IV.1.2-MAT STEPS TO FOLLOW WHEN COMPUTING MEAN AREAL TEMPERATURE (MAT) FOR MODEL CALIBRATION

Purpose

This Section describes the steps follow when computing mean areal temperature (MAT) values used for model calibration.

If these steps are followed then the resulting MAT estimate should have a minimal variability about the 'true' value considering the available data network and should be a reasonably unbiased estimate.

Introduction

MAT data values are required by the snow accumulation and ablation model and is used for two purposes:

- to determine the form of precipitation (rain or snow)
- as the index to energy exchange across the snow-air interface

Since air temperature has a dominant influence on the volume and timing of snow cover outflow, good simulation results depend on reasonably accurate and unbiased estimates of MAT. A large bias in the MAT data can have several detrimental effects on model calibration and the subsequent use of the model for river forecasting:

- Model parameter values will likely be distorted in order to try to compensate for the biased data. Among other things, this could cause large errors in simulated streamflow under abnormal conditions.
- It could be much more difficult to insure that MAT estimates derived from the operational network have the same long-term means as the MAT data used for calibration.
- The problem of incorporating new stations into the operational network will be more difficult.

It is assumed that the user has a basic knowledge of temperature variations that exist in nature and of the factors which cause these variations and that certain decisions have been made prior to beginning the MAT computational process. These decisions include:

- selecting the basin to be calibrated
- selecting the period of record
- delineating any subareas

In addition the user should:

- obtain inventories of the daily observation data

- be familiar with the techniques described in Chapter II.7 for computing MAT values

Step 1: Determine the Relationship between Temperature and Elevation on a Regional Basis

If you are sure that the long-term mean temperature at all locations within the basin can be considered to be the same, then omit this step. If any doubt exists as to whether the means can be assumed equal, then do this step.

The results of this regional temperature-elevation analysis will enable the user to decide if the mean temperature can be considered to be the same at all locations, so that all basins within the region can be treated as non-mountainous areas in terms of MAT computations. On the other hand, if the mean temperature cannot be considered as being equal at all locations, then all basins within the region should be treated as mountainous areas.

In order to analyze variations of any meteorological variable over a given area, a logical first step is to plot the data against the factor which is most highly correlated with the variability. The major factor influencing temperature variations in mountainous areas is elevation. Since temperature-elevation relationships are normally similar (have the same general slope) over fairly large areas, this analysis can be performed on a regional basis. Thus, one regional relationship can be applied to may basins:

- as a base relationship for each basin within the region
- to determine mean monthly maximum and minimum temperatures at synthetic stations
- to obtain lapse rates for input to the MCP program

Step 1.1

Pick a region over which the temperature-elevation relationships should be reasonably similar and which includes the basin or basins to be calibrated. For example, in Colorado a temperature-elevation relationship was developed for the Colorado River above Grand Junction, Colorado (including the Gunnison River), an area of about 17,5000 square miles.

Step 1.2

Select stations with long records and very little missing data. A description of how to determine the available stations and the information about each station is given under Step 2 of this section. Use a period of about 15 to 20 years for this analysis. Most stations should have data for the entire period to insure a stable relationship. However some high elevation stations with data for less than the entire

period should be included if there is a shortage of high elevation data.

Step 1.3

Pick those stations that seem to have the best data, taking into account that the entire elevation range of the region should be represented to the fullest extent possible. Program DLYTRAN can be used to get the time series data for the selected stations and is available at the Internet address http://dipper.nws.noaa.gov/hdsb/data/archived/index.html. From that program, obtain a summary of both maximum and minimum air temperature for each station. Also, obtain punched cards containing the mean monthly maximum and minimum temperatures for the entire period. Do not transfer the data into a data file.

Step 1.4

Run program TAPLOT (Chapter III.8-TAPLOT). This program uses the cards generated by DLYTRAN and plots mean monthly maximum and minimum temperatures against elevation for each station. A separate plot is generated for each month of the year.

It might seem that the consistency of the maximum-minimum temperature data at each station should be checked first. However consistency checks are probably not needed if stations with long, uninterrupted records and very little missing data are used. Any inconsistencies would, in most cases, only change the long-term mean temperature by less than 1 degree Centigrade at any given station. This would have a negligible effect on the regional temperature-elevation relationship.

Step 1.5

Using the plots generated by program TAPLOT, draw the temperature versus elevation relationship for the region. For each month there will be one relationship for maximum temperature and one for minimum temperature. Figure 1 shows such a relationship. In this case, the maximum temperature-elevation relationship is linear over the elevation range. However the minimum temperature data shows a decreasing lapse rate with elevation. It is common in most mountainous areas for the lapse rates to vary seasonally. This is the main reason for analyzing the variations in temperature on a monthly rather than annual basis.

