
APPENDIX E

**RECONSTRUCTION OF HYDRAULIC MANAGEMENT OF A
WATER-DISTRIBUTION SYSTEM USING GENETIC ALGORITHMS**

BY M.M. ARAL, J. GUAN, M.L. MASLIA, AND J.B. SAUTNER

Appendix E published by:

Multimedia Environmental Simulations Laboratory
Report No. MESL-01-01
Georgia Institute of Technology
School of Civil and Environmental Engineering
Atlanta, Georgia

October 2001

RECONSTRUCTION OF HYDRAULIC MANAGEMENT OF A WATER-DISTRIBUTION SYSTEM USING GENETIC ALGORITHMS

by M.M. Aral,¹ J. Guan,² M.L. Maslia,³ and J.B. Sautner⁴

ABSTRACT

An epidemiologic study of childhood leukemia and central nervous system cancers that occurred in the period 1979 through 1996 in Dover Township, N.J., is being conducted. The study is exploring a wide variety of possible risk factors; one being the exposure to groundwater contaminants that occurred through private and community water supplies (*i.e.*, the water-distribution system serving the area). For this purpose a model of the complex water-distribution system has been developed and calibrated through an extensive field investigation. The components of this water-distribution system, such as number of pipes, number of tanks, and number of supply wells in the network, have changed significantly over a 35-year period (1962 through 1996)—the time frame established for the epidemiologic investigation. For the completion of the study, information on monthly distribution of water in the networks, based on a management strategy, is necessary for the period 1962 through 1996. For certain months of the analysis period, some data are available to estimate the operational characteristics and the management strategy employed to operate the water-distribution system. For other months, there are no data to predict the characteristics of this management system. Further complicating the study is that within a given month of any year, the water-distribution system may have operated under peak-, winter-, or average-demand conditions. Manual reconstruction of this management system is time consuming, labor intensive, and costly, given the complexity of the system and the time constraints imposed on the study. In an effort to reduce the required computational time, the problem was formulated as an optimization problem. For each month of the investigation, the management strategy was arrived at by obtaining a solution to the optimization problem. Thus, in this study, it is assumed that the water-distribution system was operated in an optimum manner at all times to satisfy the minimum and maximum pressure constraints and tank level constraints in the system. Given these assumptions, we have used Genetic Algorithms along with the EPANET2 water-distribution network solver to solve the optimization problem and to develop the historical management strategy used by the water-distribution system serving the Dover Township area, N.J. This process reduced the required solution time significantly and generated a historically consistent management strategy for the system.

¹Director and Professor, Multimedia Environmental Simulations Laboratory, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332.

²Research Engineer, Multimedia Environmental Simulations Laboratory, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332.

³Research Environmental Engineer and Project Officer, Division of Health Assessment and Consultation, Agency for Toxic Substances and Disease Registry, Atlanta, Georgia 30333.

⁴Environmental Health Scientist, Division of Health Assessment and Consultation, Agency for Toxic Substances and Disease Registry, Atlanta, Georgia 30333.

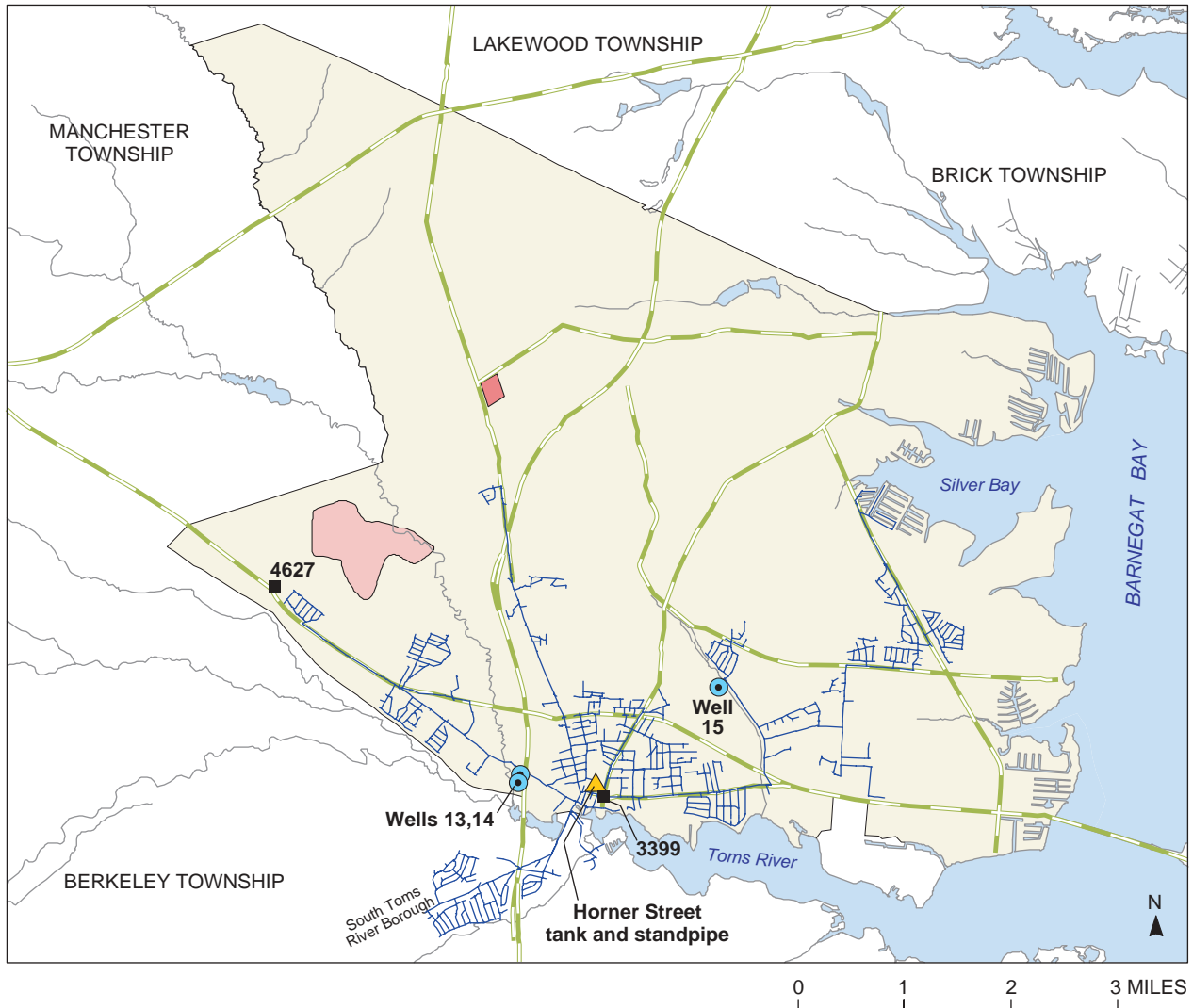
1. INTRODUCTION

An epidemiologic study of childhood leukemia and central nervous system cancers that occurred in the period 1979 through 1996 in Dover Township area of New Jersey is being conducted [Maslia *et al.*, 2000a, 2000b]. The study is exploring a wide variety of possible risk and exposure factors. Groundwater is the source of potable water in the area. Approximately 85% of the current Dover Township area residents are served by a public water supply system which relies on groundwater resources. Potential groundwater contamination of water-supply wells in Dover Township has generally consisted of volatile organic compounds (VOCs) such as trichloroethylene (TCE) and semi-volatile organic compounds (SVOCs) such as styrene-acrylonitrile (SAN) trimer [NJDHSS, 1999]. The possibility of human exposure to these contaminants exists through the groundwater pathway which was distributed by the water-distribution system. Thus, an analysis of exposure of the public to contaminants through the water-distribution system is necessary. For this purpose a computational model of the complex water-distribution system has been developed and calibrated through an extensive field investigation by the Agency for Toxic Substance and Disease Registry (ASTDR) [Maslia *et al.*, 2000a, 2000b]. For the completion of the exposure study, information on monthly distribution of water in the network based on a management strategy is necessary for the period 1962 through 1996. The components of this water-distribution system, such as number of pipes, number of tanks, and number of supply wells in the network, have changed significantly over the past 35-year period (Figure E-1 and Figure E-2). For certain months of the study period, some data are available to estimate the characteristics of the management strategy employed to operate the water-distribution system. For other months and years, these data are not available to predict the characteristics of the management system. Manual reconstruction of this management system is time consuming, labor intensive, and costly, given the complexity of the system and the time constraints imposed on the study. In an effort to reduce the time to reconstruct the management strategy, the problem was formulated as an optimization problem by defining the goal of the management system as the objective function. The goal of the management system is to satisfy the minimum and maximum pressure constraints for each junction and water level constraint for each tank in the system. For each month of the investigation period, the management strategy was arrived at by obtaining an optimal solution to the problem formulated. A Progressive Optimality Genetic Algorithm (POGA) is developed to solve the optimization problem, which utilizes the EPANET2 water-distribution network model as a simulator. This approach is employed to reconstruct a historically consistent management strategy for the water-distribution system serving the Dover Township area, New Jersey.











2. Formulation of the Optimization Problem

2.1 Decision Variables and Constraints

A model of the water-distribution system serving Dover Township, N.J., has been developed and calibrated through an extensive field investigation. An EPANET2 input data file has been generated for each month for the period 1962 through 1996 by ATSDR; the data file generates an error-free run for the pipe network system. These EPANET2 input data files are the basis of our reconstruction of the management strategy. The management strategy of the water-distribution system includes the operation of tanks, booster pumps and water-supply wells in the network. In addition to other parameters defined in an EPANET2 input data file, the management strategy of tanks and booster pumps are represented by the “control rules.” The management strategy of the water-supply wells are represented by the “patterns” associated with these wells [Rossman, 2000]. The “control rules” represent the status (close, open, and pumping efficiency) of the booster pumps during the operation of the water-distribution system. A “pattern” in an EPANET2 input data file represents the water demand over time at a junction (node) of the network based on the “base demand” of the junctions [Rossman, 2000]. A “pattern” in an EPANET2 input data file includes one or more “pattern factors.” The number of such “pattern factors” in a “pattern” is generally equal to the number of time intervals in the operation period. In this study, the number of time intervals is always defined to be 24 in a one-day period. For instance, if the “base demand” of a junction is 100 gallons per minute (gpm) and the “pattern factor” of the junction at a time interval t is 0.5, the actual water demand of the junction at the time interval t is $(100 \times 0.5 = 50)$ gpm. In an EPANET2 input data file for the Dover Township area water-distribution system, the “pattern” is assumed to be the same for all junctions except the junctions representing water-supply wells, and each water-supply

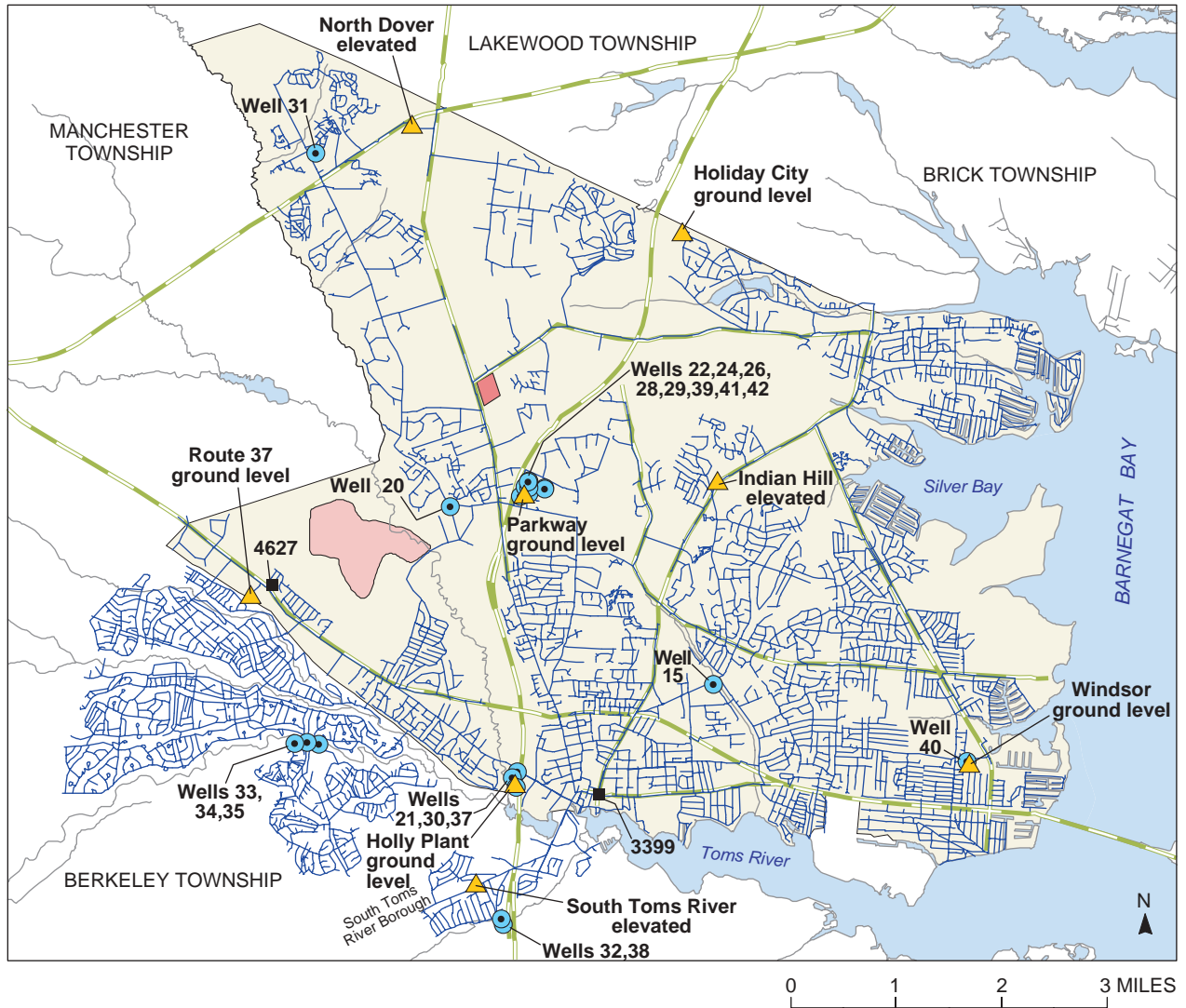


EXPLANATION

	Reich Farm NPL Site		Water pipeline		Municipal well
	Ciba-Geigy NPL Site		Major road		Storage tank
	Dover Township		Hydrography		4627 Junction location and number
	Water body				

- Notes:
- (1) Water pipelines range in diameter from 2 inches to 12 inches
 - (2) Roads, hydrography, and boundaries based on 1995 TIGER/Line data
 - (3) Pipeline from water-utility database (Flegal 1997)

Figure E-1. Water-distribution system serving the Dover Township area, New Jersey, 1962.



