Benefits of Meta-Model Validation for Real-Time Sewer System Decision Support

Andrea Zimmer[1], Barbara Minsker[1], Arthur Schmidt[1], Avi Ostfeld[4]

[1]Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, [4]Civil and Environmental Engineering, Technion - Israel Institute of Technology

ABSTRACT

Large-scale combined sewer systems are susceptible to overflows (CSOs) during heavy storm events. Management strategies that partition water flow into the sewers or nearby waterways may be based on conservative operational rules designed to prevent possible flow instabilities. However, these operations may not effectively utilize system storage capacity for all types of storm events. Real-time adaptation of system operating rules can reduce overflows while continuing to avoid hydraulic conditions that lead to transients and geysers. In this study, real-time genetic algorithm (GA) optimization is evaluated for its success in minimizing CSOs for a test case modeled after a portion of the Chicago Tunnel and Reservoir Plan (TARP).

INTRODUCTION

Combined sewer overflows are prevalent in many older urban areas (EIP, 2005; Calhoun et al. 2007; EPA NPDES). Because CSO waterway contamination can be expensive to eliminate through construction of additional sewers, cities have alternatively chosen to invest in real-time control of the combined sewer system (Pleau et al. 2005) or deep tunnels (Razak and Christensen, 2001; Dalton and Rimkus, 1985). If a deep tunnel is incorporated into the sewer system, potential overflows are directed to the tunnel via sluice gates and dropshafts.

A meta-model that depicts CSOs for real-time optimization must efficiently incorporate deep tunnel and combined sewer hydraulics (Schutze et al. 2004.) Combined sewer systems have previously been represented for real-time CSO control (Duchense et al., 2001; Darsono and Labadie, 2007; Vanrolleghem et al., 2005) without incorporating a deep tunnel. Increased popularity of deep tunnels to mitigate CSO contamination requires their representation in hydraulic models for real-time control.

In order to accelerate hydraulic computations for this study, graphical depictions of the backwater equations are utilized for online lookup. Interceptor hydraulics are based on mass and momentum conservation through the Hydraulic Performance Graph (HPG) and the Volumetric Performance Graph (VPG), which contain backwater curves that yield water surface elevations and conduit storage values for specific flow rates (Yen and Gonzales-Castro, 2000; Hoy, 2005.) Water flow through sluice gates that connect the combined sewers to the deep tunnel is based on the USGS FEQ equations (Franz and Melching, 1997). Newton-Rhapson iteration is used to solve the system of equations that represent the combined sewer system (Martin, 2010). An EPA SWMM model of the deep tunnel system is used to generate inflow and water surface elevation boundary conditions for the interceptors and deep tunnel. Overland flow to the combined sewer system is simulated through a cell model (Diskin et al. 1984), in which each watershed is represented as two linear reservoirs in series.

A GA is integrated into a Model Predictive Control (MPC) or moving window framework for real-time optimization (Hu and Chen, 2005; Celeste et al., 2004.) Decision variables are real-coded and represent the position of the sluice gates that partition water to the deep tunnel or the CSO as well as a wastewater treatment plant pumping rate that controls water levels in the interceptor lines. Each decision variable is coded on an interval from 0 to 1 in units of 0.1. Additional alterations to the genetic algorithm allow the large search space (each decision variable has 11 possible values) to be evaluated within a short time. These modifications include reductions in decision variable coding, enhanced algorithm memory, and changes to genetic operators.

Within all three types of GA adjustment, convergence to an accurate solution is aided through meta-model validation with the original, more computationally intensive, fitness function (Jin and Branke, 2005). Rao and Salomons (2007) note that during extreme events, the capabilities of a detailed physics-based model may be necessary to portray system conditions. An EPA SWMM model of the interceptor lines is used for validation. Online verification takes place when individual chromosomes (management solutions) are identified for re-evaluation with the actual fitness function. A specified sampling rate determines how many individuals in each generation are evaluated with the actual fitness model (Yan and Minsker, 2006.) This study explores the use of individual-based testing methods with the three different GA modifications.

