Shuffled Complex Evolution Model Calibrating Algorithm: Enhancing its Robustness and Efficiency

Muttil, Nitin and Jayawardena, Arlene (2008) Shuffled Complex Evolution Model Calibrating Algorithm: Enhancing its Robustness and Efficiency. Hydrological Processes, 22 (23). pp. 4628-4638. ISSN 0885-6087

Hydrological Processes, Vol. 22, Issue 23, pp. 4628-4638 (2008)
The published article can be downloaded from: http://dx.doi.org/10.1002/hyp.7082

Shuffled Complex Evolution Model Calibrating Algorithm: Enhancing its Robustness and Efficiency

Nitin Muttil[1] and A. W. Jayawardena[2]

[1] School of Architectural, Civil and Mechanical Engineering and Institute for Sustainability and Innovation, Victoria University, PO Box 14428, Melbourne, VIC, Australia, 8001.
[Corresponding author] Email: nitin.muttil@vu.edu.au

[2] International Centre for Water Hazard and Risk Management, Public Works Research Institute, 1-6, Minamihara, Tsukuba, Ibaraki 305-8516, Japan.
(Formally from the Department of Civil Engineering, The University of Hong Kong)

Abstract

The Shuffled Complex Evolution (SCE-UA) has been used extensively and proved to be a robust and efficient global optimization method for the calibration of conceptual models. In this paper, we propose two enhancements to the SCE-UA algorithm, one to improve its exploration and another to improve its exploitation capability of the search space. A strategically located initial population is used to improve the exploration capability and a modification to the downhill simplex search method enhances its exploitation capability. This enhanced version of SCE-UA is tested, first on a suite of test functions and then on a conceptual rainfall-runoff model using synthetically generated runoff values. It is observed that the strategically located initial population drastically reduces the number of failures and the modified simplex search also leads to a significant reduction in the number of function evaluations to reach the global optimum, when compared to the original SCE-UA. Thus, the two enhancements significantly improve the robustness and efficiency of the SCE-UA model calibrating algorithm.

Keywords: Evolutionary computation; Optimization; Calibration; Hydrologic models.

Introduction

With the advent of digital computers, a generation of models known as conceptual models has been developed. The successful application of a model heavily depends on how well it is calibrated. There is a substantial body of research documenting problems encountered during model calibration, especially with conceptual models (Duan et al., 1992; Gan and Biftu, 1996; Kuczera, 1997). Duan et al., (1992) pointed out five characteristics that complicate the optimization of conceptual models, which are as below:

i) Several major regions of attraction

ii) Numerous local optima in each region of attraction

iii) Rough objective function surface with discontinuous derivatives

iv) Parameters may exhibit varying degree of sensitivity and a great deal of interaction, which may be non-linear

v) Response surface is often non-convex with long curved ridges

The most important of these characteristics is the presence of multiple local optima.

To deal with the problem of multiple local minima, global search methods are applied. These methods are global in the sense that they constitute a parallel search of the search space (as opposed to a point by point search) by using a population of potential solutions. This capability of such techniques for effective "exploration" of the search space makes them less probable to get trapped into local minima. Popular global search methods are the so-called population-evolution-based search strategies such as the Shuffled Complex Evolution-University of Arizona (SCE-UA) (Duan et al., 1992) and Genetic Algorithms (GA) (Wang, 1991). This study aims to significantly improve the robustness and efficiency of the SCE-UA algorithm. The SCE-UA has been extensively used for the calibration of various rainfall-runoff models. These include the NAM/MIKE 11 model (Madsen, 2000), the Sacramento model (Sorooshian et al., 1993; Duan et al., 1994; Gan and Biftu, 1996; Yapo et al., 1996;

Gan et al., 1997; Ajami et al., 2004), the Tank model (Tanakamaru and Burges, 1996; Cooper et al., 1997), and the Xinanjiang model (Gan and Biftu, 1996; Gan et al., 1997; Cheng et al., 2002; Cheng et al., 2006; Jayawardena et al., 2006). Other than for the calibration of hydrologic models, the SCE-UA has also been used in other water related fields. Atiquzzaman and Liong (2004) employed the SCE-UA for the rehabilitation of water distribution networks, whereas Cui and Kuczera (2003) used it for optimizing urban water supply headworks systems. For calibration of groundwater models, He et al. (2007) used the SCE-UA to calibrate a groundwater prediction model for a coastal plain in Japan, Contractor and Jenson (2000) used it to calibrate an unsaturated flow model, whereas Agyei and Hatfield (2006) tested the SCE-UA on several inverse modeling problems involving the estimation of parameters in a coupled numerical groundwater flow and contaminant transport model.

A brief introduction to the SCE-UA algorithm is first presented in the next section. This is followed by the two proposed enhancements to the original SCE-UA. A comparison of the enhanced SCE-UA with its original counterpart on a suite of popular test functions and also a conceptual rainfall-runoff model is then presented.

The Shuffled Complex Evolution Algorithm

The SCE-UA algorithm was developed at the University of Arizona (Duan et al., 1992) to deal with the peculiarities of parameter estimation in conceptual rainfall-runoff (CRR) models. It combines the best features of "multiple complex shuffling" and "competitive evolution" based on the simplex search method (Nelder and Mead, 1965). The use of multiple complexes and their periodic shuffling provide an effective exploration of different promising regions of attraction within the search space. In other words, the partitioning into and shuffling of complexes facilitates a freer and extensive exploration of the search space in different directions, thereby reducing the chances of the search getting trapped in local optima. This effective exploration is coupled with the competitive complex evolution (CCE) algorithm, which is provided by the downhill simplex algorithm. The simplex search provides a robust technique for evolving each complex independently and thus directs the evolution in an improvement direction. This "competitive evolution" of the simplexes provides effective exploitation within the search space. Thus, the "multiple complex shuffling" and "competitive evolution" features provide an effective balance of exploration and exploitation in the SCE-UA as compared with other contemporary optimization strategies.

In essence, the SCE-UA begins with an "initial population" of points sampled randomly from the feasible space. The population is partitioned into one or more complexes, each containing a fixed number of points. Each complex evolves based on a statistical "reproduction" process that uses the "simplex" geometric shape to direct the search in the correct direction. Periodically, the entire population is shuffled and points are reassigned to complexes to ensure information sharing. As the search progresses, the entire population tends to converge toward the neighborhood of the global optimum, provided the initial population size is sufficiently large. For a lucid explanation on the details of the algorithm, the reader is referred to Duan et al. (1994).

