WATER RESOURCES RESEARCH, VOL. 28, NO. 4, PAGES 1015–1031, APRIL 1992

Effective and Efficient Global Optimization for Conceptual Rainfall-Runoff Models

QINGYUN DUAN, SOROOSH SOROOSHIAN, AND VIJAI GUPTA

Department of Hydrology and Water Resources, University of Arizona, Tucson

The successful application of a conceptual rainfall-runoff (CRR) model depends on how well it is calibrated. Despite the popularity of CRR models, reports in the literature indicate that it is typically difficult, if not impossible, to obtain unique optimal values for their parameters using automatic calibration methods. Unless the best set of parameters associated with a given calibration data set can be found, it is difficult to determine how sensitive the parameter estimates (and hence the model forecasts) are to factors such as input and output data error, model error, quantity and quality of data, objective function used, and so on. Results are presented that establish clearly the nature of the multiple optima problem for the research CRR model SIXPAR. These results suggest that the CRR model optimization problem is more difficult than had been previously thought and that currently used local search procedures have a very low probability of successfully finding the optimal parameter sets. Next, the performance of three existing global search procedures are evaluated on the model SIXPAR. Finally, a powerful new global optimization procedure is presented, entitled the shuffled complex evolution (SCE-UA) method, which was able to consistently locate the global optimum of the SIXPAR model, and appears to be capable of efficiently and effectively solving the CRR model optimization problem.

1. INTRODUCTION

Conceptual rainfall-runoff (CRR) models are designed to approximate within their structures the general physical mechanisms which govern the hydrologic cycle. For this reason, CRR models have found favor with many practicing hydrologists and engineers. Among the more widely used and studied CRR models are the soil moisture accounting portion (SAC-SMA) of the National Weather Service River Forecast System (NWSRFS) [*Burnash et al.*, 1973; *Brazil and Hudlow*, 1981], and the various versions of the Stanford Watershed Model (SWM) [*Crawford and Linsley*, 1966].

CRR models generally represent the soil moisture accounting phase of the hydrologic cycle as several interconnected subsystems, each representing a certain component in the processing of a hydrologic event. Empirically or heuristically determined, but physically realistic functions are used to describe the internal operation of these processes. The types of functions and the amount of detail used to represent a watershed system determine the degree of realism and sophistication of a CRR model. For any CRR model to have practical utility, it is important to be able to identify proper values for the parameters which govern these functions; the procedure for doing this is called model calibration.

The successful application of a CRR model depends on how well the model is calibrated. In recent years, automated approaches to calibration have received much attention, and several difficulties in the application of such methods have been reported [e.g., *Ibbitt*, 1970; *Johnston and Pilgrim*, 1976; *Pickup*, 1977; *Larimore*, 1981; *Sorooshian and Gupta*, 1983; *Gan and Burges*, 1990*a*, *b*]. These reports indicate that it is typically difficult, if not impossible, to obtain a unique set of optimal parameters for a CRR model using automatic calibration methods. While such problems may be partially attributable to limitations inherent in the calibration and verification data, the nonlinear structural characteristics typical of CRR models also lead to the existence of multiple optima (i.e., more than one solution). However, the automatic calibration procedures in current use are incapable of finding the globally optimal parameter estimates with any reasonable degree of confidence. This translates into uncertainty regarding the accuracy of the model forecasts. Furthermore, unless the best set of parameters associated with a given calibration data set can be found, it is difficult to determine how sensitive the parameter estimates (and hence the model forecasts) are to factors such as input and output data error, model error, quantity and quality of data, objective function used, and so on.

Most attempts at automatic calibration of CRR models have used local-type direct search optimization methods [e.g., *Dawdy and O'Donnell*, 1965; *Nash and Sutcliffe*, 1970; *Chapman*, 1970; *Ibbitt*, 1970; *Monro*, 1971; *Johnston and Pilgrim*, 1976; *Pickup*, 1977; *Sorooshian*, 1978; *Sorooshian et al.*, 1983; *Gupta and Sorooshian*, 1985; *Hendrickson et al.*, 1988]. Such procedures are not designed to handle the presence of multilocal optima, discontinuous derivatives, and other problems encountered in the calibration of CRR models. The convergence problems encountered by local search algorithms have been well documented in the literature and therefore will not be repeated here. For a historical perspective, we refer the reader to *Gupta and Sorooshian* [1985] and *Hendrickson et al.* [1988].

To our knowledge, there have been only three published reports of the use of globally based optimization methods for CRR model calibration [*Ibbitt*, 1970; *Brazil and Krajewski*, 1987; *Wang*, 1991]. *Ibbitt* [1970] reported that a random search method [*Karnopp*, 1963] was unable to obtain good estimates of the global optimum of the SWM model, but could (in some cases) provide parameter values which were good starting points for a subsequent local search. *Brazil and Krajewski* [1987] investigated the use of uniform random search and the adaptive random search (ARS) method

Paper number 91WR02985.
0043-1397/92/91WR-02985$05.00

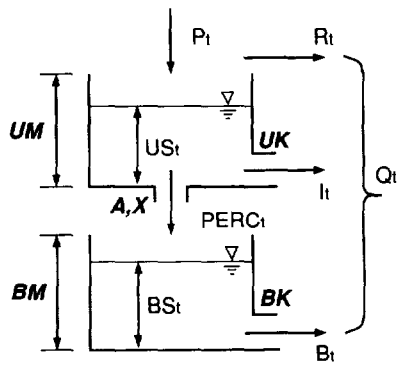

Fig. 1. The research CRR model SIXPAR.

[*Masri et al.*, 1978; *Pronzato et al.*, 1984] for fine-tuning of the parameters of the SAC-SMA model (as part of a three-stage calibration process) and concluded that the ARS method was "an attractive alternative to nonrandom search techniques" (see also *Brazil* [1988]). *Wang* [1991] reported that the genetic algorithm [*Holland*, 1975] with further tuning by a standard local search method can provide an "efficient and robust means for calibration of the Xinanjiang mode."

The theory and practice of global optimization has progressed rapidly during the last decade, and a wide variety of different algorithms are available [e.g., *Torn and Zilinskas*, 1989; *Rinnooy Kan and Timmer*, 1989]. The major reason that more research into the use of global optimization methods for CRR model calibration has not been conducted is the demand that such methods place on computational resources. However, computer technologies have now improved to the point that computationally intensive methods are practical and affordable and have made it feasible for us to conduct the investigations reported in this paper.

2. Scope

There is a clear need to have an effective and efficient optimization procedure that can help in the automatic identification of a unique and realistic set of optimal parameters for a CRR model. Such a procedure should (1) reflect the state of the art in global optimization, and (2) be carefully designed to handle the specific difficulties encountered in CRR model calibration. This paper has three main topics. In section 4 we use two computationally intensive techniques to establish clearly the nature of the problem of multiple optima for the research CRR model SIXPAR. In section 5 we test the performance of three existing global search optimization procedures. In section 6 we present a new global optimization procedure, entitled the shuffled complex evolution (SCE-UA) method, and show that it is capable of efficiently and effectively identifying the optimal values for the model parameters.

3. The Model Used in This Study

The CRR model used throughout this study is the SIXPAR model, which is a simplified research version of the SAC-SMA model. SIXPAR is a representation of some of the major modeling concepts of the SAC-SMA model, duplicating the two layer structure and the percolation feature of the SAC-SMA, while deleting some components such as evapotranspiration and tension water reservoirs (see Figure 1). The parameters UM and BM (units of length) act as thresholds that limit the sizes of upper and lower zone storages, respectively. The parameters UK and BK (units of time^{-1}) control the rates of the recession. The parameters A and X (dimensionless) relate to the nonlinear percolation process. This model was first introduced by *Gupta* [1982] and has been presented by *Gupta and Sorooshian* [1983].

For the model SIXPAR, a 200-day synthetic sequence of daily rainfall and streamflow data associated with a specified "true" parameter set (UM = 10, BM = 20, UK = 0.5, BK = 0.2, A = 0.31, X = 3) was constructed (Figure 2). In choosing the rainfall data, we have ensured that all modes of model operation were activated. If desired, synthetic "measurement" noise can be introduced into the data. 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), X = (0, 10). Note that evapotranspiration data are not required as SIXPAR has no evapotranspiration component, and that the streamflow output was not routed.

To prevent confusion, we wish to emphasize that the SIXPAR model is not intended for use in an operational setting, but was developed as a tool for the study of problems that arise in the implementation of fully automatic calibration methodologies. This simple model retains some of the important characteristics of the more complex SAC-SMA model, and we have used it to gain very useful insights into problems associated with the full-scale model. We believe that if a proposed procedure is incapable of successfully solving the problem of calibrating the SIXPAR model, it is unlikely to perform well on the full-scale model.

4. The Problem of Multiple Optima in CRR Models

The numerous difficulties that have been reported with CRR model calibration suggest that these problems may have their own peculiar characteristics. Therefore, the aim of this section is to establish clearly what the features of the problem are, how difficult the problem is, and why current procedures cannot solve the problem. This understanding will be helpful in designing a successful solution strategy.

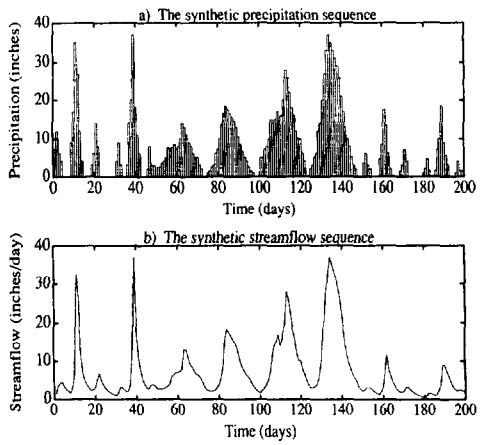

Fig. 2. The data used to calibrate model SIXPAR (1 inch equals 2.54 cm).

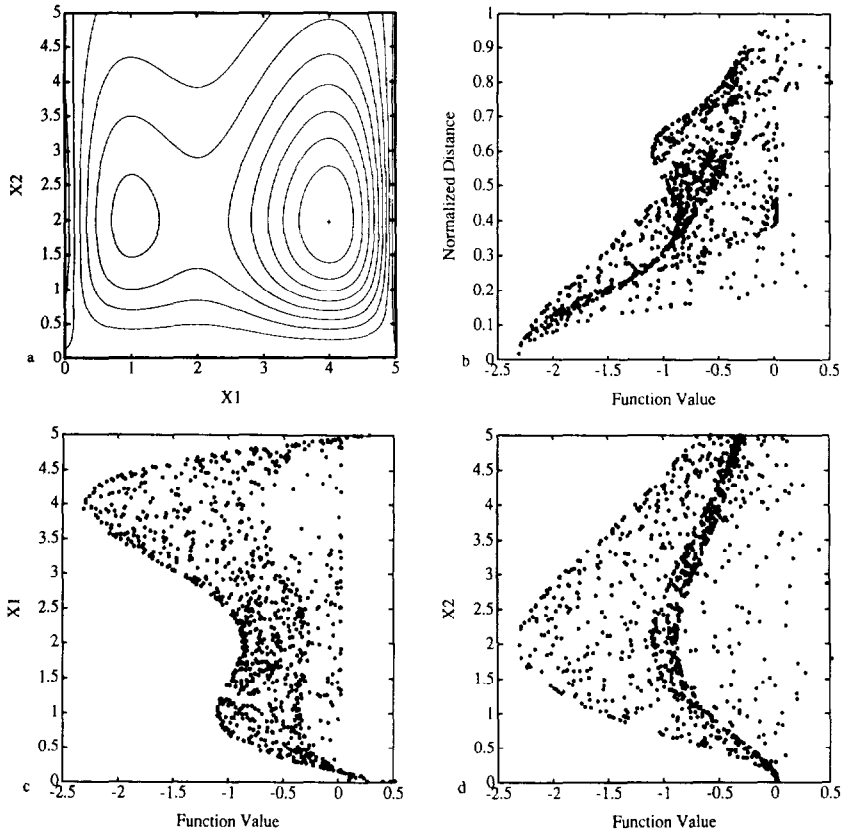

Fig. 3. Uniform random sampling demonstrated on the Hosaki function. (*a*) Contour plot of the Hosaki function. (*b*) NORD plot for the Hosaki function. (*c*) Parameter X1 PARV plot for the Hosaki function. (*d*) Parameter X2 PARV plot for the Hosaki function.