METHODS

The objective of the CSO optimization is to minimize the predicted volume of overflow, V, calculated at sequential time increments Δt, to be evaluated out to time horizon $t+(\Delta t \times N)$ based on system information available at time t. In the following equation, the variable N represents the number of time increments in the operating horizon.

$$\text{minimize} \sum_{k=1}^{N} \left\{ V\left(t+(k\times\Delta t)\right) \middle| t \right\}$$

The optimization is limited by the physical constraint that upstream flows entering the deep tunnel cannot exceed the capacity of any downstream tunnel conduit C.

$$\frac{\sum_{pipe=C}^{MostUpstream} Q(pipe)}{Capacity(C)} \le 1 \qquad\qquad \forall C$$

A second constraint is based on the Illinois Transient Model (ITM, Leon 2009). This unsteady hydraulic model solves for possible flow instabilities due to high pressures or water acceleration in the deep tunnel. Offline model runs have shown that rapid fluctuations in sluice gate positions lead to increases in water levels that cause backflow to the CSO outlets. A constraint is imposed that prohibits any sluice gate from going from a completely closed position (0) to a fully open position (1) in any two consecutive time intervals.

$$\sum_{k=1}^{N-1} \begin{cases} \text{if } Gate(DS,k) == 1, Gate(DS,k+1) > 0 \\ \text{if } Gate(DS,k) == 0, Gate(DS,k+1) < 1 \end{cases} \qquad\qquad \forall DS$$

Vasconcelos and Wright (2006) show that the risk of geysering may be reduced if the tunnel is operated to prevent pressurized air pockets from rising through water-filled dropshafts. The third constraint restricts flows from entering the tunnel when the most downstream water surface is pressurized. Flows from upstream toward a pressurized boundary would fill dropshafts with water and trap air between the water wave and boundary; propagation of the wave back upstream results in the release of entrained air and water through the dropshafts.

$$\sum_{k=1}^{N} \text{if } \frac{WSE(downstream,k)}{Invert(downstream)+Diameter} \geq 1, Q(DS,k)=0 \qquad \forall DS$$

The third constraint may be substituted with a fourth one below to yield less conservative operational results. Zhou et al. (2001 and 2004) have shown that particular ratios of dropshaft ventilation to cross-sectional areas yield the highest observed tunnel water pressures and should be avoided. Dropshaft ventilation for this study is represented by the opening formed by the sluice gate position.

$$\sum_{k=1}^{N} \frac{A_{sluicegate}}{A_{dropshaft}} \leq 0.15, \frac{A_{sluicegate}}{A_{dropshaft}} \geq 0.3 \qquad \forall DS$$

A meta-model based on water mass and momentum conservation is used to evaluate the objective function and the constraints. The Hydraulic Performance Graph (HPG), which ensures conservation of momentum, and the Volumetric Performance Graph (VPG), which ensures conservation of mass, are comprised of sets of constant flow curves that identify upstream water levels and conduit storage values for downstream water levels and flow rates for each conduit. The HPG plots upstream water surface elevation in the conduit versus downstream water surface elevation. The different curves represent different volumetric flow rates in cubic feet per second. The VPG plots storage volume in the conduit versus the downstream water surface elevation, and the lines represent different flow rates.

Families of curves are established offline for each conduit (identified by different geometry) within the interceptors by simulating the gradually varied flow, or backwater, equation until pressurized flow conditions are met. Once maximum water levels and inflows at the upstream and downstream end of the conduits are reached, the Darcy-Weisbach equation is utilized to continue recording upstream water elevations and storage values. Storage may be included in the nodes between conduits. To solve for mass and momentum conservation within the conduit system in real-time, the water elevation and flow rate at conduit junctions, weirs, and inflows are solved for via Newton-Rhapson iteration. An EPA Storm Water Management Model (SWMM) of the interceptor conduits will be used for offline calibration.

The genetic algorithm operates over a search space of all possible values that each decision variable can take (the alphabet). Alphabet cardinality may be reduced if a threshold sluice gate position can be identified above which open positions may not need to be considered. Online evaluation of the maximum water elevation under the sluice gate during the optimization interval involves calculating the free weir flow. The computed maximum elevation may help eliminate wider gate openings from alphabet consideration. Goldberg (1990) has shown that higher cardinality alphabets are prone to blocking; partial search of only a subset of the alphabet. Reducing the alphabet size allows the subset initially selected by the search algorithm to account for more of the search space, and leads to a reduction in blocking.

Gene shifting can be used to retain memory of good solutions within the algorithm. During model predictive control, an operational strategy for the forecast horizon is developed during the first time interval. The first interval of the optimized strategy is implemented while a new forecast is obtained and the next strategy is found. A strategy based on that proposed by Onnen et al. (1997) is used in which the algorithm population at each interval is initialized by shifting the genes from the two best individuals from the last evolution (time step) while the other population members are initialized randomly. The retained solutions for the two individuals are shifted by one interval (the time duration already considered.)

