Generating alternative watershed-scale BMP designs with evolutionary algorithms

K. S. Artita[1], P. Kaini[2], and J. W. Nicklow[3]

[1] Graduate Research Associate, Department of Civil and Environmental Engineering, Southern Illinois University Carbondale (SIUC), Mail Code 6603 Carbondale, IL 62901; Ph: (618) 536-2368; Fax: (618) 453-3044, Email: kartita@siu.edu
[2] Graduate Research Associate, Department of Civil and Environmental Engineering, SIUC, Email: pkaini@siu.edu
[3] Associate Dean and Professor, College of Engineering, SIUC, Email: nicklow@engr.siu.edu

Abstract

The first part of a two-step decision-making framework for watershed-scale stormwater runoff control (Kaini et al., 2008, this meeting) involves identification of the most cost-effective combination of structural BMPs that meet target peak flow and sediment reduction criteria. This paper presents the second half of that framework: simultaneous generation of near-optimal alternative design strategies using a Euclidean distance metric. Structural BMPs included in this model include detention ponds, infiltration ponds, field borders, grade stabilization structures, and grassed waterways. Alternative designs are identified by coupling Soil and Water Assessment Tool (SWAT) and a Species Conserving Genetic Algorithm (SCGA). In addition, we demonstrate SCGA's flexibility and efficiency at generating alternative designs as well as varying numbers of alternatives. The model is demonstrated on Silver Creek watershed, a sub-watershed of the larger Lower Kaskaskia watershed in southern Illinois.

Introduction

Recent studies in water resources planning and management show a gradual shift in the state of the art from numerous on-site BMP designs to watershed-scale design approaches (e.g., Yeh and Labadie, 1997; Harrell and Ranjithan, 2003; Zhen et al., 2004; Perez-Pedini et al., 2005). For example, Emerson et al. (2005) showed that for detention-based systems, a series of on-site BMPs were not particularly effective in collectively reducing watershed-wide peak flows and potentially increased peak flow rates. Instead, structural BMPs are most cost-effective when designed and implemented in regionally-strategic combinations to meet related stormwater treatment goals.

Biologically-inspired, evolutionary algorithms (EAs) are increasingly popular in engineering applications and have proven effective in search and optimization. Typically, EA-based methods focus on identifying an optimal solution to single objective problems and a set of non-dominated solutions in multi-objective problems. However, in real world applications, such as the design of a storm water control system, a mathematically optimal solution may not be the ideal solution for decision makers and stakeholders. There may exist alternative suboptimal solutions that are also cost-effective but that perform differently with respect to unmodeled and unquantifiable objectives (e.g. personal preferences, aesthetics, ecological impacts, etc.). Thus, for practical reasons, it is desirable to identify a relatively small but diverse set of high-quality alternative solutions.

Genetic algorithms (GAs) have been used to locate multiple optima in multimodal functions through two main techniques: iterative methods and parallel subpopulation methods. Iterative methods, as the name implies, repeatedly applies the same algorithm to a problem to locate multiple optima. In successive iterations, portions of the decision space are eliminated from the search to prevent convergence to previously discovered optima (e.g. Sequential Niche Technique, Beasley et al., 1993; Tabu search, Glover, 1989). Parallel subpopulation methods divide the population into several subpopulations that evolve simultaneously (e.g. Muhlenbein et al., 1991; Chen and Chang, 2007). To prevent convergence to the same optimum, subpopulations are designed to “communicate” with one another, however, at the expense of reducing solution diversity (Davidor, 1991).

Zechman and Ranjithan (2004, 2007) introduced the Evolutionary Algorithm to Generate Alternatives (EAGA) to identify multiple, maximally different solutions to engineering and water resources and environmental management problems. EAGA relies on the spatially-explicit nature of the decision space to determine differences between near-optimal solutions using a centroid- or elitist-based distance metric. The primary subpopulation seeks out the global optimum. Meanwhile, the other subpopulations, corresponding to the number of desired alternatives, are evolved either iteratively or in parallel. One major drawback of EAGA is its potential to “fail” when attempting to optimize problems with multiple global optima. It is important to note that EAGA is not a multimodal function optimizer: that is, it locates any solution that is maximally different and above a predefined objective function value threshold. In addition, each additional alternative requires another subpopulation and therefore an increase in computational expense.

