Analysis of NWSRFS Mean Areal Temperature Computations

Eric Anderson, May 2004

Introduction

In the National Weather Service River Forecast System (NWSRFS) different procedures are used to compute Mean Areal Temperature (MAT) estimates for historical analysis and operational forecasting. For historical MATs, which are used for model calibration, values are based solely on maximum and minimum temperature observations from climatological networks. Assumptions are made as to when the observed max and min values occurred and a fixed diurnal variation is used to weight the max and min values to compute 6 hourly MAT estimates. Operationally the MAT input data and procedures vary depending on whether computations are being done for the observed or forecast period. For periods with observed data, instantaneous temperature observations are used along with max/min data to generate MAT estimates. The instantaneous values define the shape of the diurnal temperature variation. For stations with only max/min values, the diurnal variation is based on surrounding stations with instantaneous observations. For future periods, forecast max/min temperatures for the 12Z to 12Z period are used along with a fixed diurnal variation that doesn't vary from one time zone to another. Also, the weights used to define the pattern aren't the same as used for the Historical procedure. The use of different data and procedures potentially produces a temperature bias between the time series used to calibrate hydrologic models (primarily snow accumulation and ablation computations) and the MAT values used operationally. This report describes a method of estimating the magnitude of such a bias and how the bias could vary depending on observation time, season, and time interval during the day. The report uses data from Fairbanks, Alaska to illustrate the method.

Background

The procedure used to compute MAT based on historical data for use in model calibration was developed initially by Anderson (1973). At that time, due to cost and other limitations, the Office of Hydrology (OH) decided that only the daily climatological data and hourly precipitation records would be procured for historical computations, thus a procedure that utilized only max and min temperature observations was required to generate MAT time series. Since the observation times of the stations in the National Climatic Data Center (NCDC) network vary throughout the day, the first step was to assign the time of the max and min values based on when the data were observed. It was assumed that no matter when during a given day the observation occurred that the minimum value occurred during the early morning on that day. For maximum values it was assumed that when the observation was made in the morning, the max had actually occurred during the early afternoon on previous day and when a station had a p.m. observation time, the max occurred on the day when the value was recorded. Max/min values were then estimated for any stations with missing data using a quadrant based procedure and weighting factors based on the inverse of the distance and the elevation difference between stations. Six-hourly average temperatures for the local time periods from midnight to midnight were determined by applying weights to the max and min values for each station based on a fixed diurnal variation determined from spring time data from the Central Sierra Snow Laboratory near Donner Summit in California and the NOAA-ARS Snow Research Station in northeastern Vermont. The 6-hourly estimates for the stations were then used to compute MAT values by assigning weights to the stations for each area of interest. A complete description of the historical NWSRFS MAT procedure is in Section II.7-CALB-MAT of the User=s Manual.

The historical MAT procedure obviously will not properly reproduce the actual daily variation in temperature when the diurnal pattern varies from the typical one of a max in the early to mid afternoon and a min in the early morning hours. In addition, there are problems in other cases even when the max and min occur when expected. When observations are made in the morning and the temperature at that time is much warmer than on the previous morning, the previous day=s minimum is also the minimum observed and used for the current day. When observations are observed in the afternoon and the observed max is much cooler than the temperature 24 hours earlier, the warmer temperature from the previous day is used as the max for the current day. These problems are depicted in detail in Figures 6-4-1 and 6-4-2 of the Calibration Manual prepared by Anderson (2002).

The current procedures for computing MAT for operational forecasting were developed in the late 1980's as part of the complete redesign of the Operational Forecast System (OFS). For the observed data period, 1, 3, or 6 hour instantaneous temperature measurements are used along with max/min data for the period from 12Z to 12Z (the hydrologic day adopted by all River Forecast Centers (RFCs)). The use of instantaneous values allows for determining when the max and min occurred each day and the shape of that day=s diurnal temperature variation. Little consideration was given at that time to the potential for a bias between the OFS MAT values and the historical values used for calibration. There was some hope that instantaneous data would be procured and the historical MAT procedure revised sometime in the near future though this has not happened. For OFS MAT computations for future periods, the most readily available and most numerous predicted values in the late 80's were max/min estimates, thus it was decided to compute future MAT values solely from these meteorological forecasts. The meteorological forecasts consisted of a prediction of the max temperature to occur in the afternoon and the min that would occur in the morning. During periods of steadily rising or falling temperatures the min for a given hydrologic day could be greater than the max. The weights used to convert max and min values into 6-hourly estimates were based on data from the two previously mentioned sites and a synoptic station in the upper Midwest as best that can be remembered. A single set of weights were used even though 12Z varies from 3 a.m. local standard time in Alaska to 7 a.m. in the eastern U.S.. A complete description of the operational NWSRFS MAT procedures is in Section II.7-OFS-MAT of the User=s Manual.

Even a relatively small bias in the MAT values used operationally as compared to those used when the models were calibrated with historical data can produce a significant difference in streamflow response. In the NWSRFS snow model, SNOW-17, MAT is used to determine the form of precipitation (i.e. rain or snow) and the snow cover energy exchange. A bias in the MAT values will operationally result in more or less snow versus rain than would be expected based on historical simulations and cause the melt and thus snow cover runoff to occur too early or too late. Figure 1 shows the effect of a 2EF change in temperature on the timing of snowmelt runoff (the effect of the temperature difference on the form of precipitation is negligible in this case as precipitation seldom occurred when temperatures were near the threshold temperature value used by the model).