A third approach used to decrease GA convergence time is the micro-GA (Krishnakumar, 1989; Coello and Pulido, 2001). A very small population is evolved with an increased mutation rate in order to account for the large search space of a real-coded genetic algorithm. In this study, the micro-genetic algorithm evolves a small population of 20 individuals through 40 generations. Within each generation, the population is evolved through 10 sub-generations where, at each sub-generation, 6 of the best individuals are retained and the rest are re-initialized.

Three ways of online meta-model validation with SWMM are explored for the best of the previously described genetic algorithms. The first attempt randomly identifies a predetermined number of individuals for reevaluation at each generation; the second trial evaluates a predetermined number of individuals having the best fitness as determined by the meta-model (Jin et al. 2000, Bull 1999); the third experiment reevaluates a number of the most representative individuals whose fitness is closest to the mean fitness of all individuals (Bhattacharya and Lu, 2003.)

CASE STUDY

The combined sewer system used in this study is based on the North Branch portion of the Chicago Tunnel and Reservoir Plan (TARP.) Data that identified this area as a location with a high number of combined sewer overflows were provided by the Metropolitan Water Reclamation District of Greater Chicago (MWRD). Figure 1 shows the system. The layout consists of five sewersheds flowing to an interceptor on the left side of the river and deep tunnel, and two sewersheds flowing to an interceptor on the right side. Watershed rainfall is converted to runoff through the cell model. Once water from all sewersheds is collected, the interceptors connect and flow to a waste water treatment plant. Weirs installed along the interceptor allow water (once over a certain level) to flow to connecting structures. These connecting structures partition water either to the combined sewer overflow point, or through sluice gates and dropshafts to a deep tunnel. The North Branch tunnel flows to its confluence with the main tunnel section and continues to a downstream pumping station and reservoir. The entire length of the deep tunnel is included in the SWMM model used to generate boundary conditions, but is not depicted in the meta-model. The sluice gate positions determine how water is partitioned between the deep tunnel and the overflow point for each connecting structure, and are used as the decision variables in the optimization. The last decision variable is

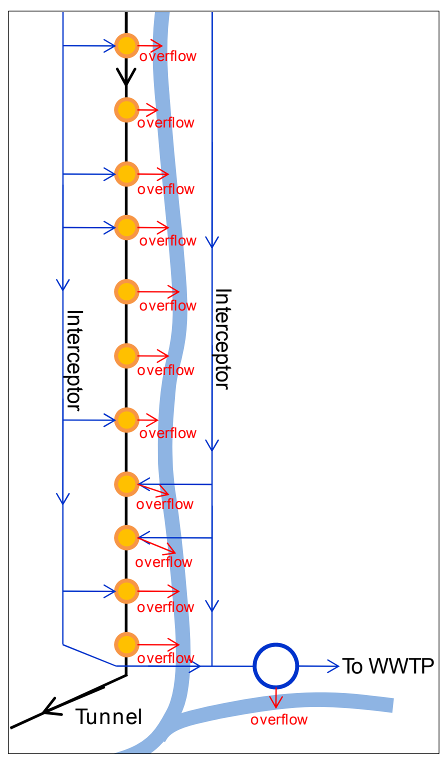

Figure 1: System Layout

the wastewater treatment plant intake rate that controls interceptor water levels; a high treatment plant pumping rate lowers interceptor water levels and allows less water flow over the weirs to the CSO or the deep tunnel.

RESULTS AND DISCUSSION

NEXRAD radar rainfall data for two storms (shown in Figure 2) are used to determine rainfall hyetographs for each watershed. The July 2010 storm shows high rainfall intensities and the deep tunnel is pressurized for the latter half of the event. The August 2007 storm has lower rainfall intensities and the tunnel does not pressurize.

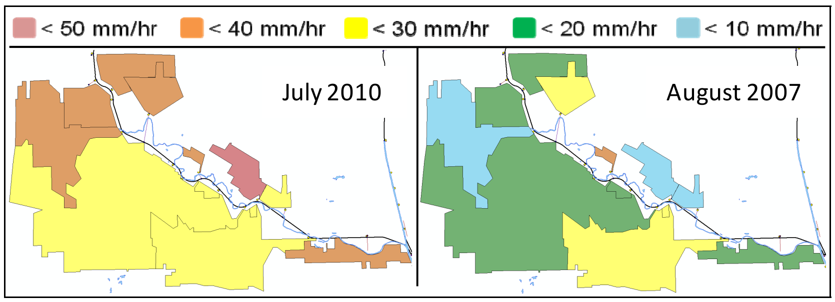

Figure 2: Rainfall Analyzed

Figure 3 compares the convergence of two types of genetic algorithms for the two storm events. Solid lines plot convergence for the gene shifting approach for memory conservation, and dotted lines depict the reduced alphabet method.

