A Multistep Automatic Calibration Scheme for River Forecasting Models

TERRI S. HOGUE, SOROOSH SOROOSHIAN, AND HOSHIN GUPTA

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

ANDREA HOLZ AND DEAN BRAATZ

North Central River Forecast Center, NOAA/National Weather Service, Chanhassen, Minnesota

(Manuscript received 5 October 1999, in final form 17 May 2000)

ABSTRACT

Operational flood forecasting models vary in complexity, but nearly all have parameters for which values must be estimated. The traditional and widespread manual calibration approach requires considerable training and experience and is typically laborious and time consuming. Under the Advanced Hydrologic Prediction System modernization program, National Weather Service (NWS) hydrologists must produce rapid calibrations for roughly 4000 forecast points throughout the United States. The classical single-objective automatic calibration approach, although fast and objective, has not received widespread acceptance among operational hydrologists. In the work reported here, University of Arizona researchers and NWS personnel have collaborated to combine the strengths of the manual and automatic calibration strategies. The result is a multistep automatic calibration scheme (MACS) that emulates the progression of steps followed by NWS hydrologists during manual calibration and rapidly provides acceptable parameter estimates. The MACS approach was tested on six operational basins (drainage areas from 671 to 1302 km^2) in the North Central River Forecast Center (NCRFC) area. The results were found to compare favorably with the NCRFC manual calibrations in terms of both visual inspection and statistical measures, such as daily root-mean-square error and percent bias by flow group. Further, implementation of the MACS procedure requires only about 3–4 person hours per basin, in contrast to the 15–20 person hours typically required using the manual approach. Based on this study, the NCRFC has opted to perform further testing of the MACS procedure at a large number of forecast points that constitute the Grand River (Michigan) forecast group. MACS is a time-saving, reliable approach that can provide calibrations that are of comparable quality to the NCRFC's current methods.

1. Introduction and scope

The development of conceptual rainfall–runoff models for use in hydrologic predictions has been driven by several needs, one of which is operational flood forecasting (Sorooshian 1997). These conceptual models vary in complexity, but nearly all have parameters for which values must be estimated. In spite of the advances in ''physically'' based modeling, there is general agreement among scientists and practicing hydrologists that a certain level of calibration is required to obtain successful streamflow predictions. Although some parameters can be derived directly from knowledge of physical watershed characteristics, others must be adjusted or ''tuned'' to get acceptable simulations of observed streamflows.

The traditional and most widespread approach to model calibration involves ''manual'' (also called ''expert'') adjustment of the parameter values to improve the model response, based on visual inspection of the observed and simulated hydrographs. The hydrologist will typically attempt to reproduce the hydrograph peaks (amount and timing), flood volumes, recession slopes, and base flow. However, for models such as the Sacramento Soil Moisture Accounting Model (SAC-SMA; Burnash et al. 1973; Burnash 1995) and ''SNOW-17'' (Anderson 1973, 1978), used by the National Weather Service (NWS) for flood forecasting, this approach requires considerable training and experience. Further, it is typically laborious and time consuming, particularly when numerous parameters with interacting effects must be adjusted.

The NWS is currently undergoing a process of modernization, including installation of the Advanced Hydrologic Prediction System (AHPS). Under AHPS, the river forecast centers (RFCs) have been issued deadlines for the calibration of the SAC-SMA and SNOW-17 models to the roughly 4000 forecast points in this system (J. Ingram 1996, personal communication). The North

Corresponding author address: Terri S. Hogue, Dept. of Hydrology and Water Resources, Bldg. 11, Room 122, The University of Arizona, Tucson, AZ 85721-0011.
E-mail: hoguets@hwr.arizona.edu

Central RFC (NCRFC), one of the NWS's largest river forecast centers, is in the process of calibrating these models for approximately 900 basins. To help in meeting these deadlines, the NCRFC (among others) has been investigating the availability of reliable methods to reduce the calibration time.

During the past three decades, considerable research has been performed on the development of automated methods to aid the model calibration process. The classical single-objective automatic approach, based on optimization theory, requires the definition of a mathematical measure (an objective function such as least squares or maximum likelihood) of the differences between the observed and simulated hydrograph. An optimization algorithm is then used to adjust the parameters toward values that minimize (or maximize, if appropriate) this function. Although the method is fast and objective, it has not received widespread acceptance among operational hydrologists. For example, NWS personnel have explored the use of automatic calibration methods (with various single-objective functions) but found that visual inspection of the hydrograph reveals areas of concern, such as poor matching of recessions and unacceptable flow biases. Use of different objective functions has not helped to resolve this problem. Therefore, poor confidence in the capabilities of automatic methods has inhibited their usage at RFCs for speeding model calibration.

The work reported in this paper is an outcome of discussions between University of Arizona (UA) researchers and NWS personnel at calibration workshops and training sessions organized by the NWS Office of Hydrology (Arkansas–Red Basin RFC in September 1998, NWS Hydrologic Research Laboratory in April 1999, and Northwest RFC in July 1999). The discussions motivated interest in the development of ''hybrid'' methodologies that would combine the strengths of the manual (expert) and automatic calibration strategies. The goal is to satisfy the operational need for a calibration scheme that will rapidly provide parameter estimates that are acceptable to NWS operational hydrologists (i.e., ''first-cut'' calibrations that provide acceptable simulations at the forecast points) and that can be refined manually later if necessary. As a result of this collaborative effort, a multistep automatic calibration scheme (MACS) procedure is under development that seeks to emulate the operational progression of steps followed by NWS hydrologists during manual calibration.

This paper presents the results of a study in which the MACS approach was applied to six operational forecast points in the NCRFC forecast area. The procedure compares favorably with the manual calibrations conducted by NCRFC hydrologists. Based on this study, the NCRFC has opted to perform further testing of the MACS procedure at a large number of forecast points that constitute the Grand River (Michigan) forecast group. Although the MACS technique reported here has been developed with specific reference to the SAC-SMA and SNOW-17 models of the NWS, these models are generally representative of the class of conceptual watershed models, and the methods presented here are, therefore, generally appropriate for calibration of various hydrologic models of this type.

A background discussion on automatic calibration is presented in the next section. The study area and forecast points used for this study are specified in section 3. Details of the SAC-SMA and SNOW-17 models are described in section 4, along with a brief overview of the UA Shuffled Complex Evolution (SCE-UA) scheme. The MACS methodology is presented in detail in section 5. Results are presented in section 6, and conclusions are summarized in section 7.

2. Automatic calibration

Automatic calibration methods have evolved significantly since early endeavors reported by Dawdy and O'Donnell (1965), Nash and Sutcliffe (1970), Ibbitt (1970), Ibbitt and O'Donnell (1971), Monroe (1971), and Johnston and Pilgrim (1976). During the last two decades, the evolution of these methods has been motivated by 1) the need to simplify and speed up the calibration process, 2) the need to assign some objectivity and confidence to the calibration process (and hence, model predictions), and 3) the lack of numerous expert calibrators available for each watershed model (Sorooshian and Gupta 1995).

Various issues that have arisen in the context of research into automatic calibration methods have included conceptually unrealistic parameter values, poor model performance on validation period (vs calibration period), and the inability of the algorithms to find a ''single'' best parameter set [see discussions by Gupta and Sorooshian (1994), Duan et al. (1993), and Gupta et al. (1998)]. More recent research has resulted in the development of global search procedures (L. E. Brazil and W. F. Krajewski 1987, personal communication; Brazil 1988; Duan et al. 1992, 1993; Sorooshian et al. 1993) and multicriteria optimization schemes (Gupta et al. 1998; Yapo et al. 1998), providing calibrators with better tools for obtaining a ''best'' parameter set for their watershed models.

The implementation of an automatic calibration procedure includes the selection of:

1) an objective function,
2) an algorithm to search the parameter space,
3) a period of historical data against which to calibrate the model, and
4) termination criteria used to determine when to stop the search.

The objective function chosen by the modeler represents the computation of a numerical measure of the difference between the simulated output and the observed (measured) values. The goal of automatic calibration is

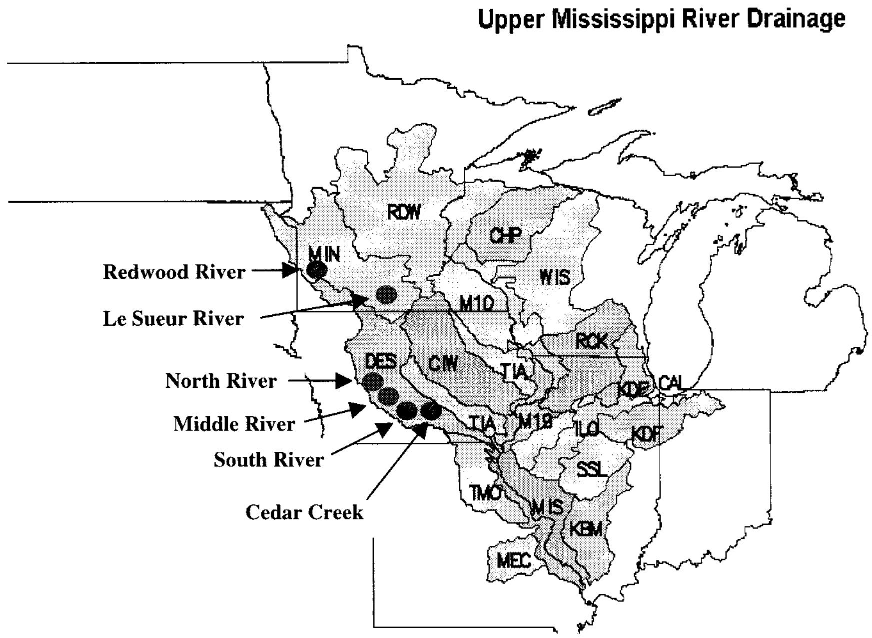

FIG. 1. NCRFC upper Mississippi River drainage with approximate locations of study basins displayed as ●.

to optimize (maximize or minimize, as appropriate) this objective function. Traditional measures used in single-objective optimization include least squares [or daily root-mean-square (DRMS)] and, more recently, maximum likelihood functions.

The selected objective function implicitly gives rise to a response surface, and an iterative optimization algorithm or search procedure is used to search this surface for an ''optimum'' point, within the user-defined constraints (usually parameter ranges). Optimization methods have been categorized as ''local search strategies,'' such as the pattern search method (Hooke and Jeeves 1961), the Rosenbrock method (Rosenbrock 1960), and the simplex method (Nelder and Mead 1965), or as ''global search strategies,'' including the random search method (Karnopp 1963), the adaptive random search (Masri et al. 1978; Pronzato et al. 1984), and, more recently, the SCE-UA method (Duan et al. 1992).

TABLE 1. Basins used in calibration study (average values based on 1948–93 data).

	Size (km^2)	Mean daily precipitation (mm day^{-1})	Mean daily flow (m^3 s^{-1} day^{-1})
Des Moines River			
South River	1192	2.385	7.275
Cedar Creek	969	2.479	6.559
Middle River	1302	2.353	7.520
North River	904	2.328	5.243
Minnesota River			
Redwood River	671	1.6987	2.838
Le Sueur River	784	2.1102	1.795

The SCE-UA method, a global search procedure based on the nonlinear simplex method, uses concepts from the random search procedures, along with the strength of the downhill simplex method.

