

Ambient Gas Monitoring Data Quality Dashboard - README

March 2020

Summary

This set of applications provides visual Quality Assurance (QA) Data Quality Dashboards that display ambient air measurements (for one of ozone, NO, NO₂, NO_y, SO₂, or CO) for each Primary Quality Assurance Organization (PQAO); selection by Monitoring Agency is also available. Assessments may be generated for each year when corresponding data are available in the AQS database; data from years 2016 through the present year are available for review. The visual assessments are intended to assist agencies in understanding whether data from ambient air gas monitors in their networks are meeting goals for measurement quality objectives (MQOs). The assessments are expected to have three major advantages over existing assessments:

- Data are visualized to illustrate large amounts of information in a concise way.
- The PQAO of interest is compared to all other PQAOs, allowing agencies to understand whether their program's data quality is similar to other agencies or perhaps an outlier.
- The dashboard applications are expected to remain openly available so that users do not have to worry about user authentication IDs and passwords. In this way, we hope that users can efficiently use the data quality dashboard for their program's gas monitors.

The QA Data Quality Dashboards cover multiple indicators of data quality for each of the individual gases:

- One-point QC precision
- One-point QC check bias
- Routine percent completeness
- Counts of monitored days in a year
- Counts of one-point QC checks
- Tile plot of QC check and routine data results
- Time series of
 - One-point QC check percent difference
 - One-point QC check concentrations
 - Daily comparisons to the level of the National Ambient Air Quality Standards (NAAQS): 1-hr for NO₂, NO_y, and SO₂, or 8-hr for ozone and CO

Assessment Design

You can view indicators of data quality in three charts on the **PQAO** tab and on one table and three charts on the **Monitor** tab. In each dashboard, you can select a year, a PQAO, and a monitoring

method from three dropdown menus. The **Year** dropdown menu defaults to the most recent completed year. The **PQAO** dropdown defaults to the first alphabetical PQAO name. The **Method** dropdown defaults to **All Methods**, which shows the average value across all methods for each PQAO. Selecting a single method shows PQAO-specific method averages for comparison in blue squares (see [Figure 1](#)).

When you first load the app, the default landing screen is the **PQAO** tab, which shows the data quality indicators across all PQAOs that measured the gas of interest in the most recent complete year. Select the **Monitor** tab at the top left to focus on the results at a single monitoring site within the currently selected PQAO. Each tab is discussed below, using ozone as an example. Other pollutants have slight variations that are pollutant-specific; these are discussed at the end of this README document.

Finally, a **Generate PDF report** button is available above the dropdown menus. Click that button to generate a report featuring data displayed on the currently selected tab (either **PQAO** or **Monitor**).

PQAO Tab

Within the **PQAO** tab, each chart shows the ranking of all PQAOs for the particular indicator of data quality. The topmost chart shows Precision by PQAO, the middle chart shows Bias by PQAO, and the bottom chart shows the percent completeness by PQAO. Figure 1 shows the landing page for the ozone dashboard and the first two charts; scroll down to see the third chart.

Common Chart Features

Each of the three charts provided in the **PQAO** tab has common features:

- The selected agency is highlighted as an orange dot overlaid on the black dot for its PQAO, with its calculated value as an orange value above the dotted number.
- PQAOs metrics are rank-ordered, with the best performing agencies on the left and the poorest performing agencies on the right.
- The MQO for each indicator of data quality is shown as a red solid line, with the MQO goal annotated above it.
- The national average indicator value and the national method average indicator value are shown as blue annotated text on the left side of the charts and as blue dashed lines within the charts. National method average indicator values are reported as NA if a method is not selected.



