Continuous Carbon Data Dashboard

Contents

1 Summary

This application provides a data analysis and quality assurance (QA) dashboard for continuous carbon measurements that are submitted to the EPA's AQS database from sites across the country. For this purpose, the term **continuous carbon** references hourly concentrations of UV Carbon and Black Carbon measured by aethalometers and other available instrumentation. This includes the following parameters:

- 84313: Black carbon PM2.5 STP
- 84314: UV Carbon PM2.5 STP
- 88313: Black Carbon PM2.5 at 880 nm
- 88314: UV Carbon PM2.5 at 370 nm
- 88317: Black Carbon PM2.5 Corrected

Sample data from 2016 through the present year are available for review in most tabs. The visual assessments are intended to assist agencies by providing quick summary access to information at the site or network level of continuous carbon measurements. 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 site of interest is compared to all other sites, 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 dashboard for their program's monitors.

The Continuous Carbon Data Dashboard provides multiple indicators and analysis tools for the key subset of continuous carbon parameters.

- Site location and selection through a United States site map
- Yearly completeness metrics across the network
- Pollution roses
- Generic time series investigation tool – combines a time series, scatter plot, and pollution rose visualization of any parameter or combination of two parameters
- Seasonality analysis
- Method QA and Bias tab (comparison to CSN/IMPROVE measurements)
- Percentage carbon makeup of total $PM_{2.5}$ concentrations
- Black Carbon Source Apportionment

2 Tabs Overview

The user can view two types of assessments within the data quality dashboard as shown in Figure 1. Tabs colored red indicate *network level assessments* that focus on network wide quality indicators. Tabs colored purple indicate *specific site-level assessments* focused on site-specific quality indicators and analysis tools. Red tabs will always be shown. Purple tabs require that a site is selected on the Map Tab (see **Map Tab** section). The EC to BC Ratio tab and the PM2.5 Percentile Ratios tab appear only if a site is (1) selected and (2) has the required data for display.

Figure 1: Dashboard tabs, shown for the NCore dashboard, categorized by color. Red are Network-level Assessments. Purple are Specific Site-level Assessments, and require that a site is selected on the Map Tab.

3 Network-wide Quality Indicator Tabs

Four network wide tabs are available: Map, Network Concentrations, and Network Completeness.

3.1 Map

The map tab is the default landing page for a data quality dashboard. Figure 2 shows the default map display for the Continuous Carbon dashboard. Each circle on the map represents an individual monitoring site. Sites are colored to specify how continuous carbon data is reported at a site. Blue indicates that continuous carbon measurements are reported at local conditions. Yellow indicates that continuous carbon

measurements are reported at STP. Green indicates that continuous carbon measurments are reported under both conditions.

To **select a site**, a user should click a circle on the map. A user must select a site on the map to view metrics on purple *Site Specific Quality Indicator* tabs. To **deselect a site**, a user may click the background of the map (i.e. any non-circle location on the map).

The table below the map provides site details, including location, site code, site name, conditions under which continuous carbon measurments are reported, as well as the availability of EC data at a site. The table will update once a site is selected on the map to show information only for the selected site. Deselecting a site will re-populate the table with information for all sites.

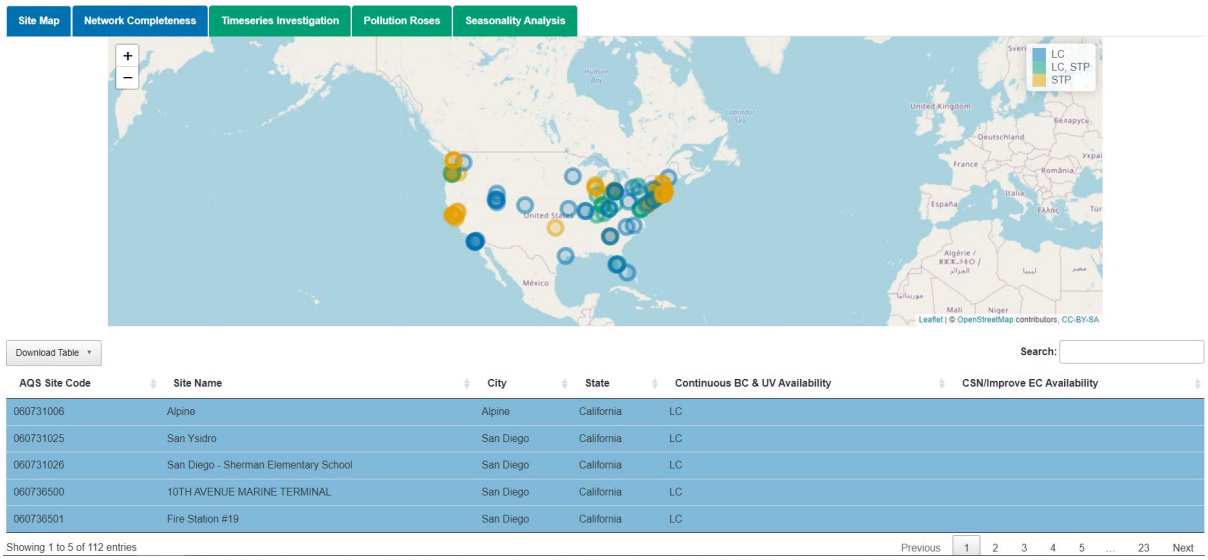


Figure 2: Map tab (default landing page) for the Continuous Carbon Dashboard. Sites shown in the map are also listed in the table below.

3.2 Network Concentrations

Figure 3 shows the Network Concentration tab display. This tab compares daily average carbon concentrations across the network. The top figure shows LC carbon measurements and the bottom figure shows UV carbon measurements.

Each box in the figures represents the distribution of daily average carbon measurements per site (as shown on the x-axis) per parameter (as shown by the box color) across the selected year. A user may select which year is displayed using the "Select year" dropdown menu. The boxplots display standard metrics, with each rectangle spanning the 25th and 75th percentiles, split by a line representing the median.

A user may select a subset of sites to display using the "Filter sites" dropdown menu. As a default, all sites with data available during the chosen year are displayed.

If a site is selected on the Map Tab, the box(es) representing measurements at the site will be outlined in orange.