As part of the automatic calibration process, a historical period of data against which to calibrate the model must also be selected. In selecting a calibration period, the goal is to choose a dataset representative of the various hydrologic phenomena experienced by the watershed. Research by Yapo et al. (1996) indicates that approximately 8 years of data, specifically including some of the wettest years, are adequate to ensure a quality calibration. This is consistent with the experience of NWS hydrologists.

Termination criteria are needed to determine when to stop the iterative search. Methods that have been used include parameter convergence, function convergence, or a maximum number of iterations. When an algorithm is unable to appreciably change parameter values and improve the objective function value, parameter convergence is achieved. Function convergence occurs when the algorithm is unable to improve the objective function beyond a predefined increment in one or more iterations. A calibrator also may set a maximum number of iterations to stop the search procedure, ensuring that the algorithm does not enter an endless loop.

3. Study area

This study involved collaboration between UA and NCRFC, located in Chanhassen, Minnesota. NCRFC's area of forecasting responsibility encompasses all of

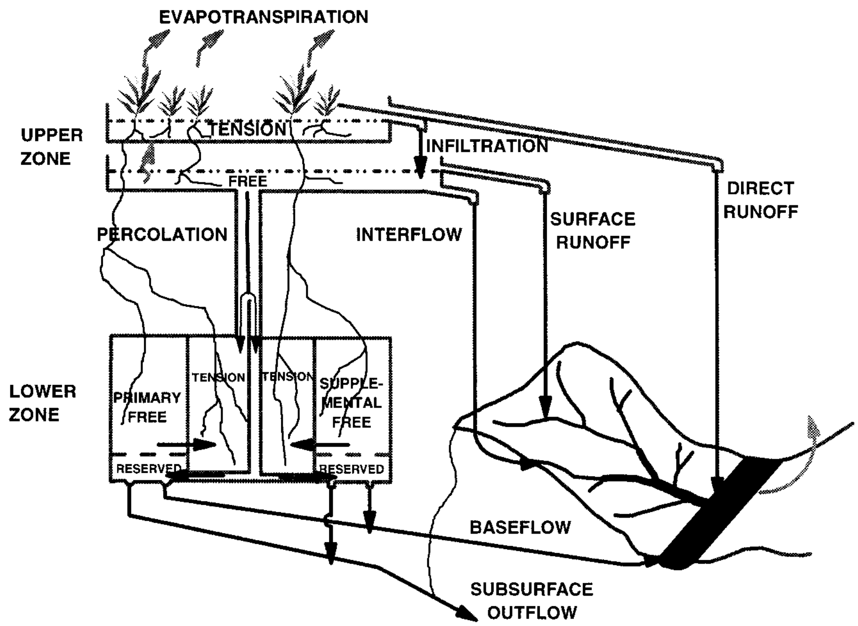

FIG. 2. NWS Sacramento Soil Moisture Accounting Model (SAC-SMA).

TABLE 2. SAC-SMA and SNOW-17 parameter descriptions and typical ranges used for calibration in the NCRFC region.

	Description	Parameter
SAC-SMA		
UZTWM	Upper-zone tension water maximum storage (mm)	20–120
UZFWM	Upper-zone free water maximum storage (mm)	10–100
LZTWM	Lower-zone tension water maximum storage (mm)	100–150
LZFPM	Lower-zone free water primary maximum storage (mm)	5–100
LZFSM	Lower-zone free water supplementary maximum storage (mm)	5–200
UZK	Upper-zone free water lateral depletion rate (day^{-1})	0.100–0.800
LZPK	Lower-zone primary free water depletion rate (day^{-1})	0.01–0.20
LZSK	Lower-zone supplementary free water depletion rate (day^{-1})	0.01–0.50
ADIMP	Additional impervious area (decimal fraction)	0.0–0.10
PCTIM	Impervious fraction of the watershed (decimal fraction)	—
ZPERC	Maximum percolation rate (dimensionless)	1.5–3.5
REXP	Exponent of the percolation equation (dimensionless)	1.5–3.5
PFREE	Fraction of water percolating from upper zone directly to lower-zone free water storage (decimal fraction)	0.01–0.40
RIVA	Riparian vegetation (decimal fraction)	—
SIDE	Ratio of deep recharge to channel base flow (decimal fraction)	—
RSERV	Fraction of lower-zone free water not transferable to lower-zone tension water (decimal fraction)	— —
SNOW-17		
SCF	Snow correction factor (dimensionless)	0.90–1.20
MFMAX	Maximum melt factor [mm °C^{-1} $(6\ h)^{-1}$]	1.0–2.0
MFMIN	Minimum melt factor [mm °C^{-1} $(6\ h)^{-1}$]	0.10–0.90
UADJ	Wind function factor [mm hPa^{-1} $(6\ h)^{-1}$]	preset to 0.100
SI	Water equivalent maximum (mm)	30–100
Areal depletion curve		
MBASE	Melt base temperature (°C)	—
NMF	Maximum negative melt factor [mm hPa^{-1} $(6\ h)^{-1}$]	—
TELEV	Elevation of temperature series (m)	—
DAYGM	Average daily ground melt (mm)	—
PLWHC	Percent liquid water-holding capacity (decimal fraction)	—
PXTEMP	Rain/snow temperature index (°C)	—
Additional parameters (usually not optimized)		
EFC	Effective forest cover (decimal fraction)	—
PXADJ	Precipitation adjustment factor (dimensionless)	—

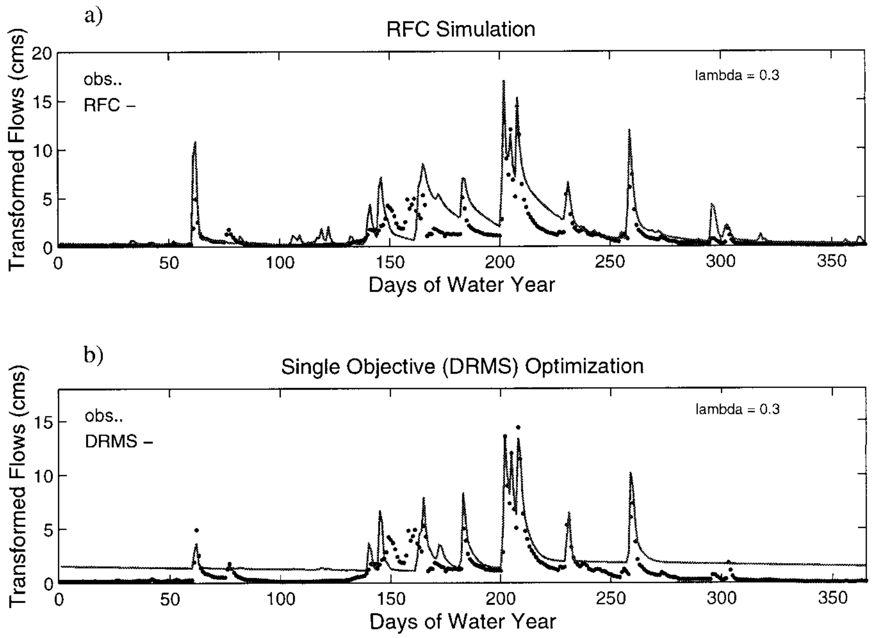

FIG. 3. One year (WY 1976) of calibration for South River. Shown are (a) RFC manual calibration and (b) least squares objective function (one step) optimization. Flows ($m^3 s^{-1}$) are in the transformed space (to observe behavior in the full range of flows better), where transformed flow = $[(\text{flow} + 1)^{\lambda} - 1]/\lambda$, and $\lambda = 0.3$.

Minnesota, Wisconsin, and Michigan, portions of North Dakota, South Dakota, Iowa, Illinois, and Indiana, along with portions of southern Canada. Major drainages include the upper Mississippi River, St. Lawrence River, and Hudson Bay.

The six headwater basins selected by the NCRFC for this study are all located within the upper Mississippi River drainage. Four are located within the Des Moines River Forecast Group: the South River, at Ackworth, Iowa; Cedar Creek, near Bussey, Iowa; the Middle River, near Indianola, Iowa; and the North River, near Norwalk Iowa. The remaining two belong to the Minnesota River Forecast Group: the Redwood River, near Russell, Minnesota, and the Le Sueur River, near Rapidan, Minnesota (Fig. 1). Although the time series data available for these watersheds may not be of ''research quality,'' they represent the type and quality of data with which NWS hydrologists must typically work. Each of the time series was checked and quality controlled by NCRFC personnel and has been used by the NCRFC hydrologists for manual calibration of the models to these watersheds. They are, therefore, representative of typical operational conditions and provide a rigorous and realistic test of the automatic methods under development.

TABLE 3. Parameters optimized during MACS process.

Model	Step 1	Step 2	Step 3
SAC-SMA	UZTWM	UZTWM	—
	UZFWM	UZFWM	—
	UZK	UZK	—
	ADIMP	ADIMP	—
	ZPERC	ZPERC	—
	REXP	REXP	—
	LZTWM	—	LZTWM
	LZFSM	—	LZFSM
	LZFPM	—	LZFPM
	LZSK	—	LZSK
	LZPK	—	LZPK
	PFREE	—	PFREE
SNOW-17	SCF	SCF	—
	MFMAX	MFMAX	—
	MFMIN	MFMIN	—
	SI	SI	—
Objective function	LOG	DRMS	LOG

The study basins, along with relevant data, are listed in Table 1. The available data include basin precipitation, temperature, and streamflow records from 1948 to 1993. Average midmonth evapotranspiration (ET) values are used to estimate the ET-demand curve in the SAC-SMA. An 11-yr period, 1971–81, was used for automated calibration of the models. The entire period from 1948 to 1993 was used for evaluation of the results.

