

Optimal Design of Urban Drainage Systems using Genetic Algorithms

Paul F. Boulos, Trent Schade, Chris Baxter and Misgana Muleta

Control of sewer overflows is vital to reducing risks to public health and protecting the environment from water pollution. Sewer overflows are a leading cause of water pollution in the nation's lakes, streams and inland bays. The untreated sewage from these overflows contains microbial pathogens, suspended solids, toxics, nutrients, trash, and other pollutants that deplete dissolved oxygen and can contaminate our waters, causing serious water quality problems and threatening drinking water supplies, fish and shellfish. This sewage can also back up into basements, causing property damage and creating threats to public health for those who come in contact with the untreated sewage. There are about 19,500 sewer systems nationwide designed to handle an average daily flow of roughly 50 billion gallons of raw sewage (Nicklow et al., 2004, 2006).

Sanitary sewer overflows (SSOs) or combined sewer overflows (CSOs) may release partially treated or untreated sewage to surface waters. High wet weather flows from rainfall-derived inflow and infiltration (RDII) can exceed system capacity, resulting in an SSO. The measured volume for this type of SSO is typically much greater than other causes of SSO. SSOs are most frequently caused by grease and debris blockage. Other causes for

SSOs include sediments buildup, pipe breaks, leaking manholes, offset joints, equipment failures, undersized sewer pipes, power outages, and other reasons. When an SSO occurs, sewage flows into streets, playgrounds and streams.

CSOs occur in older combined sewer systems that were designed to carry both sanitary sewage and storm water runoff to a wastewater treatment plant (WWTP). Under dry conditions, the WWTP treats the sewage and then discharges it to a water body. During periods of heavy rainfall or snowmelt, however, the wet weather volume in the combined sewer system exceeds the available hydraulic capacity of the sewer system or treatment plant. This leads to the discharge of excess wastewater directly to nearby streams, rivers, or other water bodies. Combined sewer systems in the United States serve roughly 746 communities containing about 40 million people. Although there are combined sewers in 32 states and the District of Columbia, they are mostly located in the Northeast and Great Lakes regions, and the Pacific Northwest (U.S. EPA, 2004).

With the growing expectations by the public for quality services, the U.S. Environmental Protection Agency (EPA) under the authority of the Clean Water Act adopted by Congress has implemented pollution control programs and set wastewater standards for the industry. In order to meet these requirements, comprehensive modeling and analysis of these sewer systems becomes necessary for developing sound cost-effective solutions for enhancing system integrity and performance to reliably convey sewer flows without surcharging, overflows, flooding, and backups.

Today, many wastewater utilities and engineering consulting companies utilize drainage network simulation models to plan improvements and design better systems. Technology to achieve these improvements includes: the addition of new sewer pipes or treatment capacity as well as increasing conduit capacity (bigger interceptors), more storage volume, and pumping capacity. Current practice involves a tedious trial-and-error evaluation procedure that seldom leads to the most effective or most economical solutions for upgrading collection systems. This requires using the drainage network simulation model to evaluate the hydraulic performance of the existing system with different design alternatives (modifications) under a range of loading conditions. The design that meets the target hydraulic criteria for the lowest cost is then selected from among the alternative designs. The complexity of this manual trial-and-error procedure increases exponentially with the number of proposed system modifications and corresponding operating conditions. It is important to point out, however,

that even if the operating specifications are met, the trial-and-error procedure has no inherent feature that assures that the solutions reached are cost optimum or even cost effective. Therefore, good engineering procedure dictates that the iterations continue until a number of promising alternatives have been evaluated. The cost of each feasible alternative is then computed to arrive at a recommended solution. Time restraint or limited understanding of sewer network hydraulics generally prevents engineers from obtaining hydraulically reliable recommendations. Given the vast number of possible combinations of system enhancements, it is unlikely that even the most experienced modeler will be able to determine the least-cost improvement alternative. This process is also not able to ensure that the final design could perform adequately under all possible loading and operating conditions. The result of using the traditional trial-and-error evaluation approach is often inefficient performance at greater cost.

