9.3 What are the Part 75 missing data procedures for CEMS?

In general, the Part 75 missing data procedures for CEMS are designed to provide conservatively high substitute data values, to ensure that emissions are not underestimated during monitor outages. Application of the missing data procedures begins either: (a) at the date and hour of provisional certification, when the CEM systems have passed all required certification tests; or (b) when the certification test deadline expires, if the monitoring systems have not yet passed all of the required tests.

Two distinct sets of CEMS missing data algorithms are described in Part 75---the "initial" and the "standard" missing data routines. The initial missing data algorithms in §75.31 are temporary "spin-up" procedures that are used for a specified period of time, after which the standard missing data algorithms in §§75.33 through 75.37 begin to be applied. For both the initial and standard missing data procedures, all of the appropriate substitute data values are calculated and applied automatically by the DAHS. If a missing data period extends past the end of a quarter, it is treated as two separate missing data periods, one terminating at the end of the current quarter and one starting at the beginning of the next quarter.

The initial missing data procedures in §75.31 are used until a certain number of hours of quality-assured CEM data have been obtained. For SO_2, CO_2, O_2, and moisture, this number is 720 hours, and for NO_x and flow rate, it is 2,160 hours. The initial missing data algorithms are simple and the substitute data values derived from them are likely to be close to the actual values. For example, the algorithm for SO_2 is the arithmetic average of the SO_2 concentrations from the hour before and the hour after the missing data period. For NO_x and flow rate, the substitute data value for each hour is an arithmetic average of the available historical data at similar load levels.

Once the requisite number of hours of quality-assured data has been obtained (i.e., 720 or 2,160), use of the initial missing data procedures ceases and the standard missing data procedures begin to be applied.62 The standard missing data routines use a tiered approach, that takes into account both the percent monitor data availability63 (PMA) and the length of the missing data period.

When the PMA is high (≥ 95%) and the missing data period is relatively short (≤ 24 hr), the standard missing data algorithms are nearly identical to the initial missing data routines---consequently, the substitute data values are generally not punitive. However, as the PMA decreases, the substitute data values become increasingly conservative, to ensure that emissions are not under-reported.

For example, when the PMA of an SO_2 or NO_x monitoring system is between 80% and 90%, the substitute data value will be the maximum value observed by looking back through the last 720 hours (for SO_2) or 2,160 hours (for NO_x) of historical, quality-assured emissions data64 (except as otherwise noted below, for units with add-on SO_2 or NO_x emission controls). But if the PMA drops below 80%, regardless of the length of the missing data period, the maximum potential SO_2 concentration or the maximum potential NO_x emission rate must be reported (except as otherwise noted below, for units with add-on SO_2 or NO_x emission controls).

The initial and standard missing data algorithms for NO_x and stack gas flow rate are load-based, in order to provide more representative substitute data values. Appendix C of Part 75 requires the owner or operator to establish 10 load ranges or "load bins", by dividing the entire load range of the source (e.g., 0 to 500 megawatts) into 10 equal parts65. Then, during periods of missing NO_x or flow rate data, the substitute data value for each hour is calculated using historical quality-assured data in the corresponding load bin.

However, certain non-EGUs (e.g., cement kilns and refinery process heaters) that were in the NO_x Budget Program, and that may be brought into CAIR, do not produce electrical or steam load. To accommodate these non load-based sources, EPA added a series of special missing data algorithms for NO_x and flow rate to Part 75 in 2002. The algorithms are structurally similar to the standard NO_x and flow rate missing data routines, except that they are not load-based. To alleviate industry concerns that the substitute data values determined in this manner may not be

representative, the rule allows the affected sources to define "operational bins" corresponding to different process operating conditions, and to populate each bin with CEM data. The substitute data value for each missing data hour is then drawn from the appropriate operational bin.

For units with add-on SO_2 or NO_x emission controls, the use of the initial and standard missing data routines is conditional. The condition is that parametric data must be available to document that the add-on controls are working properly during the missing data period. For any hour in which this parametric evidence is unavailable, the maximum potential SO_2 concentration or the maximum potential NO_x emission rate must be reported.

In June 2002, EPA revised the standard missing data routines in §§75.33 and 75.34 to allow certain sources to report more representative substitute data values. Specifically:

• Affected sources that burn different types of fuel were given the option to separate their historical CEM data according to fuel type and to apply the standard missing data procedures on a fuel-specific basis; and

• For a unit that: (a) is subject to the CAIR NO_x ozone season program; and (b) is equipped with add-on NO_x controls; and (c) reports emissions data year-round, the owner or operator may separate the NO_x emission data into ozone season and non-ozone season data "pools". Then, depending on the time of the year that the missing data period occurs (i.e., inside or outside the ozone season), the substitute data values are drawn from the appropriate data pool. This missing data option is advantageous when the NO_x emission controls are operated only during the ozone season, or if the controls are operated less efficiently in the off-season.

More recently, in January 2008, EPA revised §75.34 to provide units with add-on SO_2 and NO_x emission controls a measure of relief from reporting both the maximum value in a look back period (when the PMA is between 80 and 90%) and the maximum potential value (when the PMA is below 80%)---provided that proper operation of the emission controls can be documented.66 In the first case, where the PMA is between 80 and 90%, you may report the maximum controlled SO_2 concentration67 or the maximum controlled NO_x emission rate in the look back period instead of the maximum value. In the second case, instead of reporting the maximum potential value when the PMA is below 80%, you may report the following substitute data values:

• For SO_2 concentration, the greater of: (a) the maximum expected concentration (MEC); or (b) 1.25 times the maximum controlled concentration in the look back period ; or

• For NO_x emission rate, the greater of: (a) the maximum controlled emission rate (MCR)68; or (b) 1.25 times the maximum hourly controlled NO_x emission rate in the look back period.