Our objective is to determine what can be learned about the structural identifiability properties of CRR models. In particular, we are interested in (1) detecting the location and number of multiple optima; and (2) obtaining global information about parameter sensitivity and the structure of the objective function response surface.

Because of the highly complex and nonlinear nature of CRR models, a theoretical analysis of the model equations to obtain the information we require is difficult, if not impossible. Therefore, we have chosen to use computational methodologies that employ information sampled from the entire parameter space, taking care to adequately span the space and provide a sufficient density of coverage. The two procedures we have used are: (1) uniform random sampling (URS) of the parameter space; and (2) exhaustive gridding (EG) of the parameter space.

Uniform Random Sampling (URS) Method

The method of uniform random sampling (URS) is a primitive probabilistic approach to global optimization. In this method, a prespecified number of points, N (say N = 1000), is sampled at random from the feasible parameter space using a uniform probability distribution. The objective function value is computed at each point, and the point with the best (we shall use minimum) objective function value is taken as an estimate of the optimum. In addition, however, the N sampled points contain important information about the nature and structure of the objective function response surface. This information can be extracted using appropriate analysis procedures. One such procedure is to construct graphical projections of the sampled points after having arranged them in order of increasing objective function value. The two graphical projections we investigated were (1) X-Y plot of the distance of each point from the optimum normalized by the parameter range (NORD), versus objective function value; and (2) X-Y plot of parameter value (PARV), versus objective function value.

To illustrate how these plots are useful, two simple examples are presented. First, consider the simple two-dimensional nonconvex Hosaki function (Figure 3*a*) having two optima, a global minimum at (X_1 = 4.0, X_2 = 2.0) and a local minimum at (X_1 = 1.0, X_2 = 2.0). The equation for this function is

$$f(X_1, X_2) = \left(1 - 8X_1 + 7X_1^2 - \frac{7}{3}X_1^3 + \frac{1}{4}X_1^4\right)X_2^2 e^{-X_2} \tag{1}$$

The NORD and PARV plots for the Hosaki function are shown in Figures 3*b*–3*d*. When the sampled points are close to the global optimum (small normalized distance), the NORD measure (Figure 3*b*) increases smoothly in a banded fashion characteristic of an elliptical-type response surface. After a certain number of points have been plotted, the ordered set will begin to contain points from the local optimum region which have similar function values to the ones around the global optimum. For such points, the

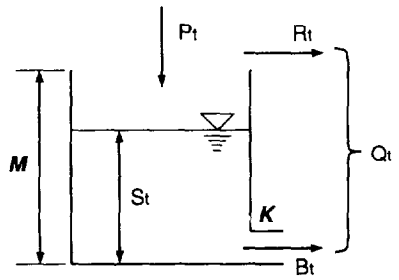

Fig. 4. The simple reservoir model TWOPAR.

normalized distance is larger, so that they cluster in a different region of the graph. In this way, the existence of the local optimum is clearly identified. In a similar manner, the local optimum and its location are identifiable on the PARV plots (Figures 3*c* and 3*d*).

Next, consider the simple two-parameter reservoir model (TWOPAR; see Figure 4) which has an extended valley on its simple least squares (SLS) function response surface (Figure 5*a*). An extended valley (extended line or area of equal function values in the parameter space) can be viewed as a special case of the multiple optimum problem. Such features have been detected in response surface studies of CRR models [*Sorooshian et al.*, 1983]. The URS method plots identify clearly both the region of the optimum and the presence of the extended valley (Figures 5*b*–5*d*). The extended valley is identified in Figures 5*b* and 5*c* as a sudden increase in the spread of the points, at a function value equal

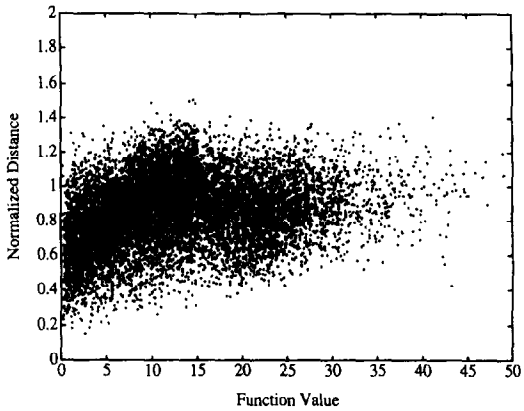

Fig. 6. NORD plot for the SIXPAR model.

to 1.5. Figure 5*c* indicates that the valley is oriented in the direction of parameter M, while Figure 5*d* indicates that the response surface in the region of the valley has a quadratic shape along parameter K.

The URS method was used to study the SLS response surface of the SIXPAR model by sampling 10,000 points from the feasible parameter space (roughly 4.5 points per parameter direction for this six-dimensional problem). The results (see Figures 6 and 7*a*–7*f*) are both interesting and disturbing. The plots all show a broad spread of points from the very beginning of the ordered data set, indicating very

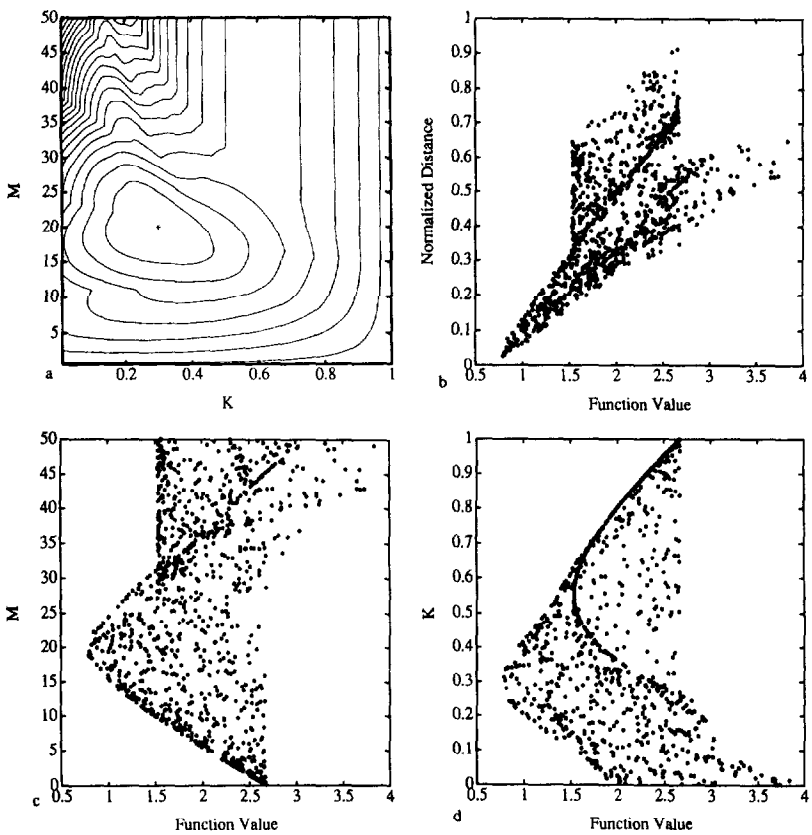

Fig. 5. URS sampling demonstrated on the model TWOPAR. (*a*) SLS contour plot for model TWOPAR. (*b*) NORD plot for model TWOPAR. (*c*) Parameter M PARV plot for model TWOPAR. (*d*) Parameter K PARV plot for model TWOPAR.

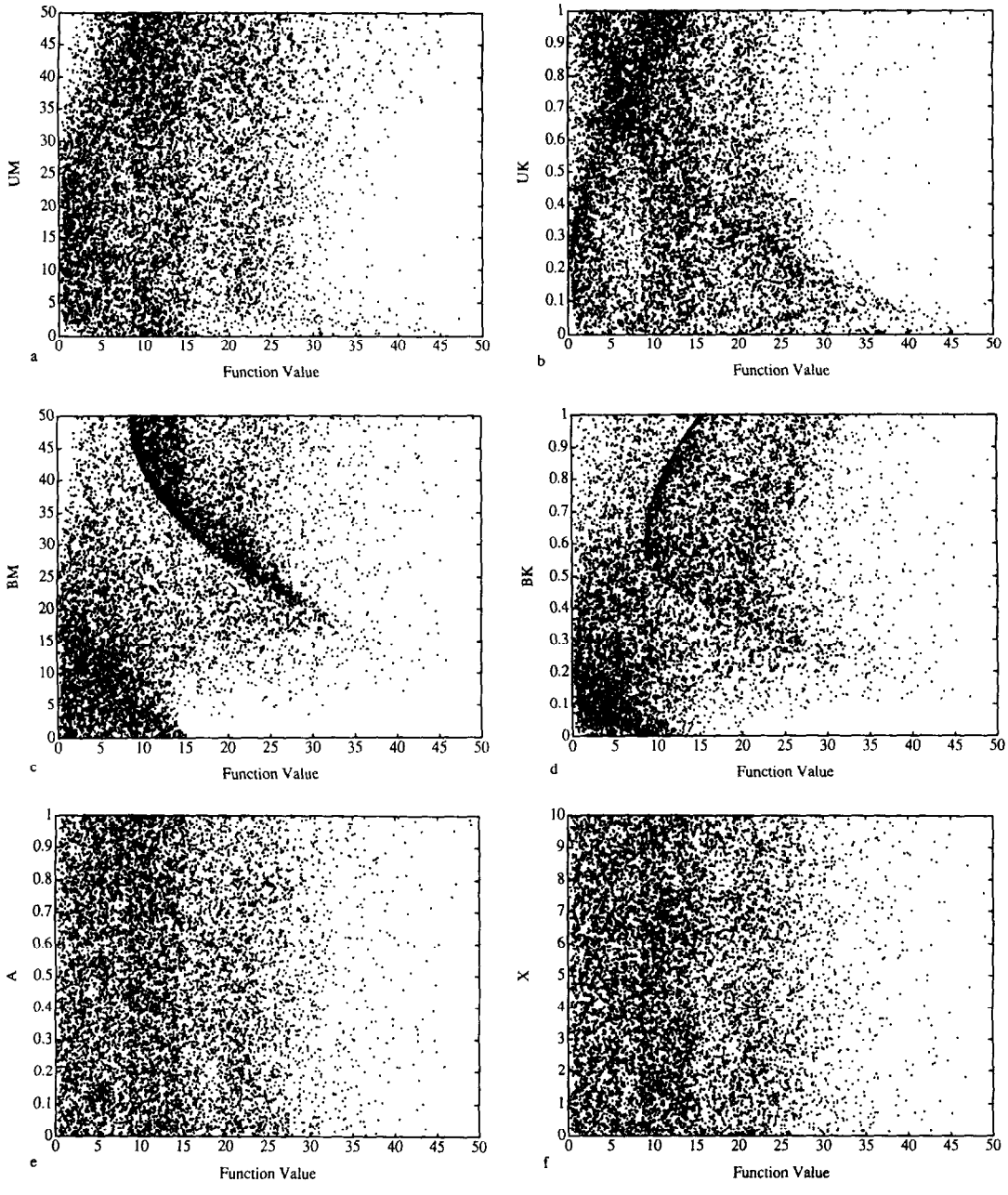

Fig. 7. PARV plots for the SIXPAR model: (*a*) parameter UM; (*b*) parameter UK; (*c*) parameter BM; (*d*) parameter BK; (*e*) parameter A; (*f*) parameter X.