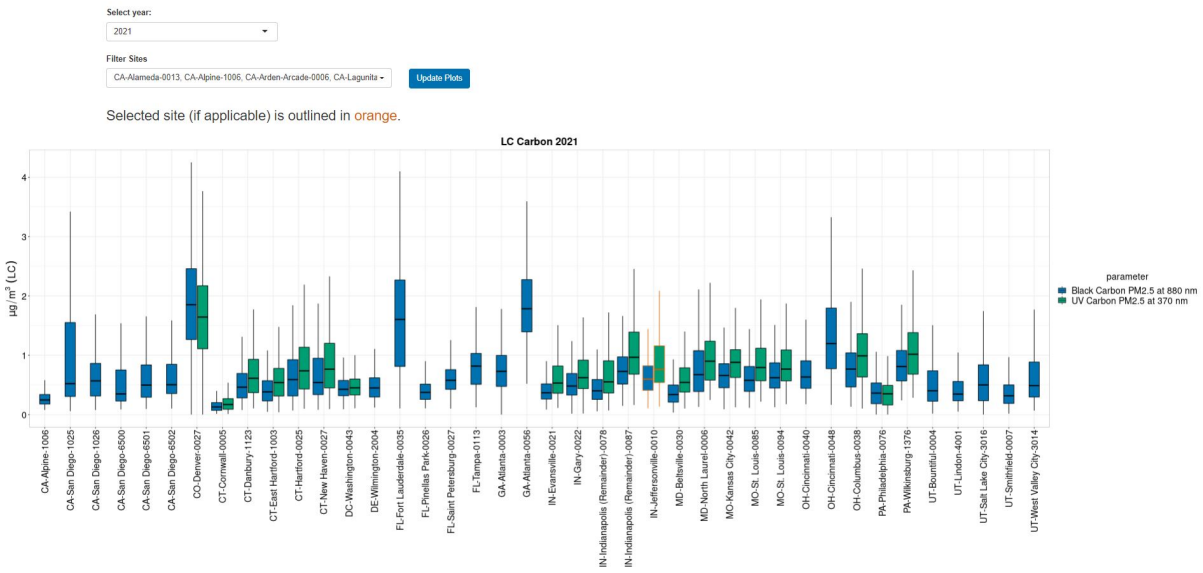


Figure 3: Network Concentration tab, showing the LC carbon plot for 2021 across the network. Boxes representing the Jefferson, Indiana site are oulined in orange, as this site was selected on the map tab.

3.3 Network Completeness

Figure 4 shows the Network Completeness tab display for required parameters at program monitoring stations. The figure shows the percent completeness of required pollutant parameters for the program network by site and year. The main table shows parameter and duration percent completeness for the site selected on the Map Tab (if one has been selected).

Year may be changed using the drop-down menu above the plot. The menu only shows complete years of monitoring data. The dropdown menu allows a user to select partial years (before reporting deadlines to AQS), but will display a warning that the completeness metrics for incomplete years are estimates (see below for methodology).

Each monitoring site is represented by a gray line segment according to rank-ordered completeness, plotted on the x-axis. Completeness (%) per parameter is plotted on the y-axis. Each parameter measured at a site is represented by a colored shape and plotted on the gray line segment representing the site.

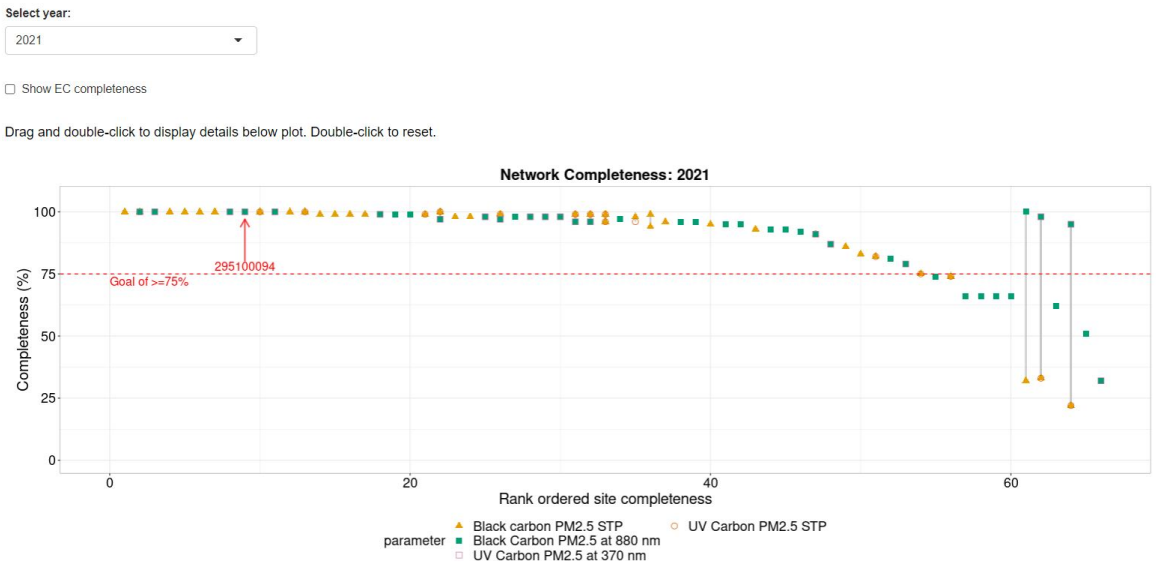
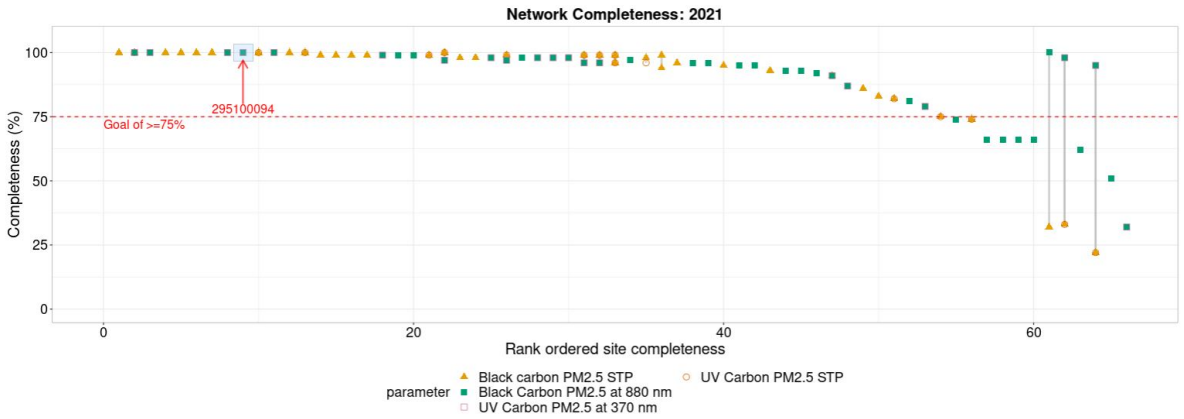


Figure 4: Network Completeness plot.

If a site is selected on the Map Tab, the segment representing that site is highlighted on the plot. A user can drag and double-click across a subset of the plot to view detailed information about the data point enclosed within the drag-box, which will appear in a temporary table between the plot and the main table (see Figure 5).



Site Rank	Site Code	POC	Year	Parameter	Method	Duration	% Complete (Parameter)	% Complete (Site Average)
9	295100094	1	2021	Black Carbon PM2.5 at 880 nm	Magee AE33/ TAPI M633 Aethalometer - Optical absorption	1 HOUR	100.00	100.00
9	295100094	1	2021	UV Carbon PM2.5 at 370 nm	Magee AE33/ TAPI M633 Aethalometer - Optical absorption	1 HOUR	100.00	100.00

Figure 5: Network completeness tab with dragbox over the representing site 090050005 (rank order site completeness = 20). The temporary table is shown below the plot.

3.4 Completeness Calculation Methodology

C statistics are reported in annual data to AQS. For complete years, the completeness statistics displayed are pulled directly from annual data.

Incomplete years are years for which data is not yet required to be submitted to AQS through the end

of the year. Annual data from incomplete years may be unavailable or inaccurate. The completeness of a datastream for incomplete years is estimated according to the methodology below.

3.4.1 Completeness estimations for incomplete years

3.4.1.1 Step 1: Identify the "required" date The deadline for submitting data to AQS is two quarters before the most recent complete quarter.

E.g. If today is June 6, 2022, the most recent complete quarter is Q1 2022 which ended on March 31, 2022. 2 quarters before this date is the end of Q3 2021, **September 30, 2021**.

