Improving the shuffled complex evolution scheme for optimization of complex nonlinear hydrological systems: Application to the calibration of the Sacramento soil-moisture accounting model

Wei Chu,¹ Xiaogang Gao,¹ and Soroosh Sorooshian¹

Received 19 February 2010; revised 4 May 2010; accepted 26 May 2010; published 25 September 2010.

[1] An innovative algorithm, shuffled complexes with principal components analysis (SP-UCI), is developed to overcome a critical deficiency of the shuffled complex evolution scheme: population degeneration. Population degeneration means that, during the evolutionary search process, the population of search particles may degenerate into a subspace of the full parameter space, thereby missing the capacity of fully exploring the parameter space. Being confined in a subspace may even lead the particle population to converge to nonstationary points, which is a fatal malfunction. To overcome this problem, SP-UCI employs the principal components analysis to detect the occurrence of population degeneration and remedy the adverse effects. The ensemble of calibrations of the Sacramento soil moisture accounting model with the SP-UCI method over the Leaf River basin, Mississippi, retrieves the optimal parameter values with the lowest recorded root-mean-squared error of simulated daily runoff against the observation. Moreover, the result also provides consistent (narrow ranges) model parameter distribution, which results in a better understanding of the model’s behavior, given the watershed’s hydrologic features.


1. Introduction

[2] The capability of hydrologic models in predicting runoff has been improved through parameter calibration [Brazil, 1988; Sorooshian et al., 1993; Vrugt et al., 2006], which has motivated the development and improvement of a number of automatic calibration methodologies over the past several decades [Gupta and Sorooshian, 1985; Duan et al., 1992; Vrugt and Gupta, 2003]. Calibration is the process where the parameters of model components, for which direct observations are not available, are estimated indirectly by minimizing the discrepancy between simulated and observed model outputs. As previous results in hydrologic literature have shown, calibration of conceptual rainfall-runoff (CRR) models has proven to be most challenging because of a variety of complexities associated with the n-dimensional parameter spaces of CRR models [Sorooshian and Gupta, 1983; Gupta and Sorooshian, 1983; Gan and Burges, 1990a, 1990b].

[3] Calibration methods for hydrologic models mainly fall into two categories: the direct search and the probabilistic estimation of parameters. The direct search method employs simulations of biological evolutionary processes [Goldberg and Holland, 1988; Eberhart and Kennedy, 1995; Storn and Price, 1997]. Typically, the procedure involves selection of particles (samples) in the parameter space through the use of competitive evolution schemes, such as the simplex scheme [Nelder and Mead, 1965], swarm intelligence [Eberhart and Kennedy, 1995], and genetic mutation/selection [Goldberg and Holland, 1988], to reproduce better offspring particles. After generations of evolutions, the population attempts to converge to a single location in the search domain with the best set of parameter values. In hydrology, CRR models have been the focus of model calibration studies for decades. In 1992, a direct search skill, the shuffled complex evolution scheme developed at the University of Arizona (SCE-UA) [Duan et al., 1992], set a milestone in CRR model calibration. Numerous studies using various CRR models, including the Hydrologic Model (HYMOD), the Six-Parameter (SIX-PAR) model [Duan et al., 1992], and the National Weather Service River Forecast System’s Sacramento soil moisture accounting (SAC-SMA) model [Sorooshian et al., 1993; Duan et al., 1994] demonstrated that SCE-UA is an effective, consistent, and efficient algorithm for global optimization of model parameters.

[4] In contrast with the direct search approach, probabilistic estimation treats model parameters as random variables. Techniques are developed to derive the joint probability distribution of parameters conditioned on how well model predictions match provided observations. In the procedure, parameter estimation is not at a single point but with
probabilistic uncertainty descriptions over the parameter domain. In favor of obtaining parameter uncertainty distribution, Beven and Binley [1992] and Beven and Freer [2001] used the generalized likelihood uncertainty estimation methodology, and Thiemann et al. [2001] applied the Bayesian recursive method. However, the successes of these methods rely heavily on the correct estimation of likelihood functions and a priori distributions. The posterior distribution will converge correctly only if the estimations are correct; otherwise, it will fail.

5 The complicated nature of rainfall-runoff processes makes it difficult to model the subprocesses precisely and derive the adequate likelihood function or a priori distributions for model calibrations. Therefore, in the above-mentioned studies, the Gaussian distribution, its derivatives, and other simple standard distributions were arbitrarily adapted as the likelihood function with uniform prior distribution. These assumptions about the stochastic properties of model parameters have been viewed as, perhaps, too simplistic to track the sophisticated response surfaces of parameters of CRR models [Duan et al., 1992]. The study by Thiemann et al. [2001], which includes comparisons between probabilistic estimations and the SCE-UA method, and this study indicate that, through a much slower procedure compared with the direct search approach, the results of probabilistic estimations usually converge to a region in the parameter space where the best points are different from the solution of SCE-UA, and these best points seldom achieved optimal parameter values as good as those achieved by SCE-UA.