One way to circumvent the previous problems is to employ optimization theory. This chapter presents a rigorous optimal design methodology, which eliminates the need of the traditional manual design technique. The problem of choosing least-cost improvements for urban drainage systems is cast as a nonlinear optimization problem and solved using optimization theory. The optimization problem consists of determining the optimal design improvement solutions that produce the minimum overall cost while satisfying target system performance requirements. The decision variables can include any selected combination of pipe slope and upsizing, storage, pumping and new piping. Performance criteria include maximum allowable depth to diameter ratio, minimum and maximum pipe velocities, maximum head loss for force mains, and minimum and maximum pipe slopes. This gives practicing engineers complete control over the solution process.

The proposed approach links a comprehensive drainage network simulator with an efficient stochastic optimization model and iterates between the simulator and the optimization model until the best solution is found. It is structured in an object-oriented framework allowing very large and complex sewer collection systems to be solved in expeditious times. The optimization procedure employed is based on the fast messy genetic algorithm (fmGA), which delivers reliable solutions in sub-quadratic time. The urban drainage network simulation is performed using an extended version of the EPA storm water management model, SWMM 5 (Rossman, 2005). The optimization model generates improved sets of decision variables that seek to minimize rehabilitation costs. These decision variables are then passed directly to the appropriate objects in the simulation program for use in quantifying system hydraulics. The hydraulic simulation results are

passed back to the optimization model for checking for any constraint violations. The iterations continue until the least cost solution is found. This allows rapid solutions to be obtained with a minimal computational overhead. The value and performance of the method are illustrated by application to an example stormwater collection system. The method should prove useful to any wastewater utility in planning and designing reliable systems and optimizing their capital improvement programs.

11.1 Optimization Model Formulation

The general statement of the optimal urban drainage system design problem is that of identifying the optimal combination of conduit sizes, storage volumes, and pumping capacities needed for selected facilities that meet target system performance requirements at minimum cost. It casts the design problem as an implicit nonlinear optimization problem subject to both implicit and explicit constraints. The sewer collection system optimization problem is thus governed by an objective function and its associated set of constraints.

11.1.1 Objective Function

The objective (fitness) function of the sewer system optimization problem is to minimize the overall design cost under a specified set of loading and operating conditions. The objective cost function can be mathematically expressed as:

$$\text{Minimize} \quad \sum_{c=1}^C C_c L_c + \sum_{s=1}^S C_s V_s + \sum_{p=1}^P C_p + \text{Penalty} \quad (11.1)$$

where c , C_c and L_c used in the left-hand side term of the equation refer to conduits and represent conduit counter, replacement cost which is a function of conduit size, and length of the conduit, respectively. Likewise the s , C_s , and V_s used in the middle term of the equation refer to storage node counter, cost per unit volume of storage and volume of the storage unit, respectively. The p and C_p used in the last term of the equation refer to pump counter and cost associated with the pump, respectively.

The decision variables, which consist of conduit, storage and pump sizes, are automatically calculated to minimize the objective cost function while satisfying three different kinds of constraints: (i) A set of implicit system constraints; (ii) a set of explicit bound constraints; and (iii) a set of explicit variable constraints.

11.1.2 Implicit System Constraints

The implicit constraints on the sewer collection system are equality constraints defining the hydraulic equilibrium state of the system. They correspond to the conservation of mass and momentum equations (Saint Venant equations) that govern the unsteady flow of water through a drainage network of channels and pipes. These equations are solved implicitly using SWMM5, by converting them into an explicit set of finite difference formulas that compute the flow in each conduit and head at each node for time $t + \Delta t$ as functions of known values at time t . Each function call to SWMM5 with a set of decision variables returns the simulated hydraulic equilibrium solution for velocities and hydraulic heads.

11.1.3 Explicit Bound Constraints

The explicit bound constraints on the optimization problem represent system performance criteria and may include constraints on sewer flow velocity (V), depth to diameter ratio (d/D), conduit slope (S), and head loss (HG) in force mains for a given sewer collection system loading and operating conditions.