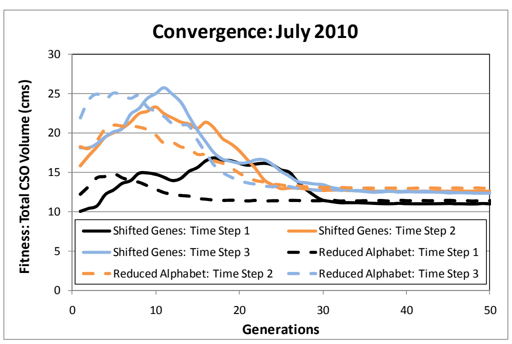

Figure 3a: GA convergence for heavy rainfall

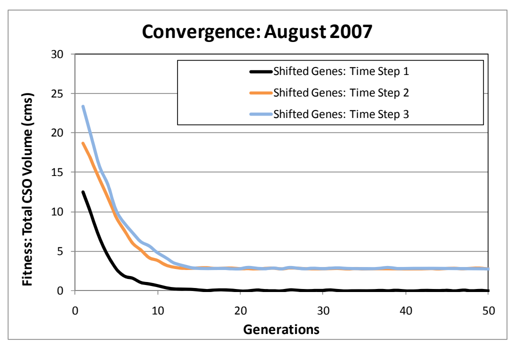

Figure 3b: GA convergence for light rainfall

Less water in the combined sewer system (for the August event) results in a lower number of generations to convergence than for the July event. For the July event, it can be seen that the reduced alphabet approach converges in fewer generations than the shifted gene approach. Better convergence gained through decision variable manipulation indicates that the use of hydraulic knowledge to enhance GA properties is more effective for this system than propagating solution memory.

The operating policy developed for the first two hours of the July storm event is shown in Figure 4. The CSO distribution indicates that overflows do not occur until water in the deep tunnel is pressurized. At the time of pressurization (interval 5), all sluice gates are closed or high tunnel water levels cause backflow.

Decisions optimized for the low intensity August event result in a single overflow; this CSO at dropshaft DS-N01 is attributed to a conveyance capacity limit in the connecting structure. All excess flow from the interceptors cannot pass through the small conduit to the deep tunnel, and water is redirected to the overflow point. Future work will locate additional points where structural capacity limitations, in addition to sluice gate positions, cause overflows. Multi-objective analysis may indicate that the immediate expense of replacing capacity limited conduits may be more beneficial than the continued risk of combined sewer overflows.

The correlation between rainfall, deep tunnel level, and combined sewer overflows is clarified when plots of overflow volumes (at the dropshafts or the wastewater treatment plant) are superimposed with rainfall rate and tunnel level in Figure 5.