poor sensitivity of the objective function to parameter variation over the entire feasible parameter space. The PARV plots for parameters UM and UK (Figures 7*a* and 7*b*) barely hint at the existence of a local optimum in the UM-UK subspace, located near (UM, UK) = (20, 0.25). However, the PARV plots for BM and BK (Figures 7*c* and 7*d*) seem to indicate the presence of a major local optimum in the BM-BK subspace near (BM, BK) = (50, 0.7). The PARV plots for A and X (Figures 7*e* and 7*f*) display no detectable patterns. Perhaps most disturbing is that of the two "best" points obtained from the sample, the second point (UM = 30, BM = 0.56, UK = 0.23, BK = 0.75, A = 0.76, X = 9.3) is far from the "true" values (UM = 10, BM = 20, UK = 0.5, BK = 0.2, A = 0.31, X = 3), while having a function value virtually indistinguishable from the first. It is not clear if this result is due to insufficient coverage of the feasible space, model structural factors, the characteristics of the data set, or some combination of the three. Further investigation of this result is desirable, but beyond the scope of this paper. However, because it is known that the sensitivity of the model output to the percolation parameters is small relative to that of the other parameters of the model, and that a significant degree of parameter interdependence exists [*Gupta and Sorooshian*, 1983], the study was repeated with the percolation parameters A and X fixed at their true values. There was no significant change in the results.

Exhaustive Gridding (EG) Method

The method of exhaustive gridding is a primitive, deterministic approach to global optimization. In this method, the

TABLE 1. Number of Local Optima in Two-Dimensional Subspace of the SIXPAR Model

	No Noise		25% Homoscedastic Error		25% Heteroscedastic Error	
Parameters	SLS	HMLE	SLS	HMLE	SLS	HMLE
UM-UK	1	2	1	1	4	2
UM-BM	2	2	7	4	11	7
UM-BK	6	2	13	31	36	3
UM-A	9	33	16	14	33	61
UM-X	42	55	62	58	77	109
UK-BM	1	1	1	1	5	5
UK-BK	3	1	1	3	2	2
UK-A	3	4	7	9	9	13
UK-X	1	1	2	4	4	1
BM-BK	2	2	6	8	11	3
BK-A	3	4	7	4	5	5
BM-X	1	1	4	5	2	5
BK-A	8	3	108	100	105	5
BK-X	2	1	102	94	101	3
A-X	11	11	113	71	127	71

number of discretization units for each parameter range is chosen, thereby specifying the number of grid points and their location, and the function value is computed at each grid point. The function value at each point is compared with those of all immediate neighboring points. The location and function value of those points for which the function value is less than or equal to that of all its neighbors are recorded. These points are part of either a global or a local optimum.

The EG method is computationally very intensive. For example, a two-dimensional problem with a grid size of 100 × 100 requires 10,000 function evaluations and 78,804 function comparisons; the method is clearly not an efficient procedure for finding the global minimum. The EG method was applied to the SIXPAR model in two- and three-dimensional subspaces only. Higher-dimensional subspaces were not studied because the computational burden became too large, even for the Convex C240 used in this study.

The two-dimensional subspaces were examined first. For each pair of parameters, a grid size of 100 × 100 was used. In addition to the case of "perfect" data (no streamflow error), the influence of two types of streamflow data error was examined: homoscedastic error (error variance does not change with the magnitude of streamflow) and heteroscedastic error (error variance changes with the magnitude of streamflow). The homoscedastic error variance value used was 25% of the mean streamflow value, while the heteroscedastic error variance used was 25% of the true streamflow value. In addition to the SLS objective function, the heteroscedastic maximum likelihood error (HMLE) objective function developed by Sorooshian [see *Sorooshian et al.*, 1983] was used so that the properties of the two objective functions could be compared.

The number of local optima found in each parameter subspace, for each objective function and for each error case, is presented in Table 1. Some of the subspaces contain quite a large number of local optima. In particular, the percolation parameters A and X are associated with larger numbers of local optima than the other four parameters. There are many local optima present on the response surface even when the data are not corrupted with errors; the introduction of errors into the streamflow data generally increases the number of local optima. When the error is heteroscedastic, the choice of objective function influences the number of optima in each subspace. The number of local optima in the UM-BK subspace for 12 different random sequences of homoscedastic streamflow data error is given in Table 2. From these results it can be seen that the randomness in the data error influences significantly the behavior of the response function, with the number of local optima varying between 25 and 62 for this case.

The objective function values obtained at each grid point were used to construct mesh surface plots. These plots show the comparative sensitivities of the parameters, the regions of roughness of the response surface, and the locations of abrupt changes in slope. Five selected surfaces are presented in Figures 8*a*–8*d* and Figure 9*a*. The locations of the local optima are indicated by artificially introducing vertical spikes on the surfaces. The plots show that the response surface can be quite steep when far from the global optimum, but is relatively insensitive to the parameters near the global optimum. Many local optima appear near the edges of the parameter bounds, while others are scattered in the general region of the global optimum. A distinct major region of attraction in the BM-BK subspace (this feature was also detected in the URS study) is clearly shown in Figure 8*c*. Note that, in addition to isolated stationary points, many of the local optima appear in clusters. The response surface in the UM-UK parameter subspace as well as three of the derivative surfaces are illustrated in Figures 9*a*–9*d*. The

TABLE 2. The Effect of Randomness on the Number of Local Optima in UM-BK Parameter Subspace

Seed Number (Value)	Number of Local Optima
1 (117)	42
2 (128)	37
3 (240)	27
4 (281)	35
5 (290)	39
6 (304)	31
7 (570)	25
8 (574)	62
9 (615)	40
10 (679)	43
11 (741)	40
12 (789)	33

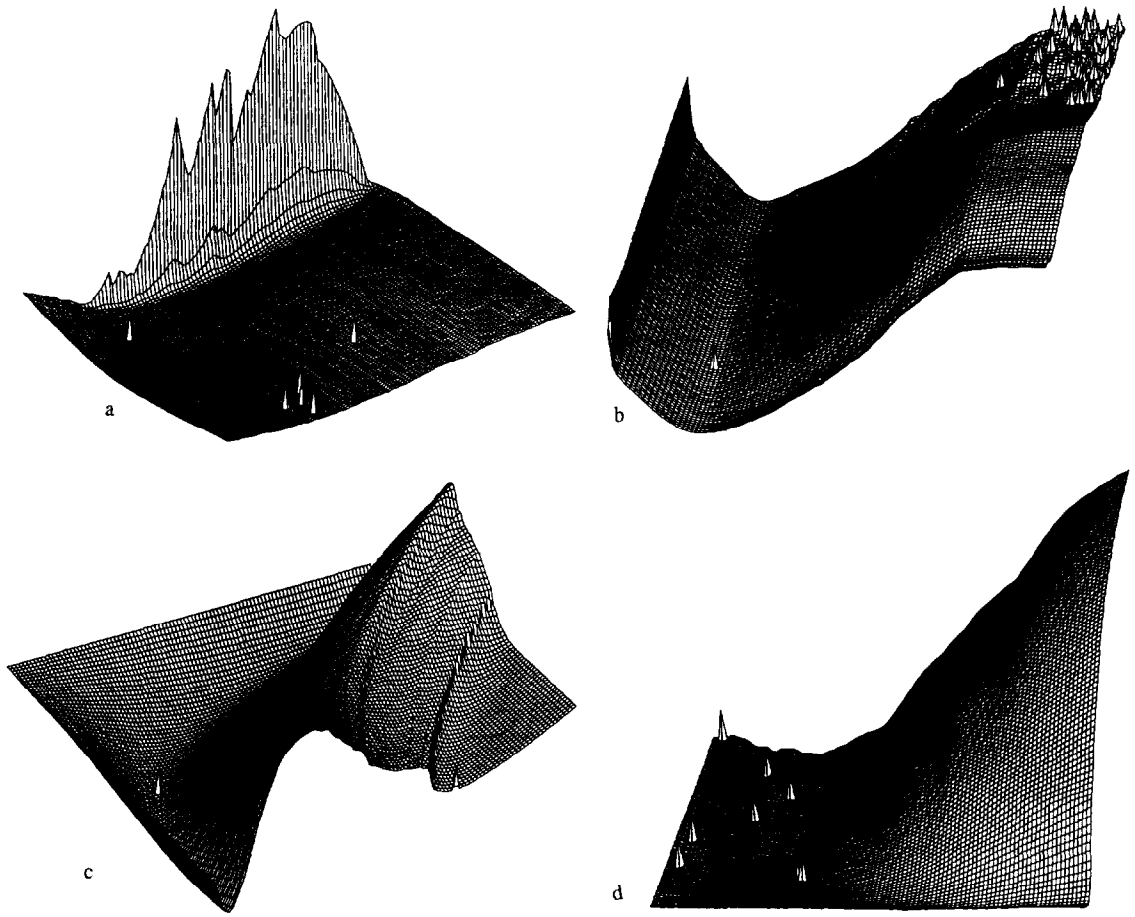

Fig. 8. Mesh surface plots for the SIXPAR model: (*a*) parameter subspace UM-BK; (*b*) parameter subspace UM-*X*; (*c*) parameter subspace BM-BK; (*d*) parameter subspace *A*-*X*.

derivatives vary in a discontinuous manner over the feasible space. This casts considerable doubt on the possibility of successfully using derivative-based (Newton or quasi-Newton type) search algorithms for parameter estimation.

The three-dimensional subspaces were examined next. A grid size of 100 × 100 × 100 was used to search each three-dimensional parameter space for local optima using the SLS objective function. The results, presented in Table 3, indicate that the number of local optima in the three-dimensional subspaces reaches as many as 812. This clearly does not bode well for local search optimization procedures. The locations of the local optima in four of these three-dimensional subspaces are displayed in Figures 10*a*–10*d*. The large circles indicate the locations of the optima in the three-dimensional space, while the small circles indicate their projections into the respective two-dimensional subspaces in which the parameter axes lie. The plots display a mixture of the following three basic patterns to the distribution of local optima: (1) sparsely scattered local optima (see Figure 10*d*); (2) dense clusters of local optima (see Figures 10*a*–10*c*); and (3) line optima (see Figures 10*c* and 10*d*).

A great many of the local optima are not located close to the global optimum. In general, the subspaces associated with parameter *X* contain the greatest number of local optima. Long ridges are also most often observed in the direction of parameter *X*.

Discussion

The URS and EG studies demonstrate clearly the severity of the optimization problem encountered when attempting to calibrate the CRR type model SIXPAR using local search procedures (e.g., simplex method, Rosenbrocks method, Newton-Raphson method, etc.); we believe these problems to be common to other CRR models such as the SAC-SMA. In Table 4, five major characteristics that complicate the optimization problem are identified. Perhaps the most important of these is that the structure of multiple optima exists on at least two scales. At the "large" scale, we find that there is more than one broad "region of attraction" into which a search strategy may converge. Such regions show up very clearly in the EG mesh surface plots, while the URS method is able to detect such regions in multiple parameter spaces. However, at the "small" scale, each major region of attraction contains numerous local minima (stationary points where the first derivatives are zero and the Hessian matrices are positive definite or positive semidefinite). These minor optima occur both close to and at various distances from the best solution. The minor local optima are not detectable using the URS method. They are also not normally visible on mesh surface and contour plots and can only be detected by numerical analyses of the gridded data. This may be why such phenomena have not been reported previously.