3.4.1.2 Step 2: Identify incomplete years Incomplete years include all years greater than or equal to the year of the required date.

Examples:

- On June 6, 2022, the required date is September 30, 2021. **Incomplete years include 2021 and 2022**.
- On July 1, 2022, the required date is December 31, 2021. **Incomplete years include only 2022** because data is required through the end of 2021.

3.4.1.3 Step 3: Pull datastream completeness statistics for complete years Completeness statistics for complete years are pulled directly from the `Observation Percent` column in annual data submitted to AQS. This completeness is delivered per unique datastream at a site (per site, poc, parameter, method).

3.4.1.4 Step 4: Estimate datastream completeness statistics for incomplete years Completeness per datastream (unique site, poc, parameter, method) for incomplete years is calculated against EITHER the required date, or the maximum date of available data if data extends beyond the required date.

The expected number of observations per data stream is calculated using the sample frequency recorded in sample data from AQS. If more than one sample frequency is available for a datastream across available years, the most recent recorded sample frequency is used.

A complete list of sample frequencies can be found here. The expected number of observations per year are calculated per sample frequency as shown in the list below (e.g. Every 4th day = 365 / 4 = 91.25 expected days). Seasonal designations are ignored in this estimation (e.g. PAMS). Sample frequencies without an obvious cadence (seasonal, random, episodic, etc.) are assigned 365. **Decimal days are rounded in subsequent steps.**

- 'EVERY DAY' = 365 expected days
- 'EVERY 24TH DAY' = 15.208 expected days
- 'EVERY 30 DAYS' = 12.167 expected days
- 'EVERY OTHER DAY' = 182.5 expected days
- 'EVERY 3RD DAY' = 121.6677 expected days
- 'EVERY 4TH DAY' = 91.25 expected days
- 'EVERY 5TH DAY' = 73 expected days
- 'EVERY 6TH DAY' = 60.833 expected days
- 'EVERY 12TH DAY' = 30.417 expected days
- 'STRATIFIED RANDOM' = 365 expected days
- 'RANDOM' = 365 expected days
- 'EVERY 90 DAYS'= 4.055555556 expected days
- 'DAILY: 24 - 1 HR SAMPLES -PAMS' = 365 expected days
- 'DAILY: 8 - 3 HR SAMPLES - PAMS' = 365 expected days

- 'DAILY: 1 - 3 HR SAMPLE - PAMS' = 365 expected days
- 'DAILY: 1 - 24 HR SAMPLE - PAMS' = 365 expected days
- 'DAILY: 4 - 6 HR SAMPLES - PAMS' = 365 expected days
- 'DAILY: 4 - 3 HR SAMPLES - PAMS' = 365 expected days
- 'EVERY 3RD DAY:24-1 HR - PAMS' = 121.667 expected days
- 'EVERY 3RD DAY:8-3 HOUR - PAMS' = 121.667 expected days
- 'EVERY 3RD DAY:1-3 HR - PAMS' = 121.667 expected days
- 'EVERY 3RD DAY:1-24 HR - PAMS' = 121.667 expected days
- 'EVERY 3RD DAY:4-6 HR - PAMS' = 121.667 expected days
- 'EVERY 3RD DAY:4-3 HR - PAMS' = 121.667 expected days
- 'EVERY 6TH DAY:24-1 HR - PAMS' = 60.833 expected days
- 'EVERY 6TH DAY:8-3 HR - PAMS' = 60.833 expected days
- 'EVERY 6TH DAY:1-3 HR - PAMS' = 60.833 expected days
- 'EVERY 6TH DAY:1-24 HR - PAMS' = 60.833 expected days
- 'EVERY 6TH DAY:4-3 HR - PAMS' = 60.833 expected days
- 'EPISODIC SAMPLING' = 365 expected days
- 'SEASONAL' = 365 expected days
- '5 OUT OF 7 DAYS (FOR PSI)' = 260.7142857 expected days
- 'EVERY 6TH DAY: 3-3 HR - PAMS' = 60.83333333 expected days

Completeness per incomplete year is estimated as follows:

$Completeness(\%) = 100 * \frac{Obs_{recorded}}{Obs_{possible}}$

Where:

- $IncompleteMaxDate$: The maximum of the required date, and the most recent observation available in AQS (within the year in question)
- $IncompleteMinDate$: The first day of the incomplete year (e.g. if the required date is 2021-09-30, Incomplete Min Date = 2021-01-01)
- $FracRequiredDays$: Expected yearly observations based on sample frequency (see list above) / 365
- $Obs_{recorded}$: Number of complete daily observations (>= 75% completeness in sample data) between $IncompleteMinDate$ and $IncompleteMaxDate$
- $Obs_{possible}$: (Number of days between $IncompleteMinDate$ and $IncompleteMaximumDate$) * $FracRequiredDays$

Notes and Assumptions

- Completeness estimates > 100% are capped at 100%. Estimates greater than 100% could result from inaccurately recorded sample frequencies.
- Completeness for seasonal datastreams may be underestimated.
- Leap years are not accounted for in this estimation
- Decimal days are rounded to the nearest day

3.4.1.5 Step 5: Calculate site completness statistics Average site completeness is calculated as the mean completenss per year across all datastreams at a site. Minimum and maximum site completeness is calculated as the minimum and maximum completeness values per year across all datastreams at a site.

4 Site Specific Quality Indicator Tabs

4.1 Time Series Investigation

The Time Series investigation tab allows a user to compare two parameters, one hourly, and the other either hourly or daily, in a time series, pollution rose, and a scatter plot. **Pollution rose and scatter plot**

comparisons are only available if both parameters are hourly.

Options at the top of the tab allow a user to select a primary hourly parameter to display, and optionally, a secondary hourly or daily parameter to compare. Once the user selects a radio button other than "None" to display a secondary parameter, a drop-down menu will appear with the available options based on chosen sample duration (e.g. 1 HOUR, 24 HOUR, etc.).

If one hourly parameter is chosen:

- The timeseries plot will display the primary parameter concentration as a solid line.
- A single pollution rose will display the primary parameter concentration.
- Scatter plot comparison is unavailable.

If two hourly parameters are chosen:

- The timeseries plot will display the primary parameter concentration as a solid line and the secondary parameter as a dashed line. A secondary y-axis (right-hand side) will be included to reference the secondary parameter values (figure 6).
- Two pollution roses will display the primary and secondary parameter concentrations, respectively (figure 7).
- A scatter plot comparison will display, with the option to overlay a one-to-one line and/or a linear regression line using check-box selections (figure 8).

If one hourly parameter and one daily parameter is chosen:

- The timeseries plot will display the primary parameter concentration as a solid line and the secondary parameter as points. A secondary y-axis (right-hand side) will be included to reference the secondary parameter values.
- A single pollution rose will display the primary parameter concentration (daily pollution rose unavailable).
- Scatter plot comparison is unavailable.

4.1.1 Plot Reactivity

A user may zoom in on a particular time range in the timeseries plot by clicking-and-dragging a box across the desired subset of data on the plot, and double-clicking within the drawn box to zoom in. If displayed, pollution roses and the scatterplot will update to reflect this zoomed time period. To reset the zoom, double-click the timeseries plot.

Two reference lines are available to overlay on a scatter plot comparison: a one-to-one line, and a linear regression line. These may be drawn by selecting one or both check boxes to the right of the scatter plot. A one-to-one line may not appear within the plot boundaries if the scales of the primary and secondary parameters do not overlap (e.g. ozone and barometric pressure).