Ozone Data Quality Dashboard

[Click Here for QA Assessment ReadMe File](#)

[Generate PDF report](#)

Choose a Single Year

2017

Choose a PQAQ

Al Dept Of Env Mgt

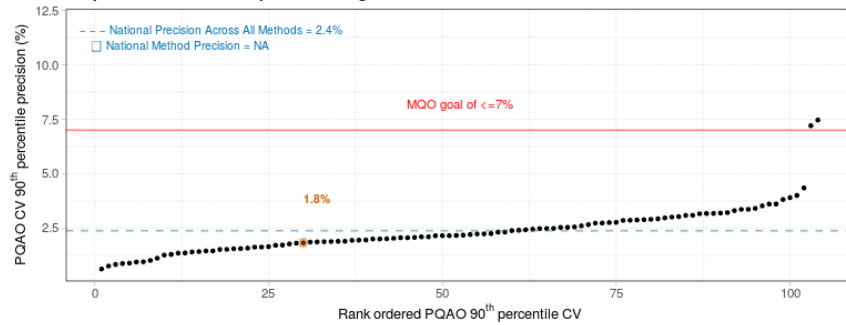
Select a Method

All Methods

PQAQ

Monitor

CV precision: PQAQ comparison - drag and double-click to zoom



Bias estimate: PQAQ comparison - drag and double-click to zoom

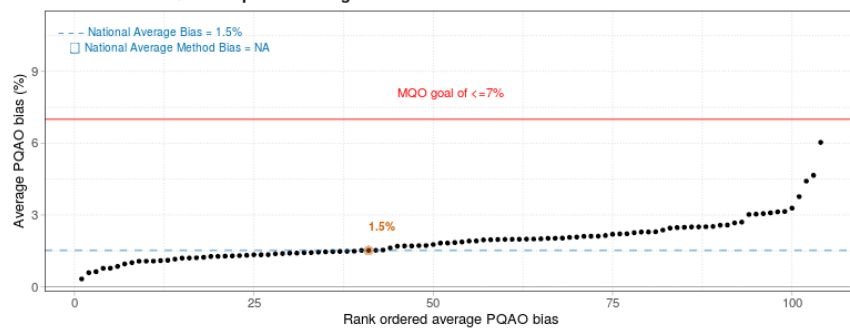


Figure 1. Default landing page for the ozone gas application. Dashboard title is on the top, followed by three dropdown menus (Year, PQAQ, Method). Below that is the tab selection (PQAQ or Monitor tab). Charts are displayed below that, with precision and bias charts shown here.

Precision

Precision is the measure of mutual agreement among individual measurements of the same property. In this application, the one-point QC check at a known concentration serves as an estimate of precision. The percent difference calculations between the monitor and the “assessment” concentrations are then used to generate a 90th percentile upper bound coefficient of variation (CV) precision. The goal for acceptable measurement uncertainty for precision is defined as an upper 90 percent confidence limit for the CV of 10 percent. The equation for gaseous pollutants is described in 40 CFR Part 58, Appendix A, Section 4.1.2 and is shown in [Equation 1](#) here.

$$CV = \sqrt{\frac{n \cdot \sum_{i=1}^n d_i^2 - \left(\sum_{i=1}^n d_i\right)^2}{n(n-1)}} \cdot \sqrt{\frac{n-1}{\chi_{0.1, n-1}^2}}$$

1)

In Figure 1, the top chart illustrates each PQAO's 90th percentile confidence limit. PQAOs are sorted from low to high, with the selected PQAO's precision calculation shown in the chart. A national precision statistic is also provided (i.e., 2.4% for year 2017). Additionally, the MQO goal (i.e., 7% for ozone as shown in Figure 1) is identified with a red horizontal line.

You can use a mouse or touchpad to "box in" any selected data on any chart for displaying more detailed information. Results appear below the chart with details on the agencies and/or data. Double-click on the boxed-in information to zoom the chart to the selected range.

Bias

Bias is defined as a systematic or persistent distortion of a measurement process that causes errors in one direction. The bias estimate is calculated using the same one-point QC checks but with a different equation. The bias estimator is an upper bound on the mean absolute value of the percent differences as shown in [Equation 2](#) (from 40 CFR Part 58, Appendix A, Section 4.1.3).

$$2) \quad |\text{bias}| = AB + t_{0.95, n-1} \times \frac{AS}{\sqrt{n}}$$

Where n is the number of one-point QC checks being aggregated, $t_{0.95, n-1}$ is the 95th quantile of a t-distribution with $n-1$ degrees of freedom, the quantity AB is the mean of the absolute percent differences, and AS is the standard deviation of the absolute value of the percent differences.