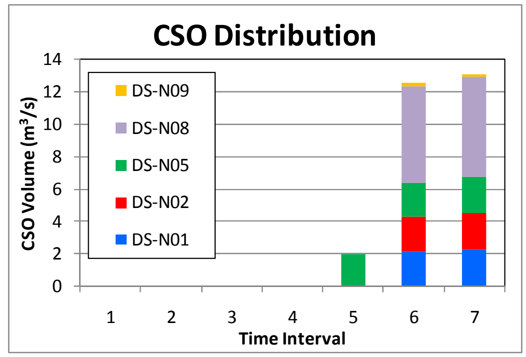

Figure 4: CSO Distribution for July 2010

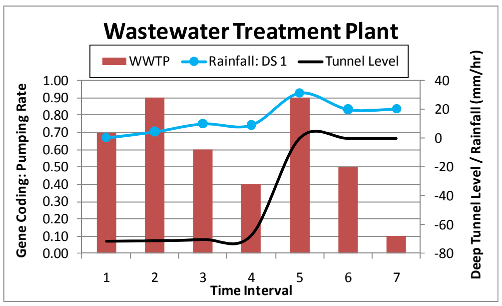

Figure 5a: Pumping Rate over Time

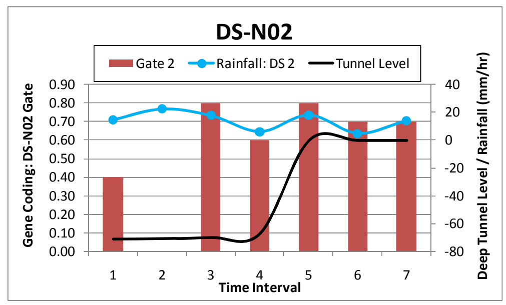

Figure 5b: DS-N02 CSO over Time

High intensity rainfall and high tunnel levels lead to open sluice gates to reduce potential overflows. The treatment plant pumping rate decreases once the deep tunnel fills, most likely to eliminate downstream CSOs when high dropshaft overflow volumes begin. System decision variables are designed to depict tradeoffs. Dropshaft CSOs indicate high interceptor levels and are an alternative to lowering interceptor levels by increasing the treatment plant pumping rate. Increasing the pumping rate above treatment plant capacity, in order to relieve dropshaft CSOs, leads to overflows at the treatment plant.

Current tunnel operating policies keep all sluice gates open until the tunnel is 70% full, then all gates are closed. Figure 4 shows that the optimized decisions resemble the current operating policy due to pressurized tunnel levels that cause high backwater elevations at the sluice gates where CSO flow is calculated. This strategy may be attributed to the third constraint listed above which prohibits water from entering the dropshaft when downstream tunnel levels are pressurized. Future work will focus on altering this constraint for a less conservative approach that continues to prohibit geysers.

CONCLUSIONS AND FUTURE WORK

Longer durations for GA convergence result when high volumes of water have entered the combined sewer system and all conduits reach their flow capacity. The reduced alphabet approach is the most successful of all genetic algorithm types compared in reducing the convergence time. The reduced alphabet GA is enhanced through the use of hydraulics to enhance GA skills. As a result, additional optimization methods with enhanced domain knowledge will be explored for high intensity rainfall events. Different storm events may yield more pronounced overflow distributions, and optimized decision results for many different types of storm events will be mined to complement genetic algorithm memory. The approaches should also be implemented with different random number seeds to ensure that the solutions are optimal or near-optimal.

Operational rules may change based on the tunnel inflow constraint used in the optimization. The two constraints designed to prevent hydraulic transients and geysering will be used with the most effective of the three types of GA to compare resulting operational strategies. Hydraulic models in addition to the meta-model, as well as verification with system operating personnel, will be used to validate the use of either constraint.

Future algorithm development may involve rainfall classification into certain storm types and magnitudes. Optimized sluice gate positions can be attributed to each storm. The previously optimized decisions can yield sets of offline decision rules specific to storm event types that can be used to effectively initialize and accelerate GA solution when rainfall is predicted.

REFERENCES

Bhattacharya, M., and Lu, G. (2003). DAFHEA: A Dynamic Approximate Fitness based Hybrid Evolutionary Algorithm. *Proc. IEEE Congress on Evolutionary Computation*, 3, 1879-1886.

Bull, L. (1999). On Model-Based Evolutionary Computation. *Soft Computing*, 3, 76-82.

Calhoun, L., Avery, M., Jones, L., Gunarto, K., King, R., Roberts, J., and Burkot, T. (2007). Combined Sewage Overflows (CSO) are Major Urban Breeding Sites for Culex quinquefasciatus in Atlanta, Georgia. *American J. of Tropical Medicine and Hygiene,* 77(3), 478-484.

Celeste, A.B., Suzuki K., and Kadota, A. (2004). Genetic algorithms for real-time operation of multipurpose water resource systems. *Journal of Hydroinformatics*, 6(1), 19-38.

Coello Coello, C.A., and Pulido, G.T. (2001). A micro-genetic algorithm for multiobjective optimization. *Proc. Genetic and Evolutionary Computation Conf.*, GECCO, 274-282, San Francisco.

Dalton, F., and Rimkus, R. (1985). Chicago Area's Tunnel and Reservoir Plan. *Journal of the Water Pollution Control Federation*, 57(12), 1114 – 1121.

Darsono, S., and Labadie, J. (2007). Neural-optimal control algorithm for real-time regulation of in-line storage in combined sewer systems. *Environmental Modeling and Software*, 22, 1349-1361.

Diskin, M.H., G. Wyseure, and J. Feyen. (1984). Application of a Cell Model to the Bellebeek Watershed. *Nordic Hydrology*, 15, 25-38.