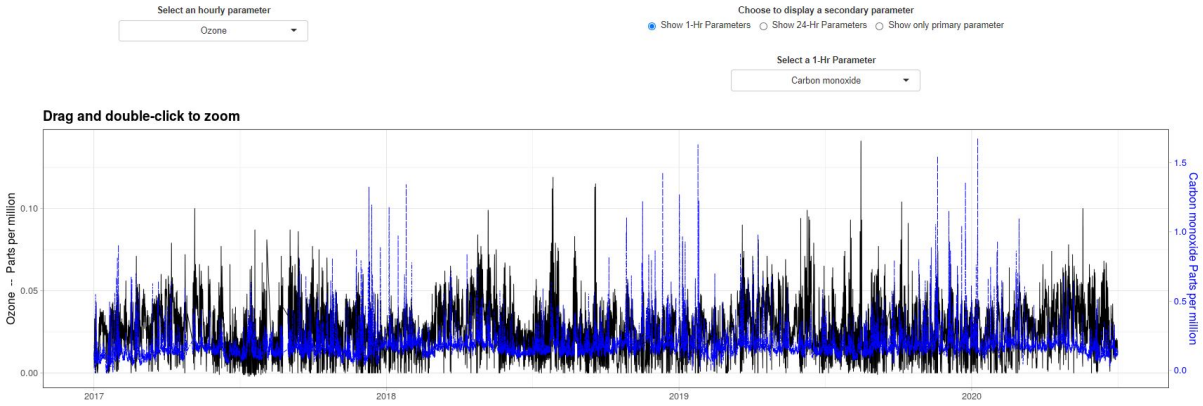


Figure 6: Time series plot a primary (hourly) and secondary parameter (hourly). The right-hand y-axis represents the secondary parameter.

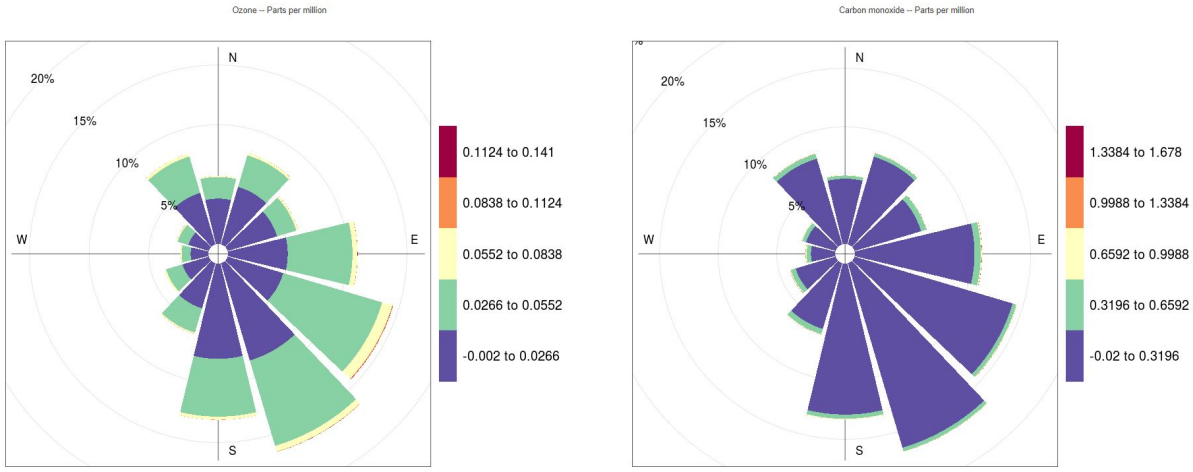


Figure 7: Pollution roses for primary (hourly) and secondary parameters (hourly).

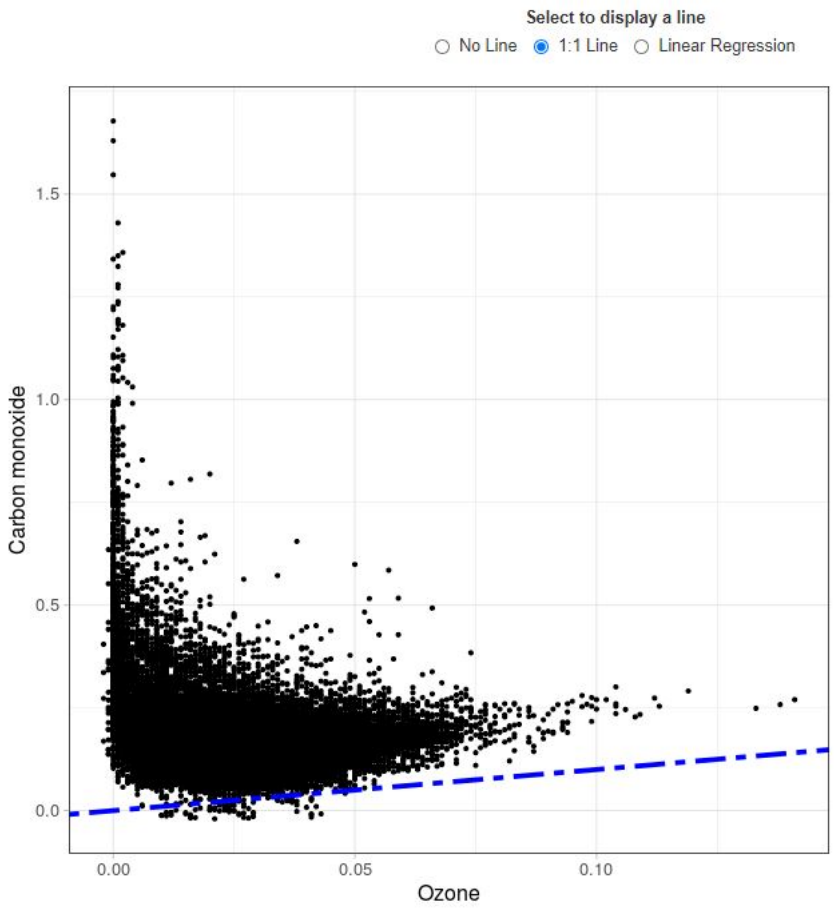


Figure 8: Scatter plot comparing primary (hourly) and secondary parameter (hourly), and one-to-one and linear regression lines shown. A scatter plot comparison is only available when two hourly parameters are chosen.

4.2 Pollution Roses

The pollution roses tab shows pollution roses for continuous UV carbon and black carbon measurements at the monitoring site selected on the Map Tab. Figure 9 shows the layout of the tab. There are drop-down menus to select the displayed parameters, year (or year range), day of the week (or range), hour of the day (or range) and meteorological season (or seasons). The seasons are defined as:

- Spring: March, April, and May
- Summer: June, July, and August
- Fall: September, October, and November
- Winter: December, January, and February

Below the two pollution roses is an urban-scale (~25 km) map showing the site location within its local area.