A number of studies have been conducted to compare the SCE-UA and other global and local search procedures for model calibration (Duan et al., 1992; Gan and Biftu, 1996; Kuczera, 1997; Franchini et al., 1998). Duan et al. (1993) compared the SCE-UA method with controlled random search (CRS2) method and a multi start algorithm (MSX) on seven well established test functions from the literature and demonstrated the superiority of the SCE-UA method. Sorooshian et al. (1993) used a "true" parameter set for the Leaf River watershed to generate "synthetic" streamflows to calibrate the Sacramento model with several optimization algorithms, including the SCE-UA and the multi-start simplex (MSX) methods. They found that the SCE-UA located the "true" parameter values at a 100% success rate while the MSX failed to locate the global optimum in all trials. Cooper et al. (1997) investigated the performance of three probabilistic optimization techniques for calibrating the Tank model. These methods were the SCE-UA, genetic algorithms (GA) and simulated annealing (SA) methods. They found that out of the three global optimization methods, SCE-UA provided better estimates of the optimal solution than GA and SA methods. SCE-UA was also the best in terms of efficiency as expressed by the number of iterations for convergence. They concluded that the superior performance of SCE-UA was perhaps due to its strategy of concurrently exploring several different promising regions of attraction. Kuczera (1997) compared four algorithms, namely the SCE-UA method, the GA (with traditional crossover) and multiple random starts (using either simplex or quasi-Newton local searches) for parameter identification of the modified Surface inFiltration Baseflow (SFB) model (Boughton, 1984). In his case study, the SCE-UA algorithm was found to be most robust and efficient. Thyer et al. (1999) compared the performance of two probabilistic global optimization methods: SCE-UA and the three-phase simulated annealing algorithm (SA-SX). Both algorithms were used to calibrate 2 parameter sets (a reduced, well-identified parameter set and the full parameter set) of a modified version of the SFB model using data from 2 Australian catchments that have low and high runoff yields. For the reduced, well-identified parameter set, the algorithms have a similar efficiency for the low-yielding catchment, but SCE-UA is almost twice as robust. Although the robustness of the algorithms is similar for the high-yielding catchment, SCE-UA is six times more efficient than SA-SX. Ndiritu and Daniell (2001) observed that the simple GA is inappropriate for global optimum location and hence they modified the simple GA by using 3 strategies for its improvement. They further compared their improved GA with the SCE-UA and noted that SCE-UA outperformed the improved GA on 2 of the 3 optimization problems that they investigated. Thus, they concluded that SCE-UA would be the first preference for the optimization of unfamiliar continuous variable problems including rainfall-runoff model calibration. These studies by various researchers have demonstrated that the SCE-UA method is a robust and efficient search algorithm.

Enhancing the SCE-UA Algorithm

Despite the good performance of the SCE-UA, the question remains as to whether the robustness and efficiency of the calibrating algorithm can be further improved. Only a few studies have been attempted to enhance the optimizing capability of the SCE-UA algorithm. Agyei and Hatfield (2006) coupled the SCE-UA with the Gradient based Lavenberg-Marquardt (GBLM) algorithm, with the aim of combining the global search power of the SCE-UA with the local search capability of the GBLM algorithm. The resultant hybrid algorithm (which they called SCEGB) was compared with the SCE-UA and GBLM algorithms on several inverse-modeling problems involving the estimation of parameters of a nonlinear numerical groundwater flow model. Using perfect (i.e., noise-free) data and also data corrupted with noise, they mention that the SCEGB and SCE-UA outperform the GBLM by producing more accurate parameter estimates. As far as the comparison of SCEGB and SCE-UA is concerned, in all simulations both were equally robust but the SCEGB was computationally more efficient. As an improvement over the SCE-UA, the Shuffled Complex Evolution Metropolis (SCEM-UA) algorithm (Vrugt et al., 2003) was developed, which adopts Markov chain Monte Carlo theory (MCMC) and uses the Metropolis-Hastings algorithm (MH) to replace the downhill simplex method to obtain a global optimal estimation. With the aim of enhancing the exploration capability of the SCE-UA algorithm, Muttil and Liong (2004) presented a systematic way of generating the initial population. The following sub-sections describe the two enhancements to the SCE-UA algorithm presented in this study.

Enhancing robustness by increasing initial population diversity

Various population-evolution-based search strategies, including the SCE-UA use a random data generator to generate the initial population. As the search proceeds, the population converges towards an optimum in one of the many possible regions of attraction. If this region of attraction does not contain the global optimum, then the search converges to a local optimum. The reason for such local minimum convergence could be insufficiently large initial population size or an initial population that is not well spread in the search space.

In order to prevent such convergence into local minima, it is essential to maintain the diversity in the population, which increases the capability of exploration in different regions of the search space. Actually, researchers in the field of evolutionary computation had realized the importance of population diversity as early as in 1970. Bremermann et al. (1966) identified that keeping individuals with lower fitness can lead to improved capability to escape local optima, and the lack of diversity is one of the major source of stagnation at ridges on the fitness landscapes. It was also proposed by Galar (1985) that evolutionary innovations may be more likely to happen from "mutations of mutants" than from the variations of the superior individuals. More discussion of population diversity is available in Fogel, (1998).

Various mechanisms have been suggested to achieve population diversity. A class of diversity preserving approaches is the crowding-based techniques. De Jong (1975) proposed the crowding approach, where each offspring is compared to a randomly selected set of K (crowding factor) individuals from the original population and the most similar (according to a distance function) individual is replaced by the offspring. Many variations of the crowding approach have been proposed (for example, Mengshoel & Goldberg, 1999). A second class of diversity preserving approaches is inspired by the biological niche concept (Holland, 1975). Goldberg and Richardson (1987) developed the first widely accepted fitness sharing technique, in which the perceived fitness of an individual is modified according to both the actual performance fitness and its neighborhood information. The more crowded the area the individual is located in, the more its fitness is degraded. This creates a number of artificial niches, each having a carrying capacity proportional to the general quality of the solutions in the niche. One of the drawbacks of fitness sharing is that it requires the definition of a neighborhood radius σ_{share} which is hard to estimate. A third class of diversity preserving approaches works by explicit segregation of competition according to some rules. A widely used technique is the "island" parallel model, in which individuals are allocated to a number of subpopulations and competition occurs only among individuals within the same subpopulation. Another way to implement segregation of competition is by defining some kind of distance metric and only allowing individuals in the range of a specified distance to compete and mate (Davidor, 1991). The underlying assumption of these spatial separation approaches is that different niches will converge to different areas of the search space and thus the diversity of the whole population can be maintained. Yet another approach to achieve population diversity is to actively generate diversity, either by restarting the search or by just increasing the mutation rate. In the restarting approach, when the population shows the signal of getting trapped in local optima, an entire new epoch is started and the population is filled with new random individuals (Fukunaga, 1997). The CHC algorithm (Eshelman, 1990) (CHC stands for Cross-generational selection, Heterogeneous recombination and Cataclysmic mutation) employs the so-called Cataclysmic mutation to pursue sustainable search. When the population gets converged, a highly disruptive mutation is applied to the best individuals for multiple times to reinitialize the population by mutating some percentage of its bits (e.g. 35%).

