17th European Symposium on Computer Aided Process Engineering – ESCAPE17
V. Plesu and P.S. Agachi (Editors)
© 2007 Elsevier B.V. All rights reserved.

Optimal Location of Booster Disinfection Stations in a Water Distribution System: A Two-Stage Stochastic Approach

Vicente Rico-Ramirez[a], Francisco Gomez-De la Cruz[a], Gustavo Iglesias-Silva[a] and Salvador Hernandez-Castro[b]

[a]*Instituto Tecnológico de Celaya, Av. Tecnológico y García Cubas S/N, Celaya, 38010, Mexico, vicente@iqcelaya.itc.mx*
[b]*Universidad de Guanajuato, Facultad de Ciencias Químicas, Col. Noria Alta S/N, Guanajuato, 36050, hernasa@quijote.ugto.mx*

Abstract

Secondary or booster disinfection consists of the addition of disinfectant at distinct locations distributed throughout a water distribution system. This work describes a stochastic approach for the optimal location of booster disinfection stations in water distribution networks. The model minimizes the expected total cost involving the installation of booster stations and the mass of disinfectant needed to satisfy the residual concentration constraints within the network. Inherent uncertainties such as water demand and the chemical reactions of the disinfectant taking place on the system indirectly incorporate uncertainties on the model parameters. Hence, the problem has been reformulated as a first-stage-integer two-stage stochastic mixed integer linear program with recourse. The parameters needed for the solution are obtained through the water quality simulator EPANET 2.0. The resulting formulation has been solved through a generalized framework based on the stochastic decomposition algorithm. The framework integrates the GAMS modeling environment, the EPANET simulator, sampling code (FORTRAN) to handle uncertainties and a C++ master program. The paper describes the model and the solution framework, and compares deterministic and stochastic optimal solutions.

Keywords: Two-stage stochastic programming, stochastic decomposition algorithm, water distribution systems, booster disinfection

1. Introduction

Control of microorganisms in drinking water is generally accomplished by the addition of disinfectant (usually chlorine) at the distribution systems. There are two general approaches to drinking water disinfection: *i*) the primary or conventional method, which involves the addition of disinfectant only at the source supply or treatment station; *ii*)secondary or booster disinfection, which implies the addition of disinfectant at distinct strategic locations distributed throughout a water distribution network. In principle, the first method might cause high concentrations (health risk) of disinfectant at distribution nodes near the source and low residual concentrations (below disinfection requirements) at the far extreme nodes of the distribution system. On the other hand, the second method has proved potential for achieving the appropriate compromise between disinfectant dosage for microorganism control and health risks due to excessive concentration. An additional issue, however, arises in the booster disinfection approach; the location and dosage of the booster stations have to be determined so that the optimal disinfectant mass (optimal cost) is utilized. The dosage scheduling and station location problems have been addressed by various works in the literature [1-5]. A common feature of these approaches is that the models assume a deterministic behavior of the variables involved in the formulations. Nevertheless, uncertainties on water demand, the physical condition of the network and the chemical reactions taking place within the system indirectly incorporate uncertainties on the model parameters. This paper describes a stochastic approach for the optimal location of booster disinfection stations in water distribution networks.

2. Mathematical programming approaches to booster disinfection

Literature reports interesting approaches based on mathematical programming to the scheduling and facility location problem for booster disinfection [1-5]. A LP model for dosage scheduling [1], a MILP model for minimizing the number of stations [5] and a MILP model for the optimal location of the stations [4] are among the most relevant formulations. Such approaches are linear due mainly to a basic assumption (linear superposition) that considers each disinfectant concentration (at given time and location) as a linear summation of the individual effects of the dosage injections at the various nodes. Periodicity of the dosage rate and concentrations (and, therefore, periodicity on model parameters) is also assumed. In particular, the original MILP model provided by Boccelli et al. [4] minimizes the average disinfectant mass injected to the network and determines the optimal location of the booster stations. Their linear model parameters (composite response coefficients) are computed through the water network simulator EPANET [6]. As explained, the reported model neglects the uncertainties inherent to some of the key model variables. Our proposed stochastic version of such model is described next.

3. Our stochastic approach to the facility location model

Uncertainties on the prediction of model parameters for disinfectant kinetics and hydraulics will influence the results of a water quality network model. In this paper we intend to quantify the effects of potential uncertainties in the facility location model but still to keep the simplicity of a linear formulation. To that end, we propose to incorporate uncertainties in the linear coefficients of the MILP model provided by Boccelli et al [4], since those coefficients are the direct results of the water quality network simulations. As a consequence, the problem becomes a stochastic problem that we reformulate as a first-stage-integer two-stage mixed integer stochastic linear program with recourse.