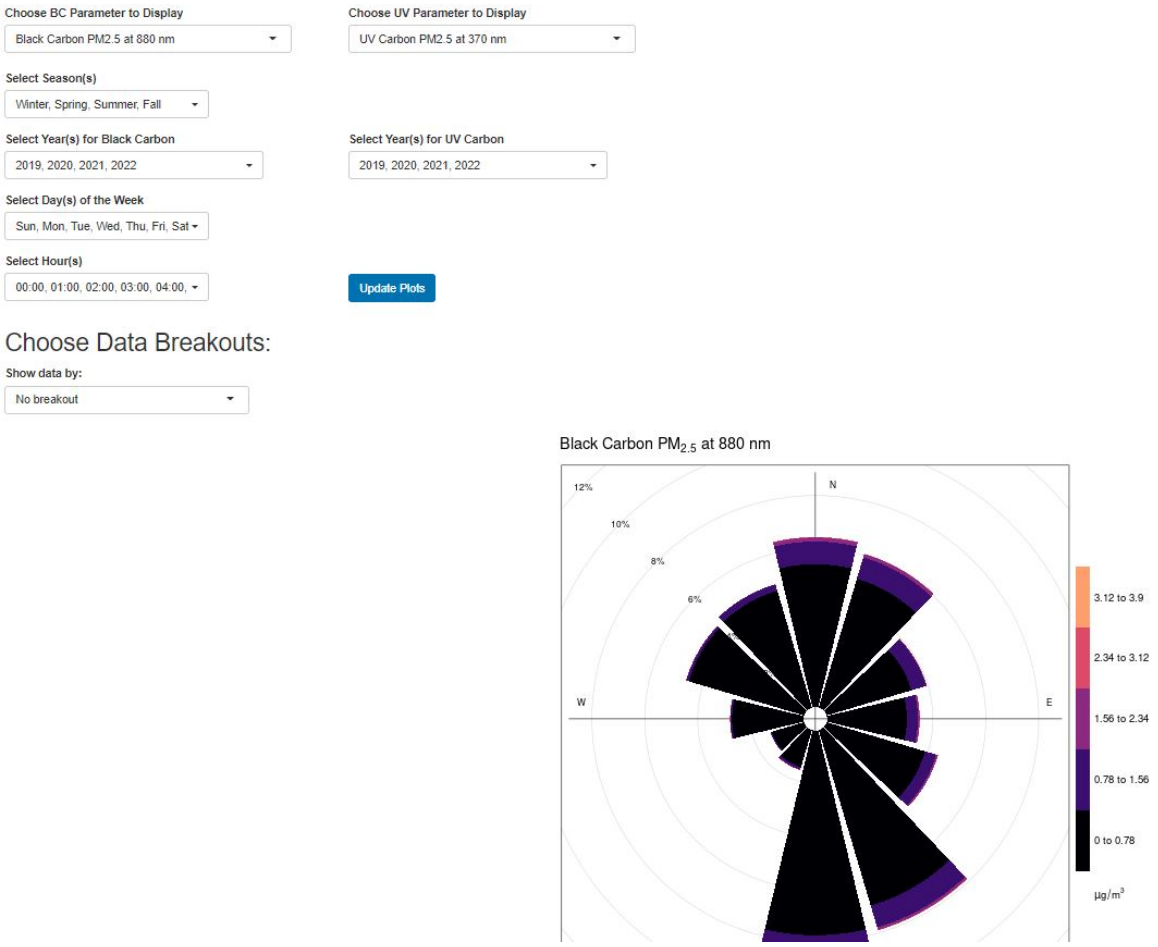


Figure 9: Pollution Roses Tab showing concentrations of continuous carbon, along with a zoomable map of the site and its surroundings.

Plots may further be broken down within the data selections to show a side-by-side comparison between seasons, months, days and hours included in the data selections (10).

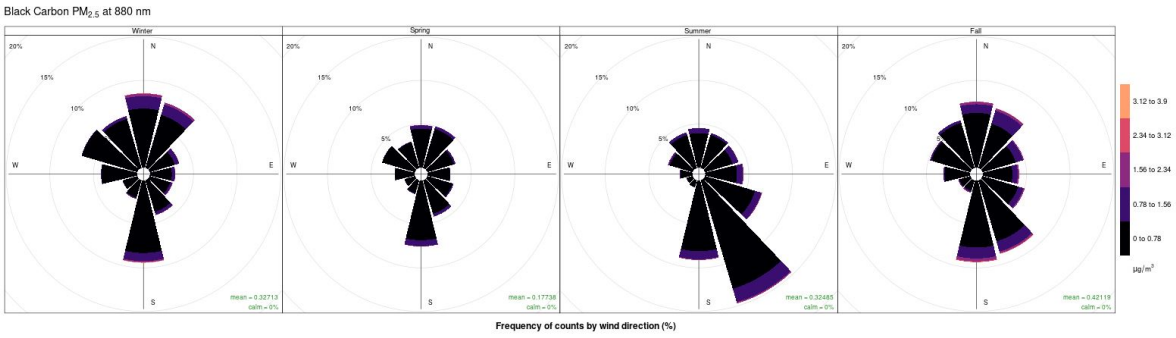


Figure 10: Pollution Rose broken out by season

4.3 Seasonality Tab

The seasonality tab allows a user to compare concentrations across hours of the day, days of the week, and months of the year. Figure 11 shows the layout of the tab. There are drop-down menus to select the displayed parameter, year (or year range), day of the week (or range), and hour of the day (or range). The generated collection of plots represents data filtered by these selections. For example, if multiple years are selected, the average concentrations per month include all selected years. .

Figure 11 Plot descriptions:

- Top: Average hourly concentrations (diurnal profile) by day of the week for the selected parameter
- Lower-left: Average hourly concentrations (diurnal profile) for the selected parameter
- Lower-middle: Average concentration per month for the selected parameter
- Lower-right: Average concentration per day of the week for the selected parameter

On each plot, the red line connects the mean value per time break, and the ribbon or box shows the extent of the 95% confidence interval around the mean.

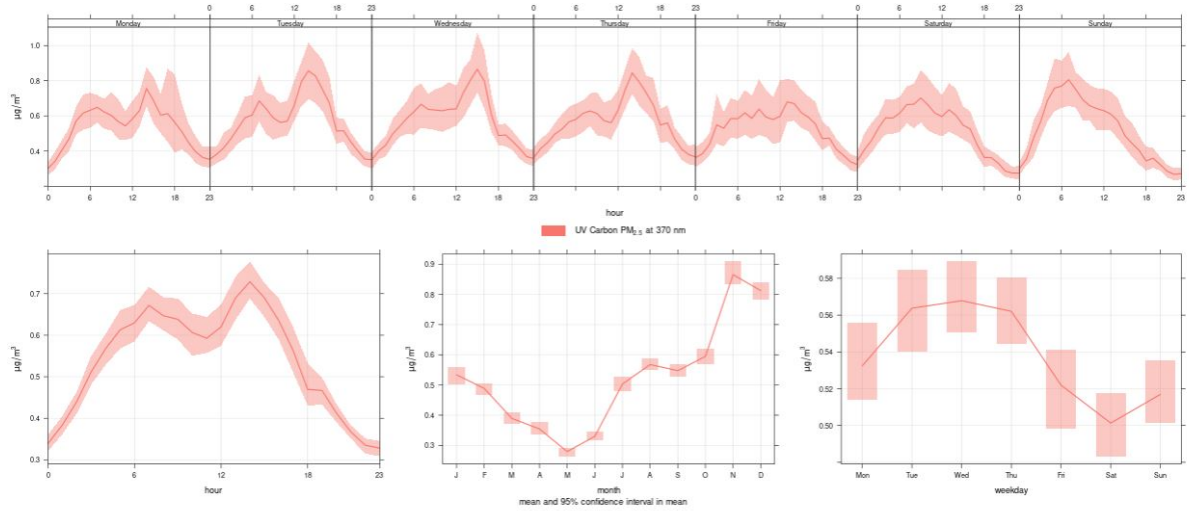


Figure 11: Seasonality plots showing variations in average concentrations by hour of the day, day of the week, and month of the year

4.4 Method QA Tab

The Method QA tab displays the percent difference between (1) continuous black carbon [BC] measurements and collocated Elemental Carbon [EC] measurements and (2) continuous UV carbon measurements and collocated Organic Carbon [OC] measurements. Collocated EC and OC measurements are part of the CSN/IMPROVE networks. This tab will only appear if the selected site has sufficient CSN/IMPROVE measurements (some sites that measure continuous carbon do not also measure EC/OC).

