Investigation of global optimum seeking methods in water resources problems

by Andreas Efstratiadis

Introduction

Optimization has become a valuable tool in most of hydroinformatics applications. Given that these problems are intrinsically nonlinear and multimodal, they do not exist deterministic optimization methods that can locate the globally optimal solution. During the last two decades, probabilistic schemes have been developed for solving global optimization problems. These methods use a combination of random and deterministic steps, without generally requiring restrictive conditions on the nature of the objective function.

The scope of this postgraduate thesis is a detailed investigation of global optimization techniques, focusing on methods used in water resources engineering and management. Within the framework of this study, a new optimization scheme has been also developed which has been proved suitable for most of the theoretical as well as real-world applications that have been examined.

Posing the global optimization problem

The nonlinear unconstrained optimization problem

The nonlinear unconstrained optimization problem with continuous variables is defined as the seeking for a vector $\mathbf{x}^* = (x_1^*, \ldots, x_n^*)$ such that:

$$f(\mathbf{x}^*) = \min f(\mathbf{x}) \tag{1}$$

The main assumptions for this problem are that control variables $\mathbf{x}$ are continuous and bounded and that all external constraints are handled either using penalty functions or via simulation. The trivial case is when f is convex and therefore a single optimum exists. However, generally the function is not convex and thus the global optimum point has to be located among many local optima.

In order to cover all possible cases (usually the real-world applications), we have also to assume that the partial derivatives of f are not calculable and a numerical approximation of them is impractical. Moreover, the analytical expression of the objective function itself may not be available and thus, there is no information about its structure and properties.

Finally, in many of the problems that are met in practice, a highly accurate solution is neither possible nor feasible; it may be impossible because of uncertainties and errors in the underlying model or data, or it may be unfeasible because of the unacceptably high cost required to attain it. Thus, the crucial issue is to find a "good" answer for the optimization problem without extremely high computational effort.

Local search techniques

Local search techniques are deterministic methods, suitable for the optimization of unimodal (single-extremum) functions. These techniques can be classified into two major categories, gradient-based and direct search methods (e.g., Pierre, 1986· Schwefel, 1994).

Gradient methods are applicable only if the analytical expression of the objective function is available and its partial derivatives are easy to compute. These are numerical algorithms that start from an initial feasible point and converge to the nearest local optimum, employing subsequent line optimizations, where they use the gradient information. The typical gradient-based methods are (Press et al., 1992):

- steepest descend methods;
- conjugate gradient methods;
- Newton methods;
- quasi-Newton or variable metric methods.

We note that Newton methods are of second order (they use first as well as second derivatives), whereas quasi-Newton methods employ numerical approximations of second derivatives.

Direct search methods are characterized by the fact that they do not compute neither approximate derivatives; the only information that they need are the values of the objective function. Instead of using derivatives, they typically consider a natural alternative, by exploring each direction of the Euclidean space in a linearly independent set of n search directions. The direct search algorithms are distinguished both by the way in which the set of the directions is chosen and by the way exploratory steps are taken (Torczon, 1991).

One of the most popular direct search methods is the *downhill simplex* algorithm of Nelder and Mead (1965). The simplex is the geometrical figure, consisting of $n + 1$ points (vertices) that span the n-dimensional space. After the definition of the initial simplex, the algorithm takes a series of iterative steps, most steps just reflecting from the worst vertex. When it can do so, the simplex is expanded to take larger steps. When it reaches a valley floor, it is contracted in the transverse direction and tries to ooze down the valley. If the simplex tries to pass through the eye of the needle, it shrinks in all directions, pulling itself around the best vertex (Figure 1). An appropriate sequence of such steps will always converge to a local minimum of the function.