Another relatively new parallel subpopulation method inspired by ecology, Species Conserving Genetic Algorithm (SCGA), is not prone to loss of solution diversity and does not require division of the population into subpopulations. SCGA has been proven an effective multimodal function optimizer (Li et al., 2002, 2003) and method of generating alternative designs (Li et al., 2007). This study presents the second half of a decision-making framework for watershed-scale stormwater runoff control that couples SCGA with USDA's Soil and Water Assessment Tool (SWAT). In addition to identifying the most cost-effective combination of structural best management practices (BMPs) that meet target peak flow and sediment reduction criteria, the SCGA-SWAT optimization model simultaneously generates near-optimal alternative design strategies using a Euclidean-based distance metric. Structural BMPs considered in this model include detention ponds, infiltration type detention ponds, borders, grade stabilization structures, and grassed waterways. The SCGA-SWAT model is applied to Silver Creek watershed, a sub-watershed of the larger Lower Kaskaskia watershed in Illinois.

Species Conserving Genetic Algorithm

Island or multi-population model EAs are inspired by the formation of ecological niches and speciation as observed in nature. Isolated environments, such as islands, nurture evolutionary phenomena. Favorable ecological niches are generally sought for abundance of resources and support rapid colonization, whereas, unfavorable niches tend to promote species extinctions. Island model EAs have the advantage of searching a decision space more effectively than traditional EA methods and are less prone to stagnation at a local optimum through an increased capability in maintenance of population diversity. One of the major shortcomings of the EAGA, which is one type of island model EA, is its inability to maintain diversity amongst individuals within a subpopulation. Lack of diversity is especially troublesome in the case of multimodal problems. Explicit methods of diversity maintenance are fitness sharing (Goldbert and Richardson, 1987; Della Cioppa et al., 2004) and crowding (Mahfoud, 1995).

The SCGA is best described as a non-conventional island model EA. Speciation is the process by which individuals adapt in order to occupy new and different environmental niches. The entire population in SCGA is finite; however, species (subpopulations) are not absolute. That is, population sizes of each species are variable from generation to generation: individuals belonging to one species may migrate to another in subsequent generations and species can be created or destroyed. In this study, the distance between two individuals, x_i and x_j, is defined as the Euclidean distance between their representative vectors in decision space,

$$d\left(x_i, x_j\right) = \sqrt{\sum_{k=1}^{n} \left(x_i^k - x_j^k\right)^2}$$

where each individual is represented by an n-dimensional array. Each species is dominated and centered on its species seed, the best individual that dominates all other individuals in the specie. For example, species S_i is centered on its species seed x^* such that for every individual y in S_i

$$d\left(x^*, y\right) < \sigma_s \tag{1}$$

where σ_s is the species distance or neighborhood radius. The species distance is taken as the Euclidean distance between solutions in the decision space (i.e., the number and location of BMPs).

Structure of SCGA

Operation of SCGA requires three special procedures embedded in the traditional GA framework (for greater detail, see Li et al., 2002, 2003):

1. Determination of species seeds. Every individual in the population in generation *G(t)* is considered as a potential species seed in the set X_s in decreasing order of fitness. Each individual is compared with existing species seeds and if there is no seed closer than the species distance, σ_s, the individual becomes a species seed and is added to X_s. Procedure is repeated for all individuals.
2. Conservation of species seeds. After generation *G(t+1)* has been created through selection, crossover, and mutation, species seeds in X_s may be conserved thereby allowing exact duplicates of certain species seeds to remain in the population. The importance of this step is straightforward: some species may not survive the genetic operations. These species tend to have lower

fitness but remain essential to the search for multiple optima by maintaining diversity. Species seeds are conserved by identifying individuals of the same species in *G(t+1)* and replacing the worst of these if the seed has a higher fitness. In the event there are no other members of the species in *G(t+1)*, the seed replaces the individual with the lowest fitness in *G(t+1)*.

3. Identification of alternative "best" solutions. The optimal solution(s) and alternative near-optimal solutions to the problem are identified by considering species seeds in X_s. Seeds above an user-define threshold, r_f ($0 < r_f < 1$), are then accepted as the best design alternatives such that

$$f(x) \geq f_{\max} \times r_f \tag{2}$$

SCGA, in this study, creates *G(t+1)* solutions through binary-tournament selection, uniform crossover, and gene-wise, bit-flip type mutation. Potential BMP designs are represented as binary strings where 1 represents placement of a BMP and 0 represents no BMP. Species seeds within 25% of f_{max} ($r_f = 0.25$) are chosen as design alternatives.