Only sites that measure continuous carbon at LC are included because CSN/IMPROVE measurements are taken at LC. The percent difference is calculated between the daily CSN/IMPROVE measurement and the daily continuous measurement (arithmetic mean).

For an observation pair to be considered valid:

- The daily measurements must meet 75% completeness criteria (for hourly measurements, at least 18 of 24 possible hourly measurements in a day)

- The CSN/IMRPOVE measurement must be greater than the average daily detection limit (as reported in AQS sample data).

Daily percent difference is calculated via the following equation:

$$\text{Percent Difference} = 100 * \frac{\text{Continuous}_{Daily} - \text{CSN/IMPROVE}_{Daily}}{\text{CSN/IMPROVE}_{Daily}}$$

A Negative percent difference indicates that the daily CSN/IMPROVE value was greater than the daily continuous value.

The Network-wide Method Bias boxplot (Figure 12) shows the distribution of daily percent difference between continuous and CSN/IMPROVE carbon measurements by continuous measurement method across all sites with valid collocated measurements. A user may display this plot for a subset of sites and years.

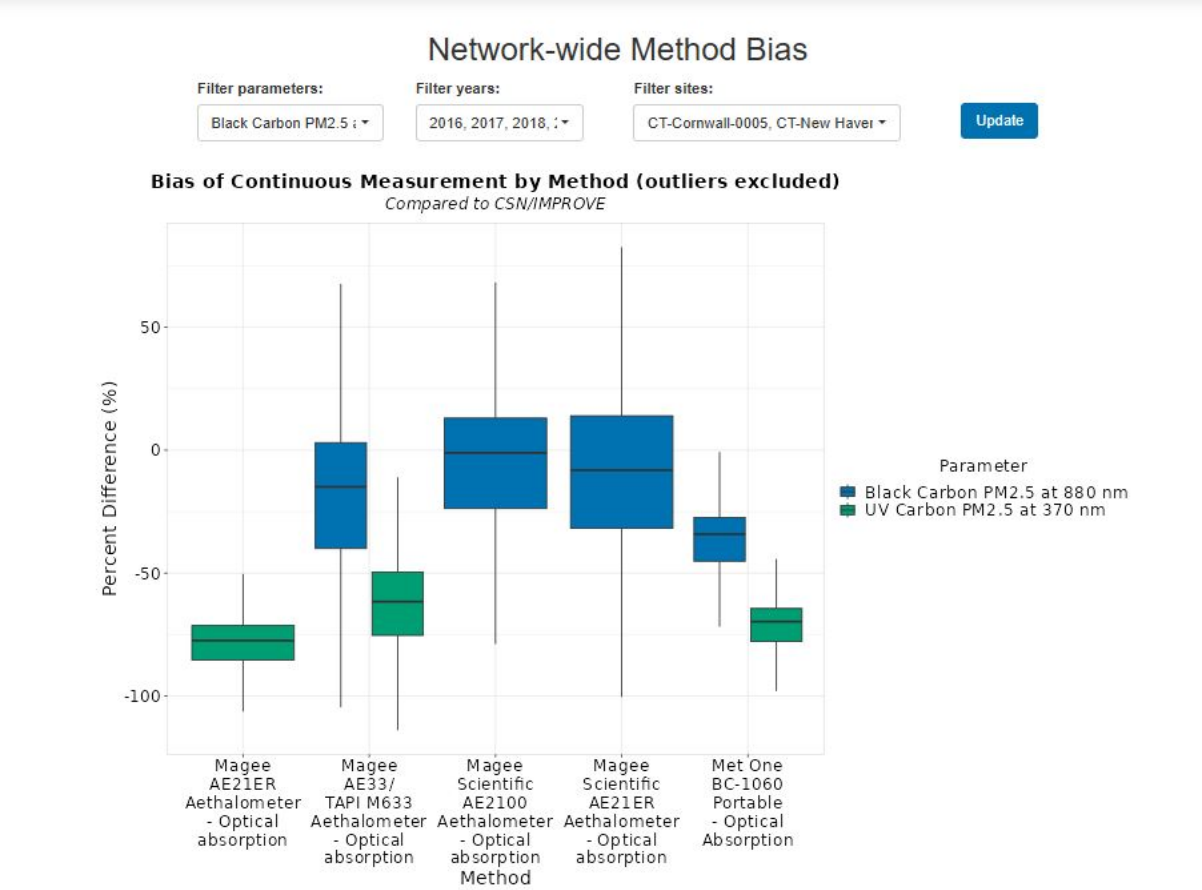


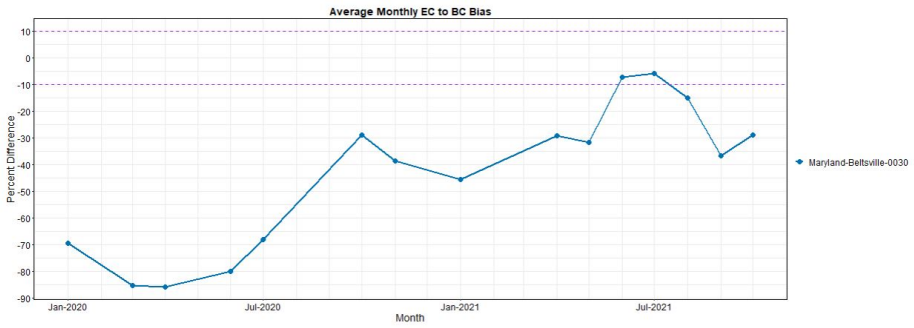
Figure 12: Continuous to CSN/IMPROVE network0wide comprison boxplot

An example of the Method QA comparison plot is shown in Figure 13. This line plot displays shows either the daily percent difference, or the mean daily percent difference per month for months with at least 10 valid observation pairs (see above for criteria). These plots are separated into panels per CSN/IMPROVE measurement method.

A scatterplot shows the linear regression between collocated monitors at each selected site.

The Plot Data table, shown in Figure 13 shows the monthly average CSN/IMPROVE concentration, the monthly average continuous concentration, and the monthly average percent difference (data shown in plot).

Note that the monthly average percent difference is an average of daily percent difference, and cannot be replicated by inserting the monthly average concentrations into the formula above.



Plot Data

Month	Mean Monthly EC ($\mu g/m^3$)	Mean Monthly BC ($\mu g/m^3$)	Mean Monthly Percent Difference (%)
Jan-2020	0.52	0.19	-69.43
Mar-2020	0.33	0.06	-85.27
Apr-2020	0.37	0.07	-85.8
Jun-2020	0.4	0.08	-80.08
Jul-2020	0.43	0.14	-68.1
Oct-2020	0.48	0.37	-28.77
Nov-2020	0.76	0.46	-38.56
Jan-2021	0.56	0.28	-45.61
Apr-2021	0.46	0.34	-29.05

Figure 13: Bias plot per site and accompanying data table

4.5 PM2.5 Ratio Tab

The $PM_{2.5}$ Ratio Tab allows a user to compare the percent carbon makeup of total $PM_{2.5}$ on high $PM_{2.5}$ days and low $PM_{2.5}$ days. This tab only appears if a site (1) records measurements of $PM_{2.5}$ at local conditions (AQS parameter code 88101), and (2) records measurements of continuous carbon at local conditions.