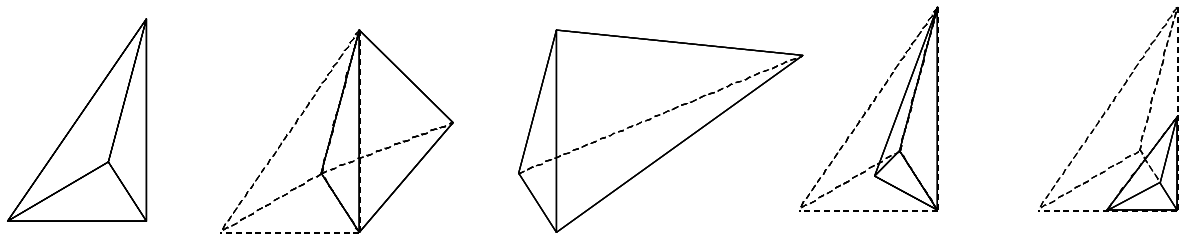

Figure 1: Possible outcomes for a step in the downhill simplex method. The simplex at the beginning of the step, here a tetrahedron, is shown left. At the end of the step, the simplex can be any one of (a) a reflection away from the worst point, (b) a reflection and expansion away from the worst point, (c) a contraction along one dimension from the worst point, or (d) a multiple contraction (shrinkage) along all dimensions towards the best point.

As mentioned before, local search methods are not appropriate for the optimization of non-convex, multimodal functions. However, because of their efficiency in simple search spaces, their principles are commonly applied in several global optimization algorithms.

Overview of global optimization techniques

Global optimization techniques aim at identifying the global optimum solution of a function that need not be convex or differentiable. These methods involve the evaluation of the function usually at a random sample of points in the feasible parameter space, followed by subsequent manipulations of the

sample using a combination of *deterministic* and *probabilistic* rules. An important issue is that they guarantee only *asymptotic convergence* to the global optimum.

Deterministic methods

The *uniform grid sampling* method is a primitive deterministic approach to the global optimization problem and consists of an examination of specific uniformly spaced values of control variables (Loucks et al., 1981). The grid search method is computationally very expensive, as the number of function evaluations increases exponentially with the dimension n. An improvement of the above technique is the application of the search procedure in the form of successive steps, with grids that are nested to each other and become progressively finer (Nalbantis and Koutsoyiannis, 1997).

Stochastic methods

The simplest probabilistic scheme is *random sampling*. In this method, a prespecified number of points are sampled at random from the feasible parameter space, using any probability distribution function (commonly a uniform distribution). The objective function value is computed at each point, and the point with the best function value is taken as an estimate of the optimum (Rubinstein, 1986).

Pure random search schemes do not use any of the information gained during sampling to direct the search in a logical manner. A better approach is to make the generation of evaluation points in a more systematic way; then the previously chosen points and function values are taken into account when the next point is chosen. These techniques are called *adaptive* because they guide the random search adaptively toward the region of the global optimum. The general strategy of adaptive methods consists of the generation of a random perturbation at the vicinity of the current solution, which is accepted only if it improves the value of the objective function; otherwise another trial point is generated.

Different adaptive search schemes depend on the choice of the generator of perturbations. The so-called *adaptive random search* methods aim to gradually reduce the "variance" of the examined solutions to implement a more detailed investigation when the global optimum solution is approached. Another category is the *controlled random search* techniques, where each perturbation is generated on the basis of previously generated points.

Price (1965, 1983) proposed several versions of the controlled random search algorithm that have become the basis of many modern global optimization methods. Here, the concept of a population of feasible solutions is introduced. At each step a simplex is formed from a sample of the population and a new trial point (i.e., the perturbation) is generated as a reflection about the centroid of the remaining points; the new trial point replaces the worst point in the original set. Krivy and Tvrdik (1995) and Brachetti at al. (1997) developed improved versions of the above algorithm.

Another stochastic approach to deal with multiple optima is to run several trials of a local search optimization algorithm from different starting points. This simple search strategy is called a *multistart* procedure. Assuming that the failure probability of a unique local search is p, then for r independent runs of the algorithm this rate reduces to p^r (Rubinstein, 1986). In an ideal case, the multistart methods aim at starting the local search procedure once in every *region of attraction* of local optima. The strategy of identifying appropriate initial solutions that belong to different regions of attraction is called *cluster* or *pattern recognition* analysis (Solomatine, 1995, 1999).

Evolutionary and genetic algorithms

The family of evolutionary algorithms is inspired from the mechanics of natural evolution of biological organisms, although these models are crude simplifications of biological reality.

Evolutionary algorithms introduce some important modifications to random search and they use the terminology from biology and genetics.