Given the importance of population diversity, with the aim of increasing the diversity of the initial population of points, a scheme to strategically locate the initial population so that the points are well spread in the search space was proposed in a previous study (Muttil and Liong, 2004). Figure 1 shows the locations of points in the initial population, suggested by the proposed scheme, for a 2-dimensional search space. This systematic initial population consists of 2^N non-axial points and (2N+1) axial points (where N is the dimension of the search space).

The points in the initial population for a 2-dimensional search space are presented in Table 1 and in Table 2 for a 3-dimensional search space (with lower and upper bounds of the parameters being -1.0 and +1.0 respectively). The strategically located initial population is expected to increase the exploratory capability of the search algorithm. On a suite of test functions, Muttil and Liong (2004) demonstrated that the strategically located initial population significantly reduced the number of failures in locating the known global optimum, thus enhancing the robustness of the SCE-UA.

A limitation of the proposed initial population is that when the dimension of the search space increases, the number of points in the initial population becomes excessively large. For a problem with 8 variables (i.e., N = 8), the suggested number of points is 273 [= 2^8 + (2*8 + 1)] and for a 10-dimensional problem, the population size increases to 1045. To overcome this limitation, we propose to select the points in the initial population in such a way that first, the axial points are included in the initial population and if the "population size" is larger than the number of axial points, then the non-axial points are included in the initial population. For a 2- and 3-dimensional search space, the sequence of selecting the points in the initial population is as presented in Table 1 and Table 2 respectively. Considering the 3-dimensional search space, if the "population size" is 10, then it is made up of the 7 axial points and the top three non-axial points from Table 2. If the "population size" is even larger than the number of axial and non-axial points, then the remaining points in the initial population are randomly generated from within the search space.

Enhancing efficiency using improved simplex search method

With the aim of improving the exploitative capability of the SCE-UA, a second enhancement to the SCE-UA is presented in this section, which deals with modifying the simplex search method. In the simplex search method employed in the original SCE-UA, the new points are generated by reflecting (or contracting) the worst point (X_w in Figure 2) in a simplex about the centroid of the remaining points (X_c). We propose to shift the newly generated reflected (or contracted) point towards the best point in the simplex (X_b), with the aim of directing the simplex towards the optimum using a lesser number of function evaluations. Thus, not only the worst point, but also the best point in a simplex is used to generate the reflected (or contracted point), thus making better use (or better exploitation) of the already available information.

The reflected and contracted points are shifted towards the best point using a parameter θ, which is defined as below for reflection and contraction, respectively.

$$X_{new} = ((1.0 - \theta) * X_{ref}) + (\theta * X_b) \quad (1)$$

$$X_{new} = ((1.0 - \theta) * X_{con}) + (\theta * X_b) \quad (2)$$

where X_{new} is the newly generated point, X_{ref} is the reflected point, X_{con} is the contracted point and X_b is the best point in the simplex. The parameter θ can take values between 0.0 and 1.0. The higher its value, the more is the exploitation pressure, since the new point (X_{new}) moves closer to the best point (X_b). When $\theta = 0.0$, the newly generated point, X_{new}, remains at X_{ref} (for the reflection step) and at X_{con} (for the contraction step). When $\theta = 0.5$, the new point is in the middle of the reflected (or contracted) point and the best point (X_b), which is shown in Figure 2 (a) and (b), for reflection and contraction steps respectively.

Experiment on Test Functions

This section presents a performance comparison between the original SCE-UA and the enhanced SCE-UA on a suite of ten test functions. These test functions include the seven that were originally used by Duan et al. (1993) to demonstrate the robustness and efficiency of the SCE-UA algorithm. The ten test functions used in this study can be found in Table 3. Of the three new test functions that are included for the comparison, two are 2-dimensional functions

(Griewank and Schwefel functions) and the third is a 10-dimensional function (Neumaier no. 3 function; also called Trid function) (Neumaier, 2008). Like the Rastrigin test function, the Griewank function is a non-linear multi-modal function, with a huge number of local optima contained in a single region of attraction. Figure 3 shows the multi-modal nature of the Griewank function, with many small peaks and valleys spread throughout the response surface. The Schwefel test function is also a non-linear multi-modal function, but with relatively lesser number of local optima. It is included because it has a deceptive response surface in that the global minimum is located far from the second-best minimum, and thus search algorithms are potentially prone to get trapped in the wrong direction. Moreover, the global minimum is near the bounds of the domain. This function will test a search algorithm's ability to overcome potentially deceptive response surfaces. The Neumaier no. 3 function being a 10-dimensional function will also serve as an additional performance gauge of the search algorithms' ability to perform in a high-dimensional environment. The details of the three new test functions can be found in the Appendix.

For the two types of SCE-UA algorithms, the default values of the parameters are used, as recommended by Duan et al. (1994). These default parameter values are the number of points in a complex being (2N+1); the number of points in a sub-complex is (N+1) and the number of evolution steps taken by each complex before shuffling is (2N+1), where N is the dimension of the search space. The performance criteria used are: (i) the number of failures (*NF*) out of 100 trials; and (ii) the average number of function evaluations (*AFE*) resulting from successful trials. The criterion *NF* measures robustness while *AFE* describes the efficiency of the algorithm.

The stopping criteria used is the same as that used by Duan et al. (1993). A trial is deemed a success as soon as the best function value in the sample became less than 10^{-3}. However, if the trial reached 25,000 function evaluations without reducing the best function value below 10^{-3}, the trial is deemed a failure. Exceptions to the stopping criterion of 25,000 function evaluations are the Neumaier no. 3 function and the Griewank 10D function, which being 10-dimensional functions, are expected to require higher number of function evaluations. As such, the maximum number of function evaluations for these 2 test functions is set to 50,000.