3.1. The Stochastic Model

The model minimizes the expected total cost involving the installation of booster stations and the mass of disinfectant needed to satisfy the residual concentration constraints within the network. The two stages of proposed stochastic model are represented by Eq. (1) and Eq. (2). The first stage (Eq.(1)) corresponds to the minimization of the cost of the booster stations installation and the expected cost of the recourse function. The recourse function *Q*, or objective function of the second stage (Eq.(2)), represents the minimization of the disinfectant mass required to maintain concentration residuals.

$$\begin{aligned} &\min_{\delta} \sum_{i=1}^{n_b} C_i \delta_i + E_{\omega}\left[Q(\delta,\omega)\right] \\ &s.t. \quad \sum_{i=1}^{n_b} \delta_i \le n_b^{\max} \\ &\qquad \delta_i \in \{0,1\} \end{aligned} \tag{1}$$

$$\begin{aligned} &Q(\delta,\omega) = \min_{x} \sum_{i=1}^{n_b} \left(\frac{1}{\Delta T_i} \sum_{k=1}^{n_i} D\, M_i^k(\omega) x_i^k \right) \\ &s.t. \quad \sum_{i=1}^{n_b} \sum_{k=i}^{n_i} \alpha_{ij}^{km}(\omega) x_i^k \le u_j;\ j = 1,\dots,n_m;\ m = M,\dots,M+n_\alpha - 1 \\ &\qquad \sum_{i=1}^{n_b} \sum_{k=i}^{n_i} \alpha_{ij}^{km}(\omega) x_i^k \ge l_j;\ j = 1,\dots,n_m;\ m = M,\dots,M+n_\alpha - 1 \\ &\qquad x_i^k \le X_i^k(\omega)\delta_i \\ &\qquad x_i^k \ge 0 \end{aligned} \tag{2}$$

δ_i is a binary variable (first stage decisions) representing the installation of a booster station at node *i*, n_b^{max} is the maximum number of stations, x_i^k is a dosage multiplier (second stage decisions), M_i^k is the disinfectant mass associated to dosage period k, X_i^k is the maximum value of x_i^k, and α_{ij}^{km} are the composite response coefficients; each composite coefficient represents the response at node j and monitoring time m to the dosage provided at node i and period k. Parameters α_{ij}^{km}, M_i^k and X_i^k are computed through water quality simulations in EPANET and are assumed as functions of the uncertainties (ω).

3.2. Solution Approach

Calculations of uncertain parameters through successive use of the EPANET simulator requires a data file containing the network nodes, pipe connectivity, user demands and monitoring time interval. The procedure for the calculation of the composite response coefficients has been described by Boccelli et al [1,4]. In particular, the calculation of the composite coefficients and other model parameters is sensitive to the value of a parameter needed to calculate the dosage rate in terms of the total flowrate into each network node [4] (flow proportional dosage concentration, β). We assume uncertainty in the values of the parameter β, defining probability distribution functions to it. That approach allows us to quantify the potential effect of uncertainties on the water quality simulations (and, therefore, on the location of booster disinfection stations) while still preserving the simple structure of the model. Also, the proposed two-stage formulation is advantageous since it can be solved through a generalized framework based on the stochastic decomposition algorithm (SD) (See Fig. 1). The SD algorithm was developed by Higle and Sen [7] and a step by step description has been provided elsewhere [8]. We have implemented the SD