Through a typical iterative step, some points of the current population $P(t)$ are recombined according to a *recombination operator*, thus creating new points (*individuals*) that are called *offsprings*. The coordinates of some points of $P(t)$ are randomly modified according to a *mutation operator*. Then, a *selection operator* is applied to choose points for the next population $P(t + 1)$.

Historically, evolutionary algorithms have been developed is three variations:

- evolution strategies;
- evolutionary programming;
- genetic algorithms.

Schwefel (1994) and Banzhaf et al. (1998) give an overview of these approaches that differ mainly in the types of recombination, mutation and selection operators. Among them, the last have received the most attention in continuous global optimization.

In *genetic algorithms* (Holland, 1975· Goldberg, 1989· Michalewich, 1992) the control variables are represented on a chromosome-like (usually binary string) structure. A fitness value is assigned to each individual, expressing the quality measure of the corresponding solution. The selection operator is applied by choosing the fittest individual strings to be recombined in order to produce better offsprings. A probabilistic mechanism (i.e., a roulette wheel) is used, allocating greater survival to best individuals. The crossover (recombination) operation implements the exchange of genes of randomly selected pairs of individuals with a certain probability. Finally, according to the mutation procedure, some of the genes in the chromosomes are randomly changed with a certain (small) probability, thus keeping the population diverse and preventing form premature convergence onto a local optimum.

Simulated annealing

The method of simulated annealing is based on an analogy with a thermodynamical process called annealing. For *slowly* cooled thermodynamical systems (metals or liquids), nature is able to find the minimum energy state, while the system may end in an amorphous state having a higher energy if it is cooled quickly. The principle can be expressed by the Boltzmann probability distribution function:

$$p(E) \sim \exp\left(-\frac{E}{\kappa T}\right) \tag{2}$$

where κ is a constant. For a system in thermal equilibrium at a given temperature T, its energy is probabilistically distributed among all different energy states E. Therefore the system may switch to a new energy state, irrespective of whether it is higher or lower. In other words, nature's minimization strategy is to allow the system sometimes to go uphill as well as downhill, so that it has a chance to escape from a local energy minimum in favor of finding a better, more global one. Note that the lower the temperature, the less likely is any significant uphill step.

Metropolis et al. (1953) transferred the principles of the annealing process into statistical mechanics. Their concept was to use a probabilistic criterion, analogue to the Boltzmann function, in order to allow some uphill steps, whereas all downhill steps should be acceptable. To apply the above concept into other than thermodynamical systems, one should provide the following elements (Press et al., 1992):

- A description of possible system configurations.
- A generator of random changes in the configuration of the system.
- An objective function (analogue of energy) to be minimized.

- A control parameter T (analogue of temperature) and an *annealing cooling schedule*, which describes the gradual reduction of T.

Simulated annealing has been first applied with quite a lot of success in combinatorial optimization (Kirkpatrick et al., 1983). Yet, there are still many difficulties in the application of simulated annealing to problems with continuous parameter space. The most problematical element is the random generator of possible moves, which is extremely inefficient narrow valleys, leading almost always in uphill moves. To cope with this handicap, several heuristic schemes have been proposed (e.g., Corana et al., 1987· Goffe at al., 1994· Cardoso et al., 1996).

The shuffled complex evolution method

The shuffled complex evolution method (Duan et al., 1992) is a new, heuristic global optimization scheme that combines the strength of the *downhill simplex* procedure of Nelder and Mead (1965) with the concepts of *controlled random search* (Price, 1965), *competitive evolution* (Holland, 1975) and *complex shuffling*. This method has become the most popular among hydrologists and many references about it can be found in the water resources literature (Duan et al. 1994a, Gan and Biftu, 1996· Cooper et al., 1997· Kuczera, 1997· Yapo et al., 1998· Freedman et al., 1998· Thyer et al., 1999).

The algorithm begins by randomly selecting a population of feasible points that are sorted and partitioned into a number of communities (*complexes*), each one containing at least $n + 1$ points. Each of the complexes is allowed to evolve in the direction of global improvement, using competitive evolution techniques that are based on the downhill simplex method. At periodic stages in the evolution, the entire set of points is shuffled and reassigned to new complexes to enable information sharing. This process is repeated until some stopping criteria are satisfied.