Other than the method of calculation, most features of the bias estimate figure are identical to CV precision. PQAOs are sorted from low to high with the selected PQAO's bias calculation shown in the chart. A national bias statistic is also provided (i.e., 1.5% for year 2017). Additionally, the MQO goal (i.e., 7% for ozone as shown in Figure 1) is identified with a red horizontal line.

Completeness

Completeness is the measure of the expected number of samples available at a monitoring site per year. Here it is defined as the number of valid reported daily values divided by the number of daily values expected, multiplied by 100. For all gas pollutants other than ozone, this is a full year of monitoring data, either 365 or 366 days. For ozone, monitoring seasons are defined by regulation in Appendix D to 40 CFR Part 58. Monitoring seasons are often consistent throughout a state, but may

vary for specific counties within a state; expected days are calculated on an individual state and county basis for ozone. The maximum percent completeness is 100%, and the MQO goal for routine completeness is 75% for all pollutants. It is commonly expected that a gas monitor will have less than 100% completeness due to QA/QC checks and instrument maintenance.

In contrast to the precision and bias plots, the completeness plot is ordered from highest completeness (100%) to lowest, as shown in [Figure 2](#). All other plot features are the same in concept to the bias and precision plots described above.

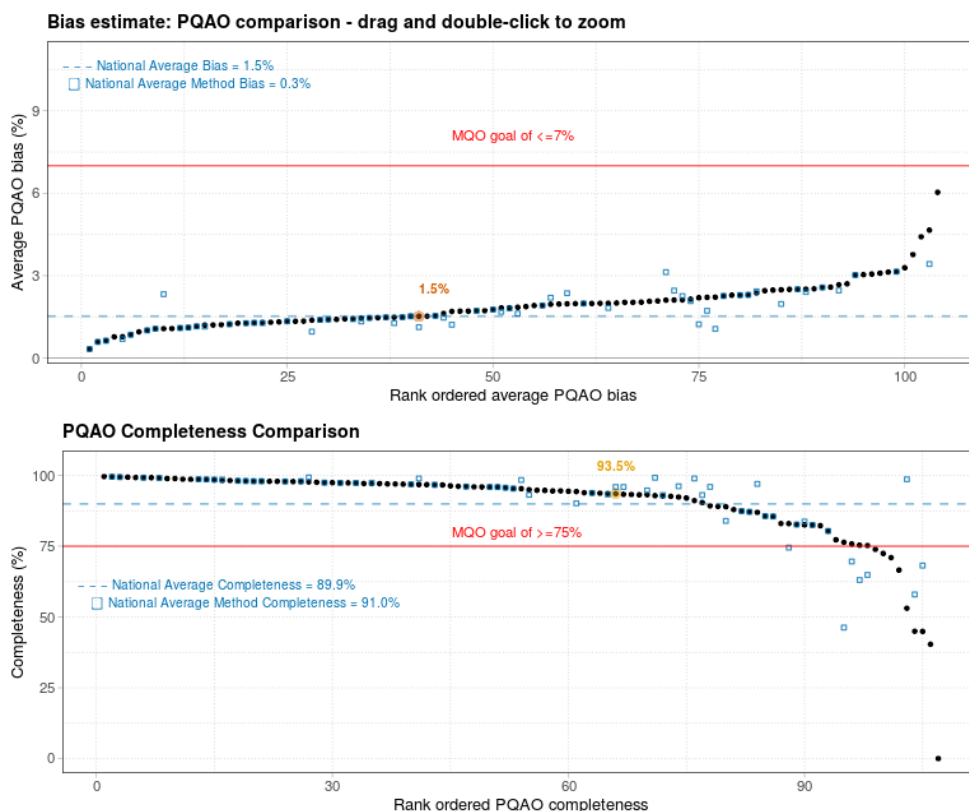


Figure 2. Screenshot of bias estimate and routine percent completeness figures from the ozone dashboard.