4. Models

Since 1985, the majority of RFCs have been using a highly integrated, centralized software system known as the NWS River Forecast System (NWSRFS; Page 1996). The system is a highly interdependent batch-operating system containing various models and algorithms to aid the RFCs in their hydrologic forecasting responsibilities. The operational part of the system contains several river forecasting models that the

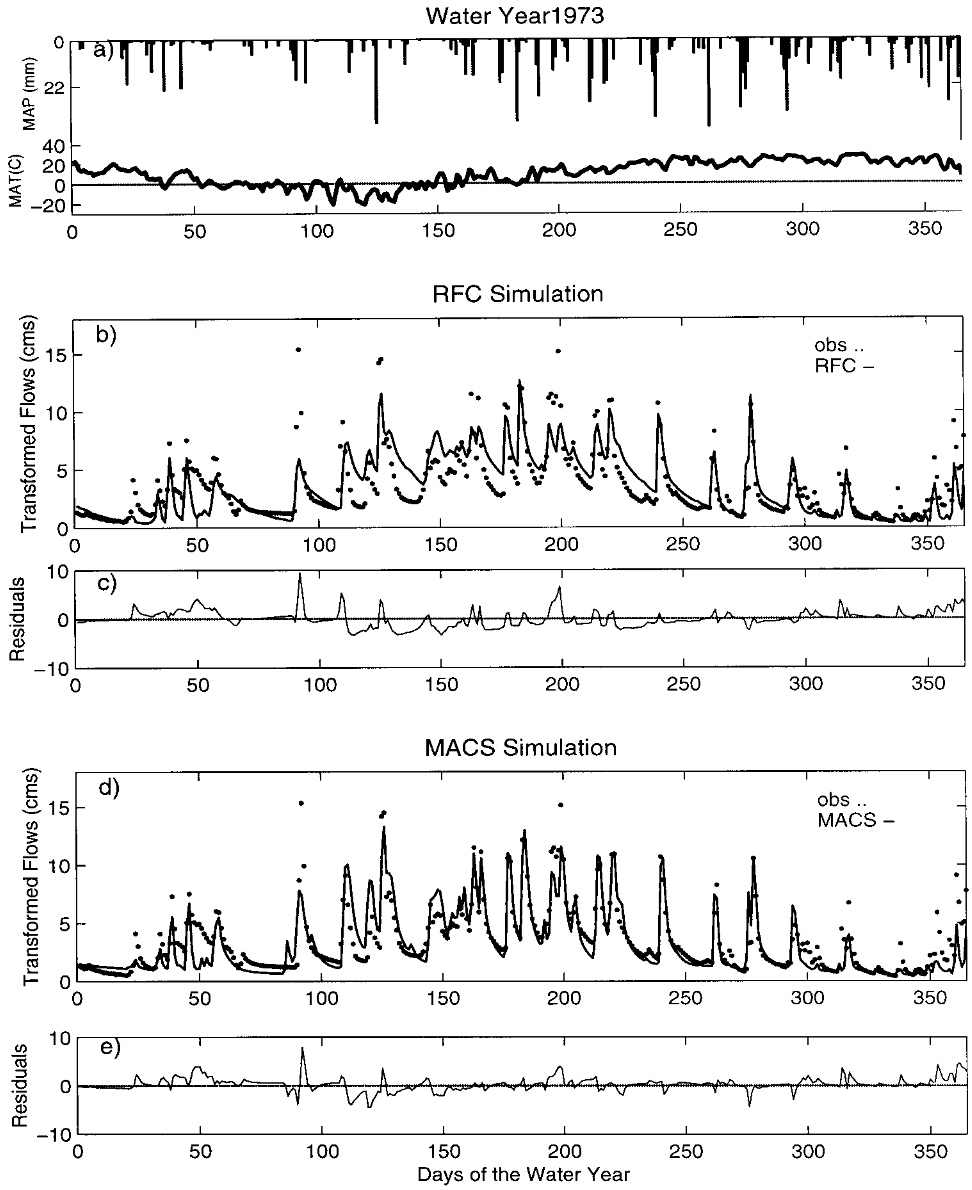

FIG. 4. MACS step-1 calibration for the South River, WY 1973. (a) MAP and MAT time series, (b) RFC simulation, (c) RFC residuals, (d) MACS simulation, and (e) MACS residuals. Plots (b) and (d) are transformed flows as described in Fig. 3.

RFC may use, including the Antecedent Precipitation Index (API) and the SAC-SMA, along with a snowmelt model, SNOW-17.

a. SAC-SMA

The SAC-SMA is a continuous, conceptual rainfall–runoff model that uses two soil-moisture layers to account for flow through the subsurface (Fig. 2). The upper zone represents surface soil regimes and interception storage, and the lower zone represents deeper soil layers containing the majority of the soil moisture and groundwater storage (Brazil and Hudlow 1981). Percolation of water from the upper to the lower zone is controlled by a complex nonlinear process dependent on the contents of upper-zone free water and the deficiencies in lower-zone storages. When rainfall exceeds interflow and percolation capacities, upper-zone storage will be full, and saturated or overland flow will occur. The model has 16 parameters, along with an ET-demand curve (or adjustment curve). A separate unit hydrograph is then used to convert channel inflow to streamflow. Inputs to the model typically include mean areal precipitation [MAP; mm $(6\ h)^{-1}$] and potential evapotranspiration (mm

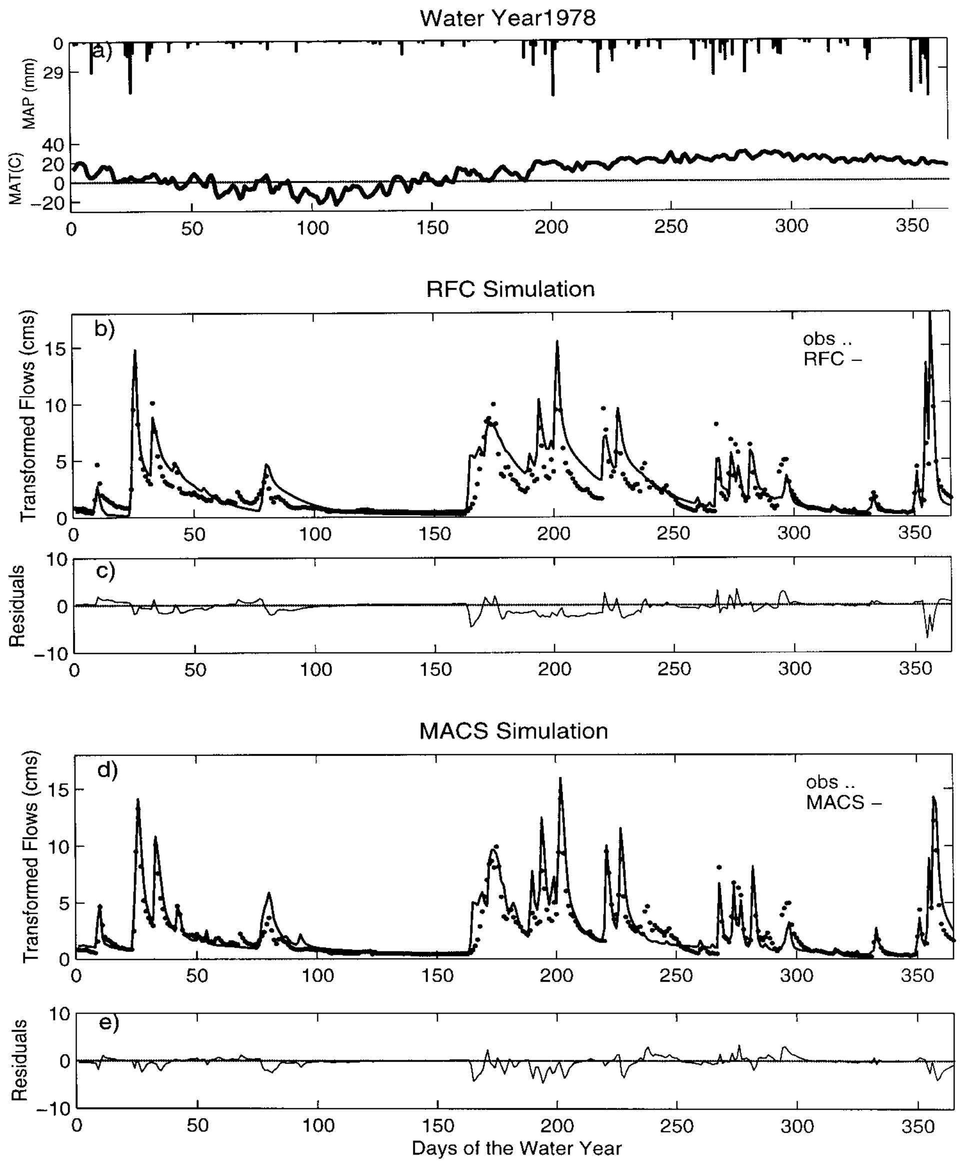

FIG. 5. Same as Fig. 4 but for 1978.

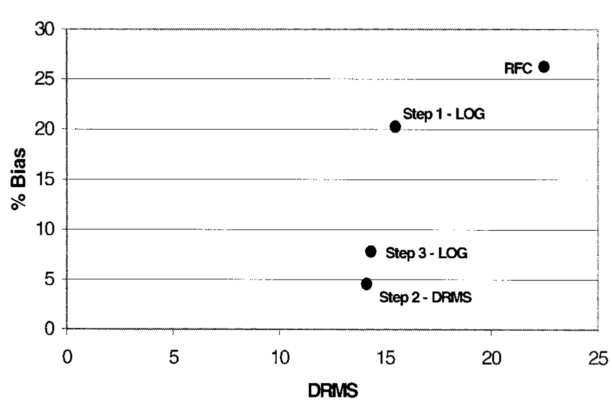

FIG. 6. MACS steps-1, -2, and -3 statistics for the South River (calibration period).

day^{-1}). Outputs are estimated evapotranspiration (mm day^{-1}) and channel inflow.

b. SNOW-17

The SNOW-17 model uses a temperature index method developed by Anderson (1973) to simulate the energy balance of a snowpack and model accumulation and ablation of the snow cover. SNOW-17 models the energy exchange at the snow surface, heat storage and heat deficit within the snowpack, liquid water retention and transmission through the snowpack, and heat exchange at the ground surface (Anderson 1973). The model uses an areal depletion curve to estimate the extent of snow cover in a basin and subsequently determine rain-on-snow or rain-

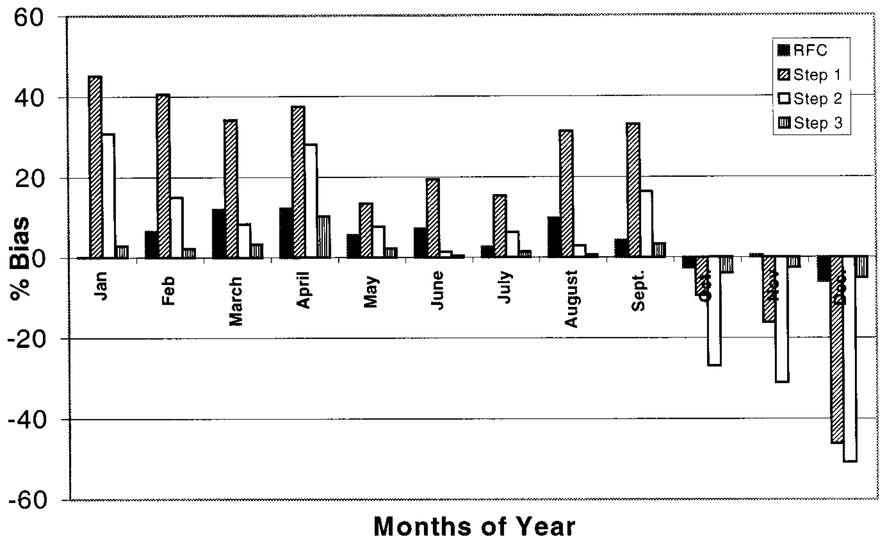

FIG. 7. Monthly flow biases for MACS process on the South River (11-yr data period).

on–bare ground model simulations. For this study, inputs to the SNOW-17 model included MAP [mm (6 h)$^{-1}$] and mean areal temperature [MAT; °C (6 h)$^{-1}$ or °F (6 h)$^{-1}$]. Outputs included a rain/melt time series [mm (6 h)$^{-1}$] and snow water equivalent (mm) of the snowpack. When used in conjunction with the SAC-SMA, input to the SAC-SMA is the rain-melt output from the SNOW-17 model. The model has 12 parameters, of which 6 are considered to be major parameters (more of an effect on simulations) and 6 are minor parameters (less of an effect). A list of parameters for the SAC-SMA and SNOW-17 models is provided in Table 2.