EXPLANATION

- | | | |
|---------------------|----------------|-----------------------------------|
| Reich Farm NPL Site | Water pipeline | Municipal well |
| Ciba-Geigy NPL Site | Major road | Storage tank |
| Dover Township | Hydrography | 4627 Junction location and number |
| Water body | | |

Notes: (1) Water pipelines range in diameter from 2 inches to 16 inches
 (2) Roads, hydrography, and boundaries based on 1995 TIGER/Line data
 (3) Pipeline from water-utility database (Flegal 1997)

Figure E-2. Water-distribution system serving the Dover Township area, New Jersey, 1996.

well has its own pattern. The major task for the reconstruction of the management strategy is to determine these patterns for the water-supply wells to satisfy the minimum and maximum pressure constraints for each junction and water level constraint for each tank in the network.

If there are N water-supply wells in the water-distribution system, we may denote these wells as $n = 1, 2, \dots, N$. Each water-supply well has its own supply pattern. The pattern is defined for a 24-hour period of one day. Each pattern consists of 24 “pattern factors” and each “pattern factor” represents the water supplied by the water-supply well during that hour. Therefore, there are 24 “pattern factors” for each water-supply well. We may denote these “pattern factors” as $x_i(t)$ ($i = 1, 2, \dots, N, t = 1, 2, \dots, 24$), where subscript i represents the well number and t represents the time (hour). Because the total water supplied by each water-supply well in the period of one day is known, the following constraints should then be applied to these “pattern factors.”

$$\sum_{t=1}^{24} x_i(t) = b_i \quad i = 1, 2, \dots, N \quad (\text{E-1})$$

where b_i is the sum of the “pattern factors” of the i^{th} well during a one-day period, which is a known constant because the total supply of the well is known for a day.

If the i^{th} water-supply well is not operated at hour t , the “pattern factor” of the well i at hour t is zero, $x_i(t) = 0$. If the i^{th} water-supply well is in operation at the hour t , the “pattern factor” of the i^{th} well at hour t , $x_i(t)$ should be larger than a minimum value and smaller than a maximum value representing the operational range of the water-supply well. These constraints for each “pattern factor” $x_i(t)$ ($i = 1, 2, \dots, N, t = 1, 2, \dots, 24$) can be represented as

$$x_i(t) = 0 \quad \text{or} \quad x_{i_{\min}} \leq x_i(t) \leq x_{i_{\max}} \quad (\text{E-2})$$

where $x_{i_{\min}}$ is the minimum and $x_{i_{\max}}$ is the maximum “pattern factor” of the i^{th} water-supply well. The pattern factors can be written in vector form as

$$\mathbf{X} = [x_1(1), x_1(2), \dots, x_1(24), \dots, x_i(1), x_i(2), \dots, x_i(24), \dots, x_N(1), \dots, x_N(24)]^T \quad (\text{E-3})$$

In this study, we assume the “control rules,” which control the operation of the booster pumps are predefined based on the hydraulic characteristics of the equipment, and the variables described above (pattern factors) are the only decision variables for our optimization problem to reconstruct the management strategy of the water-distribution system.

2.2 Objective Function

The goal of the reconstruction of the management strategy is to determine the “pattern factors” for the water-supply wells to satisfy the water pressure requirements at every junction (node) and water level constraints at every tank in the water-distribution system. Suppose there are M junctions in the system, one of the objectives is to satisfy the minimum and maximum pressure constraints as

$$P_{\min} \leq P_i(t) \leq P_{\max} \quad i = 1, 2, \dots, M, \quad t = 1, 2, \dots, 24 \quad (\text{E-4})$$

where $P_i(t)$ is the water pressure at the i^{th} junction at hour t , P_{\min} is the minimum pressure required at a junction, and P_{\max} is the maximum pressure a junction may bear.

The other objective is to satisfy the water level constraints in tanks. The tank operation rules require that the water level in any tank should always be lower than its highest allowable water level and higher than its lowest allowable water level. If we assume there are T tanks in the system, then, we can define these constraints as

$$L_{jmin} \leq L_j(t) \leq L_{jmax} \quad j = 1, 2, \dots, T; \quad t = 1, 2, \dots, 24 \quad (E-5)$$

where $L_j(t)$ is the water level in the j^{th} tank at hour t and L_{jmin} and L_{jmax} are the lowest and highest allowable water levels of the j^{th} tank.

In addition to the above two objectives, the initial ($t = 0$) water level in a tank is constrained to be equal to the water level in the tank at the end of the day ($t = 24$), because continuous simulations over a 24-hour period will be required during subsequent analysis for the epidemiologic study. We can define these constraints as

$$L_j(24) = L_j(0) \quad j = 1, 2, \dots, T \quad (E-6)$$

In Equations (E-4) to (E-6), the variables, $P_i(t)$ and $L_j(t)$ are functions of the decision variables, $x_i(t)$. They can be obtained by solving the hydraulic simulation of the water-distribution system using EPANET2 water-distribution model. Due to the complexity of the network and the EPANET2 water-distribution model, $P_i(t)$ and $L_j(t)$ cannot be represented as explicit functions of the decision variables. Thus, it may not be feasible to enforce the constraints of Equations (E-4) to (E-6) directly in the hydraulic management-reconstruction problem. In order to best satisfy the constraints, we will formulate these constraints into an objective function by imposing a penalty for the violation of the constraints. By minimizing the objective function with the constraints of Equations (E-1) and (E-2), we obtain the “pattern factors” of the water-supply wells in the system with minimized violation of constraints (Equations (E-4) to (E-6)).

If the water pressure of a junction at time t , $P_i(t)$, satisfies Equation (E-4), there is no penalty, *i.e.* the “penalty” is assumed to be 0. If the water pressure at the junction is out of the constraint range, a penalty will be imposed. We may define the “penalty” at the i^{th} junction at time t , $C_i(t)^J$, to be

$$C_i(t)^J = \begin{cases} 0 & P_{\min} \leq P_i(t) \leq P_{\max} \\ [P_{\min} - P_i(t)]^2 & P_i(t) < P_{\min} \\ [P_i(t) - P_{\max}]^2 & P_i(t) > P_{\max} \end{cases} \quad (E-7)$$

Therefore, the “penalty” for the i^{th} junction for a day can be defined as

$$C_i^J = \sum_{t=1}^{24} C_i(t)^J \quad i = 1, 2, \dots, M \quad (E-8)$$

Similar to junctions, if the water level in the j^{th} tank at time t , $L_j(t)$ satisfies Equation (E-5), the “penalty” is defined as zero. Otherwise, a penalty is imposed. Thus the “penalty” function for the j^{th} tank at hour t , $C_j(t)^T$, can be defined as,

$$C_j(t)^T = \begin{cases} 0 & L_{jmin} \leq L_j(t) \leq L_{jmax} \\ a_T [L_{jmin} - L_j(t)]^2 & L_j(t) < L_{jmin} \\ a_T [L_j(t) - L_{jmax}]^2 & L_j(t) > L_{jmax} \end{cases} \quad (E-9)$$

where a_T is the penalty factor for a tank violating the water level requirement. Similarly the total penalty for the j^{th} tank for a day can be defined as,

$$C_j^T = \sum_{t=1}^{24} C_j(t)^T \quad j = 1, 2, \dots, T \quad (\text{E-10})$$

If the water level in the j^{th} tank at time $t = 24$ hour is equal to the water level in the tank at time $t = 0$, Equation (E-6) is satisfied. Otherwise, a penalty is imposed and the penalty function, C_j^{TI} is defined as,

$$C_j^{TI} = a_{TI} [L_j(24) - L_j(0)]^2 \quad j = 1, 2, \dots, T \quad (\text{E-11})$$

where a_{TI} is the penalty factor for violating the constraint of Equation (E-6).

The objective function of the optimization problem for hydraulic management reconstruction is defined in terms of the total penalty function for the system, which can be given as,

$$C = \sum_{i=1}^M C_i^J + \sum_{j=1}^T (C_j^T + C_j^{TI}) \quad (\text{E-12})$$

2.3 EPANET2 Water-Distribution System Model

EPANET2 is a computer program that performs extended period simulation of hydraulic and water quality behavior within pressurized pipe networks [Rossman, 2000]. The EPANET2 has been used to simulate water-distribution systems in numerous field applications in the literature including [Aral 1996, 2001; and Maslia, *et. al*, 2000a, 2000b]. In our reconstruction of the management strategy, the EPANET2 code is being used to model the water-distribution system serving the residents of the Dover Township area. The EPANET2 toolkit will be used to obtain the hydraulic simulation parameters of the water-distribution system. Mathematically, the EPANET2 toolkit works as a function that associates the decision variables (the “pattern factors”), X , with junction pressures, $P_i(t)$ ($i = 1, 2, \dots, M; t = 1, 2, \dots, 24$) and water levels in tanks $L_j(t)$ ($j = 1, 2, \dots, T; t = 1, 2, \dots, 24$). For the given values of the decision variables X , the $P_i(t)$ and $L_j(t)$ can be obtained by the EPANET2 toolkit. Thus, we have mathematically

$$\begin{aligned} P_i(t) &= F_{it}^J(X) \quad i = 1, 2, \dots, M, \quad t = 1, 2, \dots, 24 \\ L_j(t) &= F_{jt}^T(X) \quad j = 1, 2, \dots, T, \quad t = 1, 2, \dots, 24 \end{aligned} \quad (\text{E-13})$$

2.4 Optimization Model

Based on the above discussion, the objective function C of Equation (E-12) is a function of the decision variables, X . Thus, we may write the objective function as

$$C = f(X) \quad (\text{E-14})$$

The reconstruction of the optimal hydraulic management strategy can now be formulated as

(E-15)

$$\begin{aligned}
 & \text{Minimize } \{ f(X) \} \\
 & \sum_{t=1}^{24} x_i(t) = b_i \quad i = 1, 2, \dots, N \\
 & x_i(t) = 0 \quad \text{or} \quad x_{i \min} \leq x_i(t) \leq x_{i \max} \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, 24
 \end{aligned}$$

3. Non-Uniqueness of Optimal Solutions

The analysis presented above will be used to generate the optimal operational strategy for the water-distribution system. During a subsequent study, this operational strategy will be used to evaluate the contaminant transport characteristics of the water-distribution system. The purpose in that study is to evaluate the effect of the operation of the water-distribution system on the contaminant transport conditions that will be generated in the system. So far we have described the operation of the water-distribution system in terms of two patterns. These patterns are generated using the manual trial and error approach and the optimization approach. However, the management strategy obtained from the above described mathematical model may not be unique. Thus, it may be argued that there are other hydraulic management strategies that may also satisfy the recorded constraints, which may not be optimal. In turn, this management strategy may yield different results, when proportionate contribution of water analysis for the system is evaluated. Therefore, based on the model described above, another mathematical model is used to identify another solution which yields a pattern distribution as different as possible from the first solution, while still satisfying the constraints of the problem. Eventually, the purpose is to find out if this new solution would yield a significantly different proportionate contribution of water distribution pattern within the system and, if it does, what would be the difference.

To start this formulation, we begin with the optimal solution obtained from the first model, defined as

$$X^0 = [x_1^0(1), x_1^0(2), \dots, x_1^0(24), \dots, x_i^0(1), x_i^0(2), \dots, x_i^0(24), \dots, x_N^0(1), \dots, x_N^0(24)]^T \quad (E-16)$$

and the corresponding optimal objective function value as $f(X^0)$ or f^0 . The new solution is still defined as given in Equation (E-3) and the objective function is expressed using Equation (E-14) as $f(X)$ or f . As expected, the new solution has an objective value that is equal to or less than $f(X^0)$. However, we would like to have the new solution pattern for all water-supply wells to be different from X^0 as much as possible. Therefore, this requirement can be quantitatively stated as a new objective function,

$$\text{Maximize } \{ F(X) \} = \text{Maximize } \left\{ \beta \sum_{i=1}^N \sum_{t=1}^{24} (x_i(t) - x_i^0(t))^2 - G(X) \right\} \quad (E-17)$$

where $G(X)$ is the term related to the previous objective value and defined as

$$G(X) = \begin{cases} (f(X) - f(X^0))^2 & f(X) - f(X^0) > \delta \\ 0 & \text{otherwise} \end{cases} \quad (E-18)$$

β is a trade-off coefficient, δ is an allowable error. Equation (E-18) reflects the tolerance for a new solution. In order to analyze and compare the solutions obtained from both models, the constraints of the new model must be the same as the ones used in the previous model. Thus, the optimization model for selecting a new solution can be mathematically stated as

$$\begin{aligned}
 & \text{Maximize } \{ F(X) \} \\
 & \sum_{t=1}^{24} x_i(t) = b_i \quad i = 1, 2, \dots, N \\
 & x_i(t) = 0 \quad \text{or} \quad x_{i \min} \leq x_i(t) \leq x_{i \max} \quad i = 1, 2, \dots, N \quad t = 1, 2, \dots, 24
 \end{aligned} \tag{E-19}$$

4. Solution Algorithm

The mathematical models given in Equations (E-15) and (E-19) are nonlinear multi-stage multi-dimensional dynamic optimization problems. In general, for the solution of dynamic problems, the Dynamic Programming (DP) method [Bellman, 1957], provides a powerful tool because it has attractive features such as global optimal solution and its unique generality. However, in this problem, the dimension of the optimal problem is defined in terms of the number of the water-supply wells. The number of the water-supply wells in the water-distribution system is very large, thus the “curse of dimensionality” in DP programming methods becomes a serious limitation. Furthermore, the computational time for the evaluation of objective function mainly depends on the scale of water-distribution system and is almost independent of the number of water-supply wells in the system. For a large-scale water-distribution system, the question as to how to reduce the computational time required in the evaluation of the objective function becomes important. This is necessary to improve the optimization algorithm performance. It should also be noted that the objective function is a non-continuous and non-separate implicit function of the decision variables. Thus, the nonlinear gradient-based search algorithms cannot be applied to this problem efficiently. In recent years, a new breed of heuristic combinatorial search techniques—genetic algorithms (GAs)—have aroused the interest of scientists in searching optimal solutions to complicated optimization problems. Genetic algorithms are established through the combination of artificial survival of the fittest concept with genetic operations abstracted from nature [Holland, 1975]. The most important feature of genetic algorithms is that they use the objective function itself, instead of derivative information on the objective function and constraints, to search the optimal solution. It is well known that genetic algorithms have proved to be very efficient in solving problems with bounds. For the problems with equality and inequality constraints, additional transformations should be done [Guan and Aral, 1999a, 1999b]. Considering the complex nature of both optimization problems and the features of genetic algorithms, a new approach, that embeds genetic algorithm (GA) in progressive optimality algorithm (POA) is developed. The resulting algorithm is identified as Progressive Optimality Genetic Algorithm (POGA) and is proposed as the method to solve the problems described above. Considering that the difference between both models is only in the definition of the objective functions, and has no effect on the optimization algorithms used, the model described by Equation (E-15) is chosen as an example to illustrate the principles and processes included in POGA.