Monitor Tab

Clicking the monitor tab takes you to a separate view of a table and four figures. The **Monitor** tab provides monitor-specific quality information for the selected PQAQ. [Figure 3](#) shows the default landing screen on the monitor tab for Alabama Department of Environmental Management ozone monitors in 2018. The top two displays are the selection table of summary statistics by monitor and the tile plot showing data quality indicators and routine data on a color/symbol-coded calendar.

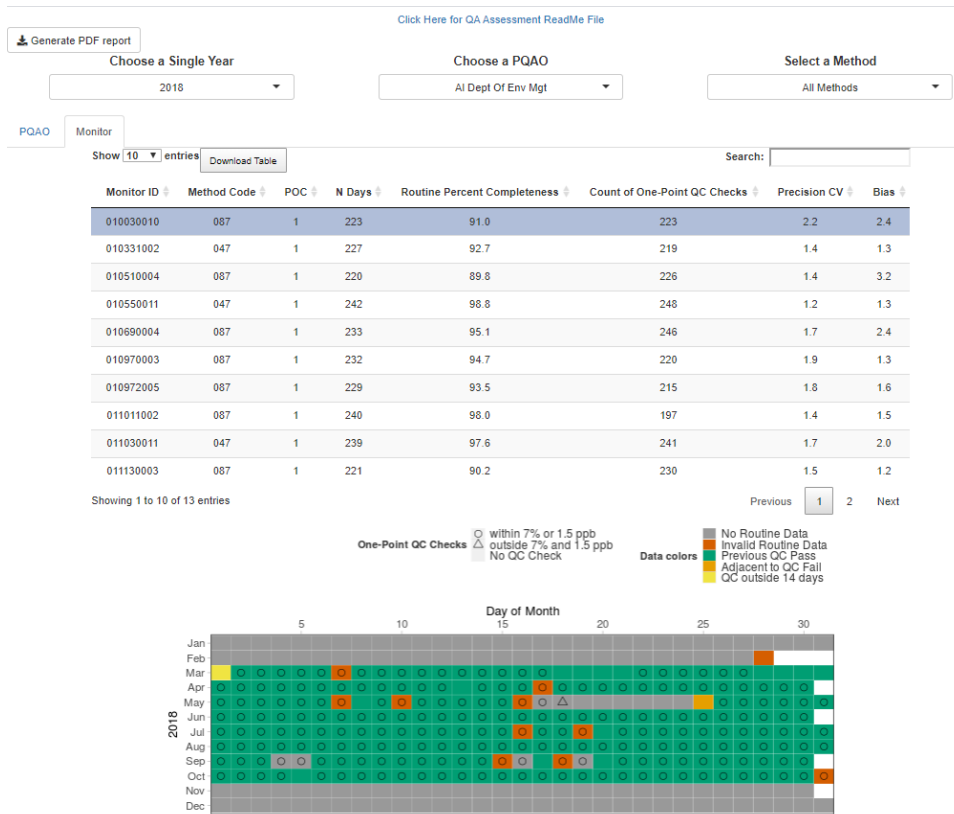


Figure 3. Default landing tab for the ozone QA data dashboard on the **Monitor** tab. Additional figures are available by scrolling down.

Monitor Table

The first feature on the **Monitor** tab is the selectable table of monitor summary statistics based on the currently selected PQAO. Figure 3 shows the table of **Monitor IDs**, which are the state, county, and site ID FIPS codes for the monitor of interest. **Method Code** and **POC** are also available as individual codes to further distinguish monitors. **N Days** indicates the number of routine monitoring days of data available for comparison to the NAAQS for the year selected. **Routine Percent Completeness** indicates the percentage of routine data available compared to the number of expected days. **Count of One-Point QC Checks** indicates the total number of checks; note that there can be multiple one-point QC checks per day and this is not a count of days with one-point QC checks. **Precision CV** and **Bias** indicate the calculated monitor values for the CV and Bias metrics. Click on a row of the table to select the monitor of interest and highlight that row. You can select only one monitor at a time, and the first row is always selected first by default. Subsequent figures on the **Monitor** tab display the charts for that monitor.