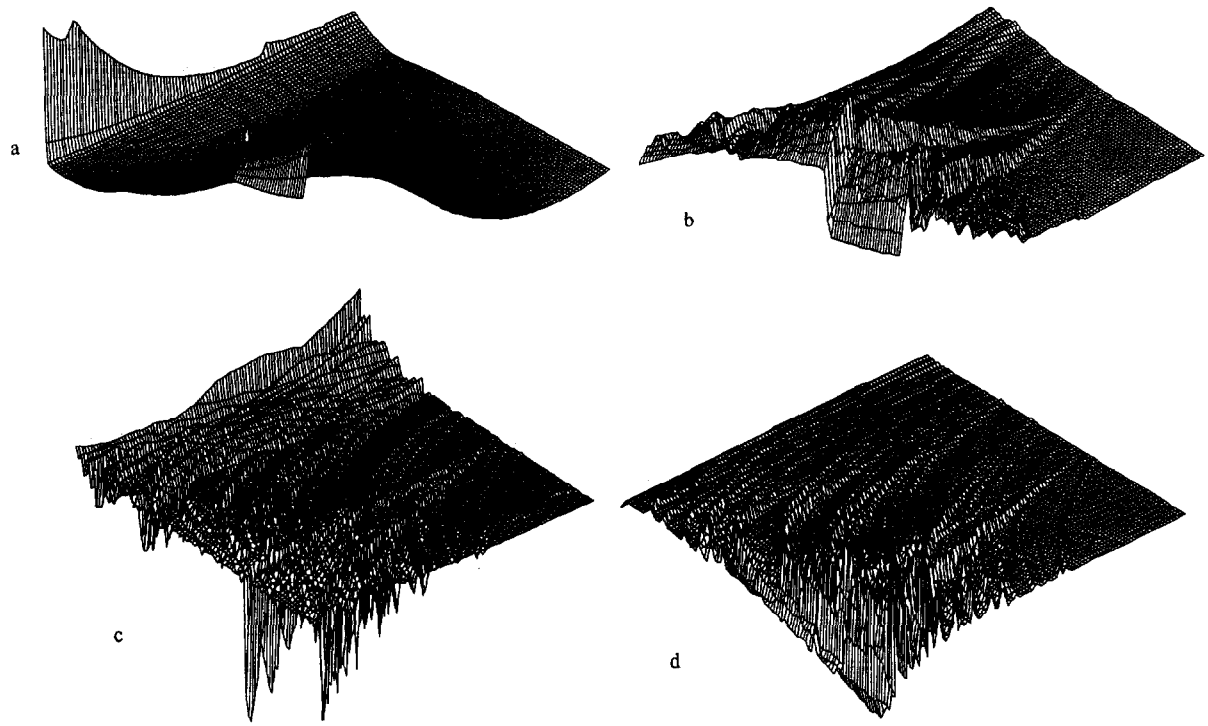

Fig. 9. Derivative surfaces in the UM-UK parameter subspace: (*a*) function surface with optimum marked; (*b*) first derivative with respect to UM; (*c*) second derivative with respect to UM; (*d*) second derivative with respect to UM and BK.

The large number of minor optima is the most probable reason for the numerous reports in the literature of the inability to find unique "optimal" parameter values. In a local search procedure, when a line optimum or a stationary point is encountered, the optimization procedure will generally stop. Small step size local searches will be unable to run the maze of minor optima, inevitably failing to reach the global optimum. Because many of the minor optima can be found quite far from the global optimum, the search may terminate without even finding an approximate solution.

TABLE 3. Number of Local Optima in Three-Dimensional Subspaces of the SIXPAR Model

Parameters	Number of Local Optima
UM-UK-BM	10
UM-UK-BK	354
UM-UK-A	122
UM-UK-X	812
UM-BM-BK	217
UM-BM-A	23
UM-BM-X	396
UM-BK-A	116
UM-BK-X	647
UM-A-X	525
UK-BM-BK	137
UK-BM-A	16
UK-BM-X	3
UK-BK-A	24
UK-BK-X	29
UK-A-X	20
BM-BK-A	72
BM-BK-X	107
BM-A-X	12
BK-A-X	227

In addition to the presence of optima at different scales, the objective function surface in the multiparameter space is not smooth and has discontinuous derivatives that vary in an unpredictable manner through the parameter space. This helps explain why derivative-based local optimization methods have not performed well. Furthermore, it indicates that, for any global optimization to be successful, it must not depend on smooth and continuous derivatives either. Finally, the response surface in the region of the global optimum is not necessarily convex, exhibits widely varying degrees of sensitivity to the model parameters, and indicates the existence of a great deal of nonlinear parameter interaction and compensation.

The combination of the five features mentioned above makes the optimization problem difficult to solve. The task, therefore, is to design an optimization procedure that is capable of dealing with these various difficulties. To deal with multiple regions of attraction, the algorithm must be globally based and possess global convergence properties. The algorithm must be able to avoid being trapped by minor optima, and it must not require the availability of explicit analytic expressions for the objective function in terms of its parameters or for the derivatives. It must be robust in the presence of parameter interaction and nonconvexity of the objective function surface. Finally, because CRR models usually involve a large number of parameters, the algorithm must be efficient in the presence of high dimensionality.

5. Evaluation of Three Existing Global Search Procedures

To date, only a few methods for finding the global solutions to multioptima problems have been developed [*Rin-*

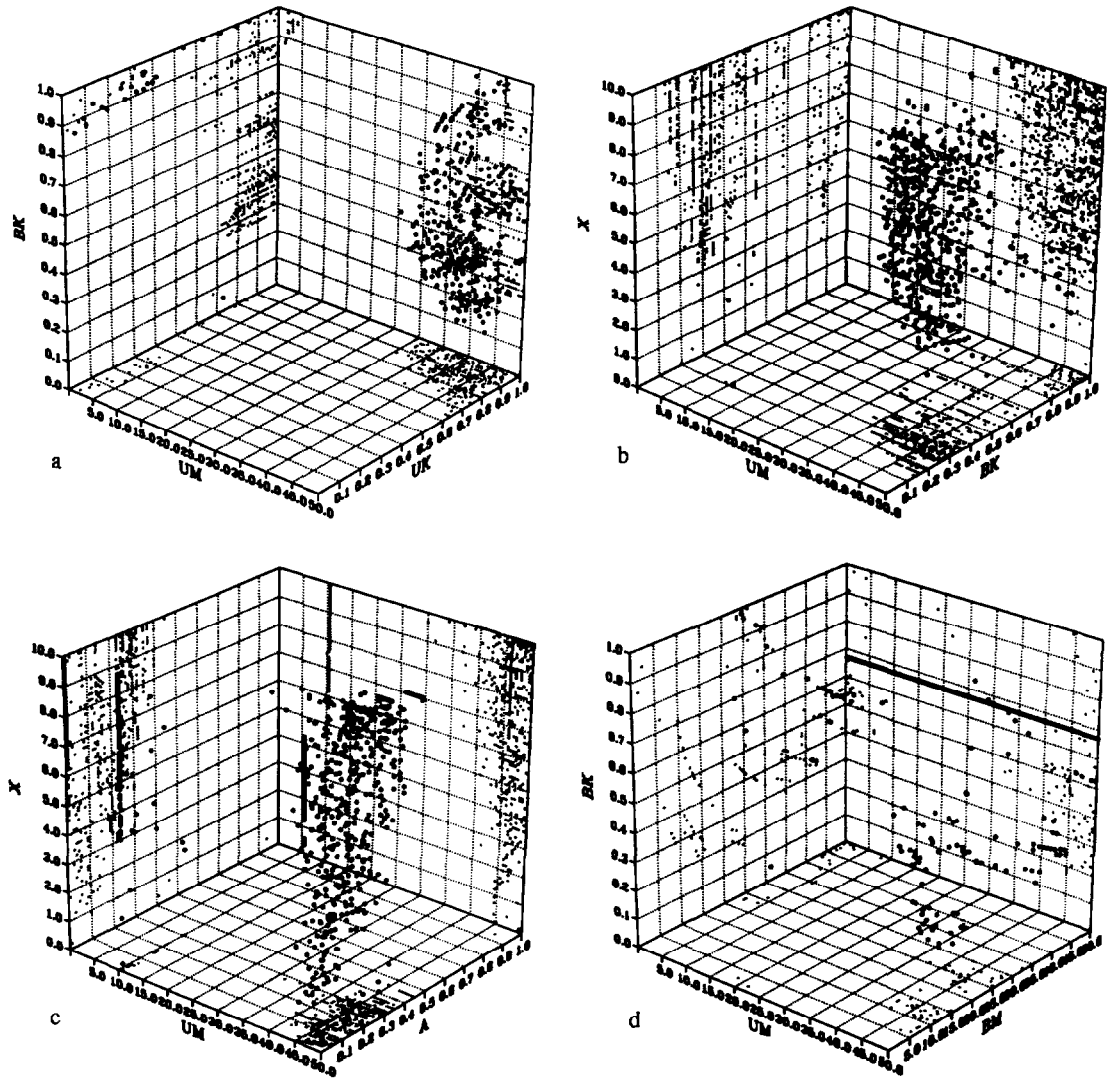

Fig. 10. Locations of local optima in three-dimensional parameter subspaces: (*a*) parameter subspace UM-UK-BK; (*b*) parameter subspace UM-BK-*X*; (*c*) parameter subspace UM-*A*-*X*; (*d*) parameter subspace UM-BM-BK.

nooy Kan and Timmer, 1989] in comparison to the multitude of methods that aim for a local optimum. In general, these methods may be classified as deterministic or probabilistic. A detailed review of the advantages and disadvantages of various global strategies appears in the work by *Duan* [1992]; therefore only a brief summary is provided here.

In essence, deterministic methods can provide a rigid guarantee of success, but they do so at the expense requiring that the function satisfy certain restrictive conditions (e.g., continuity, differentiability to second order, Lipschitz condition, etc.) that cannot be guaranteed for CRR models. Further, they are typically inefficient [*Torn and Zilinskas*, 1989] and slow in converging to the optimum. Probabilistic methods, on the other hand, involve the evaluation of the function 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. While probabilistic methods can guarantee convergence only in a probabilistic sense, they are quite efficient in practice and have the major advantage that they do not usually impose restrictive conditions on the nature of the function. Because many probabilistic methods can be employed when the function is discontinuous and when derivative information is difficult or impossible to obtain, this makes them natural candidates for the optimization of CRR models. Therefore, we chose not to investigate deterministic global optimization methods for CRR model calibration, and all the algorithms discussed in this paper belong to the class of probabilistic global optimization methods.