Application of SCGA to Silver Creek

Silver Creek is a predominantly agricultural watershed that is a part of the larger Kaskaskia watershed in southern Illinois. Silver Creek was divided into 159 subbasins where each subbasin is equivalent to a hydrologic response unit (HRU). Each HRU is characterized by a dominant land use and soil type. Of the 159 HRUs, 77 are pasture lands (48.4%), 70 are agricultural lands (44.0%), 9 are wetlands (5.7%), and 3 are deciduous forests (1.9%).

Data needed to run and calibrate SWAT were collected from various sources within the public domain. Precipitation data for the study area was obtained from the National Climate Data Center (NCDC). A 30-m resolution Digital Elevation Model (DEM) was downloaded from the United States Geographical Survey (USGS) website. Daily streamflow data for hydrologic station, 05594800 (Freeburg, St. Clair County, IL) was also collected from USGS. Digital soil maps and land use maps were obtained from the Natural Resources Conservation Service (NRCS).

SWAT was calibrated for daily streamflow at the outlet of Silver Creek watershed using a multi-objective particle swarm optimization (MOPSO; Gill et al., 2006). The calibration objectives were to simultaneously minimize the root mean squared error (RMSE) between observed and simulated daily flow, given by:

$$RMSE = \sqrt{\frac{1}{N} \times \sum_{j=1}^{N} \left(O_j - S_j\right)^2} \tag{3}$$

where O_j and S_j are the jth observed and simulated streamflow values; and minimize the log error (LOGE) expressed as:

$$LOGE = \sqrt{\frac{1}{N} \times \sum_{j=1}^{N} \left(\log\left(\frac{O_j}{S_j} \right) \right)^2} \tag{4}$$

RMSE gives strong emphasis on fitting the higher or peak output values while LOGE tends to emphasize fitting lower output values (Bekele and Nicklow, 2007). The calibration was run for 3 years (1992-1994), initialization for 2 years (1990-1991), and verification for 2 years (1995-1996). The set of streamflow parameters chosen from the MOPSO-generated non-dominated set for use in this study had a RMSE of 1.3656 and LOGE of 2.2540 for the calibration years and a RMSE of 1.7522 and LOGE of 2.6934 for the verification years. Note that SWAT was not calibrated for sediment due to lack of data.

Once calibrated, SWAT was coupled with SCGA with the objective of identifying alternative watershed-scale BMP designs (type, size and location) that:

Minimize: Total cost of BMPs

Subject to the following constraints:

1. Water balance in the watershed
2. Detention pond sizes less than a user-defined maximum limit
3. Land use constraints
4. Peak flow and sediment load reduction criteria
5. BMP placement constraints (i.e., no BMPs can be placed in subbasins with wetlands or forests as its dominant land use type)

Structural BMPs considered include detention ponds, infiltration type detention ponds, field borders, grade stabilization structures, and grassed waterways. Optimal BMP designs found with a simple genetic algorithm included few parallel terraces, thus they are not considered in this example application. Table 1 lists BMP parameters for this study. For additional details on BMP implementation and for a breakdown of BMP unit costs, the reader is referred to Kaini et al., 2008 (this meeting).

Results

The SCGA-SWAT model was run with 200 individuals for 100 generations with $\sigma_s = 8.0$. Maximum daily peak flow and annual sediment yield constraints were both set to 5% reduction of base values (calibrated case with no BMPs implemented). In the final generation, SCGA-SWAT identified 48 of the 200 individuals as species seeds. The least cost BMP design (species seed 1) and species seeds within $r_f = 25\%$ are listed in Table 2. Detention and/or infiltration ponds are not included in any of the best alternative designs and are not implemented in a design until there is a 90% cost increase above the optimal design cost. Figure 1 shows the spatial distribution of BMPs for the top 6 alternative designs.

