Best Practices for Review and Validation of Ambient Air Monitoring Data

This Page Intentionally Left Blank

EPA-454/B-21-007
August 2021

Best Practices for Review and Validation of Ambient Air Monitoring Data

U.S. Environmental Protection Agency
Office of Air Quality Planning and Standards
Air Quality Assessment Division
Research Triangle Park, NC

Disclaimer

The statements in this document, with the exception of referenced requirements, are intended solely as guidance. This document is not intended, nor can it be relied upon, to create any rights enforceable by any party in litigation with the United States. This guidance may be revised without public notice to reflect changes in EPA's approach to implementing 40 CFR Parts 50, 53, and 58.

Mention of commercial products or trade names should not be interpreted as endorsement. Some types of instruments currently in use may be described in text or in example figures or tables. Sometimes these products are given as a typical and perhaps well-known example of the general class of instruments. Other instruments in the class are available and may be fully acceptable.

Table of Contents

List of Figures

Acronyms and Abbreviations

AMTIC	Ambient Monitoring Technical Information Center
ADQ	audit of data quality
AQS	Air Quality System
ARM	Approved Regional Method
CFR	Code of Federal Regulations
COC	chain of custody
CV	coefficient of variation
DAS	data acquisition system
DASC	Data Assessment Statistical Calculator
DQA	data quality assessment
DQIs	data quality indicators
DQOs	data quality objectives
EPA	Environmental Protection Agency
FEM	federal equivalent method
FRM	federal reference method
IT	information technology
LDL	lower detectable limit
LIMS	laboratory information management systems
MDL	method detection limit
MQOs	measurement quality objectives
NAAQS	National Ambient Air Quality Standards
NCore	National Core Network
NIST	National Institute of Standards and Technology
NPAP	National Performance Audit Program
OAQPS	Office of Air Quality Planning and Standards
OGC	Office of General Counsel
ORD	Office of Research and Development
PE	performance evaluation
PEP	Performance Evaluation Program
ppb	parts per billion
ppm	parts per million
PQAO	primary quality assurance organization
QA	quality assurance
QA/QC	quality assurance/quality control
QAGD	Quality Assurance Guidance Document
QAM	quality assurance manager
QAO	quality assurance officer
QAPP	quality assurance project plan
QMP	quality management plan
SLAMS	state or local air monitoring stations
SLT	state, local or tribal
SOP	standard operating procedure

SPM	special purpose monitor
TSA	technical systems audit
US	United States
ZPS	zero, precision, span

Acknowledgements

In January 2018, an EPA Data Validation Workgroup formed with a goal to develop a tool that could be used to assist personnel in any state, local, or tribal (SLT) monitoring organization with performing data review and validation techniques. Workgroup members included EPA quality assurance (QA) and ambient air monitoring technical staff primarily responsible for conducting Technical Systems Audits (TSAs) and Audits of Data Quality (ADQs). The document was peer-reviewed by additional EPA QA staff in the Regional Offices and the Office of Air Quality Planning and Standards (OAQPS). The following EPA staff are acknowledged for their contributions:

Region 1: **Robert Judge, Catherine Taylor**
Region 3: **Verena Joerger**, Kia Long
Region 4: **Anthony (Tony) Bedel, Richard Guillot, Stephanie McCarthy**, Keith Harris
Region 5: **Chad McEvoy**, Saphique Thomas
Region 6: Kara Allen, David Anderson
Region 7: Leland Grooms
Region 8: Ethan Brown, Adam Eisele
Region 9: **Mathew Plate, Bilal Qazzaz**, Randy Chang, Dena Vallano, Roseanne Sakamoto
Region 10: Chris Hall
OAQPS: **Greg Noah, Trish Curran**
ORD: Robert Vanderpool
OGC: David Orlin, Sonja Rodman

Names that appear in bold are the members of the EPA Data Validation Workgroup.

Appendix A of this document contains data review checklists. These checklists were reviewed and field-tested by the following state and local ambient air monitoring organizations, and are acknowledged and thanked for their assistance: the Metro Public Health Department of Nashville/Davidson County, Tennessee; the Arizona Department of Environmental Quality; and the Wisconsin Department of Natural Resources.

Additionally, we are grateful for some of the figures contained in this document, which were provided by monitoring organizations.

Preface

Intent of Document

Data review is covered in Section 17 of the 2017 Environmental Protection Agency (EPA) *Quality Assurance Handbook for Air Pollution Measurement Systems, Volume II* (also referred to as the QA Handbook or Redbook)[1]. Data validation templates are also presented in Appendix D (updated March 2017)[2] of the referenced Handbook. Together, these provide guidance on data review concepts relevant to EPA's ambient air monitoring program and help users interpret and implement EPA requirements. This document is intended to supplement the existing guidance by providing a step-by-step process that air monitoring organizations can follow to validate ambient air monitoring data. This document is written to apply to monitoring of criteria pollutants in ambient air, but it may also be adapted for other air monitoring programs.

This document was created in response to requests from ambient air monitoring organizations for additional, formalized guidance to help them develop comprehensive and consistent data review programs. Quality Assurance (QA) and technical monitoring staff from the EPA Regional Offices and the Office of Air Quality Planning and Standards (OAQPS) met in Chicago in June 2017 and agreed the creation of such guidance was a priority. A workgroup to develop the guidance formed soon after that meeting. Much of the material in this document is available from other EPA guidance documents and trainings. This document is intended to consolidate and present "best practices" that, if implemented and followed, should result in a consistently validated, high quality dataset in the EPA's Air Quality System (AQS) database.

This document makes use of internet links that provide the user with access to more detailed information on a particular subject. Web links to references are included as footnotes for the reader to follow for additional information.

Document Review and Distribution

The information in this document was developed by the members of the EPA Data Validation Guidance Workgroup, representing EPA Headquarters and the EPA Regional Offices, and has been reviewed by the workgroup. The document has also been provided for review and comment to all EPA Regional Offices prior to distribution. This document has been signed and distributed by OAQPS QA staff to promote consistency across EPA and monitoring organizations in performing data review activities, including data validation. This document may be viewed on the internet and downloaded from the EPA Ambient Monitoring Technical Information Center (AMTIC) website.

Recommendations for improvement are welcome, and comments should be directed to the Data Validation Workgroup members identified in the Acknowledgements section in bold. This document will be reviewed at least every 5 years by the workgroup and revised as needed. The document may require more frequent revisions following significant rule changes and/or to keep pace with technological

[1] https://www3.epa.gov/ttn/amtic/files/ambient/pm25/qa/Final%20Handbook%20Document%201_17.pdf

[2] https://www3.epa.gov/ttn/amtic/files/ambient/pm25/qa/APP_D%20validation%20template%20version%2003_2017_for%20AMTIC%20Rev_1.pdf

advances in monitoring methodology. Appendices that contain data review checklists or examples of data review/coding scenarios may also require more frequent updates.