If the systematic initial population already consists of the known global optimum of a test function, the comparison is obviously not fair. Such an example is the Rastrigin function, whose parameter ranges are $-1 \leq X_1 \leq 1$ and $-1 \leq X_2 \leq 1$, and the global optimum is at the point (0, 0). In the systematically generated initial population, one of the proposed axial points is the point (0, 0). In such cases, the parameter ranges are changed to $-0.9 \leq X_1 \leq 1$ and $-0.9 \leq X_2 \leq 1$, so that the systematic initial population does not contain the global optimum.

The results of the comparison are presented in Table 4. For the enhanced version of the algorithm, the value of θ that gave the best results is also presented in Table 4. It is seen that best results are obtained when θ is in the range 0.1 - 0.5 and values higher than this lead to an increase in the number of failures. It is clearly seen that on the test functions, the proposed enhancements significantly reduce the number of failures (*NF*) (thus improving the robustness) and also the average function evaluations (*AFE*) (thus improving the efficiency) of the SCE-UA algorithm.

For the high dimensional test functions, namely the 6-dimensional Hartman function and the 10-dimensional Neumaier no. 3 and Griewank functions, the number of points in the proposed initial population is 77 and 1045 respectively, thus necessitating the use of 6 complexes for the Hartman function and 50 complexes for the 10-dimensional functions. As mentioned previously, in order to avoid large population sizes, it is proposed to select the axial points first into the initial population and then, if necessary, the non-axial points are included. By using such an approach for filling up the initial population, even lesser number of complexes can be used. Tables 5, 6 and 7 present the results for using lesser number of complexes than that proposed by the systematic initial population for the Hartman, Neumaier no. 3 and Griewank 10D test functions respectively. It is observed in these tables that even with a lesser number of complexes, the enhanced SCE-UA performs better than its original counterpart for all the three test functions. Especially for the Griewank 10D function, it is observed that when the value of θ is increased beyond 0.0, the number of failures tend to increase, suggesting that the SCE-UA works best with only the first enhancement.

To further test the capability of the enhanced SCE-UA on high dimensional calibration problems, it is compared to the original SCE-UA on a 30-dimensional version of the Rastrigin test function. The stopping criteria used is the same as that for the 10-dimensional test functions. The results of the comparison are presented in Table 8, which again shows significant improvement in the statistics *NF* and *AFE*, demonstrating improved robustness and efficiency of the enhanced SCE-UA in a higher dimensional search space.

Experiment on a Conceptual Rainfall-Runoff Model

The enhanced SCE-UA, which was shown to have improved robustness and efficiency on the test functions, is now applied to the calibration of a conceptual rainfall-runoff (CRR) model using synthetic runoff data, generated using a "true" (i.e. known) parameter set. The CRR model considered in this study is the SIXPAR model, which is a simplified research version of the Sacramento Soil Moisture Accounting (SAC-SMA) model. The SIXPAR model was originally used by Duan et al. (1992, 1993) to demonstrate the difficult nature of CRR model calibration, owing to the severity of the problem of multiple optima for different parameter subspaces of this model. Since this study presents an enhanced version of the original SCE-UA algorithm, we thought it important to compare the enhanced SCE-UA with its original counterpart on the SIXPAR model.

Figure 4 presents the SIXPAR model, which retains some of the major modeling concepts of the SAC-SMA model, including the two-layer structure and the percolation features, while deleting some components such as evapotranspiration and tension water reservoirs of the SAC-SMA model. The six parameters considered for calibration are UM, BM, UK, BK, A and X. The parameters UM and BM (units of length) act as thresholds that limit the sizes of the upper and lower zone storages, respectively. The parameters UK and BK (units of time^{-1}) control the rates of the recession, whereas parameters A and X (dimensionless) relate to the nonlinear percolation process (Duan et al., 1992).

For the SIXPAR model, using a known "true" parameter set (UM = 5.0, BM = 0.2, UK = 0.8, BK = 0.2, A = 0.3, X = 3.0), a 200-day synthetic sequence of daily rainfall and streamflow data was constructed. The lower and upper parameter bounds used to define the feasible parameter space are UM = (0, 50), BM = (0, 50), UK = (0, 1), BK = (0, 1), A = (0, 1) and X = (0, 10). This procedure for construction of synthetic data is similar to that employed by Duan et al. (1992).

The original SCE-UA is compared with the version with the two enhancements in locating the known global optimum of the SIXPAR model. As with the experiment on test functions, all the parameters of the SCE-UA are the default ones. The objective function that is minimized is the simple least squares (SLS). The stopping criteria used is also same as that used for the test functions, except that the maximum number of function evaluations is set to 10,000. The results of the comparison are presented in Table 9. The number of complexes was varied from 2 to 6 and further runs with higher number of complexes was not conducted because with 6 complexes, the original SCE-UA was able to locate the global optimum in all the 100 trials (i.e., *NF* = 0 in Table 9). The average function evaluations (*AFE*) needed by the original SCE-UA is 2191. The enhanced SCE-UA, on the contrary achieves *NF* = 0 with 4 complexes and with an *AFE* value of 1177, which is about 50% less number of function evaluations to locate the global optimum, when compared to its original counterpart.

Even for other values of the number of complexes, it is seen from the results presented in Table 9 that the enhanced SCE-UA is significantly more robust and efficient. For example, with 3 complexes, the original SCE-UA has 8 failures in the 100 trials (i.e., *NF* = 8) and the average function evaluations required for the successful trials is 1258 (i.e., *AFE* = 1258). The enhanced SCE-UA, with 3 complexes has *NF* = 1 and *AFE* = 999 (for $\theta = 0.2$). Thus, in calibrating the SIXPAR model also, the two proposed enhancements significantly improve the robustness (indicated by significant reduction in the number of failures to locate the global optimum) and also the efficiency (indicated by substantial reduction in the number of function evaluations in locating the global optimum) of the SCE-UA algorithm.

Conclusion and Recommendations

The present study proposes two enhancements to the SCE-UA model-calibrating algorithm, which is compared with the original SCE-UA on a suite of test functions and also on the SIXPAR conceptual rainfall-runoff model. The test functions used include non-linear, highly multi-modal, deceptive test functions with dimensions ranging from two to thirty, which are expected to provide a difficult test for any optimization algorithm. Test functions like Griewank and Rastrigin have a large number of local optima, which is expected in the response surface of a rainfall-runoff model.

