thompson, K.A., Fann, S., Kadiyala, R., Hasit, Y.J., Jacobson, G., 2009. Implementing a contamination warning system at a department of defense facility: case study port Hueneme, California. In: Proceedings CD of World Environmental and Water Resources Congress. ASCE.

Thompson, S.L., Casman, E., Fischbeck, P., Small, M.J., VanBriesen, J.M., 2007. Vulnerability assessment of a drinking water distribution system: implications for public water utilities. In: Proceedings CD of World Environmental and Water Resources Congress. ASCE.

Tune, J., Elmore, C.A., 2009. Drinking water field analyses for the detection and Enumeration of Coliform Bacteria in rural Guatemalan Highlands. In: Proceedings CD of World Environmental and Water Resources Congress. ASCE.

U.S. Census Bureau, 2009. Hardin County, Kentucky Estimated Population in 2009, April 15. United States of America Department of Commerce. http://factfinder.census.gov/home/saff/main.html?_lang=en.

Watson, J.-P., Greenberg, H.J., Hart, W.E., 2004. A multiple-objective analysis of sensor placement optimization in water networks. In: Proceedings CD of World Water and Environment Resources Conference. ASCE.

Weickgenannt, M., Kapelan, Z., Blokker, M., Savic, D.A., 2010. Risk-based sensor placement for contamination detection in water distribution systems. Journal of Water Resources Planning and Management 136 (6). doi:10.1061/_ASCE_WR.1943-5452.0000073.

Woo, H.M., Yoon, J.H., Choi, D.Y., 2001. Optimal monitoring sites based on water quality and quantity in water distribution systems. In bridging the gap: meeting the world's water and environmental resources challenges. In: Proceedings CD of the ASCE Annual Conference on Water Resources Planning and Management. ASCE.