After examining these relationships, the user should be able to decide if the region should be treated as a mountainous or non-mountainous area in terms of MAT computations.

Step 2: Locate Available Maximum-Minimum Temperature Data for Basin MAT Computations

At this point the user should consider whether MAT values are

going to be computed for a single basin or whether MAT computations for several basins are to be done at one time. If all or several basins in a given region are to be eventually calibrated, time can be saved by doing the MAT computations for several basins at once. The only limits to the number of basins that can be done at one time are the dimensions of program MAT (Section III.8-MAT) and the space needed for storage of maximum-minimum temperature and MAT data.

Step 2.1

List all of the stations in the climatological network that have recorded maximum-minimum temperature data during the period of record being used for calibration and are located in or near the basin or basins to be calibrated. This information can be obtained from 'Climatological Data for the United States' published by the National Climatological Data Center, NOAA. By looking at the annual summaries for the first, last and a few of the intermediate years, a list of all stations which collected a reasonable amount of maximum-minimum temperature data can be quickly compiled. In addition to listing the station names, the index number, elevation, latitude and longitude for each station should be tabulated.

Step 2.2

Check the daily observation data inventories to determine the exact length of record for each station and the number of months with all missing data. This can be done using program STAINV which is available at the Internet address http://dipper.nws.noaa.gov/hdsb/data/archived/index.html.

Step 2.3

Select the maximum-minimum temperature stations to be included in the MAT computations. The stations selected should meet the following criteria:

- Stations should generally be located in or near the basin.
- If no high elevation data are available within or close to the basin, high elevation stations located more distant from the basin should be included.
- Each station used should have at least 4 to 5 years of data. These data are needed to adequately define the mean monthly temperatures.

A minimum of about 5 stations should be used to insure that there will not be any days with missing data at all stations.

Step 2.4

Tabulate the observation time for each station at the beginning of the period of record. Observation times can be obtained from

the monthly summaries of 'Climatological Data for the United States'. Also tabulate, for each station, any changes in observation time that occur during the period of record. It is most important to document changes in observation time from AM to PM or vice versa.

Step 3: Transfer Maximum-Minimum Temperature Data to Data Files

Program DLYTRAN is used to transfer the maximum-minimum temperature data for the selected stations to the data files. It is necessary to transfer only that portion of the period of record for which each station actually recorded maximum-minimum temperature data. Program MAT will automatically fill in missing data for the remainder of the period of record. Besides transferring the maximum-minimum data into the data files, it is a good idea to request a data summary. Also, the option to output mean monthly maximum and minimum temperatures on cards should be used. The cards are needed as input to program MAT (mountainous regions only).

Step 4: Check the Consistency of the Maximum-Minimum Temperature Data

Program MAT (Chapter III.8-MAT) is used to check the consistency of the maximum-minimum temperature data being used in the analysis. To check the consistency of the data, run MAT with the option to write to data file set to no and the consistency check option set to yes. In this case, the plots used to check the consistency of the data are generated but MAT values are not computed.

Most of the input required by program MAT (such as station location, elevation, observation time and mean monthly temperatures) has been obtained during the previous steps. Still to be specified are station Fe values and station assignments to groups for consistency checks.

Step 4.1

For mountainous regions, set the Fe value for both maximum and minimum temperatures equal to 50 for all stations. This is a good default value and is completely adequate at this stage of the analysis. The mean monthly maximum and minimum temperature cards output by program DLYTRAN have a value of 50 in columns 16-20. The cards can be used as input to program MAT.

Step 4.2

Assign stations to groups for checking the consistency of the data. Normally, all stations can be assigned to the same group for this purpose. In cases when either there is a wide range of station elevations or the stations are scattered over a large area, it is probably preferable to separate the stations into 2 or sometimes 3 groups as a function of elevation and/or location. In any case, all stations within a group should normally be assigned positive sequence numbers so that the group base consists of a large number of stations. The more stations that are in the group base, the less effect an inconsistency in one

station will have on the consistency plots for the other stations in the group.

Step 4.3

Run program MAT in order to obtain consistency plots for maximum and minimum temperature at each station.

Step 5: Adjust for Inconsistencies in the Maximum-Minimum Temperature Data

The consistency plots generated by program MAT are used to determine if inconsistencies exist in the data for a given station, to identify the time period involved and to determine the magnitude of the adjustment required to make the data consistent.

