Global optimisation techniques in water resources management

Andreas Efstratiadis and Demetris Koutsoyiannis
Department of Water Resources
National Technical University of Athens
Parts of the presentation

- Water resources engineering and global optimization
- The nonlinear unconstrained optimization problem
- An overview of nonlinear optimization techniques
  - The downhill simplex method
  - Genetic algorithms
  - The shuffled complex evolution method
  - An evolutionary annealing-simplex algorithm
- Evaluation of global optimisation algorithms
  - Application in mathematical problems
  - Application in real-world problems
- Conclusive remarks
Global optimisation and water resources – Example 1: Hydrological models calibration

Input data:
- Precipitation
- Temperature (average, minimum, maximum)
- Potential evapotranspiration
- Geomorphology and land uses

Output data:
- Soil and ground storage
- Real evapotranspiration
- Total runoff

Problem formulation:
Identify the values of model parameters by minimising the deviation between the computed output variables and the outputs measured in the physical system.
Global optimisation and water resources –
Example 2: Hydrosystems control and management

Hydrosystem components:
- Surface and groundwater resources.
- Water conveyance systems.
- Power generators (hydroelectric stations) or consummators (pumps).

Mathematical model inputs:
- Network topology.
- Physical constraints, due to project attributes (e.g., storage capacity).
- Inflow data (historic, synthetic).
- Targets, specified from the manager of the hydrosystem (e.g., water supply, flood control, environmental protection, power production).

Problem formulation:
Maximise the performance measure of the system (expressed in terms of yield, reliability or profit), which is evaluated via stochastic simulation.
Global optimisation and water resources – Example 3: Parameters estimation of stochastic models

- In linear multivariate stochastic models of the form:
  \[ Y = a Z + b V \]
  the problem of estimating the elements of parameter matrix \( b \) arises.
- Matrix \( b \) is given from an equation of the form:
  \[ c = b b^T \]
  where \( c \) is a covariance matrix.
- If \( c \) is positive definite, the problem has infinite solutions; otherwise there are no feasible solutions (inconsistent \( c \)).
- The skewness of noise variables \( V \) depends on \( b \), i.e. \( \mu_3[V] = \xi(b) \). If some element of \( \xi \) is too high, \( \mu_3[V] \) cannot be preserved.

**Problem formulation:**
Determine \( b \) from the known \( c = b b^T \) so that the coefficients of skewness of noise variables \( V \) be as small as possible.

*For more details about decomposition of covariance matrices see: Koutsoyiannis, 1999*
The nonlinear unconstrained optimization problem

Find an optimiser \( \mathbf{x}^* \) such that:

\[
\hat{f}(\mathbf{x}^*) = \min f(\mathbf{x}), \quad a < \mathbf{x} < b
\]

**Main assumptions:**
- The control variables are continuous and bounded.
- All constraints are handled either using penalty functions or via simulation.

**Typical handicaps:**
- Due to non-convexity, \( f \) may have many local optima.
- The partial derivatives of \( f \) may not be calculable and a numerical approximation of them is usually impractical.
- An analytical expression of \( f \) may not be available.
- The evaluation of \( f \) may be very expensive or time-consuming.

In real-world applications, a highly accurate solution is neither *possible* (due to uncertainties and errors in the underlying model or data) nor *feasible* (because of the unacceptably high computational effort).
Deterministic local optimisation methods:
- Gradient methods (e.g., steepest descend, conjugate gradient, quasi-Newton or variable metric methods).
- Direct search methods (e.g., downhill simplex, rotating directions).

Global optimization methods:
- Set covering techniques.
- Pure random search.
- Adaptive & controlled random search.
- Multiple local search.
- Evolutionary & genetic algorithms.
- Simulated annealing.
- Tabu search.
- Combined algorithms (e.g., shuffled complex evolution, simplex-annealing).

Global optimisation algorithms 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. They guarantee asymptotic convergence to the global optimum.
The downhill simplex method  
(Nelder and Mead, 1965)

Description of the algorithm:

- A set of $n + 1$ points (a simplex) is generated in the $n$-dimensional space.
- At each iteration, the simplex is reflected from the worst vertex.
- When it can do so, the simplex is expanded to take larger steps.
- When the simplex 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.