Tile Plot

The tile plot is the first data visualization on the **Monitor** tab. It shows a calendar of daily routine monitoring data and one-point QC check data availability by a combination of color codes and symbols. The tile plot itself represents a calendar of months (rows-horizontal) and days (columns-vertical). Symbols within individual squares indicate one-point QC check results for the day. Colors for the day indicate the presence or absence of routine monitoring data and the interaction of the routine monitoring data with the nearest one-point QC check results. An example with most combinations is shown in [Figure 4](#).

Symbols include:

- Blank – No one-point QC check on that day
- Hollow circle – A single one-point QC check within the MQO (within acceptance criteria, e.g., within 7% or 1.5 ppb for ozone)
- Hollow triangle – A single one-point QC check outside the MQO (outside acceptance criteria e.g., outside 7% and 1.5 ppb for ozone)
- Symbol X – All one-point QC checks reported as invalid for that day
- Circle with cross – Multiple one-point QC checks, all within the MQO criteria
- Filled triangle – Multiple one-point QC checks, at least one of which is outside the MQO criteria

Colors for tiles include:

- Gray – Routine data for that day is missing or not reported
- Dark orange – Routine data for that day is flagged as invalid in AQS
- Green – Routine data is available and adjacent to one-point QC check within MQO criteria
- Light orange – Routine data is available and adjacent to any one-point QC check outside MQO criteria
- Yellow – Routine data is available and outside of 14 day one-point QC check window

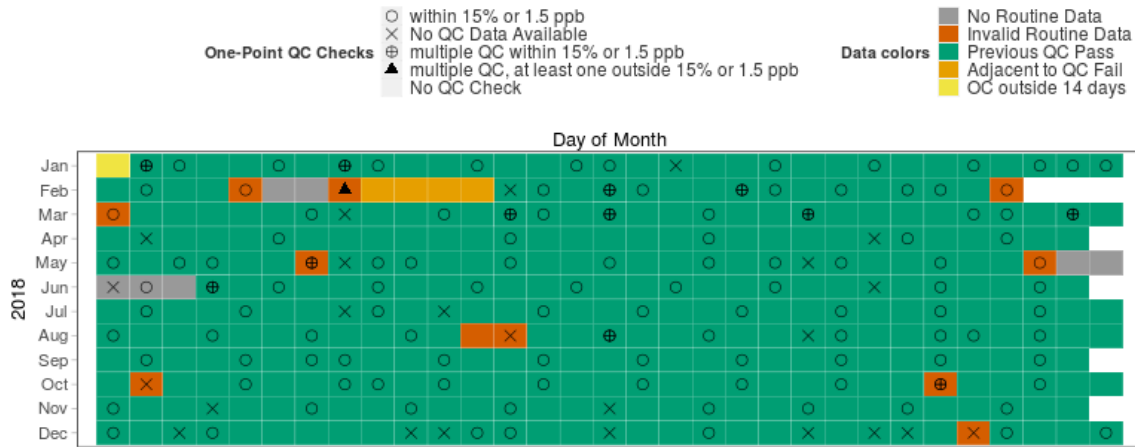


Figure 4. Tile plot example for the NO₂ QA data dashboard showing symbols and color codes.

The hierarchy for choosing colors is as follows:

- Is the routine daily data available and complete; assess **Validity Indicator** from AQS dailyData
 - NO, flag as **No Routine Data** and color it gray
 - YES, are 75% or more hourly records available and valid?
 - NO, flag as **Invalid Routine Data** and color it dark orange
 - YES, is the day at least 13 days after the most recent one-point QC check day?
 - NO, flag as **Outside 14 Day Window** and color it yellow
 - YES, is either the previous or subsequent one-point QC check day outside of MQO criteria?
 - NO, flag as **Adjacent to QC Fail** and color it light orange
 - YES, flag as **QC Pass** and color it green

Time Series of One-Point QC Checks and One-Point QC Concentrations

Below the tile plot on the **Monitor** tab are two time series figures displaying one-point QC check percent difference values or one-point QC concentration values, respectively, as shown in Figure 5. The top part of the plot shows the percent differences relative to the MQO goal. In the image for NO₂ in 2018, the MQO goal is for the one-point QC check to be within a band of $\pm 15\%$. As in other time series, click-and-dragging around the points will yield a table of individual one-point QC check dates, concentrations, and percent differences for the selected date range.