The first enhancement is a scheme to systematically, instead of randomly, generating the initial population, which leads to a much better exploration of the search space, demonstrated by a significant reduction in the number of failures in locating the global optimum. The second enhancement, a modification to the downhill simplex search method leads to enhanced exploitation, which is demonstrated by a reduction in the number of function (or model objective function) evaluations to reach the global optimum. In this study, the parameter θ is used to control the exploitative power of the simplex search method. For both the test functions and the SIXPAR model, different values of θ are analyzed and it is observed that in general its value ranging from 0.1 – 0.3 lead to reductions in the number of function evaluations to reach the global optimum. In some cases, as in the SIXPAR model, with two complexes, a value of θ = 0.3 leads to an increase in the number of failures and therefore it seems safe to conclude that a value of θ up to 0.2 would lead to improvements in efficiency. Further study would be undertaken for delineating some guidelines on how to accurately choose the value of θ.

Thus, in conclusion, this study proposes two enhancements to the original SCE-UA model-calibrating algorithm to improve its robustness and efficiency, which is demonstrated on a suite of test functions and also on the SIXPAR rainfall-runoff model.

Acknowledgment

The authors wish to thank Dr. Q. Duan of the NOAA (National Oceanic and Atmospheric Administration, USA) for kindly providing the source codes for the SCE-UA (version 2.2), the SIXPAR model and the data used in this study.

Appendix

Details of the test functions other than the ones used by Duan et al. (1993) are presented below:

a) The Schwefel function (Schwefel, 1981) is a 2-dimensional test function, which is given by the equation:

$$f(\mathrm{x}) = \sum_{i=1}^{2} -x_i \,.\, \sin(\sqrt{|x_i|})$$

The upper and lower bounds are -500 $\leq x_1, x_2 \leq$ 500 and the global minimum is $f(\mathrm{x})$ = –837.9658 at (x_1 = 420.9687, x_2 = 420.9687).

b) The Griewank 2-dimensional function (Griewank, 1981) is given by the following equation:

$$f(\mathrm{x}) = \sum_{i=1}^{2} (x_i^2/200) - \prod_{i=1}^{2} \cos(x_i/\sqrt{i}) + 1$$

The upper and lower bounds are -550 ≤ x_1, x_2 ≤ 600 and the global minimum is f(x) = 0 at (x_1 = 0, x_2 = 0). For this test function, the lower bound is set to -550 (instead of -600) so that the systematically generated initial population does not contain the global minimum.

c) The Neumaier No. 3 function (Neumaier, 2008), also called Trid function is a 10-dimensional function and is given by the following equation:

$$f(\mathrm{x}) = \sum_{i=1}^{10} (x_i - 1)^2 - \sum_{i=2}^{10} x_i \, . \, x_{i-1}$$

The upper and lower bounds are –100 ≤ x_i ≤ 100 and the global minimum is f(x) = –210 at (10, 18, 24, 28, 30, 30, 28, 24, 18, 10).

d) The Rastrigin 30-dimensional function (Rastrigin, 1974) has the equation as below:

$$f(\mathrm{x}) = 30 + \sum_{i=1}^{30} [x_i^2 - \cos(2\pi x_i)]$$

The upper and lower bounds are –0.9 ≤ x_i ≤ 1 and the global minimum is f(x) = 0 at x_i = 0 for i = 1 to 30. For this test function also, the lower bound is set to -0.9 so that the systematically generated initial population does not include the global minimum.

References

1. Agyei, E. and Hatfield, K. 2006. Enhancing gradient-based parameter estimation with an evolutionary approach. *Journal of Hydrology*, 316, 266-280.

2. Ajami, N. K., Gupta, H., Wagener, T. and Sorooshian, S. 2004. Calibration of a semi-distributed hydrologic model for streamflow estimation along a river system, *Journal of Hydrology*, 298 (1-4), 112-135.
3. Atiquzzaman, M. and Liong, S. Y. 2004. Application of shuffled complex evolution to water distribution network rehabilitation. *Proceedings of the 6th international conference on Hydroinformatics*, June 21-25, Singapore, Vol. I, 882-889.
4. Boughton, W. C. 1984. A simple model for estimating the water yield of ungauged catchments. Australian Civil Engineering Transactions, 26(2), 83-88.
5. Bremermann, H. J., Rogson, M. and Salaff, S. 1966. Global properties of evolution processes. Natural Automata and Useful Simulations. H.H. Pattee, E.A. Edelsack, L. Fein and A.B. Callahan (eds), Spartan Books, Washington D.C., 3-41.
6. Cheng, C.T., Ou, C.P. and Chau, K.W. 2002. Combining a fuzzy optimal model with a genetic algorithm to solve multiobjective rainfall-runoff model calibration. *Journal of Hydrology*, 268 (1-4), 72-86.
7. Cheng, C.T., Zhao, M.Y., Chau, K.W. and Wu, X.Y. 2006. Using genetic algorithm and TOPSIS for Xinanjiang model calibration with a single procedure. *Journal of Hydrology*, 316 (1-4), 129-140.
8. Contractor, D.N. and Jenson, J.W. 2000. Simulated effect of vadose infiltration on water levels in the Northern Guam lens aquifer. *Journal of Hydrology*, 229 (3-4), 232-245.
9. Cooper, V. A., Nguyen, V. T. V. and Nicell, J. A. 1997. Evaluation of global optimisation methods for conceptual rainfall-runoff model calibration. *Water Science Technology*, 36(5), 53-60.
10. Cui, L. J. and Kuczera, G. 2003. Optimizing urban water supply headworks using probabilistic search methods. *Journal of Water Resources Planning and Management*, ASCE, 129 (5), 380-387.

11. Davidor, Y. 1991. A naturally occurring niche and species phenomenon: The model and first results. In Proceedings of the Fourth International Conference on Genetic Algorithms, San Mateo, California, USA: Morgan Kaufmann Publishers. 257-263.
12. De Jong, K. A. 1975. An Analysis of the behavior of a class of genetic adaptive systems. PhD thesis. University of Michigan. Dissertation Abstracts International 36(10), 5410B. (University Microfilms No. 76--9381).
13. Duan, Q., Sorooshian, S., and Gupta, V.K. 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research*, 28(4), 1015-1031.
14. Duan, Q.A., Gupta, V.K., and Sorooshian, S. 1993. Shuffled complex evolution approach for effective and efficient global minimization. *Journal of Optimization Theory and Application*, 76(3), 501-521.
15. Duan, Q.A., Sorooshian, S., and Gupta, V.K. 1994. Optimal use of the SCE-UA global optimization method for calibrating watershed models. *Journal of Hydrology*, 158, 265-284.
16. Eshelman, L. J. 1990. The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination. 265-283, Morgan Kauffman, San Mateo, CA.
17. Fogel, D. B. 1998. Evolutionary Computation: The Fossil Record. (eds). Wiley-IEEE Press.
18. Franchini, M., Galeati, G. and Berra, S. 1998. Global optimisation techniques for the calibration of conceptual rainfall-runoff models. *Journal of Hydrologic Science*, 43(3), 443-458.
19. Fukunaga, A. 1997. Restart scheduling for genetic algorithms. In Proceedings of the Seventh International Conference on Genetic Algorithms, Thomas Back (eds.).