1.0 Introduction

Monitoring organizations are required to establish quality systems for their air monitoring programs. A quality system is the framework by which an organization applies sufficient quality control (QC) and quality assurance (QA) practices to ensure program results meet or exceed expectations. Figure 1 provides a basic illustration of the flow path and elements of an ambient air monitoring quality system. It is based upon a "Plan-Do-Check-Act" cyclical model that includes planning the work, implementing what is planned, assessing the results against performance criteria, reporting on data quality, and then making improvements if necessary. The figure shows data verification, validation, and data quality assessment as fundamental components of the system. Therefore, a vital element of any ambient air monitoring program is the establishment and implementation of a structured data review process, where data examination can be performed in a standardized, consistent manner.

The data review process is a multi-step, multi-layered process to ensure data has been recorded, transmitted, and processed correctly and meets the needs of the end data user. It is best performed as a tiered process, with different people and perspectives responsible for the different stages of data review. Data review incorporates various verification and validation techniques, which are important, distinct aspects of data management that require asking critical questions and using well-informed judgment to determine the quality of environmental data. In accordance with 40 CFR 58.16(c), ambient air quality monitoring data submitted to the EPA's Air Quality System (AQS) database must be **validated**; as such, data collected as part of the National Ambient Air Quality Monitoring Program must undergo a comprehensive data review process prior to AQS submittal. This data review process is performed by the monitoring organization and will be the focus of this document.

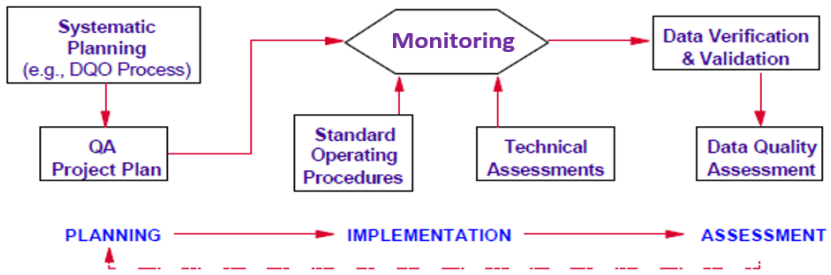


Figure 1: Elements of an Ambient Air Monitoring Quality System

The term "data validation" has been used synonymously with "data review" in some publications. However, this document will define and differentiate between the terms associated with data review, as they relate to an ambient air monitoring program. Towards that end, data validation means evaluating whether the data being gathered are useful for their intended purpose(s), i.e., the monitoring objective(s). Therefore, data validation includes evaluating whether the data meet specifications established in: (1) the Code of Federal Regulations, or CFR[3]; (2) the monitoring organization's Quality Assurance Project Plan (QAPP) and Standard Operating Procedures (SOPs); (3) the specific analytical method utilized; (4) the instrument's Federal Reference Method (FRM) or Federal Equivalent Method (FEM) designation; and (5) the Measurement Quality Objectives (MQOs) for the specific pollutant. Data validation examines the data collection records and supporting documentation to ensure compliance with these requirements can be demonstrated.

[3] https://www.ecfr.gov/cgi-bin/ECFR?page=browse

When reviewing data, it is important to recognize that all data has a "chain-of-custody" and is influenced by numerous personnel and processes. Figure 2 provides a generalized illustration of how data flows in an ambient air monitoring program. Data review and validation start at the monitor level, to confirm whether MQOs for individual pollutant monitors are achieved. The accuracy of values from individual monitors must be defensible. It is also critically important that the data under evaluation be compared to actual events, as described in the EPA document *Guidance on Environmental Data Verification and Validation* (*EPA QA/G-8*)[4]. After validation, data is entered into the EPA AQS database. From that point, assessments are performed. Assessments, as defined in *ANSI/ASQC-E4* and EPA's document, *Guidance on Technical Audits and Related Assessments for Environmental Data Operations (EPA QA/G-7)*[5], are evaluation processes used to measure the performance or effectiveness of a system and its elements. Assessment is an all-inclusive term used to denote any of the following: audit, performance evaluation, systems review, peer review, inspection, or surveillance. For the National Ambient Air Quality Monitoring Program, types of assessments include network reviews (i.e., annual and 5-year), performance evaluations (i.e., audits), technical system audits (TSAs), and data quality assessments (DQAs). For the purposes of this document, however, only data assessments, such as annual data certification, will be discussed.

The data review process utilized by the monitoring organization should be documented and performed using specified techniques to accept data, to reject data as invalid for a particular purpose, and/or to qualify, or "flag", data in a consistent and objective manner. 40 CFR 58.16(c) states that the procedures for editing and validating data are described in the AQS Data Coding Manual[6] and in each monitoring organization's QAPP. Therefore, the procedures, people involved, and frequency of data review must be fully explained in the monitoring organization's QAPP and relevant SOPs. It is important that data be

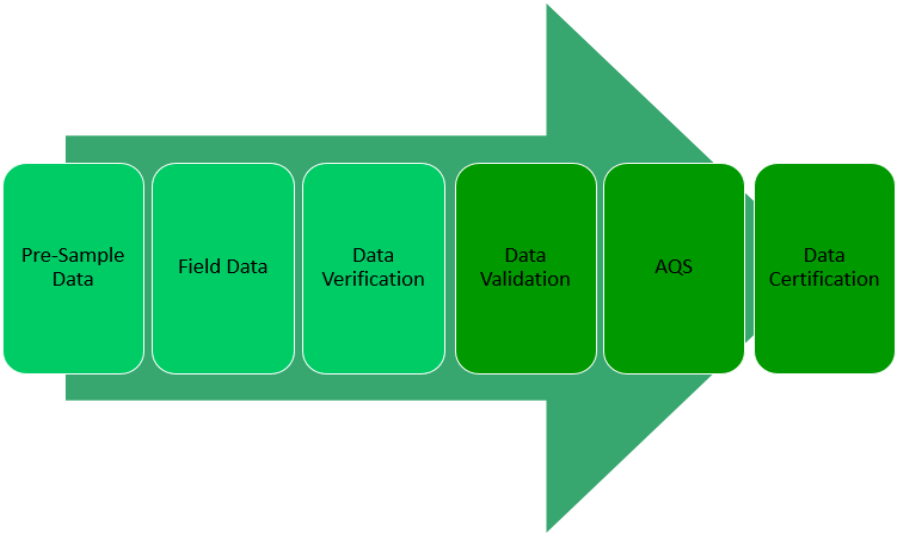


Figure 2: Generalized Ambient Air Monitoring Data Flow Path

[4] https://www.epa.gov/quality/guidance-environmental-data-verification-and-data-validation
[5] https://www.epa.gov/quality/guidance-technical-audits-and-related-assessments-environmental-data-operations-epa-qag-7
[6] https://www.epa.gov/aqs/aqs-manuals-and-guides

reviewed on a frequent, ongoing basis. Systematically reviewing smaller sets of data every few weeks or sooner helps identify problems early, before they can affect data completeness or significantly compromise real-time data reporting.