Step 5.1

Determine if inconsistencies exist in the maximum or minimum temperature data at any of the stations. Inconsistencies show up as distinct breaks in the slope of the plots. The change in slope should, in general, persist for at least a year before it can definitely be concluded that there was a shift in the temperature data. Changes in slope which last for shorter periods are often the result of climatic fluctuations. If you doubt that a change in slope indicates an inconsistency in the data, ignore that case because the effect on the final MAT data will undoubtedly be negligible.

It should also be noted that estimated data cannot be adjusted directly. However if all the actual data are consistent, the estimated values will be consistent. Thus if an estimated period of data appears inconsistent on the plots look at the stations being used as estimators. At least one of the estimators should also be inconsistent.

If no inconsistencies exist go on to Step 6.

Step 5.2

Determine the exact period which needs to be adjusted for each inconsistency that exists. Adjustments to the maximum or minimum temperature data should begin with the month when the slope first changes. The approximate month can be obtained from the plots. However usually it is difficult to determine exactly when the change occurs. This is not normally critical. Since most inconsistencies result from changes in location, the exact date can possibly be determined by checking if and when the station was moved. Changes in station location are listed at the end of the annual summaries of Climatological Data. It should be noted that changes in station location often will affect only the maximum temperatures or more likely only the minimum temperature data. Thus there is no reason to automatically expect that if a station has an inconsistency in the minimum temperature record the maximum temperature data will also be inconsistent.

Step 5.3

Determine the magnitude of each inconsistency in the maximum or minimum temperature data. Figure 2 shows two examples of how to determine the magnitude of the adjustment. In one case the entire record is adjusted to conform the most recent measurements. In the other case the most recent measurements are adjusted to agree with the earlier data. In general the data should be adjusted to conform to the most recent measurements. This is because the most recent measurements are more likely to be consistent with the current operational network. Only if the most recent measurements are in complete disagreement with the data from the other stations should the most recent measurements be adjusted to agree with the earlier data.

Step 5.4

Make the adjustments to the maximum-minimum temperature data and rerun program MAT to be sure that all the stations now have consistent data. Only those adjustments which are greater than about one degree Centigrade per month need to be included. Adjustments less than one degree Centigrade per month will undoubtedly have only a minor effect on the computed MAT data.

It should be noted that large adjustments, especially those that include a sizable portion of the period of record, can have a significant effect on the mean temperature at the station. Thus, in mountainous areas where the mean monthly temperatures are used in the computations, the monthly means should be corrected to account for the effect of the adjustment. For example, if an adjustment of -4 degrees Centigrade per month is applied to minimum temperatures for one-half of the period of record, then the mean monthly minimum temperatures for that station should be adjusted by -2 degrees Centigrade per month.

Step 6: Determine if Elevation is the Dominant Factor Affecting the Variation in Temperature Values

If this is a non-mountainous region, omit this step and go to step 9. If it is determined that elevation is the dominant factor affecting temperature variations, the process of determining station weights is simplified.

Step 6.1

Plot the mean monthly maximum and minimum temperatures for each station on the regional temperature-elevation relationship that was established during Step 1. Some of the stations currently in the basin analysis may have been included in the regional analysis. However it is likely that additional stations are now included, plus the period of record probably is not the same.

Step 6.2

Establish the temperature-elevation relationship for the basins

currently being analyzed for the period of record to be used. The relationship for maximum and minimum temperature data should be drawn for each month. The basin relationships may or may not be parallel to the regional relationships. However the regional relationships should be a guide, especially in extrapolating the basin relationships beyond the elevation range of the existing stations.

Step 6.3

Examine the anomalies from the basin temperature-elevation relationship. The anomaly for each station is its departure from the basin temperature-elevation relationship. Determine if there is a pattern to the anomalies. In the case of temperature data, it is usually not necessary to plot the anomaly values on a map unless a visual examination of the plots strongly suggests that such a pattern exists. Normally, the temperature anomalies are more or less randomly distributed, indicating that elevation is the dominant factor. In this case, the anomalies are mainly a result of the local environment of each station.

Even though elevation is the dominant factor affecting the variation in temperature, the mean minimum temperatures for some stations may be much lower than indicated by the basin temperature-elevation relationship. This is caused by a significant amount of cold area drainage at the particular location of these stations. The minimum temperatures for these stations can be raised by applying an adjustment to the data. The adjustment is made just like adjustments for inconsistencies, except that it is applied to the entire period of record. However if other stations are available at the same relative elevation such an adjustment is not necessary.

Step 7: Determine Station Weights for MAT Computations in Mountainous Regions

A method is described for obtaining station weights when elevation is the dominant factor affecting the variation in temperature. This is the case in the vast majority of basins. If it was determined that there is a pattern to the temperature anomalies, the method of determining the factors causing the pattern and how to use them to compute station weights is left up to the user.