Conduit Constraints

Conduit constraints describe limits on the hydraulic behavior of the conduit groups. These limits include the ratio of maximum depth to full section depth, minimum and maximum velocities, maximum headlosses for force mains, and maximum and minimum invert elevations. Each of these limits is based on design criteria consistent with local regulations or engineering practice. The ratio of maximum depth to full section depth ensures that the conduit maintains sufficient hydraulic capacity to reduce the risk of sewage discharge through lateral connection. The minimum velocity constraint ensures that a sufficient velocity is maintained to keep solids in suspension and encourage passage of sewer floatables. The maximum velocity is critical for pressurized mains to ensure that momentum forces are not objectionably high on the system especially at joints. For gravity mains, the maximum

velocity is usually not a significant concern; however, under some conditions, the transition from supercritical to subcritical flows can create a large hydraulic jump and may be of concern to the engineer. The maximum head loss constraint ensures that pressurized mains have sufficient lift to carry the water to the upstream location.

Occasionally local codes specify geometric requirements for gravity sewer systems, which typically include slope and embedment depth. For constructability purposes, a minimum slope (and inter alia velocity) must be maintained. The embedment depth requirement accommodates areas where freezing weather or surcharge surface loading (beneath railways or streets) is of concern. The invert elevation at the downstream end of each conduit (expressed as an offset from the node invert elevation) constrains the design to meet the slope and embedment depth requirements.

The conduit constraints can thus be expressed as:

$$V_{min_k} \leq V_k \leq V_{max_k} \quad \forall k, \forall t \quad (11.2)$$

$$(d / D_k) \leq (d / D_{max_k}) \quad \forall k, \forall t \quad (11.3)$$

$$S_{min_k} \leq S_k \leq S_{max_k} \quad \forall k \quad (11.4)$$

$$HG_k \leq HG_{max_k} \quad \forall k, \forall t \quad (11.5)$$

where:

- V_k = the flow velocity of conduit k at time t ,
- V_{max_k} = the maximum allowable flow velocity for conduit k ,
- V_{min_k} = the minimum allowable flow velocity for conduit k ,
- d/D_k = the depth to diameter ratio of conduit k at time t ,
- d/D_{max_k} = the maximum allowable depth to diameter ratio for conduit k ,
- S_{max_k} = the maximum allowable slope for conduit k ,
- S_{min_k} = the minimum allowable slope for conduit k ,
- HG_k = the head loss for force main k at time t , and
- HG_{max_k} = the maximum allowable head loss for force main k .

When a design solution does not satisfy an implicit bound constraint, a penalty method is used to handle the constraint violation.

Penalty Cost (PC)

A penalty cost is added to the objective function to penalize an infeasible solution (degrade its fitness) and force the search procedure towards the region of feasible solutions. The penalty cost function is defined as the divergence (distance) of the computed solution from the feasible region or:

$$PC = \sum_{i=1}^N v_i |C_i(x) - C_i| \quad (11.6)$$

where:

- N = the number of constraints;
- v_i = a weighting factor associated to constraint C_i ;
- $C_i(x)$ = the value of the i^{th} constraint; and
- C_i = the constraint limit.

It is expected that different values of the penalty costs will result in different solutions and also affect the efficacy of the optimization calculation. Therefore, a number of trial optimization runs with different penalty costs may be required for better exploring the solution space and narrowing in on the lowest cost solutions. If a feasible solution (a solution that satisfies all the constraints) is not found, the optimization model will return the lowest-cost solution with the minimum constraint violation.

11.1.4 Explicit Variable Constraints

The explicit variable constraints are used to set minimum (lower) and maximum (upper) limits on the sewer sizes and to specify the discrete (commercially available) diameter values for the new conduits. Conduits should be lumped together in separate logical design groups based on the known physical characteristics of the associated conduits such as size and location. As such all conduits within a group will possess an identical size.

For each conduit group, the conduit size is bound by an explicit inequality constraint as:

$$Dmin_n \leq D_n \leq Dmax_n \quad \forall n, \forall D_n \in D^0 = \{d_k^0, k = 1, \dots, K\} \quad (11.7)$$

where:

- $Dmin_n$ = the lower bound (the minimum value) of conduit sizes for conduit group n ;

D_{max_n} = the upper bound (the maximum value) of conduit sizes for conduit group n ; and
 D_n = the conduit size for conduit group n and selected from a set of available conduit sizes of D^0 .