The ordinates of the unit hydrograph (channel routing) were provided by NCRFC and were not calibrated as part of this study. There are 28 parameters remaining that need to be estimated: 16 for the SAC-SMA and 12 for the SNOW-17. Of the 16 SAC-SMA parameters, 3 of the parameters, RIVA, SIDE, and RSERV, are typically set to established values (Burnash et al. 1973; Burnash 1995). The PCTIM parameter also can be established from regional maps and local hydrologic information. For the SNOW-17 model, the minor parameters, along with the areal depletion curve, were set at values obtained from NCRFC and were not optimized. These parameters can be estimated from model documentation (Anderson 1973, 1978) or obtained from historical snow data for the basin. An additional three parameters, EFC, PXADJ, and UADJ, were not optimized, being set to 0.0, 1.00, and 0.1, respectively, for all basins based on values used in other areas of NCRFC. In summary, a total of 16 parameters, 12 from the SAC-SMA model and 4 from the SNOW-17 model, were considered for optimization in this study.

c. NWSRFS automatic calibration program

The calibration system within NWSRFS contains manual and automatic calibration programs to aid in the estimation of parameters for their hydrologic models. The majority of hydrologists within the NWS use a manual approach that relies on an interactive visual interface to estimate parameters for the SAC-SMA and SNOW-17 models. The calibration system also contains an automatic optimization program (OPT3; NWS 1999) that can be used to estimate parameters for watersheds automatically. OPT3 is a single-objective calibration procedure that allows the user to choose one of six possible objective functions as the optimization criterion. The procedure also allows the user to select one of three well-established optimization strategies: the pattern search algorithm, the adaptive random search algorithm, or the SCE-UA algorithm. A maximum of 16 parameters can be estimated simultaneously within a single optimization run. These characteristics of OPT3—available objective functions, single-criterion optimization algorithms, and maximum number of optimizable parameters—defined the parameters (limitations) of this study; the goal was to develop a viable automatic calibration strategy that uses only these resources. This paper describes an approach that accommodates the limitations of OPT3 into a functional optimization scheme that gives results comparable to those that can be obtained by manual means.

d. SCE optimization method

The SCE-UA, available within OPT3, is a global search optimization method designed to handle difficult, nonlinear response surfaces encountered in the calibration of conceptual watershed models. The SCE-UA algorithm has been tested extensively in the last few years (Duan et al. 1992, 1993; Sorooshian et al. 1993; Gan and Biftu 1996; Kuczera 1997; Cooper et al. 1997; Franchini et al. 1998; Freedman et al. 1998; Thyer et al. 1999) and is found to be efficient and consistent in finding the global optimum. A detailed description of

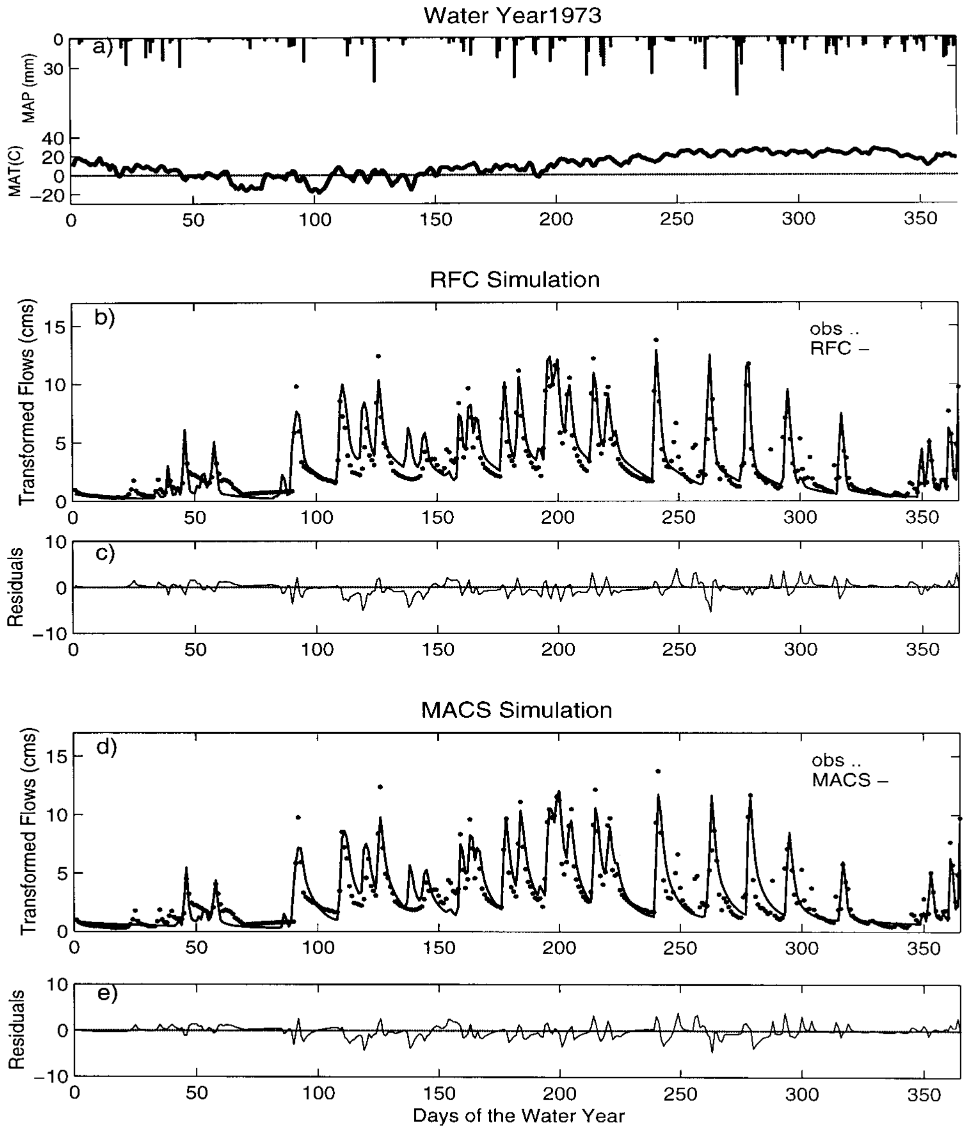

FIG. 8. MACS step-3 calibration on Cedar Creek, WY 1973. (a) MAP and MAT time series, (b) RFC simulation, (c) RFC residuals, (d) MACS simulation, and (e) MACS residuals. Plots (b) and (d) are in the transformed space as described in Fig. 3.

the method appears in Duan et al. (1992). A summary of the algorithm is presented here.

1) A ''population'' of points is selected randomly from the feasible parameter space.
2) The population is partitioned into several complexes, each with $2n + 1$ points, where n is the number of parameters to be optimized.
3) Each complex is evolved independently based on the downhill simplex method.
4) The population is periodically shuffled to share information, and new complexes are formed.
5) Evolution and shuffling are repeated until the specified convergence criteria are satisfied.

In this study, 13 complexes were used {this gave a population size of $13[(2 \times 16) + 1)] = 429$ points}, with a convergence criterion of 0.1% (change in objective function) in five loops. The search was terminated after 20 000 iterations (if an optimum was not reached).

5. Multistep automatic calibration scheme

a. Introduction

As mentioned in the previous section, the goal of this study was to develop a viable automatic calibration strategy that uses only the resources currently available (within the OPT3 procedure) to NWS hydrologists and that gives results comparable to those that can be obtained by manual means. To set the stage for this goal,

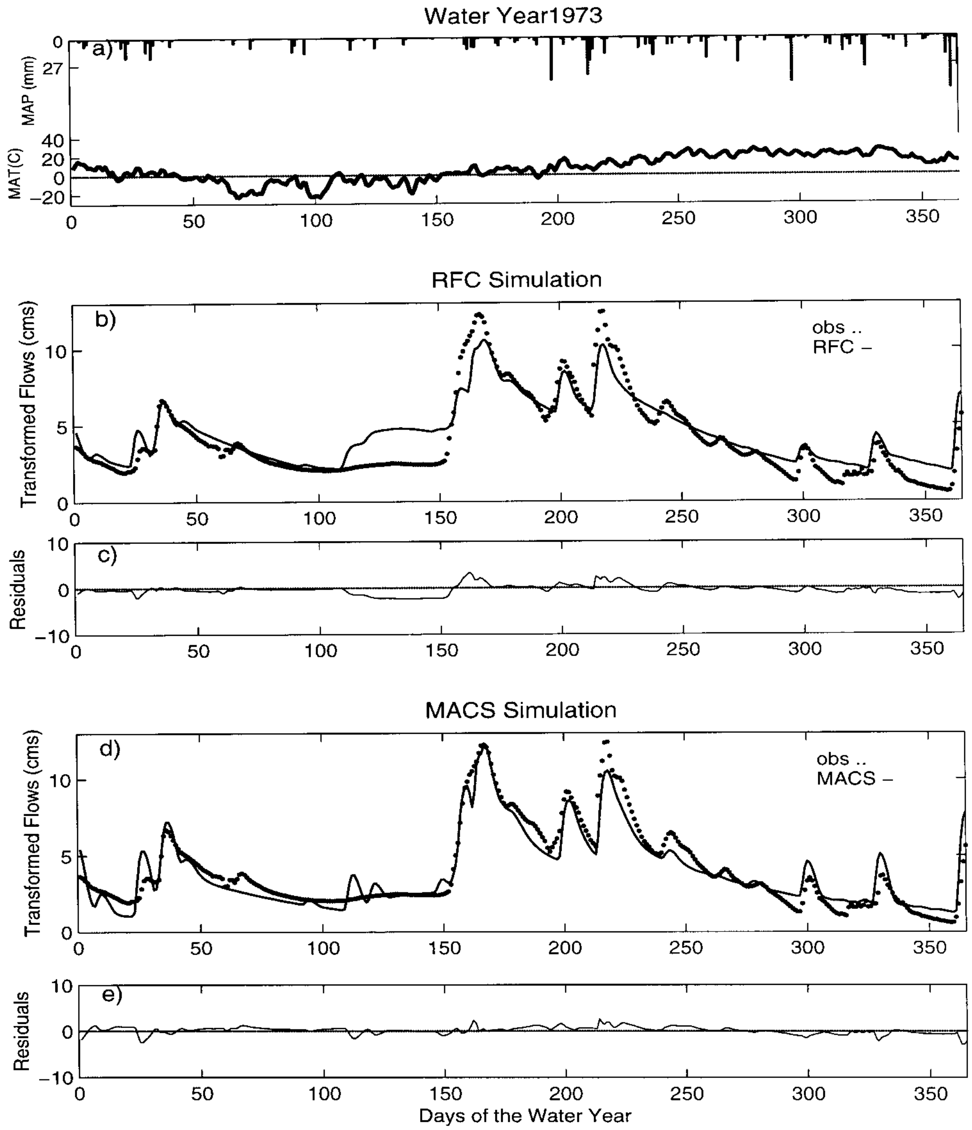

FIG. 9. Same as Fig. 8 but on the Le Sueur River.

an initial series of independent single-step calibration runs was performed using each of the six objective functions available within NWSRFS. The results of a DRMS calibration are shown in Fig. 3. The RFC manual calibration run for the 1976 water year (WY) is shown in Fig. 3a, and the single-step optimization run is displayed in Fig. 3b. The tendency of the DRMS to try to match the peaks while providing strongly biased simulations of other parts of the hydrograph is clearly apparent. The DRMS calibration never adequately returns to observed base flow. Forecasters rely on reasonable simulations of hydrograph recessions to have confidence that the soil-moisture states are being simulated correctly in preparation for a subsequent flood event. Other criteria gave varying results, but all had similarly unacceptable biases in various portions of the hydrograph.

