

Chapter 7

Step 5 - Calibration of Hydrologic Models

Introduction

Now that all the needed information is available, the variability of physical and climatological factors have been assessed, the calibration locations and periods of record to use were selected, and the data have been checked, analyzed, and put in the proper form, it is finally time to calibrate the hydrologic models. In order to simulate conditions over an entire river basin many models and procedures must be used. Most of these contain variables that must be determined. Some can be determined directly by analyzing physical or experimental data, such as reservoir storage - elevation relationships and spillway rating curves or the drainage area of a particular watershed or local area. Other variables, which constitute the majority, are model parameters that vary from one area to another based on changes in physical factors and climatology. Methods exist for estimating the parameters for many models in an *a priori* fashion based on various information.

Such parameter estimates may be satisfactory for some applications, but seldom provide the accuracy needed for river forecasting. To use the models to produce reliable forecasts both in the short term and for extended periods into the future, requires that a thorough calibration be conducted to determine the appropriate values of the parameters.

This manual will primarily focus on strategies and procedures for calibrating the NWSRFS SNO W-17 and Sacramento soil moisture models, however, in order to compute the flow in the rivers, a number of other models must be used. Even for a headwater drainage with no complications, a model of the channel system is required to convert the runoff into the channel network to discharge at the streamgage location. For downstream local areas, some kind of routing model is needed to translate the flows from upstream to the downstream locations. For points with more complexity within their drainage area, models of reservoir operations, irrigation demands, glacial effects, etc. could be required. These other models will be mentioned in this chapter though not in the same detail as the snow and Sacramento models.

There are many references in the literature to calibration procedures and techniques. The vast majority of these deal with the calibration of models to an individual drainage area, usually a headwater area with few complications. While this manual includes strategies and procedures for calibrating an individual drainage, the primary emphasis is the calibration of an entire river basin which eventually leads to the calibration of the entire area of responsibility of an RFC. To calibrate the models needed for an entire RFC area for river forecasting applications, requires a calibration process that is efficient, that results in spatially consistent parameter values, and that accurately simulates streamflow and other variables under a full range of climatic conditions. If calibration of an entire RFC area is conducted as a series of largely independent efforts with various individuals working on one watershed at a time with only minor coordination, the results will not only fall short of the objectives, but will take much longer to complete than necessary. The strategies and procedures given in this chapter are aimed at fully meeting calibration objectives in the minimum amount of time.

Objectives

There are three basic objectives when calibrating conceptual hydrologic models to an entire river basin for river forecasting applications.

1. Produce a good reproduction of the observed hydrograph at each individual point on the river system. The aim is to achieve a fit that contains the minimum amount of bias possible, i.e. all errors are random. This includes all types of bias including overall bias, bias related to the magnitude of flow, seasonal bias patterns, and bias related to specific snow and soil moisture conditions such as during an abnormally large snow accumulation year or after a long dry spell. Also intermediate variables such as snow water equivalent and soil moisture deficits should compare realistically to any observations of these variables. The amount of random error should be largely a function of the random error associated with the input variables, especially precipitation. Errors in the amount of precipitation, as categorized by the typical spatial variability of this input variable, are the primary reason that lumped models do not produce satisfactory results in some areas of the country as was discussed back in Chapter 1 and illustrated in Figure 1.1.

2. The parameters of the models should function as they are intended. Both the SNOW-17 and Sacramento models are conceptual models which represent, although in a simplistic fashion, the main physical processes that occur in nature. These models were designed to have a physical basis and the parameters control portions of the models that represent specific components of the overall process. The parameters of the Sacramento model were designed to represent items such as the timing and maximum contribution of various runoff components, the maximum soil moisture deficits that can occur, and the rate of the movement of water within the soil profile with changing moisture conditions. The snow model parameters represent such items as the seasonal variation in melt rates when the area is completely snow covered, the areal depletion pattern as the snow melts, and the amount of liquid water that can be held within the snow cover. The effects of each parameter are designed to be reflected in specific sections of the simulated hydrograph under specific soil moisture or snow cover conditions. In order to be consistent with the physical basis of the models and to produce results that will not only best reproduce the full range of historical observations, but also be most likely to extrapolate correctly beyond what was observed in the available historical record, each parameter of the models should be used as it was intended.

3. There should be a realistic variation in parameter values from one area (headwater, local, or subdivision within a drainage) to another within the river basin and with areas just across the divide in adjacent river basins. Changes in parameter values from one area to the next should be explainable based on changes in physiographic factors, climatic conditions, or hydrograph response. Not only is this objective reasonable from a physical point of view, but if adhered to, makes it much easier to monitor and understand operational variations and run time adjustments to state variables.