Similarly, pump curves are selected from various user-specified pump curves to represent pump station expansion. Storage volumes are computed based on shape of storage, defined either as a mathematical function or in tabular form (head vs. area), and are bound by user-specified minimum and maximum storage depths.

11.2 Solution Methodology

In order to solve the optimal sewer design/rehabilitation problem as formulated in the above equations, a dual-level solution methodology was employed by means of which SWMM5 drainage network simulator was directly embedded into an optimization model. Starting with an initial feasible set of decision variables, it is passed to SWMM5 for use in explicitly satisfying the implicit system constraints and in evaluating the implicit bound constraints. The SWMM5 solution (velocity and hydraulic head) is then passed back to the optimization model for use in quantifying the objective function and any violations in the implicit bound constraints. This information is then utilized to produce an improved set of decision variables that automatically satisfies the explicit variable constraints and that seeks to minimize the objective function. This iterative process is repeated until the best solution is found. The fast messy Genetic Algorithm was adopted as the optimization routine. Figure 11.1 illustrates the sewer collection system optimization process.

11.3 Fast Messy Genetic Algorithm

Genetic algorithms (GAs) are stochastic numerical search procedures inspired by biological evolution, cross-breeding trial solutions and allowing only the fittest solutions to survive and propagate to successive generations. They deal with a population of individual solutions, which undergo constant changes by genetic operations of reproduction, crossover, and mutation.

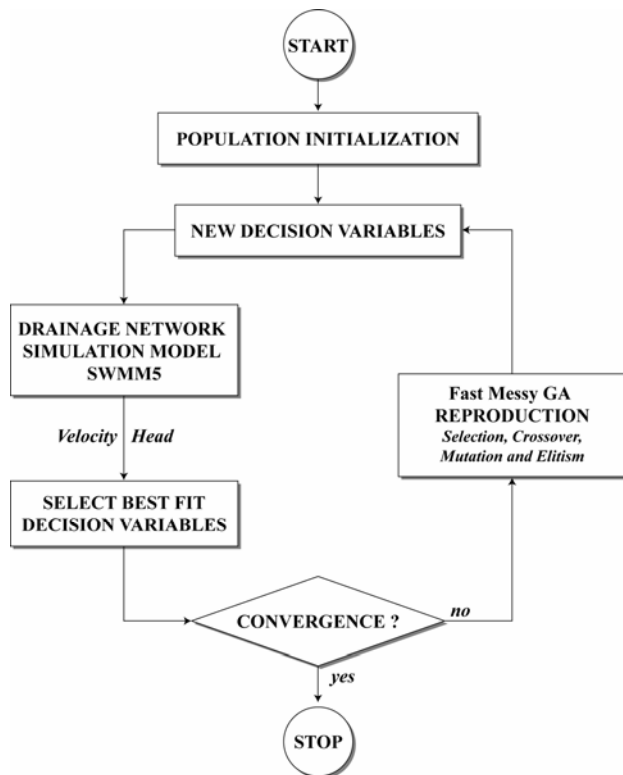


Figure 11.1 Flowchart of urban drainage system optimization process.

The algorithm ranks the solutions from best to worst in terms of their optimality (their fitness with respect to the objective function) where the fit individuals are more likely to reproduce and propagate to the next generation. Based on their fitness values, the algorithm selects individuals (parents) for reproduction of the next generation and exchanges genetic information to form children (crossover). The parents are then removed and replaced in the population by the children to keep a stable population size. The result is a new generation with (normally) better fitness. Occasionally, mutation is introduced into the population to prevent the convergence to a local optimum and help generate unexpected directions in the solution space.

The more GAs iterate, the better their chance to generate an optimal solution. After a number of generations, the population is expected to evolve artificially, and the (near) optimal solution will be reached. The measure of success is the convergence to a population with identical members. The principal advantage of GAs is their inherent ability to intelligently explore the solution space from many different points simultaneously enabling higher probability for locating global optimum without having to analyze all possible solutions available and without requiring derivatives (or numerical approximations) or other auxiliary knowledge.

