Percent monitor data availability in EPA CAMD's Power Sector Emissions Data

July 21st 2021

Coal-fired power plants that are subject to the Acid Rain Program and part 75 emission monitoring, reporting, and recordkeeping provisions[1] must operate continuous emission monitoring systems (CEMS) to measure CO_2 and SO_2 concentrations, NO_X rate (a combination of NO_X and a diluent (CO_2 or O_2) concentrations), and stack gas flow. Many natural gas- and oil-fired power plants use CEMS to measure NO_X rate. Over the past 25 years, these CEMS have proven to be reliable, providing accurate and timely data. To ensure the systems are providing valid data, power plants must conduct regular quality assurance (QA) tests on the monitors.

What is percent monitor data availability?

Percent monitor data availability (PMA) is a measure of the reliability of a CEMS, expressed as the percent of operating hours that a monitor is providing valid and quality assured emissions data.

When a CEMS is not providing valid or quality assured data, the PMA is used to determine which data substitution algorithm to use.[2] The lower the PMA, the more conservative the data substitution (i.e., the more likely to overestimate emissions). Because the emissions data, including substitute data determined by PMA, is used to assess compliance with emission trading programs, emissions sources have an incentive to keep a high PMA to avoid conservative data substitution.

The PMA calculation uses a lookback period of 8,760 operating hours for year-round emission reporters (3,672 for ozone-season-only reporters) going back up to three calendar years (26,280 clock hours). Because the PMA is a rolling average, a new PMA is calculated and reported every unit operating hour.[3]

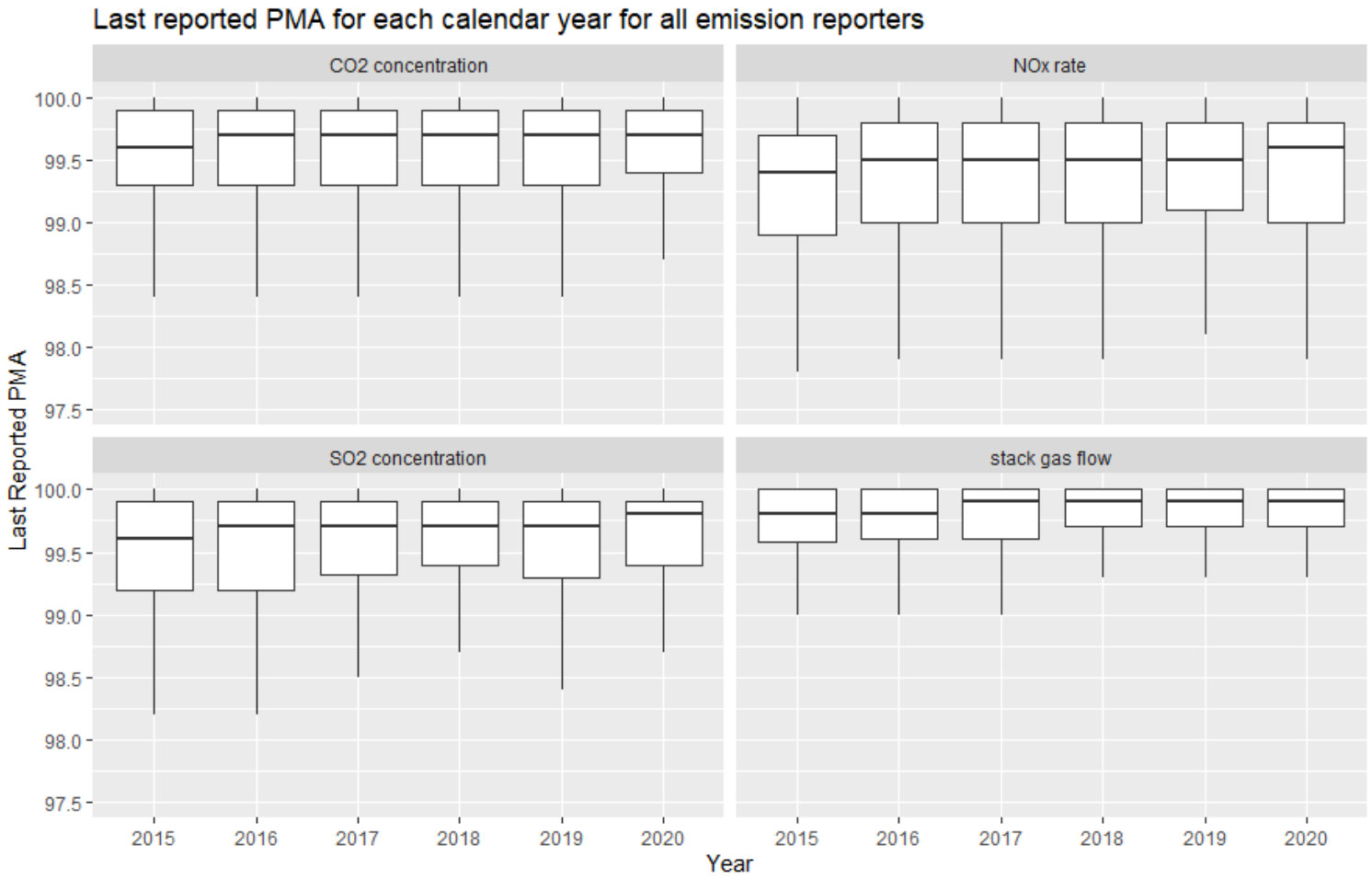
$$\text{PMA} = \frac{\text{Total lookback QA operating hours}}{\text{Total lookback operating hours}} \times 100$$

[1] To learn more about 40 CFR part 75, refer to the Plain English Guide to Part 75.