One question that is often asked, especially by those first learning how to calibrate conceptual hydrologic models, is “when am I done.” Basically the answer is that you are done when the objectives have been met, i.e. when all possible bias has been removed, when the portions of the hydrograph controlled by each parameter are checked to make sure that the parameter is acting as intended, and when any changes in parameter values from one area to another are consistent with the assessment that was made of the spatial variability of hydrologic factors across the region.

When trying to judge whether the objectives have been met, it is important to remember that there is often considerable noise in the input data being used for calibration, especially the precipitation on input. The amount of noise varies with the region of the country and the gage network. There may not only be considerable noise in the data for a given watershed, but the amount of noise can vary from one watershed to another depending on the number of gages available, as well as their location and accuracy of the measurements. Noise in the input data can make it not only difficult to determine the appropriate parameter values for a given watershed, but the variation in the amount of noise from one watershed to another can affect the spatial consistency of the results. This is why the strategy recommended in this chapter starts with calibrating the watershed that has the least noise in the data record. This is also one of the reasons why a sufficiently long period with considerable climatic variability is needed for calibration. Such a period should minimize “curve fitting”, though there still may be considerable uncertainty in the value of parameters that control portions of the models that are seldom activated. Verifying the results on independent data periods should help to reduce this uncertainty and further minimize any “curve fitting”.

After the initial watershed in the basin is calibrated, one must be careful that realistic spatial consistency patterns are not destroyed by “curve fitting” during the calibration of the other drainages in the basin. In order to achieve the proper balance between the calibration objectives, given the amount and variation of noise in the data, the reproduction of the observed hydrograph, at least as measured by goodness of fit statistics, may have to be sacrificed somewhat in order to achieve spatial consistency of the parameters.

It is stated in objective one that all possible bias should be removed. When working with a lumped model, there are certain types of bias that are inherent in how the model is applied. In addition, there are certain model limitations that will cause trends in the results. These factors result in bias that typically cannot be removed including:

- An under simulation of the highest flows. A lumped application of a model uses the average amount of precipitation over an area, whereas in nature the amount of precipitation is seldom uniform. Since the rainfall-runoff process is non-linear, those portions of an area that have rainfall or snowmelt amounts that are greater than the mean areal value will produce relatively more runoff than parts of the area where the amounts are below the mean. This results in the actual runoff being greater than what would be produced by applying the mean value to the entire area. Adjustments to some model parameters, especially the percolation curve in the Sacramento model, can partly adjust for this tendency, but especially in regions where there is typically a large variation in intensity levels during storms over individual drainage areas, a lumped application of a model will under compute runoff during high flow events.

- A over simulation of low flows. When most models are applied in a lumped fashion it is typically assumed that baseflow is being generated over the entire area based on the contents of the groundwater storages, at least that is the case with the Sacramento model. Under the lowest flow conditions, this is likely not the case in nature. Thus, there is a tendency for a lumped application of the Sacramento model to over simulate the lowest flow levels.

- In mountainous areas where snowmelt dominates runoff production, the simulated spring snowmelt typically occurs too early during years with a much below normal snow cover. This occurs because the snow primarily only covers the highest portion of the upper elevation zone, whereas the lumped application assumes the snow is distributed over the entire zone. This situation is described further in Section 6-1 and illustrated in Figure 6-1-3.

- Some biased results when snowmelt is not occurring over the entire area. The model assumes that either melt is taking place over the entire snow covered area or it is not. Especially in mountainous areas, early in the melt season snowmelt may only be occurring at the lowest elevations and on south facing slopes. This can result in some bias in simulating runoff during the first week or so of the snowmelt period.

- The largest snowmelt runoff events are typically under simulated, particularly in regions where high winds and dew-points are associated with major snowmelt situations. This is partly due to the lumped application of the model, but primarily due to using an index to compute snowmelt. In the SNOW-17 model air temperature is used as the sole index to snowmelt.

While temperature is a good indicator of melt under most conditions, during some extreme snowmelt situations the typical relationship between temperature and melt doesn't hold. Especially in the northeast, major snowmelt events are associated with high dew-points and wind speeds. This causes large amounts of latent and sensible energy exchange and alters the normal relationship between air temperature and melt. There are other situations when the relationship between temperature and melt varies from the normal, but these are not associated with a particular level of melt and thus tend to randomly affect flow interval bias computations.

- Rainfall events that occur late in the snowmelt season on watersheds with a prolonged snow depletion period, generally mountainous areas, typically are over simulated. In these situations the soil has dried out in portions of the area that have been bare of snow for sometime, whereas as long as the areal average melt computed by the model, which is coming from only a small part of the area, exceeds the evaporation rate, the soil will remain wet. This can be minimized by using additional elevation zones, but the use of too many zones can cause operational difficulties (see Section 6-1).

- Baseflow recharge can't be modeled consistently when there is a lower zone tension water deficit in the Sacramento model. This is caused by the model assuming that a constant fraction of the area (PFREE parameter) contributes to recharge during this situation when in reality the fraction should undoubtedly vary depending on the size of the deficit, i.e. the ratio of LZTWC/LZTWM. If the ratio is 0.0, little recharge should occur. As the lower zone beco

mes more saturated, the amount of recharge should increase.