Conclusions

A biologically-inspired algorithm, SCGA, was coupled with the SWAT to simultaneously generate optimal watershed-scale BMP designs and several near-optimal alternatives. BMPs considered in this model were detention and infiltration ponds, grassed waterways, field borders, and grade stabilization structures. SCGA-SWAT was demonstrated on Silver Creek watershed and yielded several best alternative designs that did not include any detention or infiltration ponds.

Several key factors appear to determine the quality and number of alternatives generated by SCGA-SWAT. These include the number of HRUs in the watershed, the

Table 1. SWAT parameters for BMP implementation

BMP	SWAT parameter	Parameter description	HRU value	
			without BMP*	with BMP*
Detention pond	*pnd_k*	hydraulic conductivity	0.0	0.0
	pnd_fr	fraction of HRU draining to pond	0.0	0.9
	pnd_ESA	emergency spillway area (fraction of HRU)	0.0	0.01
Infiltration pond	*pnd_k*	hydraulic conductivity	0.0	0.0
	pnd_fr	fraction of HRU draining to pond	0.0	0.9
	pnd_ESA	emergency spillway area (fraction of HRU)	0.0	0.01
Field border	*filterw*	filter width (m)	0	5
Grassed waterway	*ch_cov*	channel cover	0.4855[b]	0.0
	ch_erod	channel erodibility	0.2157[b]	0.0
	ch_n(2)	channel roughness	0.01[a]	0.24
Grade Stabilization Structure	*ch_erod*	channel erodibility	0.2157[b]	0.0
	ch_s(2)	channel slope steepness	-	§

[*] Unless otherwise noted, values are adopted from Bracmort et al., 2006 and Arabi et al., 2006.

[a] Calibrated value

[b] Un-calibrated value

§ $ch\_s(2)_{new} = ch\_s(2)_{SWAT} - \frac{D}{ch\_l(2)}$, where D is the height of the structure (1.2 m).

Table 2. Results of SCGA-SWAT showing species seeds within top 25%

Seed	**Cost ($ million)**	**% increase**	**Ponds**	**Waterways**	**FBs**	**GSSs**	**Total BMPs**
1	0.523		0	9	18	20	47
2	0.590	0.128	0	11	27	28	66
3	0.604	0.154	0	12	27	25	64
4	0.620	0.185	0	8	25	33	66
5	0.622	0.188	0	10	27	27	64
6	0.627	0.197	0	10	28	28	66
7	0.629	0.202	0	10	28	28	66
8	0.633	0.208	0	13	29	27	69
9	0.641	0.224	0	10	28	30	68
10	0.648	0.239	0	10	27	30	67
11	0.649	0.240	0	8	26	37	71
12	0.653	0.247	0	10	29	30	69