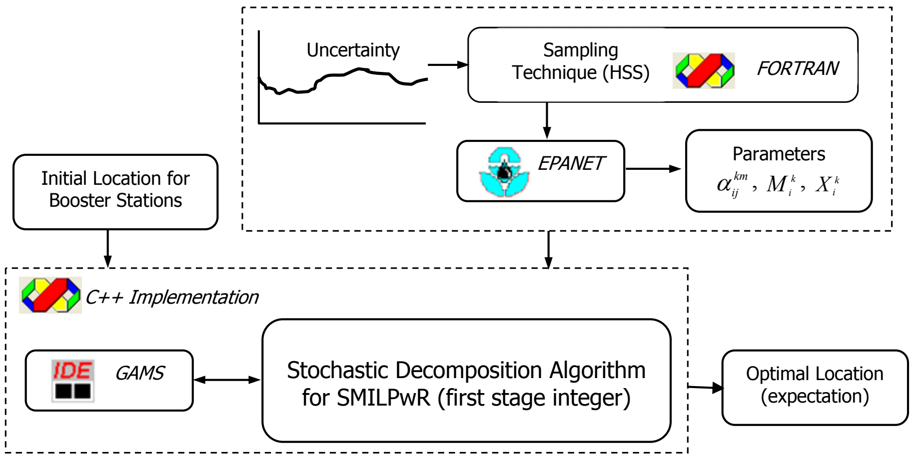

Fig. 1 Solution Approach

algorithm through a generalized framework that integrates the GAMS modeling environment, the EPANET simulator, sampling code (FORTRAN) to handle uncertainties and a C++ master program. Probability distribution functions are then defined for the β parameters of each of the nodes of the network. The Hammersley sequence sampling is used to sample the values of such parameters. The β parameters are fed to EPANET, as well as the data representing the distribution network. As a result, EPANET provides the linear coefficients of the model. The coefficients are next used by the stochastic decomposition algorithm in order to solve the stages of the model and to generate the optimality cuts to be added to the first stage. Each iteration of the stochastic decomposition algorithm involves the formulation of GAMS models (through C++ master program) which are solved through the solver OSL. The procedure continues for a large number of samples (iterations of the SD algorithm) of the β parameter for each node until convergence is achieved.

3.3. Case study

To show the quantitative effect of uncertainties in the facility location model of booster disinfection, the approach has been applied to the water distribution network provided by the EPANET simulator as Example 2 [6]. The example involves 36 nodes (34 sink nodes, 1 source node, and one pump station). Although several case studies have been solved, the case we show here assumes 1 dosage period during a monitoring time of 24 hours. The stochastic model contains 36 first stage decision variables (binary), 36 second stage decision variables and 1764 constraints. Given the distribution function of the 36 uncertain parameters, 31104 composite response parameters are computed through EPANET for each set of sampled parameters (each iteration). Three different types of probability distribution functions were assumed for the uncertain parameter: normal, triangular and uniform. Also, low and high station installation costs were analyzed to study their effect on the resulting configuration. The maximum number of stations was set as 7.

3.4. Results and analysis

As an illustration, Table 1 shows the optimal location of booster disinfection stations for the deterministic and the stochastic cases (nodes to locate the stations). Results refer to an example where the 36 uncertain parameters of the stochastic case are represented by triangular distribution functions and when the installation cost of the stations is significant. For simplicity, the defining parameters of the distributions functions are omitted here. Table 1 also presents a comparison between the stochastic and deterministic values of the objective function. Note that the value of the stochastic solution (VSS) lies between 2 and 4 %. However, there are cases in which VSS values are as high as 112%; showing the potential effect of uncertainties in the formulation.

Table 1 Illustrative results

	Stochastic (Uniform Distribution)	*Stochastic (Triangular Distribution)*	*Deterministic*
Location (nodes)	1, 9, 29, 33	1, 4, 5, 7, 14, 23, 27	1, 22, 25
Total Mass (Kg)	1.10135	0.66844	1.28344
Total Cost	1.192970E+09	1.181280E+09	1.224078E+09
VSS (%)	2.6076	3.6230	

4. Concluding remarks

This paper describes an extension to the facility location model for booster disinfection provided by Boccelli et al. [1,4]. Our model incorporates uncertainties to the model, and reformulates it as a two-stage stochastic program which is solved through a generalized computational framework base on the SD algorithm linked to the EPANET water quality simulator. The case studies considered so far confirm a significant impact of uncertainties on the optimal location of booster disinfection stations and the disinfectant mass utilized and, therefore, on the cost associated to the physical implementation. Results show VSS higher than 100% for some instances of the model parameters.

Acknowledgements

V. Rico-Ramirez acknowledges CONACYT, DGEST and ITC for their support.

References

1. D. L. Boccelli, M. E. Tryby, J. G. Uber, L. A. Rossman, M. L. Zeirof and M. M. Polycarpou, J. Water Resour. Plann. Manage., 124 (1998) 99
2. D. L. Boccelli, M. E. Tryby, J. G. Uber, L. A. Rossman and A. Lewis, In the Proceedings of the Annual Water Resour. Plann. Manage Conference, (1998) 266
3. M. E. Tryby, D. L. Boccelli, L. Dominic,M. T. Koechling, J. G. Uber, James G., R. S. Summers, L. A. Rossman and A. Lewis, J Am Water Works Assoc 91 (1999) 95
4. D. L. Boccelli, M. E. Tryby, J. G. Uber and L. A. Rossman, J. Water Resour. Plann. Manage., 128 (2002) 322
5. P. Subramaniam, Optimal Locations of Booster Stations in Water Distribution Systems, M. S. Thesis, University of Cincinnati, 2001
6. L. A. Rossman, EPANET 2: User Manual, Enviromental Protection Agency, USA, 2000
7. J. L. Higle and S. Sen, Stochastic Decomposition, Kluwer Academic Publisher, USA, 1996
8. J. M. Ponce-Ortega, V. Rico-Ramirez, S. Hernandez-Castro, and U. M. Diwekar, Comp Chem Eng, 28 (2004) 767