5 Vrugt et al. [2003, 2006] developed a revised Markov chain Monte Carlo (MCMC) approach named SCEM-UA for hydrological model calibrations. In SCEM-UA, the SCE-UA procedure is strictly followed, except that the Metropolis scheme replaces the simplex scheme as the search kernel. Assisted by the shuffled complex strategy, SCEM-UA outperforms the traditional MCMC in CRR model calibration [Vrugt et al., 2003]. However, SCEM-UA still uses Gaussian (or other simple standard) distribution as the proposal distribution, which jeopardizes its validity when applied to complex problems. Therefore, experiments in this study demonstrate that SCE-UA achieves much better parameter sets than that achieved by SCEM-UA.

7 As evident by its popularity [Thyer et al., 1999], SCE-UA has resulted in a successful and objective strategy to cope with difficult problems in global optimization, which makes it superior to other direct search and probabilistic estimation methods for hydrological model calibration. This study intends to strengthen and expand the robustness of SCE-UA by overcoming a critical deficiency that SCE-UA experiences in high-dimensional or complex cases.

8 Our recent study in experimenting with high-dimensional problems revealed that the SCE-UA method suffers from a critical problem: “population degeneration”. This is mainly due to overlooking a fundamental requirement of direct search methods, which is that the searching particles should keep the capability of searching the full parameter space through the entire search process. Population degeneration refers to the phenomenon that all searching particles are driven into a subspace (or hyperplane) of the original parameter space. As illustrated in Figure 1, if there are three parameters to be calibrated, the parameter space is a three-dimensional (3-D) space. In this case, population degeneration means that all the particles fall on a single plane, which is a subspace of the parameter space. Since the SCE-UA uses linear operations (the Nelder-Mead [1965] simplex method) on current particles to generate new particles, the population will be confined on this plane in the remaining search. If the global optimum is not located on this plane, the degenerated population will miss the global optimum and has the fate of misconvergence or even stagnation at nonstationary points.

8 To overcome the adverse consequences of population degeneration suffered by SCE-UA, a new method, named Shuffled Complex strategy with Principal Component Analysis (SP-UCI), is developed. This new algorithm is formulated by integrating Principal Components Analysis (PCA) and some stage-of-the-art techniques of evolutionary computation with SCE-UA. As revealed in this study, PCA has the potency of identifying lost dimensions and restoring searches in the full parameter space. In another study (Chu et al., A new evolutionary search strategy for global optimization of high-dimensional problems, submitted to Information Sciences, 2010), we demonstrated that this method excels some prevailing direct search algorithms on optimization of high-dimensional or complex problems. In the current study, this method is applied to calibrate the SAC-SMA model and study the parameter uncertainties. Results in this study show that the proposed algorithm outperforms SCE-UA in the following aspects: (1) It retrieves better parameter values which further reduce the model simulation’s root-mean-squared error; (2) The SP-UCI method is more robust; (3) The ensemble of optimized parameters retrieved by SP-UCI better delineates the uncertainty distributions of model parameters. The latter helps the modeler and model users to better understand the model behavior.

10 This paper is organized as follows: In section 2, the population degeneration phenomenon is discussed, and our approach to address it is articulated. Section 3 compares the results of SCA-SMA model calibration over the Leaf River
2. Population Degeneration Phenomenon

Before introducing the SCE-UA algorithm and discussing its problem in high-dimensional searches, we exemplify the task of model calibration as follows.

Given a model,

\[ \hat{f}_i = f(x_i, \theta), \]

where \( x_i \) is the time series of input vector, \( \theta \) is the parameter vector with \( d \) components (parameters), and \( \hat{f}_i \) is the time series of simulated outputs.

The task is the estimate,

\[ \theta = \text{Arg min} \left( y = F(f, \hat{f}_i) \right), \]

where \( f_i \) is the time series of observed outputs and \( F(f, \hat{f}_i) \) is the objective function.