As seen in Figure 2, data validation is only one component of the data review process needed to ensure collected air monitoring data are of high quality and suitable for decision-making purposes. This document defines the various stages and levels of the data review process and presents EPA's recommended best practices for verifying and validating ambient air monitoring data. The document highlights technical ambient air monitoring requirements that must be examined during data review and provides the background and rationale as to why these are significant. The audience for this document includes QA and Air Program Managers, as well as the individuals who perform data verification and validation activities, including site operators (field technicians), data analysts, and QA staff. To help users of this document locate specific information, the document is structured as follows:

- Section 1.2 provides the basis for the data review requirements, including data quality regulations and supporting fundamentals.
- Section 2 provides insight into the resources necessary to build an effective data review program.
- Section 3 offers basic step-by-step instruction on verification and validation techniques.
- Section 4 provides a brief overview of assessments.
- Appendix A includes comprehensive data review checklists for verification and validation.
- Appendices B and C provide real-world monitoring examples that illustrate how to code data in AQS, as well as how to evaluate data validity based on weight of evidence.

1.1 Definitions

The following includes a list of significant terms that will be used in this document. Understanding these key terms is important for applying the concepts described herein. Definitions for additional terms commonly used in the National Ambient Air Quality Monitoring Program can be found in the monitoring regulations (see 40 CFR 50.1 and 40 CFR 58.1), as well as in other EPA guidance documents. It is important to note that some of the terms that follow, although commonly used, may be defined and applied differently in other programs and quality systems. This document defines these terms for use in the EPA Ambient Air Quality Monitoring Program.

- **Action (Warning) limit** is a percentage of the minimum and maximum values of a defined acceptance criterion that is allowed before an instrument calibration or other corrective action measure is warranted. Action limits should be defined in QAPPs/SOPs and set lower (i.e., more restrictive) than the control limits (i.e., MQOs). Corrective measures should be taken when an action limit is exceeded, in order to prevent data loss.

- **As-Found** is a term used to describe data recorded prior to an instrument adjustment being made or, if an adjustment has not been made, the conditions of an instrument upon receipt.

- **As-Left** is a term used to describe data recorded after an instrument adjustment has been made or, if an adjustment has not been made, the conditions of an instrument when all services have been completed.

- **Assessment,** also referred to as **data quality assessment (DQA),** is the process of evaluating the *aggregated* data set's ability to meet the intended objectives (i.e., data quality objectives, or DQOs). QA/QC data can be statistically assessed at various levels of aggregation to determine whether the DQOs have been attained. Assessments are performed on *validated* data. Ultimately, DQAs determine how well validated data can support their intended use.

- **Best practice** is a procedure that is accepted as being the most correct based on widely accepted scientific practices and/or experience throughout the ambient air monitoring community.

- **Chain-of-custody (COC)** is defined as an unbroken trail of accountability that ensures the physical security of samples, data, and records. It is a legal term that refers to activities guaranteeing that no tampering has occurred for measurements or data, at any point in the process of measuring, recording, transferring, and reporting the results. It is vital that measurements, especially when comparing to standards, have records necessary for completely verifying the integrity of the data.

- **Compelling evidence** (reason) is data that concretely establishes instrument performance or validity of a QA/QC check. It includes, but is not limited to, data generated from independent audit point(s), multi-point verifications, and/or a prior zero/span check. This data establishes whether the analyzer was operating within its acceptance limits. It also indicates whether a QC check itself is considered valid or invalid.

- **Control limit** is the maximum value (threshold) for which a defined acceptance criterion is considered acceptable, and above which associated data are considered "out of control". During data review, specifications in the data validation templates (i.e., MQO tables) are considered control limits.

- **Data Quality Objectives (DQOs)** are qualitative and quantitative statements derived from the systematic planning process (see Figure 1) that clarify the purpose of the study, define the most appropriate type of information to collect, determine the most appropriate conditions from which to collect that information, and specify tolerable levels of potential decision errors. In short, they are the specifications needed to determine the type, quantity, and quality of data needed to make defensible decisions or to make creditable estimates with an acceptable level of certainty. DQOs provide a goal on which to build a quality system. The qualitative DQOs for the Ambient Air Quality Monitoring Program are identified in 40 CFR Part 58. The quantitative DQOs for the criteria pollutants are specified in 40 CFR Part 58, Appendix A, Section 2.3.1. (See Figure 5 for a comparison of DQOs, DQIs, and MQOs.)

- **Data Quality Indicators (DQIs)** are quantitative and qualitative attributes associated with data. DQIs include representativeness, comparability, sensitivity (i.e., detection limit), precision, bias, and completeness. (See Figure 5.)

- **Data review** is the examination of data; a multi-step, multi-layered process to ensure data has been recorded, transmitted, and processed correctly and ultimately meets the needs of data users. Data review incorporates various verification and validation techniques which are used to accept, reject, or qualify data in an objective and consistent manner.

- **Data validation** is a data review technique designed to ensure that reported values meet the quality goals of the environmental data operation, in this case ambient air monitoring operations. It can be further defined as the examination, through the provision of objective evidence, that the particular requirements for a *specific intended use* (i.e., monitoring objectives) are fulfilled (see Figures 3 and 4). Validation includes the evaluation of data for compliance with specified QC requirements, such as whether the acceptance limits for various performance specifications were achieved.

