IV.4.2-OPT3 CALIBRATION SYSTEM AUTOMATIC PARAMETER OPTIMIZATION PROGRAM (OPT3)

Purpose

The Automatic Parameter Optimization Program (OPT3) is designed to enable users of NWSRFS models to further refine parameter estimates previously developed through manual calibration procedures.

These automatic optimization procedures are not a substitute for the manual calibration process.

The following optimization strategies are currently available:

- o the Pattern Search (PATSERCH) algorithm [Monro, 1971] (see Section IV.4.2-OPT3-PATSERCH)
- o the Adaptive Random Search (ARS) algorithm [Brazil, 1989] (see Section IV.4.2-OPT3-ARS)
- o the Shuffled Complex Evolution (SCE) algorithm [Duan, 1991] (see Section IV.4.2-OPT3-SCE)

Background

An automatic optimization technique consists of two main components:

- o a search algorithm
- o an objective function or optimization criterion

The objective function is a statistical measure of the difference between the observed and simulated hydrographs. If N parameters are being optimized, the optimization criterion is a function of N dimensional space. The surface created by this criterion is called the response surface. The automatic optimization procedure searches this surface to locate a minimum point, corresponding to an optimal set of parameter values. The response surface is usually quite irregular with many bumps and dips, thus, it is difficult to know if the point at which the search algorithm stops is the global optimum or a local optimum corresponding to an inferior parameter set.

The optimization strategies are of two main types:

- o global procedures, such as the ARS and SCE techniques, sample an entire portion of the response surface
- o local search procedure, such as the Pattern Search technique, tries to find an optimum by starting at a particular point on the surface

Thus, the ARS and SCE techniques are less likely to be trapped by local optima (i.e., small valleys on the response surface) than the

Pattern Search technique. The primary difference between the ARS and SCE techniques is that the ARS technique approaches the parameter search process largely in a pure random manner, while the SCE technique conducts the parameter search using a strategy that has both probabilistic and deterministic elements. The advantage of the SCE approach over the ARS approach is that the SCE method can efficiently utilize information gained in the search process and information about the response surface. Because the ARS and SCE techniques search the entire parameter space while the Pattern Search technique searches about a given point, the number of trials required by the ARS and SCE routines (typically several hundred per parameter being optimized) is much greater than that required by the Pattern Search routine (typically 10 to 20 per parameter). Even with so many trials, the chances of finding the global optimum are relatively small when the ARS scheme is used [Armour, 1990, and Wienig, 1991].

The Pattern Search algorithm should only be used to refine parameter values. The range to be searched for each parameter should be relatively small when using the Pattern Search technique. The SCE and ARS algorithms should be used if a larger range of possible values needs to be examined. However, the specified range should be as restrictive as possible so that only the portion of the response surface that is most likely to contain the optimum is searched. The efficiency of the ARS technique is particularly sensitive to the size of the response surface. For all of the techniques, only parameters with a significant amount of uncertainty should be included in order to reduce the complexity of the response surface as much as possible. Parameters that are reasonably well known and parameters whose value is relatively insensitive to the objective function being used should not be included in an automatic optimization run.

The advantage of these optimization techniques is that they are able to test a large number of parameter sets, ultimately producing a set with a lower optimization criterion than the initial set. The great disadvantage of these techniques is that they evaluate the goodness of fit of a simulation based on a single value. The variability of all the hydrograph components are summarized by a single number, the objective function.

General Features

The following objective functions (optimization criterion options) are available:

- o Daily RMS Error (CMSD):

  Criterion=((mean_daily_QSIM-mean_daily_QOBS)**2/num_days)**.5

  This is the most frequently used option.

- o Monthly Volume RMS Error (MM):

  Criterion=((month_vol_QSIM-month_vol_QOBS)**2/num_months)**.5

This option is recommended for situations where monthly volume errors are more critical than daily flow errors.