The SCE-UA procedure searching for the optimal \( \theta \) includes five steps:

1. Start the \( k \)-th generation of the population \( P_k \) including \( M \) particles \( \theta_i \) and their function evaluation values \( y_i \):

\[ P_k = \{ (\theta_i, y_i) \}_{i=1}^{M}, \quad \theta_i \in \mathbb{R}^d, \ y_i \in \mathbb{R} \]

2. Order particles \( \{\theta_i\} \) according to evaluation values such that \( y_1 \leq \cdots \leq y_M \) (\( l = 1 \) is the best particle):

\[ P_k = \{ (\theta_i, y_i) \}_{i=1}^{M}, \quad y_1 \leq \cdots \leq y_M \]

3. Distribute the ordered particles into \( p \) complexes \( C_i \), each contains \( m \)-ordered particles (note that \( M = m \times p \)).

4. For each complex \( C_i \), \( i = 1, \ldots, p \).

5. Check whether the stop criteria is satisfied. If not, break all of the complexes into \( M \) individual particles and go to 1. If yes, Stop!

The strength of the SCE-UA stems from the synthesis of concepts that have proven successful for global optimization. In the procedure, step 4 performs competitive evolution using the multistart Nelder-Mead [Nelder and Mead, 1965] simplex scheme. By searching its vicinity, a simplex can find qualified offspring effectively but with the risk of converging to a local minimum. In steps 2 and 3, the complex shuffling scheme is designed to rescue the multi-start simplexes from being trapped by local regions of attraction (minima) and put them back on a course to search for the global minimum. As presented by Duan et al. [1992] and Thyer et al. [1999], the response surfaces of CRR models usually possess a number of complex features, such as nonconvexity with numerous regions of attraction having multiple optima; discontinuous derivatives and roughness on the response surface, where the roughness generates numerous minor local optima; and interaction between parameters, resulting in high correlations. In the SCE-UA, a combination of the simplex and shuffled complex schemes makes “systematic evolution of a complex of points spanning the space, in the direction of global improvement” (Duan et al., 1992).

However, there is a fundamental requirement underlying the SCE-UA procedure: the particles in a population (complex) should span the whole parameter space in order to be able to explore the attractive region that leads toward the global optimum. Here, “span the whole search space” means that the space spanned by the particle population should possess dimensionality equal to that of the parameter space. At an intuitive level, consider using a single simplex for parameter calibration: it is well-known that, for searching a two-parameter space, the simplex should have three vertices (particles) forming a triangle (not on the same line) so that the particles (vertices) span a two-dimensional (2-D) space (i.e., in any 2-D coordinate system formed by a basis, projections of the particles on each axis have no nonzero variance); for a three-parameter space, the simplex should have four vertices forming a tetrahedron (not on a plane) so that the particles span a 3-D space (i.e., in any 3-D coordinate system formed by a basis, projections of the particles on each axis have a nonzero variance), and so on.

This requirement is usually satisfied when the SCE-UA with multiple complexes is applied to low-dimensional cases. However, for problems with high dimensionalities or for circumstances when the model parameters possess high correlations (which is often the case for models that simulate real-world processes, such as the CRR models), this assumption tends to be violated easily. When the violation occurs (i.e., population degeneration), all of the particles of a population degenerate into a subspace of the original parameter space. Geometrically, there exists at least one basis in the space and, on one or more vectors of this basis, all of the particles are projected at the same point (zero variance). These vectors are considered as lost dimensions for the population. In the SCE-UA, because the offspring particles are generated by linear combinations of their parent particles, once population degeneration occurs, the new particles will remain in the same subspace. Although the SCE-UA involves a random resampling operation (in the simplex scheme), it will not work until the linear operation fails and its efficacy decreases geometrically with dimensionality. Therefore, after population degeneration, the fate of the search will largely depend on whether the global optimum is still located inside the subspace. If the global optimum is excluded from the subspace, then a subsequent search will lead the population to converge to the “optimum” point within the subspace. This point could be a local optimum or even a nonstationary point in the full parameter space (i.e., a point with a nonzero gradient).
[28] We employed a composition benchmark function designed by Liang et al. [2005] to demonstrate how easily this nonconfigurative phenomenon will occur in high-dimensional cases and the fatal consequences. This composition benchmark function has been widely used to benchmark direct search algorithms [Suganthan et al., 2005]. It allows the user to change dimensionality, the number of minima, the locations of minima, and many other features of the response surface.

\[
f(x) = \sum_{i=1}^{10} w_i \left[f_i((x - o_i)/\lambda_i) + \text{bias}_i\right], \text{ with bias}_i = (i-1) \times 100,
\]

where \(x\) and \(o_i\) are vectors in \(R^d\), and \(d\) is the dimensionality of the problem.