Different GAs may vary in the way they create the mating pool, choose parents, generate children, or display their population dynamics, and several versions have been proposed with varying level of success. The fast messy genetic algorithm (fmGA) implemented herein is one of the most competent type of GA delivering reliable solutions in sub-quadratic time (Goldberg, 1989; Goldberg et al., 1989, 1990, 1993; Wu et al., 2001; Nicklow et al., 2004; Boulos et al., 2006). It differs from other GA paradigms in solution representation, initialization, and reproduction.

The fmGA employs a flexible scheme of gene representation with variable length strings. The length of strings varies from one individual to another. It allows short strings, namely, partial solutions, to be generated and evaluated during the GA optimization. The short strings retain *building blocks* or clusters of genes that form the good solutions. The fmGA proceeds in two phases of genetic operations. It identifies the building blocks first before exchanging them effectively. The fmGA starts with an initial population of full-length strings and follows by a gene filtering and reproduction process. It identifies short strings of higher fitness by randomly deleting genes from the initial strings. The gene filtering process continues until the length of strings is reduced to a certain order prescribed. The identified short strings are used to form individual solutions. An individual solution is produced by *cut* and *splice* operations. *Cut* and *splice* are different from the standard GA crossover and mutation. *Cut* divides one string into two strings while *splice* concatenates two strings to form one individual solution. Both genetic operations are designed to effectively exchange the building blocks for composing the (near) optimal solution.

The fmGA combines the gene identification and exchange phases into one artificial evolution process, and continues over a number of outer iterations of gene initialization, filtering and reproduction. The iterations are continued until the (near) optimal solution is identified in a population.

11.4 Example Applications

In order to demonstrate the utility and capabilities of the present approach it was applied to the sewer collection system shown in Figure 11.2. This network can be considered as a simplified but reasonably typical municipal combined sewer system. It comprises 64 conduits, 59 junctions, 4 storage, 2 pumps, and 3 outfalls. Two cases are analyzed. The first case solves for the lowest cost to replace system components to meet a 20-y design storm load. The second case solves for the lowest cost to expand the existing system to accommodate new development.

11.4.1 Evaluate Existing System

To estimate existing system performance, we simply load the network with a 20-y return-period design storm. The notation in Figure 11.2 shows the volume of water leaving the flooded nodes and the number of minutes each conduit experiences a surcharged condition resulting from the application of the 20-y design event to the existing stormwater system. Clearly, this system is significantly undersized for a twenty-year return period event.

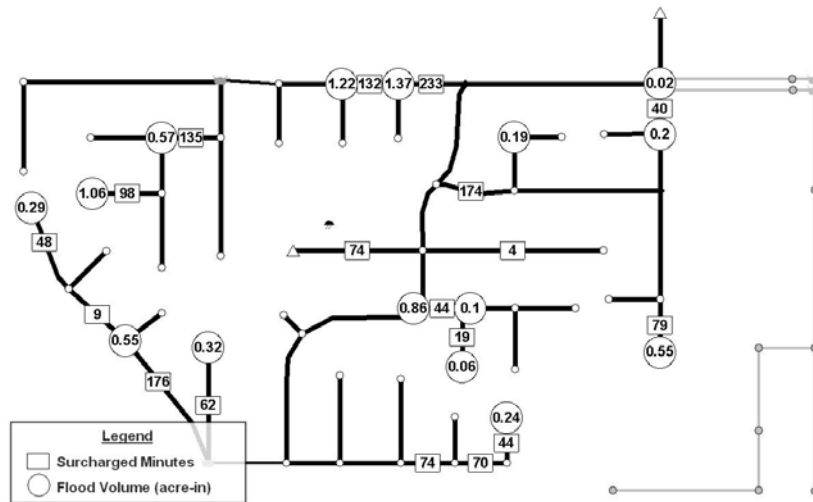


Figure 11.2 Surchage time and flood volume under 20-y return period storm.

Setting Conduit Constraints

For this example, the conduit constraints must reflect community standards for flooding under a 20-y design storm. For this community, the conveyance system must carry a 20-y storm with sufficient additional capacity at any point to receive stormwater flows from directly connected downspouts in some areas. To meet this goal in our design, we set the ratio of maximum depth to full section depth to 0.75 for areas with directly connected downspouts and 0.85 for other areas. We set other conduit constraints to meet generally acceptable engineering design values for minimum and maximum velocities, maximum headloss for force mains, and minimum and maximum invert elevations.

