

Fact Sheet for EPA-Provided, State-Level Hourly Sulfur Dioxide (SO₂) Data

Background

EPA's final 1-hour SO₂ National Ambient Air Quality Standard (NAAQS) provides the opportunity for States to demonstrate attainment through ambient air quality monitoring and/or emissions dispersion modeling. To assist states that choose to use emissions dispersion modelling, EPA is providing emissions data that is collected under the Part 75 continuous emission monitoring regulation. These data include stack characteristics and hourly emissions data provided by coal-fired sources that participate in the Acid Rain Program, Cross-State Air Pollution Rule, or former NO_x Budget Trading Program. The data are primarily from electric generating units (EGUs) with some industrial boilers included. All of the sources collect emissions data using continuous emission monitoring systems (CEMS) and electronically report this information to the EPA on a quarterly basis. The data being provided cover calendar years 2012 through 2014.

Data Dictionary for EPA-Provided Data

The EPA-provided, State-level data is comprised of two Comma Separated Value (CSV) files. "Location Information" contains the characteristics of the facility (including the facility name, location, SO₂ emission controls (if any), and characteristics of the stack (shape, height, construction, etc.). "Emissions Information" contains the hourly emissions data for each location described in the Location Information dataset.

The following is a description of the fields included in each dataset.

Location Information

- **EPA Region** – the EPA Region in which the facility is located
- **State** – the two letter abbreviation for the state in which the facility is located
- **Facility Name** – the name of the facility, as reported to EPA
- **ORIS Code** – the numeric identifier assigned to this facility for EPA reporting purposes
- **Latitude** – Latitude to a location inside the facility. Note: this is not the latitude of an emissions location, but the facility itself.
- **Longitude** – Longitude to a location inside the facility. Note: this is not the longitude of an emissions location, but the facility itself.
- **Location ID** – the identifier for the location of the emissions data. For single stack configurations this will be the unit ID, as registered with the CAMD Business System. For monitored common or multi-stack configurations, this will be the stack identifier that the facility

has assigned to the stack, beginning with the prefix “CS” (for common stacks) or “MS” (for multiple stacks or ducts)¹.

- **SO₂ Controls** – a code that represents the technology used to control SO₂ emissions from the unit. See Table 3, “Control Codes and Descriptions” in the *ECMPS Reporting Instructions - Monitoring Plan*²
- **Stack Height** – the height of the stack exit, in feet above ground level
- **Stack Shape** – the shape of the stack at the flow monitoring location, either round (ROUND) or rectangular (RECT)
- **Material Code** – a code that describes the material from which the inner wall of the duct or stack is constructed at the flow monitoring location, either brick (BRICK) or some other material (OTHER)
- **Cross Area Exit** – the inside cross-sectional area, in square feet, of the stack at the flue exit
- **Cross Area Flow** - the inside cross-sectional area, in square feet, of the stack at the flow monitor location
- **Ground Elevation** - the elevation of ground level, in feet above sea level, at the base of a stack or unit

Emissions Information

- **State** – the two letter abbreviation for the state in which the facility is located
- **ORIS Code** - the numeric identifier assigned to the facility for EPA reporting purposes
- **Location ID** – the identifier for the location of the emissions data. For units in single stack configurations this will be the unit ID, as registered with the CAMD Business System. For common or multi-stack configurations this will be the stack identifier that the source has assigned to the stack, beginning with the prefix “CS” (for common stacks) or “MS” (for multiple stacks or ducts).

Note: **Location IDs** are only unique to a facility. A global unique identifier for an emissions source would be comprised of both the **ORIS Code** and the **Location ID**.

¹ Note that when the “MS” prefix is used, the emissions monitoring may be done at the stack or duct level. There are three possibilities: (1) the unit has multiple exhaust stacks that are monitored individually, and stack gas exits through all of the stacks continuously; (2) the unit has only one exhaust stack, and monitoring is done in the ductwork, generally in split breechings; or (3) the unit is equipped with SO₂ controls and has a main exhaust stack and a bypass stack, both of which are monitored. Therefore, when the data from a source are reported using “MS” prefixes, it is important for modelers to find out which monitoring configuration is being used.