\[
f_i(x) = ||x||^2
\]

\[
f'_i(x) = 2000f_i(x)/f_{\text{max}}, \quad f_{\text{max}} = \max\{f_i(x), i = 1,...,10\}
\]

\[
w_i(x) = \exp\left(-f_i(x - o_i)/2000\sigma_i^2\right)
\]

\[
\rightarrow w_i = \begin{cases} w_i & \text{if } w_i = \max\{w_i\} \\ w_i \left(1 - \max\{w_i\}\right)^{10} & \text{if } w_i \neq \max\{w_i\} \end{cases}
\]

[29] As shown in equation (3), this function is composed of 10 transformed sphere functions, \(f'_i\), with the minimum value of bias, located at \(o_i\). Therefore, \(f(x)\) has its global minimum (= 0) at \(o_1\) and nine major local minima (= \(i \times 100, i = 1, \ldots, 9\) at \(o_{2-10}\)). Locations randomly generated in the search space are \(o_{1-9}\), whereas \(o_{10}\) is set at the origin in order to trap algorithms that take advantage of this special location. To stretch (\(\lambda_i > 1\)) or compress (\(\lambda_i < 1\)) function \(f'_i\), \(\lambda_i\) is used. The variable that controls the coverage range of \(f'_i\) is \(\sigma_i\); a small \(\sigma_i\) gives a narrow range. For CF1, \(\sigma_{1-10} = 1\) and \(\lambda_{1-10} = 0.05\).

Figure 2 shows the response surface of the composition function in a 2-D case. This function has a very irregular response surface, evident regions of attraction, and multiple minima. Indicated by equation (3), this function is smooth with continuous first-order derivatives everywhere. When applied to such a smooth function, the SCE-UA is expected to converge to one of the minima where the gradients are zero. In this experiment, the SCE-UA was applied to the function with 100 dimensions for 30 times. The result shows that no run converged to any of the minima (Figure 3), and all runs terminated at points with nonzero gradients (Figure 4). Further examination confirms that all 30 ending points stopped on the slope of the attractive region toward the global minimum.

[31] To eliminate the severe effects of population degeneration, the dimensionality of the population should be monitored, and the lost dimensions should be recovered promptly. Principal component analysis (PCA) is selected to serve this purpose. PCA is an orthogonal linear transformation of the coordinates of a given data set (particles) to a new coordinate system such that the data projections have the greatest variance on the first coordinate (axis) which is called the first PC, the second greatest variance on the second coordinate (the second PC), and so on. The PC coordinate system is uniquely determined by the given data set (particles), and it has orthogonal (uncorrelated) coordinate axes which are ordered according to the variances of the data projections. If the data set has lost \(n\) independent dimensions, the data projection variances equal zero on the last \(n\) PCs. Therefore, PCA can determine the dimensionality for a given data set and, by adding new particles along the PC with zero variance, will help immediately recover the lost dimension. Based on this understanding, a PCA scheme is introduced in the SP-UCI procedure at each evolution loop. Mathematically, PCA can be easily calculated as the eigenvectors of the data covariance matrix [Dunteman, 1989].

4 of 12
SP-UCI was tested through the same experiments on the 100-D composition benchmark function. In Figure 3, all 30 randomly initiated runs of SP-UCI successfully converged to the global minimum without losing efficiency. Additional experiments comparing SP-UCI with other popular direct search algorithms, such as particle swarm optimizer [Eberhart and Kennedy, 1995] and differential evolution [Storn and Price, 1997] can be found in Chu et al. (A new evolutionary search strategy for global optimization of high-dimensional problems, submitted to Information Sciences, 2010).

3. Application to the SAC-SMA Model

The SP-UCI algorithm was also applied to the calibration of the SAC-SMA model. This model is the major component of the National Weather Service (NWS) River Forecast System and is currently serving as the operational model for flood and river flow forecasting over the United States. As mentioned above, SAC-SMA has been chosen for many CRR model calibration studies; therefore, results from previous studies are available for comparison. In Table 1, parameter and daily root-mean-square error (DRMS) values retrieved by SCE-UA and BaRE are the results by Thiemann et al. [2001], the parameter values retrieved by SCEM-UA are the results reported by Vrugt et al. [2006], and the DRMS value is calculated for the same time period (described in the following section) as that in other studies.

3.1. Study Watershed

To keep with previous studies, we selected the Leaf River basin, Mississippi, as the study watershed. This basin has an area of 1944 km$^2$, receives an annual mean precipitation of around 1300 mm, and produces mean river outflow at the rate of ~ 30 m$^3$/s. As an intensively studied area, the basin has abundant, accessible, long-term hydrological data. In this study, we obtained the input and observation data from the NWS Hydrologic Research Laboratory. These data include the observed time series of basin-averaged rainfall (mm per 6-hourly), the potential evapotranspiration (millimeters per day) computed using the Penman equation, and the daily streamflow (cubic meters per second). As was the case in previous studies [Thiemann et al., 2001], an 11 year (1 January 1953 to 31 December 1963) data set is used in calibration. In addition, another 10 year (1 January 1964 to 31 December 1973) data set is used to validate the calibrated model parameters.