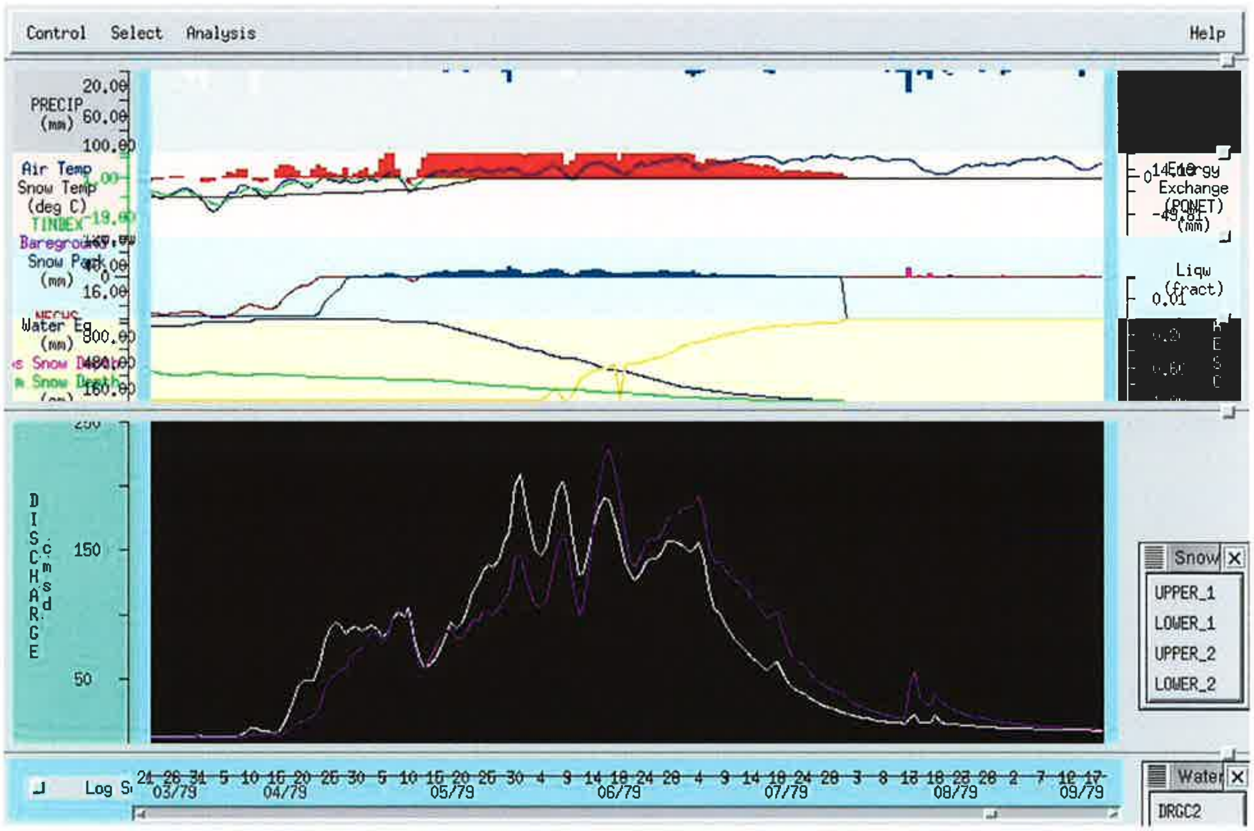

Figure 1. Effect of a 2EF difference on streamflow response - Animas River nr Durango, CO

Comparison Methods

There are several ways that historical and operational MAT values could be compared. This author is not aware that any of these types of comparisons have been made to determine the actual magnitude of a possible bias. Three possibilities for comparing MAT values are:

1. Compare archived operational MAT time series with ones generated for the same period using the historical procedure. Such a comparison would indicate the overall difference between the operational and historical values. The results would not only include differences caused by the different data and procedures used in each case, but also could contain discrepancies due to an improper definition of station means and inconsistencies

at individual sites. These discrepancies may occur when the mean monthly values are defined for operational stations, especially in the case of stations that were not part of the historical analysis, and when stations are relocated or equipment is changed. It would also be difficult to make comparisons for individual 6-hour intervals in many regions since the operational Z time intervals only line up with historical standard time intervals for the Central time zone

2. Generate MAT values using each procedure from historical instantaneous and max/min temperature records for a group of stations. Any data inconsistencies would be removed and mean monthly values properly defined at all sites. The exact same stations would be used for each procedure and the same stations and weights would be used to compute MAT values, thus the comparison would only reflect the difference caused by using different data and computational procedures. The effect of observation times on the estimates produced by the historical procedure would be dependent on historical stations used for the analysis. Again exact individual 6-hour period comparisons would only be possible for the Central time zone.

3. Generate 6-hour temperature estimates using each procedure from historical hourly temperatures at a single station. Max/min values would be determined from the hourly data. The observation time for historical max/min computations could be varied to see how the results were affected. To eliminate problems with directly comparing 6-hour values for most time zones, each estimate could be compared to a >true= value computed by averaging the appropriate hourly temperatures (>true= values would be difficult to determine for the second method described and impossible to determine for the first method).

This report uses method #3 to attempt to get an estimate of magnitude of the differences between the various NWSRFS procedures for computing MAT values.

<u>Description of Method</u>

MAT values are computed using each of the 3 NWSRFS procedures, i.e. historical, operational observed data period, and operational future period, using a complete (no missing data) hourly temperature record and then compared to >true= values. The >true= values are computed by averaging the hourly temperatures for the appropriate 6-hour intervals. The hourly temperatures at the each end of the interval are weighted by 1/12 and the hours within the interval by 1/6. A description of the computations for each of the procedures follows.

1. <u>Historical Procedure</u> (Input is the observation time to be used for computing the daily max and min temperatures)
    a. Compute max and min that would be measured for each day as the highest and lowest hourly value that occurs from the defined observation time on one day to that time on the next day.

   b. Assign max and min values to the day that they would be assumed to have occurred based on the observation time. Minimum values always assigned to the day that they were observed. Maximum temperatures observed in the morning are assigned to the previous day while those observed in the afternoon are assigned to that day.
   c. Compute 6-hour temperatures for each time interval from the max/min values using the equations given in Section II.7-CALB-MAT of the User=s Manual.

2. Operational Observed Data Procedure (Input is local standard time hour corresponding to 12Z and the time interval of instantaneous data to be used, either 3 or 6 hours)
   a. Compute observed max and min values for each day as the highest and lowest hourly values from 12Z on the previous day through 12Z on the current day.
   b. Compute initial 6-hour mean value from the 3 or 6 hour instantaneous values. If 3 hour data, values at the end of each interval are weighted by 0.25 and the value at the middle of the interval is weighted by 0.5. If 6 hour data assumed, the hourly values at the start and end of each interval are averaged.
   c. Adjust 6-hour means based on when the max and min most likely occurred. The max is assumed to have occurred during the 6-hour interval with the highest initial mean and the min during the interval with the lowest initial mean. When 3 hour instantaneous data are being used, the initial mean is weighted by 0.75 and the max or min by 0.25 to get a revised mean. For 6 hour instantaneous data, the initial mean and the max or min are averaged to get the revised mean.

3. Operational Future Data Procedure (Input is the local standard time corresponding to 12Z as input for the observed data procedure)
   a. Compute assumed forecast max and min values for each hydrologic day from hourly temperature data. Max is assumed to be the highest hourly temperature from 10 a.m. to 6 p.m. local time. Min is assumed to be the lowest hourly temperature from 1 a.m. to 9 a.m. local time.
   b. Compute 6-hour mean temperatures by weighting the max and min values using the equations in Section II.7-OFS-MAT of the User=s Manual.

Once 6-hour mean temperatures are computed for each procedure they are compared to the >true= 6-hour mean values (separate 'true' values are computed for the Historical and Operational procedures due to differences in timing) and various statistics computed. The statistics generated are:

1. For all 6-hour intervals and for just those when the temperature exceeds 32EF (computed for the entire period of record and on a monthly basis):
   a. Bias
   b. RMS error
   c. Average absolute error

2. For each 6-hour interval of the day during the melt season (defined by user) when the temperature exceeds 32 EF:
    a. Bias
    b. Number of cases

In addition, the computer program written to implement the comparison method will compute values for the weights used to convert max and min temperatures to 6-hour means that should minimize the resulting bias. The weights are computed for the two procedures, historical and operational future period, that use such a scheme to generate 6-hour values. Weights are computed considering all days and also using just days during the defined melt season. The values computed are the weight to be assigned to the max temperature for each of the 6-hour intervals during the day. The weights are computed by taking the fraction that the >true= temperature for a given 6-hour interval lies between the appropriate max and min and then weighting the fractions for each time interval by the magnitude of the difference between the max and min and calculating the average weighted fraction. The user then has the option to use these weights instead of the weights given in the User=s Manual. For intervals when 2 minimum temperatures are given weight (noon to 6 p.m. for the historical procedure and 18Z to 0Z for the OFS forecast procedure), half of 1.0 minus the max temperature weight is assigned to each of the min values when user defined weights are applied.