Calibration Methods

There are two basic methods used for the calibration of hydrologic models. The first is a guided trial and error procedure where the users knowledge of the model and how each parameter affects the results are used to control changes to parameter values. Decisions as to which parameters to change are made primarily by comparing simulated versus observed values, especially hydrograph plots. This procedure is most effective when interactive, graphical software is available to view the results and make parameter changes. The calibration is finished when the user subjectively determines that the objectives have been met.

The second method is automated calibration [Gupta *et al.*, 2000]. In this method various computer algorithms are used to achieve the best simulated reproduction of observed values, typically mean daily discharge. The algorithms contain strategies for varying the values of user specified parameters in an attempt to obtain an optimal fit. Typically the user can apply limits on the range over which parameter values can vary in the hope of obtaining more physically realistic results. The quality of the reproduction is often determined by a single statistical objective function, such as minimizing the daily root mean square error. Sometimes a series of steps are used where different groups of parameters and different objective functions are used at each step (e.g. the objective function used for parameters affecting low flows may differ from the function used for parameters that primarily control storm runoff) [Hogue *et al.*, 2000]. In some approaches multiple objective functions are used to try to find a group of parameter sets that will produce good results based on several criteria [Gupta *et al.*, 1998]. Then the user can choose subjectively from this group of parameter sets. Automatic optimization has been primarily used for the calibration of individual watersheds, mainly headwater drainages. There are limited strategies available for using automated optimization over entire river basins.

Table 7-1 summarizes some of the features of the interactive trial and error method versus automated optimization of parameters (this table was prepared for a calibration training video developed by the NWS in conjunction with the Hydrologic Research Center [Hydrologic Research Center, 1999]). The biggest difference between the two methods is that the interactive method allows for the user to maintain the physical basis of the models, whereas the automated method relies on various algorithms to achieve a statistical best fit determination of the parameter values. NWS RFS contains software for both methods of calibration. The Interactive Calibration Program (ICP) allows for the user to make parameter changes and view the results graphically. The automatic optimization program (OPT) contains 3 algorithms and several objective functions for use in determining best fit parameter sets. The OPT program was quite useful prior to the time when ICP was first available. Back then trial and error calibration was being done in a time consuming batch mode.

INTERACTIVE vs. AUTOMATED CALIBRATION

INTERACTIVE	AUTOMATED
<ul style="list-style-type: none">• Emphasis on Component Process Representation	<ul style="list-style-type: none">• Emphasis on Overall Model Fit to Data
<ul style="list-style-type: none">• Requires Good Knowledge of Physical Model Basis	<ul style="list-style-type: none">• Treats Model as Nonlinear Regression
<ul style="list-style-type: none">• Person Intensive	<ul style="list-style-type: none">• Low Personnel Requirements
<ul style="list-style-type: none">• Use of a Multitude of Performance Criteria	<ul style="list-style-type: none">• A Small Number of Statistical Criteria
<ul style="list-style-type: none">• Less Affected by Data Quality Problems	<ul style="list-style-type: none">• Sensitive to Data Quality
<ul style="list-style-type: none">• Requires Well Designed Graphical Interfaces	<ul style="list-style-type: none">• Requires Robust Optimization Methods
<ul style="list-style-type: none">• Likely to Produce Parameter Estimates Which Would Allow Reliable Simulations of Future Events	<ul style="list-style-type: none">• Likely to Produce Parameter Estimates with Uncertain Value for the Simulation of Future Events

Table 7-1. Comparison of Interactive and Automated Calibration Methods.

It is this author's belief, based on personal experience, that automated optimization methods can not be used to meet the 2nd and 3rd objectives listed earlier in this chapter for the calibration of conceptual hydrologic models. Because of this assessment, the author has not used automated methods enough in recent years to feel qualified to include a discussion of strategies for using such procedures for calibrating conceptual models in this manual. Other references, such as those mentioned earlier in this subsection, are available for those who want to try automated methods. The biggest obstacle for the successful use of the interactive trial and error method is the time required to develop the knowledge of the model structure and how to isolate the effects of each parameter. While automated methods can in many cases achieve a good reproduction of streamflow for an individual watershed quicker than the interactive trial and error method, by following the procedures and strategies in this chapter, interactive calibration should produce not only a good reproduction of observed conditions, but preserve the physical nature of the models and do so in a very efficient manner for a large area like an entire river basin. After the calibration of the initial headwater area within a river basin, the strategies given in this chapter should produce parameter sets for subsequent drainages that meet all the listed objectives at least as fast as automated methods can meet only the first objective.