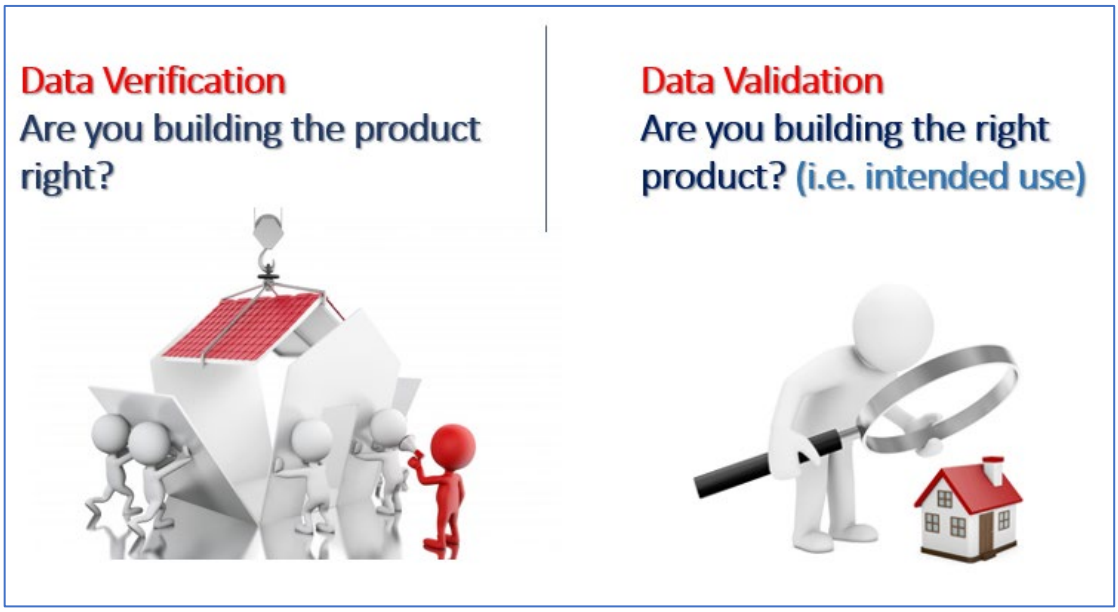


Figure 3: General Illustration to Distinguish Between Verification and Validation

- **Data verification** is a process of comparing how the data were gathered to the data collection plan (QAPP/SOPs). It is a data review technique that evaluates the completeness, correctness, and conformance of data against method, procedural, and/or contractual specifications. It can be further defined as the confirmation, through the provision of objective evidence, that specific requirements have been fulfilled. Verification usually consists of checking that SOPs were followed and QC activities were performed. (See Figures 3 and 4.)

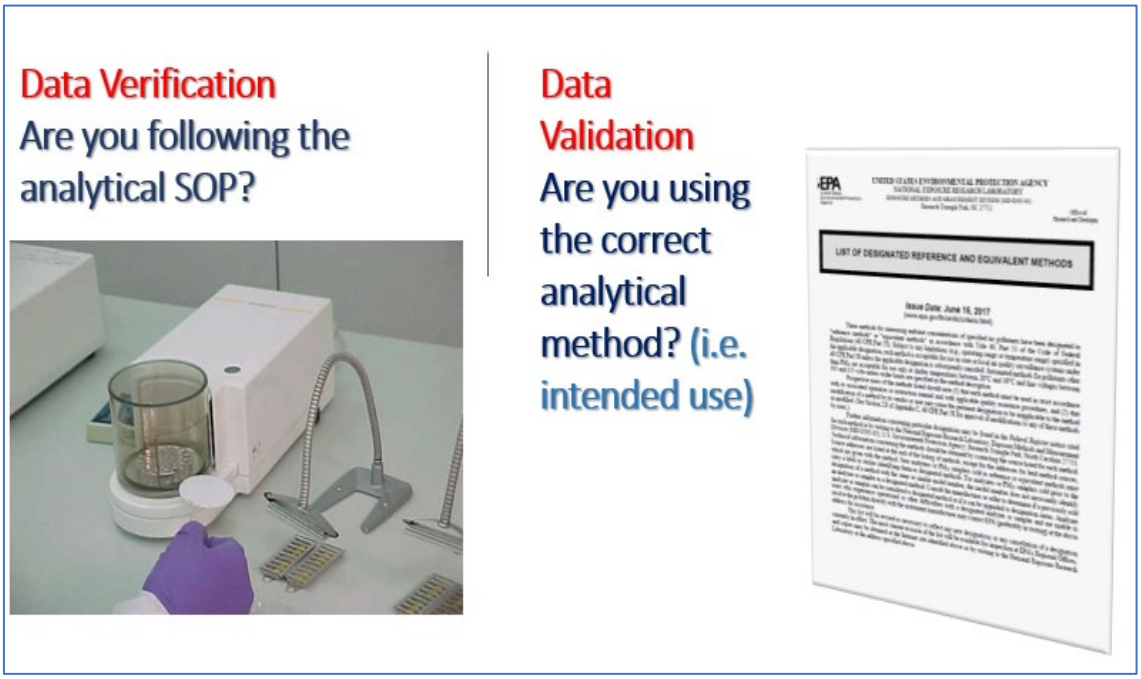


Figure 4: Air Monitoring-Specific Illustration to Distinguish Between Verification and Validation

- **Informational codes** are types of AQS-qualifiers used to alert users to data that may have been impacted by exceptional events (or other unique situations). Like QA qualifier codes, these codes do not invalidate data, but rather provide a means to tell a more complete story about the events which may have impacted the data. Examples of how and when to apply specific informational codes should be prescribed in QAPPs/SOPs.

- **Integrity**. As defined in the EPA Information Quality Guidelines (IQG), integrity refers to security, such as the protection of information (data) from unauthorized access or revision, to ensure that the information is not compromised through corruption or falsification. Therefore, an important element of data review is to evaluate the integrity of the collected data.

- **Measurement Quality Objectives (MQOs)** are designed to evaluate and control various phases (e.g., sampling, transportation, preparation, and analysis) of the *measurement process* (i.e., measurement/instrument level) to ensure that total measurement uncertainty is within the range prescribed by the DQOs. MQOs can be defined in terms of the DQIs. MQOs serve as control limits in the data review process. (See Figure 5.)

- **NIST-traceability** (see Traceability, below). National Institute of Standards and Technology (NIST)-traceability of field instruments is verified with documentation (i.e., calibration certificates) that demonstrates comparison against a NIST standard, directly or indirectly. NIST is the US authority on metric quantities, for commerce and research. All ambient monitoring measurements should be traceable to NIST.

- **Null codes**, also referred to as null qualifiers, are alphanumeric codes used within the AQS database to invalidate data. They are also required when submitting a null (i.e., nothing was collected) sample measurement. Based on the descriptions in AQS, null codes should be used to inform data users as to why valid data are not available, to the extent possible. Examples of how and when to apply specific null codes should be prescribed in QAPPs/SOPs.

- **Primary Quality Assurance Organization (PQAO)**. Established in 40 CFR Part 58, Appendix A, Section 1.2, a PQAO is defined as a monitoring organization or a group of monitoring organizations that is responsible for a set of stations that monitors the same pollutant and for which DQAs will be pooled. PQAOs are defined such that measurement uncertainty among all stations in the organization can be expected to be reasonably homogeneous as a result of common factors, which are explained within the regulation. Since DQAs are made and data certified at the PQAO level, the monitoring organization identified as the PQAO will be responsible for the oversight of the quality of data of all monitoring organizations within the PQAO.