Results and Discussion

Results were generated using hourly temperatures for Fairbanks, Alaska for the period January 1998 through September 2003. These data were provided by the Alaska Pacific RFC (APRFC). Results are presented for the following variations of the procedures:

1. H-24 B Historical procedure using an observation time of midnight
2. H-18 B Historical procedure using an observation time of 6 p.m.
3. H-7 B Historical procedure using an observation time of 7 a.m.
4. O-3 B OFS observed data period procedure using 3 hour instantaneous data
5. O-6 B OFS observed data period procedure using 6 hour instantaneous data
6. F B OFS Forecast period procedure

The melt season is defined as April through June for Fairbanks. 12Z is 3 a.m. local standard time, thus the operational time intervals are 3 to 9 a.m., 9 a.m. to 3 p.m., 3 to 9 p.m., and 9 p.m. to 3 a.m. local time. Historical intervals are, of course, midnight to 6 a.m., 6 a.m. to noon, noon to 6 p.m., and 6 p.m. to midnight.

Figure 1 shows the overall bias (i.e. as compared to the >true= values) produced by each procedure for all time periods and for periods when the >true= temperature was greater than 32EF (typical melt base temperature when using the SNOW-17 model). The O-3 procedure has no bias in either case. The overall all period bias for the H-18 and F procedures is also zero. The H-24 and H-7 procedures have a similar large negative bias in both cases, whereas when an observation time of 18 is used, the overall bias is quite small (the reason for this is described later). The Forecast period (F procedure) bias is much larger for temperatures above freezing, whereas for the other procedures the magnitude of the overall bias is similar for above freezing periods as compared to all periods.

Figure 2 shows how the bias for all periods varies seasonally for each of the procedures. The pattern is similar for all of the procedures with the largest negative bias occurring during the melt season. The O-3 procedure has the smallest seasonal bias variation. The Historical and Forecast procedures all have a large seasonal bias pattern. The maximum negative bias of -1.4°F occurs in May and June for the H-24 procedure. Seasonal bias values when considering only periods when the temperature exceeded 32EF were about the same as the values for all intervals except during the November through March period when the temperature is seldom above freezing at Fairbanks, thus there were only a few cases available to compute a bias.

Figure 3 illustrates how the RMS and average absolute errors vary from one procedure to another. The O-3 procedure not only has minimal bias (Figure 1), but also exhibits little variation about the true values. The Historical procedure produces similar RMS and average absolute error values no matter what the observation time of the max and min temperatures even though the bias was much smaller when the observation time was 6 p.m.. The OFS Forecast procedure, like the Historical procedures, exhibits considerable variability about the >true= values. The error values are somewhat smaller for intervals when the temperature is above freezing than when all intervals are considered.

Figure 4 shows how the bias varies during the day from one time interval to the next. As mentioned earlier these bias values are for the melt season and only for intervals when the >true= temperature is above freezing. The O-3 procedure shows very little diurnal pattern in bias values. The Historical procedures all have a similar pattern with a somewhat negative bias for the first interval (midnight to 6 a.m.), a somewhat positive bias for the 2nd and 3rd intervals, and a larger negative bias for the 4th interval of the day. The diurnal bias pattern is the worse for the Forecast procedure with a large positive bias during the 1st interval and a very large negative bias for the 3rd interval. Much of this is due to the fact that same weights are used for this procedure no matter which time zone is being considered. For example, the 3rd period for Fairbanks is from 3 p.m. to 9 p.m local time while in the Eastern time zone it is from 7 p.m. to 1 a.m. The current NWSRFS weights for this procedure seem to be more logical for the eastern portions of the country.

Figure 5 shows the computed weights for the Forecast procedure for each time interval for Fairbanks and how they compare with the weights currently in the OFS MAT preprocessor. As expected the computed weights for the 1st and 3rd intervals are much different than the weights given in the User=s Manual.

Figure 6 shows the diurnal bias pattern during the melt season that results when the computed weights are applied for the Forecast procedure and how this pattern compares to that generated using the weights currently in OFS. The computed weights produce a much improved pattern.

Figure 7 illustrates how the computed weights for the Historical procedures compare with those in the User=s Manual. These weights don=t differ as much from the values currently in use as those for the Forecast procedure. The difference in the computed weights from the current NWSRFS weights is as would be expected from examining Figure 4.

Figure 8 shows how the overall error statistics vary for the Historical and Forecast procedures when computed weights are applied as opposed to using the weights given in the User=s Manual. In this figure the computed weights are based on all days. While the overall bias improves somewhat or stays essentially the same, the amount of variability in the 6-hour temperatures about the >true= values is reduced in all cases when computed rather than User Manual weights are used.

The program written to perform the MAT analysis also has an option to generate NWSRFS datacard images of the MAT time series. These MAT time series can then be used as input to the NWSRFS hydrologic models. Figure 9 shows the results from using each of the Historical procedure time series ('true', H-24, H-7, and H-18) as input to the SNOW-17 model. Precipitation data needed by SNOW-17 were generated using hourly and daily precipitation gage records from Fairbanks. The parameters of the snow model were then determined by comparing simulated and observed water-equivalent and snow depth. For Figure 9 and the subsequent 2 figures, the precipitation data were multiplied by 2 in order to have a longer snowmelt period over which to compare the results produced by using the different MAT time series.

In Figure 9 the MAT time series generated with the H-24 and H-7 procedures produce rain+melt (RAIM) values that are fairly similar to that produced by the 'true' temperatures on almost all days. The only difference is that the H-24 and H-7 procedures have a negative bias as compared to the 'true' values (expected based on Figure 2 values for late April and early May) thus, it takes several additional days for all the snow to melt. The snow cover is gone by May 8^{th} for the 'true' case while the snow doesn't disappear until May 11^{th} for the H-24 and H-7 procedures. For the H-18 procedure the snow disappears at about the same time as when using the 'true' temperatures. This is somewhat expected considering that the H-18 procedure shows less bias than the other two Historical procedures in Figures 1 and 2. In looking closely at the RAIM plot in Figure 9 it becomes apparent as to why the H-18 procedure has minimal overall bias and yet exhibits RMS and average absolute errors of the same magnitude as the H-24 and H-7 procedures. On several of the days the H-18 procedure generates much greater RAIM values than any of the other procedures. This occurs on days when the afternoon temperature is quite a bit cooler than on the day before even though the typical diurnal variation exists (on Figure 9 this is apparent because the RAIM values for the other procedures for these days are much lower than on the previous day). On such days the maximum for the 24 hour period ending at 6 p.m. occurs at the very start of the interval, i.e. at 6 p.m. on the previous day. When the max/min values are observed at 7 a.m. or midnight this situation doesn't occur.

Figures 10 and 11 illustrate how the different Historical procedure MAT time series could affect streamflow simulations. The SNOW-17 output produced with the different MAT time series was used as input to the Sacramento model. An artificial areal depletion curve was added to the snow model to represent watershed conditions. The runoff from the Sacramento model was then passed through a unit hydrograph to generate streamflow. Parameters for the Sacramento model and the unit hydrograph were selected to illustrate the possible response and were not based on any calibration. Figure 10 show that the different MAT time series produce a variation in the timing of the snowmelt runoff response. There was essentially no rain during the melt period that year (there was a rain event right after the snow cover disappeared). During 2003, as shown in Figure 11, there was a precipitation event early during the melt period. Most of the precipitation for this event was typed as rain for the 'true', H-24, and H-18 procedures; however, the temperatures during the event must have dropped below the rain-snow threshold for the H-7 procedure as a large portion of the precipitation was assumed to be in the form of snow.

Conclusions

Based on the analysis of MAT produced using the different NWSRFS procedures at Fairbanks, Alaska the following conclusions seem justified.