Interactive Calibration Program (ICP)

The software used to perform interactive trial and error calibration within NWSRFS is the Interactive Calibration Program (ICP). ICP provides an interactive, graphical interface for the Manual Calibration Program (MCP). MCP contains all of the models and other procedures that are used to generate a simulation and output the results. The input for MCP specifies all of the time series that are to be used, the sequence of operations (i.e. models and other procedures including displays), and parameters and initial state variables for each operation. ICP allows the user to select a watershed, run MCP, display the results, make parameter changes, save specified simulated time series, and resubmit a new MCP run with modified parameter values. This process is continued until the calibration is judged to be finished.

ICP currently contains two graphical displays. These are linked to the WY-PLOT and PLOT-TS operations of MCP. The WY-PLOT display allows the user to graphically plot mean daily flow time series, both simulated and observed. The WY-PLOT display also includes panels that show many of the state variables and internal computations for the SNOW-17 and Sacramento models. These panels are described in detail at the beginning of section 7-7 for the snow model and 7-8 for the Sacramento model. Pull down, tear off menus allow the user to rapidly switch between the snow and Sacramento panels, as well as from one subarea to another for watersheds with multiple zones. The same method can be used to quickly go back and forth between multiple WY-PLOT displays if they are specified in the MCP control file. The user can scroll through the period of record, change the length of the period displayed, alter the range of the plots, and switch from arithmetic to semi-log scales. A single simulated time series can be saved for display

after a subsequent run. This allows the user to compare the results from one run to another.

The PLOT-TS display allows the user to stack from one to six plots on top of each other, all showing the same period. The time series included on each plot must have the same units, but can have different time intervals. The flexibility of the PLOT-TS display allows the user to construct plots to help in visualizing a variety of situations. Just like the WY-PLOT display, the user can scroll through the run period, change the length of the period shown at any one time, alter the range of the plots, and switch from arithmetic to semi-log scales. One time series on each plot can be saved and replotted after a subsequent run.

Other output generated by the operations in MCP can be viewed via ICP in text format. This includes the statistical computations generated by the STAT-QME operation. Graphical displays likely will be added for some of these operations in the future.

ICP contains interactive windows for changing values of selected parameters for the snow, Sacramento, and unit hydrograph operations. Plots are provided for changing multiple valued parameters like the ET-Demand curve for the Sacramento model, the areal depletion curve for the snow model, and the shape of the unit hydrograph. Other changes to the MCP control file are made by editing the file directly.

River Basin Calibration Strategy

In order to meet the calibration objectives in an efficient manner it is necessary to have a clear strategy for calibrating each of the headwater and local areas within a river basin. The recommended strategy is as follows:

1. First calibrate the headwater area with the best data and least complications. The aim of calibration is to determine the model parameter values that will provide the best possible forecasts of future conditions. Also the strategy recommended for a river basin ties the parameter values for all the other drainages within the basin to the parameters determined for the initial headwater area. Therefore, it is critical to start with an area where one has the best possible chance of determining appropriate parameters for the snow and Sacramento models. Complications such as reservoirs, diversions, irrigation, power plants, etc. and noise in the input data make it much more difficult to determine proper parameter values. The more noise caused by errors in the input data and complications that affect the observed output signal, the more difficult it is to see through the noise and determine the best values for the model parameters. Thus, the first area to calibrate should be the one with the best data and least complications. Section 7-1 describes in detail the selection criteria and the strategies and procedures that are recommended for use in calibrating the initial headwater area.
2. Calibrate other areas with minimal complications. These are other headwater areas, including reservoir inflows where the inflow hydrograph doesn't contain significant noise, as well as local areas for which a good definition of the local contribution can be determined. It includes areas where the flows have been adjusted to natural flow conditions and those where t

he flows were not adjusted but the man-made controls don't have a dominating effect. Once the local natural flow contribution is separated out by subtracting routed flows or correcting for man-made features, parameters for these areas can be determined in a similar fashion as the initial headwater area. The difference is that during the calibration of these areas *only the parameter values that clearly needed to be changed should be altered*. It is important not to make parameter changes for these areas merely to improve goodness of fit statistics or to make changes based on single events. Parameter changes should be based on clear evidence over many periods or events. If this philosophy is followed, the result should be a realistic variation in parameter values across the river basin. The assessment of spatial variability done in step 2 of the calibration process (see chapter 4) should provide an insight into the magnitude and types of parameter changes that should be expected. Section 7-2 describes details for determining local area hydrographs and strategies to follow when calibrating these drainage areas.

3. Determine parameters for the remaining headwater and local areas. These are areas where a good definition of the local natural flow cannot be obtained due to the effect of man-made controls or excessive noise in the hydrograph after subtracting routed upstream flows from the discharge at the downstream point. In these areas a true calibration of the snow and soil moisture models is not possible. This step also includes areas for which the data are not available to make adjustments for or model the effect of control structures, thus a historical simulation is not possible. Parameters in both of these cases are generally obtained from nearby areas that were calibrated. Only a few simple adjustments are then made (if even possible) to remove any bias. Section 7-3 describes procedures to follow for determining the appropriate parameters for these remaining drainage areas.