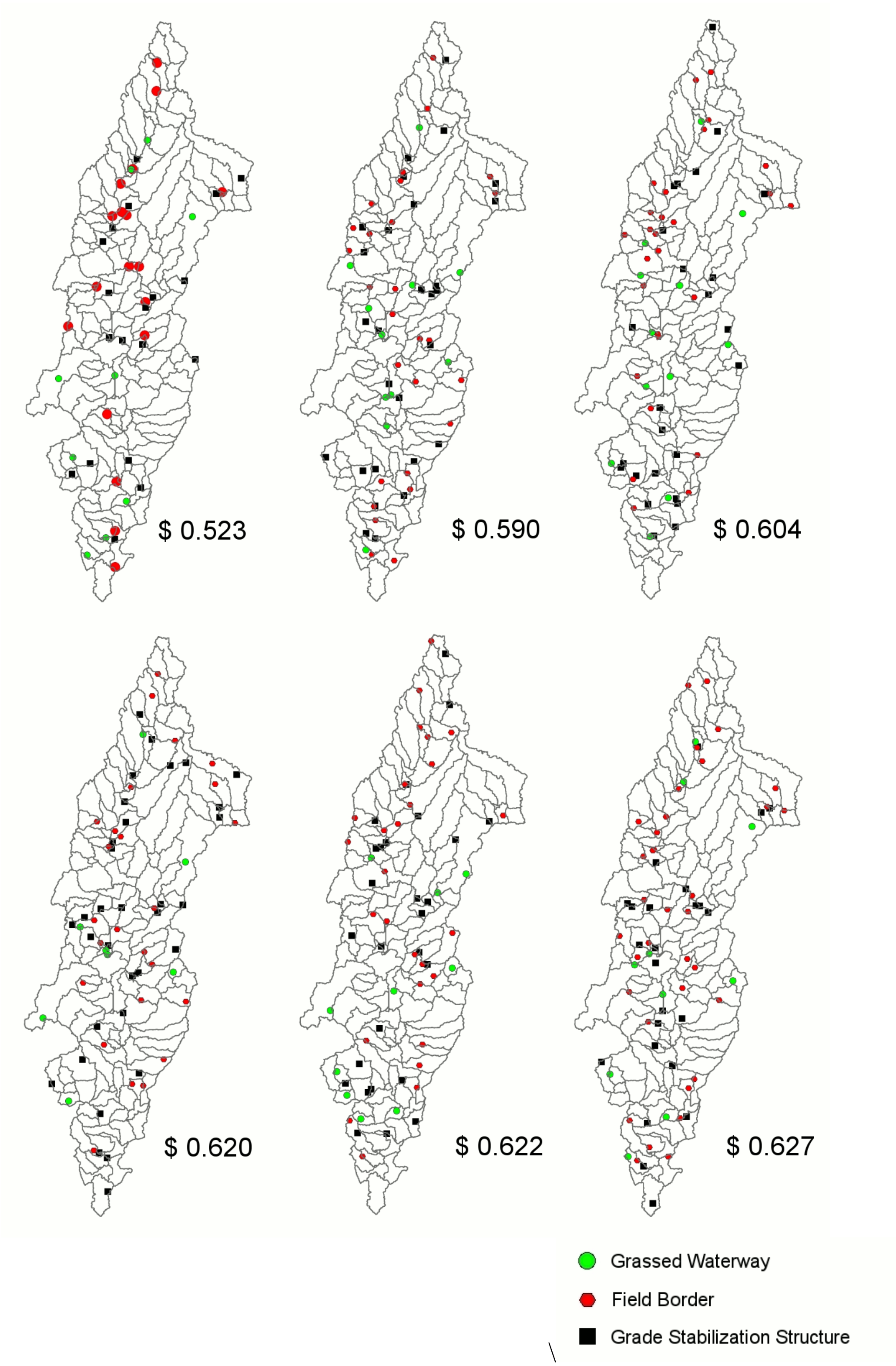

Figure 1. Optimal watershed-scale BMP designs and near-optimal alternatives. Costs listed are in millions USD.

number of different BMP types considered, the number of variables associated with each BMP, the size of the algorithm population, the maximum number of generations, and the choice of species distance. Future work will include increasing the number of BMP design variables (e.g., variability in pond area or fraction of watershed), investigating the sensitivity of SCGA-SWAT to population size, maximum generations, and species distance, and generating alternatives with different distance metrics (i.e., changing how the species distance is measured).

In addition, results of SCGA-SWAT will be presented in a regional educational outreach seminar targeting city and county planners, local stakeholders, local interest groups, and decision makers. We hope to demonstrate the appeal of watershed-scale BMP design that may affect future water resources planning and management in the southern Illinois area.

Acknowledgement

This project is supported by the U.S. Department of the Interior, the U.S. Geological Survey, and the Board of Trustees of the University of Illinois under Grant No. 06HQGR0185. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Government.

References

Arabi, M., Govindaraju, R. S., and Hantush, M. M. (2006).”Cost-effective allocation of watershed management practices using a genetic algorithm.” *Water Resour. Res.,* 42, W10429, doi:10.1029/2006WR004931.

Beasley, D., Bull, D. R., and Martin, R. R. (1993). “A sequential niche technique for multimodal function optimization.” *Evol. Comp.,* 1(2), 101-125.

Bekele, E. and Nicklow, J. W. (2007). “Multiobjective Automatic Calibration of SWAT using NSGA II” *J. Hydrol.,* 341, 165-176.

Bracmort, K.S., Arabi, M., Frakenberger, J. R., Engel, B. A., and Arnold, J. G. (2006). “Modeling long term water quality impact of structural BMPs.” *American Society of Agricultural and Biological Engineers* 49(2), 367-374.

Chen, L., and Chang, F. J. (2007). “Applying a real-coded multi-population genetic algorithm to multi-reservoir operation.” *Hydrol. Process.*, 21, 688-698.

