

Data analysis for 2021 IMPROVE sample year

Scott Copeland 6_8_2023

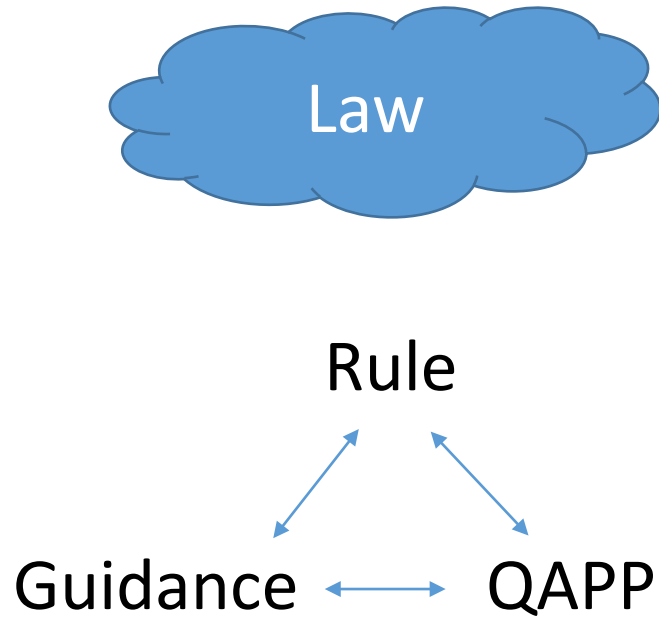
Data Analysis and Reporting Subcommittee – Jenny Hand, Chair

The principal products of the IMPROVE monitoring program are the speciated data and related visibility metrics. Routine and novel data analyses and products are regularly generated, such as the RHR metrics generated for each complete calendar year of data and the IMPROVE report. These data and all associated metadata and reports are made available through the IMPROVE (<https://vista.cira.colostate.edu/Improve/>) and FED (<https://views.cira.colostate.edu/fed/>) websites. Data are also available from EPA's AQS system (<https://www.epa.gov/aqs>)

This subcommittee is responsible for developing the policies for generating and distributing the IMPROVE data, metadata, and data products. It reviews and oversees all quality control assessments as defined in the QAPP and each year reviews irregular/suspect data and makes recommendations for its disposition. The subcommittee oversees the generation of the routine reports and data products and helps with novel IMPROVE data analyses and assessments. This is done in close collaboration with the communications subcommittee for the distribution of the IMPROVE data, metadata, data analyses and reports via the IMPROVE FED, and AQS websites.

- Oversee data availability to ensure timely submission to EPA AQS and CIRA FED;
- Produce and distribute RHR metrics via the FED and IMPROVE websites;
- Produce unique data analysis products, as needed; and
- Produce the IMPROVE Report on a schedule of every 5 years.

When data processing considerations



- Start from assumption that data is as reported.
- Make decisions consistent with sound science.
- What would stakeholders want?
- What are alternatives?
- What is the end data use?

Deliver notes below only affect previous year data at BLIS1 which has changed only site code to BLIS2 effective 1/1/2020. RHTS="BLIS_RHTS". All other changes are with respect to preliminary data from 2021.

Various	Various	Various	N/A	N/A	For various months and various analyses, field blank statistics used in the initial deliveries were invalidated and were recalculated for redelivery because individual filters changed after original delivery (e.g. due to reanalysis/reweigh, status change, swaps) or insufficient number of field blanks were available at the time of original processing.
BLIS1/2	1/21/2020 - 11/30/2021	All	N/A	N/A	This site moved to a temporary location at the beginning of 2020. However, the samplers are still at their temporary location so it was decided to make this a new site. The 2021 data has been updated accordingly to be associated with the new site. Data in AQS has been moved from being associated with BLIS1 monitors to being associated with BLIS2 monitors between the two dates. NPS has been contacted to determine how to handle the data.
Various	January 2021 - June 2021	A	XX	NM	A total of 71 filters were originally delivered with the elements Al, Cl, P, S, and Si invalidated with the XX status because of a suspected analysis issue. The source of the analysis issue is known and is suspected to impact more filters than the 71 filters identified during validation of 2021 data, however the impact on the reported data is thought to be minimal. For the redelivery 68 of the originally identified 71 filters had the XX status removed from the Cl, P, S, and Si elements, leaving Al as it was originally delivered. For three filters, all elements were re-validated as they were deemed to not be impacted by the issue. Investigations and analysis are ongoing.
All	June 2021 - December 2021	A	N/A	N/A	HIPS data (fAbs parameter) have been excluded from the skinny redelivery files between June 2021 and December 2021 while investigations are ongoing into suspect results. The data will likely be redelivered at a later date. Please do not use the fAbs data in the wide format files between June and December 2021.

- By request, group means files now include a new variable “Number_obs”, the number of valid observations used for each reported mean value.*
- Explicitly added site specific Rayleigh (ss_Rayleigh) values to group means files.
- No other changes to data or algorithm for this process.

Update for version “_12_22_2”:

- Four new sites were added to the Impairment Metrics files. The sites are FCPC1, KPBO1, SYCA2, and LTCC1.
- These sites have the requisite 5 years of complete data to calculate the e95 and 2064 end point values, but the algorithm was not properly including them.
- Years affected are

KPBO1	2016-2021
LTCC1	2015-2021
FCPC1	2017-2021
SYCA2	2016-2021

What do we need to calculate impairment?

- Episodic minimum 95th %ile carbon and dust.
- Defined in guidance over 2000-2014, but not defined for new sites.
 - Because of the “power” of e95, I chose *5 year* minimum required to calculate.
 - **The value is allowed to change over up to first 15 years.**
- Natural conditions “nc2” values for sites without a 2000-2004 “baseline” are from first 5 years of data at that site.
- 2064 endpoints use up to first 15 years of data from a site.
 - **The value will change over up to first 15 years.**

Regional Haze Metrics

- Impairment metrics here:

<https://drive.google.com/drive/folders/0Bxfj1vyyXeDYWVpfeUo4NEYtTU0?resourcekey=0-d0Bn5HHHEkghiHZvIQWaLQ&usp=sharing>

- Haziest Day metrics here:

<https://drive.google.com/drive/folders/0Bxfj1vyyXeDYTjNLellwUUx0TTg?resourcekey=0-yt0EY9maDLhRmcE8s5tC6A&usp=sharing>

- History of RHR metric changes since 10/2019.

https://docs.google.com/presentation/d/1H3OpvZ6LBVrNmQyB-2vqS7GTJx1Uzbd9/edit?usp=share_link&oid=116534812255078445612&rtpof=true&sd=true



Scott Copeland