By following this strategy one should be able to meet the calibration objectives in an efficient manner. The calibration of the initial headwater area must be done carefully as the parameters for all the other areas are tied to the values determined for this area. The calibration of this area should take longer than any of the other areas. The calibration of the other areas with minimal complications should go quite quickly as long as one knows how to identify the portion of the hydrograph affected by each model parameter and thus can fairly easily determine which parameters clearly need to be altered. The time required for the completion of the remaining headwater and local areas is primarily a function of how long it takes to determine routing parameters and/or remove or model man-made complications.

Statistics to Monitor

Decisions regarding which parameters need to be changed and the overall reproduction of the observed conditions are primarily determined by examining graphical plots of simulated and observed time series that are produced by the WY-PLOT and PLOT-TS displays of ICP. In addition to viewing these plots, there are various statistics that are helpful to monitor when doing interactive trial and error calibration. The currently available statistics are computed from mean daily flow time series by the STAT-QME operation and are not in graphical form (eventually it would be beneficial to be able to generate statistics for any data type, display many of the results graphic

ally, and compute a wider variety of statistics than are included in STAT-QME). STAT-QME will optionally generate statistics on a water year basis in addition to the multi-year statistics for the entire run period. It is recommended that the user focus on the total run period statistics in order to see trends that may exist in the overall simulation. Yearly statistics are seldom of value. Of the currently available multi-year statistics there are several that are very helpful to periodically check during a calibration (Figures 7-1 and 7-2 show how these statistics appear in the output -- in the ICP program the STAT-QME output is included in the 'Edit Wide Listing' display).

These are:

- Seasonal and overall bias – These are shown in terms of both percentage and depth of runoff. Both columns are useful. The overall bias shows whether the current data and parameters are generating a near zero water balance or whether the models are producing too much or too little overall runoff. For the calibration period the overall bias should at least be less than 5%. The monthly figures indicate how these quantities vary seasonally. The aim is to have a random variation of reasonably small deviations from zero. A definite pattern in the seasonal bias indicates that something needs to be adjusted. Several months in a row of a double digit percent bias in one direction likely indicates a problem.

Table 7-1. Illustration of Overall (Year Avg. Percent), Monthly (Percent and Runoff Depth), and Flow Interval Bias (Percent) Output from the STAT-QME Operation.

MULTIYEAR STATISTICAL SUMMARY										
FRENCH BROAD-ROSMAN			AREA (SQ KM) = 176.00		WATER YEARS 1954 TO 1965					
MONTHLY	SIMULATED MEAN (CMSD)	OBSERVED MEAN (CMSD)	PERCENT BIAS	MONTHLY BIAS (SIM-OBS) (MM)	MAXIMUM ERROR (SIM-OBS) (CMSD)	PERCENT AVERAGE ABSOLUTE ERROR	PERCENT DAILY RMS ERROR	MAX MONTHLY VOLUME ERROR (MM)	PERCENT AVG ABS MONTHLY VOL ERROR	PERCENT MONTHLY RMS ERROR
OCTOBER	5.644	5.493	2.75	2.299	42.470	13.25	47.10	15.322	6.40	8.17
NOVEMBER	4.425	4.579	-3.36	-2.267	-11.530	14.79	25.55	-15.238	11.49	13.38
DECEMBER	6.120	6.064	0.92	0.847	-17.333	12.80	26.10	14.508	6.77	8.36
JANUARY	6.568	6.380	2.94	2.856	-15.996	13.19	23.62	21.007	7.88	9.50
FEBRUARY	8.370	8.202	2.04	2.301	-8.448	11.83	17.65	12.042	4.56	5.31
MARCH	8.773	8.618	1.79	2.351	5.101	11.68	16.26	17.723	6.33	7.07
APRIL	10.218	10.594	-3.55	-5.535	-21.689	10.66	18.01	-36.210	5.48	8.13
MAY	7.310	7.404	-1.27	-1.429	-7.120	11.85	17.46	-13.731	6.12	7.29
JUNE	5.725	5.687	0.68	0.566	5.939	14.34	20.32	26.752	11.70	15.71
JULY	4.560	4.557	0.08	0.054	-5.751	16.22	23.76	-25.436	13.62	18.12
AUGUST	4.762	4.740	0.47	0.336	6.125	12.60	19.76	18.309	7.95	10.39
SEPTEMBER	4.366	4.292	1.74	1.098	11.797	16.24	27.37	26.773	12.99	17.26
YEAR AVG	6.369	6.350	0.31	3.477	42.470	12.88	23.79	-36.210	7.80	10.18
DAILY RMS ERROR (CMSD)	DAILY AVERAGE ABS ERROR (CMSD)	AVERAGE ABS MONTHLY VOL ERROR (MM)	MONTHLY VOLUME RMS ERROR (MM)	CORRELATION COEFFICIENT	LINE OF BEST FIT OBS = A + B*SIM					
1.510	0.818	7.403	9.663	0.9696	A: 0.4717	B: 0.9229				
FLOW INTERVAL	NUMBER OF CASES	SIMULATED MEAN (CMSD)	OBSERVED MEAN (CMSD)	PERCENT BIAS	BIAS (SIM-OBS) (MM)	MAXIMUM ERROR (CMSD)	PERCENT AVG ABS ERROR	PERCENT RMS ERROR		
0.00 - 2.00	305	1.712	1.634	4.75	0.0381	0.931	13.64	18.71		
2.00 - 4.00	1213	2.966	2.979	-0.41	-0.0060	2.120	13.82	17.46		
4.00 - 6.00	916	4.881	4.917	-0.73	-0.0177	4.169	14.00	18.10		
6.00 - 9.00	968	7.250	7.303	-0.73	-0.0261	5.268	12.68	16.23		
9.00 - 14.00	464	11.169	10.813	3.29	0.1745	7.041	11.90	16.76		
14.00 - 30.00	212	18.751	18.440	1.69	0.1526	7.290	11.38	14.29		
30.00 AND ABOVE	32	44.073	46.414	-5.04	-1.1492	42.470	15.20	23.90		