Based on this initial analysis, two objective functions were chosen as suitable for use in the development of a multistep automatic calibration approach. The LOG function [Eq. (1)] was selected for calibration of the parameters that influence the hydrograph recessions (lower-zone parameters), and the DRMS function [Eq. (2)] was selected for calibration of the parameters affecting high-flow events. These functions are defined as follows:

$$\mathrm{LOG} = \sum (\log Q_{\mathrm{sim},t} - \log Q_{\mathrm{obs},t})^2 \quad \text{and} \qquad (1)$$

$$\mathrm{DRMS} = \left\{\frac{1}{n}\left[\sum_{t=1}^{n} (Q_{\mathrm{sim},t} - Q_{\mathrm{obs},t})^2\right]\right\}^{1/2}, \qquad (2)$$

where $Q_{\mathrm{sim},t}$ = simulated flows, and $Q_{\mathrm{obs},t}$ = observed flows at time step t.

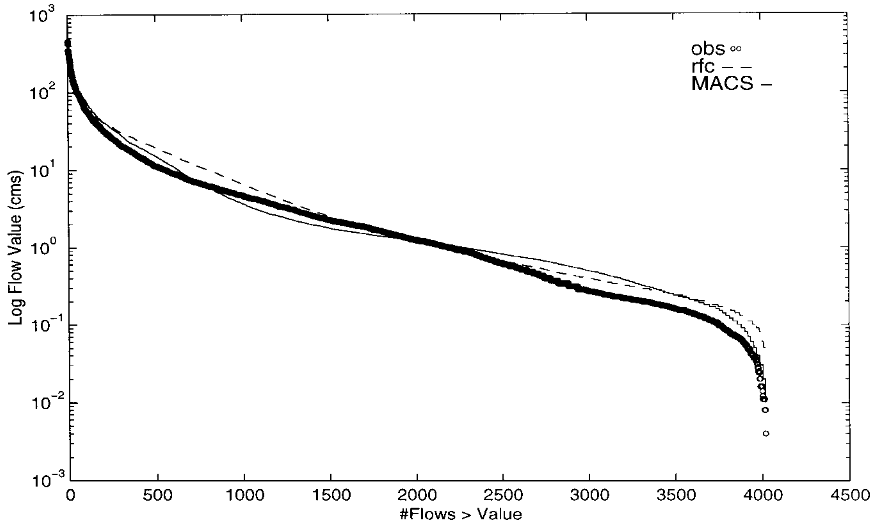

FIG. 10. Flow duration curve for the South River (calibration period). Shown are observed (circles), MACS (solid line), and RFC (dashed line).

Automatic stepwise schemes to calibrate hydrologic models have been investigated by other researchers. Harlin (1991) developed the process-oriented calibration (POC) automatic calibration scheme to apply to the Hydrologiska Byråns Vattenbalansavdelning (HBV) model. The POC method involves two steps, calibrating different sets of parameters with different objective functions over relevant subperiods of the data. Brazil (1988) developed an interactive multistage semiautomated (manual) method to calibrate the SAC-SMA. However, discussions with NWS personnel indicate that this method has never been integrated into NWS operations.

b. Methodology

The multistep automatic calibration procedure reported in this paper has been designed to mimic the manual process used by NWS operational hydrologists. The parameter ranges (optimization bounds) reported in Table 2 were provided by the NCRFC and represent ''realistic'' values for the area of study. The MACS stepwise procedure follows.

1) *Step 1.* In the manual approach, an attempt is first made to adjust the parameter values related to the lower soil zone, in order to obtain acceptable reproduction of the base flow portions of the hydrograph. The first step in MACS is, therefore, to optimize the parameters of the lower zone to minimize errors in base flow by using a criterion that emphasizes fitting of the hydrograph recessions and lower-flow values. An optimization run is made on all 16 parameters (12 SAC-SMA and 4 SNOW-17) using the LOG objective function over the 11-yr calibration period (Table 3). The midpoint values of the parameter ranges are used as the starting point for the optimization algorithm. Notice that the use of the LOG criterion places strong weighting on the low-flow portions of the hydrograph to provide good estimates of the lower-zone parameters. However, by computing the criterion over the entire hydrograph (thereby weighting the higher flows less strongly) and optimizing *all* of the parameters, this step also helps to constrain loosely the remaining (upper zone) model parameters into the region that provides coarse fitting of the peaks.
2) *Step 2.* The next step of the manual process is to estimate the parameters that influence higher-flow events. The second step of the MACS procedure builds on step 1 by fixing the lower-zone parameters constant at the values estimated during step 1 and optimizing the remaining parameters (SAC-SMA upper-zone and percolation parameters and SNOW-17 model parameters; see Table 3). The DRMS objective function is used to provide stronger emphasis on reproduction of the peak flows. Once these upper-zone and snow parameters are estimated, they may be fine-tuned manually or held as estimated, but they are not optimized further.
3) *Step 3.* After completion of the second step, an additional run is made to allow refinement of the lower-zone parameters. Only the lower-zone parameters are optimized, using the LOG criterion, holding the upper-zone parameters at the values estimated in step 2.
4) *Step 4.* Last, an optional fourth step is to adjust manually the monthly ET parameters to minimize monthly flow biases.

6. Results

a. Calibration period

The 11-yr period of WY 1971–81 was selected for model calibration because it consisted of several high-

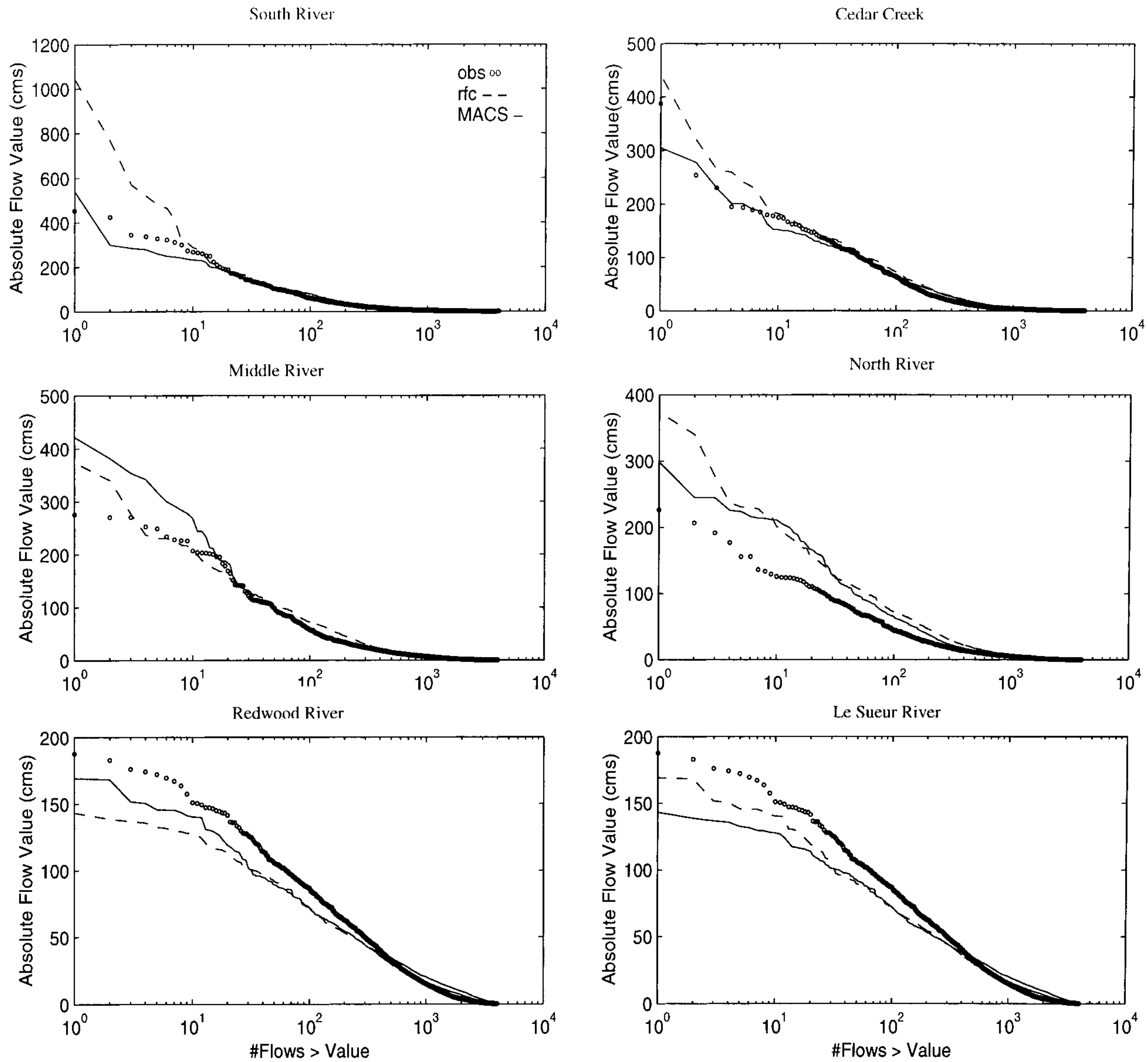

FIG. 11. Modifed flow duration curves for all six study basins (calibration period). Shown are observed (circles), MACS (solid line), and RFC (dashed line).

flood events along with periods where drier conditions prevailed. The parameter estimates provided by both the MACS approach and the NCRFC manual calibrations were compared and evaluated by running simulations over the entire historical record of data, 1948–93.