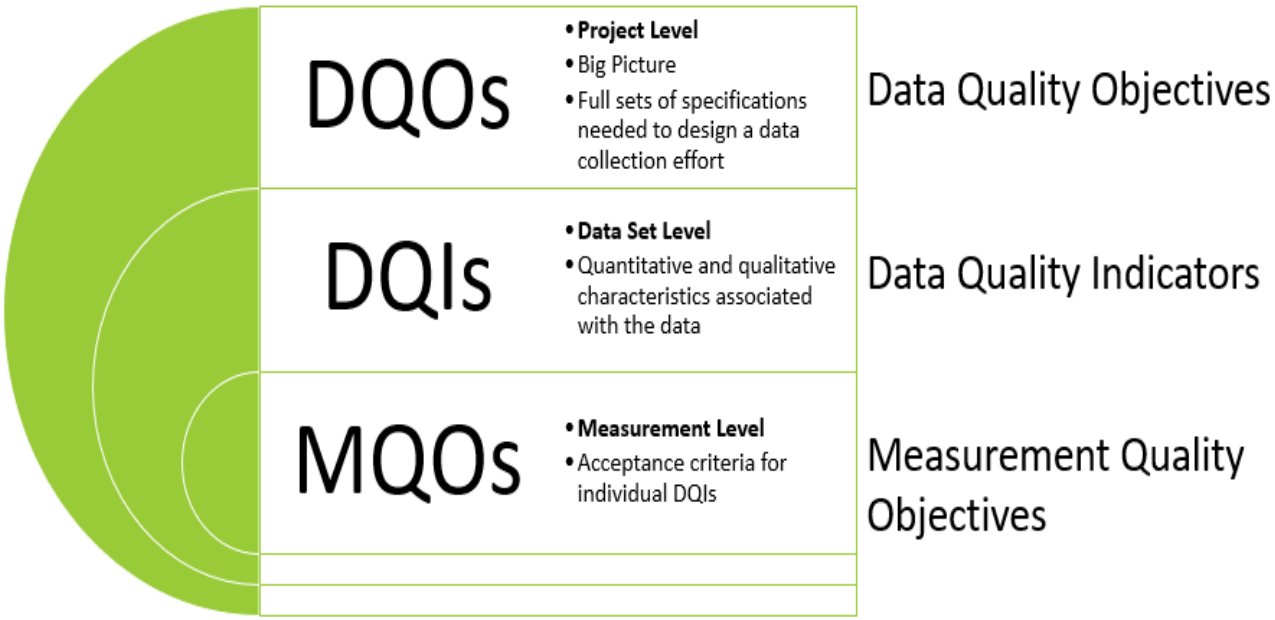


Figure 5: Comparison of DQOs, DQIs, and MQOs

- **Quality Assurance (QA)** is a series of management activities, including planning, implementation, and assessment, necessary to ensure the quality and defensibility of the final product (e.g., data). Examples of QA activities include developing QAPPs and SOPs.

- **QA Qualifier Codes** are used when data are valid, but additional commentary is needed in the AQS database to support and explain the validity decision. As its name suggests, QA qualifier codes qualify data, alerting users that specific QA/QC issues were identified with the flagged data. QA qualifier codes are alphanumeric codes in the AQS database. Examples of how and when to apply specific QA qualifier codes should be prescribed in QAPPs/SOPs.

- **Quality Control (QC)** is the system of technical activities conducted to measure the attributes and performance of a process against defined standards. QC provides a reasonable level of checking (verification) at various stages of the data collection process to ensure quality is maintained. Examples of QC activities include calibrations and precision checks. Although sometimes used synonymously with QA, QA and QC are significantly different concepts.

- **Reconciliation** is the evaluation of the aggregated data set's *and* the specified objective's ability to meet the users' needs. It may also include a re-evaluation of the users' needs. Reconciliation represents the completion of the quality cycle. It is a process by which data quality improvement is considered and recommendations are made for data quality planning and updates to data quality objectives.

- **Traceability** is the property of a measurement result whereby the result can be related to a stated reference through a documented unbroken chain of calibrations/comparisons, each contributing to measurement uncertainty. Traceability also refers to the ability to verify the history, location, or application of an item (or air monitoring calibration standard, for example), by means of documented recorded identification.

1.2 Guiding Principles

Two primary objectives of the Ambient Air Quality Monitoring Program's quality system are to produce credible data and support sound, defensible decisions. Monitoring organizations and EPA work together to achieve these end goals, sharing core principles that guide data quality decision-making processes. This section highlights the policies and premises that provide that foundation. Many monitoring organizations receive assistance grants from EPA, and as a result, must adhere to EPA's quality system requirements. Although monitoring organization staff do not have to be fluent in the quality policies and national/international standards that have been used to build EPA's quality system – and ultimately drive EPA's QA/QC recommendations for the Ambient Air Quality Monitoring Program – a core understanding of these policies and requirements will help managers and data reviewers make well-informed validity decisions. With this in mind, **this section is written primarily to assist QA and Air Program Managers in understanding these principles of establishing data quality.** Ideally, monitoring organizations' Level 3 data reviewers (see Section 3) should also be knowledgeable of these principles, including where the requirements originate. Section 3 of this document will offer specific instructions on how to verify and validate data in a manner that incorporates these fundamentals.

EPA encourages monitoring organizations to train staff – site operators and data reviewers – on these fundamental data quality principles. Ultimately, monitoring organization staff responsible for performing any level of data review should be fluent in the "part" of the process for which they are responsible and understand how errors identified during their review impact the process as a "whole". Referenced links within this section could be added to training plans as required reading, at a minimum. The APTI SI-470 course[7] also offers modules on EPA policy that discuss these standards and guiding principles as they apply to monitoring networks and data review procedures.

[7] https://www.apti-learn.net/LMS/EPAHomePage.aspx

1.2.1 Standards for Data Usability

EPA's Ambient Air Quality Monitoring Program for criteria pollutants produces data that is intended to be used by EPA, States, Locals, and Tribes for policy and regulatory decisions. Data must meet specifications dictated by the EPA quality system, the Clean Air Act, the Information Quality Act (IQA), and associated regulations and guidance. In addition, good laboratory practices and scientific protocols for data collection and handling should be followed to ensure the integrity of decision-making processes.

Since the success of the ambient air quality monitoring program's objectives rely heavily on the data and their interpretation, it is critical that the data available to users be reliable, of known and discernible quality, and aggregated in a manner that is acceptable for its primary use. In order to accomplish this activity, data must be collected and handled in a consistent manner that protects and ensures its integrity. Hence, the data review process should be designed to verify that these essential elements of data quality are in place and that the data set is usable for its intended purpose.

Standards for data usability are included below to help data reviewers understand where requirements and guidelines originally came from. It is also important to note that a discussion on data usability – establishing how the monitoring organization will consider and evaluate data's "fitness for use" – is a required element in the monitoring organization's QAPP.