20. Galar, R. 1985. Handicapped individuals in evolutionary processes. *Biological Cybernetics*, 51, 1-9.
21. Gan, T.Y. and Biftu, G. F. 1996. Automatic calibration of conceptual rainfall-runoff models: optimization algorithms, catchment conditions, and model structure. *Water Resources Research*, 32(12), 3513-3524.
22. Gan, T. Y., Dlamini, E. M., and Biftu, G. F. 1997. Effects of model complexity and structure, data quality, and objective functions on hydrologic modeling. *Journal of Hydrology*, 192, 81-193.
23. Goldberg, D.E. and Richardson, J. 1987. Genetic algorithms with sharing for multimodal function optimization. In Proceedings of the Second International Conference on Genetic Algorithms, 41–49.
24. Griewank, A. O. 1981. Generalized descent for global optimization. *Journal of Optimization Theory and Applications*, 34, 11-39.
25. He, B., Takase, K. and Wang, Y. 2007. Regional groundwater prediction model using automatic parameter calibration SCE method for a coastal plain of Seto Inland Sea. *Water Resources Management*, 21, 947–959.
26. Holland, J.H. 1975. Adaptation in natural and artificial systems. First edition. University of Michigan Press. Reprint MIT Press, 1992.
27. Jayawardena, A.W., Muttil, N. and Lee, J.H.W. 2006. Comparative analysis of a data-driven and a GIS based conceptual rainfall-runoff model. *Journal of Hydrologic Engineering*, ASCE, 11 (1), 1-11.
28. Kuczera, G. 1997. Efficient subspace probabilistic parameter optimization for catchment models. *Water Resources Research*, 33(1), 177-185.
29. Madsen, H. 2000. Automatic calibration of a conceptual rainfall–runoff model using multiple objectives. *Journal of Hydrology*, 235, 276-288.

30. Mengshoel, O. J. and Goldberg, D. E. 1999. Probabilistic crowding: deterministic crowding with probabilistic replacement. In Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-99), 409-416.
31. Muttil, N. and Liong, S.Y. 2004. A superior exploration-exploitation balance in shuffled complex evolution. *Journal of Hydraulic Engineering*, ASCE, 130 (12), 1202 - 1205.
32. Ndiritu, J. G. and Daniell, T. M. 2001. An improved genetic algorithm for rainfall-runoff model calibration and function optimization. *Mathematical and Computer Modelling*, 33, 695-706.
33. Nelder, J. A., and Mead, R. 1965. A simplex method for function minimization. *Journal of Computing*, 7, 308-313.
34. Neumaier, A. 2008. Hard Global Optimization Test Problems, www-document as of 2008, http://www.mat.univie.ac.at/~neum/glopt/my_problems.html.
35. Rastrigin, L. A. 1974. Systems of Extremal Control (in Russian), Nauka, Moscow.
36. Schwefel, H. P. 1981. Numerical optimization of computer models. John Wiley & Sons, Chichester, UK.
37. Sorooshian, S., Duan, Q. and Gupta, V. K. 1993. Calibration of rainfall-runoff models: application of global optimization to the Sacramento soil moisture accounting model. *Water Resources Research*, 29(4), 1185-1194.
38. Tanakamaru, H., and Burges, S.J. 1996. Application of global optimization to parameter estimation of the Tank model. *Proceedings of the International Conference on Water Resources and Environment Research*, 29-31 October, Kyoto, Japan, Vol. II, 39-46.
39. Thyer, M, Kuczera, G., and Bates, B. C. 1999. Probabilistic optimisation for conceptual rainfall-runoff models: A comparison of the shuffled complex evolution and simulated annealing algorithms. *Water Resources Research*, 35(3), 767-773.

40. Vrugt, J A, Gupta H V, Bouten W, and Sorooshian S. 2003. A shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resources Research*, 39 (8): SWC11-SWC116.
41. Wang, Q. J. 1991. The genetic algorithm and its application to calibrating conceptual rainfall-runoff models. *Water Resources Research*, 27(9), 2467-2471.
42. Yapo, P., Gupta, V. K. and Sorooshian, S. 1996. Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data, *Journal of Hydrology*, 181, 23-48.

List of Tables:

Table 1. Proposed initial population for a 2-dimensional search space

	X_1	X_2
Axial points [a]	0.0	0.0
	+ 0.5	0.0
	– 0.5	0.0
	0.0	+ 0.5
	0.0	– 0.5
Non-axial points [b]	+ 0.75	+ 0.75
	– 0.75	+ 0.75
	+ 0.75	– 0.75
	– 0.75	– 0.75

[a] 2N+1 (i.e. 5) axial points

[b] 2^N (i.e. 4) non-axial points

Hydrological Processes, Vol. 22, Issue 23, pp. 4628-4638 (2008)
The published article can be downloaded from: http://dx.doi.org/10.1002/hyp.7082

Table 2. Proposed initial population for a 3-dimensional search space

	X_1	X_2	X_3
Axial points [a]	0.0	0.0	0.0
	+ 0.5	0.0	0.0
	– 0.5	0.0	0.0
	0.0	+ 0.5	0.0
	0.0	– 0.5	0.0
	0.0	0.0	+ 0.5
	0.0	0.0	– 0.5
Non-Axial points [b]	+ 0.75	+ 0.75	+ 0.75
	– 0.75	+ 0.75	+ 0.75
	+ 0.75	– 0.75	+ 0.75
	– 0.75	– 0.75	+ 0.75
	+ 0.75	+ 0.75	– 0.75
	– 0.75	+ 0.75	– 0.75
	+ 0.75	– 0.75	– 0.75
	– 0.75	– 0.75	– 0.75

[a] 2N+1 (i.e. 7) axial points

[b] 2^N (i.e. 8) non-axial points

Hydrological Processes, Vol. 22, Issue 23, pp. 4628-4638 (2008)
The published article can be downloaded from: http://dx.doi.org/10.1002/hyp.7082

Table 3. Test functions considered

Function name	Dimension
Goldstein-Price *	2
Rosenbrock *	2
Six-hump camelback *	2
Rastrigin *	2
Griewank 2D	2
Schwefel	2
Shekel *	4
Hartman *	6
Griewank 10D *	10
Neumaier no. 3	10