Duan et al. (1992) compared the global search without sharing information to giving a number of competent persons a difficult problem to solve without conferring with each other. It is obvious that it is much better for the people to first work independently (individually or in small groups) and then get together periodically to share information about their progress. The combination of competitive evolution and complex shuffling ensures that the information gained by each of the individual complexes is shared throughout the entire population. This results in a robust optimization algorithm that conducts an efficient search of the parameter space.

Tabu search

The natural system on which tabu search is based is the human memory process. The modern form of tabu search derives from Glover (1986). The basic principle is to maintain a so-called *tabu list* of recent moves (transitions from point to point) in order to prevent the search from moving back to where it was previously. The basic steps of the algorithm consist of starting from a feasible point and then moving to the best neighbor. This is similar to a hill-climbing search, except for the fact that it may move to a worse solution from the current one.

Although tabu search has been successfully applied to a variety of combinatorial problems, very few works deal with its application to continuous optimization. Among them, we can distinguish the works of Al-Sultan and Al-Fawzan (1997), Siarry and Berthiau (1997) and Chelouah and Siarry (2000).

The evolutionary annealing-simplex scheme

Literature review

An interesting category of global optimization techniques is this that aims to combine the robustness of simulated annealing in rugged problems with the efficiency of hill-climbing methods in simple search spaces. The strengths and weaknesses of the two approaches are complementary. Simulated annealing avoids local optima by jumping away from them, but it sacrifices efficiency (i.e., computational effort) by doing so; on the other hand, hill-climbing methods converge quickly to the nearest local optimum, but they have no way of getting out of it.

One typical local optimizer is the downhill simplex algorithm of Nelder and Mead (1965). However, there are only few references in literature on how to incorporate the simulated annealing strategy to this method. Press et al. (1992) introduced a logarithmically distributed random variable proportional to the temperature into the objective function associated with every vertex of the simplex. In this way, the simplex behaves between a random walk and an ordered downhill motion, depending on the ratio $\Delta f / T$, where Δf is the difference in values of the function at two vertices and T is the temperature. On the basis of the above scheme, Pan and Wu (1998) introduced some improvements; the most important is the use of some follow-up strategies to escape from local optima. Kvaniscka and Pospichal (1997) proposed a different scheme, based on the controlled random search method of Price (1965), where the construction of reflection points is randomized and their returning to the population is implemented according to a probabilistic criterion. A parallel version of the above algorithm uses a decomposition of the whole population into disjoint subpopulations, for witch independent simulated annealings are done. During the evolution, the subpopulations randomly interact so that between two subpopulations their best points found so far are exchanged, whereas the worst ones are eliminated.

Description of the algorithm

The evolutionary annealing-simplex algorithm, which was developed within the framework of this postgraduate thesis, uses ideas from several methodological approaches, enhancing them with some original elements. The main concept is based on a *controlled random search* optimization scheme, where a generalized *downhill simplex* methodology is coupled with a *simulated annealing* procedure.

A brief description of the algorithm is the following: An initial population P is randomly generated into the feasible space. At each step a simplex is formulated, by randomly choosing a set of $n + 1$ points from P. The simplex is reflected from a randomized "worst" vertex $\mathbf{x}_w$ (this can be any vertex except of the current best). If either the reflection point $\mathbf{x}_r$ is not accepted according to a probabilistic criterion or $f(\mathbf{x}_r) < f(\mathbf{x}_w)$, the simplex is moved downhill following a generalized Nelder-Mead strategy, where the lengths of all types of movements (expansions and contractions) are randomized. If $\mathbf{x}_r$ is accepted albeit being worse than $\mathbf{x}_w$, trial expansion steps are taken along the uphill direction in order to jump from the local optimum, pass the ridge and explore the neighboring area. If any trial step successes, a random point is generated on the boundaries of the population and replaces $\mathbf{x}_r$ according to a mutation probability p_m. The procedure is repeated until one of the termination criteria is satisfied.

The above algorithm contains several original issues referring to:

- the generation of the initial population;
- the automatized regulation of the annealing schedule;
- the generalization of the Nelder-Mead procedure, in order to be more competitive and stochastic;
- the implementation of a simple line minimization procedure to accelerate the downhill search;
- the use of a heuristic follow-up strategy to escape from local optima;
- the definition of the mutation function;

- the introduction of a reannealing capability, in case of early convergence.