Step 7.1

Plot the area-elevation curve for each area for which MAT values are to be computed.

Step 7.2

Plot the available stations on the area-elevation curve for each area. Plot only those stations which lie within the area or are reasonably close to the area. That is, plot all stations which could legitimately receive weight. Stations with cold area drainage problems, which were not previously adjusted, should not be plotted. These stations will still remain in the analysis for

use in estimating missing data, but should not be assigned any weight. The same is true for high elevation stations which are located at a considerable distance from the subarea. These stations are not assigned any weight, but may be used to estimate the maximum-minimum temperature values at a synthetic station.

Step 7.3

Determine the portion of the area which can reasonably be assigned to each of the stations. Figure 3 shows an example of how to use the area-elevation curve to determine the portion of the area to be assigned to each station. Make sure that the station elevation is reasonably close to the mean elevation of the portion of the area assigned to that station. If two or more stations are at the same relative elevation, assign a portion of that elevation band to each station. When there is a choice, stations located within the area should receive somewhat more weight than those which lie outside of the area. In some cases, 100 percent of the area can be assigned to the available maximum-minimum temperature stations, but in most mountainous regions the available stations will not represent the entire area.

Step 7.4

Create synthetic stations to represent those portions of the area that are not assigned to the available maximum-minimum temperature stations.

1. Determine the elevation of each synthetic station from the area-elevation curve. As in the case of real stations, the elevation of the synthetic station should be approximately the same as the mean elevation of that portion of the area assigned to the synthetic station. Figure 3 also shows an example of how to use the area-elevation curve to determine the elevation of a synthetic station.

2. Obtain the mean monthly maximum and minimum temperatures for each synthetic station from the basin temperature-elevation relationships. These relationships were established during Step 6.2. By knowing the elevation of the synthetic station, the mean maximum and minimum temperature at that elevation for each month can be obtained directly from the plotted basin temperature-elevation relationships.

3. Select a location and Fe values for each synthetic station such that a data record will be created by using the most logical estimators. In general, the most logical estimators are those real stations which are nearest in elevation to the synthetic station. This keeps the elevation range over which maximum-minimum temperature data must be extrapolated to a minimum. If an appropriate location cannot be found by using the default Fe value of 50, the estimators can possibly be

changed by varying the Fe value. Step 8 elaborates further on the selection of Fe values.

Step 8: Select Final Fe Values for Each Station in Mountainous Area

As has been mentioned previously, Fe=50 is a good default value. This value means that a user, in estimating missing data at a given station, would prefer to use a station which is up to 50 kilometers further in terms of distance rather than a station which has a 1000 meters greater elevation difference. (In English units, this is about the same as equating 10 miles of distance to 1000 feet of elevation difference.) This seems like a reasonable value for Fe and should be completely adequate for most stations. However for stations which have a significant amount of missing data and a large weight in the MAT computations, an Fe value of 50 may not insure that the most logical estimators will be used or that a nearly optimum estimate will be computed. For such stations, the user may want to select or compute an Fe value which should improve the estimation of missing maximum-minimum temperature data. Step 8.1 describes how to compute an Fe value for a given station so that the most logical estimators will be used when the data are missing. Step 8.1 is applicable to both synthetic and real stations. Step 8.2 describes how to determine an optimum Fe value. Optimum Fe values can be computed only for stations with some actual maximum-minimum temperature observations.

Step 8.1

Select Fe values so that the most logical estimators possible will be used to estimate missing values at critical stations which have a significant amount of missing maximum-minimum temperature data. This step primarily applies to synthetic stations. Synthetic stations usually are assigned a significant weight, plus the data record at a synthetic station consists of all missing values.

The procedure for estimating missing values at a given station uses the estimator in each quadrant with the largest weight (Chapter II.7-CALB-MAT). In mountainous areas, the weight function for a given estimator (W_i) is defined as:

$$W_i = \frac{1.0}{D_{i,X} + F_e * \Delta E_{i,x}} \qquad (1)$$

where
- X is the station for which data are being estimated
- I is the estimator station
- $D_{i,x}$ is the distance between stations I and X (KM)
- $\Delta E_{i,x}$ is the absolute difference in elevation, expressed in 1000 M, between stations I and X

Since $D_{i,x}$ and $\Delta E_{i,x}$ are fixed (through several locations may be tested for a synthetic station before a final location is selected), Fe is the only factor in Equation 1 that can be varied in order to change the weight functions of each estimator. The

Fe value for which the weight functions of two estimators are the same is:

$$Fe = \frac{D_{2,x} - D_{1,x}}{\Delta E_{1,x} - \Delta E_{2,x}} \qquad (2)$$

where the numbers 1 and 2 refer to the two estimators

If Fe is greater than the value computed by Equation 2, the estimator with the smaller $\Delta E_{i,x}$ value will have the larger weight. If Fe is smaller than the Equation 2 value, the estimator with the larger $\Delta E_{i,x}$ value will be used. If the Fe computed by Equation 2 is negative, it indicates that the estimator which is closer in terms of distance also is nearer to station X in elevation. In this case there is no choice as to which estimator is used.