* Test functions used in Duan et al. (1993)

Hydrological Processes, Vol. 22, Issue 23, pp. 4628-4638 (2008)
The published article can be downloaded from: http://dx.doi.org/10.1002/hyp.7082

Table 4. Performance of SCE-UA algorithms on test functions

Function name (dimension)	(No. of complexes; population size)	Original SCE-UA		Enhanced SCE-UA (with both enhancements)		
		NF	*AFE*	*θ*	*NF*	*AFE*
Goldstein-Price (2D)	(2; 10)	2	162	0.5	0	86
Rosenbrock (2D)	(2; 10)	0	274	0.2	0	214
6-hump camelback (2D)	(2; 10)	0	162	0.5	0	87
Rastrigin (2D)	(2; 10)	34	340	0.0	20	303
Griewank (2D)	(2; 10)	12	355	0.2	9	289
Schwefel (2D)	(2; 10)	53	257	0.3	14	177
Shekel (4D)	(3; 27)	23	494	0.2	0	415
Hartman (6D)	(6; 78)	10	673	0.4	0	469
Neumaier no. 3 (10D)	(50; 1050)	0	20,989	0.6	0	9,760
Griewank (10D)	(50; 1050)	0	28,843	0.6	0	15,438

Table 5. SCE-UA runs for Hartman function with lesser number of complexes (than that required by the systematic initial population)

(No. of complexes; population size)	Original SCE-UA		Enhanced SCE-UA (with both enhancements)		
	NF	*AFE*	θ *	*NF*	*AFE*
(2; 26)	22	422	0.0	4	420
			0.1	5	381
			0.2	4	345
			0.3	2	307
			0.4	1	279
(3; 39)	16	495	0.0	4	458
			0.1	6	414
			0.2	11	379
			0.3	3	345
			0.4	2	304
(4; 52)	9	518	0.0	0	496
			0.1	0	453
			0.2	0	422
			0.3	0	370
			0.4	0	334

* When $\theta = 0.0$, only the first enhancement, namely the systematic initial population is being used by SCE-UA

Table 6. SCE-UA runs for Neumaier no. 3 function with lesser number of complexes (than that required by the systematic initial population)

No. of complexes	Original SCE-UA		Enhanced SCE-UA (with both enhancements)		
	NF	*AFE*	θ *	*NF*	*AFE*
(2; 42)	0	1261	0.0	0	1164
			0.1	0	1088
			0.2	0	1006
			0.3	1	2846
(3; 63)	0	1258	0.0	0	1244
			0.1	0	1163
			0.2	0	1363
			0.3	13	5693

* When $\theta = 0.0$, only the first enhancement, namely the systematic initial population is being used by SCE-UA

Table 7. SCE-UA runs for Griewank 10D function with lesser number of complexes (than that required by the systematic initial population)

(No. of complexes; population size)	Original SCE-UA		Enhanced SCE-UA (with both enhancements)		
	NF	*AFE*	θ *	*NF*	*AFE*
(2; 42)	24	1823	0.0	2	1593
			0.1	10	1837
(3; 63)	24	1802	0.0	19	1674
			0.1	16	1961
(4; 84)	23	1801	0.0	16	1784
			0.1	22	1961
(9; 189)	5	2671	0.0	3	2689
			0.1	12	2648
(10; 210)	8	3000	0.0	8	2978
			0.1	15	2886
(11; 231)	2	3300	0.0	1	3279
			0.1	9	3138
(12; 252)	2	3600	0.0	4	3605
			0.1	13	3430
(13; 273)	0	3989	0.0	0	3939
			0.1	6	3788

* When $\theta = 0.0$, only the first enhancement, namely the systematic initial population is being used by SCE-UA

Table 8. Performance of SCE-UA algorithms on the Rastrigin 30D test function

(No. of complexes; population size)	Original SCE-UA		Enhanced SCE-UA (with both enhancements)		
	NF	*AFE*	θ *	*NF*	*AFE*
(2; 122)	89	8334	0.0	0	3706
			0.1	0	4016
			0.2	0	3293
(3; 183)	76	6777	0.0	0	5157
			0.1	0	4678
			0.2	0	3948
(4; 244)	66	7162	0.0	0	5467
			0.1	0	5081
			0.2	0	4287
(5; 305)	56	6667	0.0	0	5771
			0.1	0	5414
			0.2	0	4663
(6; 366)	26	7281	0.0	0	6183
			0.1	0	5839
			0.2	0	5104
(7; 427)	20	7305	0.0	0	6628
			0.1	0	6440
			0.2	0	5599
(8; 488)	5	7740	0.0	0	7334
			0.1	0	7094
			0.2	0	6123
(9; 549)	2	8443	0.0	0	8065
			0.1	0	7748
			0.2	0	6746

* When $\theta = 0.0$, only the first enhancement, namely the systematic initial population is being used by SCE-UA

Table 9. Performance of SCE-UA algorithms on the SIXPAR model

(No. of complexes; population size)	Original SCE-UA		Enhanced SCE-UA (with both enhancements)		
	NF	*AFE*	θ *	*NF*	*AFE*
(2; 26)	17	945	0.0	9	1142
			0.1	12	1066
			0.2	22	1662
			0.3	53	1869
(3; 39)	8	1258	0.0	4	1139
			0.1	3	1015
			0.2	1	999
			0.3	2	1079
			0.4	18	1430
(4; 52)	2	1566	0.0	1	1390
			0.1	3	1297
			0.2	0	1177
			0.3	4	1068
			0.4	4	1020
(5; 65)	4	1832	0.0	1	1754
			0.1	0	1548
			0.2	2	1444
			0.3	2	1284
			0.4	4	1175
(6; 78)	0	2191	0.0	0	2064
			0.1	0	1813
			0.2	0	1626
			0.3	2	1508
			0.4	1	1374

* When $\theta = 0.0$, only the first enhancement, namely the systematic initial population is being used by SCE-UA