1. There is definitely a difference between the MAT time series produced using the Historical, Operational Observed Period, and Operational Forecast Period procedures. There is also a difference between the results from the Historical procedure when different observation times for max/min temperatures are assumed. When examining the diurnal pattern of the 'true' MAT time series for Fairbanks during the melt seasons, it is apparent that there are very few days when the max and min temperatures occur at a time other than mid afternoon and early morning, respectively. Also, being at a far northern latitude, Fairbanks has some potential differences in the shape of the diurnal temperature pattern than sites in the lower 48 states. This means that one can't infer that the results from Fairbanks will necessarily apply at other locations.

2. The Operational Observed Period procedure generates MAT time series that are warmer than the Historical procedure during the melt season. The difference varies from about 0.5˚F to around 1.2˚F depending on the observation time used for the Historical procedure. This would indicate that snowmelt runoff would generally occur somewhat earlier for operational forecasts than was experienced during calibration.

3. The Operational Observed Period procedure produces results that are closest to the 'true' MAT. This procedure not only has a minimal bias, but also exhibits very little variability about the 'true' temperatures. This indicates that using instantaneous, in addition to max/min, temperature data will improve the results. It is also noted that using 3 hour instantaneous values produces better results than 6 hourly observations. Ideally the same data and algorithms that are currently used for the Operational Observed Period procedure should be used for all the NWSRFS MAT computations.

4. Improvements are possible with the procedures that use only max/min data, i.e. the Historical and OFS Forecast procedures, by allowing the users to input the weights

applied during each 6 hour period to compute the MAT value based on max/min observations. More appropriate weights for a given region could likely be determined by running this type of analysis on one or more stations with hourly temperature data within the region. Allowing user specified weights would be a relatively easy change to make to the Historical MAT processing program and the OFS Forecast computations.

5. If nothing else is done, at least the equations used by the OFS Forecast procedure should be changed to account for the variation in time zones between the RFCs. The current weights used in the equations are clearly inappropriate for use in Alaska. The current equations seem most likely to produce more realistic values in the eastern portion of the country. The further west you go, the greater the chance that there will be a significant bias generated during some periods of the day.

Since it is not known whether all the results from the analysis for Fairbanks would apply at other locations, it is recommended that hourly temperature data be obtained for a number of other sites with varying climatic conditions before making more definitive general statements about the differences between the various procedures.

References

1. Anderson, Eric A., "National Weather Service River Forecast System – Snow Accumulation and Ablation Model," *NOAA Technical Memorandum NWS HYDRO-17*, U.S. Dept. of Commerce, Silver Spring, MD, 217 pp., November 1973.
2. Anderson, Eric, "Calibration of Conceptual Hydrologic Models for Use in River Forecasting," (can be obtained on the web at: http://www.nws.noaa.gov/oh/hrl/calb/calbmain.htm), August 2002.
3. National Weather Service River Forecast System (NWSRFS) User's Manual, (can be obtained on the web at: http://www.nws.noaa.gov/oh/hrl/nwsrfs/users_manual/htm/formats.php)