3.2. Calibration

The purpose of this experiment is to compare SP-UCI with SCE-UA in estimating the model parameters and parameter uncertainties. Therefore, both methods are deployed to carry out 50 independently initialized runs under the same experimental setup. We use the DRMS error as the objective function. This objective function emphasizes high

![Figure 3](image1.png)

**Figure 3.** Fitness curves for 30 SCE-UA runs (black) and 30 SP-UCI runs (red) on the 100-D composition benchmark function.

![Figure 4](image2.png)

**Figure 4.** Histogram of gradient norms at final points of 30 SCE-UA runs on the 100-D composition benchmark function (numerically estimated with $\Delta x = 1 \times 10^{-6}$).
flows, which suits the mission of the SAC-SMA model as a part of the NWS River and Flood Forecasting System. Many other objective functions [Sorooshian et al., 1983] exist which can provide more comprehensive measurements of model performance. However, the purpose of this study is to compare two methods, and a detailed study on selecting the objective function is beyond the scope of this research.

As with previous studies, 13 parameters in the model are calibrated, with the others set at default values. The upper and lower bounds of the parameters are adopted from Brazil [1988], as listed in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>UZTWM</th>
<th>UZFWM</th>
<th>UZK</th>
<th>PCTIM</th>
<th>ADIMP</th>
<th>ZPERC</th>
<th>REXP</th>
<th>LZTWM</th>
<th>LZFSM</th>
<th>LZFPMM</th>
<th>LZSK</th>
<th>LZPK</th>
<th>PFREE</th>
<th>DRMS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lower Bound</strong></td>
<td>150.00</td>
<td>150.00</td>
<td>0.500</td>
<td>0.100</td>
<td>0.400</td>
<td>250.00</td>
<td>5.00</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>0.250</td>
<td>0.025</td>
<td>0.600</td>
<td></td>
</tr>
<tr>
<td><strong>Upper Bound</strong></td>
<td>1.00</td>
<td>1.00</td>
<td>0.100</td>
<td>0.000</td>
<td>0.000</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.010</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Brazil</td>
<td>9.00</td>
<td>39.80</td>
<td>0.200</td>
<td>0.003</td>
<td>0.250</td>
<td>250.00</td>
<td>4.27</td>
<td>240</td>
<td>40</td>
<td>120</td>
<td>0.200</td>
<td>0.006</td>
<td>0.024</td>
<td>20.3</td>
</tr>
<tr>
<td>SCE-UA</td>
<td>14.09</td>
<td>63.83</td>
<td>0.100</td>
<td>0.000</td>
<td>0.363</td>
<td>249.97</td>
<td>2.46</td>
<td>238</td>
<td>3.19</td>
<td>99.8</td>
<td>0.019</td>
<td>0.021</td>
<td>0.001</td>
<td>19.2</td>
</tr>
<tr>
<td>BaRE</td>
<td>33.61</td>
<td>76.12</td>
<td>0.332</td>
<td>0.016</td>
<td>0.266</td>
<td>117.30</td>
<td>4.95</td>
<td>236</td>
<td>132</td>
<td>124</td>
<td>0.089</td>
<td>0.015</td>
<td>0.146</td>
<td>21.8</td>
</tr>
<tr>
<td>SCEM-UA</td>
<td>10.70</td>
<td>32.55</td>
<td>0.39</td>
<td>5.1 × 10⁻⁴</td>
<td>0.10</td>
<td>241.46</td>
<td>1.66</td>
<td>261</td>
<td>16.96</td>
<td>45.08</td>
<td>0.19</td>
<td>0.01</td>
<td>0.10</td>
<td>21.6</td>
</tr>
<tr>
<td>SP-UCI</td>
<td>6.46</td>
<td>45.68</td>
<td>0.109</td>
<td>0.0001</td>
<td>0.363</td>
<td>241.08</td>
<td>1.80</td>
<td>260</td>
<td>11.9</td>
<td>80.1</td>
<td>0.25</td>
<td>0.017</td>
<td>0.0001</td>
<td>17.98</td>
</tr>
</tbody>
</table>