The evaluations are based both on subjective visual inspection of the modeled versus simulated hydrographs and on several statistics used routinely by the NWS RFCs to evaluate model performance. These statistics include monthly and flow-interval percent biases, overall percent biases, and overall DRMS values. The DRMS is calculated using Eq. (2), and the percent bias is defined as follows:

$$\%\text{Bias} = \left[\sum_{t=1}^{n} (Q_{\text{sim},t} - Q_{\text{obs},t}) \Big/ \sum_{t=1}^{n} Q_{\text{obs},t}\right]100. \quad (3)$$

DRMS and %Bias statistics were computed for all model runs during this study, using a statistical package within NWSRFS (''STAT-QME''). Visual inspections of hydrographs included comparisons of peak simulations, recession slopes, precipitation typing, and timing of spring snowmelt. These comparisons were performed for both the 11-yr calibration period (WY 1971–81) and the entire 46-yr evaluation period (WY 1948–93).

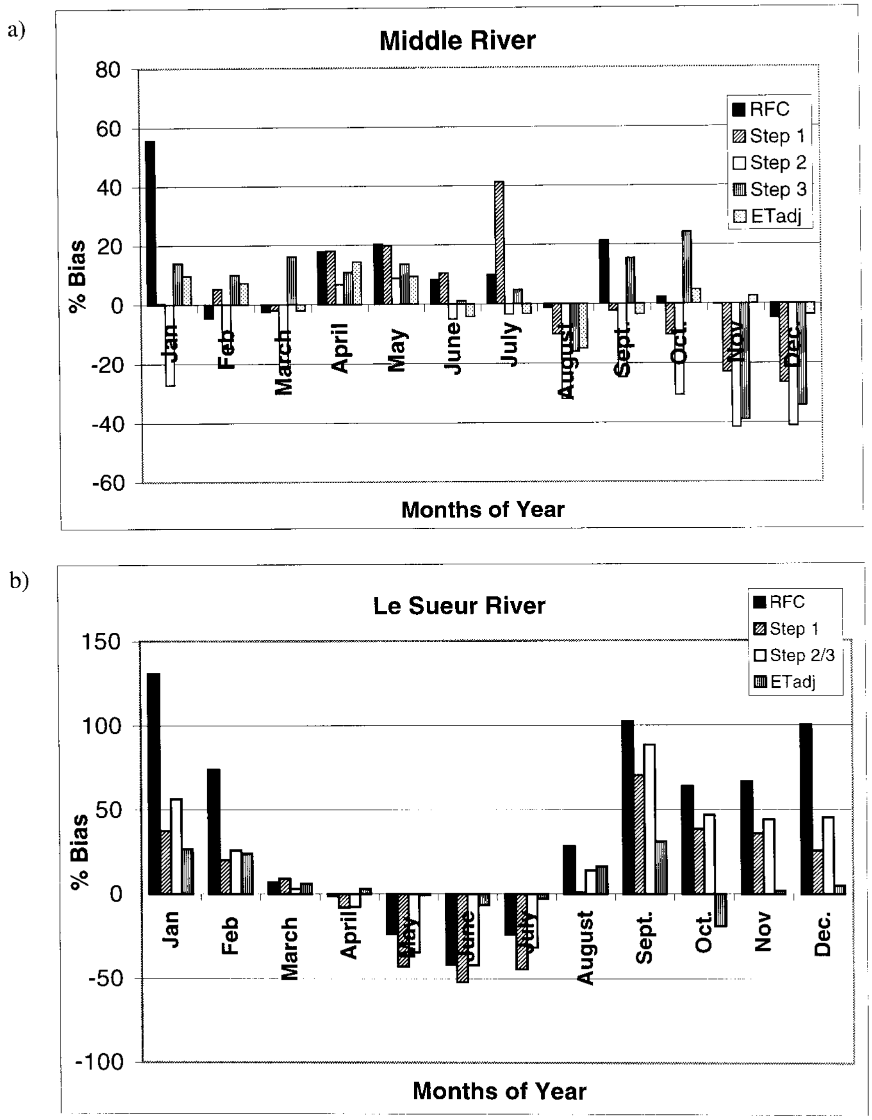

FIG. 12. Monthly flow biases for all MACS steps on (a) the Middle River and (b) the Le Sueur River for the 11-yr calibration period.

1) STEP 1: ESTIMATION OF BASE FLOW PARAMETERS

Figure 4 illustrates the results for the South River basin for WY 1973. Daily MAP and MAT time series are at the top of Fig. 4a, with hydrographs of the RFC manual calibration versus observed flows shown in Fig. 4b and MACS Step 1 calibration versus observed flows displayed in Fig. 4d. Calibration residuals (the difference between observed and simulated flows) for each of the simulations are displayed in Figs. 4c and 4e. The hydrographs depicted in Figs. 4b and 4d are presented using a simple transformation of the flows that expands the lower end of the flow scale (for a better view of the recessions) while still maintaining a reasonable visual perspective of the higher flows. The transformed flow Q is calculated as follows:

$$Q_{\text{transformed}} = \frac{(Q_t + 1)^{\lambda} - 1}{\lambda}, \qquad (4)$$

where $\lambda = 0.3$ in this transformation. (Note: a value of $\lambda = 0$ equates to a log transformation, and $\lambda = 1$ equates to no transformation. The transformation $\lambda = 0.3$ was chosen after evaluating the visual effect provided by a full range of λ values.)

Both the MACS step-1 and RFC calibrations slightly underestimate peak flow values; this underestimation is especially evident for the flow event near day 90, where both calibration methods have missed this event. Other

TABLE 4. Parameter values of RFC manual calibration and MACS scheme.

	South River		Cedar Creek		Middle River		North River		Redwood River		Le Sueur River	
	RFC	MACS	RFC	MACS	RFC	MACS	RFC	MACS	RFC	MACS	RFC	MACS
SAC-SMA												
UZTWM	50	119	40	72	100	120	40	120	83	119	50	40
UZFWM	15	46	25	57	35	16	25	37	40	66	30	43
UZK	0.3	0.534	0.6	0.527	0.4	0.23	0.6	0.33	0.34	0.1	0.15	0.1
LZTWM	120	140	175	160	80	118	225	107	170	150	120	150
LZFSM	60	47	30	18	50	28	45	38	100	29	100	66
LZFPM	10	13	10	16	30	11	10	7.3	25	53	20	14
LZSK	0.1	0.047	0.1	0.248	0.07	0.044	0.1	0.044	0.07	0.068	0.035	0.027
LZPK	0.007	0.008	0.005	.015	0.005	0.007	0.005	0.01	0.004	0.002	0.005	0.002
ADIMP	0	0.1	0	0.1	0	0.1	0	0.099	0.01	0	0.02	0.098
ZPERC	20	199	150	199	47	200	150	101	82	199	40	62
REXP	3	1.5	1.75	1.5	2.8	1.5	1.75	1.5	1.8	2.4	2	2
PFREE	0.03	0.171	0.05	0.068	0.3	0.139	0.1	0.12	0.3	0.156	0.2	0.047
SNOW-17												
SCF	1	0.90	1.1	0.90	0.85	0.8	1.1	0.9	1.1	0.92	1	0.90
MFMAX	1.8	1.13	1.6	1.35	1.8	1.86	1.6	2	1.5	2	1.8	1.99
MFMIN	0.2	0.43	0.6	0.51	0.9	0.23	0.6	0.36	0.6	0.1	0.8	0.55
SI	50	64	35	76	50	91	35	90	90	32	90	86

high-flow events for WY 1973 are somewhat better estimated by MACS than by the RFC calibration. The MACS step-1 simulation also does a better job of simulating recession flows than does the RFC manual calibration. Figure 5 shows the results for a different water year (WY 1978). Both simulations capture most of the high-flow events (near days 200 and 355), but the MACS method models recession of the flood peak at day 200 much better than does the RFC simulation. Other flow events during this water year (days 40 and 220) are generally estimated more accurately (particularly the overall recessions) with MACS than with RFC calibrations.

2) STEP 2: CALIBRATION OF UPPER-ZONE SAC-SMA AND SNOW-17 PARAMETERS

Step 2 of MACS is designed to estimate higher-flow events better and to match peak flows while reducing overall DRMS and %Bias on flows. Figure 6 shows the evolution of the DRMS and %Bias statistics for the different steps in the MACS procedure, as well as a comparison with the RFC manual calibration results for the South River basin. The second step in MACS can be seen to reduce dramatically the overall %Bias (from 20% to 4.9%) while slightly improving the overall DRMS (this is typical of the results for all six basins).

3) STEP 3: REFINEMENT OF BASE FLOW PARAMETERS

MACS step 2 provides a significant improvement in model performance when viewed in terms of the overall DRMS and %Bias statistics, but the *monthly* percent biases for most of the months (see Fig. 7) are still considerably larger than those provided by the RFC calibration. This difference is especially evident during the winter months, where MACS step-2 %Biases run as large as 40%. MACS step 3 performs a fine-tuning adjustment of the base flow parameters, while keeping the upper-zone and snow parameters fixed at the values estimated through steps 1 and 2. Figures 6 and 7 show that the monthly %Biases for the South River are significantly improved through this step, with only slight deterioration in the overall %Bias statistic. The monthly biases are now less than approximately 5% and much more in line with RFC simulations. However, this step provided significant improvements for only some of the study basins.

Figures 8 and 9 display the WY 1973 hydrographs resulting from MACS step 3 and the RFC manual calibrations for the Cedar Creek (Iowa) and Le Sueur (Minnesota) River basins, respectively. Note that the hydrologic behavior of the Minnesota River basins is different from that of the Iowa River basins. On the Cedar Creek basin, the MACS calibration (Fig. 8d) provides results that are nearly identical to the RFC manual calibration (Fig. 8b). Both simulations slightly underestimate most

TABLE 5. Model performance of RFC and MACS methods.

Basin	DRMS		% Bias		R. Coeff.	
	RFC	MACS	RFC	MACS	RFC	MACS
Calibration period (1970–80)						
South	22.48	14.33	26.24	7.79	0.7770	0.8496
Cedar	13.49	11.79	23.74	11.95	0.8101	0.8295
Middle	15.27	13.63	17.73	10.12	0.7674	0.8456
North	10.48	9.58	10.40	11.88	0.7982	0.8093
Redwood	2.40	3.35	47.26	16.38	0.6862	0.7059
Le Sueur	13.93	12.30	15.42	6.03	0.8008	0.8414
Verification period (1948–93)						
South	18.72	15.94	17.84	4.55	0.7845	0.8315
Cedar	13.60	14.05	11.36	0.14	0.8437	0.8314
Middle	13.29	12.60	9.69	2.14	0.7879	0.8295
North	9.30	9.90	–2.14	1.10	0.7773	0.7395
Redwood	3.51	3.96	–4.36	–25.84	0.7922	0.7412
Le Sueur	17.19	15.12	6.67	–4.24	0.8271	0.8757

of the flow events while fitting the recessions fairly well. Both methods miss the small event near day 250, indicating possible problems with the precipitation data. On the Le Sueur River basin, the MACS calibration (Fig. 9d) does a slightly better job than the RFC calibration does (Fig. 9b) in matching the observed flows for most of this water year (Fig. 9). Both calibrations have a slight problem near day 110, but MACS does a better job of estimating the first big runoff event on day 160. Both simulations underestimate the next big flow event, near day 225. The MACS and RFC procedures were both able to give reasonable quality calibrations for the basins belonging to the Iowa River forecast group, but they both had difficulty providing good-quality calibrations for the two Minnesota River basins (Le Sueur and Redwood). We speculate that this result may be due to a combination of problems with the data (precipitation inputs) and/or difficulties in modeling the Minnesota terrain (topographically very flat, low flows, deep snowpack, etc.).