To produce high quality, usable environmental information, data should:

- **Meet regulatory requirements**
  - Follow monitoring methods defined in regulation, including EPA FRM/FEM specifications
  - Follow procedures detailed in EPA's quality system and approved QAPPs
- **Be technically sound**
  - Consistent with validated methods and accepted standards of quality
  - Supported by measurements that include standard materials that are traceable to an authoritative source (NIST or equivalent), and calibrations checked by a second, independent standard to verify the integrity of the standardization process
  - Systematically reviewed to verify and validate data usability against program objectives
- **Be defensible**
  - Ensure all data collection steps are documented and this documentation and associated raw data are retained and NIST-traceable
  - Ensure data integrity and reliability
  - Maintain physical chain-of-custody (COC)
  - Ensure unethical practices are not occurring and are actively prevented.

The sections that follow will provide brief summaries of the regulatory requirements, technical requirements, and defensibility elements that should be examined during data review to ensure ambient air monitoring data is usable – accurate, reliable, and legally sound. These standards should support and inform data quality decisions made using a weight of evidence approach.

1.2.1.1 Regulatory Requirements

There is strong precedent in EPA's ambient air monitoring program to only use data for EPA National Ambient Air Quality Standards (NAAQS) decisions that meets requirements established in regulation. Consideration to accept data should be made based on evaluating compelling evidence. EPA must also consider adherence to the IQA, where influential information is held to a higher standard of quality/transparency. The Office of Management and Budget (OMB) notes that information is *influential* if:

> *"…the agency can reasonably determine that dissemination of the information will have or does have a clear and substantial impact on important public policies or important private sector decisions."*[8]

Considering this, non-compliance with regulatory requirements and uncertainty about data quality, integrity, and defensibility, usually result in the inability of EPA to use data that do not meet regulatory requirements to determine compliance with the NAAQS.

Therefore, when discussing standards of data usability, the first significant consideration is that the data meet regulatory requirements. Data quality regulations and policy will be summarized first, to provide background and clarity on EPA's quality system requirements. At the highest level, these standards and regulations determine (or set) what level of QA is required for the monitoring program and, therefore, set the stage for program and project-specific guidance from EPA. Ambient air monitoring regulations will be summarized afterwards.

1.2.1.1.1 Data Quality Regulations and Policy

When EPA develops its QA policy, it considers adopting national consensus standards similar to the American National Standard Institute's (ANSI) standards or the standards developed by the International Organization for Standardization (ISO). Monitoring organizations that might already be complying with these national or international standards will likely find it easier to comply with EPA policies. Ultimately, it is EPA policy (see EPA Order 2105.1)[9] that all environmental programs performed by EPA, or through EPA-funded extramural agreements (e.g., state and local assistance grants), shall be supported by individual quality systems that comply with the 2014 American National Standard *Specifications and Guidelines for Quality Systems for Environmental Data Collection and Environmental Technology Programs* (ANSI/ASQC E4-2014). Data quality is also governed by the IQA and related Information Quality Guidelines (IQG)[10]. The IQG requires that information supporting EPA decisions meet EPA quality requirements and be documented and transparent to the public and the regulated community.

QA regulations for environmental data, collected/used under grants and agreements, are found in 40 CFR Part 35 and 2 CFR Section 1500.12; and QA regulations for data collected under EPA contracts are found in 48 CFR Part 46. The QA requirements are also reiterated and clarified in EPA Environmental

[8] Section 6.2 (Page 19) at https://www.epa.gov/sites/production/files/2020-02/documents/epa-info-quality-guidelines_pdf_version.pdf
[9] https://www.epa.gov/irmpoli8/environmental-information-quality-policy
[10] https://www.epa.gov/quality/epa-information-quality-guidelines

Information Quality Policy CIO 2105.1. These stress that information shall be generated from documented quality systems that follow national and international standards for quality.

EPA data quality requirements for quality systems are also documented in the policy documents for Quality Management Plans (QMPs)[11] and QAPPs[12]. The requirement for approved QMPs and QAPPs for the ambient air monitoring program is also reiterated in 40 CFR Part 58, Appendix A. As such, these documents and their contents, and by extension any associated SOPs, reflect regulatory requirements. Where QMPs and QAPPs have not been developed, are not approved, or are outdated, the quality of the data collected may not be appropriate for EPA decisions. Where there are inconsistencies between QAPP/SOP quality commitments and explicit regulatory requirements, data should be reviewed based on the regulatory requirements and in consultation with EPA.

1.2.1.1.2 Ambient Air Monitoring Regulations

While the Clean Air Act[13] contains language pertaining to air monitoring data quality, the regulations pertaining to ambient air monitoring are found in 40 CFR Parts 50, 53, and 58. EPA and monitoring organizations reference and utilize these specific regulations most frequently. The following summarizes the regulations.

- 40 CFR Part 50 Appendices: Reference methods for collection and analysis of criteria pollutant data. Each of these methods define operational and calibration approaches that must be followed to meet FRM requirements.

- 40 CFR Part 53: Analytical method and instrument validation requirements. Procedures for establishing both FRMs and FEMs are defined here. It is important to note that Part 53 is most relevant to instrument vendors (applicants) and the EPA Office of Research and Development (ORD) staff who review the applications for those candidate methods.

- 40 CFR Part 58: The general requirements for ambient air monitoring. The monitoring network operation and design elements are included in the main text of Part 58 and in Appendix D. **For the purposes of data review/validation, most of the applicable requirements in Part 58 are presented in the QA system requirements, Appendix A, and in the siting and probe design requirements, Appendix E.**

- 40 CFR Part 58 Appendix A: General quality system requirements, including establishing a PQAO with independent QA and defining QMPs and QAPPs as required documents. EPA and PQAOs are instructed to use a weight of evidence approach when evaluating data quality; however, the final evaluation of data applicability for regulatory decisions is reserved for EPA. Appendix A also defines specific, minimum QA/QC checks that must be implemented as part of the ambient air monitoring quality system.

- 40 CFR Part 58 Appendix E: Monitoring probe placement, obstructions, trees, roadway distance, probe material, and residence time. It is assumed that most of these requirements will be met prior

[11] https://www.epa.gov/quality/epa-qar-2-epa-requirements-quality-management-plans
[12] https://www.epa.gov/quality/epa-qar-5-epa-requirements-quality-assurance-project-plans
[13] https://www.epa.gov/clean-air-act-overview

to the initiation of monitoring. However, this may not be the case and/or circumstances will change over time, requiring periodic reviews and potential data actions, such as qualification of impacted data using AQS QA qualifier codes specific to siting issues. There is also a provision for EPA to waive these requirements in rare, limited circumstances.