$PM_{2.5}$ at local conditions is measured with either a federal required method (FRM) or a federal equivalent method (FEM). The display will default to showing ratios calculated against the "top ranked" $PM_{2.5}$ data stream if multiple data streams are available at a site. $PM_{2.5}$ data streams are ranked as follows, with 1 being the highest rank:

- Data streams are first ranked by method type with FRM methods ranked above FEM methods
- Next, data streams are ranked by POC with lower numbered POCs ranked above higher numbered POCs

E.g. site has 3 PM2.5 instruments, POC1 FEM, POC2 FRM, POC2 FEM. The data streams will be ranked (1) POC2 FRM, (2) POC 1 FEM, (3) POC2 FEM.

A user may choose up to 5 sites to compare to the selected site. Comparison sites available in the dropdown menu are filtered by the selected carbon parameter. Ratios displayed for comparison sites are calculated as the chosen carbon parameter to the top-ranked PM2.5 method available at the site (see ranking methodology above).

A user may filter data to a year (or range of years), and either by month or season.

High $PM_{2.5}$ days are defined as days with a daily $PM_{2.5}$ measurement greater than the 75th percentile daily $PM_{2.5}$ concentration. Low carbon days have a daily $PM_{2.5}$ measurement lower than the 25th percentile daily $PM_{2.5}$ concentration. 75th and 25th percentile thresholds are calculated per site (if multiple sites chosen) on the subset of filtered data chosen by the user (filter options described below).

Ratios shown only include PM2.5 measurements $> 3\ \mu g/m^3$. Because of this, certain data selections may cause a site to omit a "low pm2.5" display. This is because the 25th percentile for the selected data is <= 3 ug/mˆ3 at that site.

The first plot shows a boxplot of the range of Carbon:PM2.5 ratios per site (within the subset of data filters), separated by high carbon days and low carbon days (figure 14). The box and whiskers represent standard metrics, with the box boundaries spanning the 25th to 75th percentile and whiskers spanning 1.5*IQR from the box boundaries. The median is shown as a black line within the box, and a notch is included which approximates a 95% confidence interval around the median value. Outliers are excluded from the plot.

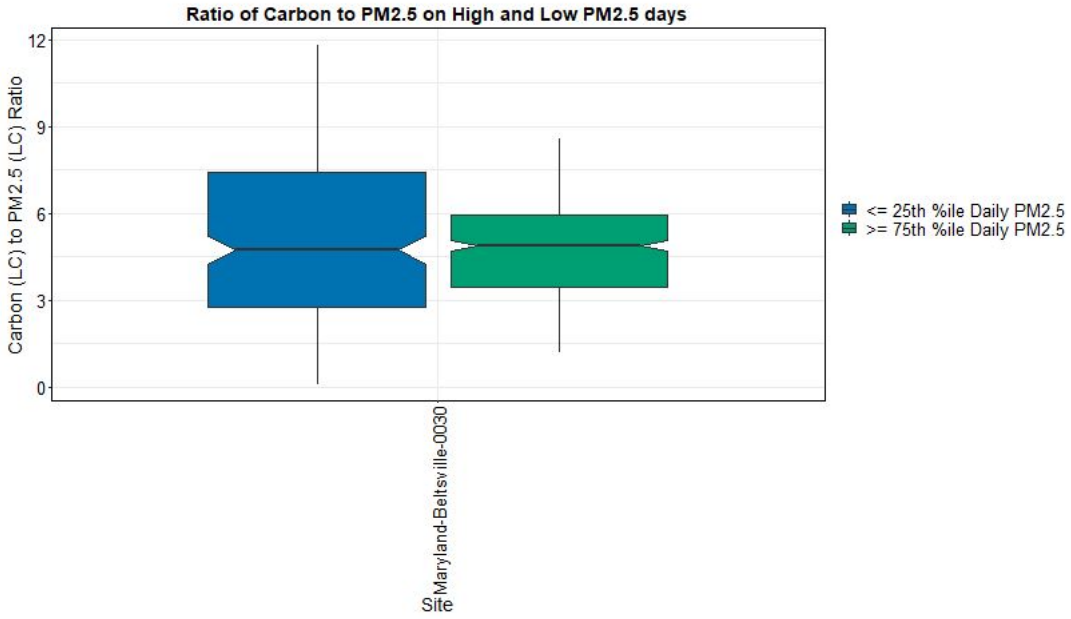


Figure 14: Boxplot showing the range of Carbon:PM2.5 ratios at a given site, separated by high and low PM2.5 days

The second set of plots shows a scatter plot for each box included in the boxplot above ((figure 15). A scatter plot is generated for the selected site, and any comparison sites added to the display. The plots on the left show low PM2.5 days, and the plots on the right show high PM2.5 days. Daily Average PM2.5 is shown on the x-axis against the daily carbon measurement on the right.

Note that the low PM2.5 day plot(s) may show stepwise increases in daily PM2.5 due to the combination of a relatively restricted scale (between 0 and the 25th percentile threshold) and the precision of collected measurements (to the 10ths place).

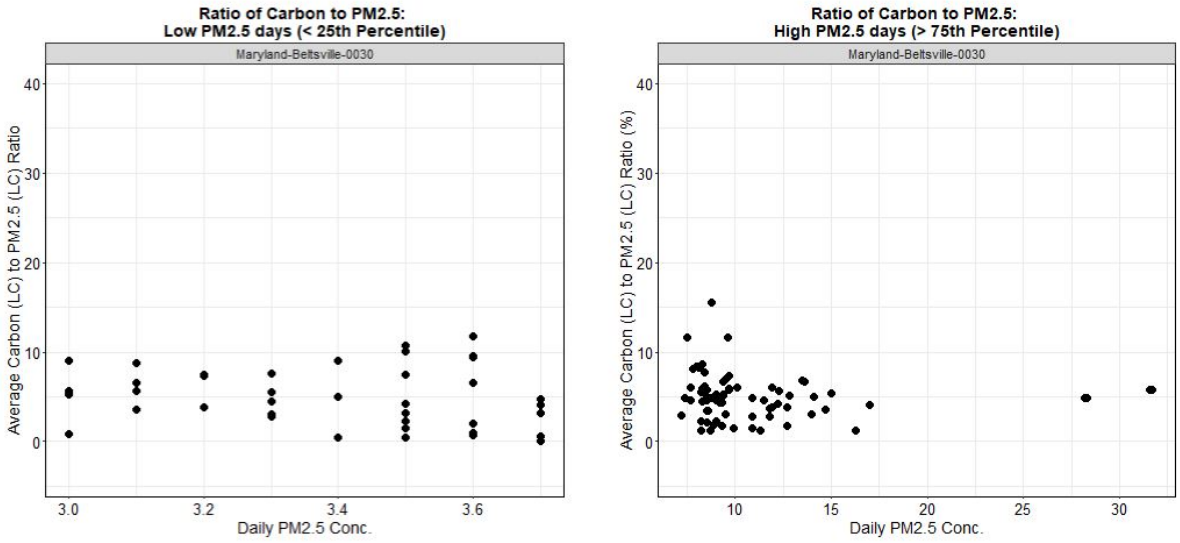


Figure 15: Seasonality plots showing variations in average concentrations by hour of the day, day of the week, and month of the year

4.6 Source Apportionment Tab

The Source Apportionment Tab estimates the daily average percent/concentration of total black carbon (BC) sourced by biomass burning and fossil fuel combustion. Biomass-burning-sourced black carbon is referred to with the subscript $_{bb}$, and fossil-fuel-sourced black carbon is referred to with the subscript $_{ff}$. **This tab only appears if a site reports concurrent observations of continuous carbon measured at local conditions**, i.e. the parameters listed below.