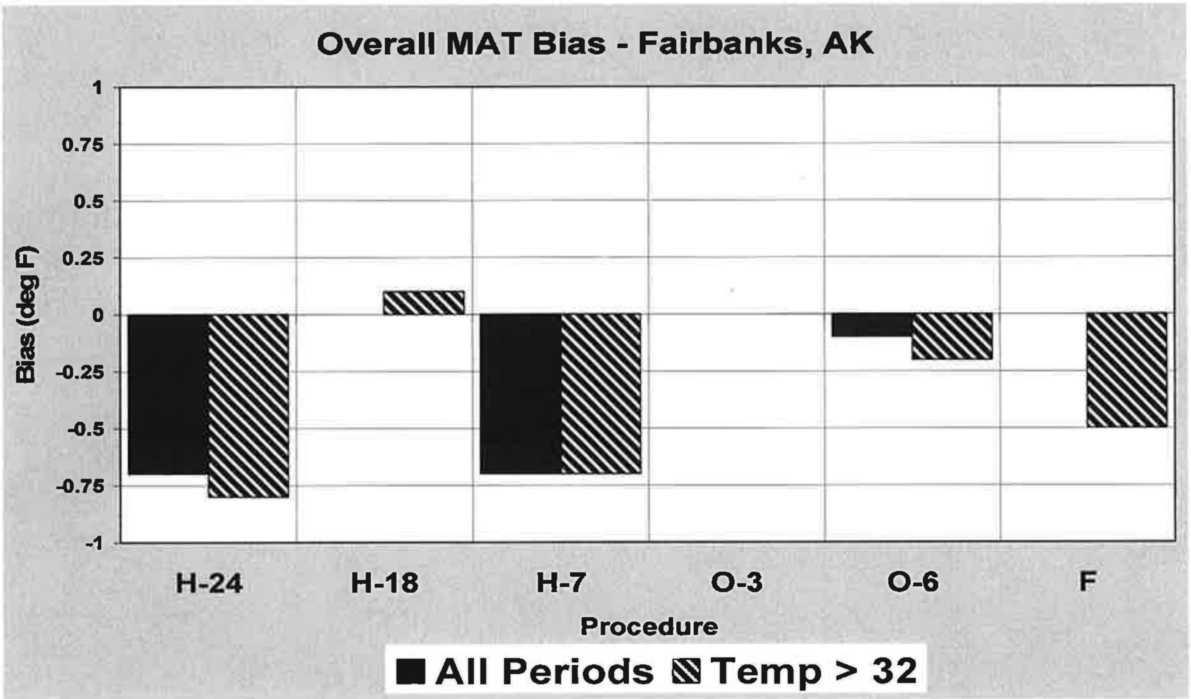

Figure 1 - Overall bias comparison

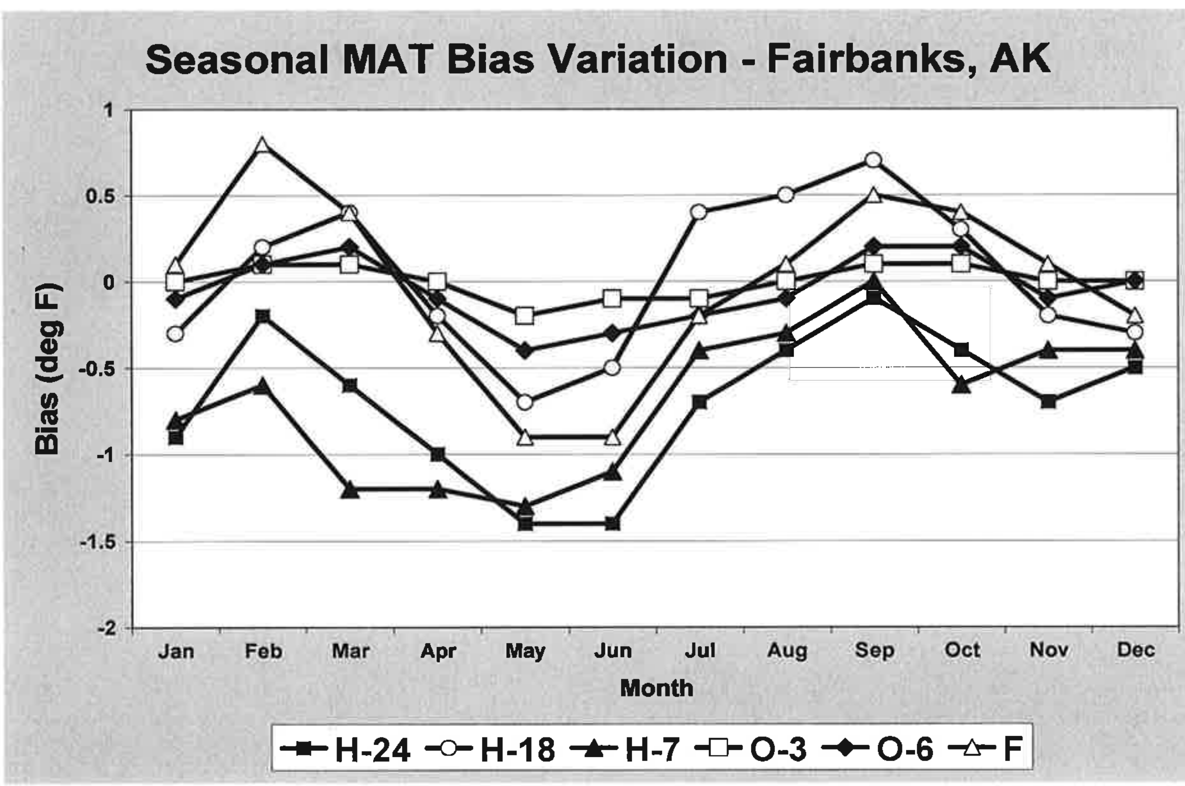

Figure 2 - Seasonal bias comparison considering all time periods.

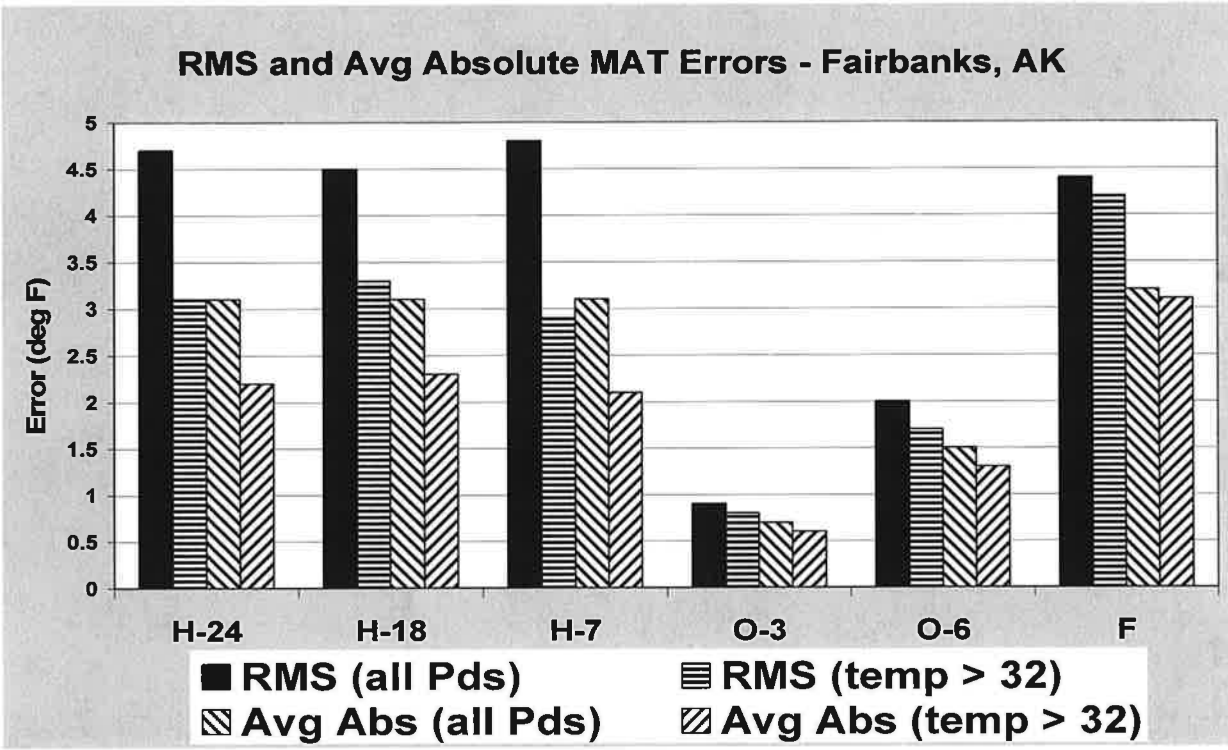

Figure 3 - RMS and Average Absolute error comparison

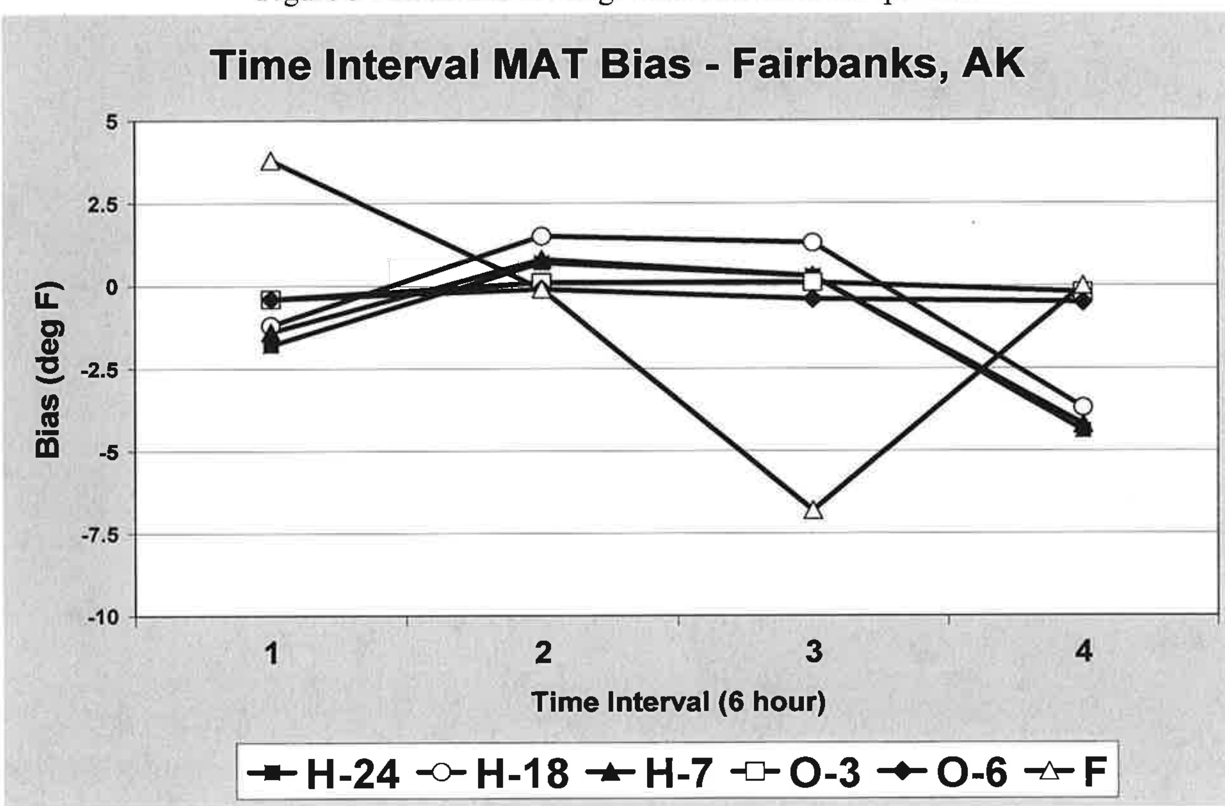

Figure 4 - Time interval bias for each of the procedures - Melt season, temperature > 32EF.