Although the downhill simplex is not a global optimisation method, its principles are commonly applied in several global optimisation algorithms.
Genetic algorithms

Main concepts:

• Inspired from the process of natural selection of biological organisms.
• Representation of control variables on a chromosome-like (usually binary string) structure.
• Search through a population of points (individuals), not a single point.
• A fitness value is assigned to each solution, expressing its quality measure.
• Genetic operators are applied in order to create new generations.

Genetic operators:

• **Selection**: Chooses 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.

• **Crossover**: Recombines (exchanges) genes of randomly selected pairs of individuals with a certain probability.

• **Mutation**: Randomly changes genes in the chromosomes with a certain (small) probability, thus keeping the population diverse and preventing form premature convergence onto a local optimum.
The shuffled complex evolution method
(Duan et al., 1992)

Main concepts:

- Combination of probabilistic and deterministic approaches.
- Systematic evolution of a complex of points spanning the parameter space.
- Competitive evolution.
- Complex shuffling.

Description of the algorithm:

- A random set of points (a “population”) is sampled and partitioned into a number of complexes.
- 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.
Simulated annealing

Principles of the annealing process in thermodynamics:

- For *slowly cooled* thermodynamical systems (e.g., metals) 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.
- Nature’s minimisation 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 minimum.
- For a system at a given temperature $T$, its energy is probabilistically distributed among all energy states $E$ according to the Boltzmann function:
  \[
  \text{Prob}(E) \sim \exp\left(-\frac{E}{kT}\right)
  \]
  - The lower the temperature, the less likely is any significant uphill step.

Necessary components of a simulated annealing algorithm:

- A generator of random changes in the configuration of the system.
- An objective function (analogue of energy) to be minimised.
- A control parameter $T$ (analogue of temperature) and an annealing cooling schedule, which describes the gradual reduction of $T$. 

An evolutionary annealing-simplex algorithm

Main concepts:

- Combination of the robustness of simulated annealing in rugged problems with the efficiency of local optimisation methods in simple search spaces.
- Generalisation of the simplex method to be competitive and stochastic.
- Introduction of follow-up strategies to escape from local optima.

Description of the algorithm:

- An initial population $P$ is randomly generated into the feasible space.
- At each iteration a simplex is formulated, by choosing $n + 1$ points from $P$.
- The simplex is reflected from a randomised “worst” vertex $x_w$.
- If the reflection point $x_r$ is either not accepted or $f(x_r) < f(x_w)$, the simplex is moved downhill according to the Nelder-Mead criteria performing randomised expansion, contraction or shrinkage steps.
- If $x_r$ is accepted albeit being worse than $x_w$, trial expansion steps are taken along the uphill direction in order to “climb” the hill and explore the neighboring area. If any trial step success, a random point is generated far from the population and replaces $x_r$ according to a mutation probability.
Evaluation and comparison of optimisation methods

General methodology:

- Multiple runs of each problem, starting from stochastically independent initial conditions (e.g., different initial population).
- Evaluation of the *effectiveness* (i.e., probability of locating the global optimum) and *efficiency* (i.e., convergence speed) of each algorithm.

Differences between real-world and mathematical applications:

- The properties of the response surface as well as the citation of the global optimum are not known a priori.
- Due to the computational effort for each function evaluation, it is likely to stop the optimisation procedure before convergence criteria are satisfied.

Algorithms examined:

- Downhill simplex (source code adapted from Press et al., 1992).
- Simple genetic algorithm (source code adapted from Goldberg, 1989).
- Shuffled complex evolution (source code adapted from Duan et al., 1994).
- Evolutionary annealing-simplex (original code).
## Mathematical applications