4.1 Progressive Optimality Genetic Algorithm (POGA)

Progressive optimality algorithm is based on “principle of optimality” [Bellman, 1957]. Its basic idea of searching optimal solution is straightforward. In this approach, the optimization process is divided into two steps: decomposition and iteration. In the decomposition step, POGA decomposes a multi-stage multi-dimensional dynamic problem into a series of two-stage multi-variable static problems, which can be solved by static optimization methods such as the linear and nonlinear programming and genetic algorithms. In the iteration step, POGA applies the iteration mechanics to obtain a converged solution. For simplicity and without loss of generality, an example with one water-supply well with index “ i ” is chosen to illustrate the computational procedures of the POGA.

Step 1. A feasible solution is assumed as initial solution and is denoted as $[x_i(1)^{(k)}, x_i(2)^{(k)}, \dots, x_i(24)^{(k)}]^T$, where superscript k indicates index of iterations, where initially $k = 0$.

Step 2. Under the condition of fixed $x_i(\tau)^{(k)}, \forall \tau$ except $x_i(t)^{(k)}$ and $x_i(t+1)^{(k)}$, one adjusts $x_i(t)^{(k)}$ and $x_i(t+1)^{(k)}$ to minimize the objective function. The resultant new solution is $[x_i(1)^{(k)}, x_i(2)^{(k)}, \dots, x_i(t)^{(k+1)}, x_i(t+1)^{(k+1)}, \dots, x_i(24)^{(k+1)}]^T$.

Step 3. For $t = 1$ to 23, one repeats step 2 and obtains a new solution denoted as $[x_i(1)^{(k+1)}, x_i(2)^{(k+1)}, \dots, x_i(24)^{(k+1)}]^T$.

Step 4. Compare two consecutive solutions $[x_i(1)^{(k)}, x_i(2)^{(k)}, \dots, x_i(24)^{(k)}]^T$, and $[x_i(1)^{(k+1)}, x_i(2)^{(k+1)}, \dots, x_i(24)^{(k+1)}]^T$. If the difference between these two solutions is smaller than a pre-specified accuracy limit, the computation process terminates. Otherwise, one selects $[x_i(1)^{(k+1)}, x_i(2)^{(k+1)}, \dots, x_i(24)^{(k+1)}]^T$ as the new initial solution, *i.e.*, $k = k + 1$ and continues with steps 2 through 4 until convergence is achieved.

As demonstrated in the application of this algorithm, the objective function values in the iteration process are monotonically decreasing, and the minimum objective function value in this problem is zero if all constraints are satisfied. This indicates that the iteration process is convergent.

4.2 Optimization Algorithm in Two-Stage Problems

In two-stage optimization problems, genetic algorithm is applied to search for the optimal solution. Assuming that the two stages are t and $t+1$, $X(t)$ and $X(t+1)$ are two vectors denoted as $X(t) = [x_1(t), x_2(t), \dots, x_N(t)]^T$, and $X(t+1) = [x_1(t+1), x_2(t+1), \dots, x_N(t+1)]^T$, then the original optimization problem can be restated as

$$\text{Minimize } \{C = f(X(t), X(t+1); X(\tau), \tau \neq t, t+1)\}$$

$$x_i(t) = b_i - x_i(t+1) - \sum_{\substack{\tau=1 \\ \tau \neq t, t+1}}^{24} x_i(\tau) \quad (\text{E-20})$$

$$x_{i \min}(t) \leq x_i(t) \leq x_{i \max}(t)$$

$$x_{i \min}(t+1) \leq x_i(t+1) \leq x_{i \max}(t+1)$$

$$i = 1, 2, \dots, N$$

where $x_i(\tau), \tau = 1, 2, \dots, 24$, but $\tau \neq t$ and $t+1$, are constants evaluated by previous iteration or initial solution. The model can be further simplified as

$$\text{Minimize } \{C = f(X(t), X(t+1))\}$$

$$x_i(t) = d_i - x_i(t+1)$$

$$x_{i \min}(t) \leq x_i(t) \leq x_{i \max}(t) \quad (\text{E-21})$$

$$x_{i \min}(t+1) \leq x_i(t+1) \leq x_{i \max}(t+1)$$

$$i = 1, 2, \dots, N$$

where

$$d_i = b_i - \sum_{\substack{\tau=1 \\ \tau \neq t, t+1}}^{24} x_i(\tau) \quad (\text{E-22})$$

Through this approach, the original problem becomes a static optimization problem with $2 \times N$ variables. Because of the equality constraints, the number of independent variables is N . Without loss of generality, $x_i(t+1) \forall i$ are taken as the independent variables, and $x_i(t) \forall i$ as the dependent variables, which can be directly calculated using the equality constraints. In the application of GA, the random selection of $x_i(t+1)$ must yield $x_i(t)$ variables that satisfy the bounds of this variable. To achieve this condition, the constraint for $x_i(t+1)$ can be given as,

$$d_i - x_{i \max}(t) \leq x_i(t+1) \leq d_i - x_{i \min}(t) \quad (\text{E-23})$$

At the same time, $x_i(t+1)$ must also satisfy $x_{i \min}(t+1) \leq x_i(t+1) \leq x_{i \max}(t+1)$. Synthetically considering these two constraints, a new constraint can be obtained as,

$$x_{i \text{low}}(t+1) \leq x_i(t+1) \leq x_{i \text{up}}(t+1) \quad (\text{E-24})$$

where

$$\begin{aligned} x_{i \text{low}}(t+1) &= \max\{d_i - x_{i \max}(t), x_{i \min}(t+1)\} \\ x_{i \text{up}}(t+1) &= \min\{d_i - x_{i \min}(t), x_{i \max}(t+1)\} \end{aligned} \quad (\text{E-25})$$

Therefore, the two-stage problem can be simplified as,

$$\begin{aligned} & \text{Minimize } \{C = f(X(t), X(t+1))\} \\ & x_{i \text{low}}(t+1) \leq x_i(t+1) \leq x_{i \text{up}}(t+1) \\ & x_i(t) = d_i - x_i(t+1) \\ & i = 1, 2, \dots, N \end{aligned} \quad (\text{E-26})$$

This is a bounded optimization problem, which can be solved by genetic algorithms efficiently [Guan and Aral, 1999a, 1999b]. The application of GA is straightforward for the optimal solution of this problem. GA starts with an initial population of members generated randomly. Three operators, identified as “selection,” “crossover,” and “mutation,” abstracted from biologically evolutionary process, are applied to this population in improving the objective function value to form the new generation, in which the members have a higher “quality.” Each member in the population corresponds to a solution in the solution space. Member’s quality is represented by its fitness associated with the objective function value. The principle of “survival of the fittest” is taken as a rule in the search process. Members of the population are improved from generation to generation until an optimum condition is satisfied [Goldberg, 1989]. These procedures of the genetic algorithm used in the solution of this problem are described below for completeness.

Division of Variables: variables $x_i(t)$ and $x_i(t+1) \forall i$ are divided into two parts, *i.e.*, the independent and dependent variables. In this formulation $x_i(t+1) \forall i$ are defined as the independent variables, $x_i(t) \forall i$ are defined as the dependent variables.

Coding the problem: in GAs, the independent variables are encoded as binary strings. The bits of the encoded variables are combined to simulate the biological chromosomes of a natural environment. If a k -bit binary variable is used to represent one variable $x_i(t+1)$, the integers of the binary variable would range from 0 to (2^k-1) . This range can be mapped linearly into the solution space $[x_{i\ low}(t+1), x_{i\ up}(t+1)]$ and the variable range will be discretized into 2^k points. The discrete interval can now be given as

$$\Delta x_i(t+1) = \frac{x_{i\ up}(t+1) - x_{i\ low}(t+1)}{2^k - 1} \quad (E-27)$$

If the integer of binary variable is “m,” then the corresponding variable $x_i(t+1)$ value can be given by

$$x_i(t+1) = x_{i\ low}(t+1) + m\Delta x_i(t+1) \quad (E-28)$$

If $x_i(t+1)$ is less than the minimum load constraint of water-supply wells, then let $x_i(t+1) = 0$.

Forming the initial population: to start the solution, an initial population with m members is created. For the independent variables $x_i(t+1) \forall i$, integer values are generated randomly in the interval $[0, 2^k-1]$ with uniform distribution. These integer values are then mapped into real area of $x_i(t+1)$ using Equation (E-27) given above. The dependent variables $x_i(t)$ are calculated using the equality constraint, together with $x_i(t+1)$, forming the initial population. In order to guarantee the objective function to be monotonic, the initial population should include the last solution.

Evaluation of members in the population and choosing parents: the “quality” of the members of a population is measured by their fitness values. The fitness of a member is a nonnegative function, which is determined by the objective function value. In this problem, objective function is minimized. The fitness of each member can be defined by

$$F_j = f_{\max}(X(t), X(t+1)) - f_j(X(t), X(t+1)) \quad (E-29)$$

where j indicates the index of the members in the population. Relative fitness or probability of selection (ps_j) is defined by

$$ps_j = \frac{F_j}{\sum_{l=1}^m F_l} \quad (E-30)$$

The probability of selection value may then be used to select strings from the current population to form a mating pool. The larger the fitness, the larger the probability that the corresponding member is selected as parents. Thus, members with higher relative fitness values are more likely to be selected for mating. The roulette wheel selection may be taken as a method for choosing members of the mating pool.

Forming a new population: The next population is formed by using the three standard operations of GAs. These include direct selection, crossover and adaptive mutation processes.

Direct Selection Operation: Members with the highest fitness enter directly to the next population at some proportion. These members may also be selected as parents for the mating pool. Thus, using this selection, the best members are protected, and at the same time this selection assures the monotonic increase of the objective function values of the best member during each generation.

Crossover Operation: Crossover is a principle genetic operation. It is the splicing of two parent members from mating pool, at a randomly chosen point, to create two children strings each made up of “genetic material” from both parents. The crossover operation can be easily realized by computer bit operations.

Adaptive Mutation Operation: Mutation operation randomly changes bits within newly created strings. In this operation zero/one bits are changed to one/zero bits respectively in a random manner based on a mutation probability value. Mutation maintains variability in the population, and reduces the chance that the population will prematurely converge on one possible sub-optimal solution. For standard GAs, the probability of any member in the population to be selected to undergo mutation is the same, that is,

$$p_i = \frac{1}{m} ; \quad i = 1, 2, \dots, m \quad (\text{E-31})$$

This process may also be improved if a member with smaller fitness value has a larger probability of mutation than one with a larger fitness value. Therefore, the following steps are used to determine the probability of mutation in GA; (i) sort the members of the population according to their fitness value from the smallest to the largest; and (ii) assign the largest probability value to first member by,

$$p_i = \frac{c}{m} ; \quad c > 1 \quad (\text{E-32})$$

and the smallest probability value to final member in the list by,

$$p_m = p_{\min} = \frac{2-c}{m} \quad (\text{E-33})$$

The probability values for other members of the population are linearly interpolated as,

$$p_j = p_1 + \frac{p_m - p_1}{m-1}(j-1) = \frac{c}{m} + \frac{2(1-c)}{m(m-1)}(j-1) \quad (\text{E-34})$$

In these equations the condition given below will be observed,

$$\sum_{j=1}^m p_j = 1; \quad p_j \geq 0 \quad (\text{E-35})$$

Because the interval for c is $1 \leq c \leq 2$, then selecting $c = 1$ implies $p_j = 1/m$, which is the standard genetic algorithm rule. Selecting $c = 2$ yields

$$p_1 = \frac{2}{m}; \quad p_m = 0; \quad p_j = \frac{2(m-j)}{m(m-1)} \quad (\text{E-36})$$

If this equation is used to assign the probabilities for the selection of members that will undergo mutation, then it is assured that the member with the lowest fitness has the highest probability of mutation, and the member with the highest fitness has a zero probability of mutation. This implies that the worst members are more likely to be selected for mutation and the best members do not undergo mutation. This adaptive mutation process helps to protect best members from mutation and improves worst members through mutation.

The three operations are applied to the current population to form next population of $x_i(t+1)$. For each member in next population, $x_i(t)$ can be computed by the equality constraint, forming a new generation with higher quality, together with $x_i(t+1)$. $x_i(t)$ and $x_i(t+1)$, together with $x_i(\tau)$, $\tau = 1, 2, \dots, 24$, but $\tau \neq t$ and $t+1$ determine the fitness of each member in next generation.

The steps described above are repeated from generation to generation for a specified number of generations or until some stopping criterion, such as convergence of the population to a single solution is achieved. These procedures yield a very effective solution strategy for optimization problems, where the solution domain is bounded. The flowchart of the procedure for progressive optimality genetic algorithm is shown in Figure E-3.