In some instances, regulations may reference guidance documents, consensus standards, or methods that must be followed. When this occurs, these documents are considered an extension of the regulation.

Note that QA/QC tasks may not have objective acceptance criteria published in regulation. In these cases, the data reviewer should view the conducting of the task as required. For example, 40 CFR Part 58, Appendix A, Section 3.1.2 states:

A performance evaluation must be conducted on each primary monitor once a year.

No criteria for evaluating the results of the audits are presented in this citation. Section 3.1.2.1 to Appendix A elaborates further about the number of concentration points that must be conducted during a performance evaluation, along with a range for the concentrations of each audit point – but again, no acceptance criteria for the audit results are provided. With this in mind, the absence of specific audit result acceptance criteria in the regulation does not mean such evaluations are less critical or do not need to be performed. Instead, the reviewer should interpret the audit frequency and audit concentrations specified in regulation as required and evaluate data for conformance with these requirements.

It is important to note that data collected that meets established quality objectives for public notification or research, but does not meet regulatory requirements, may be used and reported as such. There is an expectation that this data will be reported to EPA in a manner that excludes it from NAAQS decision-making. This could include reporting data to AQS with appropriate qualification (such as designating the data as non-regulatory) or not reporting data to AQS. If the latter, data that is not reported to AQS should be shared with EPA in an alternative format such as a report and/or direct data deliverable. Monitoring organizations are cautioned that reporting data that does not meet EPA quality standards to AQS without appropriate qualification may lead to erroneous NAAQS decisions, which could result in significant consequent actions.

1.2.1.2 Technical Expectations

When discussing standards of data usability, the second significant consideration is that data should be scientifically and technically sound. Towards that end, there are expectations inherent to collection of sound environmental analytical data which extend to environmental samples collected in the field. These expectations are reflected in various EPA guidance documents, but generally relate to addressing the data quality indicators of precision, bias, representativeness, comparability, and sensitivity. As a result, many of these technical requirements are addressed in the ambient air monitoring regulations. However, several topics are not fully addressed in regulation and are essential for ensuring that data sets are technically adequate. These include standardization/traceability and data review/validation, the subject matter of this document.

While substantially addressed by ambient air monitoring regulation, the accepted scientific practices – method validation, traceability, calibration, evaluation of uncertainty/accuracy (bias and precision), preservation, sensitivity, and demonstration of proficiency – should be addressed for all data. Therefore,

ambient air monitoring data review should include evaluations of the technical acceptability of each data set, based on these scientific principles. **Where there are indications that a data set does not meet basic technical requirements, but meets regulatory requirements, appropriate data actions should be taken.** For example, consider a scenario where an instrument reports an abnormally constant low concentration of a pollutant in the environment, but passes all quality assurance checks, calibrations, and verifications. A subsequent in-depth review of the data determines the concentration reported is greater than the instrument detection limit, but this concentration is not expected in the environment. Further investigation shows the sensitivity of the instrument is impaired. This latter determination is further supported by an elevated zero reading during a routine audit. Consequently, the impacted data is concluded to be invalid (despite meeting regulatory requirements). In this scenario, the weight of evidence indicates the data was not technically sound – because the instrument's sensitivity was impaired. Weight of evidence considers *overall compliance* with Part 58 and will be discussed in more detail in Section 2.2.1.3 of this document.

1.2.1.2.1 FRM/FEM Requirements

The process for establishing FRMs and FEMs is defined in 40 CFR Part 53. Instruments and analytical methods must be reviewed by ORD for compliance with requirements in 40 CFR Part 53 to demonstrate their ability to meet FRM or FEM status and to produce data which are comparable to the Federal Reference Method, as defined in regulation. The candidate method testing may result in a unique set of hardware configurations, software configurations, instrument/method specific QA/QC, environmental conditions, and/or operational settings used to achieve FRM/FEM status. These requirements are summarized in published designation specifications[14], once FRM/FEM status is granted. In some cases, method updates may result in the need for instrument operational parameters or manuals to be changed, and these changes would need to be reflected in existing instruments, prior to their deployment in the NAAQS network. Where these parameters are deemed necessary in the FRM/FEM demonstration, they must be carried into the routine operations of these methods. Additionally, data quality criteria must be established for FRM/FEM methods where they differ from or supplement regulation and guidance.[15]

40 CFR Part 58, Appendix C, Section 2.1 states that criteria pollutant monitoring methods used for making NAAQS decisions must be a reference (FRM) or equivalent method (FEM) as defined in 40 CFR 50.1. However, CFR does not include specific requirements or QA/QC acceptance criteria for individual makes/models of instrumentation based upon their designation status. Instead, these criteria are typically included in instrument user manuals or other guidance. In order for data produced by the instrument to be technically sound, the instrument must be operated in accordance with its FRM/FEM specifications and user manual requirements. As part of the data review process, then, where these criteria are deemed a necessary part of the FRM/FEM demonstration, they should be interpreted as critical criteria. Deviations from FRM/FEM operational parameters or criteria, and/or method changes, should be approved by EPA and reflected in QAPPs and SOPs.

[14] https://www.epa.gov/sites/production/files/2019-08/documents/designated_reference_and-equivalent_methods.pdf
[15] See 40 CFR 53.4

1.2.1.2.2 Traceable Measurements

The ambient air monitoring regulations specify that NIST-traceable standards be used for certain measurements. Traceability of standards is not defined or specified, however, for other measurements (such as in the gravimetric laboratory requirements for $PM_{2.5}$ analysis). To perform any field or laboratory operation that produces scientifically and technically-sound results, the best practice is to utilize accurate, traceable standards.

The National Environmental Laboratory Accreditation Conference (NELAC) Quality Systems Standard (co-published by EPA) includes EPA's guidance for measurements. It states:

> *All equipment used for environmental tests, including equipment for subsidiary measurements (e.g. for environmental conditions) having a significant effect on accuracy or validity of the result of the environmental test or sampling shall be calibrated before being put into service and on a continuing basis.*[16]

These calibrations must be referenced to national and/or international standards or reference material. Where no standard is available, an adequate alternative must be approved by EPA through guidance and/or in an organization's QAPP.

Technical requirements for traceability apply to all parameters that support a measurement. For a gaseous pollutant, for example, this would include: the calibration gas, the dilution gas (zero air), the flow sensors, mass flow controllers, temperature sensors (including those that monitor environmental/shelter conditions), and potentially pressure sensors. Similarly, for particulates, this would include flow rate standards and support equipment (thermometers, barometers, manometers), and for the laboratory, devices such as temperature and humidity devices, mass reference standards (i.e., check weights), and the microbalance. Records should be available to support the traceable standards, and subsequently, to support the traceability of the resulting data. The impact on data quality for having missing or expired traceability will vary depending on the standard's purpose in supporting monitoring. Expired primary standards used to calibrate an instrument could lead to data being unusable for technical decisions; however, this may be mitigated if the instrument calibration was verified with a non-expired secondary source standard.