- o Criterion=(mean_daily_QSIM-mean_daily_QOBS))**Exponent

- o Criterion=(log(mean_daily_QSIM)-log(mean_daily_QOBS))**Exponent

- o Modified Correlation Coefficient [McCuen and Snyder, 1975]

  Criterion=1.0-Modified_R

- o Maximum Likelihood Estimator for the Heteroscedastic Error Case [Sorooshian et al, 1983]
  This option may have advantages over the daily RMS error criterion in some cases. See Section IV.4.2-OPT3-HMLE.

where QSIM is the simulated discharge
QOBS is observed discharge

Options 3, 4 and 5 were developed mainly for use in research; however, there may be circumstances that warrant their use in calibration work. Additional objective functions can be added to the program in the future.

Data can be excluded from the optimization criterion by the following methods:

1. A buffer period ranging from 0 to 6 months at the beginning of the period of record can be specified.

2. A maximum of 10 data periods can be excluded by entering the dates when computation of the optimization criterion should stop and resume.

3. Low flow data may be excluded by specifying a discharge value (CMSD) below which the objective function will not be computed.

Parameters for the following Operations can be optimized:

1. SAC-SMA : PXADJ, PEADJ, UZTWM, UZFWM, UZK, PCTIM, ADIMP, RIVA, ZPERC, REXP, LZTWM, LZFSM, LZFPM, LZSK, LZPK, PFREE, RSERV, SIDE, EFC, ETHIGH, ETLOW. (ETHIGH and ETLOW are adjustment parameters for the ET-Demand or PE-Adjustment curve and always have an initial value of 1.0.)

2. SNOW-17 : PXADJ, ELEV, SCF, MFMAX, MFMIN, UADJ, SI, NMF, TIPM, MBASE, PLWHC, DAYGM, TAELEV.

3. UNIT-HG : UGH, UGV (These parameters reflect the horizontal and vertical adjustment of the unit hydrograph and always have an initial value of 1.0).

4. API-SLC : REC, LLMT, A, I, WN, WX, EI, E2, CP, K, M, POW, PSIG

5. XIN-SMA : K, IMP, WUM, WLM, WDM, SM, B, EX, C, KSS, KG+KSS, CI,

CG, ETHIGH, ETLOW

6. API-CONT : AIXW, AIXD, CW, CD, CS, SMIX, FRSX, BFIM, AICR, CG, CP, CT, CSOIL

When a parameter in an Operation is optimized, only the value of the parameter in the specified Operation is changed (and carryover if specified by the carryover transfer rules). An option exists in the program to allow the same parameter in other Operations of the same type to change also. The other parameter values will be computed based on the ratio or difference between their initial values and the initial value of the parameter being optimized, depending on the type of parameter. Currently, ratios are maintained between parameters, for all Operations that can be optimized, except SAC-SMA and API-CONT. All parameters in SAC-SMA use ratios except for UZTWM, UZFWM, and LZTWM, which use differences. In API-CONT differences are used for AIXW, AIXD, SMIX, and CP.

Input Summary

The input data for program OPT3 consists of the input for program MCP3 and the input needed to define the optimization options, the time series to be used for computing the objective function, and the parameters to be optimized.

Card Group A consists of the MCP3 input data cards (see Section IV.4.1-MCP3).

Card Group B contains the optimization options and the information to define the time series to be used in computing the objective function. The B1 and B2 cards contain general optimization control parameters. The format of the B3 card is dependent on the scheme selected on B1. Formats are given for each available scheme. The B4 and B5 cards also contain general control information for the program.

Card	Format	Column	Contents
B1	A8	1-8	Optimization scheme to be used: 'PATSERCH' = Pattern Search algorithm 'ARS' = Adaptive Random Search algorithm 'SCE-UA' = Shuffled Complex Evolution algorithm
	I5	11-15	Number of months in buffer period (0 through 6).
	I5	16-20	Number of data periods to be dropped (maximum 10).
	A4	22-25	Option to display manual calibration output using parameters from the run with the best optimization criterion. Enter 'MCP3' to the specify option.