- Flow interval bias – This table indicates whether mean daily discharges are over or under computed for 7 intervals. The preferred option is for the user to specify the intervals (possibly using physically meaningful values such as bankfull and flood flow), though the STAT-QME operation will generate interval ranges if user specified values are not entered. Again percentage and depth of runoff columns are included, but only the percentage column is worth examining. The aim is to have a random variation of reasonably small deviations from zero, though as mentioned under the objectives section of this chapter there is a tendency with lumped models to under simulate the highest flows and over estimate the lowest flows.

- Accumulated flows and errors – This optional table tabulates the accumulated simulated and observed flows and their difference on a quarterly basis (produced when the 'QUAR' input

field is included). Ideally the errors for each period should be random and the accumulated error should meander back and forth around zero. Trends in the quarterly accumulated errors indicate that the relationship between simulated and observed discharge is changing over time. This may indicate an inconsistency in the input or discharge data or a physical change within the watershed that is not being considered. This table doesn't need to be monitored as frequently as the overall, seasonal, and flow interval bias values.

File Edit Search Preferences Shell Windows Help						
/home/aaa/calb_report/figures/accum.temp line 1, col 0, 5135 bytes						
I ACCUMULATED FLOW IN MM						
PERIOD	YEAR	OBSERVED	SIMULATED	ACC ERROR	ERROR THIS PERIOD	
OCTOBER TO DECEMBER	1953	168.06	160.61	-7.45	-7.45	
JANUARY TO MARCH	1954	542.81	525.73	-17.09	-9.63	
APRIL TO JUNE	1954	799.86	780.28	-19.58	-2.50	
JULY TO SEPTEMBER	1954	881.93	865.58	-16.35	3.23	
OCTOBER TO DECEMBER	1954	987.47	966.18	-21.28	-4.93	
JANUARY TO MARCH	1955	1211.08	1214.44	3.36	24.64	
APRIL TO JUNE	1955	1593.08	1584.05	-9.04	-12.40	
JULY TO SEPTEMBER	1955	1838.30	1789.53	-48.77	-39.73	
OCTOBER TO DECEMBER	1955	1961.83	1894.51	-67.32	-18.55	
JANUARY TO MARCH	1956	2220.75	2139.16	-81.58	-14.27	
APRIL TO JUNE	1956	2490.30	2378.58	-111.72	-30.14	
JULY TO SEPTEMBER	1956	2622.37	2484.45	-137.92	-26.20	
OCTOBER TO DECEMBER	1956	2792.93	2644.84	-148.09	-10.17	
JANUARY TO MARCH	1957	3146.74	2982.87	-163.88	-15.79	
APRIL TO JUNE	1957	3607.68	3428.88	-178.80	-14.92	
JULY TO SEPTEMBER	1957	3806.87	3642.86	-164.01	14.79	
OCTOBER TO DECEMBER	1957	4262.06	4084.87	-177.19	-13.18	
JANUARY TO MARCH	1958	4653.86	4462.38	-191.48	-14.30	
APRIL TO JUNE	1958	5062.24	4857.98	-204.25	-12.77	
JULY TO SEPTEMBER	1958	5272.48	5084.24	-188.24	16.01	
OCTOBER TO DECEMBER	1958	5397.99	5212.71	-185.28	2.95	
JANUARY TO MARCH	1959	5671.75	5510.14	-161.60	23.68	
APRIL TO JUNE	1959	6081.86	5935.07	-146.79	14.82	
JULY TO SEPTEMBER	1959	6303.80	6162.87	-140.93	5.86	
OCTOBER TO DECEMBER	1959	6657.06	6541.30	-115.75	25.18	
JANUARY TO MARCH	1960	7095.50	6988.56	-106.94	8.82	
APRIL TO JUNE	1960	7432.55	7322.72	-109.83	-2.89	
JULY TO SEPTEMBER	1960	7638.30	7518.04	-120.25	-10.43	
OCTOBER TO DECEMBER	1960	7825.32	7691.68	-133.65	-13.39	
JANUARY TO MARCH	1961	8138.75	8026.83	-111.92	21.73	
APRIL TO JUNE	1961	8484.82	8369.68	-115.14	-3.22	
JULY TO SEPTEMBER	1961	8878.04	8742.62	-135.43	-20.28	
OCTOBER TO DECEMBER	1961	9243.02	9120.68	-122.34	13.09	
JANUARY TO MARCH	1962	9689.95	9611.56	-78.39	43.95	
APRIL TO JUNE	1962	10068.47	10002.94	-65.52	12.87	
JULY TO SEPTEMBER	1962	10193.77	10156.20	-37.58	27.95	
OCTOBER TO DECEMBER	1962	10349.19	10341.75	-7.44	30.14	
JANUARY TO MARCH	1963	10611.38	10629.90	18.53	25.97	
APRIL TO JUNE	1963	10833.53	10862.24	28.71	10.18	