Evaluation of the Adaptive Random Search Method

As mentioned in section 4, the uniform random sampling approach is a primitive probabilistic optimization method. The URS approach does not use any of the information (regarding the nature of the response surface) gained during sampling to direct the search in a logical fashion. For this reason, several strategies have been proposed to guide the random search adaptively toward the region of the global optimum. One such strategy is the "adaptive random search (ARS)" method proposed by *Masri et al.* [1978] and modified by *Pronzato et al.* [1984], which was used by *Brazil and Krajewski* [1987] for parameter fine tuning of the SAC-SMA model as part of a multilevel calibration strategy. To imple-

TABLE 4. Summary of the Five Major Characteristics Complicating the Optimization Problem in CRR Model Calibration

Characteristic	Reason for Complication
1. Regions of attraction	more than one main convergence region
2. Minor local optima	many small "pits" in each region
3. Roughness	rough response surface with discontinuous derivatives
4. Sensitivity	poor and varying sensitivity of response surface in region of optimum, and nonlinear parameter interaction
5. Shape	nonconvex response surface with long curved ridges

ment the ARS method it is first required that the user select a portion of the parameter space in which the automatic search is to be conducted; we shall call this the "feasible space." For instance, the feasible space may be defined by specifying upper and lower bounds on each of the parameters. Such bounds can be estimated by analysis of the hydrologic data, from knowledge of the physiographic characteristics of the watershed and by manual calibration procedures; *Brazil* [1988] describes some procedures that have been developed for the SAC-SMA model. The ARS strategy is as follows:

1. Choose a focal point (for example, this can be the "best" point obtained in the preliminary process of defining the parameter space, or it can be some arbitrary point such as the centroid of the feasible space).

2. Generate a set of N points randomly distributed in the entire feasible space (for example, according to a uniform or normal distribution) and centered on the focal point. Store the location of the point with the best function value.

3. Repeat step 2 a prespecified number of times, on the ith time using the initial parameter range divided by 10^i and centered on the focal point (Figure 11) to restrict the search space. Each time, store the location of the point with the best function value.

4. Compare all the stored points and determine the point with the best function value. Redefine this point to be the new focal point. Record in which range level this point was found.

5. Repeat steps 2–4 until the best point is found in the smallest range level a user-specified successive number of times (say three). This point is chosen as the optimal parameter set.

Reports in the optimization literature indicate that the ARS strategy works well in practice. We therefore conducted detailed tests of the ARS strategy on the SIXPAR model, using the following experimental approach. The centroid of the feasible parameter space was used as the initial focal point. One hundred independent optimization runs were conducted using the ARS algorithm, each using a different, randomly selected initial seed value; this consists of 100 statistically independent trials of the algorithm. We know from experience with the SIXPAR model that a function value of 10^{-3} guarantees that the "true" parameter values have been found to within a few decimal places. Therefore, for all of the studies reported in this paper, we have defined an optimization run to be a success as soon as a function value of 10^{-3} is achieved. The number of successes out of 100 is a measure of the "effectiveness" of the algorithm, and the average number of function evaluations taken over all successes is a measure of the "efficiency" of the algorithm. A uniform sampling distribution was used, the number of range levels used was four, and the test was repeated for N (number of sample points in each range level) equal to 10, 100, and 200.

The best results were obtained for the smallest value of N ($N = 10$; this is consistent with Masri's findings [*Masri et al.*, 1978]), and the results for this case are presented in Figure 12. In this figure, a point closer to the origin indicates one which is relatively more effective and more efficient than one farther away. The solid line indicates how the performance of the ARS algorithm varies with the maximum possible number of function evaluations. The algorithm is clearly neither effective nor efficient on this problem, with the best result being an approximately 30% success rate (70 failures out of 100) at 25,000 function evaluations. The manner in which the best function value improved during the search for each of the 100 optimization runs is displayed in Figure 13; the circles indicate the individual values, the solid line indicates the average taken over all 100 values, the dash-dot lines indicate the maximum and minimum values, and the dotted line indicates the function convergence criterion used. The corresponding parameter values are displayed in a similar fashion in Figure 14 (the dotted line indicates the "true" parameter values). The plots indicate that after about 1000–5000 function evaluations, the marginal benefit of further sampling is quite small. This is not surprising, since the probability of finding a lower function value diminishes as the search proceeds. The parameter value plots show that the ARS algorithm is only successful at improving the estimated range (in a statistical sense) of the threshold parameters UM and BM while having little impact on the others. Given what we now know about the nature of the response surface (see the earlier section on the problem of multiple optima), the poor performance of the ARS algorithm is easy to understand.

Evaluation of a Combined ARS/Simplex Method

It has been suggested by both *Ibbitt* [1970] and *Brazil* [1988] that a random search strategy could be used to obtain

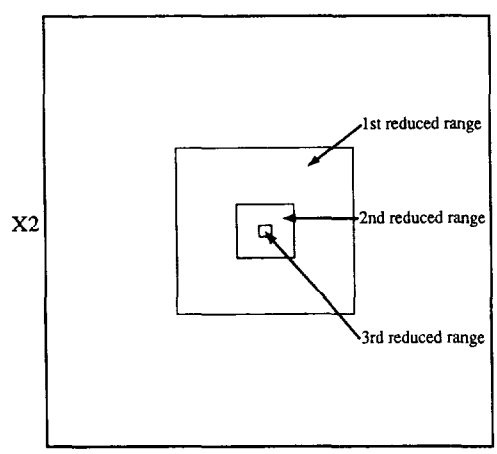

Fig. 11. Method for reducing parameter search range in the ARS algorithm.

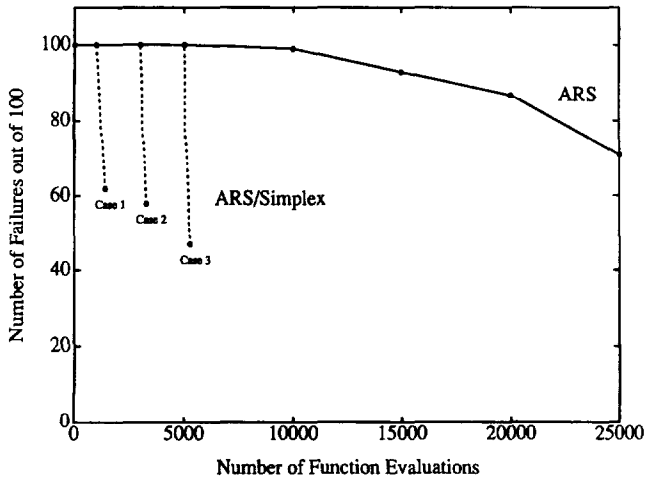

Fig. 12. Effectiveness versus efficiency plot for the ARS and ARS/simplex algorithms (solid line, ARS; dashed line, ARS/simplex).

an initial point from which a local search procedure could then be started. Hence, we tested three cases of a combined global-local ARS/simplex search strategy. In the first case, for each of the 100 independent runs of the ARS algorithm, the best point obtained after 1000 function evaluations was used as the starting point for the "simplex" local search optimization method [*Nelder and Mead*, 1965]. Similarly, cases 2 and 3 used the best points obtained after 3000 and 5000 function evaluations, respectively, of the ARS. The simplex method was selected because of its abilities to adaptively adjust its shape and size to the response surface and to not be easily trapped by minor optima, and because it has been used previously for CRR model calibration [see *Johnston and Pilgrim*, 1976; *Sorooshian and Dracup*, 1980; *Sorooshian and Arfi*, 1982]. The procedure of the simplex method has been described in the above references and will not be repeated here. Each initial simplex used the ARS point as one of its vertices, with the other points generated using perturbations in each parameter direction of ±5% of the parameter range.

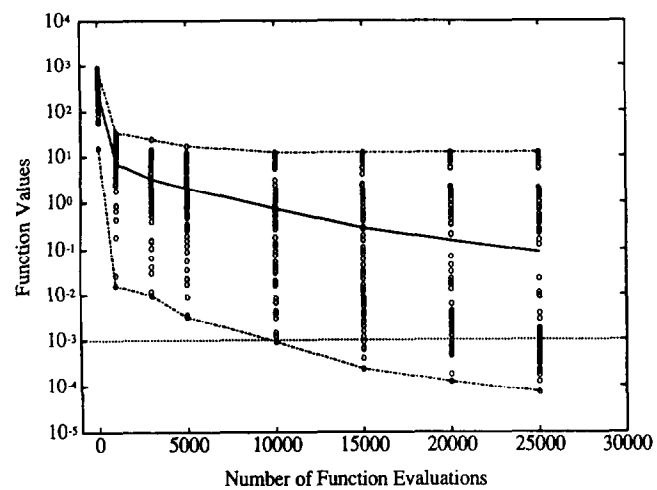

Fig. 13. Function values plotted against number of function evaluations for each of the 100 runs of the ARS algorithm.

The results are presented in Figure 12; the solid line shows the performance of the ARS method alone, and the dashed lines show the performance of the three cases of the ARS/simplex method. In each case the failure rate has reduced considerably (from 100% to 62% for case 1, from 100% to 58% for case 2, and from 100% to 45% in case 3) with only marginal increases in the average number of function evaluations; however the failure rate is still high. Figure 15 shows the beginning and ending parameter values for the 100 simplex searches started after 5000 ARS function evaluations; while 55 of the 100 simplex local searches terminated at the global optimum (horizontal dotted line), there has been no significant reduction in the distribution of the parameter values for the remaining 45 optimization runs.

Evaluation of a Multistart Simplex Procedure

A method for dealing with multiple optima that has been suggested in the hydrologic literature [e.g., *Johnston and Pilgrim*, 1976] is to run several trials of a local search optimization method from different starting points in the feasible space. The validity of such an approach can be demonstrated by the following arguments. To have good confidence in the results of any probabilistic optimization procedure, we require that it have a relatively small failure probability on the problem of interest. Let us say that we run a procedure once on a problem beginning from some randomly selected location in the feasible search space, and the probability of failure is P_f (out of 100 independent tests of the method, we expect that $100 \times P_f$ of them will fail). If we then rerun the procedure r times from r independent randomly selected locations, the overall failure probability will decrease according to the equation $P_f(r) = P_f(1)^r$ and tend to zero as r becomes large. If P_f is 0.65 (65 failures out of 100), then r equal to 12 will give a failure rate of less than 1 in 100. This simple global search strategy of repeating the search from many different locations is called a "multistart" procedure.

The efficiency of the multistart procedure varies nonlinearity with P_f, so that the number of restarts r required to achieve an overall failure probability of $P_f(r)$ is given by $r = \ln(P_f(r))/\ln(P_f(1))$. The curve of r versus $P_f(1)$ is plotted in Figure 16 for the cases of $P_f(r)$ equal to 0.01 (one failure in 100) and 0.05 (five failures in 100). Clearly, for single-start failure probabilities $P_f(1)$ of less than approximately 0.8 (80 failures in 100), we do not require a very large number of restarts. However, as $P_f(1)$ increases above 0.8 toward 1.0, the number of restarts required rapidly increases toward infinity, making the procedure impractical.