Equation 2 can be used to help determine the Fe value that should be assigned to a given station so that those stations which the user has selected as being the most logical estimators will be used to compute missing temperature values. If may not be possible to find a single Fe value such that all of the most logical estimators are used. However an Fe value can be readily found which insures that the estimators which are used are reasonably similar to those the user thinks would give the best results.

Step 8.2

Compute the optimum Fe values are critical real stations which have a large amount of missing maximum-minimum temperature data. In most cases, no stations will meet these criteria, so this step can be omitted. If the closest estimators for stations which do meet these criteria also have the smallest elevation difference, the value of Fe will not have a significant effect on the estimation of the missing values. Therefore, in such cases this step can also be omitted.

Program MAT allows the user to input a value of Fe for estimating maximum temperature different from the value used for minimum temperature. Optimum values of Fe for a given station can be determined through a iterative technique, utilizing the available valid data from the station. Program MAT contains a feature designed to perform this iterative procedure. The use of this feature is referred to as a TEMPCK run in program MAT input summary. The user selects various Fe values and program MAT computes error statistics and generates a plot of observed versus estimated temperature for each Fe value. Equation 2 should be used in selecting the Fe values since the largest changes in the accuracy of the estimate occur when one estimator replaces another. The Fe values which produce the lowest RMS error are the values to use for this station.

Step 9: Prepare Basin Definitions Used in Computing Station Weights in Non-Mountainous Regions

Station weights in non-mountainous regions are computed by the grid-point weight method (Chapter II.7-CALB-MAT). The station weights are computed within program MAT using the basin definitions.

Step 10: Make Final Run of program MAT

All of the information needed in order to compute MAT time series for each of the areas included in the analysis should now be available. Thus, the only remaining step is submitting the final run of program MAT. The MAT input deck will require several changes from the deck used on previous runs of program MAT:

1. The number of maximum-minimum temperature stations needs to be increased if synthetic stations are included.

2. The option to output MAT time series data to a data file should be set.

3. Cards must be added with the elevation, coordinates, Fe values and mean monthly temperatures for each synthetic station.

4. Cards must be added for each area for which MAT is to be computed. These cards contain area name, identification code and either station weights or a grid map.

5. If consistency checks are made, remember to add any synthetic stations to one of the groups. Synthetic stations should be assigned negative sequence numbers for the consistency check part of the program.

The details of how to make these changes are described in Section III.8-MAT.

If all stations have missing data for a given 6-hour period, a message will be printed. The MAT time series will contain missing data for all such periods. Such cases should occur very infrequently. When such cases do occur, the MAT values for those periods must be estimated by the user.

If all of the previous steps were done carefully, the resulting MAT time series should contain reasonably unbiased estimates of the mean areal temperature for each of the areas.

<u>Summary of Steps to Follow in Computing Mean Areal Temperature</u>

Step 1: If a mountainous region, establish a regional temperature-elevation relationship. If a non-mountainous region, go to Step 2.

1.1 Select region over which the temperature-elevation relationships have the same general slope.

1.2 Select stations within the region with long records and

little missing data.

1.3 Obtain mean monthly maximum and minimum temperatures for each station using program DLYTRAN.

1.4 Run program TAPLOT.

1.5 Draw the regional temperature-elevation relationship.

Step 2: Locate available maximum-minimum temperature data for basin MAT computations (Chapter III.8-MAT for maximum number of basins, stations and years of data).

2.1 List all maximum-minimum temperature data stations in or near the basin.

2.2 Check inventories for length of record and amount of missing data.

2.3 Select stations for use in the MAT computations (each should have at least 4 to 5 years of data).

2.4 Tabulate observation times at the beginning of the record and any changes for each station.

Step 3: Transfer maximum-minimum temperature data into the disk files using program DLYTRAN.

Step 4: Check the consistency of the maximum-minimum temperature data.

4.1 Set Fe to default value of 50 (mountainous regions only).

4.2 Assign stations to groups for consistency computations.

4.3 Run MAT program to get consistency plots.

Step 5: Adjust the maximum-minimum temperature data for any inconsistencies.