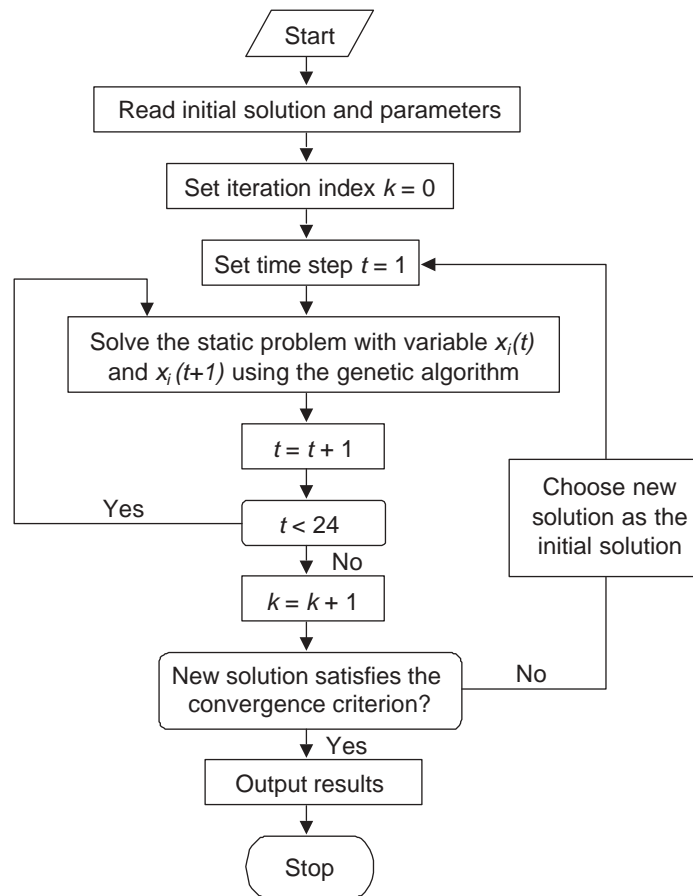


Figure E-3. Flowchart of progressive optimality genetic algorithm.

5. Integrated System of POGA and EPANET2

The POGA has been employed as an effective optimization algorithm to reconstruct hydraulic management strategy in a water-distribution system and to select new patterns based on the optimized solution for analyzing the effect of non-uniqueness of solutions. In POGA, the pressure distribution on the node network and the operation processes of tanks are calculated using the functions in EPANET2 dynamic link library. EPANET2 software is a powerful tool for performing extended period simulation of hydraulic and water quality behavior within pressurized pipe networks. EPANET2 is taken as a simulator, which can be either operated independently or embedded in POGA software for analyzing the effect of optimized results on hydraulic and water quality behavior. The manually constructed historical datasets are the starting point of the optimization process. The optimized datasets provide a consistent management strategy for reconstruction of the hydraulic management of the water-distribution system. The optimized results are used as the input to reconstruct the historical contaminant distribution pattern in the water-distribution system. The components of POGA are integrated to form an integral decision support system. The system architecture is shown in Figure E-4.

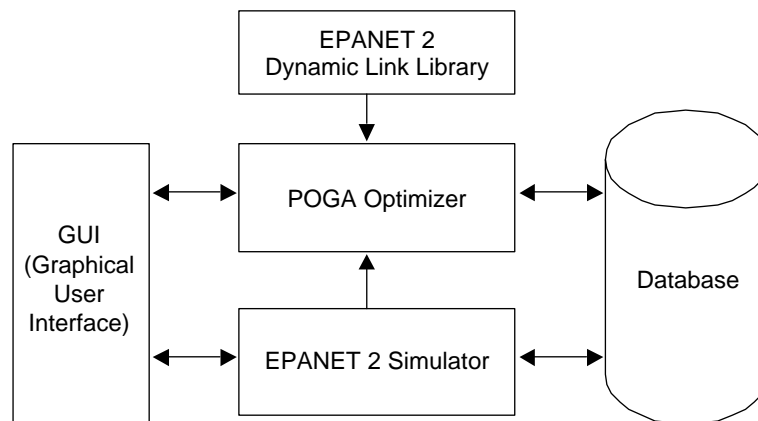


Figure E-4. The integrated system architecture for the Progressive Optimality Genetic Algorithm (POGA).

6. Numerical Results

To demonstrate the models and algorithms described above, the integrated system of POGA and EPANET2 is applied to the water-distribution system serving the Dover Township area to reconstruct its hydraulic management strategy. The study period spans from 1962 through 1996. The components of the water-distribution system have been significantly changed during this period. For example, the water-distribution system contained only one water storage tank and three water-supply wells in 1962, as shown in Figure E-1, and became a very complex system with 9 water storage tanks and 20 water-supply wells in 1996, as shown in Figure E-2. Through extensive field investigation, an initial EPANET2 input file has been generated for each month of every year by ATSDR. A corresponding text file is also generated for each month to include the specific constraint information for the water-supply wells and tanks in the water-distribution system for the particular period. After the POGA and EPANET2 are applied to the monthly data, both an optimized EPANET2 input data file and a text file including input parameters and information on iteration process are generated for the purpose of further analysis. The optimized EPANET2 input data files represent the optimized hydraulic management strategy for each month. The optimized hydraulic management strategy generally results in much smaller water pressure violation at the network junctions in Equation (E-7) as well as water level violation in tanks in Equations (E-9) and (E-11) than the initial management strategy presented in the initial EPANET2 input data files. Both the calculation of optimal pattern and the selection of new patterns are addressed in this section.

Parameter Setting: The integrated system requires two categories of parameters—water-distribution system parameters and POGA parameters. According to the pressure requirements for the Dover Township water-distribution system, the minimum pressure limit, P_{min} , is selected as 15 pounds per square inch (psi), and the maximum pressure limitation, P_{max} , is selected as 110 psi. The penalty coefficient for water-level violation in a tank in Equation (E-9), a_T , is selected to be 1.0 and the penalty coefficient for the initial water level violation in a tank in Equation (E-11), a_{TL} , is selected to be 5000. The selection of these penalty coefficients may affect the optimization results because the objective function is dependent on these coefficients. The larger the penalty coefficients a_T and a_{TL} are, the more the water-level violation in the tanks weighs in the objective function. Other GA parameters used in POGA are listed in Table E-1.

Table E-1. Parameters used in Progressive Optimality Genetic Algorithm (POGA)

Parameter	Value
Population size	5
String length for each variable	10
Maximum generation	2
Crossover ratio	0.9
Direct selection ratio	0.1
Mutation ratio	0.1
Beta (β)	100
Delta (δ)	0

Calculation of Optimal Patterns: The optimal patterns are calculated using two data sets. The first data set represents the water-distribution system serving the Dover Township in May 1962, which contained 2,272 junctions, 2,403 pipe links, one water storage tank and 3 water-supply wells as shown in Figure E-1. Optimization calculation takes four iterations. The initial and optimal operation patterns of water-supply wells are shown in Figure E-5. The optimal patterns are obtained through slight adjustment based on the initial patterns using POGA. However, junction pressure violation cost is largely reduced. The relationship between objective function, time step and iterations are shown in Figure E-6. In this case, tank violation cost is very small, junction violation cost is 3436.8 initially. The first iteration reduces junction violation cost from 3436.8 to 470.25. Consecutively these values are reduced to 61.19, 8.91, and finally to 0.21. The objective function is convergent and also monotonically reduces throughout the computation. The results imply that the water pressure of each junction in the system would stay in the allowed water pressure range (15 psi to 110 psi) if the optimized hydraulic management strategy were employed in the system operation. The water pressure as function of time in a one-day period for the junction #3399 and junction #4627 clearly shows that the optimized data set is much better than the initial data set in terms of water pressure violation as shown in Figures E-7 and E-8. Using the initial patterns, the water pressure at junction # 3399 is lower than the lower pressure limit (15 psi) from 8:00 AM to 12:00 PM and from 10:00 PM to 12:00 AM. However, using the optimal patterns, the water pressure at that junction is always higher than the lower pressure limit (15 psi). Similarly, the water pressure at junction # 4627 is lower than 15 psi from 6:00 AM to 14 PM using the initial patterns, while the water pressure is higher than 15 psi all the time using the optimal patterns. For both the initial and the optimal patterns, the water level in tank 33934-HORTA stays within its lower level and upper level limits at all times, and the water level in the tank at the beginning of the day equals to that at the end of the day for this case as shown in Figure E-9. However, the tank operates on a higher water head using the optimal patterns than using the initial patterns, increasing the efficiency of the tank.

The second data set represents the water-distribution system serving the Dover Township area in October 1996. This network shows significant development and now contains 14,965 junctions, 16,048 pipe links, 8 water-storage tanks and 20 water-supply wells (Figure E-2). Optimization calculation also takes four iterations. The initial and optimal operation patterns of three of the water-supply wells are shown in Figure E-10. Because there is a large decision space in this case, the optimal patterns have significantly changed when compared with the initial patterns. In this case, junction violation, tank violation costs and objective value are initially 42,728.6, 1585.72 and 44314.3,

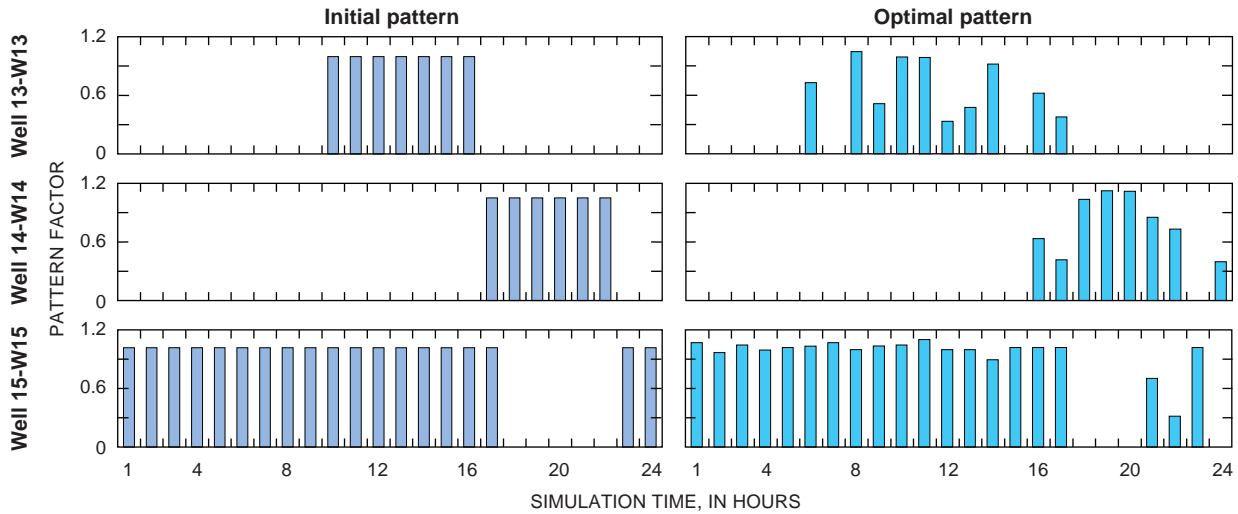


Figure E-5. Initial and optimal operation patterns of water-supply wells, May 1962. (See Figure E-1 for well locations.)

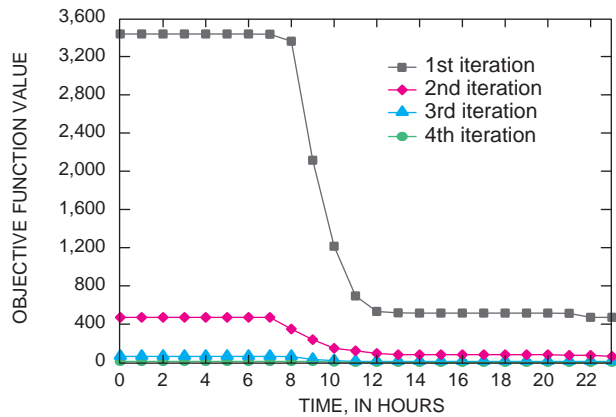


Figure E-6. Convergence of objective function as function of iteration number and time step, May 1962.

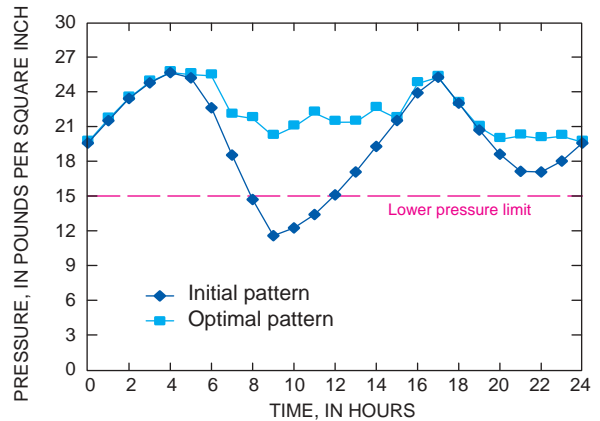


Figure E-7. Water pressure at Junction #3399 using two different patterns, May 1962. (See figure E-1 for location.)

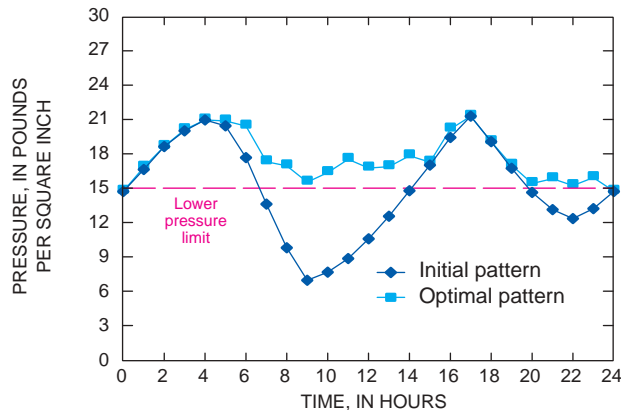


Figure E-8. Water pressure at Junction #4627 using two different patterns, May 1962. (See figure E-1 for location.)