|                |       |       |     |       |       |       |      |       |       |        |      |      |       |       |
| **Brazil**     | 9.00  | 39.80 | 0.200| 0.003 | 0.250 | 250.00| 4.27 | 240   | 40    | 120    | 0.200| 0.006| 0.024 | 20.3  |
| **SCE-UA**    | 14.09 | 63.83 | 0.100| 0.000 | 0.363 | 249.97| 2.46 | 238   | 3.19  | 99.8   | 0.019| 0.021| 0.001 | 19.2  |
| **BaRE**      | 33.61 | 76.12 | 0.332| 0.016 | 0.266 | 117.30| 4.95 | 236   | 132   | 124    | 0.089| 0.015| 0.146 | 21.8  |
| **SCEM-UA**   | 10.70 | 32.55 | 0.39 | 5.1 × 10⁻⁴| 0.10 | 241.46| 1.66 | 261   | 16.96 | 45.08  | 0.19 | 0.01 | 0.10  | 21.6  |
| **SP-UCI**    | 6.46  | 45.68 | 0.109| 0.0001| 0.363 | 241.08| 1.80 | 260   | 11.9  | 80.1   | 0.25 | 0.017| 0.0001| 17.98 |

aUnit of all the capacity variables is mm and unit of DRMS is m³/s. Values estimated by Brazil [1988], shuffled complex evolution scheme developed at the University of Arizona (SCE-UA) and BaRE [Thiemann et al., 2001], and revised Markov chain Monte Carlo approach (SCEM-UA) [Vrugt et al., 2006].

bDRMS stands for daily root-mean-square error.

3.3. Results and Discussion

Performances of SP-UCI and SCE-UA in the 50 individual runs are plotted in Figure 5. Similar to Figure 3, all of the SCE-UA runs terminated earlier than SP-UCI with worse DRMS values. In the SP-UCI runs, the best DRMS is 17.98 m³/s, which is better than the best one achieved by SCE-UA (18.53 m³/s) and other previous studies, including Brazil’s study [1988], BaRE [Thiemann et al., 2001], and SCEM-UA (see Table 1). Since the SAC-SMA model is not a smooth function, we cannot calculate the gradients at the ending points to determine if the SCE-UA runs stop at

![Figure 5](image-url)
Figure 6. The result of the principal component analysis scheme in a randomly selected SP-UCI run. If the start falls on the line of Yes, it means that, in the corresponding loop, population degeneration occurs, and the SP-UCI procedure finds evident slope along one or more of the lost dimensions.

In Figure 6, the record of one SP-UCI run shows that, in six of the total 40 evolution loops, the PCA scheme identified and successfully remedied the lost dimensions. This observation is universal in all 50 SP-UCI runs, and it indicates that, in the SCA-SMA model calibration, population degeneration indeed occurs and can be remedied by the SP-UCI method.

In Figure 7, histograms of the final DRMS values for 50 SCE-UA runs and 50 SP-UCI runs are plotted, respectively. In comparison to the SCE-UA runs, the SP-UCI runs converge to smaller values of DRMS with a much reduced uncertainty range (SP-UCI: 17.98–18.54 m$^3$/s; SCE-UA: 18.53–23.67 m$^3$/s). This improvement indicates that the SP-UCI solution is more consistent and reliable.

Figure 7. Histograms of daily root-mean-square (DRMS) values from the calibration runs of SCE-UA and SP-UCI, respectively. The solid line and the dashed line indicate the values of runs with parameter sets of Brazil [1988] and SCEM-UA, respectively.
The consistency and superior performance of SP-UCI are also witnessed when examining the hydrographs. In general, model simulations, generated with parameters obtained from SP-UCI, capture times and magnitudes of streamflow peaks more accurately than those generated with parameters obtained from SCE-UA. Figure 8 shows the simulated and observed streamflows for the period of January–April, 1961. This time period contains the largest peak in streamflow through the entire 11 year period. From the plot, it can be clearly observed that the model with parameters retrieved by SP-UCI simulates both peak flow and low flow better than the model with parameters retrieved by SCE-UA. Model residuals of the same period are plotted in Figure 9. All model simulations underestimate the two largest peaks and overestimate the low flows. In keeping with the hydrographs in Figure 8, models calibrated by SP-UCI have smaller residuals compared with those calibrated by SCE-UA.