² Page 20; *ECMPS Reporting Instructions – Monitoring Plan* is available at <http://www.epa.gov/airmarkets/participants/monitoring/reporting-instructions.html>

- **Date** – the date on which a particular pollutant concentration or flow rate was recorded (mm/dd/yyyy)
- **Hour** – the clock hour of the day in which a particular pollutant concentration or flow rate was recorded (range of values 00-23)
- **OP Time** – the fraction of the clock hour during which the unit combusted any fuel (or, for a monitored common stack the fraction of the clock hour during which flue gases flowed through the stack).
- **Heat Input** – the calculated heat input rate for the hour in million Btu/hr
- **SO₂ MODC** – a “method of determination code” that indicates whether a reported hourly average SO₂ concentration is quality-assured or is a substitute data value. See Table 19 in the *ECMPS Reporting Instructions – Emissions*³ for information about the codes.
- **SO₂ Unadjusted** – the unadjusted, quality assured SO₂ concentration for the hour, as measured by the SO₂ monitor, in parts per million (ppm).
- **SO₂ Adjusted** – the bias-adjusted SO₂ concentration calculated by multiplying the SO₂ unadjusted value by the bias adjustment factor (BAF) determined during the relative accuracy test audit (RATA).
- **SO₂ Mass** – the calculated SO₂ mass emission rate in lbs/hr.
- **Flow MODC** – a “method of determination code” that indicates whether a reported hourly average flow rate reading is quality-assured or is a substitute data value. See Table 19 in the *ECMPS Reporting Instructions – Emissions* for information about the codes.
- **Flow Unadjusted** – the unadjusted, quality-assured stack gas volumetric flow rate for the hour, as measured by the flow monitor, in standard cubic feet per hour (scfh)
- **Flow Adjusted** - the bias-adjusted flow rate, calculated by multiplying the unadjusted value by the bias adjustment factor (BAF) determined during the RATA.

³ Pages 55-56; *ECMPS Reporting Instructions – Emissions* is available at <http://www.epa.gov/airmarkets/participants/monitoring/reporting-instructions.html>

Frequently Asked Questions

- ***Where do the data come from?***

The facility information and stack-level hourly emission and operation data are collected by EPA under 40 CFR Part 75 for assessing compliance with the Acid Rain Program, Cross-state Air Pollution Rule, and other state and federal programs.

- ***How are the emissions measured/calculated?***

The emissions are measured using continuous emissions monitoring systems (CEMS). A CEMS consists of all the equipment needed to measure and provide a permanent record of the emissions from an affected unit. Examples of CEMS components include:

- Sample probes
- Sample (“umbilical”) lines
- Sample pumps
- Sample conditioning equipment (e.g., heaters, condensers, gas dilution equipment)
- Pollutant concentration monitors (e.g., SO₂ monitors).
- Diluent gas monitors, to measure %O₂ or %CO₂
- Stack gas volumetric flow rate monitors
- Data loggers or programmable logic controllers (PLCs)
- Data acquisition and handling system (DAHS) components that electronically record all measurements and automatically calculate and record emissions and heat input in the required units of measure.

- ***What is a Bias Adjustment Factor (BAF)?***

Sources that report under Part 75 are required to periodically perform RATAs to test their CEMS against a reference method. If the CEMS readings are found to be statistically biased low compared to the reference method, a bias adjustment factor (BAF) is calculated and applied to any subsequent hourly emissions data. BAFs are used for SO₂ concentration and stack gas flow rate, but not for CO₂ concentration.

- ***What is the difference between unadjusted and adjusted flow rate and SO₂ concentration values?***

The unadjusted values are the raw measured values. The adjusted values are determined by applying the BAF to the unadjusted value to correct for statistically-significant low bias in the measured values.

- ***What is a method of determination code (MODC)?***

The MODC indicates whether a reported gas concentration or stack flow rate is a quality-assured value measured by a monitor or is a substitute data value. Table 19 in the *ECMPS Reporting Instructions – Emissions* (available at <http://www.epa.gov/airmarkets/participants/monitoring/reporting-instructions.html>) lists the available codes.

- ***What are substitute data?***

Part 75 requires that facilities report emissions and operations data for every hour that an affected unit is operating (i.e., combusting fuel). If for some reason (e.g. CEMS malfunction) quality-assured data are not available for a particular operating hour, the Part 75 missing data routines must be used to provide substitute data for that hour. The longer the period of missing quality-assured data or the more frequently quality-assured data are not available, as measured by a CEMS' percent monitor availability (PMA), the more conservative (i.e., higher) the substitute data.

Substitute data in the dataset can be identified by the method of determination code (MODC). Substitute data will be “flagged” with a MODC of 06, 07, 08, 09, 10, 11, 12, 13, 15, 21, 23, or 55.

- ***Why is stack temperature not included in the EPA-provided data?***

Sources collect hourly stack temperature data but are not required to submit this data under Part 75. If hourly stack temperature data is required, states might choose to contact the sources directly to request the data.