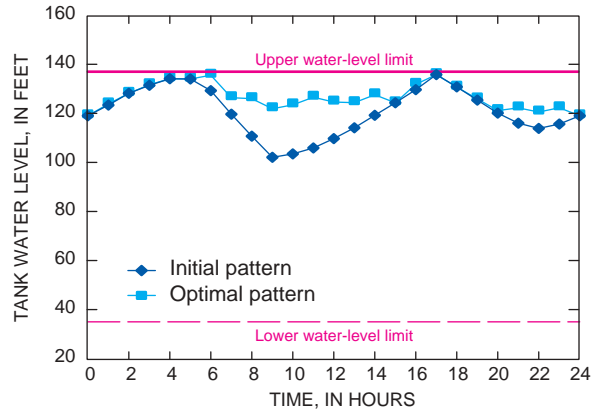


Figure E-9. Water level at Horton Street tank and standpipe using two different patterns, May 1962. (See figure E-1 for location.)

respectively. Through optimization calculation of four iterations, they are reduced to 2910.2, 0.98 and 2911.2 respectively. Figures E-11(a), E-11(b) and E-11(c) show the convergence of the objective function, junction violation cost and tank violation cost as a function of time step and iterations. Although the tank-violation cost has some fluctuation during the iteration, the objective function is monotonically reduced. In this case the water pressure violations in many junctions are very large when the initial patterns are used. The water-pressure violations are significantly improved although the violation still exists in some junctions using the optimal patterns, as shown in Figures E-12, and E-13. These violation junctions are at the suction sides of the pumps and not at demand nodes, and are expected to occur. The water pressure at junction #55024 is much higher than the upper pressure limit (110 psi) from 0800 to 1000 hours and at a maximum reaches 240 psi using the initial patterns. This situation should be prohibited in real operation of the water system distribution. In contrast, the water pressure at the junction stays within the lower pressure limit and upper pressure limits at all times using the optimal patterns (Figure E-12). Similar observation can be seen from Figure E-13 for the junction #55022. The water level in the tank is within its lower level and upper level limits at all times for both the initial and the optimal patterns, which is enforced by the EPANET2 water-distribution system model as shown in Figures E-14 and E-15. However, the initial water level in a tank at the beginning of the day may not be equal to the water level at the end of the day. The water levels in tank 33564-IHTA at the beginning and the end of the day are 196 ft and 195.92 ft, respectively, resulting in a difference of 0.08 ft using the initial patterns, and the difference reduces to 0.01 ft using the optimal pattern (Figure E-15). For tank 33566-NDTA, the difference reduces from 0.48 ft (203–202.52) in the initial patterns to 0.01 ft (203–202.99) in the optimal patterns (Figure E-14). In the reconstruction simulation, the optimal patterns selected will be continuously operated for 1200 hrs. Even though the difference in water levels in a tank between the beginning and end of the day is small, the long term simulations may result in a large difference in the water level in the tank between the beginning and the end of the simulation period. This may affect the flexibility of the patterns, and create negative pressures at junctions and cause water delivery problems. This is the reason why the penalty coefficient for this term is selected to be large compared with other terms in the objective function.

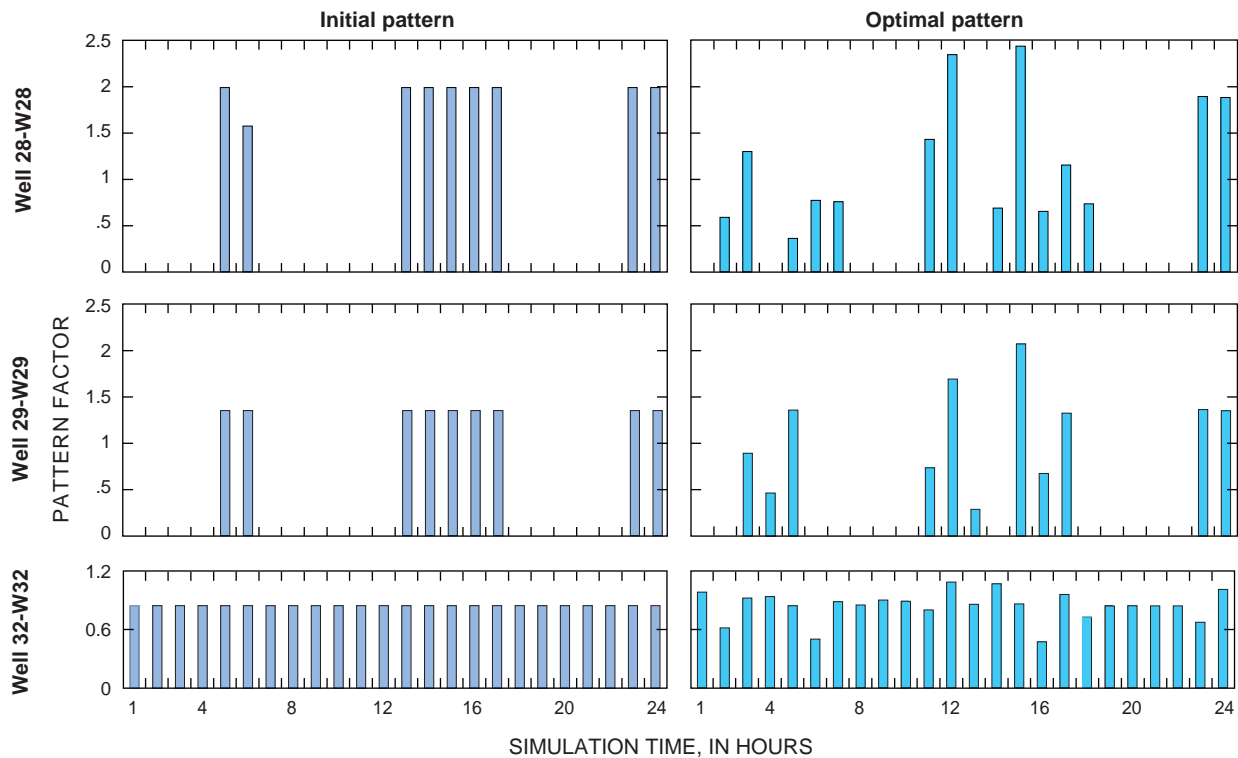


Figure E-10. Initial and optimal operation patterns of water-supply wells, October 1996. (See Figure E-2 for well locations.)

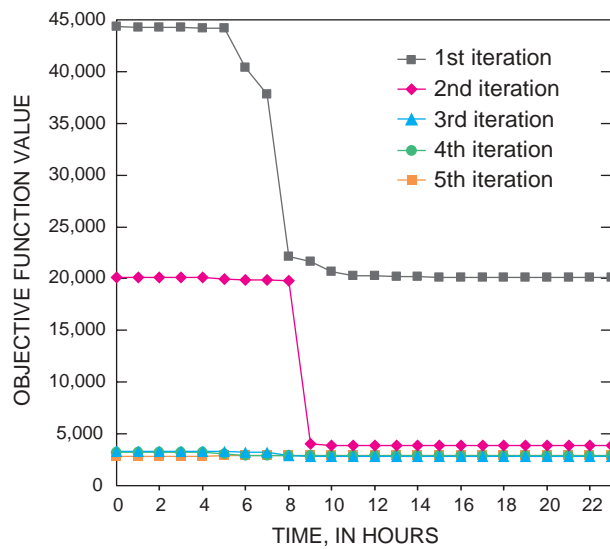


Figure E-11a. Convergence of objective function as function of iteration number and time step, October 1996.

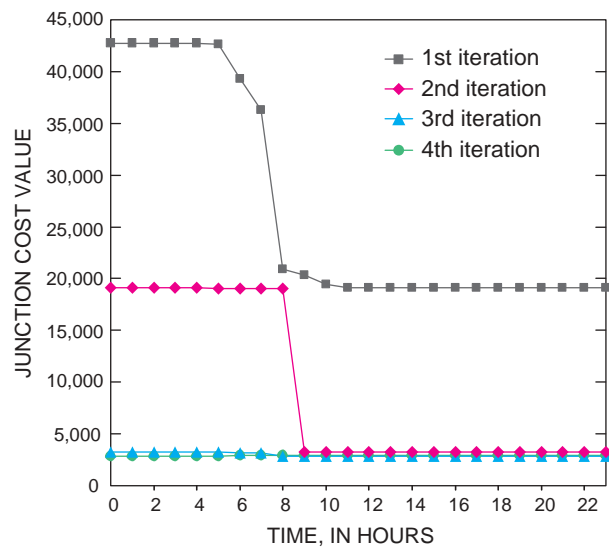


Figure E-11b. Convergence of junction cost as function of iteration number and time step, October 1996.

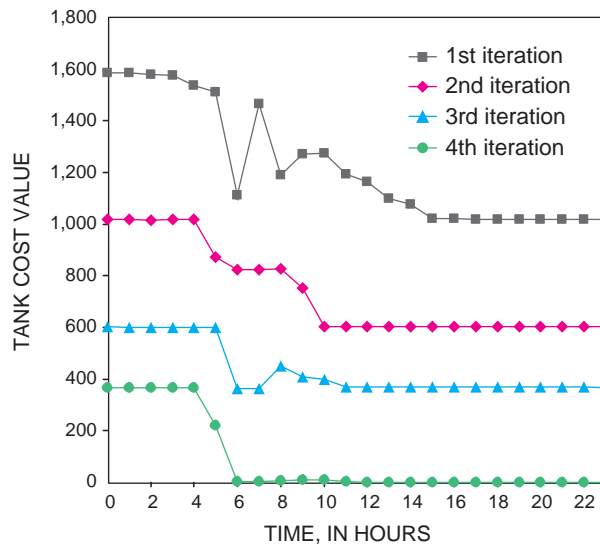


Figure E-11c. Convergence of tank cost as function of iteration number and time step, October 1996.

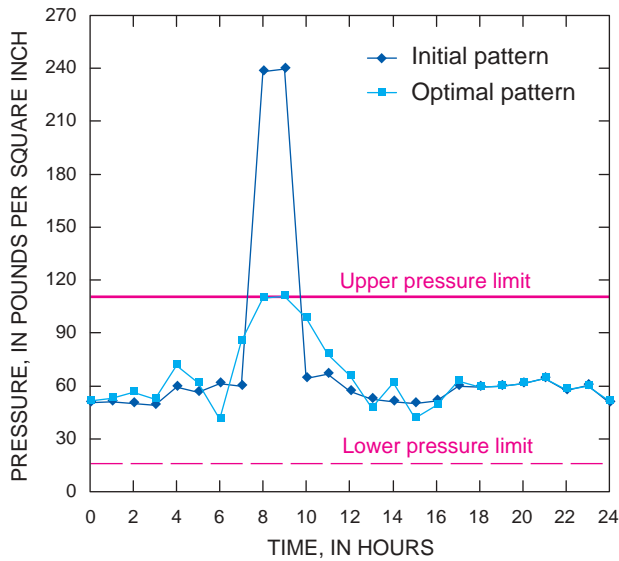


Figure E-12. Water pressure at Junction #55024 (Parkway Well 24) using two different patterns, October 1996. (See figure E-2 for location.)

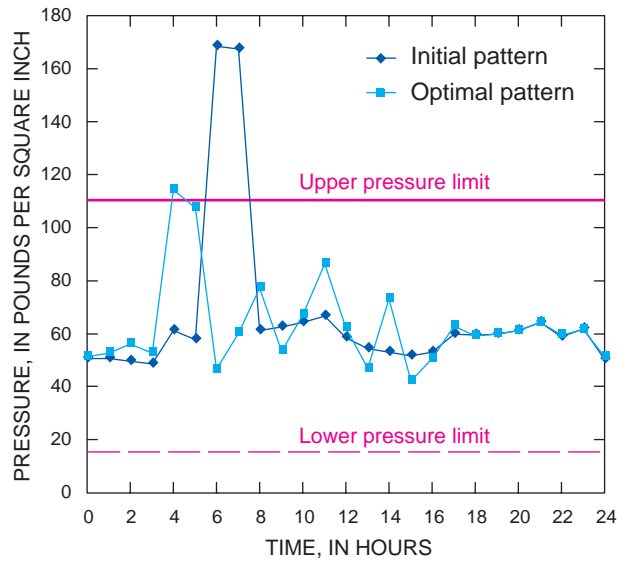


Figure E-13. Water pressure at Junction #55022 (Parkway Well 22) using two different patterns, October 1996. (See figure E-2 for location.)

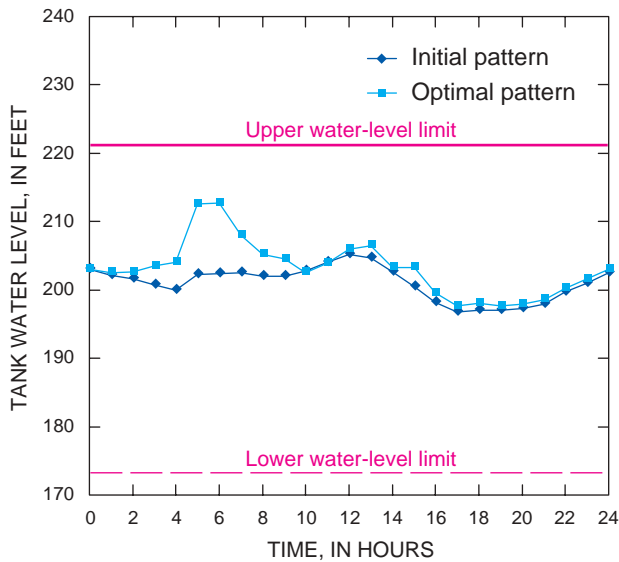


Figure E-14. Water pressure at North Dover elevated storage tank using two different patterns, October 1996. (See figure E-2 for location.)

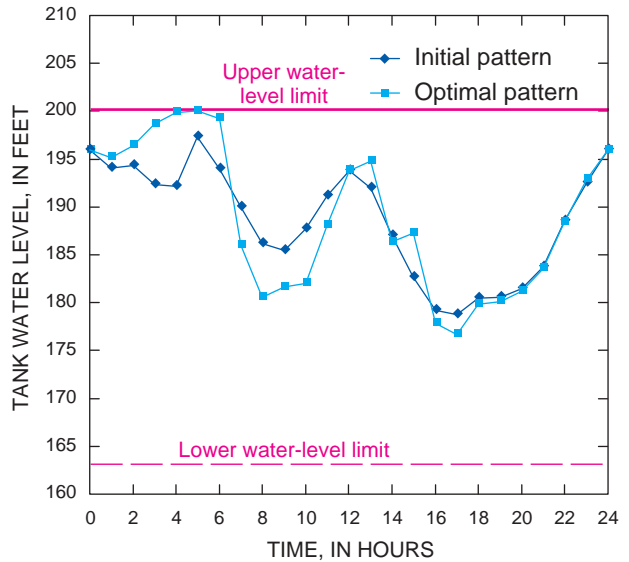


Figure E-15. Water pressure at Indian Hill elevated storage tank using two different patterns, October 1996. (See figure E-2 for location.)