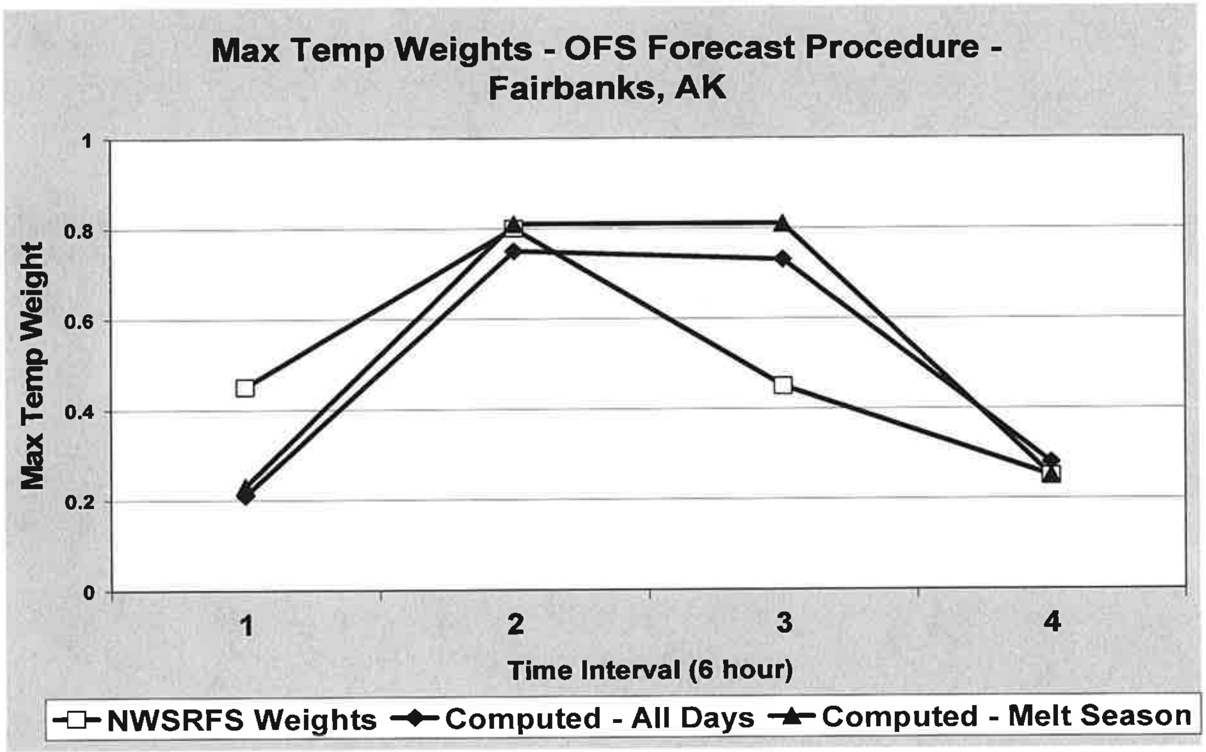

Figure 5 - Comparison of computed and User Manual weights for the OFS Forecast Procedure.

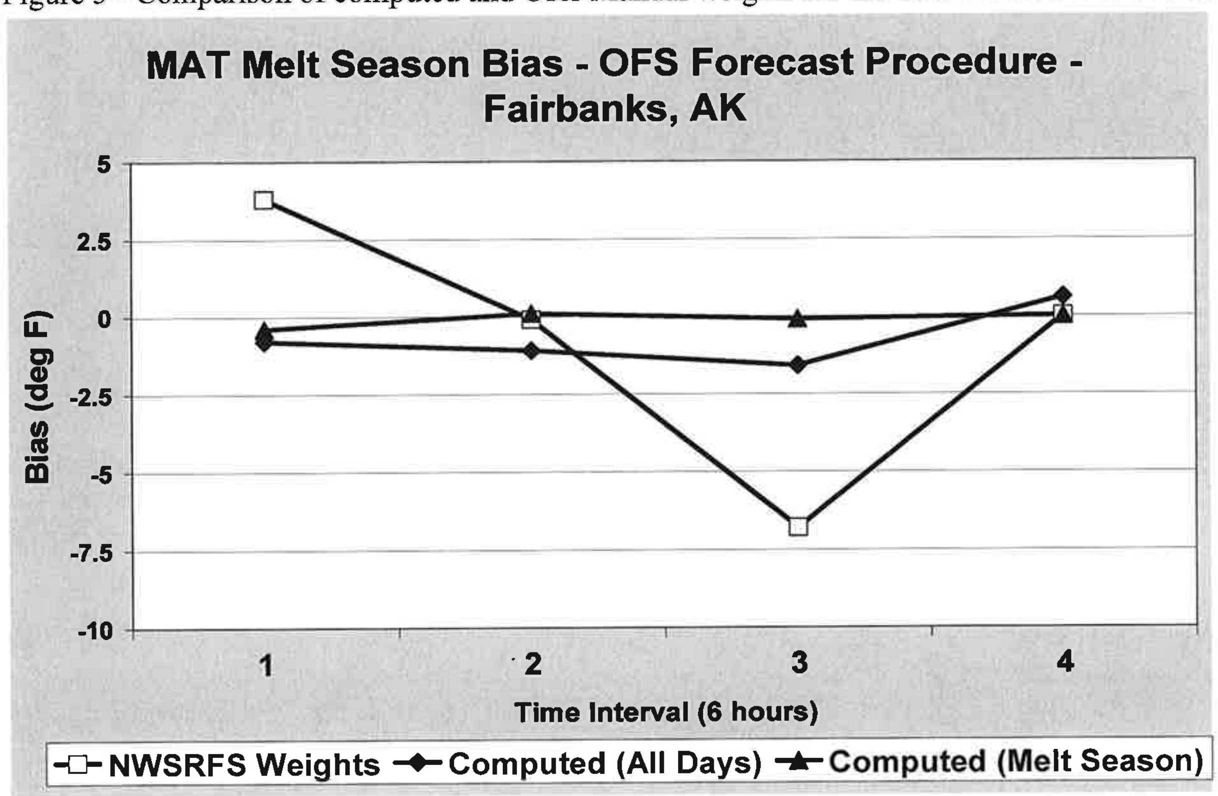

Figure 6 - Time interval bias comparison using various weights - OFS Forecast procedure.