5.1 Check to see if any inconsistencies are present. If none exist, go to Step 6.

5.2 Determine the periods that need to be adjusted.

5.3 Determine the magnitude of the adjustments.

5.4 Make adjustments and rerun MAT program to make sure data are now consistent.

Step 6: If a mountainous region, determine if elevation is the dominant factor affecting temperature variations over the basin area. If a non-mountainous region, go to Step 9.

6.1 Plot mean monthly maximum and minimum temperatures for each station on the regional temperature-elevation relationship.

6.2 Establish the temperature-elevation relationship for the basin being analyzed.

6.3 Examine anomalies to determine if elevation is the dominant factor affecting temperature variations.

Step 7: Determine station weights for mountainous regions.

7.1 Plot area-elevation curves for each area.

7.2 Plot available stations on the area-elevation curve.

7.3 Determine the portion of the area to be assigned to each station.

7.4 Create synthetic stations to represent the portions of the area not assigned to the real stations.

Step 8: Select final Fe values for each station. (This step is critical only for stations with a lot of missing data. Default value is adequate for most stations.)

8.1 Select Fe values for synthetic stations so that the data record will be computed using the most logical estimators.

8.2 Compute optimum Fe values for real stations using the TEMPCK feature of program MAT.

Step 9: Prepare basin definitions for use in computing station weights in non-mountainous regions.

Step 10: Make final MAT run to write MAT time series to a data file.

Figure 1. Sample regional temperature/elevation relationship

Letters refer to maximum-minimum temperature stations.

Relationship is for the Upper Colorado and Gunnison River basins for the month of June, based on water years 1952 to 1971.