Setting Unit Costs

The proposed optimization model decouples the network objects unit-cost functions from the network object constraints. This independence provides flexibility in defining both the unit costs and conduit constraints. Whereas the two categories of conduit constraints are applied according to the presence/absence of downspout connections, the unit costs vary spatially with other engineering considerations such as the traffic condition, soil type, terrain slope, etc. Figures 11.3 and 11.4 show the spatial distribution of the network unit costs, and the ten unit cost curves applied to these areas.

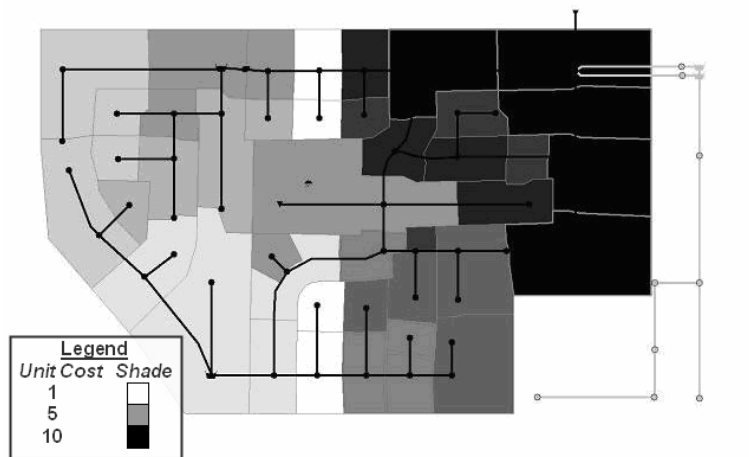


Figure 11.3 Spatial distribution of unit cost categories (1-10).

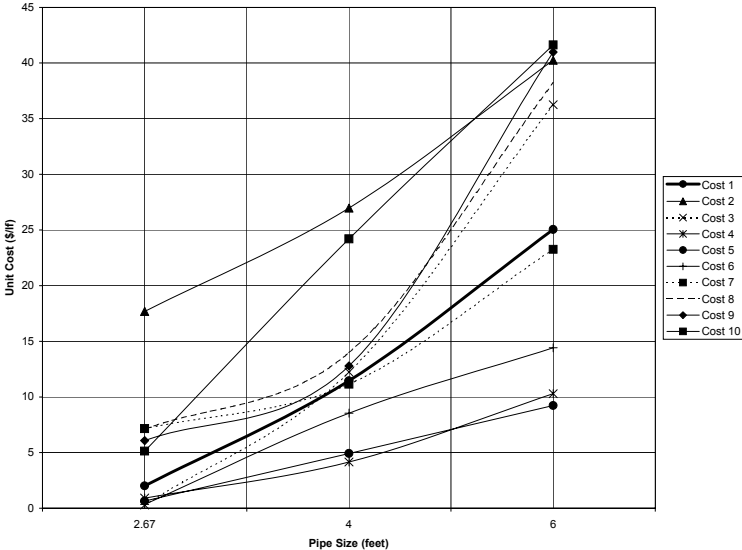


Figure 11.4 Unit cost curves.

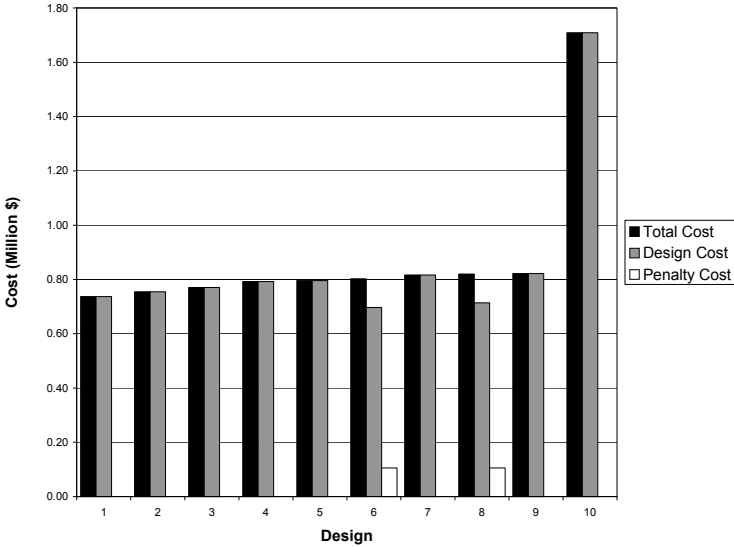


Figure 11.5 Optimization results.

