Section 6-6

Mean Daily Flow Data Checks and Adjustments

<u>Consistency Checks</u>

Like precipitation, temperature, and evaporation data that are used to estimate input time series for model calibration, the streamflow data that are used to verify the simulation results should also be checked for consistency. Such a check can serve two purposes. First, a consistency check can reveal inconsistencies caused by factors that are not known to the user. Second, the check can determine whether known changes within the basin have a significant effect on the amount of runoff measured at the gage. Inconsistencies in streamflow data can occur for a variety of reasons including such things as changes in land use (e.g. rural to suburban, agricultural to forest, drainage of wetlands), the construction of new control structures or modifications in the operation of these structures, changes in the amount of water diverted into or out of the watershed, the effect of large forest fires, alterations in agricultural practices (e.g. changes in crops being grown, addition of drain tiles, modifications to irrigation practices), constantly shifting rating curves that aren't updated as often as they should, and changes to measurement methods or practices. Consistency checks could reveal information on both the magnitude and timing of these effects and determine whether the data may need to be adjusted or whether only portions of the period of record should be used for determining model parameters.

Consistency checks of mean daily flow data would be done using double mass analysis as it is for precipitation data. Prior to generating the double mass plots, the daily flow data needs to be converted to depth of runoff or at least the values from all the gaging sites need to be scaled to a common area. As with precipitation, the deviation of the accumulated runoff from the group average for each station would be plotted against the accumulated average runoff for the group. Unlike precipitation, only selected stations should be included in computing the group average. The group average should be computed using only stations that have consistent data, i.e. stations with known significant changes in land use, diversions, control structures, agricultural practices, etc. should not be part of the group average. Stations should be grouped geographically on individual plots as there can be shifts in the runoff relationships between watersheds due to changes in climatic patterns. Such changes are real and no adjustments are needed. Such real shifts would be indicated by having several stations in the same general area showing similar patterns of change in their double mass plots.

In most cases inconsistent runoff data would not be corrected by applying a simple multiplying factor as is done with precipitation. Generally the double mass plots would indicate that an inconsistency might exist and give an idea of its magnitude and timing. Then the reason for the inconsistency needs to be determined by looking into the history of changes that have occurred within the watershed. Once the reason is known, then a decision can be made as to how to adjust the data or whether only a portion of the data record can be used for calibration.

Even though the direct capability for making consistency checks on mean daily flow data doesn't currently exist within the NWSRFS software, there is a way that such checks might be made with some limitations using the existing programs. This would involve the following steps:

1. Convert the mean daily flow data to depth of runoff and create time series. This can be done by using a LOOKUP operation for each streamflow site. The LOOKUP operation allows the user to specify the relationship between any two variables without having to maintain the same dimension or units. Two points are sufficient in this case to define the relationship (e.g. the 0.0,0.0 point and a point that indicates the flow volume that represents 1.0 inch of runoff). The resulting runoff time series would be declared as OUTPUT time series.

2. Use these runoff time series as input to the PXPP program which will generate the information needed to produce a double mass plot (plot displayed either by PXPP or IDMA). The PXPP program will treat these time series as if they were precipitation data. In general this shouldn't matter as far as the double mass analysis is concerned except that the procedure in PXPP for estimating periods of missing data may not be totally appropriate for runoff data, thus one needs to avoid missing data periods when drawing conclusions regarding the consistency of the streamflow data. The missing data estimation procedure in PXPP may work in regions where most of the runoff comes from rainfall, but will undoubtedly have more problems in regions where the timing of snowmelt varies considerably from one gaging site to another. Runoff consistency checks in these regions may be possible using PXPP only during periods when all stations included in the analysis have observed data. Another problem in using PXPP to check the consistency of runoff data, is that PXPP doesn't currently have the capability to eliminate certain stations from the computation of the group average (this feature exists in the MAP program, but the MAP strategy for estimating missing data involving both hourly and daily stations will not work with runoff data derived from mean daily streamflow records). Thus, discharge stations with known consistency problems should not be included as input to PXPP.

Streamflow Data Adjustments

There are two general cases that need to be dealt with regarding possible adjustments to streamflow data. The first case is where there are changes in runoff generation that occur slowly over time, such as those resulting from changes in land use or agricultural practices or from natural events like a large forest fire (fire occurs suddenly, but recovery is slow). The second case is where there are modifications to natural flow conditions caused by man-made changes or natural transfers of water across watershed divides.

When the amount of runoff generated changes slowly over time, it is very difficult to make adjustments to the data. In these cases, the normal solution is to use only the period of record that most closely reflects the current state of the watershed when determining model parameter values via calibration. Generally this is the most recent period, though in some cases the most recent period could be an aberration. For example, if a large forest fire occurred a few years ago and now the watershed has just recovered from the effects of the fire, it would be more appropriate to use the period prior to the fire for calibration.

When there are modifications to flow conditions due to man-made control structures or natural transfers across watershed divides, there are two general alternatives. The first option is to model all of the flow modifications so that the simulation results can be compared directly to the observed streamflow data. This would include modeling of such things as reservoir operations, effects of large lakes, diversions across watershed divides whether controlled by man or by nature, irrigation usage and return flow, and municipal or industrial use within the watershed boundaries. The capabilities for modeling a number of these exist within NWSRFS. If modeling these flow modifications is attempted during calibration, the data required should be available not only historically, but also in real time so that the calibration results can be directly applied operationally. If the real time operation of a control structure differs from the rules used during calibration, the changes would need to be included in the operational setup. There are definite advantages to modeling the flow modifications whenever possible. One advantage is that during operational use there is much less reliance on agencies that operate reservoirs or diversions to provide information on releases or diversion amounts. This results in much more timely forecasts at downstream locations. Another advantage is that probabilistic predictions can be extended further downstream, though the uncertainty of each modeled feature needs to be reflected in the total uncertainty of the prediction at each location.