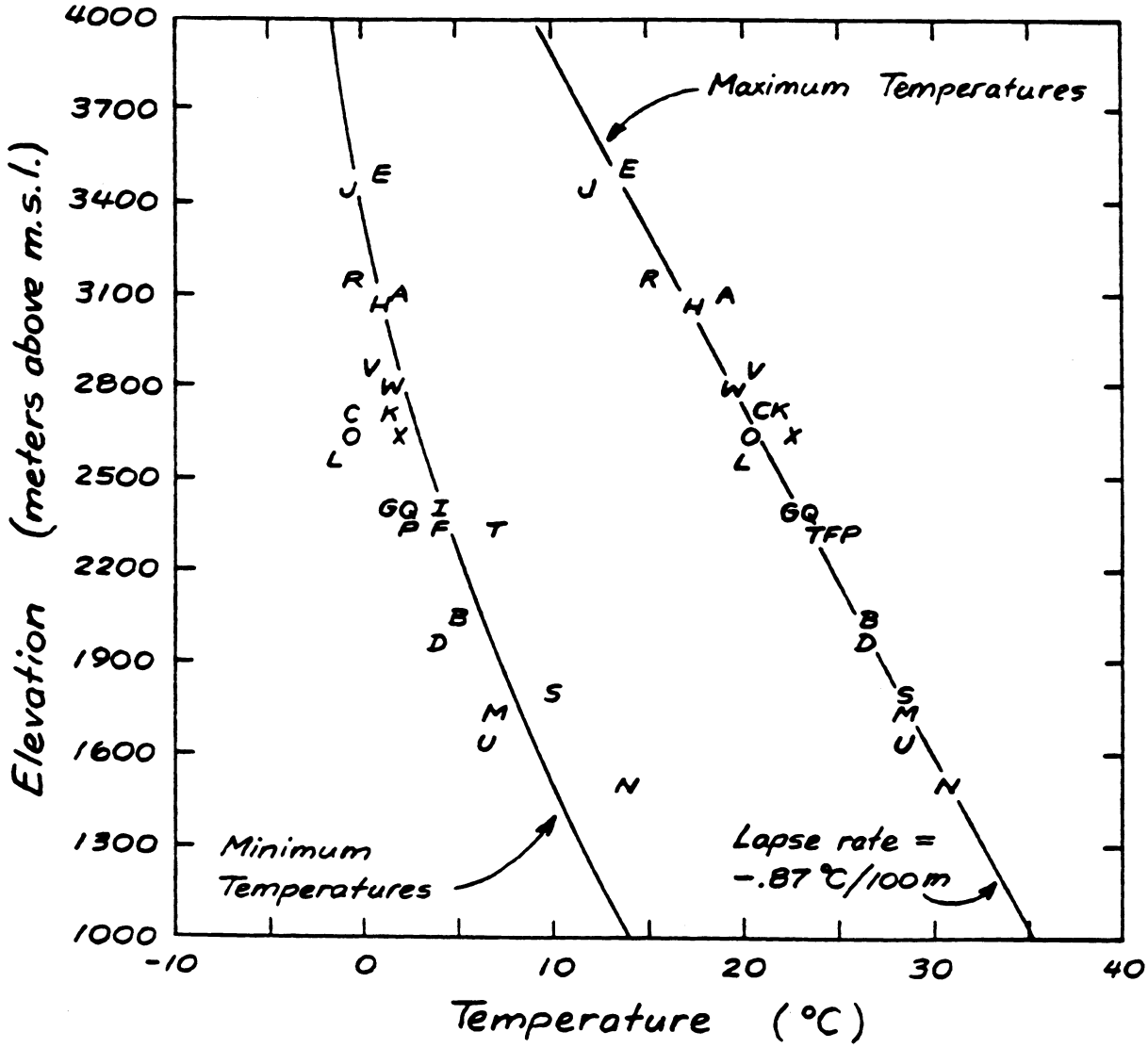


Figure 2. Examples of how to calculate MAT consistency adjustments

Solid lines represent original consistency plots.
Dashed lines show what plots will look like after adjustment.
Stations 1 and 3 are adjusted to conform to the most recent data.
Station 2 is adjusted to agree with earlier data.
Starting date and adjustments for each are:
Station 1: 10/66, -1.0 DEGC ; 04/69, 0.0 DEGC
Station 2: 10/70, 3.0 DEGC
Station 3: 10/66, 2.0 DEGC ; 10/67, -1.5 DEGC ; 10/69, 0.0 DEGC