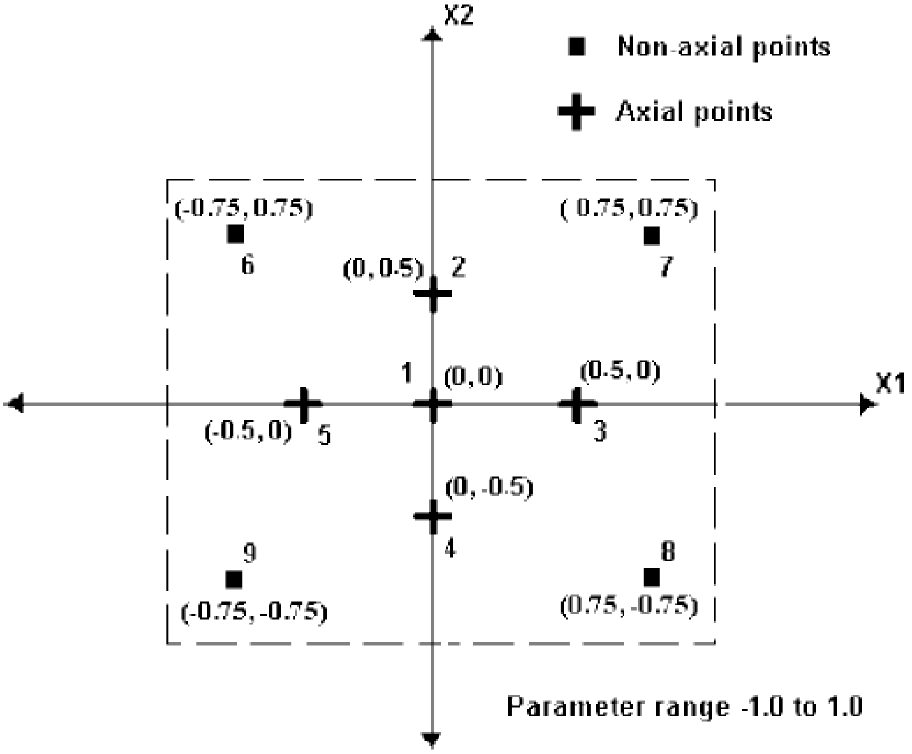

Figure 1. Systematically generated initial population

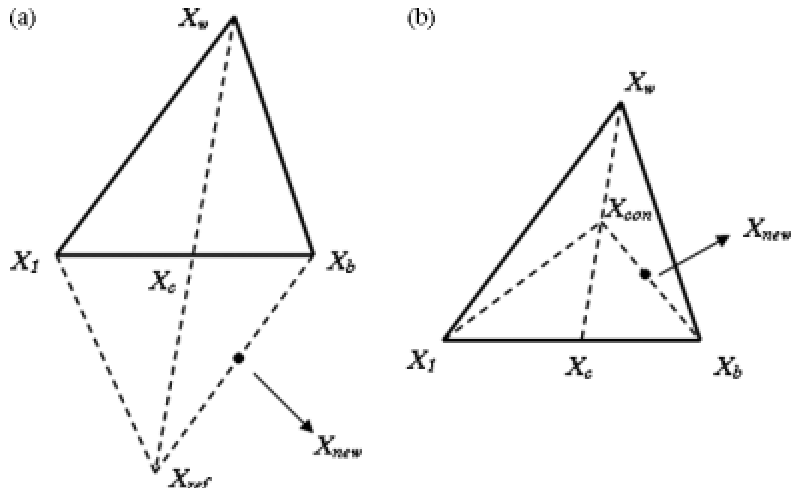

Figure 2. Reflection and contraction steps when θ = 0.5

Hydrological Processes, Vol. 22, Issue 23, pp. 4628-4638 (2008)
The published article can be downloaded from: http://dx.doi.org/10.1002/hyp.7082

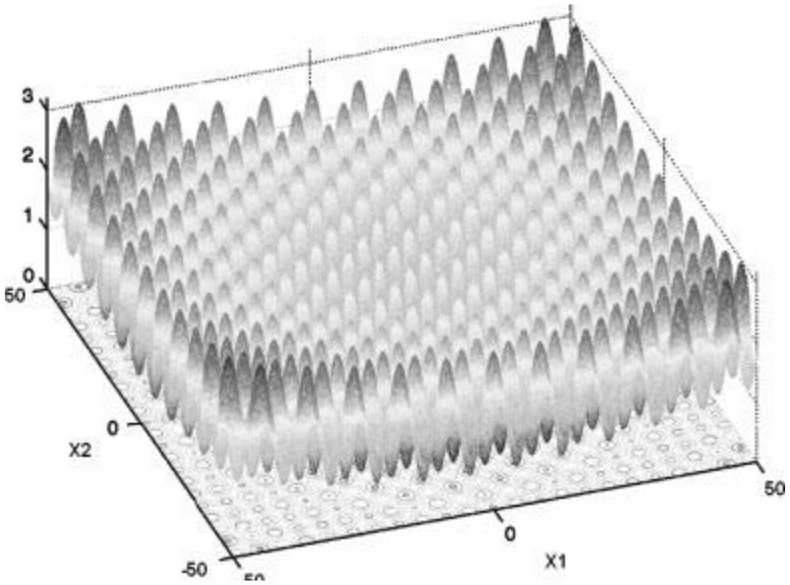

Figure 3. Multi-modal nature of the Griewank test function

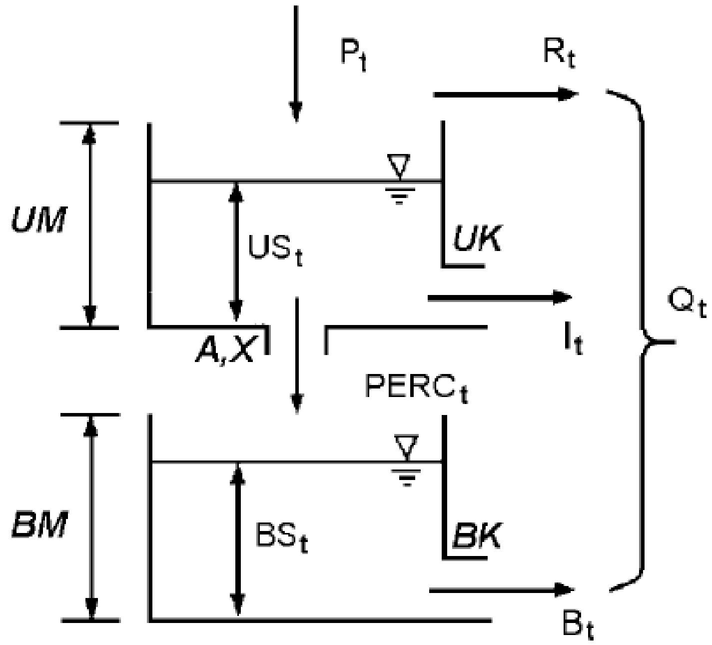

Figure 4. The CRR model SIXPAR