Duchense, S., Mailhot, A., Dequidt, E., and Villeneuve, J. (2001). Mathematical modeling of sewers uncer surcharge for real time control of combined sewer overflows. *Urban Water*, 3, 241-252.

Environmental Integrity Project (EIP). (2005). Backed-Up. Cleaning up Combined Sewer Systems in the Great Lakes.

Environmental Protection Agency (EPA) National Pollutant discharge Elimination System (NPDES). http://cfpub.epa.gov/npdes/home.cfm?program_id=5

Franz, D., and Melching, C. (1997). Full Equations (FEQ) model for the solution of the full, dynamic equations of motion for 1-D unsteady flow in open channels and through control structures. USGS Water-Resources Investigation Report 96-4240. Washington, D.C.: U.S. Geological Survey.

Goldberg, D. E. (1990). Real-coded Genetic Algorithms, Virtual Alphabets, and Blocking. illiGAL Report No. 90001, The Illinois Genetic Algorithms Laboratory.

Hoy, M. (2005). Unsteady Flow Routing Using Predetermined Solutions to the Equations for Conservation of Mass and Momentum. MS Thesis, University of Illinois at Urbana-Champaign.

Hu, X., and Chen, W. (2005). Genetic algorithm based on receding horizon control for arrival sequencing and scheduling. *Engineering Applications of Artificial Intelligence*, 18, 633-642.

Jin, Y., and Branke, J. (2005). Evolutionary Optimization in Uncertain Environments – A Survey. *IEEE Transactions of Evolutionary Computation*, 9(3), 303-317.

Jin, Y., Olhofer, M., and Sendhoff, B. (2000). On evolutionary optimization with approximate fitness functions. *In: Proceedings of the Genetic and Evolutionary Computation*, 786-792.

Krishnakumar, K. (1989). Micro-genetic algorithms for stationary and non-stationary function optimization. *SPIE Proceedings (1196): Intelligent Control and Adaptive Systems*, 289-296.

Leon, A., Choi, N., Schmidt, A., and Garcia, M. (2009). The Illinois transient Model. V.Y. Chow Hydrosystems Lab, Dept. of Civil and Env. Eng., University of Illinois at Urbana Champaign.

Onnen, C., Babuska, R., Kaymak, U., Verbruggen, H.B., and Isermann, R. (1997). Genetic Algorithms for Optimization in Predictive Control. *Control Eng. Practice*, 5(10), 1363-1372.

Pleau, M., Colas, H., Lavallee, P., Pelletier, G., and Bonin, R. (2005). Global optimal real-time control of the Quebec urban drainage system. *Environ. Modelling and Software*, 20, 401-413.

Rao, Z., and Salomons, E. (2007). Development of a real-time, near-optimal control process for water-distribution networks. *Journal of Hydroinformatics*, 9, 25-37.

Razak, I., and Christensen, E. (2001). Water quality before and after deep tunnel operation in Milwaukee, Wisconsin. *Water Research*, 35(11), 2683-2692.

Schutze, M., Campisano, A., Colas, H., Schilling, W., and Vanrolleghem, P. (2004). Real time control of urban wastewater systems – where do we stand today? *J. of Hydrology*, 299, 335-348.

Vanrolleghem, P., Benedetti, L., and Meirlaen, J. (2005). Modelling and real-time control of the integrated urban wastewater system. *Environ. Modeling and Software*, 20(4), 427-442.

Vasconcelos, J., and Wright, S. (2006). Mechanisms for Air pocket Entrapment in Stormwater Storage Tunnels. *Proc. 2006 World Environmental and Water Resources Congress*, May 21-25, 2006, Omaha, Nebraska.

Yan, S., and Minsker, B. (2006). Optimal groundwater remediation design using an Adaptive Neural Network Genetic Algorithm. *Water Resources Research*, 42.

Yen, B.C. and Gonzales-Castro, J.A. (2000). Open-Channel Capacity Determination Using Hydraulic Performance Graph. *Journal of Hydraulic Engineering*, 126(2), 112-122.

Zhou, F., Tsang, D., Hicks, F., Steffler, P., and Ward, C. (2001). Pressure Oscillation Induced by Trapped Air in a Trunk-Shaft Sewer System. *Proc. 15th Canadian Hydrotechnical Conference*, May 30 – June 2, Victoria, British Columbia.

Zhou, F., Hicks, F., and Steffler, P. (2004). Analysis of effects of air pocket on hydraulic failure of urban drainage infrastructure. *Can. J. Civ. Eng*, 31, 86-94.