Card	Format	Column	Contents
	I5	26-30	Option to punch Operations Table with parameters from run with the best optimization criterion. Enter '1' to specify option.
B2	I5	1-5	Optimization criterion: 1 = Daily RMS Error 2 = Monthly Volume RMS Error 3 = (S-O)**Exponent 4 = (logS-logO)**Exponent 5 = 1.0-ModifiedR 6 = ML Estimator (HMLE)
	I5	6-10	MAXN Maximum number of trials allowed before optimization is terminated. The purpose of MAXN is to stop an optimization scheme before too much CPU time is expended. MAXN should be set large enough so that optimization is generally completed before MAXN trials are performed. Recommended values for MAXN are based on the optimization scheme used and are: PATSERCH - 20 times the number of parameters being optimized ARS - 1000 times the number of parameters being optimized SCE-UA - 40 times the number of parameters being optimized times the number of complexes used
	F5.2	11-15	Exponent to be used with optimization criterion option 3 or 4. Default is 2.0.
	F10.2	16-25	Drainage area of basin (square kilometers).
	F10.2	26-35	Minimum observed discharge for which the objective function will be computed. Default value is zero (objective function will be computed for all flows). This can be used to remove baseflow from the objective function when optimizing parameters only affecting high flows.

B3 Card for Pattern Search scheme:

Card	Format	Column	Contents
B3	I5	1-5	Type of parameter increment to be used. Enter '0' for fixed quantity (absolute value). Enter '1' for percentage parameter value. Generally it is recommended that percentages be used.
	I5	6-10	Maximum number of resolutions allowed

Card	Format	Column	Contents
			before optimization is terminated. Recommended value is 3.
	I5	11-15	Number of trials in which the criterion value must change by the specified percentage or else optimization will be terminated. Recommended value is 3.
	F5.2	16-20	Percentage by which the criterion value must change in the specified number of trials or the optimization is terminated (Use decimal equivalent: Percentage/100). Recommended value is 0.01.

B3 Card for Adaptive Random Search scheme:

Card	Format	Column	Contents
B3	I5	1-5	IF1 Number of range levels in each cycle. Recommended value is 4.
	I5	6-10	IF3 Number of trials in the first range level. Recommended value is 200.
	I5	16-20	IF5 Number of successive cycles that optimization must be in the minimum (smallest) range to stop. IF5 determines when the optimization process is finished. Recommended value is 3.
	I5	21-25	Seed value for random number generator. The choice of the seed value may affect CPU time (speed of convergence) but the final result should not vary much. Recommended value is any odd integer.
	I1	30	Output flag set equal to 1 to obtain full output. Default generates abbreviated output. Recommended value is the default.

B3 Card for Shuffled Complex Evolution scheme:

Card	Format	Column	Contents
B3	I5	1-5	NGS Number of complexes used for optimization search. Minimum value is 1. Recommended value is based on the number of parameters being optimized as follows:

No. of Parameters	No. of Complexes
1 - 2	1
3 - 4	2
5 - 6	4

Card	Format	Column	Contents
			7 - 8 4 9 - 10 8 > 10 8 The more complexes used, the more robust is the procedure, though the CPU time is increased.
	I5	6-10	Minimum number of complexes required, if the number of complexes is allowed to reduce as the optimization search proceeds. The default value is equal to NGS.
	I5	11-15	Random seed used in optimization search. Enter any integer number. Recommended value is any large integer.
	I5	16-20	Number of shuffling loops in which the criterion must improve by the specified percentage or else optimization will be terminated. Recommended value is 5.
	F5.2	21-25	Percentage by which the criterion value must change in the specified number of shuffling loops or the optimization is terminated (Use decimal equivalent: Percentage/100). Recommended value is 0.01.
	I5	26-30	IDEFLT Flag for setting the control variables of the SCE-UA algorithm. Enter '0' for default setting. Enter '1' for user specified setting. If option '1' is chosen, type an optional input card B3a.