JULY TO SEPTEMBER	1963	10980.88	10996.57	15.68	-13.02	
OCTOBER TO DECEMBER	1963	11120.88	11123.67	2.79	-12.89	
JANUARY TO MARCH	1964	11546.26	11537.09	-9.17	-11.96	
APRIL TO JUNE	1964	11951.35	11912.78	-38.57	-29.40	
JULY TO SEPTEMBER	1964	12240.48	12260.09	19.61	58.18	
OCTOBER TO DECEMBER	1964	12811.19	12850.57	39.38	19.77	
JANUARY TO MARCH	1965	12811.19	12850.57	39.38	0.00	
APRIL TO JUNE	1965	12811.19	12850.57	39.38	0.00	
JULY TO SEPTEMBER	1965	12811.19	12850.57	39.38	0.00	

Table 7.2. Sample Accumulated Flow and Error Table from the STAT-QME Operation.

While these statistics don't indicate which parameters or data need adjustments, they do give some insight into trends that exist and possibly the parameters and model components that should be examined. Many of the other statistics that are produced by the STAT-QME operation only give an indication of the overall goodness of fit and may be useful when evaluating the results, but they aren't very helpful in providing insights into making changes to parameters.

Flows to Route Downstream

After calibrating each point within the river basin, time series need to be produced and saved for use in calibrating and evaluating the simulation at the next location downstream. In order to best calibrate these downstream locations and to properly evaluate the ability to reproduce historical conditions at downstream gages, the following two time series should be generated at each point:

- Adjusted instantaneous discharge – these time series are used when calibrating the downstream local areas. Ideally a quality controlled observed instantaneous discharge time series whose volume matched that of the USGS mean daily flows would be routed downstream, but such a time series is rarely available. As a substitute an adjusted instantaneous discharge time series can be generated with the ADJUST-Q operation. This operation uses the simulated

instantaneous and observed mean daily flows to create adjusted instantaneous discharge values. Periods of observed instantaneous discharge can also be included. The adjusted instantaneous discharge time series uses the shape information from the simulated time series, but adjusts the daily volume to match the observed mean daily flow. These adjusted instantaneous flow time series are then routed to the next downstream point and subtracted from the observed flow at that location to get the local area contribution which is then used to calibrate the local area. By adjusting to the observed volume, errors in upstream simulations are not propagated downstream. The calibration of the local areas are then based on observed discharges just like the headwater drainages. If observed daily flows are not available for some periods at an upstream gage, the adjusted time series will contain just the simulated values. Such periods should be avoided when calibrating the local area.

- Simulated instantaneous discharges – these time series are used to evaluate the ability to forecast at downstream locations. When forecasting, whether it be a short term or extended forecast, there are no observed data in the future (forecast values may be modified somewhat based on the last available observations of streamflow). To determine how well one can reproduce conditions at downstream points in the forecast mode, simulated discharge for the entire drainage area above the gage needs to be compared to observed flow. If only adjusted discharges are routed downstream, errors and bias from upstream locations will be removed from downstream simulations. Relatively minor errors and bias at upstream points could add up to larger errors and bias downstream.

These two time series are generated and saved after completing the calibration at each location.