Evaluation and comparison of optimization algorithms

Two factors, the *effectiveness* and the *efficiency* characterize the performance of any optimization algorithm (Schwefel, 1994). The first one indicates the probability of locating the global optimum starting from any random initial solution (or population of solutions) whereas the second one indicates the speed of the algorithm. A measure of the effectiveness of an algorithm in a specified problem is the number of successes out of a predefined number of independent runs and a measure of its efficiency is the average number of function evaluations that are needed to converge to the optimum.

Mathematical applications

Four representative methodologies have been tested on a set of several optimization problems, including traditional benchmark functions used in optimization literature. These methodologies were:

- a multistart downhill simplex procedure of Nelder and Mead (1965);
- a simple, binary-coded genetic algorithm, based on the source code of Goldberg (1989);
- the shuffled complex evolution method of Duan et al. (1992);
- the original evolutionary annealing-simplex scheme.

For each of the above algorithms, a detailed analysis has been implemented, in order to investigate their performance in relation with some critical input arguments, like the population size. In Table 1, the results obtained for the best only set of these arguments are presented.

Table 1: Effectiveness and efficiency (in parenthesis) indices for the global optimization algorithms.

Test function	*n*	Number of optima	Multistart simplex [1]	Genetic algorithm	SCE-UA	Annealing-simplex
Sphere	2	1	100 (4266)	100 (45463)	100 (5159)	100 (4831)
Hozaki	2	2	99 (137687)	81 (26731)	100 (296)	100 (303)
Goldestein-Price	2	4	100 (91557)	96 (26731)	99 (449)	100 (419)
Rozenbrock	2	1	100 (183588)	65 (27374)	100 (1191)	100 (583)
Rozenbrock	10	1	7 (7532)	0 (45463)	99 (11105)	33 (12635)
Griewank	10	> 1000	100 (10524)	89 (52853)	100 (5574)	99 (7567)
Michalewicz	2	> 100	35 (210534)	31 (27048)	44 (438)	58 (1373)
Step	10	1	17 (528120)	4 (45463)	1 (2350)	78 (6691)
Mean effectiveness			69.8	58.3	80.0	83.5

(1) Results obtained after 20 independent runs of the downhill simplex algorithm.

The main conclusions of the analysis were the following:

- The effectiveness of the multistart technique that implements the downhill simplex scheme is proportional to the number of independent runs of the local search algorithm; the consequence is a low efficiency of the method, especially for difficult optimization problems.
- The performance of the simple, binary-coded genetic algorithm was not satisfactory, not only due to the low value of the effectiveness index but mainly because of the extremely large number of function evaluations needed for convergence.
- Both the shuffled complex evolution method and the evolutionary annealing-simplex scheme faced with success almost all optimization problems; the latter was proved slightly more effective but slightly more time-consuming.

- All methods were sensitive to the corresponding algorithmic input arguments.

Real-world applications

In real-world optimization problems, the properties of the response surface of the objective function as well as the location of the global optimum are not known a priori. Moreover, due to the effort needed for the computation of function values, it is likely to terminate the optimization before convergence criteria are satisfied, in order not to exceed a maximum number of function evaluations.

Three global optimization problems were examined, taken from the domain of water resources engineering and management. The algorithms that were used to solve the problems were those who demonstrated the best performance through the test functions-based analysis, namely the shuffled complex evolution method and the evolutionary annealing-simplex method.

The first problem was the calibration of a mathematical model, more precisely a simple water balance model. In such an application, the objective is to identify the values of the model parameters by minimizing the departure between the computed output variables and the outputs measured in the physical system. The water balance model was developed by Mantoudi (2000) and applied to the river basin of Kremasta. The number of parameters was 4; specifically the imperviousness of the basin, the storage capacity of the soil moisture reservoir and the recession coefficient of the soil moisture and groundwater. In addition, two other control variables were assumed; namely the initial values of soil and ground storage. Input data was the precipitation and the potential evapotranspiration for a simulation period of 93 months. Two case studies were examined; the first one was based on real output data (i.e., historical runoff series) whereas the second one was based on synthetic output data, where the runoff was calculated assuming arbitrary parameter values. The reason of using synthetic data was the elimination of errors due to the data as well as the model structure. The results are shown in Figures 2 and 3, respectively. For both problems, the annealing-simplex method managed to achieve the same effectiveness rate (16%), which was better than the rate achieved via the shuffled complex evolution algorithm (20% and 28%, respectively).