<table>
<thead>
<tr>
<th>Function name</th>
<th>$n$</th>
<th>Number of optima</th>
<th>Downhill simplex</th>
<th>Genetic algorithm</th>
<th>SCE-UA</th>
<th>Annealing-simplex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sphere</td>
<td>10</td>
<td>1</td>
<td>93 (212)</td>
<td>100 (45463)</td>
<td>100 (5159)</td>
<td>100 (4128)</td>
</tr>
<tr>
<td>Hozaki</td>
<td>2</td>
<td>2</td>
<td>4 (18205)</td>
<td>81 (26731)</td>
<td>100 (296)</td>
<td>100 (324)</td>
</tr>
<tr>
<td>Goldestein-Price</td>
<td>2</td>
<td>4</td>
<td>49 (5028)</td>
<td>96 (26731)</td>
<td>99 (449)</td>
<td>100 (552)</td>
</tr>
<tr>
<td>Rozenbrock</td>
<td>2</td>
<td>1</td>
<td>85 (6560)</td>
<td>65 (27374)</td>
<td>100 (1191)</td>
<td>100 (619)</td>
</tr>
<tr>
<td>Rozenbrock</td>
<td>10</td>
<td>1</td>
<td>0 (372)</td>
<td>0 (45463)</td>
<td>99 (11105)</td>
<td>26 (10847)</td>
</tr>
<tr>
<td>Griewank</td>
<td>10</td>
<td>&gt;1000</td>
<td>73 (603)</td>
<td>89 (52853)</td>
<td>100 (5574)</td>
<td>91 (2768)</td>
</tr>
<tr>
<td>Michalewicz</td>
<td>2</td>
<td>unknown</td>
<td>2 (27518)</td>
<td>31 (27048)</td>
<td>44 (438)</td>
<td>51 (1409)</td>
</tr>
<tr>
<td>Integer step</td>
<td>10</td>
<td>1</td>
<td>0 (48011)</td>
<td>4 (45463)</td>
<td>1 (2350)</td>
<td>100 (3324)</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td></td>
<td></td>
<td>38.3</td>
<td>58.3</td>
<td>80.0</td>
<td>83.5</td>
</tr>
</tbody>
</table>

Effectiveness index (number of successes out of 100 stochastically independent trials)  
Efficiency index (average number of function evaluations required)
Real-world applications (1)

Calibration of a simple water balance model

- Model parameters (6 in total):
  - percentage of imperviousness surface
  - storage capacity of soil moisture reservoir
  - recession coefficient of soil moisture and ground water
  - initial values of soil and ground storage


Even for the simple hydrologic model with only 6 parameters, a large number of function evaluations is required in order to achieve a relatively acceptable failure rate.
Real-world applications (2)

Maximisation of benefit from energy production

- The system consists of two hypothetical parallel hydroelectric reservoirs.
- A 16-year (192 months) synthetic inflow data is used.
- Reservoir target releases are assumed as control variables of the model; thus the total number of parameters is $2 \times 192 = 384$.
- Results are compared to those of a low-dimensional methodology, where parametric operation rules are used (Nalbantis and Koutsoyiannis, 1997).
Real-world applications (3)

Decomposition of covariance matrices of a PAR(1) model

- The objective function consists of 3 components, for the preservation (or approximation) of covariances, variances and skewness, respectively.
- The derivative of the function has an analytical expression, thus enabling the usage of a fast and accurate gradient-based optimisation method.
- Model parameters are $8^2 = 64$ (monthly rainfall and runoff in 4 locations).

Almost exact preservation of historical sample statistics.
Satisfactory approximation of historical statistics.
Unable to preserve any of historical sample statistics.
Concluding remarks

- The current trend in global optimisation research is the combination of strategies obtained from diverse methodological approaches (including classical mathematics), in order to develop more robust search schemes.

- After comparing three representative algorithms, the main conclusions are:
  
  ✓ In most cases, the simple, binary-coded genetic algorithm is neither effective nor efficient enough.

  ✓ The shuffled complex evolution method is obviously robust and efficient and should be preferred, especially when the fast location of a good solution is desired.

  ✓ The evolutionary annealing-simplex scheme, although not very fast, seems to be the most appropriate for hard optimisation problems with pathogenic characteristics (e.g., many local optima).

- The performance of all methods depends, less or more, on the algorithmic input arguments (e.g., population size), usually calibrated experimentally.

- In spite of the development of robust and fast optimisation techniques, the parsimony of parameters still remains a significant requirement of the mathematical models building.
References


• Duan, Q., S. Sorooshian, and V. Gupta, Distribution diskette for the shuffled complex evolution (SCE-UA) method, 1994.