[2] To learn more about substitute data, refer to the Monitoring Insights substitute data analysis.

[3] Refer to 40 CFR 72.2 for definitions, including quality-assured monitor operating hour, and 40 CFR 75.32 for information about calculating PMA after initial certification (i.e., before the system operates 8,760 hours).

PMA values for each parameter are concentrated above 99%



For information about how to read these figures, refer to page 6

PMA in the Power Sector Emissions Data is increasing or stable for CO_2 and SO_2 concentration

Over the past five years, the PMA, as measured by the last reported PMA for the calendar year, has improved for several of the measured parameters. The median PMA values for CO_2 and SO_2 concentration, NO_X rate, and stack gas flow have improved or remained stable. In 2020, the median PMA values for the four parameters was above 99.5 percent. Over the same five years, approximately 75 percent or more of the PMA values for CO_2 and SO_2 concentrations, and NO_X rate were above 99 percent. For stack gas flow, 100 percent of the last reported PMA values were above 99 percent.

In addition to higher median values, the range of PMA values is narrowing and shifting higher for three of the parameters—CO_2 and SO_2 concentration, and stack gas flow, The range of PMA values for NO_X rate has remained relatively stable during the period. The NO_X rate system is dependent on two monitors (NO_X concentration and a diluent—CO_2 or O_2 concentration). This could contribute to the slightly lower median and wider range of values for NO_X rate relative to the other parameters.

These values indicate that power plants have continued to operate CEMS effectively and perform the required quality assurance activities.

For more information about the data or this analysis…

EPA's part 75 monitoring and reporting program

- 40 CFR part 75—Continuous Emission Monitoring
- Plain English Guide to Part 75 (PDF)
- EPA CAMD power sector programs—progress reports

Power Sector Emissions Data

- CAMD's Power Sector Emission Data
- CAMD's Power Sector Emissions Data Guide (PDF)

Contact information

Stacey Zintgraff
EPA's Clean Air Markets Division
202-564-2204
zintgraff.stacey@epa.gov

Analytical methodology

This analysis was completed in R. If you would like to review the code or source data, contact Stacey Zintgraff. To complete this analysis, we …

1. Compiled the last reported PMA for each calendar year for each monitoring system at affected power plants. The last reported PMA was chosen because PMA is a rolling average, it would be redundant to use annual averages.
2. Grouped the PMA values by parameter. Note that NO_X concentration and O_2 concentration were excluded because they are very small data sets.
3. Calculated the quartile PMA values (i.e., min, 25th percentile, median, 75th percentile, and max).
4. Created box and whisker plots for each parameter using R. Note that outliers were omitted from the boxplot, but all parameter had a small number of outliers in the data set.

By the numbers

Median PMA for last operating hour in 2020

- CO_2 concentration: 99.7%
- NO_X rate: 99.6%
- NO_X concentration: 99.1%
- O_2 concentration: 99.6%
- SO_2 concentration: 99.8%
- Stack gas flow: 99.9%

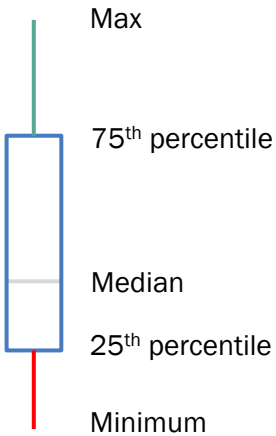
Number of CEMS for each parameter over the six years used in this analysis

- CO_2 concentration: 657 (4,088 values)
- SO_2 concentration: 632 (3,927 values)
- Stack gas flow: 664 (4,039 values)
- NO_X rate: 2945 (17,020 values)
- NO_X concentration: 70 (361 values)
- O_2 concentration: 26 (132 values)

How to read a box plot

The annual PMA for all reporters are displayed using a “box plot” also known as a “box-and-whisker plot”.

Box plot key



A box plot is a method to depict groups of numerical data in quartiles. It illustrates the distribution, central tendency, and variability.

In the example box plot on the left:

- The green vertical line, or top whisker, represents the values between the 75th percentile and the maximum value.
- The blue box represents the middle half of all values—also known as the interquartile range (IQR)—those that fall between the 25th and 75th percentile.
- The grey horizontal line represents the median value (i.e., the 50th percentile value).
- The red vertical line, or bottom whisker, represents the values between the minimum and the 25th percentile.

Each box plot provides visual representations of both the magnitude and variability of PMA values for all reporters in a given year in a single chart.