Selection of New Patterns: As described in Section 3, the solution of the optimization model in Equation (E-15) may not be unique. The data set in October 1996 is chosen as an example to address the effect of non-uniqueness of optimal solutions on the operation of the water-distribution system. The model in Equation (E-19) and POGA algorithm is employed with the optimal patterns obtained from Equation (E-15) as its initial solution to select new patterns. The new patterns should have an original objective value calculated from Equation (E-14) close to the optimal patterns and they also must have a different pattern schedule when compared to the original optimal pattern. In this optimization, maximum iteration is selected as 8 considering the characteristics of the model in Equation (E-19). The resulting patterns are shown in Figure E-16 for three water-supply wells in this case. It might be noted that new patterns have a large difference when compared with the original optimal patterns. The original objective value of the new patterns reduces to 2900.45 from the value of 2911.09 for the original optimal pattern, resulting in a small difference 10.6 (0.37%). The convergence process of the new objective function in Equation (E-17) as a function of iterations and time steps is shown in Figure E-17(a). The new objective values are monotonically increasing as iterations and time steps. In the iterations, the original objective function is fluctuating within $\pm 1\%$ range as shown in Figure E-17(b). The water pressure in the representative junctions using the new patterns has similar values when compared to the original optimal patterns although there is a significant difference between the two patterns. Figures E-18 and E-19 show the water pressure at junction #55024 and #55022 using the two different patterns. It should also be noticed that the operation of tanks is also similar for original optimal solution and the new patterns as seen in Figures E-20 and E-21. Therefore, we may conclude that the effect of using different patterns on contaminant transport solution would be negligible.

Computational Cost: In POGA, the computational cost in the terms of computational time consists of two parts. First is the computational cost of calculation of the objective function and second is the calculation of GA operations, such as crossover, mutation and comparison of fitness values. The calculation of objective function accounts for more than 95% of total computational cost so that the effect of the second calculation cost can be neglected. The computational cost for calculation of the objective function depends on the scale of the water-distribution system. For example, the computational cost for each calculation of objective function needs about 2 seconds for the system in 1962, and about 10 seconds for the system in 1996 on an IBM/PC Pentium system. If we define t as the computation time of one calculation of the objective function, G as the maximum number of generations, P as the population size, N as maximum number of iterations, M as the number of time steps and T as the total time, then total time using POGA can be approximately estimated by

$$T = G \times P \times N \times M \times t \quad (\text{E-37})$$

In searching optimal patterns, $G = 2$, $P = 5$, $M = 23$, $N = 4$, $T = 920 t$. In the case of 1962 simulation, $t = 2$ s, then total computational time $T = 1840$ s = 30.7 m = 0.5 hr. In the case of 1996 simulation, $t = 10$ s, then total computational time $T = 9200$ s = 153.3m = 2.6 hrs.

From Equation (E-36), the total computational time seems to be independent of the number of decision variables. In fact, the effect of the number of decision variables on the total computational cost is implicitly reflected in population size and generations. In general, the more the decision variables there are, the larger the number of population size and generations should be set. For example, in the case of year 1996 simulation, there are 20 water-supply wells in the system. The resulting number of decision variables is ($20 \times 24 = 480$). If the GA is directly used to solve this problem, the population size should be set at least as 50 and the number of generations should be set to at least 20. The total computational time T would be at the least ($50 \times 20 \times t = 10,000$ sec = 2.8 hrs). In POGA, the optimization problem is divided into 23 sub-problems in each iteration. The number of variables in the sub-problem reduces to the number of water-supply wells. In the case of year 1996 simulation, it is 20. Therefore, we can choose a small population size and generations to reduce the computational cost. Furthermore, the equality constraints are not so easily handled as they are treated in POGA if GA is directly applied to this problem.

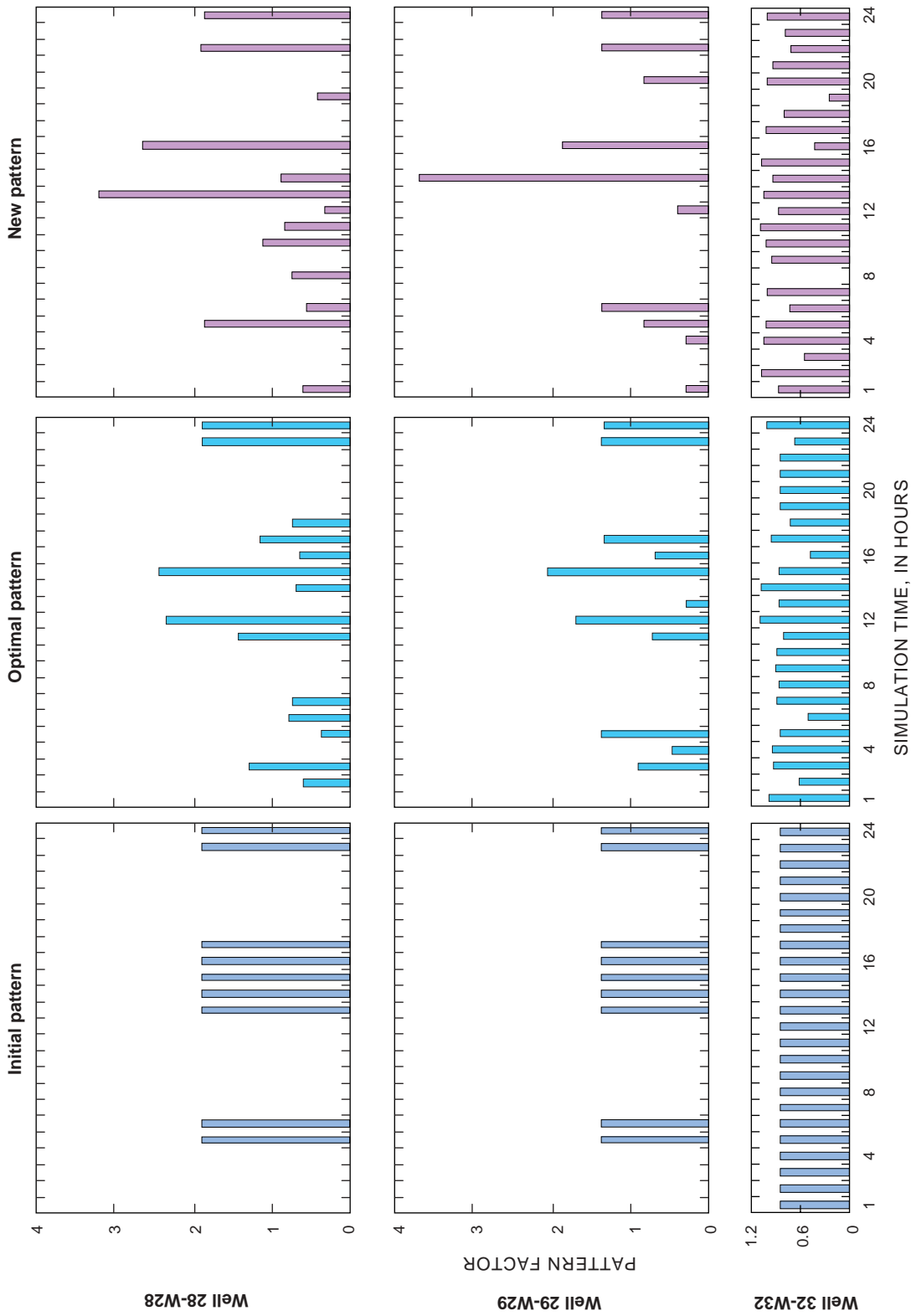


Figure E-16. Initial, optimal, and new operation patterns of water-supply wells, October 1996. (See Figure E-2 for well locations.)

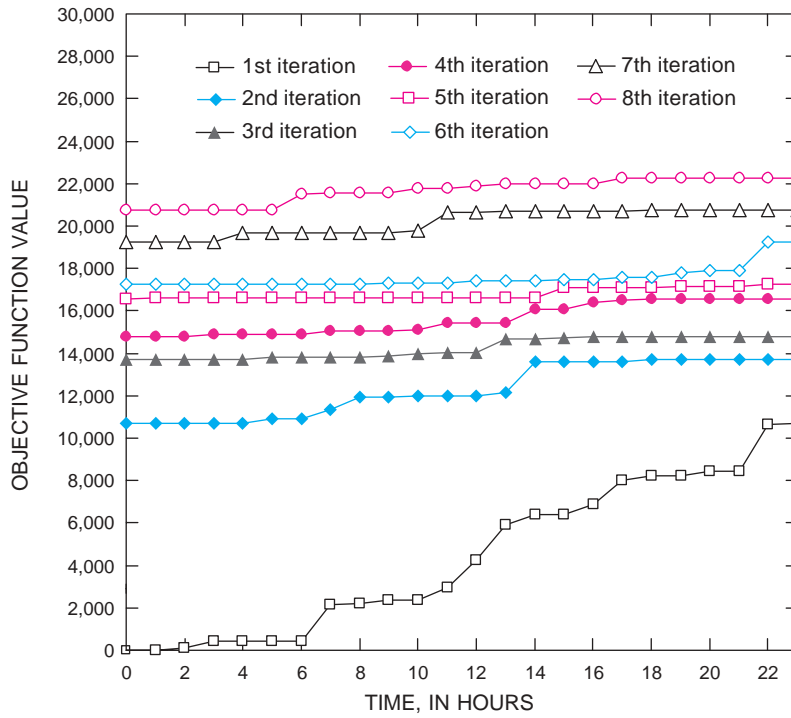


Figure E-17a. Convergence of new objective function as function of iteration number and time step, October 1996.

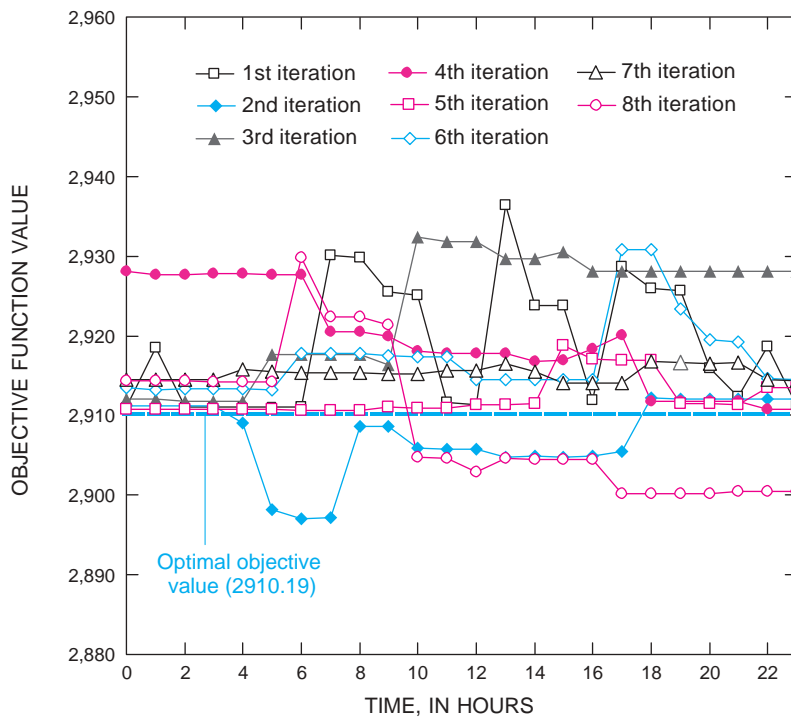


Figure E-17b. Variation of original objective function value in iteration process, October 1996.

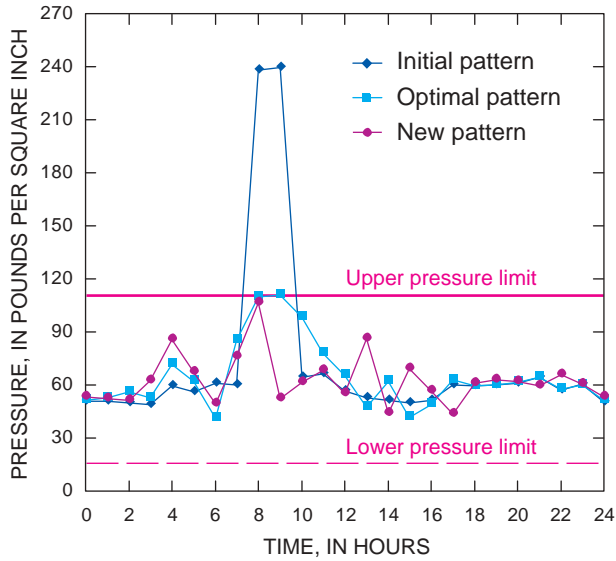


Figure E-18. Water pressure at Junction #55024 (Parkway Well 24) using three different patterns, October 1996. (See figure E-2 for location.)

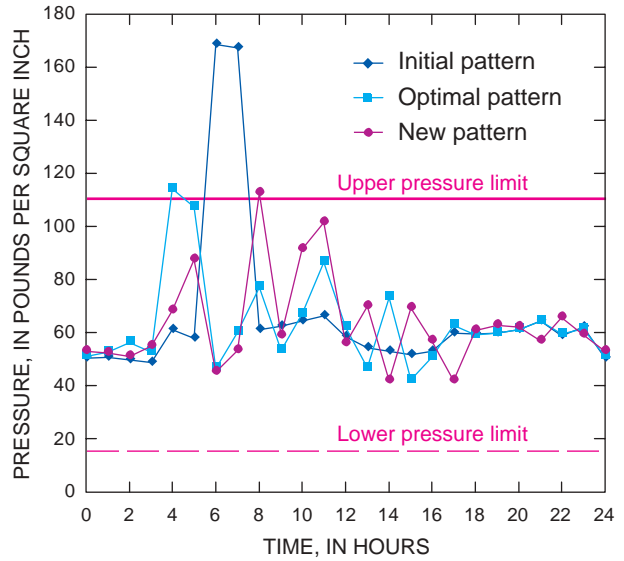


Figure E-19. Water pressure at Junction #55022 (Parkway Well 22) using three different patterns, October 1996. (See figure E-2 for location.)