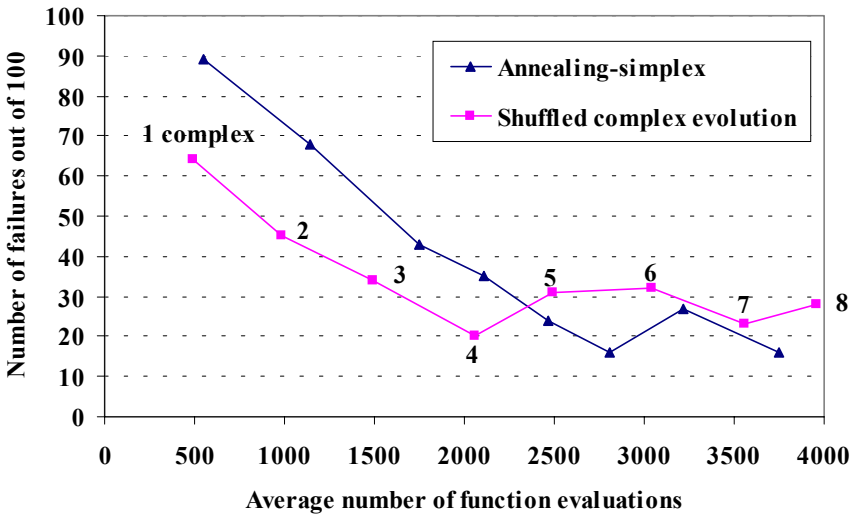

Figure 2: Comparison of the performance of the optimization algorithms used for the calibration of the simple water balance model with real runoff data.

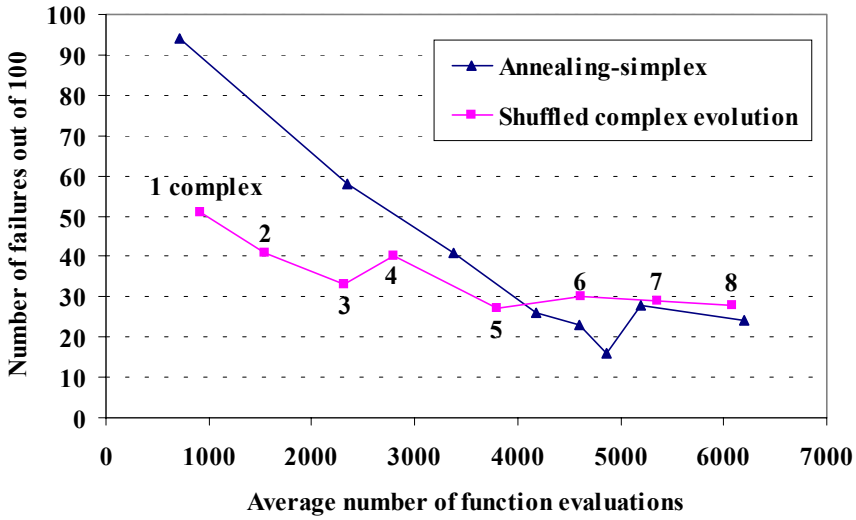

Figure 3: Comparison of the performance of the optimization algorithms used for the calibration of the simple water balance model with synthetic runoff data.

The second application was the maximization of the mean annual energy profit of a hypothetical hydrosystem, consisting of two reservoirs in parallel. A high-dimensional methodology was implemented, assuming as control variables the step-by-step reservoir target releases. Thus, for a simulation period of 16 years, the total number of control variables was 384. The results were 44.7 monetary unions for the shuffled complex evolution method and 46.5 unions for the evolutionary annealing-simplex algorithm. Note that Economou (2000), who examined the same problem assuming a low-dimensional methodology where parametric operation rules (Nalbantis and Koutsoyiannis, 1997) were used, achieved a performance index of 49.9 unions. The main characteristic of this problem was the flat-type response surface of the objective function, which is due to the use of desirable and not real magnitudes as control variables. This feature makes extremely difficult the location of the gradient of the function, especially when the number of parameters is large.