We tested 100 independent optimization runs of a multistart strategy based on the simplex local search procedure, on the SIXPAR model. The initial simplex for each run was created by randomly selecting seven points (number of parameters plus 1) in the feasible space. The results of the multistart simplex (MSX), presented in Figure 17, show that from a single-start failure probability of approximately 65%, the failure rate with 12 restarts falls to 1 in 100. This result is very encouraging, in view of the difficult nature of the problem. However, the number of function evaluations required to achieve this feat (approximately 10,500) is still high. The more complex SAC-SMA model, when calibrated to several years of data, requires quite significant amounts of computer time for even one function evaluation (approxi-

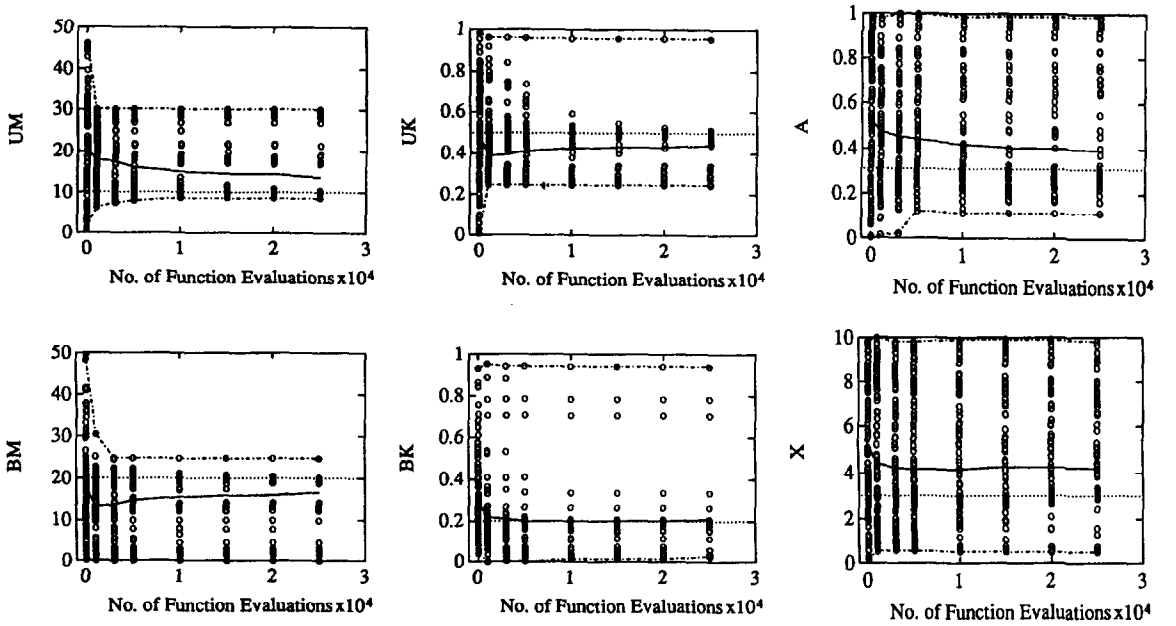

Fig. 14. Parameter value plotted against number of function evaluations for the 100 runs of the ARS algorithm.

mately 5 CPU seconds for 7 years of daily streamflow and 6-hourly precipitation calibration data on the Convex 240 supercomputer), and 10,000 evaluations would require considerable computational resources. Moreover, due to the fact that the SAC-SMA model has many more optimizable parameters than the SIXPAR model, the initial failure probability $P_f(1)$ is likely to be much higher than 0.65, so that the number of restarts required would be much larger than 12. It is, therefore, desirable that the efficiency of the search procedure be improved.

6. The Shuffled Complex Evolution Global Optimization Method

A strategy based on the use of multiple simplexes started from random locations of the search space has certain desirable properties that enable it to overcome the various difficulties encountered on the response surface of the SIXPAR model. However, it is easy to see that a source of inefficiency in the method is that each simplex search operates completely independently, with no sharing of infor-

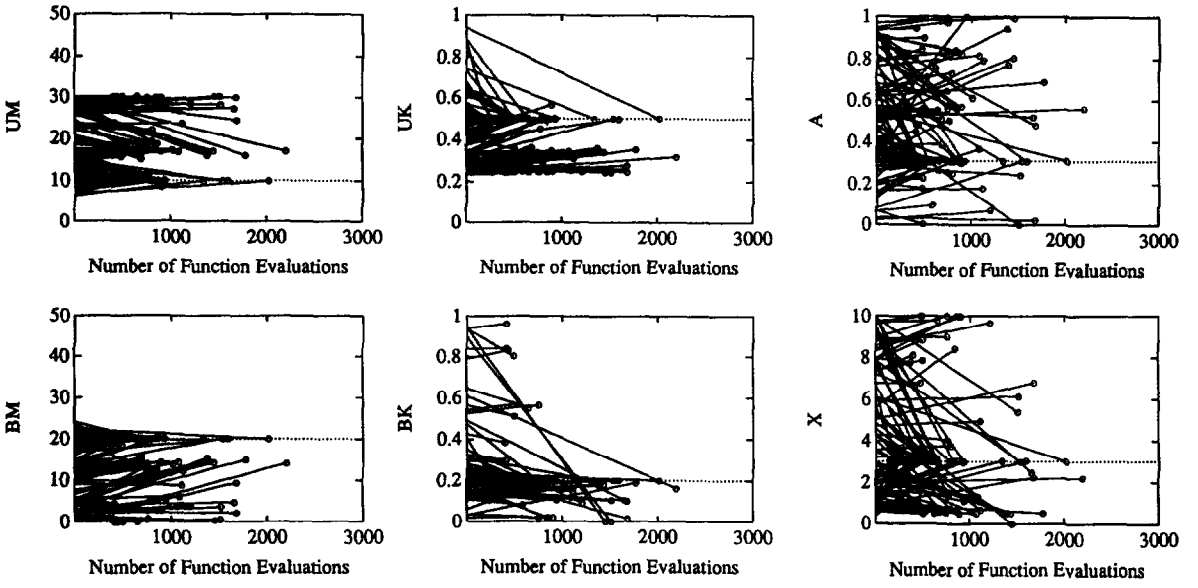

Fig. 15. Parameter value plotted against number of function evaluations for the 100 runs of the simplex algorithm started after 5000 function evaluations of the ARS algorithm.

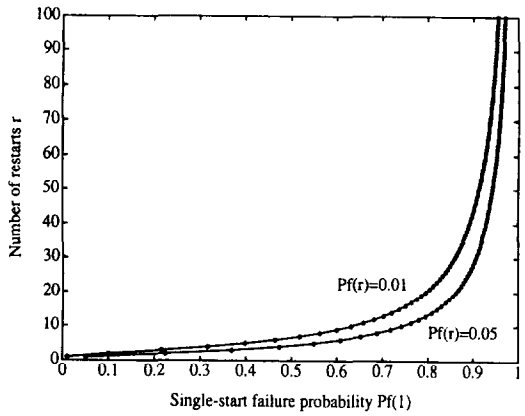

Fig. 16. Theoretical performance curve for any multistart algorithm.

mation. This is analogous to giving the same difficult problem to 12 identically capable people and asking them to solve it without conferring with each other. A more efficient strategy would clearly be for them to spend some time working independently or in small groups and to get together now and then to share information about their progress.

Motivated by the nature of the CRR model optimization problem, and based to some extent on this notion of sharing information, and on concepts drawn from principles of natural biological evolution, we have developed a global optimization strategy called the "shuffled complex evolution (SCE-UA)" method [*Duan et al.*, 1992] (computer code available from the authors). The essence of the method is as follows. We begin with a population of points sampled randomly from the feasible space. The population is partitioned into several communities, each containing $2n + 1$ points where n is the dimension of the problem. Each community is made to evolve based on a statistical "reproduction" process that uses the "simplex" geometric shape to direct the search in an improvement direction. At periodic stages in the evolution, the entire population is shuffled and points are reassigned to communities to ensure information sharing. As the search progresses, the entire population tends to converge toward the neighborhood of global optimum, provided the initial population size is sufficiently large. The algorithm is presented and discussed in more detail below.

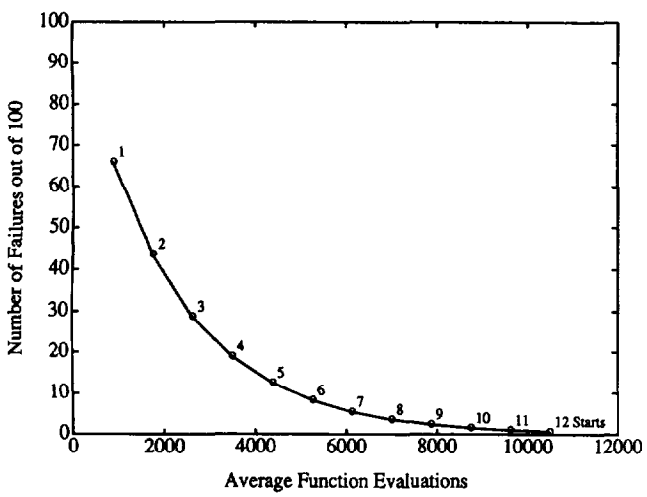

Fig. 17. Effectiveness versus efficiency plot for the MSX algorithm.

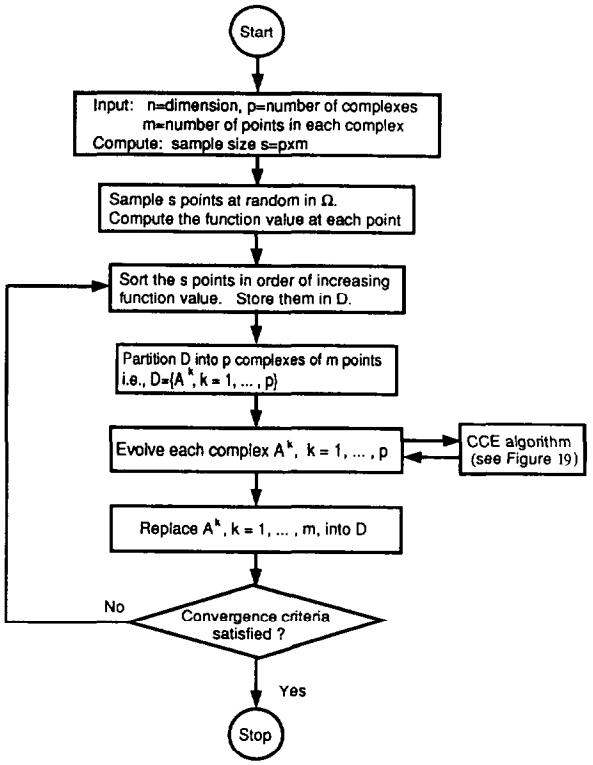

Fig. 18. Flowchart of the SCE-UA algorithm.

The Shuffled Complex Evolution Algorithm

The shuffled complex evolution (SCE-UA) method is a new global optimization strategy designed to be effective and efficient for a broad class of problems. The SCE-UA strategy combines the strengths of the simplex procedure of *Nelder and Mead* [1965] with the concepts of controlled random search [*Price*, 1987], competitive evolution [*Holland*, 1975] and the newly developed concept of complex shuffling. The SCE-UA strategy is presented below and is illustrated in Figure 18.

1. To initialize the process, select $p \geq 1$ and $m \geq n + 1$, where p is the number of complexes, m is the number of points in each complex, and n is the dimension of the problem. Compute the sample size $s = pm$.

2. Then generate a sample as follows. Sample s points $x_1, \cdots, x_s$ in the feasible space $\Omega \subset \mathbb{R}^n$. Compute the function value f_i at each point x_i. In the absence of prior information, use a uniform sampling distribution.

3. Rank the points as follows. Sort the s points in order of increasing function value. Store them in an array $D = \{x_i, f_i, i = 1, \cdots, s\}$, so that $i = 1$ represents the point with the smallest function value.

4. Partition D into p complexes $A^1, \cdots, A^p$, each containing m points, such that $A^k = \{x_j^k, f_j^k | x_j^k = x_{[\cdot]}, f_j^k = f_{[\cdot]}, j = 1, \cdots, m\}$.