Results

The optimization model yielded the top ten minimum cost designs as shown in Figure 11.5. All of these designs met the requirements for the ratio of maximum depth to full section depth. The difference in cost between solution #1 and solution #2 is US\$17,586.00 or 2.4% of project cost, and the difference between design #1 and #3 is US\$33,828 or 4.6% of project cost. Notably, two of the top ten designs failed to meet minimum velocity requirement for the 20-y storm, incurring a slight penalty.

11.4.2 Design New Development

The next example utilizes the proposed optimization model to determine the optimal solution for expanding the existing sewer network to accommodate a new development. The outfall requirements for this area stipulate that the current peak flow cannot be exceeded. Therefore, the challenge is to properly size a flow equalization basin and any necessary pumping. Figure 11.6 highlights the area of the proposed new expansion. Conduit constraints for this example are similar to the ones used in the previous application and their associated unit costs are shown in Figure 11.4 curve number 5. Unit cost curves for pumping and storage are given in Table 11.1.

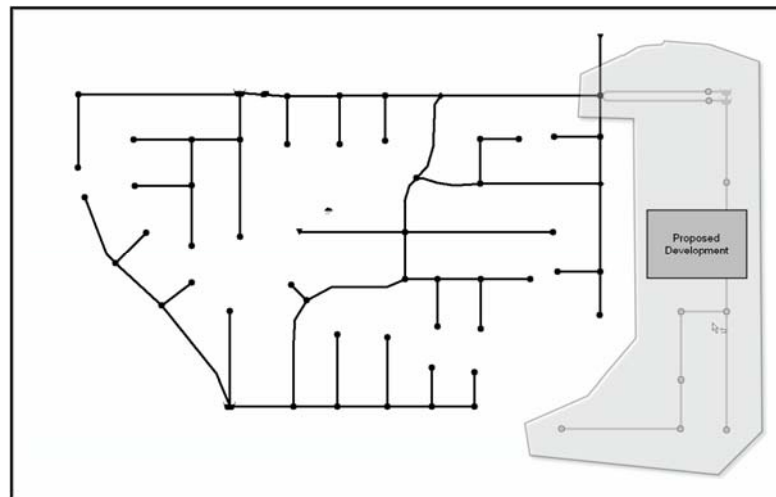


Figure 11.6 Location of proposed development.

This optimal design application takes advantage of the explicit constraints to limit the flow to the outlet by constricting the discharge line for normal flow to equal the maximum allowed discharge. A penalty occurs when the ratio of maximum depth to full section depth at the conduit leading to the outfall equals 98%. In this manner, the model is constrained based on a peak flow loading condition.

Because the decision variables for this problem involve pipe diameter, pump size and storage volume, the solution space will contain billions of possible designs. The model yielded the optimal solution in only 10,000 iterations (5 min). In just the top 10 solutions, there is a range of total project costs that span several thousands of dollars. The model also provides the needed information for an engineer to apply sound judgment in balancing the total project cost with the possibility of exceeding a constraint.

Table 11.1 Pump and storage unit costs.

Pump Unit Cost (\$/pump)	Pump Capacity (cubic feet/second)	
3,000.00	0.5	
5,000.00	1	
7,500.00	1.5	
11,000.00	2	
15,000.00	2.5	
17,500.00	3	
Storage Unit Cost (\$/cubic foot)	Diameter (feet)	Area (square feet)
2.25	8	1,670
2.40	8	3,340
2.50	8	6,685
2.60	8	10,025
2.70	10	10,695
280	10	13,370

11.5 Conclusions

Sewer collection systems are a valuable part of our nation’s infrastructure. EPA estimates that the nation’s sewers are worth a total of more than \$1 trillion (Nicklow et al., 2004, 2006). Many of these systems are experiencing chronic overflows, which are compounding the major urban pollution problems that cities are facing from non-point sources of runoff entering streams and rivers. In light of the great need and apparent commitment to

expend vast amounts of money to improve the reliability and integrity of the nation's sewer systems, it is imperative that sound analysis procedures are available to evaluate cost-effective options for upgrading and modifying these systems and eliminating unwanted overflows.