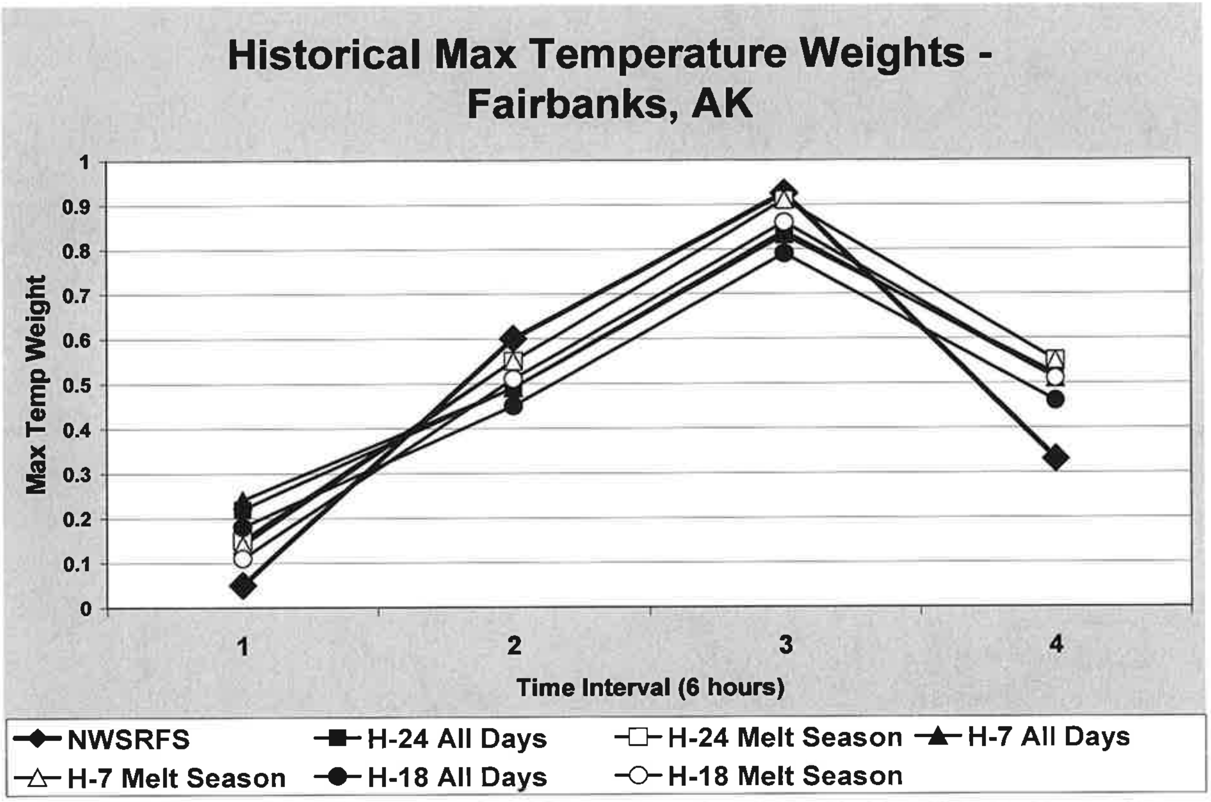

Figure 7 - Comparison of computed and User Manual weights for the Historical procedure.

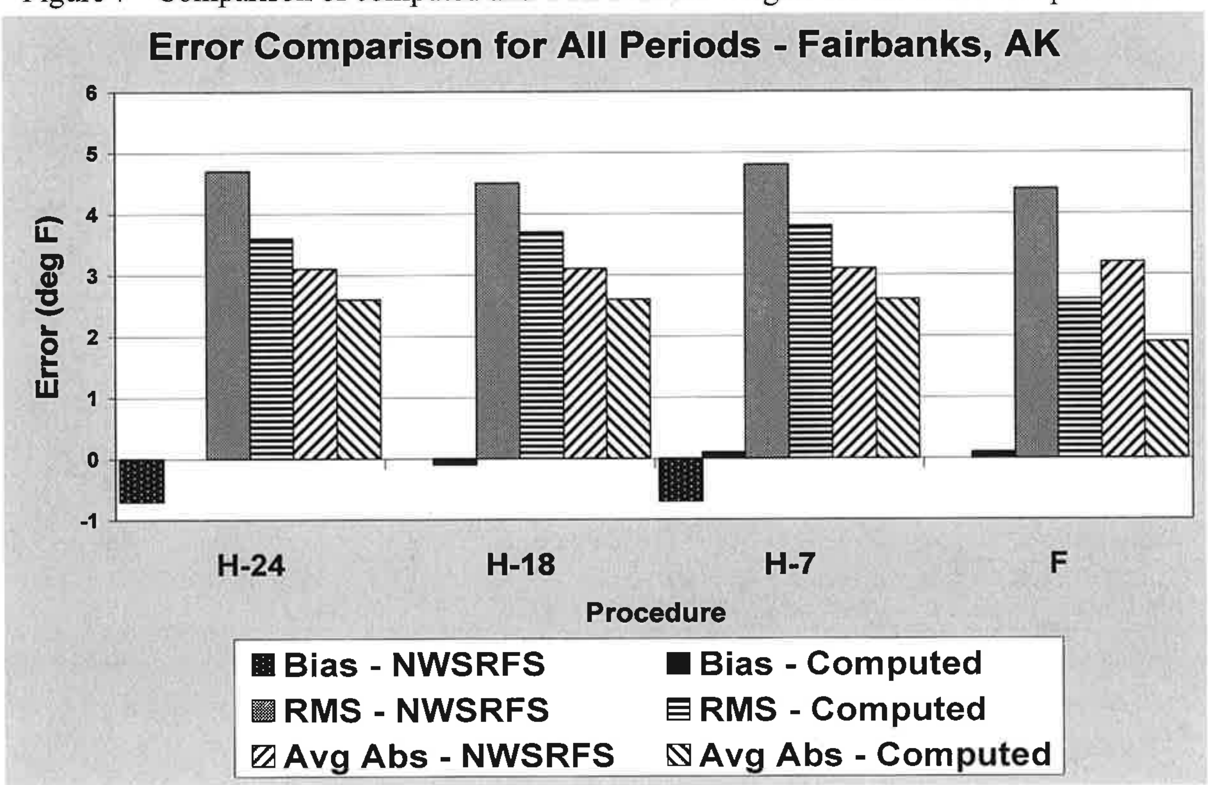

Figure 8 - Comparison of errors from computed and User Manual weights.

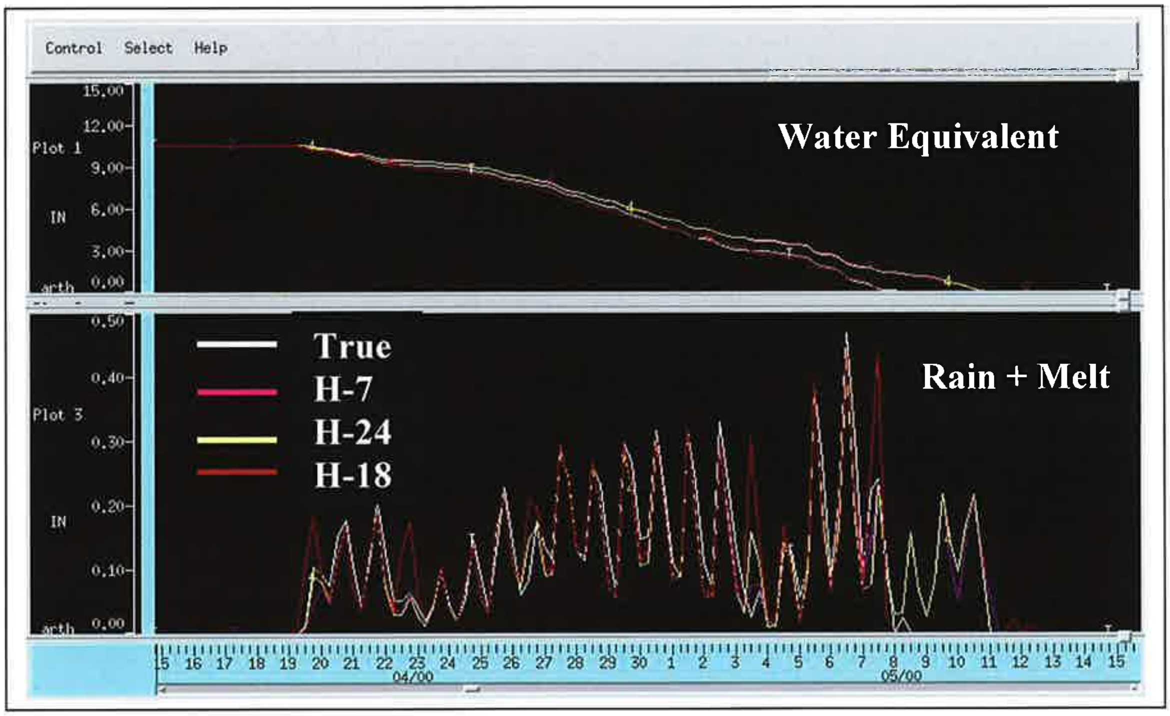

Figure 9 - SNOW-17 output generated with the Historical procedure MATs - 2001.