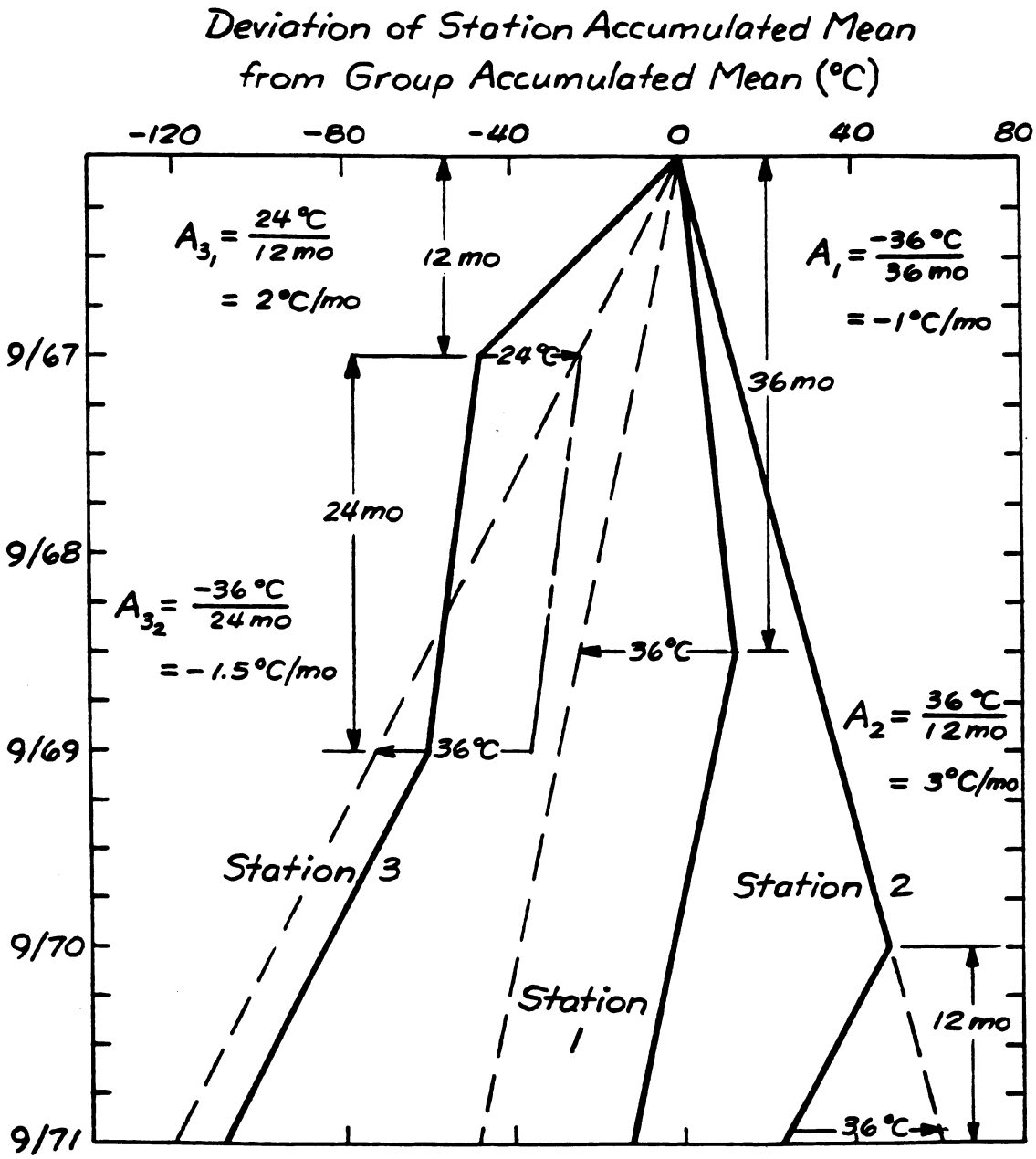
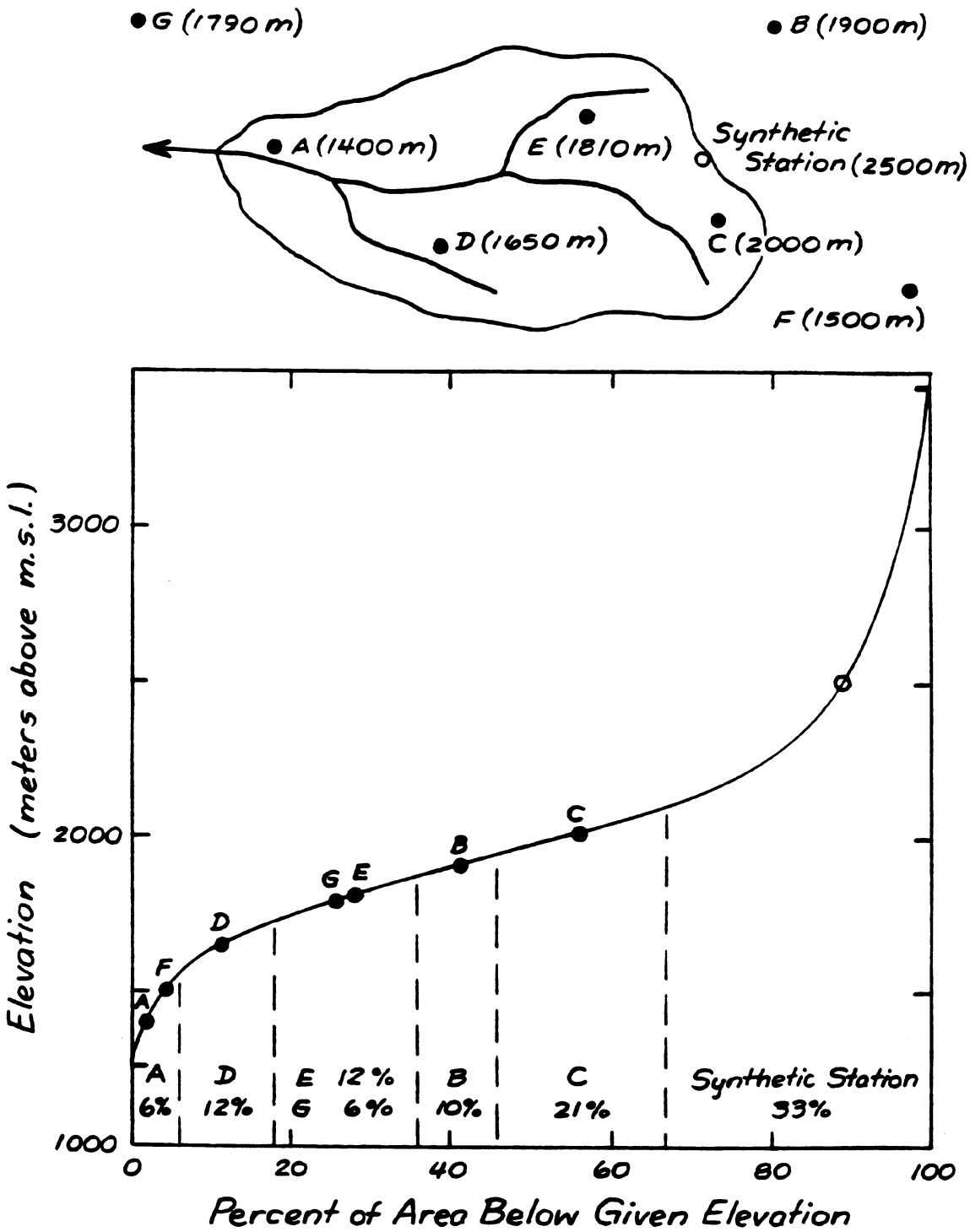


Figure 3. Example of using the area-elevation curve to determine station weights and the elevation of a synthetic station

Letters refer to maximum-minimum temperature stations.