The time-series of one-point QC concentrations shows the monitor and assessment concentrations that yield the percent differences. The absolute concentrations can be examined to determine the QC check concentration level and whether changes in absolute concentrations explain any patterns in

the percentage differences of the QC checks. Dragging-and-clicking around individual values brings up the monitor and assessment concentrations for the selected date range.

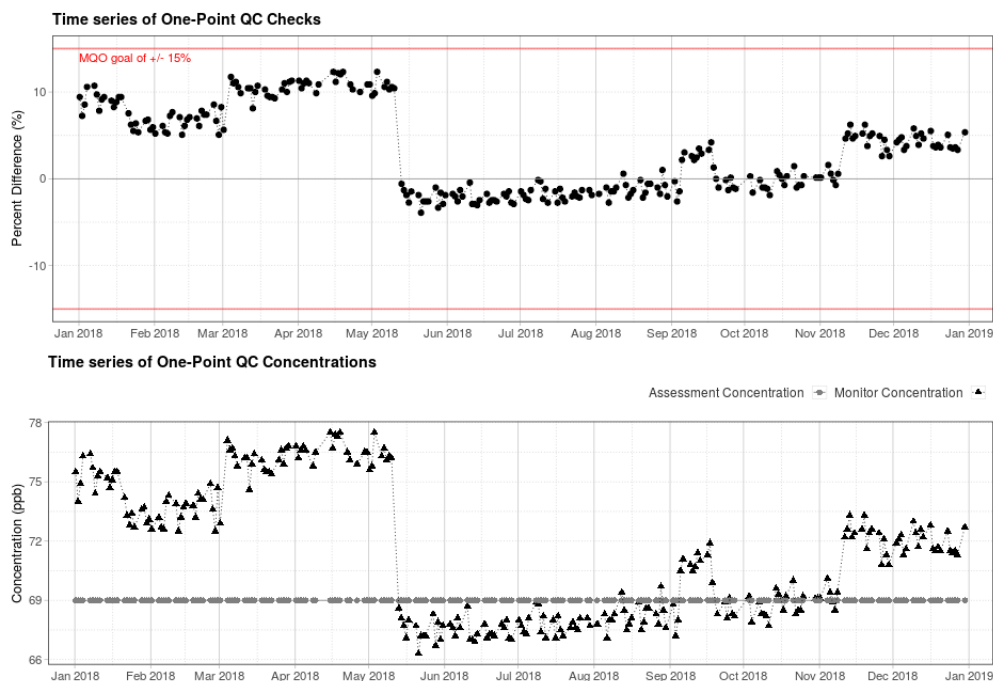


Figure 5. Monitor tab time series plots for one-point QC checks (%) and one-point QC concentrations (ppb).

Time Series of Concentrations

In the bottom figure of the **Monitor** tab, a time-series of concentrations is displayed. The durations are either 1-hr maximum concentrations (NO_2 , NO_y , SO_2) or 8-hr maximum concentrations (ozone, CO). If a pollutant has a NAAQS value, it is displayed as a red line (except for CO, because of nationally low ambient CO concentrations). As in other time series plots, drag and click over a series of points to display them in a table beneath the figure. Date ranges are lined up in the two figures, so you can compare one-point QC concentrations with the observed maximum concentrations.

Additional Notes

- The data source for the application is the EPA AQS API (https://aq5.epa.gov/aqsweb/documents/data_api.html). Data from the most recent complete 2-year period (including the current incomplete year) are updated on a weekly basis. Older data are updated on a quarterly basis. [Example: During 2020, data for 2018-2020 are downloaded weekly. Data from 2017 or before are downloaded quarterly.] The date of the most recent data acquisition is shown as a footnote on each tab. We welcome any comments on this work. Please email any comments to Tim Hanley in EPA's Ambient Air Monitoring Group (hanley.tim@epa.gov).