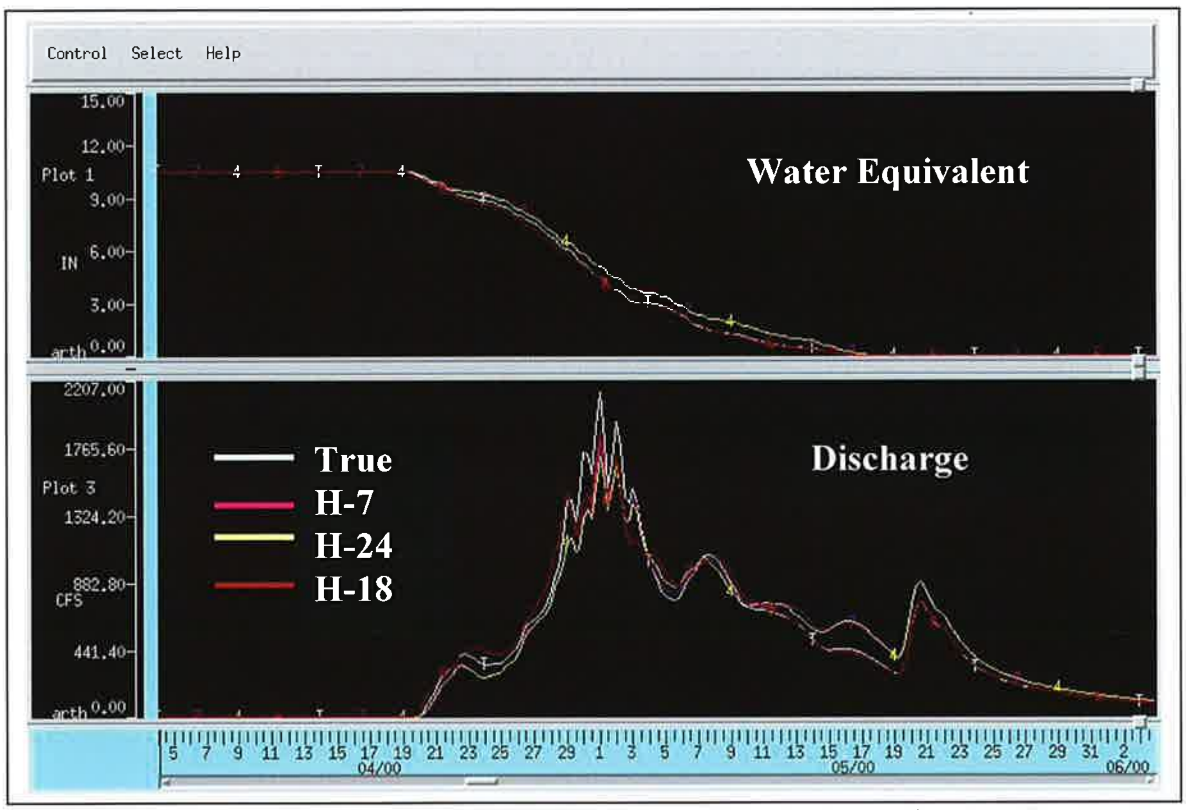

Figure 10 - Streamflow generated using the various Historical procedure MATs – 2001.

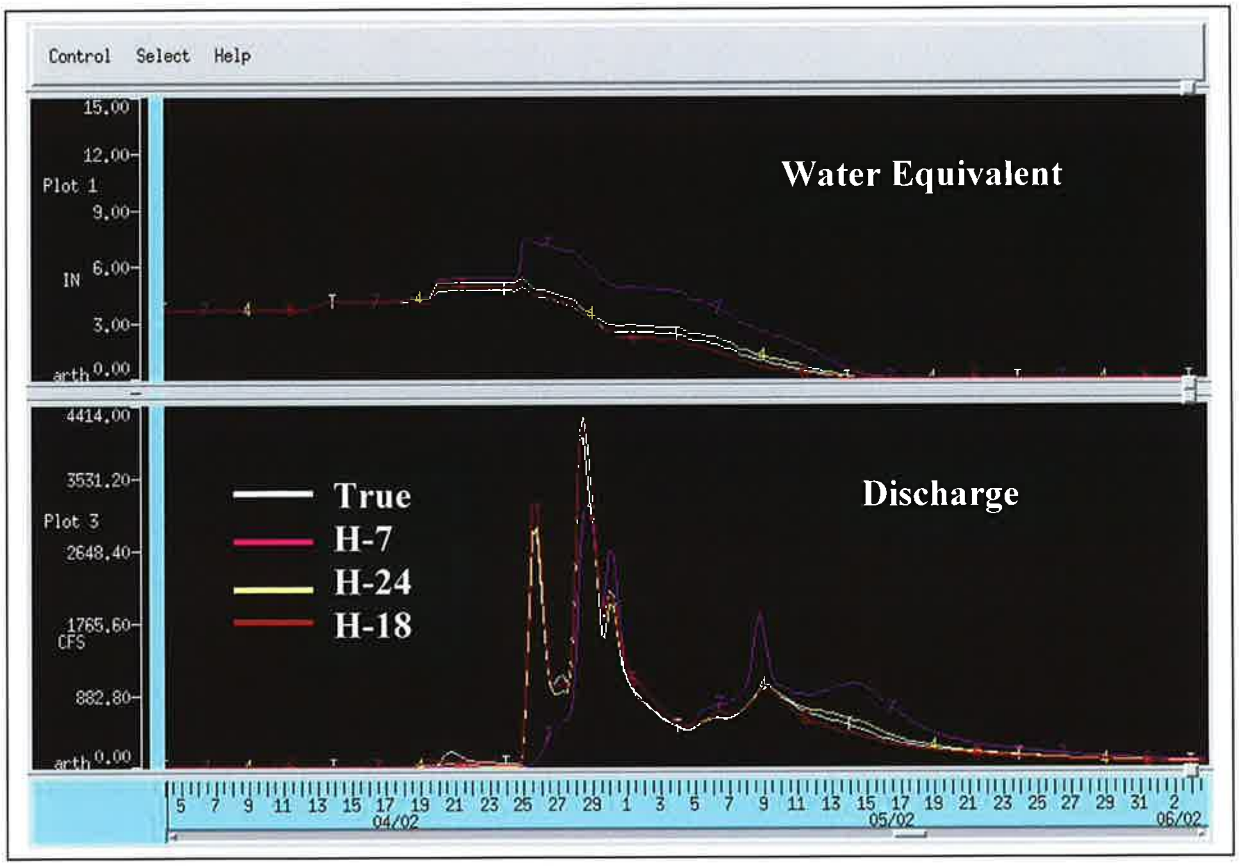

Figure 11 - Streamflow generated using the various Historical procedure MATs – 2003.