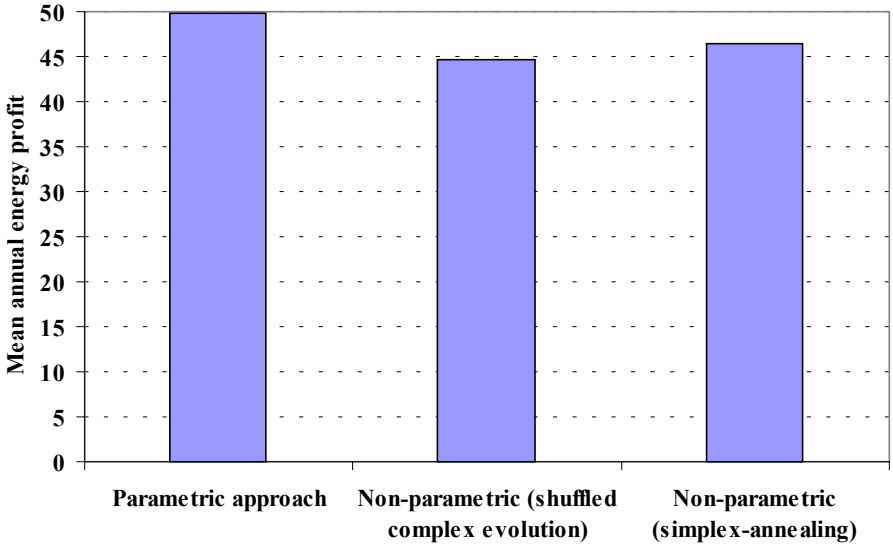

Figure 4: Results for the energy profit maximization problem.

The third application was an algebraic problem, typical in stochastic hydrology applications. The objective is the calculation of a matrix **b** such that:

$$\mathbf{c} = \mathbf{b}\,\mathbf{b}^{T} \tag{3}$$

The above problem, also known as square matrix decomposition, is met in all linear multivariate stochastic models of the form:

$$\mathbf{Y} = \mathbf{a}\,\mathbf{Z} + \mathbf{b}\,\mathbf{V} \tag{4}$$

where **Y** is a vector of variables to be generated, **Z** is a vector of known variables, **V** is a vector of noise variables (innovations) and **a**, **b** are matrices containing model parameters. In that case, **c** is the covariance matrix of the model with known elements. The problem has infinite number of solutions if **c** is positive definite, otherwise there are no feasible solutions. Moreover, the skewness coefficients of noise variables **V** depend on **b**; i.e., $\mu_3[\mathbf{V}] = \xi(\mathbf{b})$. If some element of the vector ξ is too high, $\mu_3[\mathbf{V}]$ cannot be preserved. Therefore, even if a feasible matrix **b** can be found, this may not be appropriate for the preservation of all model statistics. Koutsoyiannis (1999) proved that this can be formulated as a nonlinear unconstrained optimization problem, where the objective is to minimize the departures $\mathbf{d} = \|\mathbf{c} - \mathbf{b}\,\mathbf{b}^{T}\|$, by keeping simultaneously the skewness of **V** as small as possible. He also proved that exists an analytical expression of the partial derivatives of the objective function and therefore any gradient-based method can be used to solve the problem.

The proposed methodology was implemented for the multivariate generation of monthly rainfall and runoff series at the four basins (Mornos, Evinos, Yliki, Marathon) of the hydrosystem of Athens, using a periodic AR(1) model. Twelve independent optimization problems were examined, each one referring to the decomposition of the corresponding monthly covariance matrix. The number of control variables, namely the elements of matrices **b**, were $8 \times 8 = 64$. In addition to the shuffled complex evolution method and the evolutionary annealing-simplex algorithm, a multistart conjugate gradient scheme was applied, in order to compare their performance on a basis of a fast and accurate derivative-based method. The results are presented in Figure 5.