The second option, when there are man-made or natural flow modifications, is to adjust the historical streamflow data to reflect natural flow conditions. Then the calibration results are compared to this “observed” natural flow. Operationally either the forecast can be natural flow values, or the simulated natural flow must be adjusted in some way based on real time information that is not available historically. On many western rivers where the regulation of flows is very complex based on current irrigation, municipal, industrial, fishery, and power generation demands, the reservoir and diversion operators want a prediction of the amount of natural flow generated from rainfall or snowmelt. Based on these predictions, they determine how the control structures are to be operated. This control information is then used by the RFC to create actual flow forecasts. Also in many western areas extended predictions of natural flow are used to make advance plans on items such as which acreage can likely be planted with crops and the amount of seed, fertilizer, and other agricultural products that probably will be needed. In other regions there are reservoirs and other control structures whose operations would be difficult to model historically. Operationally the RFC computes natural flow from the headwaters and locals and relies on the operating agencies to provide release information.

Some adjustments to historical streamflow data and possible methods of making these adjustments to generate natural flow time series for use when calibrating the models are listed below.

1. Diversions - these are diversions of water across watershed divides for various uses. The ideal solution is to obtain a time series of daily diversion volumes that can be directly added or subtracted from the mean daily discharge at the downstream point. In some cases it may be appropriate to prorate how much of the diversion is applied to each day due to travel time from the diversion location to the river gage. Daily diversion time series are available in some cases along with the other streamflow data provided by the USGS. In other cases it may

be necessary to request this information from whoever operates the diversion. In cases where daily diversion data are not available, there may be tabulations of monthly volumes that are published or easier to obtain from the diversion operator than daily data. These monthly volumes can be prorated into daily time series so that the mean daily flow data at downstream points can be adjusted. While the adjustments are not correct on a daily basis, at least the correct volume adjustment is made to the river data.

2. Reservoirs and lakes - this requires either the calculation of the natural inflow to the pool or the adjustment of downstream flows for storage changes and direct gains or losses such as those from rain or evaporation. The water balance equation for a lake or reservoir can be expressed as:

$$I - O + P \cdot A - E \cdot A + D = \Delta S \qquad (6\text{-}6\text{-}1)$$

where: I = Inflow,
O = Outflow,
P = Precipitation on water surface,
E = Evaporation from the water surface,
A = Water surface area,
D = Diversions into or out of the pool other than at the main outlet, and
ΔS = Change in Storage (each term is in volume units).

In some cases the amount of precipitation and evaporation may be small compared to the inflow and outflow volumes and can be neglected. Many reservoirs and most lakes don’t have any canals or pipes that divert water into or out of the pool or these volumes may be small compared to the main inflow and outflow and thus can be ignored. In other cases diversion data or estimates should be used, as well as estimates of precipitation and evaporation. An MAP time series can be generated to estimate the precipitation falling on the water surface. If the surface freezes during the winter, a rain plus melt time series may be more appropriate. Evaporation losses could be estimated from daily FWS evaporation, ideally adjusted for changes in heat storage, or from calculations of climatological mean lake evaporation.

Natural inflow is computed from Equation 6-6-1 by measurements or estimates of the other quantities. Generally, at least outflow data must be available along with either pool elevation or storage data to compute the ΔS term. To adjust streamflow data at downstream locations for the effect of reservoirs or lakes would require the computation of the combined ΔS, P, E, and D terms in the equation. This quantity would be added to the measured discharge at downstream points. In many cases the daily values of this quantity should be prorated to reflect routing effects between the reservoir or lake and the downstream locations.

3. Springs - a few watersheds in some parts of the country contain large springs which contribute a significant amount, at least at lower flow levels. Some of the water discharged by such springs undoubtedly originates within the watershed boundaries, but a large fraction may

come from outside the immediate drainage. Where such large springs exist there is frequently some information on the magnitude and seasonal variation of the discharge. Sometimes it is periodic flow measurements just downstream from the spring published in the USGS Water Resources Data reports. It is also likely that the outflow from these springs varies from year to year based on changes in precipitation patterns. If sufficient information is available, either a daily time series can be created to represent the outflow from the spring or the CHANLOSS operation can be used to at least account for the general magnitude of the outflow and how it varies seasonally.

4. Flood overflows - in some locations there are overflows across drainage divides during high flow periods. This can occur when two rivers come very close to each other, when flood control dams are built near watershed boundaries, and in regions with very flat terrain. Such occurrences are usually mentioned in the remarks section of the USGS Water Resources Data reports. These reports may indicate when such overflows occurred and may provide information that would be helpful in calculating the magnitude of the flow that crosses the divide.

5. Irrigation Losses - the net loss from irrigation, i.e. evaporation loss minus return flow, is generally difficult to determine. There may be many places within a watershed where water is removed from the river for irrigation and records of withdrawals are difficult to obtain and time consuming to process if they can be obtained. There are at least three general ways for dealing with irrigation losses. The best method is to try to model the losses and the return flows with procedures like the CONS-USE operation. Another technique is to account for the general magnitude of the irrigation effects by subtracting or adding flows on a seasonal basis even though the timing and magnitude of the losses and return flows are not the same every year. This can be done with a feature like the CHANLOSS operation or even mimicked with the Sacramento Model RIVA parameter. The other method is to just recognize that irrigation exists within a basin and not alter model parameters to match observed flows during periods with irrigation effects.