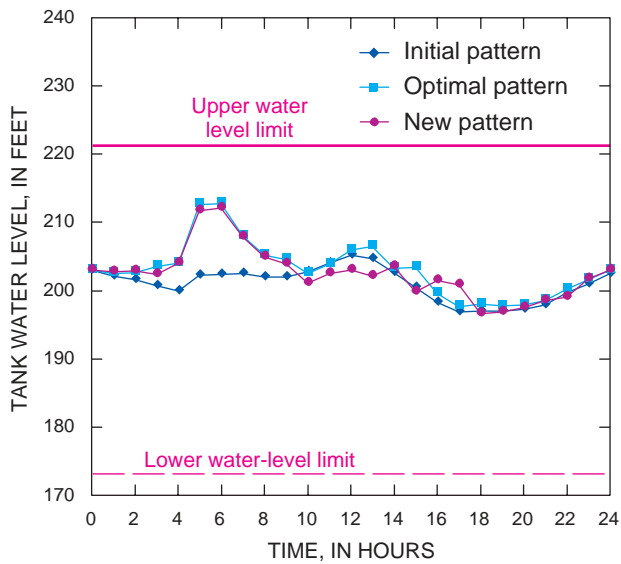


Figure E-20. Water pressure at North Dover elevated storage tank using three different patterns, October 1996. (See figure E-2 for location.)

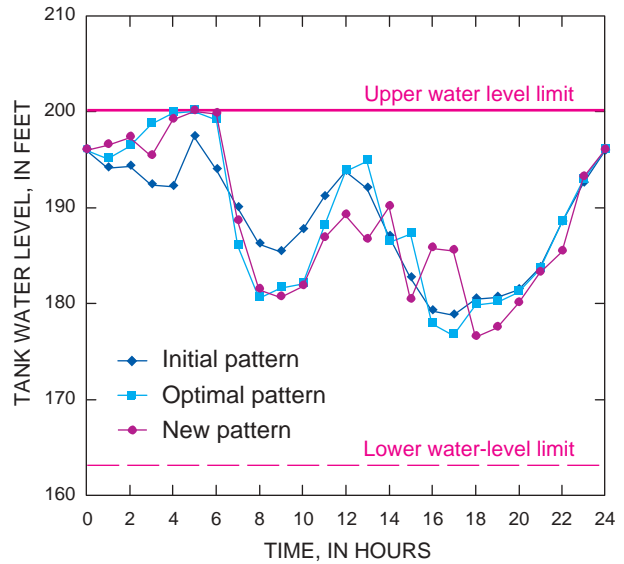


Figure E-21. Water pressure at Indian Hill elevated storage tank using three different patterns, October 1996. (See figure E-2 for location.)

7. Sensitivity Analysis

The control parameters in the model have a large effect on the optimal solution. Among these parameters, the minimum pressure limitation at the junctions in the water-distribution system and the difference in water levels in the tanks between the beginning and the end of the day play an important role in the objective function value. Therefore, a sensitivity analysis for these two parameters is necessary. The effect of minimum pressure on the operation pattern of water-supply wells can be analyzed by setting different values. In this study, three minimum pressure values, 15, 20, 30 psi, are used to calculate the optimal patterns. The effect of difference between storage starting and ending water levels in a 24-hour period on the optimal patterns can be reflected by adjusting Equation (E-6). Equation (E-6) is a tightening constraint and can be relaxed by

$$|L_j(24) - L_j(0)| \leq \Delta L \quad j = 1, 2, \dots, N \quad (\text{E-38})$$

where ΔL is an allowable error, and can be set to zero or 3 ft as is done in the sensitivity analysis phase. The dataset for October 1996 is used to present the sensitivity analysis results here, although numerous cases were analyzed. The optimal patterns of the water-supply wells are shown in Figures E-22(a) and E-22(b) by setting different minimum pressure values and the different water-level limits in tanks. It may be noted that these two parameters have a significant effect on the schedule of optimal patterns obtained. The convergence of the objective function in the first iteration using different patterns with different parameters is shown in Figure E-23, where P_{min} indicates minimum water pressure, ΔL indicates allowable water-level difference in the storage tanks. From this Figure E-23, we can observe the following: when the minimum water-pressure limit increases, more junctions violate the minimum water-pressure constraints; thus, the objective values increase. The objective values decrease when the constraint of water-level difference in tanks is relaxed, and this effect is more obvious for higher minimum water-pressure requirement. The different patterns result in the fluctuation of the water pressure for some junctions. The water-pressure variation as function of time is the same for junction #55024, however, it is has changed for the junction #55022 as shown in Figures E-24 and E-25. The change in minimum water-pressure limit has no obvious effect on the operation of tanks, although the optimal patterns obtained are different as shown in Figures E-26 and E-27. Comparing with the initial patterns, all optimal patterns tend to keep the tanks operating at a higher water level to maintain the higher “low pressure” limits. Under the same minimum water-pressure limit, different allowable water-level differences between the starting and the ending of a 24-hour period (day) in storage tanks leads to the ending water level not returning to that of the starting of the day as shown in Figures E-28 and E-29. The starting and ending water levels in storage tanks are given in Table E-2.

Table E-2. Difference between storage tank starting and ending water levels in a 24-hour period

Storage tank	Minimum water pressure (pounds per square inch)	Allowable water-level difference (feet)	Water level		Difference (feet)
			Starting	Ending	
33566-NDTA	20	0	203.00	202.97	0.03
		3	203.00	201.44	1.56
	30	0	203.00	202.96	0.04
		3	203.00	200.82	2.12
33564-IHTA	20	0	196.00	195.98	0.02
		3	196.00	196.93	-0.93
	30	0	196.00	195.98	0.02
		3	196.00	198.79	-2.79

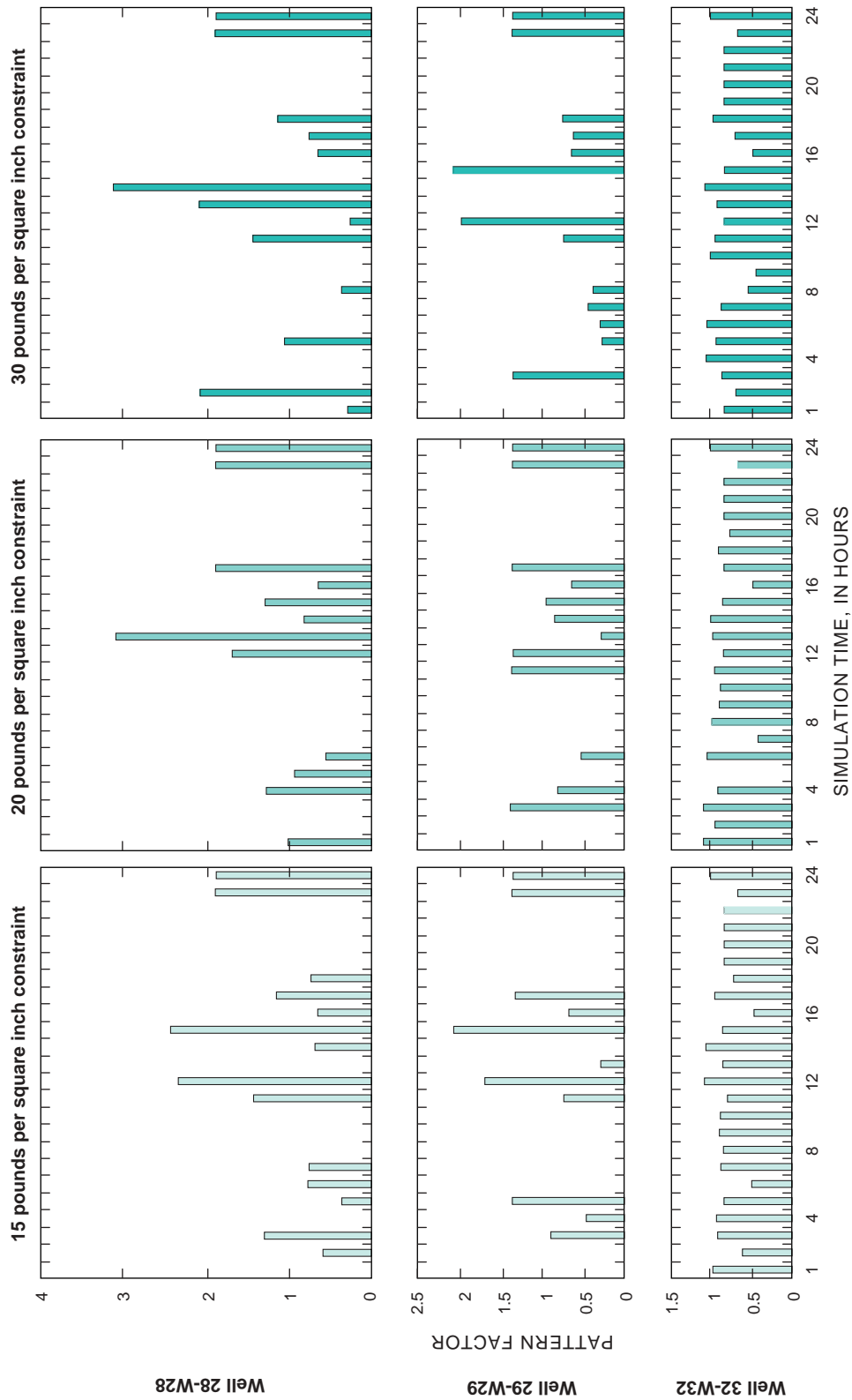


Figure E-22a. Optimal patterns of water-supply wells under different minimum pressure constraints, October 1996.
 (See Figure E-2 for well locations.)

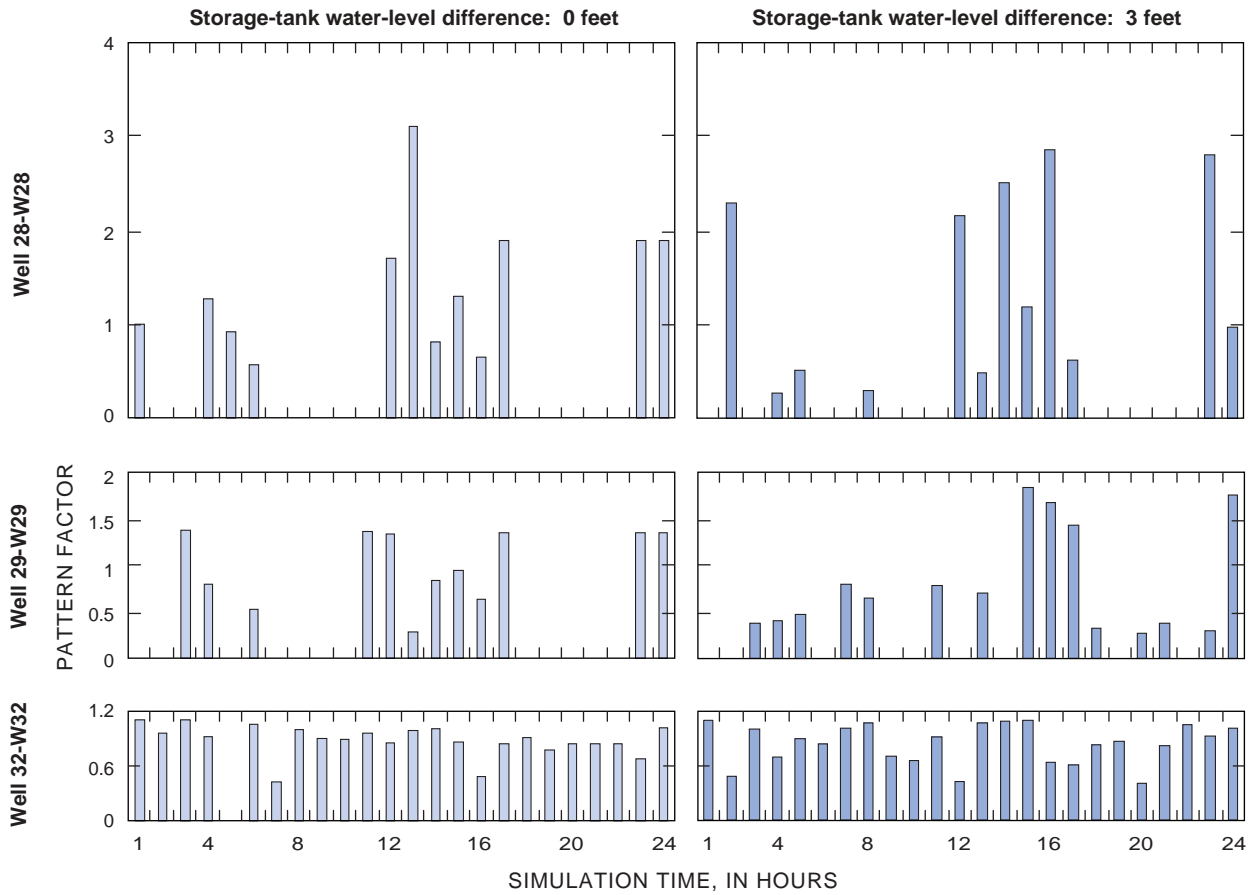


Figure E-22b. Optimal patterns of water-supply wells under minimum pressure of 20 pounds per square inch and two different tank-level constraints, October 1996. (See Figure E-2 for well locations.)

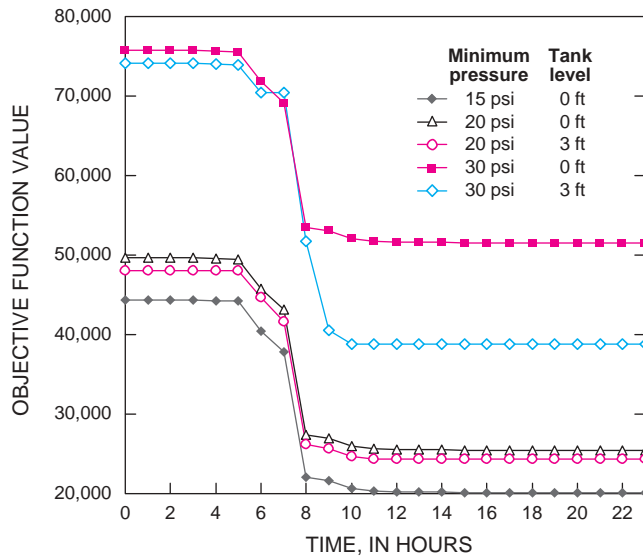


Figure E-23. Convergence of objective function in first iteration using different patterns with different minimum pressure and tank-level constraints, October 1996.