5. Evolve each complex A^k, $k = 1, \cdots, p$, according

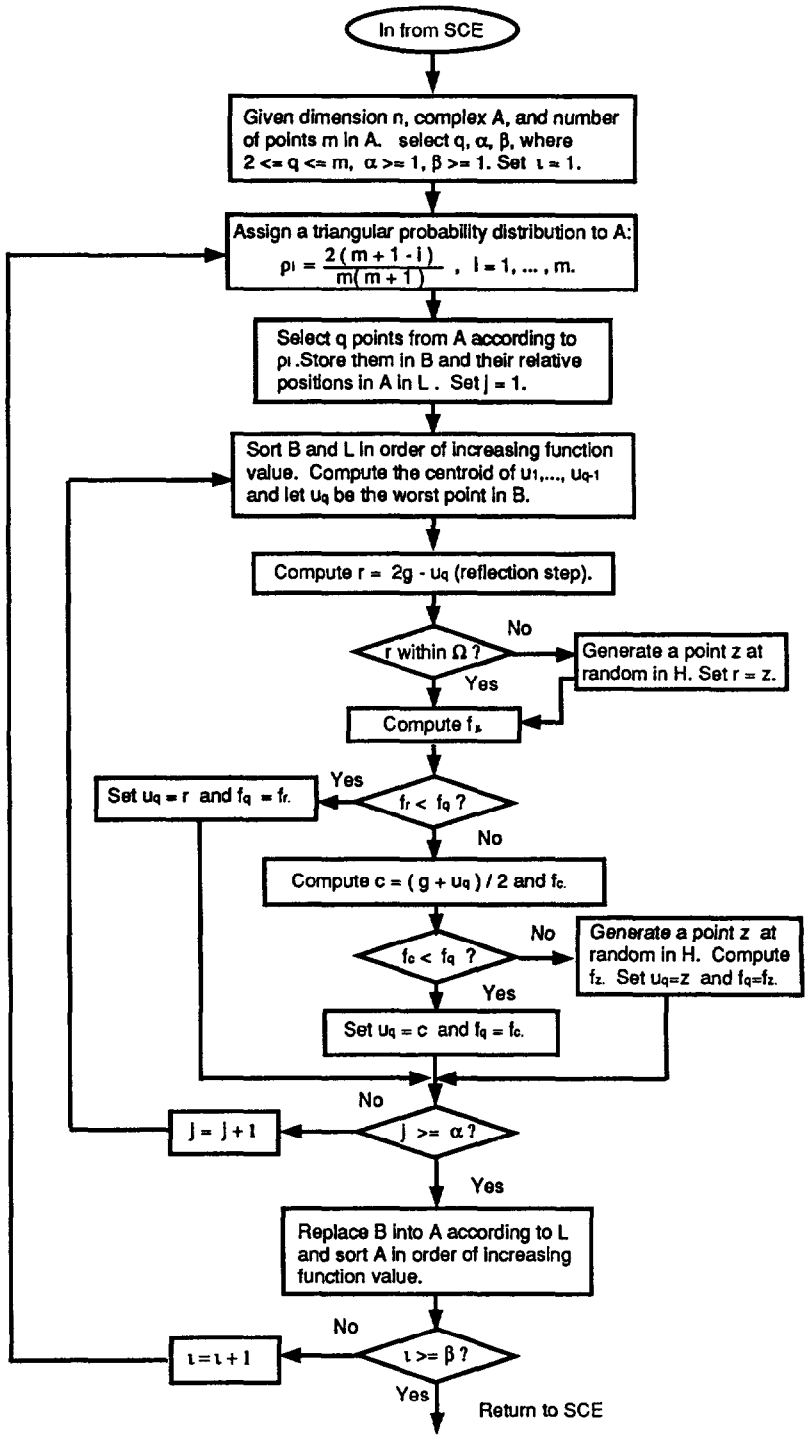

Fig. 19. Flowchart of the CCE strategy of the SCE-UA algorithm.

to the competitive complex evolution (CCE) algorithm outlined separately.

6. Shuffle the complexes as follows. Replace $A^1, \cdots, A^p$ into D, such that $D = \{A^k, k = 1, \cdots, p\}$. Sort D in order of increasing function value.

7. Check convergence. If the convergence criteria are satisfied, stop; otherwise, return to step 4.

The competitive complex evolution (CCE) algorithm required for the evolution of each complex in step 5 of the shuffled complex evolution method is presented below and is illustrated in Figure 19:

1. To initialize the process, select q, α, and β, where $2 \leq q \leq m$, $\alpha \geq 1$ and $\beta \geq 1$.

2. Assign weights as follows. Assign a trapezoidal probability distribution to A^k, i.e.,

$$p_i = \frac{2(m+1-i)}{m(m+1)}, \quad i = 1, \cdots, m \tag{4}$$

The point x_1^k has the highest probability $p_1 = 2/m + 1$. The point x_m^k has the lowest probability $p_m = 2/m(m + 1)$.

3. Select parents by randomly choosing q distinct points $u_1, \cdots, u_q$ from A^k according to the probability distribution specified above (the q points define a "subcomplex"). Store them in array $B = \{u_i, v_i, i = 1, \cdots, q\}$, where v_j is the function value associated with point u_j. Store in L the locations of A^k which are used to construct B.

4. Generate offspring according to the following procedure: (a) Sort B and L so that the q points are arranged in order of increasing function value and compute the centroid g using the expression:

$$g = \frac{1}{q-1} \sum_{j=1}^{q-1} u_j \tag{3}$$

(b) Compute the new point $r = 2g - u_q$ (reflection step). (c) If r is within the feasible space Ω, compute the function value f_r and go to step d; otherwise compute the smallest hypercube $H \subset \mathbb{R}^n$ that contains A^k, randomly generate a point z within H, compute f_z, set $r = z$ and set $f_r = f_z$ (mutation step). (d) If $f_r < f_q$, replace u_q by r, go to step f; otherwise compute $c = (g + u_q)/2$ and f_c (contraction step). (e) If $f_c < f_q$, replace u_q by c, go to step f; otherwise randomly generate a point z within H and compute f_z (mutation step). Replace u_q by z. (f) Repeat steps a–e α times, where $\alpha \geq 1$ is a user-specified parameter.

5. Replace parents by offspring as follows: Replace B into A^k using the original locations stored in L. Sort A^k in order of increasing function value.

6. Iterate by repeating steps 2–5 β times, where $\beta \geq 1$ is a user-specified parameter which determines how many offspring should be generated (how far each complex should evolve).

The version of the SCE-UA algorithm used for the optimization runs reported in this paper used the values $m = (2n + 1)$, $q = (n + 1)$, $\alpha = 1$, and $\beta = (2n + 1)$. Hence, the only variable to be specified by the user is the number of complexes p.

The SCE-UA approach treats the global search as a process of natural evolution. The s sampled points constitute a population. The population is partitioned into several communities (complexes), each of which is permitted to evolve independently (i.e., search the space in different directions). After a certain number of generations, the communities are mixed and new communities are formed through a process of shuffling. This procedure enhances survivability by a sharing of the information (about the search space) gained independently by each community.

Each member of a community (complex) is a potential parent with the ability to participate in a process of reproduction. A subcomplex selected from the complex is like a pair of parents, except that a subcomplex may consist of more than two members. To ensure that the evolution process is competitive, we require that the probability that "better" parents contribute to the generation of offspring is higher than that of "worse" parents. The use of a triangular probability distribution ensures this competitiveness. *Nelder and Mead*'s [1965] procedure is applied to each subcomplex to generate most of the offspring. This strategy uses the information contained in the subcomplex to direct the evolution in an improvement direction. In addition, offspring are introduced at random locations of the feasible space under certain conditions to ensure that the process of evolution does not get trapped by unpromising regions. This is analogous to mutation in response to stress that can occur in biological evolution. Each mutation also helps to increase the amount of information stored in the sample. Finally, each new offspring replaces the worst point of the current subcomplex. This ensures that every parent gets at least one chance to contribute to the reproduction process before being replaced or discarded. Thus, none of the information contained in the sample is ignored.

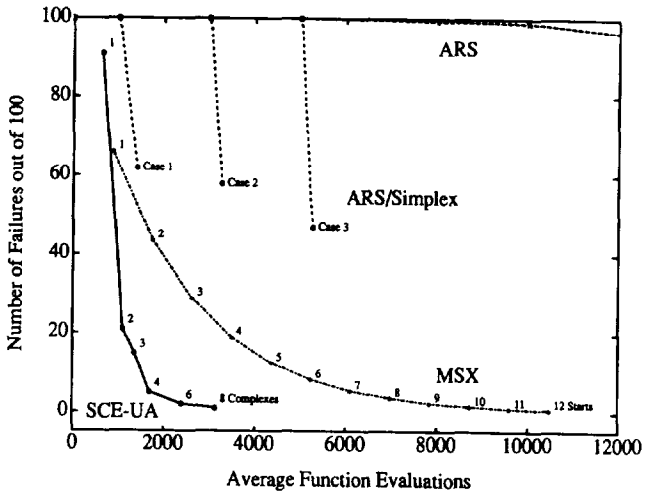

Fig. 20. Effectiveness versus efficiency plot for the SCE-UA, MSX ARS and ARS/simplex algorithms.

The processes of competitive evolution and complex shuffling inherent in the SCE-UA algorithm help to ensure that the information contained in the sample is efficiently and thoroughly exploited. They also help to ensure that the information set does not become degenerate. These properties endow the SCE-UA method with good global convergence properties over a broad range of problems. In other words, given a prespecified number of function evaluations (fixed level of efficiency), the SCE-UA method should have a high probability of succeeding in its objective of finding the global optimum.

Experimental Results

The results of 100 independent optimization runs of the SCE-UA method on the SIXPAR model are presented as the solid line in Figure 20. For purposes of comparison, the results from the other approaches (ARS, dashed line; MSX, dash-dot line, ARS-simplex, dotted line) are also included. The successive points from left to right on the SCE-UA line indicate increasing numbers of complexes (from one to eight) of 13 points. The SCE-UA method is three times more efficient than the MSX method, achieving 1 failure in 100 with an average of only 3,300 function evaluations, for an initial population size of 104 points (eight complexes of 13 points). If a 5% failure rate is considered acceptable, then the required number of function evaluations reduces to less than 2000 (using four complexes of 13 points). Figure 21 shows the beginning and ending best parameter values (connected by solid lines) for each of the 100 SCE-UA optimization runs in relation to the global optimum (horizon-

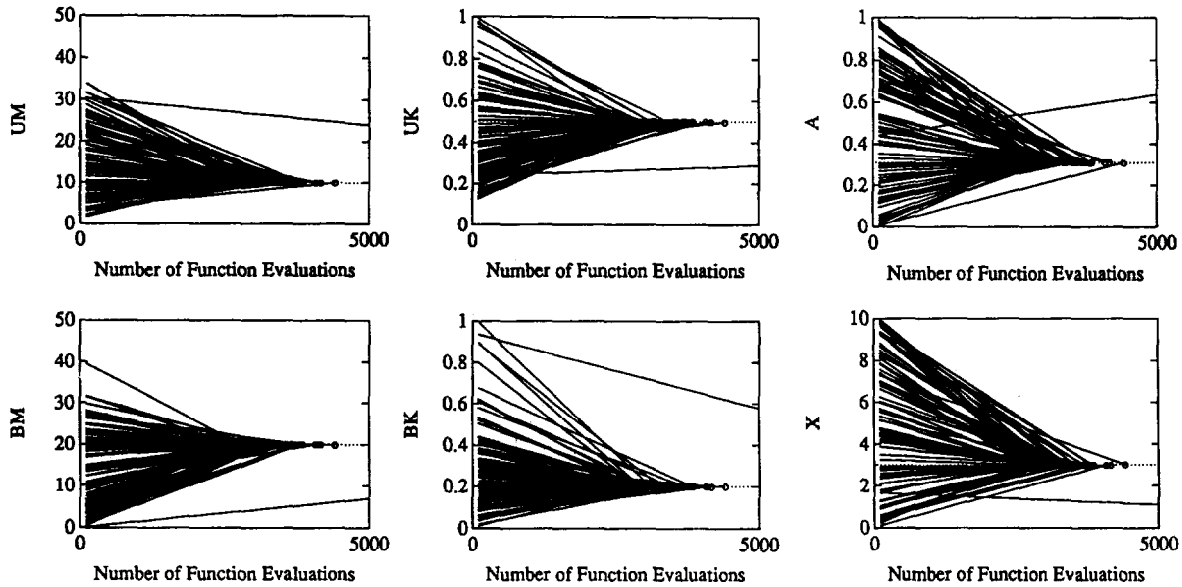

Fig. 21. Parameter value plotted against number of function evaluations for the 100 runs of the SCE-UA algorithm.

tal dotted line); 99 of the runs have terminated at the global optimum and only one has failed.