B3a Optional Input Card for the Shuffled Complex Evolution scheme (only needed if IDEFLT = 1):

Card	Format	Column	Contents
B3a	I5	1-5	NPG Number of points in each complex. NPG should be greater or equal to 2. The default value is equal to (2 * number of optimized parameters + 1).
	I5	6-10	NPS Number of points in each sub-complex. NPS should be greater or equal to 2 and should be less than NPG. The default value is equal to (number of optimized parameters + 1).

Card	Format	Column	Contents
	I5	11-15	Number of evolution steps taken by each complex before next shuffling loop. Default value is equal to NPG.
	I5	16-20	Flag to control whether the initial point is included in the starting population. Enter '1' if the initial point is to be included. Enter '0' if the initial point is not to be included. The default value is equal to '1'.
	I5	21-25	Printout control flag. Enter '0' to only display the best set of values for each shuffling loop. Enter '1' to print out every point in the entire sample population during each shuffling loop. Default value is equal to 0.
B4	A8	1-8	Internal identifier for the simulated discharge time series to be used for optimization.
	A4	11-14	Data type code of simulated discharge time series.
	I2	19-20	Time interval of simulated discharge time series. (Statistics computed only for 24-hour values at this time).
	A8	26-33	Internal identifier for the observed discharge time series to be used for optimization.
	A4	36-39	Data type code of observed discharge time series.
	I2	44-45	Time interval of observed discharge time series. (Statistics computed only for 24-hour values at this time.)

Repeat card B5 for each data period to be dropped from optimization criterion computations. (Note: The drop periods must not contain any days of the buffer period. Periods must be in chronological order.)

Card	Format	Column	Contents
B5 (optional)			Date criterion computation stops:
	I2	1-2	Month
	I2	6-7	Day
	I4	11-14	Year

Card	Format	Column	Contents
			Date criterion computation resumes:
	I2	16-17	Month
	I2	21-22	Day
	I4	26-29	Year

Card Group C contains the information on the parameters to be optimized. Repeat cards C1 and C2 for each parameter that is to be optimized (maximum of 16 separate parameters).

Card	Format	Column	Contents
C1	A8	1-8	8-character identifier for the type of Operation.
	A8	11-18	8-character user-supplied name for the Operation (same as in MCP3 input).
	A8	21-28	Parameter name (left justified).
	F5.2	31-35	Parameter increment. Only needed for Pattern Search technique. If percentage is used, recommended value is 0.01.
	F5.2	36-40	Lower constraint on parameter.
	F5.2	41-45	Upper constraint on parameter.
	I5	46-50	Number of other Operations to have ratio or difference maintained for the parameter being optimized.

A separate C2 card is needed for each additional Operation that is to be affected by the optimization of the Operation and parameter defined in card C1.

Card	Format	Column	Contents
C2	A8	1-8	Identifier for the type of Operation (see Section V.3.2).
	A8	11-18	User supplied name for the Operation (same as in MCP3 input).
C3	A4	1-4	'STOP'

Program Execution Information

See Chapter I.2 for information about how to execute the program.

Sample Input

Sample input is shown in Figure 1 and sample output is shown in Figure

2.

Output Data

OPT3 generates the following types of output:

1. Printer output consisting of the time series and Operations used, the optimization options selected, the time series to be used in the optimization criterion, the parameters to be optimized, the results of optimization trials, and the optimized parameter values. An option exists to generate MCP3 output from the Operations, using optimized parameters, on the last pass through the Operations Table.

2. Time series output to the calibration disk files on the last pass through the optimizer using the optimized parameter values. This occurs if the option is specified in card group D in the MCP3 input and if the final MCP3 display pass is specified for OPT3.

3. Punch cards consisting of the optimized input to the Operations for which punch subroutines have been written.

References

Armour, A., 'Adaptive Random Search Evaluated As A Method For Calibration Of The SMA-NWSRFS Model', MS Thesis, Department of Hydrology and Water Resources, University of Arizona, Tucson, AZ 85721, 1990.

Brazil, L.E., 'Multilevel Calibration Strategy for Complex Hydrologic Simulation Models', NOAA Technical Report, NWS 42, Silver Spring, Maryland, 1989.