As an additional test, flow duration curves were constructed for the six basins. The observed flows and results for the RFC and MACS calibrations are presented (Fig. 10) for the 11-yr calibration period for the step-3 calibration of the South River. The MACS calibration tends to match the observed flows throughout the range of flows. However, although this traditional flow duration curve plot shows the general trend of the simulations, it is difficult to visualize how well the simulation performs at the higher range of flow values. Figure 11 illustrates modified flow duration curves, with the *x* axis expanded at the higher flow values by use of a logarithmic scale. Five of the basins are plotted with parameters from step 3 of MACS, and the Le Sueur River basin is illustrated with parameters estimated after step 4 (discussed below). On four of the six basins (South River, Cedar Creek, North River, and the Redwood River), the MACS simulation tends to match higher flow values better. On the Middle River basin, both RFC and MACS tend to oversimulate flow values, and, on the Le Sueur River basin, both RFC and MACS undersimulate observed flow values at the high end.

4) STEP 4: ADJUSTMENT OF ET VALUES

As a final, but optional, step, a final check of monthly %Biases may reveal trends that call for an adjustment of previously estimated ET parameters. The current version of OPT3 does not allow for automatic optimization of the monthly ET demand curve. A manual fine-tuning or adjustment of the ET curve (ETadj), using monthly percent biases as a guide, may produce more accurate streamflows during all seasons. After the first three calibration steps with MACS, the Le Sueur and Middle River basins still had monthly biases that were unreasonably high and showed seasonal trends. Using monthly percent biases and ET values from nearby basins as a guide, monthly ET parameters for both basins were adjusted manually to bring the %Biases into a more reasonable range (ideally $<10\%$). Figure 12 illustrates final monthly %Biases for the MACS method for the Middle River and Le Sueur River basins after ET adjustment. Monthly %Biases are plotted for steps 1, 2, 3, and 4. Because there was no change in base flow parameters when running step 3 on the Le Sueur River, steps 2 and 3 are plotted together. The plot for the Middle River basin displays all four steps of the MACS method. The adjustment of the ET parameters significantly reduced the monthly biases for both basins, as evidenced by the decrease in %Biases from step 3 to step 4 (ETadj) for most of the months. This step was only necessary on the Middle River and Le Sueur River basins and was not performed on other basins.

Final parameter values obtained for the six study basins are listed in Table 4. The MACS and RFC calibrations converged to slightly different values on the majority of basins. Parameter bounds used in the calibration were considered representative for the study area and are the same bounds that the NCRFC ‘‘typically’’ uses in their calibrations. It is also noted that, although

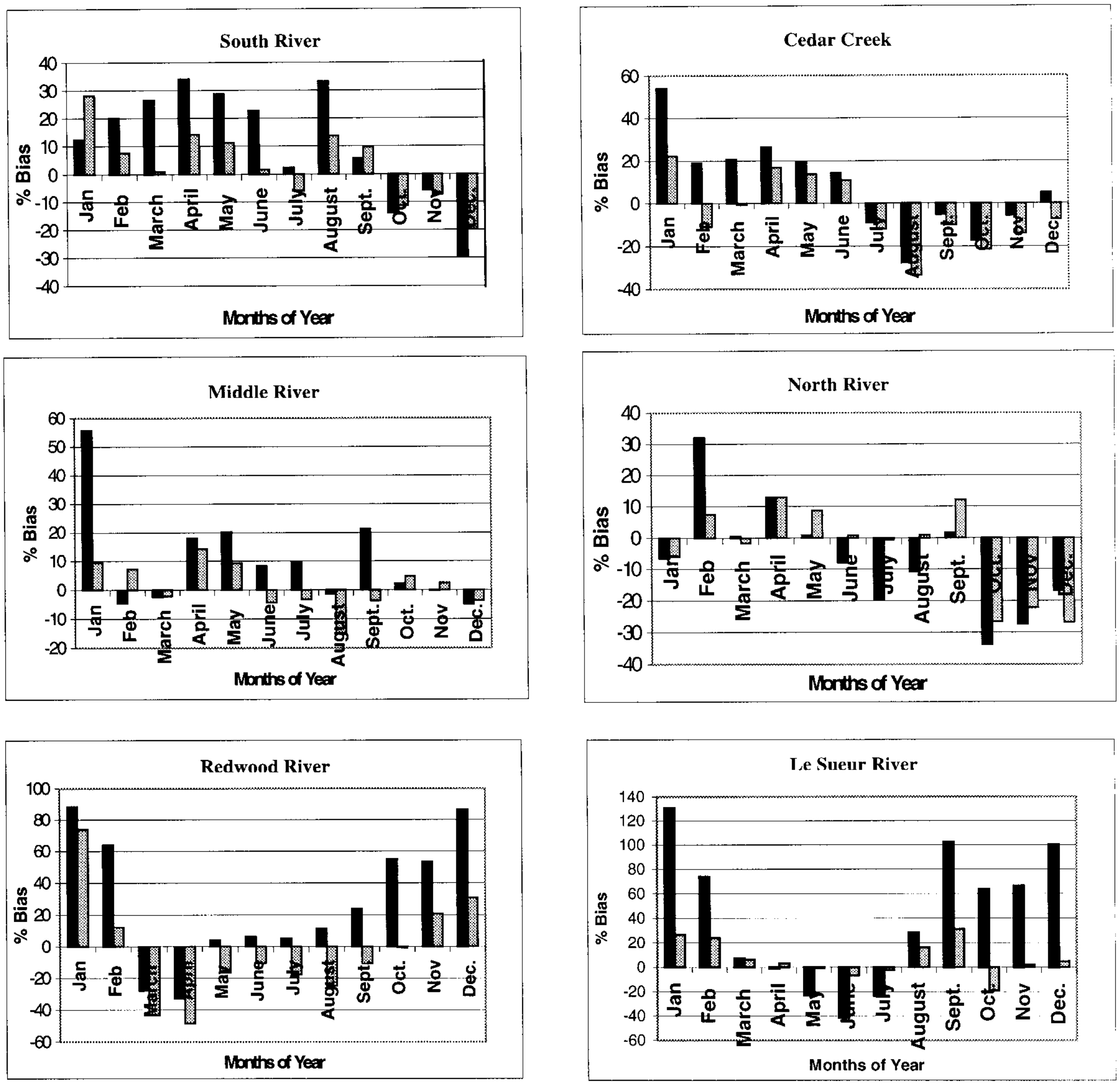

FIG. 13. Monthly %Bias for six study basins for evaluation period of 1948–93. RFC simulation is shown in black; MACS is shown in gray.

four of the headwater basins are within the Des Moines River Forecast Group (South River, Cedar Creek, Middle River, and North River), it is obvious from the RFC and MACS calibrations that the headwaters respond differently and that parameter values obtained in these four basins do vary from one another. This result lends credence to the NWS requisite that a headwater in each river basin should be calibrated before proceeding downstream in the river system.

Last, it is observed that parameters on some of the calibrations, most notably UZTWM, LZTWM, and ZPERC, tended to be at or near the bounds of the parameter ranges specified by NCRFC. Additional studies were performed by widening the limits on the abovementioned parameters; however, preliminary results from these studies were inconsistent and inconclusive. Further research into the ‘‘acceptable’’ optimization limits for specific regions is required but is beyond the scope of this paper.

b. Evaluation period

After initial calibration of all the basins using the MACS method, evaluation of the calibrated parameters was performed by running the model for the entire 46-yr data period, 1948–93. MACS (step 3) statistics for the 11-yr calibration period and 46-yr verification periods are displayed in Table 5. The same statistics are also listed for the RFC over the same calibration and evaluation periods. Overall DRMS, %Bias, and correlation coefficient (R. Coeff.) statistics are listed for each basin. For the calibration period (1971–81), the MACS

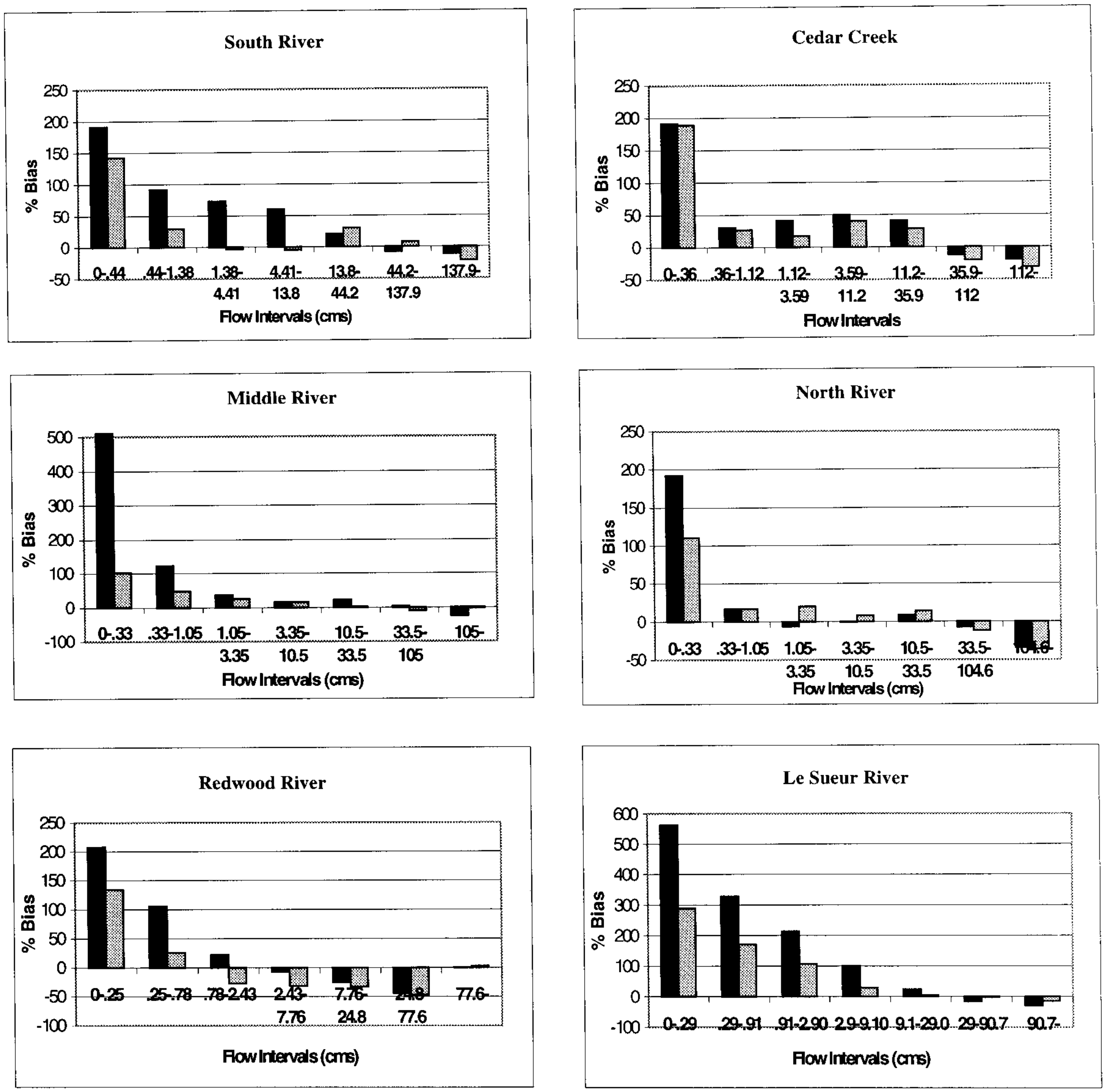

FIG. 14. Flow interval biases for six study basins for evaluation period of 1948–93. RFC simulation is shown in black; MACS is shown in gray.

procedure results in 1) better DRMS values for five of the six basins (excluding the Redwood River), 2) better %Bias values for five of the six basins (excluding the North River), and 3) better correlation statistics for all six basins. The second part of the table displays statistics for the 46-yr verification or evaluation period. In this section, the RFC and MACS methods show more comparable performance. MACS gives better DRMS and correlation coefficient (R. Coeff.) on three of the basins (South River, Middle River, and Le Sueur River), while RFC does better on the other three. However, the MACS procedure gives significantly lower %Bias for five of the six basins (the exception is the Redwood River).