Parameter distributions produced from SP-UCI can shed light on model behavior given the hydrologic features of the basin. It is noticeable that the parameters labeled PCTIM (minimum impervious area) and PFREE (fraction of percolated water into lower zone free water) are both close to their lower bounds, and the parameter labeled ZPERC (maximum percolation rate) is near its upper bound. These are in accordance with the results of Brazil [1988], which were obtained using a complicated multilevel calibration strategy based on long-term observations in the watershed. Several inferences about the hydrologic properties of the Leaf River basin can be drawn based on the SP-UCI results:

1. The value of PCTIM close to zero means that the watershed produces little immediate runoff in response to rainfall.
2. High ZPERC value indicates that a substantial amount of water percolates into the lower zone, but the PFREE close to zero allows most of the percolated water to enter the lower zone tension water storage.
3. Since most percolated water goes into tension water storage, the functionalities of lower zone free water storages, both the primary and supplemental portions, are insignificant. This fact is expressed by low storage capacities (LZPFM and LZFSM) and very insensitive lateral drainage rates (LZSK and LZPK), as shown in Figure 10.

Components of water balance in observation and simulation over the simulation period are also examined and compared. The ratios of annual runoff/precipitation, evapotranspiration (ET)/precipitation, and base flow/precipitation are presented in Table 2. For the observation, ET is calculated from water balance, $ET = \text{precipitation} - \text{runoff}$, and base flow is calculated using the digital recursive filter which is formulated as [Chapman, 1991]

$$b_k = \alpha b_{k-1} + 0.5(1 - \alpha)(f_k + f_{k+1}),$$

where $b_k$ is the base flow, and the fast flow component $f_k$ can be estimated from the total flow $y_k$ as

$$f_k = \frac{3\alpha - 1}{3 - \alpha} f_{k-1} + \frac{2}{3 - \alpha} (y_k - \alpha y_{k-1}).$$

The filter coefficient $\alpha$ has the feasible range between $0.9 \sim 0.995$.

The results show that, in average, simulations calibrated with SCE-UA and SP-UCI have similar partitions of
total runoff and ET and agree well with the observation. Difference in the baseflow components for the two groups of simulations is evident. However, averages of both of them fall within the range of base flow calculated from observation with the feasible range of $\alpha$.

It should be emphasized that the global optimization method is superior to other popular stochastic frameworks in retrieving parameter uncertainties, such as Bayesian inference and MCMC, because of the following two reasons:

1. There is no assumption or simplification regarding the studied parameter distributions, whereas many stochastic methods are expected to work correctly only if the likelihood or proposal distribution can be correctly defined, which is extremely difficult. We want to clarify that, in this study, the daily root-mean-square error is used only as a statistics measure to quantify the difference between observation and simulation. We do not imply that the error is following Gaussian distribution or that mean square error is the maximum likelihood estimator.

2. The distributions of model parameters are meaningful only if the model performs well with the retrieved parameter values. Compared with the above-mentioned stochastic methods, SP-UCI prominently enhances the performance of the SAC-SMA model over the studied basin [Thiemann et al., 2001; Vrugt et al., 2006].

3.4. Verification

To test the validity of the parameter estimations from calibration, the verification experiment was also conducted. The data set of a 10 year period (1 January 1964 to 31 December 1973) following the calibration period was adopted to verify the calibration results. The parameter sets achieved by SCE-UA and SP-UCI, as well as those of Brazil [1988] and SCEM-UA (shown in Table 1), were verified. The DRMS between model simulation and observation of output flow is illustrated in Figure 11.

Similar to the results of calibration simulations, parameter sets from the SP-UCI calibration are superior to those from the SCE-UA in terms of lower DRMS mean and variance. Furthermore, parameter sets from SP-UCI consistently outperform those of Brazil [1988] and SCEM-UA.

4. Discussion

Employment of PCA is indispensable. PCA identifies population degeneration and helps restore the population’s capability of searching the full parameter space, which is the fundamental requirement of any direct search method. The PCA procedure requires additional computation time. However, for dimensionality lower than 1000, the computation time of PCA is usually trivial compared to the com-

Figure 9. Model residuals (model output - observation) of daily runoff of the Leaf River basin over the period of January–April, 1961 (top: 50 simulations with the 50 parameter sets retrieved by SP-UCI; bottom: 50 simulations with the 50 parameter sets retrieved by SCE-UA.

Figure 10. Distributions of 13 parameters that are calibrated in this study. For each parameter, the left graph shows the distribution retrieved by the final results of the SP-UCI method and the right graph shows the distribution retrieved by the final results of the SCE-UA method. (Solid lines are estimation of Brazil [1988].)
putational time of model simulations. For instance, in the SAC-SMA experiment, the PCA procedure takes only 0.1% of the computation time that required by running the model simulation one time [54]. The SP-UCI algorithm is a significant modification of SCE-UA, even for low-dimensional problems. Intuitively, population degeneration is more prone to occur on high-dimensional problems. However, in real applications, it is hard to tell beyond what number is considered high dimensional. For example, in the SAC-SMA experiment, only 13 parameters were calibrated. But, because of the complexity of the model, population degeneration indeed occurs and poses a difficulty for the SCE-UA algorithm. Theoretically, population can occur on problems of any dimensionality greater than two. Instead of running SCE-UA with the potential risk of population degeneration, one should consider SP-UCI, which has a sound and solid population-monitoring and restoration scheme. As demonstrated in Figures 3 and 5, SP-UCI shows similar efficiency when compared with SCE-UA but with much better effectiveness and accuracy.