Duan, Q., 'A Global Optimization Strategy for Efficient and Effective Calibration of Hydrologic Models', PhD dissertation, University of Arizona, Tucson, Arizona, 1991.

McCuen, R.H. and W.M. Snyder, 'A Proposed Index for Comparing Hydrographs', Water Resources Research, December, 1975.

Monro, J.C., 'Direct Search Optimization in Mathematical Modeling and a Watershed Model Application', NOAA Technical Memorandum NWS HYDRO 12, Silver Spring, Maryland, 1971.

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 Resources Research, February, 1983.

Weinig, W., 'Calibration of the SMA-NWSRFS Model using the Adaptive Random Search Algorithm', MS Thesis, Hydrology and Water Resources Department, University of Arizona, 1991.

Figure 1. Sample Input for Program OPT3

```
                                   - Column -

    5   10   15   20   25   30   35   40   45   50   55   60   65   70   75   80
----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+
FRENCH BROAD RIVER BASIN ABOVE BLANTYRE, NORTH CAROLINA
   10 1953   01 1954
DEF-TS
BLANTYRE   MAP     6            INPUT
serfc/Blantyre_MAP
BLANTYRE   RAIM    6
BLANTYRE   SASC   24
BLANTYRE   INFW    6
BLANTYRE   ROCL   24
BLANTYRE   SMZC   24
BLANTYRE   MAT     6            INPUT
serfc/FrenchBroad_MAT
GREENVIL   PTPE   24            INPUT
serfc/E.Greenville_PTPE
ROSMAN     QINE    3            INPUT
serfc/FrenchBroad_QINE
BREVARD    QINE    3            INPUT
serfc/Davidson_QINE
BLANTYRE   SQIN    3
BLANTYRE   QME    24            INPUT
serfc/Blantyre_QME
BLANTYRE   SQME   24
BLANTYRE   QIN     6            INPUT
serfc/FrenchBroad_QIN
END
SNOW-17     BLANTYRE
FRENCH BROAD-BLANTYE 915. 35.0            YES SUMS
    6  BLANTYRE MAP      1.000       BLANTYRE RAIM
  BLANTYRE MAT     6
                                BLANTYRE SASC   24
 1.30 0.90 0.400.100 125.
 0.15 0.50  0.0  1.0 0.10 0.20
 0.12 0.17 0.20 0.22 0.25 0.30 0.38 0.50 0.70
SAC-SMA     BLANTYRE
FRENCH BROAD-BLANTYE          6  BLANTYRE RAIM       BLANTYRE INFW
  BLANTYRE SASC   24  BLANTYRE  BLANTYRE SUMS
                    1.0001.000 85.0 25.00.3000.0350.1000.100    00.250
                      6.0 1.50 180.290.01000.0.100.00500.2000.300 0.0
  GREENVIL PTPE     0.700.500.360.220.321.201.101.101.100.900.750.75
                     70.0  0.0 130.  0.0 250. 200.                          10/53
UNIT-HG     BLANTYRE
FRENCH BROAD-BLANTYE         185.0   22
  BLANTYRE INFW    6  ROSMAN   QINE    3
    0.2600    3.0000    5.2900    3.1700    1.0500    0.7900    0.6300
    0.4700    0.4000    0.3400    0.3000    0.2700    0.2400    0.2100
    0.1800    0.1500    0.1200    0.0900    0.0700    0.0500    0.0300
    0.0200
LAG/K       ROSMAN
ROSMAN   QINE  3 BLANTYRE SQIN  3     5     0
    10.000     0.0      12.000   100.000    18.000   250.000           X
    18.000   300.000     9.000   450.000
     0.0
    0
UNIT-HG     BREVARD
FRENCH BROAD-BLANTYE         125.1   21
  BLANTYRE INFW    6  BREVARD  QINE    3
    0.1800    2.8200    4.0100    1.4800    0.5100    0.3900    0.3500
    0.2900    0.2600    0.2300    0.2000    0.1900    0.1700    0.1300
    0.1100    0.0800    0.0600    0.0500    0.0400    0.0200    0.0100
LAG/K       BREVARD
BREVARD  QINE  3 BLANTYRE SQIN  3     4     0
```