7. SUMMARY

This paper has addressed the problem of obtaining the globally optimal parameters for a simplified CRR model. In section 4 we used two graphical techniques to show that the severity of the problem of multiple optima in CRR models is more difficult than had been previously thought. We identified five major characteristics that complicate the optimization problem (Table 4). The most important of these is that there are a large number of local minima located throughout the feasible parameter space. These results indicate that currently used local search optimization procedures have a low probability of successfully finding the optimal parameter sets. In section 5 we tested the performance of three existing global search optimization procedures on the research CRR model SIXPAR. One hundred independent trials of the ARS method showed that it has an unacceptably low probability of successfully finding the global optimum. The use of a local search simplex procedure [*Nelder and Mead*, 1965] to augment the ARS global search yielded better results, but still had a high probability of failure. In contrast, a multistart procedure based on the simplex algorithm (MSX) was able to achieve a success rate of 99 in 100, with 12 independent restarts. While the effectiveness of the MSX algorithm is high, it is not very efficient, requiring a large number of function evaluations. Finally, in section 6 we presented a new global optimization procedure, entitled the shuffled complex evolution (SCE-UA) method, which appears to be capable of efficiently and effectively identifying the optimal values for the model parameters. The SCE-UA procedure required only one third of the function evaluations needed by the MSX approach to achieve the same probability of success (99 in 100).

8. CONCLUSIONS

Attempts to calibrate CRR models, such as the SAC-SMA portion of the National Weather Service River Forecast System, have been typically unable to obtain unique optimal parameter estimates using conventional automatic calibration procedures. The inability to place a reasonable degree of confidence on the estimated parameter values translates into uncertainty regarding the accuracy of the model forecasts. Furthermore, unless the best set of parameters associated with a given calibration data set can be found, it is difficult to determine how sensitive the parameter estimates (and, hence, the model forecasts) are to factors such as input and streamflow data error, model error, quantity and quality of data, objective function used, and so on.

The reason that unique optimal parameter estimates are difficult to obtain is that the optimization procedures typically used are not powerful enough to do the job. The results presented in this paper suggest that, in addition to the previously known problems of parameter interaction, nonconvexity of the response surface, and discontinuous derivatives, the problem of multiple optima occurs on at least two scales. If large numbers of minor optima are indeed present on the function response surface for the SAC-SMA and other CRR models (as our experience leads us to believe), it seems likely that conventional local search optimization procedures will terminate prematurely (the simplex method was found to have a 65% probability of failure on the SIXPAR model in this study). Under such conditions, it seems reasonable to also infer that random search–based global optimization procedures such as the adaptive random search method will also fail to locate the global optimum with an acceptable probability of success.

Our results indicate that at least two optimization procedures may be capable of finding the globally optimal parameters during CRR model calibration: MSX, a multistart procedure based on the "simplex" algorithm; and SCE-UA,

the shuffled complex evolution procedure developed at the University of Arizona by the authors. The MSX algorithm is relatively easy to program and implement and can, therefore, be used when the initial failure probability $P_f(1)$ is not too large and when sufficient computer time is available. However, in the case that large amounts of data are being used in the calibration of a relatively complex CRR model, the MSX procedure may not be feasible. The shuffled complex evolution (SCE-UA) procedure appears to be about 3 times more efficient than the MSX procedure, while still being relatively easy to program and implement.

Acknowledgments. The authors gratefully acknowledge the constructive suggestions of Steve Burges and two anonymous reviewers, and the excellent editorial comments of Corla Theis. Technical support was provided by the Center for Computing Information and Technology at the University of Arizona. Partial financial support for this research was provided by the National Science Foundation (grants ECE-86105487 and BCS-8920851) and by the Hydrologic Research Laboratory of the National Weather Service (grants NA85AA-H-HY088 and NA90AA-H-HY505). Financial assistance provided to the first author by the Salt River Project Fellowship is gratefully acknowledged.

REFERENCES

Brazil, L. E., Multilevel calibration strategy for complex hydrologic simulation models, Ph.D. dissertation, Dep. of Civ. Eng., Colo. State Univ., Fort Collins, 1988.

Brazil, L. E., and M. D. Hudlow, Calibration procedures used with the National Weather Service River Forecast System, in *Water and Related Land Resource Systems*, edited by Y. Y. Haimes and J. Kindler, pp. 457–466, Pergamon, New York, 1981.

Brazil, L. E., and W. F. Krajewski, Optimization of complex hydrologic models using random search methods, paper presented at Engineering Hydrology Conference, Hydraul. Div., Am. Soc. Civ. Eng., Williamsburg, Va., Aug. 3–7, 1987.

Burnash, R. J. E., R. L. Ferral, and R. A. McGuire, A generalized streamflow simulation system, report, Joint Fed.-State River Forecast Center, Sacramento, Calif., 1973.

Chapman, T. G., Optimization of a rainfall-runoff model for an arid zone catchment, *UNESCO Publ. 96*, pp. 126–143, U. N. Educ., Sci., and Cult. Organ., Wallington, N. Z., 1970.

Crawford, N. H., and R. K. Linsley, Digital simulation in Hydrology: Stanford watershed model IV, *Tech. Rep. 39*, Dept. of Civ. Eng., Stanford Univ., Stanford, Calif., 1966.

Dawdy, D. R., and T. O'Donnell, Mathematical models of catchment behavior, *J. Hydraul. Div. Am. Soc. Civ. Eng.*, *91*(HY4), 113–137, 1965.

Duan, Q., A global optimization strategy for efficient and effective calibration of hydrologic models, Ph.D. thesis, Dep. of Hydrol. and Water Resour., Univ. of Ariz., Tucson, 1991.

Duan, Q., V. K. Gupta, and S. Sorooshian, A shuffled complex evolution approach for effective and efficient global minimization, *J. Optim. Theory Appl.*, in press, 1992.

Gan, T. Y., and S. J. Burges, An assessment of a conceptual rainfall-runoff model's ability to represent the dynamics of small hypothetical catchments, 1, Models, model properties, and experimental design, *Water Resour. Res.*, *26*(7), 1595–1604, 1990*a*.

Gan, T. Y., and S. J. Burges, An assessment of a conceptual rainfall-runoff model's ability to represent the dynamics of small hypothetical catchments, 2, Hydrologic responses for normal and extreme rainfall, *Water Resour. Res.*, *26*(7), 1605–1619, 1990*b*.

Gupta, V. K., Calibration of conceptual rainfall-runoff models: Problems caused by model structure, M.S. thesis, Syst. Eng. Dep., Case Western Reserve Univ., Cleveland, Ohio, 1982.

Gupta, V. K., and S. Sorooshian, Uniqueness and observability of conceptual rainfall-runoff model parameters: The percolation process examined, *Water Resour. Res.*, *19*(1), 269–276, 1983.

Gupta, V. K., and S. Sorooshian, The automatic calibration of conceptual catchment models using derivative-based optimization algorithms, *Water Resour. Res.*, *21*(4), 473–486, 1985.

Hendrickson, J. D., S. Sorooshian, and L. Brazil, Comparison of Newton-type and direct search algorithms for calibration of conceptual rainfall-runoff models, *Water Resour. Res.*, *24*(5), 691–700, 1988.

Holland, J. H., *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, 1975.

Ibbitt, R. P., Systematic parameter fitting for conceptual models of catchment hydrology, Ph.D dissertation, Imp. Coll. of Sci. and Technol., Univ. of London, London, England, 1970.

Johnston, P. R., and D. Pilgrim, Parameter optimization for watershed models, *Water Resour. Res.*, *12*(3), 477–486, 1976.

Karnopp, D. C., Random search technique for optimization problems, *Automatica*, *1*, 111–121, 1963.

Larimore, W. E., Applications of the maximum likelihood parameter identification to hydrologic forecasting, *Rep. TR-1480-3*, The Analytic Sciences Corporation, Reading, Mass., 1981.

Masri, S. F., G. A. Bekey, and F. B. Safford, An adaptive random search method for identification of large-scale nonlinear systems, *Identif. Syst. Parameter Estim. Proc. IFAC Symp. 4th 1976*, 246–255, 1978.

Monro, J. C., Direct search optimization in mathematical modeling and a watershed model application, *Tech. Memo NWS HYDRO-12*, Natl. Oceanic and Atmos. Admin., U.S. Dep. of Commerce, Silver Spring, Md., 1971.

Nash, J. E., and J. V. Sutcliffe, River flow forecasting through conceptual models, 1, A discussion of principles, *J. Hydrol.*, *10*(3), 282–290, 1970.

Nelder, J. A., and R. Mead, A simplex method for function minimization, *Comput. J.*, *7*(4), 308–313, 1965.

Pickup, G., Testing the efficiencies of algorithms and strategies for automatic calibration of rainfall-runoff models, *Hydrogeol. Sci. Bull.*, *22*(2), 257–274, 1977.

Price, W. L., Global optimization algorithms for a CAD workstation, *J. Optim. Theory Appl.*, *55*(1), 133–146, 1987.

Pronzato, L., E. Walter, A. Venot, and J. F. Lebruchec, A general-purpose global optimizer: Implementation and applications, *Math. Comput. Simul.*, *26*, 412–422, 1984.

Rinnooy Kan, A. H. G., and G. T. Timmer, Global optimization: A survey, in *New Methods in Optimization and Their Industrial Uses*, edited by J.-P. Penot, Birkhauser Verlag, Basel, Switzerland, 1989.

Sorooshian, S., Considerations of stochastic properties in parameter estimation of hydrologic rainfall-runoff models, Ph.D. dissertation, Univ. of Calif., Los Angeles, 1978.

Sorooshian, S., and F. Arfi, Response surface parameter sensitivity analysis methods for post-calibration studies, *Water Resour. Res.*, *18*(5), 1531–1538, 1982.

Sorooshian, S., and J. A. Dracup, Stochastic parameter estimation procedures for hydrologic rainfall-runoff models: Correlated and heteroscedastic error cases, *Water Resour. Res.*, *16*(2), 430–442, 1980.

Sorooshian, S. and V. K. Gupta, Automatic calibration of conceptual rainfall-runoff models: The question of parameter observability and uniqueness, *Water Resour. Res.*, *19*(1), 251–259, 1983.

Sorooshian, S., V. K. Gupta, and J. L. Fulton, Evaluation of maximum likelihood parameter estimation techniques for conceptual rainfall-runoff models: Influence of calibration data variability and length on model credibility, *Water Resour. Res.*, *19*(1), 251–259, 1983.

Torn, A., and A. Zilinskas, *Global Optimization*, Springer-Verlag, New York, 1989.

Wang, Q. J., The genetic algorithm and its application to calibrating conceptual rainfall-runoff models, *Water Resour. Res.*, *27*(9), 2467–2471, 1991.

Q. Duan, V. Gupta, and S. Sorooshian, Department of Hydrology and Water Resources, University of Arizona, Tucson, AZ 85721.

(Received June 24, 1991;
revised November 22, 1991;
accepted December 2, 1991.)