[55] SP-UCI has the same limitation as SCE-UA, which is that users need to adjust the algorithmic parameters, such as the population size, number of complexes, and so on, to balance the efficiency and effectiveness of the optimization process [Duan et al., 1994]. Future work will be focused on developing schemes for self-adjusting algorithmic parameters and designing specialized algorithms for individual practical problems.

5. Summary and Conclusions

[56] Based on the SCE-UA method, a new evolutionary optimization strategy, SP-UCI, is developed to overcome the population degeneration problem discovered in the application of SCE-UA to high-dimensional or complex problems. Experiments on calibrating the SAC-SMA model with 11 year historical data from the Leaf River basin demonstrate that SP-UCI is more effective and more robust compared with SCE-UA. This is further substantiated by the results of an independent verification on the data set of a

Table 2. Components of Water Balances over the Simulation Period

<table>
<thead>
<tr>
<th>Component</th>
<th>Observation</th>
<th>Mean SP-UCI</th>
<th>Mean SCE-UA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Runoff/Precipitation</td>
<td>0.3330</td>
<td>0.3566</td>
<td>0.3562</td>
</tr>
<tr>
<td>ET/Precipitation</td>
<td>0.6640</td>
<td>0.6428</td>
<td>0.6434</td>
</tr>
<tr>
<td>Base flow/Precipitation</td>
<td>0.0870–0.1452</td>
<td>0.1167</td>
<td>0.0990</td>
</tr>
</tbody>
</table>

*SP-UCI denotes shuffled complexes with principal components analysis, SCE-UA denotes Shuffled Complex Evolution scheme developed at the University of Arizona, and ET denotes evapotranspiration.

*Range corresponding to the range of $\alpha$ of [0.900–0.995].

![Figure 11](image-url)  
**Figure 11.** Histograms of daily root-mean-square (DRMS) values from the verification runs of SCE-UA and SP-UCI, respectively. The solid line and the dashed line indicate the values of verification runs of Brazil [1988] and SCEM-UA, respectively.
10 year period following the calibration period. SP-UCI elevates the performance of the SAC-SMA model in simulating the Leaf River basin to a very high level, achieved by no other method. Furthermore, the ensemble of SP-UCI calibrations can provide insight into model parameter uncertainty and, hence, can help in understanding the model behavior as expected from the historical hydrological data of a particular catchment. The results from SP-UCI are in accordance with the more detailed study carried out by Brazil [1988], who used a multilevel calibration approach.

The discrepancy between model simulation and observation may stem from several confounding sources, including input error, observation error, and model error. Furthermore, the model error can be considered as two components: model structural inadequacy and model parameter uncertainty. Global optimization methods explicitly treat the model parameter uncertainty conditioned on the existence of other errors. Improved optimization methodology results in a twofold benefit, as demonstrated by the experimental results. First, from a user’s standpoint, better optimization methods can retrieve better parameter sets, which improve a model’s simulation and predictability. Second, from a modeler’s perspective, parameter sets estimated by good optimization methodology can better inform the developers whether or not the model is functioning as expected and, if not, how to improve the model structure to better represent the underlying physical process.

This study demonstrates that SP-UCI can improve the SAC-SMA model’s performance through parameter optimization. However, the usability of this algorithm should have a much wider range in many hydrologic inverse problems.

The authors would like to provide the algorithm codes upon email requests (wchu2@uci.edu) in order to have more application tests.

Acknowledgments. The authors gratefully acknowledge the valuable and constructive suggestions provided by the anonymous reviewers. The computation support was provided by Dan Braithwaite. Corrie Thies is also appreciated for her careful and professional editing of the manuscript. This research was supported by UCOP program of University of California (grant 09-LR-09-116849-SORS), CPPA program of NOAA (grants N080AR4310876 and NA05AR4310062) and NEWS program of NASA (grant NNX06AF93G).

References


S. Chu, X. Gao, and S. Sorooshian, Department of Civil and Environmental Engineering, University of California, Irvine, CA 92697, USA. (wchu2@uci.edu)