For headwaters, both time series can be created in the same run. For downstream points, two runs are needed. To generate the adjusted time series, adjusted flows should be routed down from upstream points in addition to using the ADJUST-Q operation at the downstream location. This insures the best possible simulated instantaneous discharge at the downstream point and will use as much observed volume data as possible if there are periods of missing data at that location. To generate the simulated time series, simulated flows must be routed down from upstream points so that all the flow at the downstream location is made up of simulated values.

At some locations it is impossible to perform a historical simulation due to the lack of the necessary data to adjust or model the effect of control structures. At these locations the sequence of routing simulated flows downstream is broken. Observed flow data at these locations can be used to create instantaneous streamflow values to route downstream (generally instantaneous flows are calculated from observed mean daily discharge data using the CHANGE-T operation and, when available, observed instantaneous flows can be used) for use in calibrating the next downstream location, however, a full historical simulation of the entire drainage area below such points is not possible.

Recalibration

Over time there will undoubtedly be a need to recalibrate all or part of a river basin. Reasons for needing to recalibrate were given in chapter 2. Whether a recalibration is necessary can be the

user's choice in some cases and mandatory in others. Cases when the user must decide if a recalibration is needed include:

- calibration expertise and knowledge has increased and as a result past calibrations, as well as operational performance, are reviewed to determine at what locations a recalibration is likely to improve results (if the increase in expertise and knowledge will likely result in improved data estimates, then the affected historical data record should also be regenerated).
- operational results are unsatisfactory under certain circumstances -- after determining the situations that cause problems, the user must decide if these were the result of cases that were missing, overlooked, or modeled improperly during the calibration – if so, a recalibration is a possible solution (in many of these cases the historical data record must first be extended to include the situations that didn't occur within the previous calibration period), and
- physical changes (such as new agricultural practices, large forest fire, or modifications to land use) have occurred within portions of the basin and based on decreased operational performance it is decided that the effects are significant – if so, a recalibration is needed for the affected portions of the basin (in this case the historical data must first be extended and then only the period after the change was well established would be used to modify the model parameters).
- climatic changes have occurred and it appears that the operational output exhibits trends such as over estimation of runoff during certain seasons which might be attributed to increased evaporation rates - if so, a recalibration should be explored using an extension of the historical record. Adjustments, such as those to climatological average ET-Demand values might not require recalibrating every watershed, but only selected drainages to determine the size of the adjustments to apply over the entire area.

In the first of these cases it is probably best to totally redo the calibration, i.e. calibrate the headwater area with the best data first and then move on to the other watersheds following the recommended guidelines. Trying to start with the parameters from the previous unsatisfactory calibration will only interfere with and negatively influence the process. In the other two cases it is recommended that one start with the existing parameter values and change only those that clearly need to be modified.

Situations when a recalibration is mandatory include:

- existing operational models are being replaced with new models (such as replacing an event API type rainfall runoff model with a conceptual soil moisture accounting model) in an attempt to improve forecast results (the existing historical data may be able to be used to calibrate the new model, however, in many cases the new model may require new data types or the data may need to be in a different form (e.g. continuous as opposed to event data)),
- historical data have been reanalyzed using different procedures (such as switching from the

non-mountainous area technique to a mountainous area procedure) resulting in input time series that are biased as compared to those used for the previous calibration (in this case new model input data must be generated for the entire historical data period),

- new methods which should improve model input (such as determining precipitation estimates from a combination of gage, radar, and other data), but likely produce values that are biased compared to that used in the current calibration, have generated a sufficiently long historical record to be used to adjust the model parameters (in this case the recalibration can use only the period for which consistent data can be determined using the new method), and
- new forecast points are to be established at locations within the basin that were not part of the current calibration thus resulting in a subdivision of existing drainage areas and changes to the routing reaches (frequently the historical data must be extended in order to have as much historical streamflow data as possible at the new locations).

In the first of these cases a complete calibration of the new models is required since the new models are being applied to the basin for the first time. In addition it may be necessary to modify parameters or at least check the performance of other models that were part of the previous calibration. For example, if the rainfall runoff model is being replaced, the parameters for the snow and channel response (e.g. unit hydrograph) models may need to be modified. Even though the snow model should be independent from the rainfall runoff model, its parameter values may have been affected by how the parameters and data for the rainfall runoff model were determined in the previous calibration. The same thing holds true for the channel response model and, in addition, the function of the channel model may vary depending on the structure of the rainfall runoff model (e.g. the Sacramento model requires a different unit hydrograph than an API type model – see section 7.6). As long as the other models were properly calibrated previously, only adjustments that are clearly needed should be made to the parameters for these models.

In the other cases when a recalibration is mandatory, a complete new calibration is generally not required. It should be possible to utilize the parameters from the previous calibration as a starting point and then modify only the values of those parameters that need to be changed. For the snow and soil moisture models changes should be substantial only when there is a significant bias between the new and previous input data. Guidelines for the specific case of using radar based precipitation estimates for operational forecasting are included in chapter 8.