Figure 1. Sample Input for Program OPT3 (continued)

```
                              - Column -

    5   10   15   20   25   30   35   40   45   50   55   60   65   70   75   80
----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+
5.000     0.0       6.000    70.000     9.000   200.000           X
     6.000   350.000
     0.0
    0
UNIT-HG     LOCAL
FRENCH BROAD-BLANTYE          176.0   22
  BLANTYRE INFW    6  BLANTYRE SQIN    3
    0.2500    2.8500    5.0300    3.0200    1.0000    0.7500    0.6000
    0.4500    0.3800    0.3200    0.2900    0.2600    0.2300    0.2000
    0.1700    0.1400    0.1100    0.0800    0.0700    0.0500    0.0300
    0.0200
LAG/K       BLANTYRE
BLANTYRE SQIN  3               0     0    12
     0.0
     9.000     0.0       3.000    50.000     9.000    90.000           X
    36.000   110.000    42.000   130.000    42.000   170.000           X
    36.000   180.000    21.000   200.000    12.000   260.000           X
     6.000   340.000     4.000   400.000     3.000   500.000
    0
MEAN-Q      BLANTYRE
  BLANTYRE SQIN    3  BLANTYRE SQME   24
INSQPLOT    BLANTYRE
FRENCH BROAD-BLANTYE    2    3    1
  BLANTYRE RAIM    6        RAIM+MELT
  BLANTYRE INFW    6        RUNOFF
  BLANTYRE QIN     6        OBSERVED        +
  BLANTYRE SQIN    3        SIMULATED       *
WY-PLOT     BLANTYRE
FRENCH BROAD-BLANTYE    2         767.0 200.  YES
  BLANTYRE QME         OBSERVED        +
  BLANTYRE SQME        SIMULATED       *
  BLANTYRE RAIM    6            BLANTYRE            BLANTYRE
STOP
PATSERCH      3    0 MCP3
    1   60        767.
    1    3    3  .01
BLANTYRE  SQME    24     BLANTYRE  QME     24
SAC-SMA   BLANTYRE  UZTWM       .02  20. 150.
SNOW-17   BLANTYRE  SCF         .02   .8  2.0
STOP
```

Figure 2. Sample Output for Program OPT3

NWSRFS CALIBRATION SYSTEM [illegible]

Figure 2. Sample Output for Program OPT3 (Continued)

13. 1[illegible]

Figure 2. Sample Output for Program OPT3 (Continued)

MINOR* [illegible]

Figure 2. Sample Output for Program OPT3 (Continued)

****UNIT*HG**OPERATIONUNIT*NAME [illegible]

Figure 2. Sample Output for Program OPT3 (Continued)

LAG/K****OPERATION*LAG* [illegible]

Figure 2. Sample Output for Program OPT3 (Continued)

ORDI[illegible]

Figure 2. Sample Output for Program OPT3 (Continued)

[illegible]

Figure 2. Sample Output for Program OPT3 (Continued)

*** [illegible]

Figure 2. Sample Output for Program OPT3 (Continued)

****[illegible]

Figure 2. Sample Output for Program OPT3 (Continued)

****AUTOMATIC*PARAMETER [illegible]

Figure 2. Sample Output for Program OPT3 (Continued)

[illegible]

Figure 2. Sample Output for Program OPT3 (Continued)

4 [illegible]

Figure 2. Sample Output for Program OPT3 (Continued)

'BLANTYRE [illegible]

Figure 2. Sample Output for Program OPT3 (Continued)

[illegible]

Figure 2. Sample Output for Program OPT3 (Continued)

[illegible]

Figure 2. Sample Output for Program OPT3 (Continued)

Figure 2. Sample Output for Program OPT3 (Continued)