

Procedures for Substituting Values for Missing NWS Meteorological Data for Use in Regulatory Air Quality Models

Dennis Atkinson and Russell F. Lee

7/07/92

Introduction

The Meteorological data required for short-term air quality dispersion modeling consist of hourly surface observations, and morning and afternoon mixing heights. The RAMMET meteorological pre-processor program requires surface data in the CD144 format, and mixing height data in the TD9689 format. In order for the RAMMET program to function properly, these data must be 100% complete. Data acquired from the National Climatic Data Center (NCDC) occasionally have periods of missing data. If the lengths of these periods are not excessive, reasonable values may be substituted without seriously degrading the modeling results.

As with on-site data, a data set which is less than 90% complete should not be used for air quality modeling purposes. Substitutions for missing data should only be made to complete the data set for modeling applications, and should not be used to achieve the 90% criterion. Note that, for purposes of achieving the 90% criterion, cloud cover is valid for a given hour if a value for EITHER opaque cloud cover or total cloud cover is available.

The procedure for providing substitute values for missing data has two parts. The first part is an objective procedure, which applies to single isolated hours with missing surface data and single isolated days of missing mixing height data. Substitutions for those data are accomplished using procedures described below.

The second part is a subjective procedure which applies to longer sequences of missing data. Substitutions for those data require judgement, and should be accomplished by an air quality meteorologist based on scientific knowledge and professional experience. The procedures, described in detail below, are generally consistent with procedures used historically by EPA.

Step 1: The Objective Procedure

Surface Data

Hourly surface weather observations of opaque cloud cover, ceiling height, temperature, wind direction, and wind speed are required. Substitutions for missing surface data are made as follows:

- If opaque cloud cover is missing, then total cloud cover is used. If total cloud cover is also missing, then a "0" is used for opaque cloud cover if the ceiling is greater than 7000 feet (high level clouds are generally not opaque), and a "7" if the ceiling is less than or equal to 7000 feet. If the opaque cloud cover, total cloud cover, and ceiling are all missing, values of ceiling and opaque (or total) cloud cover from the preceding hour are used, provided that

values from both the preceding and the succeeding hours are present (i.e., only a single hour is missing). If two or more consecutive hours of cloud cover (both opaque and total) are missing, then the subjective procedure outlined below should be used.

- If ceiling height is missing, and either the opaque cloud cover or the total cloud cover is less than or equal to 5, then "---" (the code for unlimited ceiling) is used. Otherwise, if the ceiling height is missing, but values for both the preceding and succeeding hours are present, then the value from the preceding hour is used. If two or more consecutive hours of ceiling height are missing, then the subjective procedure outlined below is used.
- If temperature is missing, then a value interpolated between the preceding hour and the succeeding valid hour is substituted. If two or more consecutive hours of temperatures are missing, then the subjective procedure outlined below is followed.
- If a single hour of wind direction and speed is missing then the average direction and average speed from the four hours surrounding that hour is substituted. Except for the first two and last two hours of the year, this average is derived from the two hours before and two hours after the hour of the missing value. In the interest of simplicity, the average wind direction is obtained from the mean unit vector wind. If two or more consecutive hours of either wind direction or wind speed are missing, then the subjective procedure outlined below should be used.

Mixing Height Data

Both a morning mixing height and an afternoon mixing height are required for each day of the year. Objective substitutions for missing mixing heights are made as follows:

- If a single afternoon mixing height is missing, the value interpolated between the preceding and successive afternoon mixing heights are used. Similarly, if a single morning mixing height is missing, the value interpolated between the preceding and succeeding morning mixing heights is used.
- If two or more consecutive afternoon mixing heights or two or more consecutive morning mixing heights are missing, then the subjective procedure outlined below is used.

Step 2: The Subjective Procedure

When the objective procedure does not provide a substitute value for some parameter, the data are reviewed by an air quality meteorologist. If technically appropriate, a substitute value is identified, based on the following procedure, using sound scientific knowledge and professional experience.

Surface Data

The following procedures are recommended for substituting values for missing surface data when the objective procedure described above is not appropriate. Sound meteorological judgment and professional experience is required to developing appropriate substitute data values.

If a ceiling height value is missing, then the five ceiling height values before and after that hour are reviewed to obtain appropriate values to substitute. For an opaque cloud cover of 5 or less, the value for ceiling height must be "---" to indicate unlimited ceiling (by definition, the ceiling is unlimited when the cloud cover is 5/10 or less). If the opaque value is 6 or greater, the values before and after that hour are reviewed to derive appropriate ceiling values.

If temperature is missing, then the temperature values on either side of the data void are reviewed. Often an interpolation can be made for up to a few hours. However, these values may need to be adjusted to be consistent with the other meteorological factors. Near times of daily maxima or minima, one can adequately simulate daily maxima or minima for other days of similar meteorological conditions.

If a wind direction value is missing, the data values a few hours (perhaps five or so) on either side of the data void are reviewed. It is important to maintain not only consistency of wind direction, but also some consistency of wind variability. Otherwise, a few hours with the wind from the same direction will result, which can produce unreasonably high estimates of 8-hour and 24-hour average concentrations.

If wind speed is missing, the five values both before and after the data void are reviewed. Any obvious patterns of wind speed should be continued and adjusted, if needed, to make the substitute data realistic from an air quality modeling viewpoint.

Mixing Height Data

The following procedure is recommended for substituting values for missing morning or afternoon mixing height values, when the objective procedure described above is not appropriate. The procedure with mixing height data, as with surface data, requires careful meteorological judgment and professional experience in order to derive appropriate substitute values.

As a general rule, a sequence of missing mixing heights of five days or more are not filled in. Also, any missing sequence of four days which does not have adequate continuity (at least five full days on either side of the data void) is also not filled in. Such files cannot be used as input to the RAMMET (or PCRAMMET) preprocessor, since doing so produces values of zero for the mixing height, resulting in unrealistic modeling results.