This chapter has presented a robust and efficient computer optimization methodology for use in determining cost-effective design and rehabilitation strategies for simple and complex urban drainage systems. System improvement alternatives can include any desired combination of pipe slope and upsizing, storage, pumping and new piping to eliminate unwanted sewer overflows and achieve targeted system performance requirements. Performance criteria include maximum allowable depth to diameter ratio, minimum and maximum pipe velocities, maximum head loss for force mains, and minimum and maximum pipe slopes. The model is driven by overall system improvement cost and system performance considerations, and is solved by linkage of a stochastic optimizer with a drainage network simulator (SWMM5). Iterations between the simulator and optimizer are continued until the best solution is obtained. The search procedure is derived from emulating the evolution process observed in nature and biological organisms. This evolution process is based on the preferential survival and reproduction of the fittest member of a population with direct inheritance of genetic information from parents to offspring and the occasional mutation of genes. It is assumed that the drainage network model representation and calibration accurately reflect the actual sewer system, something all procedures require, but whose need cannot be overstated.

The utility and capabilities of the proposed method were demonstrated by application to an example combined sewer system. Two cases were analyzed. The first case illustrates the use of the optimization method to size undersized sewer mains to eliminate overflows, while the second application illustrates the ability of the method to solve for the lowest cost solution to expand the existing system to accommodate new development. The method demonstrated its efficiency by searching only a small fraction of the total search space, yielding optimal solutions in expeditious times. Most of the input requirements coincide with actual design variables for typical drainage networks making the optimization tool directly amenable for use by practicing engineers. Because the method retains the reality of existing urban drainage systems, it readily provides wastewater utilities with a powerful and effective tool for conceiving and evaluating sound and economical design, expansion or rehabilitation alternatives and optimizing their capital improvement programs.

References

- Boulos, P.F., K.E. Lansey and B.W. Karney (2006). "Comprehensive Water Distribution Systems Analysis Handbook for Engineers and Planners." Second Edition. MWH Soft Pub., Pasadena, CA, 660 pp.
- Goldberg, D.E. (1989). "Genetic Algorithms in Search, Optimization and Machine Learning." Addison Wesley, Reading, MA.
- Goldberg, D.E., B. Korb and K. Deb (1989). "Messy genetic algorithms: motivation, analysis, and first results." *Complex Systems*, **3**, pp. 493-530.
- Goldberg, D.E., B. Korb and K. Deb (1990). "Messy genetic algorithms revisited: studies in mixed size and scale." *Complex Systems*, **4**, pp. 415-444.
- Goldberg, D.E. et al. (1993). "Rapid, accurate optimization of difficult problems using fast messy genetic algorithms." In *Proceedings of the Fifth International Conference on Genetic Algorithms*, University of Illinois at Urbana Champaign, Urbana-Champaign, IL, pp. 56-64.
- Nicklow, J.W., P.F. Boulos and M.K. Muleta (2004). "Comprehensive Sewer Collection Systems Analysis Handbook for Engineers and Planners." MWH Soft Pub., Pasadena, CA, 287 pp.
- Nicklow, J.W., Boulos, P.F., and Muleta, M.K. (2006). "Comprehensive Urban Hydrologic Modeling Handbook for Engineers and Planners." 1st ed., MWH Soft Publ., Pasadena, CA, 376 pp.
- Rossmann, L.A. (2005). "Stormwater Management Model Users Manual, Version 5." U.S. Environmental Protection Agency, National Risk Management Research Laboratory, Cincinnati, OH.
- U.S. EPA (2004). "Report to Congress: Impacts and Control of CSOs and SSOs." U.S.EPA Office of Water; Washington, D.C.; August 2004.
- Wu, Z, P.F. Boulos, C.H. Orr and J.J. Ro (2001). "Using genetic algorithms to rehabilitate water distribution systems." *Journal American Water Works Association*, **93**(11), pp. 74-85.