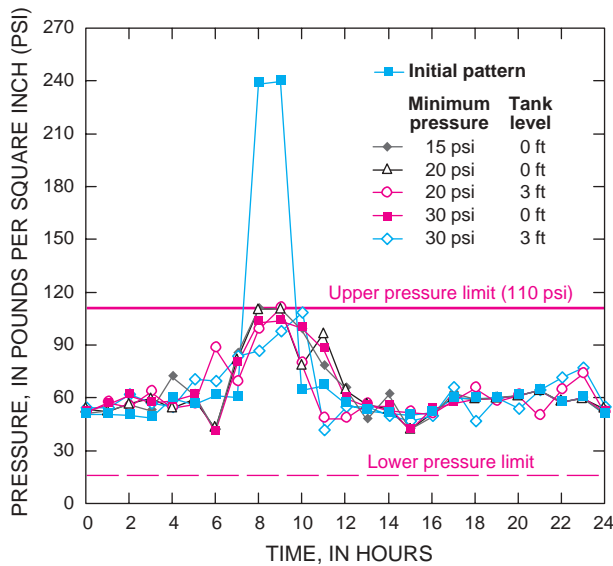


Figure E-24. Water pressure at Junction #55024 (Parkway Well 24) using different patterns with minimum pressure and tank-level constraints, October 1996. (See figure E-2 for location.)

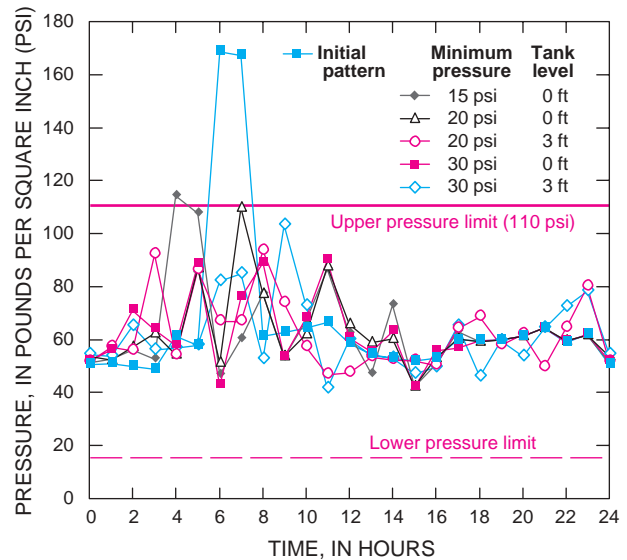


Figure E-25. Water pressure at Junction #55022 (Parkway Well 22) using different patterns with minimum pressure and tank-level constraints, October 1996. (See figure E-2 for location.)

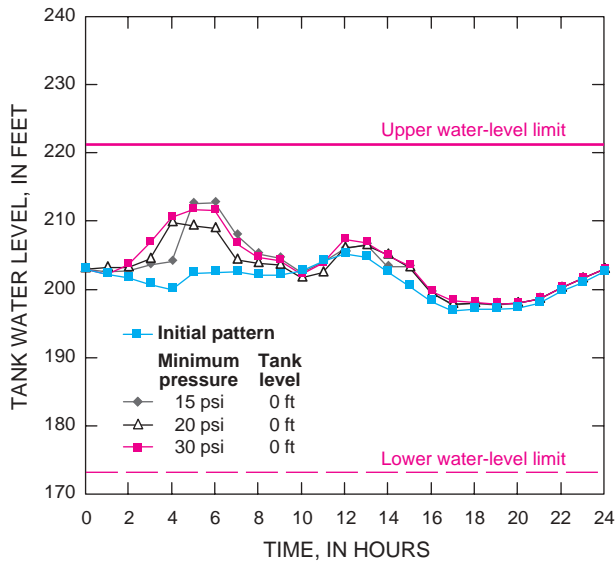


Figure E-26. Water level at North Dover elevated storage tank using different patterns with different minimum pressure constraints, October 1996. (See figure E-2 for location.)

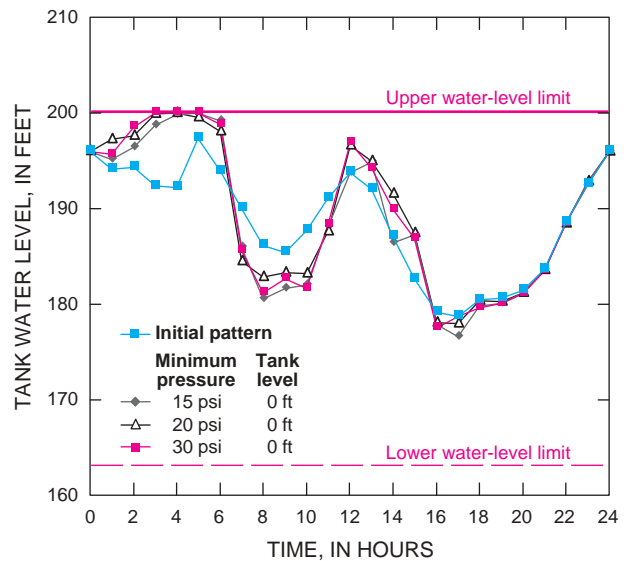


Figure E-27. Water level at Indian Hill elevated storage tank using different patterns with different minimum pressure constraints, October 1996. (See figure E-2 for location.)

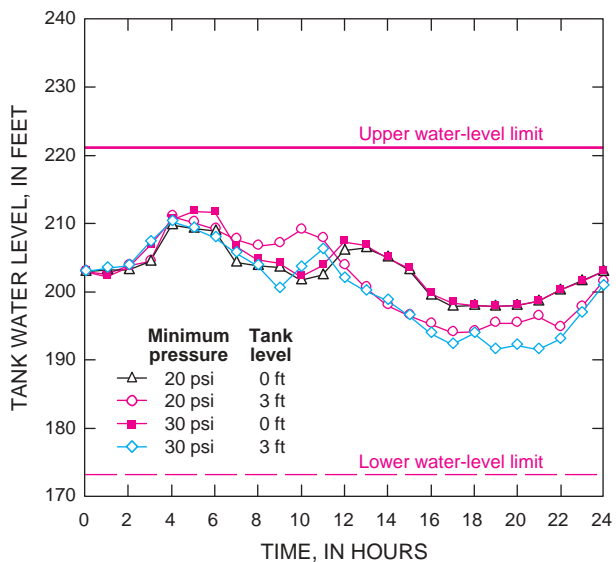


Figure E-28. Water level at North Dover elevated storage tank using different patterns with different minimum pressure and tank-level constraints, October 1996. (See figure E-2 for location.)

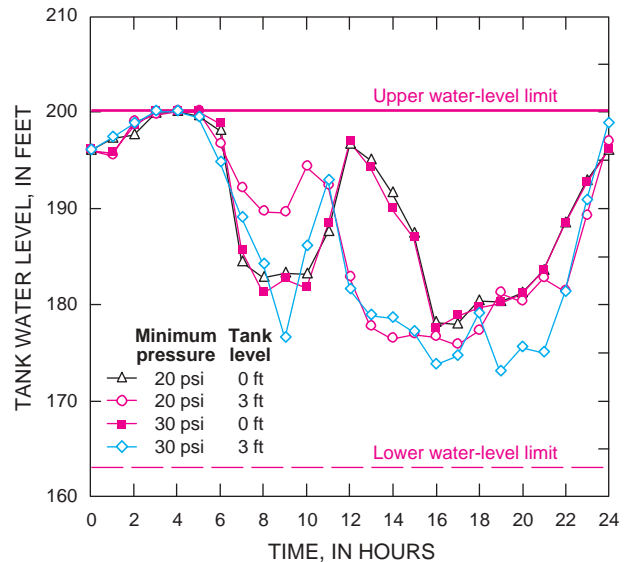


Figure E-29. Water pressure at Indian Hill elevated storage tank using different patterns with different minimum pressure and tank-level constraints, October 1996. (See figure E-2 for location.)

when the allowable water level difference in tanks is set as 3 feet, the difference in actual operation is 1.56 feet for $P_{min} = 20$ psi and 2.12 feet for $P_{min} = 30$ psi for tank 33566-NDTA; is -0.93 feet for $P_{min} = 20$ psi and -2.79 feet for $P_{min} = 30$ psi for tank 33564-IHTA. Although increasing allowable water-level difference in tanks may improve water pressure distribution at junctions and provide flexibility to the selection of patterns, the long-term continuous operation using the same patterns during a 1200 hr simulation period results in a large difference at the starting and end of the simulation period in storage tanks which would violate the tank-level constraints.

8. Conclusions

In this study, the reconstruction of hydraulic management strategies in the water-distribution system is formulated as an optimization problem. In the model, the pumping rates of water-supply wells are selected as decision variables, and the penalty functions are assigned to the violation of the management objectives, which are incorporated into the objective function. The equality and bounded limits of decision variables for each water-supply well reflect the historical operation of the system. Considering the non-uniqueness of the optimal solution of this model, a new optimal model is developed based on the optimal solution of the first model. The objective function in the new model reflects the expectation that the new patterns obtained from the second model have an objective value close to the optimal solution obtained from the first model, and also there should be a large difference between the optimal patterns. Both models have the same constraints so that the results obtained from them can be used to analyze the effect of different optimal patterns on hydraulic and contaminant transport characteristics of the system. Both models are multiple-stage nonlinear dynamic problems with equality constraints. For the solution of both models, a new algorithm, identified as Progressive Optimality Genetic Algorithm (POGA), is developed and coded. This algorithm incorporates progressive optimality algorithm and genetic algorithm. POGA is integrated with EPANET2 to construct a high integrity system, which provides a visual tool for reconstructing the hydraulic management strategies in a water-distribution system.

The integrated system has been applied to the monthly data sets of the water-distribution system serving the Dover Township area, N.J. for the study period between the years 1962 and 1996. The effect of several parameters on hydraulic management strategies in the water-distribution system has also been analyzed. The computational results show that the objective function value has been significantly improved for all of the data sets, which implies that the optimized hydraulic management strategies are better in satisfying the hydraulic management objectives than the initial hydraulic management strategy. Based on the results and analysis, we make the following conclusions:

- The models described in Equations (E-15) and (E-19) reflect the historical situation associated with the water-distribution system and can be used to reconstruct the hydraulic management strategies. The POGA is an effective and efficient algorithm for solution of both models. The idea of POGA may be applied to other multiple-stage dynamic problems.
- Optimized management of the water-distribution system may improve the hydraulic condition of the system operation, reduce the violation of water-pressure requirement, and heighten the efficiency of the system.
- The reconstruction problem may not have a unique solution. However, the different patterns satisfying optimal conditions seem to have little effect on water-pressure distribution at junctions, the operation of tanks in the water-distribution system and consequently have little effect on final contaminant distribution pattern in the system.
- The constraints on minimum water pressure at junctions and water level of the tanks at the starting and ending of the simulation period directly affect the optimal patterns. The relaxed allowable water-level difference in tanks may reduce the violation of water pressure at junctions, but may result in a large accumulated water-level difference between the starting and ending of the simulation period.

The results of analysis presented in this report provide the data that can feed into the epidemiologic study of childhood leukemia and central nervous system cancers that occurred in the period 1979 through 1996 in Dover Township area, N.J. It is expected that these results will be used in identifying the contaminant transport characteristics of the water-distribution system so that the epidemiologic studies can be completed.

9 References

- Aral, M.M., Maslia, M., Ullrich, G. and Reyes, J. J. "Estimating Exposure to VOCs from Municipal Water-Supply Systems: Use of a Better Computational Model," *Archives of Environmental Health*, Vol. 51, No. 4, pp. 300–309, 1996.
- Aral, M.M., Guan, J., Maslia, M.L., Liao, B., Sautner, J.B., Williams, R.C. and Reyes, J.J. Reconstruction of Hydraulic Management of a water-distribution system Using Genetic Algorithms, Proceedings: World Water and Environmental Resources Congress 2001, Environmental and Water Resources Institute of the American Society of Civil Engineers, May 20–24, 2001, Orlando, FL, 2001a.
- Aral, M.M., Guan, J., Maslia, M.L., Liao, B., and Sautner, J.B., "Reconstruction Hydraulic Management of a Water-Distribution System Using Optimization," Multimedia Environmental Simulations Laboratory, School of Civil and Environmental Engineering, Georgia Institute of Technology, Report No. MESL-01-01, 58 p., October 2001b.
- Bellman, R.E. Dynamic Programming, Princeton, NJ: Princeton University Press, p. 342, 1957.
- Goldberg, D.E. Genetic Algorithms in Search, Optimization, and Machine Learning: Reading, MA: Addison-Wesley, p. 412, 1989.
- Guan, J. and Aral, M.M., "Progressive Genetic Algorithm for Solution Optimization Problems with Nonlinear Equality and Inequality Constraints," *Applied Mathematical Modelling*, Vol. 23, pp. 329–343, 1999a.
- Guan, J. and Aral, M.M. "Optimal Remediation with Well Locations and Pumping Rates Selected as Continuous Decision Variables," *Journal of Hydrology*, Vol. 221, pp. 20–42, 1999b.
- Holland, J.H., *Adaptation in Natural and Artificial Systems*, Ann Arbor, MI: University of Michigan Press, p. 183, 1975.
- Maslia, M.L., Sautner, J.B., Aral, M.M., Reyes, J.J., Abraham, J.E., and Williams, R.C. "Using Water-Distribution System Modeling to Assist Epidemiologic Investigation," *ASCE Journal of Water Resources Planning and Management*, Vol. 126, No. 4, pp. 180–198, June/August 2000a.
- Maslia, M.L., Sautner, J.B., and Aral, M.M. Analysis of the 1998 Water-Distribution System Serving the Dover Township Area, New Jersey: Field-Data Collection Activities and Water-Distribution System Modeling, Atlanta, Georgia: Agency for Toxic Substances and Disease Registry, June 2000b.
- New Jersey Department of Health and Senior Services (NJDHSS), Drinking Water Quality Analyses, March 1996 to June 1999, United Water Toms River, Dover Township, Ocean County, New Jersey, Trenton, NJ: New Jersey Department of Health and Senior Services, November 1999.
- Rossman, L.A., *EPANET 2 Users Manual*, Cincinnati, OH: National Risk Management Research Laboratory, U.S. Environmental Protection Agency, p. 192, June 2000.