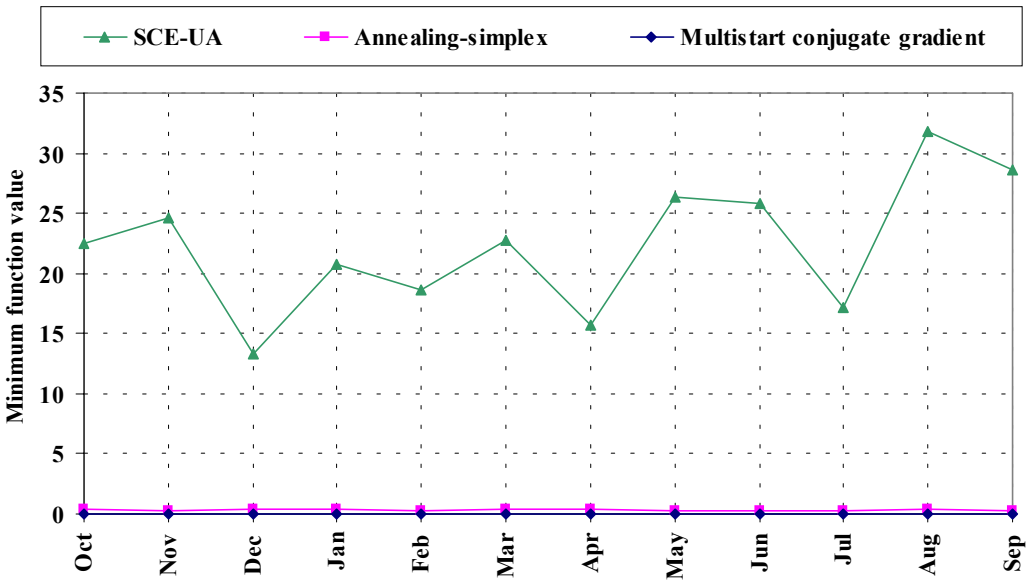

Figure 5: Comparison of the performance of the optimization algorithms used for the decomposition of covariance matrices of the PAR(1) model.

As expected, the multistart conjugate gradient method managed to find the best function values with a very little computational effort (about 2 minutes of CPU time), whereas the other two methods needed much more time (about 30 minutes of CPU time) to converge. In that case too, the evolutionary annealing-simplex algorithm had an obviously better performance. An important characteristic of this

problem was the sensitivity of the objective function with small changes of parameters' values and the existence of many local optima. The historical sample statistics were almost perfectly preserved by using the conjugate gradient method and they were also well approximated, by using the evolutionary annealing-simplex algorithm. On the other hand, the shuffled complex evolution method was proved unable to preserve those statistics.

Summary and conclusions

Through this postgraduate thesis, an integrated literature review of global optimization techniques was implemented and an original optimization scheme, the evolutionary annealing-simplex algorithm, was developed. Moreover, the performance of several methodologies was evaluated on the basis of mathematical and real-world optimization problems.

The most important conclusions from the research were the following:

- The current trend in global optimization research is the combination of strategies obtained from diverse methodological approaches (including classical mathematics), in order to develop more robust search schemes.
- The simple, binary-coded genetic algorithm, which was used only in mathematical test functions, achieved a relatively low effectiveness rate and, the most important, needed an extremely large number of computations in order to locate the optimal solution.
- On the other hand, both the shuffled complex evolution method and the evolutionary annealing-simplex scheme were proved robust and efficient, and managed to solve almost all optimization problems that were examined. However, we have to note that the annealing-simplex scheme was proved more effective in real-world applications, especially in the matrix decomposition problem.
- Through the analysis, two types of objective functions were proved the most difficult to optimize; either rough, flat-surface functions or multimodal functions, sensitive to small changes of the control variables values. For both cases, the evolutionary annealing-simplex scheme was proved more effective than the shuffled complex evolution method.
- The performance of all optimization methods depended, less or more, on some critical algorithmic input parameters, usually calibrated experimentally.
- A final conclusion is that in spite of the development of robust and fast optimization schemes, parsimony of parameters still remains a significant requirement of mathematical modeling.

Focusing on the evolutionary annealing-simplex scheme, we can propose some ideas for further development. The first one is a parallelized version of the algorithm, which can reduce significantly the time of computations. The second one is the implementation of a shuffling procedure (analogue to the procedure used in the shuffled complex evolution method), which will enable the sharing of the information gained during the optimization process. Another aspect is the incorporation of an automatic procedure for the setting of input arguments, like the population size and the annealing schedule coefficients, in order to ameliorate the performance of the algorithm and reduce user's interventions. A last point is a generalization of the evolution rules, in order to incorporate either line optimization or gradient-based methodologies, when function derivatives are available.