Tab-required Parameter codes:

- 88313 = Black Carbon PM2.5 at 880 nm
- 88314 = UV Carbon PM2.5 at 370 nm

Values for BC_{bb} and BC_{ff} shown on this tab represent **daily averages** of hourly measurements. The first plot shows a timeseries of BC_{bb}, including the estimated daily average concentration and percent of total BC sourced from biomass burning. The second plot shows the same metrics for fossil-fuel-sourced BC (BC_{ff}).

A user may use the date filter to subset the data shown in the plots and the table.

4.6.1 Methodology

The following equations are listed in Sections 5.1 and 11.4 of the Aethalometer Model AE33 user manual to calculate the fraction of total measured black carbon sourced by biomass burning (BB):

$$EQ0 : BC_{measured,\lambda} = \frac{b_{abs}(\lambda)}{\sigma_{air,\lambda}}$$

$$EQ1 : \frac{b_{abs}(\lambda_1)_{ff}}{b_{abs}(\lambda_2)_{ff}} = (\frac{370}{880})^{-\alpha_{ff}}$$

$$EQ2 : \frac{b_{abs}(\lambda_1)_{bb}}{b_{abs}(\lambda_2)_{bb}} = (\frac{370}{880})^{-\alpha_{bb}}$$

$$EQ3 : b_{abs}(\lambda_1) = b_{abs}(\lambda_1)_{ff} + b_{abs}(\lambda_1)_{bb}$$

$$EQ4 : b_{abs}(\lambda_2) = b_{abs}(\lambda_2)_{ff} + b_{abs}(\lambda_2)_{bb}$$
$$EQ5 : BB(\%) = \frac{b_{abs}(\lambda_2)_{bb}}{b_{abs}(\lambda_2)}$$

Constants/Derived Values:

- λ_1 - 370 nm (short wavelength measurement available from AQS)
- λ_2 - 880 nm (long wavelength measurement available from AQS)
- α_{ff}
- α_{bb}
- σ_{air}
- $b_{abs}(\lambda)$

See below sections for sources/derivations of each quantity.

Note: The AE33 manual lists 470 nm for short-wavelength (λ_1) and 950 nm for the long-wavelength (λ_2) measurements. We have substituted 370 nm and 880 nm respectively, as these are the wavelengths reported to AQS. This difference is accounted for in the values for α.

Determine absorption Ångström exponent (α) values

Table 3 in Zotter et.al (2016) indicates that reasonable absorption Ångström exponents (α) for fossil-fuel sourced BC (ff) and biomass sourced BC (bb) at 370nm and 880nm are:

- $\alpha_{ff} = 0.9$
- $\alpha_{bb} = 2.09$

Determine σ_{air} values

σ_{air} varies per wavelength:

Channel	Measurement wavelength (nm)	Mass absorption cross-section σ_{air} (m^2/g)
1	370	18.47
2	470	14.54
3	520	13.14
4	590	11.58
5	660	10.35
6	880	7.77
7	950	7.19

Section 5.1 of the Aethalometer Model AE33 user manual

Derive $b_{abs}(\lambda)$ values using *Equation* 0:

Rearrange EQ0:

$$b_{abs}(\lambda) = BC_{measured,\lambda} * \sigma_{air,\lambda}$$

Plug in relevant σ values:

$$b_{abs}(370) = BC_{measured,370} * 18.47 \; b_{abs}(880) = BC_{measured,880} * 7.77$$

Where $BC_{measured,370}$ is data in AQS measured under parameter code 88314, and $BC_{measured,880}$ is data in AQS measured under parameter code 88313.

Solve for $b_{abs}(880)_{bb}$ in terms of $b_{abs}(370)$, $b_{abs}(880)$, and constants:

Step 1: Re-configure Equation 2

$$b_{abs}(880)_{bb} = \frac{b_{abs}(370)_{bb}}{A_{bb}}$$

Step 2: sub eq 3 for $b_{abs}(370)_{bb}$

$$b_{abs}(880)_{bb} = \frac{b_{abs}(370) - b_{abs}(370)_{ff}}{A_{bb}}$$

Step 3: sub eq 1 for $b_{abs}(370)_{ff}$

$$b_{abs}(880)_{bb} = \frac{b_{abs}(370) - b_{abs}(880)_{ff} A_{ff}}{A_{bb}}$$

Step 4: sub eq 4 for $b_{abs}(880)_{ff}$

$$b_{abs}(880)_{bb} = \frac{b_{abs}(370) - (b_{abs}(880) - b_{abs}(880)_{bb}) A_{ff}}{A_{bb}}$$

Step 5: solve fore $b_{abs}(880)_{bb}$

$$b_{abs}(880)_{bb} = \frac{b_{abs}(370) - b_{abs}(880) A_{ff} + b_{abs}(880)_{bb} A_{ff}}{A_{bb}}$$

$$b_{abs}(880)_{bb} - \frac{b_{abs}(880)_{bb} A_{ff}}{A_{bb}} = \frac{b_{abs}(370) - b_{abs}(880) A_{ff}}{A_{bb}}$$

$$b_{abs}(880)_{bb}(1 - \frac{A_{ff}}{A_{bb}}) = \frac{b_{abs}(370) - b_{abs}(880) A_{ff}}{A_{bb}}$$

$$b_{abs}(880)_{bb} = \frac{b_{abs}(370) - b_{abs}(880) A_{ff}}{A_{bb}(1 - \frac{A_{ff}}{A_{bb}})}$$

$$b_{abs}(880)_{bb} = \frac{b_{abs}(370) - b_{abs}(880) A_{ff}}{A_{bb} - A_{ff}}$$

Note: Equation for $b_{abs}(880)_{bb}$ achieved above confirmed in Equation 5 of Xu et al. 2020

Final Equation for BB:

$$BB = \frac{b_{abs}(880)_{bb}}{b_{abs}(880)} = \frac{\frac{b_{abs}(370) - b_{abs}(880) A_{ff}}{A_{bb} - A_{ff}}}{b_{abs}(880)}$$

where BB is a decimal representing the fraction of total measured BC sourced by biomass burning.

Source Apportionment:

Once BB is calculated, concentrations of black carbon originating from each source can be estimated using the equations below from Section 11.4 of the Aethalometer Model AE33 user manual:

Fraction of total measured black carbon sourced by fossil fuel combustion:

$$FF = 1 - BB$$

Black Carbon sourced from biomass burning:

$$BC_{bb} = BB * BC$$

Black Carbon sourced from fossil fuel combustion:

$$BC_{ff} = FF * BC$$

Where BC is total measured Black Carbon as reported under parameter code 88313 (Black Carbon PM2.5 at 880 nm).