Davidor, Y. (1991). “A naturally occurring niche and species phenomenon: the model and first results.” In Belew, R. K., and Booker, L. B., Eds., *Proc. of the Fourth Int. Conf. on Genetic Algorithms*, San Mateo, CA, 257-263.

Della Cioppa, A., De Stefano, C., and Marcelli, A. (2004). “On the role of population size and niche radius in fitness sharing.” *IEEE Trans. Evol. Comp.*, 8 (6): 580-592.

Deb, K., and Goldberg, D. E. (1989). “An investigation of niche and species formation in genetic function optimization.” In Schaffer, J. D., Ed., *Proc. of the Third Int. Conf. on Genetic Algorithms,* San Mateo, CA, 42-50.

Emerson, C. H., Welty, C., and Traver, R. G. (2005). “Watershed-scale evaluation of a system of storm water detention basins.” *J. Hydrol. Eng.*, 10(3), 237-242.

Gill, M.K., Kaheil, Y.H., Khalil, A., McKee, M., and Bastidas, L. (2006). “Multiobjective particle swarm optimization for parameter estimation in hydrology." *Water Resour. Res.,* 43, W07417.

Glover, F. (1989). Tabu Search – Part I. *ORSA Journal on Computing*, 1(3), 190-206.

Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading Massachusetts, USA: AddisonWelsey.

Goldberg, D. E., and Richardson, J. (1987). "Genetic algorithms with sharing for multimodal function optimization." In Grefenstette, J. J., Ed., *Proc. of the Second Int. Conf. on Genetic Algorithms*, Hilldale, NJ, 41-49.

Harrell, L. J., and Ranjithan, S. R. (2003). "Detention Pond Design and Land Use Planning for Watershed Management." *J. Water Resour. Plan. Mgmt.*, 129(2), 98-106.

Kaini, P, Artita, K. S., and Nicklow, J.W. (2008). "Designing BMPs at a watershed-scale using SWAT and a genetic algorithm." *Proc. of the 2008 World Water and Environmental Resources Cong.*, ASCE, Tampa, FL.

Muhlenbein, H., Schomisch, M., and Born, J. (1991). "The parallel genetic algorithm as function optimizer." Parallel Comput., 17(6-7), 619-632.

Muleta, M. K. and Nicklow, J. W. (2005). "Sensitivity and Uncertainty Analysis coupled with Automatic Calibration for a Distributed Watershed Model" *J. Hydrol.,* 306, 127-145.

Li, J.-P., Balazs, M. E., and Parks, G. T. (2007). "Engineering design optimization using species-conserving genetic algorithms." *Eng. Optimiz.,* 39 (2): 147-161

Li, J.-P., Balazs, M. E., Parks, G. T., and Clarkson, P. J. (2002). "A Species Conserving Genetic Algorithm for Multimodal Function Optimization. *Evol. Comp.*, 10(3): 207-234

Li, J.-P., Balazs, M. E., Parks, G. T., and Clarkson, P. J. (2003). "Erratum: A Species Conserving Genetic Algorithm for Multimodal Function Optimization." *Evol. Comp.*, 11(1), 107-109.

Mahfoud, S. W. (1995). *Niching methods for genetic algorithms.* IlliGAL Technical Report 95001, Illinois Genetic Algorithms Laboratory, Univ. Illinois, Urbana, IL.

Perez-Pedini, C., Limbrunner J. F., and Vogel R. M. (2005). "Optimal Location of Infiltration-Based Best Management Practices for Storm Water Management." *J. Water Resour. Plan. Mgmt.*, 131(6), 441-448.

Yeh, C. H., and Labadie, J. W. (1997). "Multi-objective Watershed Level Planning on Storm Water Detention System." *J. Water Resour. Plan. Mgmt.*, 123(6), 336-343.

Zechman, RM and RS Ranjithan, 2004. "An evolutionary algorithm to generate alternatives (EAGA) for engineering optimization problems." *Eng. Optimiz.*, 36, 539-553.

Zechman, EM and RS Ranjithan, 2007. "Generating alternatives using evolutionary algorithms for water resources and environmental management problems." *J. Water Resour. Plan. Mgmt.,* 133, 156-165.

Zhen, X.-Y., Yu, S. L., and Lin, J.-Y. (2004). "Optimal location and sizing of stormwater basins at watershed scale." *J. Water Resour. Plan. Mgmt.* 130, 339-347.