Monthly %Biases and flow interval %Biases for each basin are displayed in Figs. 13 and 14, respectively. For the South River, Middle River, North River, and Le Sueur River, the MACS procedure results in lower monthly %Bias. On Cedar Creek and the Redwood River, the MACS still has some months where the %Biases are too high, and further work on these basins may be required. Flow interval %Biases (Fig. 14) reveal that the MACS and RFC are very comparable on the majority of basins, especially towards the higher flow intervals. On nearly all of the basins (excluding Cedar Creek), the MACS simulation has lower %Biases on the lower flow intervals. On the Redwood River and Cedar Creek, the MACS method has slightly higher simulations on higher-flow events than the RFC calibrations. Overall, the RFC and MACS have fairly equivalent %Biases across the range of flows.

7. Conclusions

Advances in technology during the past few decades have made improved calibration tools available to the NWS hydrologists responsible for producing timely and accurate river forecasts. There is increasing interest among RFC hydrologists in the use of automatic procedures for reducing the number of person hours required for producing model calibrations. A simple and straightforward multistep automatic calibration scheme for calibration of river forecasting models has been developed and tested. The study reported here illustrates how the MACS procedure can be incorporated into the operational setting to provide quality first-cut calibrations in a timely manner. A number of statistical measures, as well as visual hydrograph comparisons, were used in this study. Based on these measures, the MACS procedure yielded calibrations that are comparable to, or, in several cases, slightly better than, those obtained by NCRFC hydrologists using the traditional manual approach. Under the NWS modernization program, RFCs are under pressure to produce quality calibrations for all river forecast points in a relatively short period of time. A time-saving automated or semiautomated procedure such as MACS can help the RFCs to achieve this goal. The number of person hours spent using MACS was estimated to be around 3–4 hours per basin, with computer time per calibration estimated to be from 6 to 8 hours per run (IBM-RS6000, Model 591, single processor). An extra advantage of an automated procedure is that the hydrologists are free to perform other forecast duties while the computer runs are being performed. In contrast, the NCRFC hydrologists spend at least 15–20 hours for manual calibration of a forecast point.

Although the MACS procedure developed here provides acceptable first-cut calibrations for operational basins, there are limitations with the stepwise optimization approach. The multiobjective approach under development at UA (Gupta et al. 1998; Boyle et al. 2000) allows hydrologists to examine the trade-off associated with using different objective functions and is expected to provide better support for decision making in the calibration of river forecast points. However, it will be some time before the multiobjective procedure is fully understood and will be ready for operational use. In the short term, methods such as MACS can allow the NWS River Forecast Centers to exploit more fully the time-saving tools available to them within NWSRFS in the production of timely and accurate river-flow forecasts.

Acknowledgments. Partial financial support for this research was provided by the Hydrologic Research Laboratory of the National Weather Service (Grant NA 87WHO582) and the National Science Foundation Graduate Research Trainee Program (Grant DGE-9355029#3). Special thanks are due to Jim Broermann and Dan Braithwaite for providing assistance with computational resources and to Corrie Thies for proofreading.

REFERENCES

Anderson, E. A., 1973: National Weather Service River Forecast System—Snow accumulation and ablation model. NOAA Tech. Memo. NWS Hydro-17, U.S. National Weather Service. [Available from NOAA/NWS, Office of Hydrology, 1325 East-West Highway, Silver Springs, MD 20910.]

——, 1978: Snow cover energy exchange. NWSRFS Manual and Documentation II.2.1-1, NOAA Office of Hydrology, U.S. National Weather Service. [Available from NOAA/NWS, Office of Hydrology, 1325 East-West Highway, Silver Springs, MD 20910.]

Boyle, D. P., H. V. Gupta, and S. Sorooshian, 2000: Towards improved calibration of hydrologic models: Combining the strengths of manual and automatic methods. *Water Resour. Res.,* in press.

Brazil, L. E., 1988: Multilevel calibration strategy for complex hydrologic simulation models. Ph.D. dissertation, Dept. of Civil Engineering, Colorado State University, 217 pp.

——, and M. D. Hudlow, 1981: Calibration procedures used with the National Weather Service River Forecast System. *Water and Related Land Resource Systems,* Y. Y. Haimes and J. Kindler, Eds., Pergamon Press, 457–566.

Burnash, R. J. C., 1995: The NWS River Forecast System—Catchment modeling. *Computer Models of Watershed Hydrology,* V. P. Singh, Ed., Water Resources Publications, 311–366.

——, R. L. Ferrell, and R. A. McGuire, 1973: A generalized streamflow simulation system. Report of the Joint Federal–State River Forecast Center, Department of Water Resources, State of California, and National Weather Service, 204 pp.

Cooper, V. A., V. T. V. Nguyen, and J. A. Nicell, 1997: Evaluation of global optimization methods for conceptual rainfall–runoff model calibration. *Water Sci. Technol.,* **36** (5), 53–60.

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

Duan, Q., V. K. Gupta, and S. Sorooshian, 1992: Effective and efficient global optimization for conceptual rainfall–runoff models. *Water Resour. Res.,* **28,** 1015–1031.

——, ——, and ——, 1993: A shuffled complex evolution approach for effective and efficient global minimization. *J. Optim. Theory Appl.,* **76,** 501–521.

Franchini, M., G. Galeati, and S. Berra, 1998: Global optimization techniques for the calibration of conceptual rainfall–runoff models. *Hydrol. Sci. J.,* **43** (3), 443–458.

Freedman, V. L., V. L. Lopes, and M. Hernandez, 1998: Parameter identifiability for catchment-scale erosion modeling: A comparison of optimization algorithms. *J. Hydrol.,* **207** (1–2), 83–97.

Gan, T. Y., and G. F. Biftu, 1996: Automatic calibration of conceptual rainfall–runoff models: Optimization algorithms, catchment conditions, and model structure. *Water Resour. Res.,* **32,** 3513–3524.

Gupta, V. K., and S. Sorooshian, 1994: A new optimization strategy for global inverse solution of hydrologic models. *Numerical Methods in Water Resources,* A. Peters et al., Eds., Kluwer Academic, 743–751.

——, ——, and P. O. Yapo, 1998: Towards improved calibration of hydrologic models: Multiple and non-commensurable measures of information. *Water Resour. Res.,* **34,** 751–763.

Harlin, J., 1991: Development of a process oriented calibration scheme for the HBV Hydrologic Model. *Nord. Hydrol.,* **22,** 15–36.

Hooke, R., and T. A. Jeeves, 1961: Direct search solutions of numerical and statistical problems. *J. Assoc. Comput. Mach.,* **8** (2), 212–229.

Ibbitt, R. P., 1970: Systematic parameter fitting for conceptual models of catchment hydrology. Ph.D. dissertation, Imperial College of Science and Technology, University of London.

——, and T. O'Donnell, 1971: Fitting methods for conceptual catchment models. *J. Hydraul. Div., Amer. Soc. Civ. Eng.,* **97** (9), 1331–1342.

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

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

Kuczera, G., 1997: Efficient subspace probabilistic parameter optimization for catchment models. *Water Resour. Res.,* **33,** 177–185.

Masri, S. F., G. A. Bekey, and F. B. Safford, 1978: An adaptive random search method for identification of large-scale nonlinear systems. *Proc. Fourth Symp. on Identification and System Parameter Estimation,* Amsterdam, Netherlands, Int. Federation of Automatic Control, 246–255.

Monroe, J. C., 1971: Direct search optimization in mathematical modeling and a watershed model application. NOAA Tech. Memo. NWS Hydro-12. [Available from National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 1325 East–West Hwy, Silver Spring, MD 20910.]

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

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

NWS, 1999: National Weather Service River Forecast System (NWSRFS) User's Manual. [Available from NOAA/National Weather Service, Office of Hydrology, 1325 East–West Hwy, Silver Spring, MD 20910.]

Page, D., 1996: The Implementation of an Interactive River Forecast System for the National Weather Service. Preprints, *12th Int. Conf. Interactive Information and Processing System (IIPS) for Meteorology, Oceanography, and Hydrology,* Atlanta, GA, Amer. Meteor. Soc., 365–369.

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

Rosenbrock, H. H., 1960: An automatic method of finding the greatest or least value of a function. *Comput. J.,* **3,** 175–184.

Sorooshian, S., 1997: The trials and tribulations of modeling and measuring in surface water hydrology. *Land Surface Processes in Hydrology,* S. Sorooshian, H. V. Gupta, and J. C. Rodda, Eds., NATO ASI Series, Vol. 46, Springer-Verlag, 19–45.

——, and V. K. Gupta, 1995: Model calibration. *Computer Models of Watershed Hydrology,* V. P. Singh, Ed., Water Resources Publications, 23–63.

——, Q. Duan, and V. K. Gupta, 1993: Calibration of rainfall–runoff models: Application of global optimization to the Sacramento Soil Moisture Accounting Model. *Water Resour. Res.,* **29,** 1185–1194.

Thyer, M., G. Kuczera, and B. C. Bates, 1999: Probabilistic optimization for conceptual rainfall-runoff models: A comparison of the shuffled complex evolution and simulated annealing algorithms. *Water Resour. Res.,* **35,** 767–773.

Yapo, P. O., H. V. Gupta, and S. Sorooshian, 1996: Calibration of conceptual rainfall–runoff models: sensitivity to calibration data. *J. Hydrol.,* **181,** 23–48.

——, ——, and ——, 1998: Multi-objective global optimization for hydrologic models. *J. Hydrol.,* **203,** 83–97.