1.2.1.3 Defensibility

When discussing standards of data usability, the third significant consideration is that the data be defensible, especially if the data is intended to be usable for NAAQS decision-making. To be defensible, this means the data include: complete and traceable QA/QC documentation (e.g., NIST-traceable calibrations, one-point QC checks, and performance evaluations); complete COC (physical sample handling COC, as well as data handling COC); and are consistent with commitments made in grant conditions and the grant workplan, which could include demonstrations of competence[17]. Documentation is a key component of defensibility.

[16] https://nelac-institute.org/content/CSDP/standards.php
[17] https://www.epa.gov/sites/production/files/2015-03/documents/competency-policy-aaia-new.pdf

1.2.1.3.1 Documentation

There should be documentation available to support decisions made at the monitoring organization level regarding the validity of data. Logbooks, data forms, and other records must be maintained in order to justify data qualification (flagging) or invalidation. Similarly, these records must be available to support that data are valid.

Review, verification, and validation require that sufficient documentation has been collected and maintained with integrity, reliability, and defensibility. This applies to electronic records, non-electronic records, and physical samples. Where records do not exist or have not been properly produced and/or maintained, data may not be suitable for its specific intended use. Documentation can be electronic or hard-copy, and both types of records need to meet the basic requirements, including; secure storage, limited access, uniquely identified authors, all entries include date and time, and original entries are retained (not erased or discarded when revised).[18]

Documentation is critical to ensure the integrity of data sets; where data or critical activities are not documented, or documentation is not retained, the adequacy of the data cannot be verified. In some instances, lack of documentation will preclude the use of data sets for decision making. Where documentation is not complete, other lines of evidence, including raw data and information provided by the instrument technician, should be used to supplement the review and to determine if there is sufficient weight of evidence to verify that QC checks were valid, and meet all regulatory, QAPP, and SOP requirements. **If there is insufficient evidence to show that a QC activity was performed, the data should be treated as if the activity was not conducted**. A formal corrective action process should also be initiated to prevent future documentation deficiencies. Where documentation is consistently incomplete, the integrity of the data set should be evaluated, and possible data quality actions may need to be taken; if the latter occurs, the EPA Regional Office should also be informed.

Corrections to documentation prior to or during the data review process should be made using a process detailed in the *NEIC Policies and Procedures Manual*[19]. Per the manual, "Any subsequent error discovered should be corrected by the person who made the entry, the person who discovered the error, or another person familiar with the work. All subsequent corrections must be initialed and dated." For electronic records, an equivalent process, that retains and corrects the original entry, should be used. For more information, please see Appendix J of the QA Handbook (2017) and EPA's Cross-Media Electronic Reporting Rule (CROMERR)[20]. In some cases, a review may identify conflicts in documentation and/or technicians' recollections. Often there are conflicts between procedures in planning documents and procedures "as documented" during data collection. As with missing documentation, weight of evidence should be used to resolve conflicts and, subsequently, corrective actions should be initiated to prevent further conflicts and/or improve documentation.

Additionally, sufficient raw instrument data must be collected and maintained to support data review and document the data set. For instruments where hourly completeness is paramount, sub-hourly data (i.e., minute data)[21] is also important and should be retained and reviewed. It is further recommended that

[18] QA Handbook, Appendix J*, Guidance on the Use of Electronic Logbooks (2017)*
[19] https://nepis.epa.gov/Exe/ZyPDF.cgi?Dockey=9101JOP2.PDF
[20] https://www.epa.gov/cromerr
[21]EPA QA Handbook (2017), Sections 6.4.1, 10.4, and 14.2

instrument meta-data, including operational parameters such as flow, pressure, and temperatures, be collected and maintained to aid in the validation process. This information is often important in identifying instrument malfunctions and its evaluation improves the overall quality of the data reported.

1.2.1.3.2 Custody

Chain-of-custody procedures are required to maintain the integrity of sample collection. In NBSIR 85-3105 (NIST), *Principles of Quality Assurance of Chemical Measurements*, it is noted that:

> *The concept of "chain-of-custody" most often is viewed as a means for legal validation of samples, but its use for quality assurance is equally if not more important. An adequate system provides both assurance of identification of the samples that are analyzed and that all aspects of quality control required for them have been observed.*[22]

For each sample, integrity and preservation should be maintained from the time of sampling to the time of analysis and disposal. For sampling media that need to be pre-analyzed (weighed in the case of PM filters, e.g.), custody of sampling media is also required to maintain the integrity of samples. Samples under chain-of-custody must be under a specified person's control, in their physical possession or in a secure location (e.g., a container that is secure from tampering or in an area with restricted and controlled access). To demonstrate adequate custody, COC forms and sample labels must be maintained that include unique sample identifiers, names of persons collecting the sample, data and time of collection, place of collection, and preservation information. Custody forms should include signatures and custody times/dates for each sample custodian, from sampling to analysis. From an air monitoring perspective, a site operator (field technician) who handles the samples is considered a sample custodian for the time period the sample is in the operator's possession.

Although not explicitly stated, as many of the EPA quality documents are written generically to address multi-media, these same expectations for analytical data apply to ambient air monitoring data. If custody procedures are not followed, samples should not be used for decision making. If custody is incomplete, missing information should be supplemented to the custody form with a signed/dated statement from the appropriate custodian. Where gaps in custody cannot be accounted for, data should be qualified based on the weight of evidence.

2.0 Building a Data Review Program

The fundamental resources needed for establishing a data review program within an ambient air monitoring organization include personnel and tools, the latter of which includes both a physical means to collect and manage data, as well as a well-defined process to review and validate that data once collected. This section discusses these resources in more detail.

2.1 Personnel

Data collection commences in the field at the ambient air monitoring station. Accurate, scientifically-sound data collection is dependent on monitors that are configured and calibrated correctly. A site operator (field technician) is needed for this function. After monitors are calibrated, it is the responsibility

[22] https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nbsir85-3105.pdf

of the site operator to maintain the monitors, reviewing the collected data routinely to ensure it is complete and accurate. Should a monitor begin to drift from its calibration curve, it is the site operator's responsibility to perform corrective actions. The site operator is responsible for conducting QC activities, such as conducting or reviewing zero/span and precision checks of the monitors, as well as performing required maintenance procedures. The site operator must document all QC and maintenance activities. Other local events that may influence the monitor's collected dataset (prescribed burns, e.g.) should also be documented by the site operator. With a first-hand knowledge of the site, the monitors, and activities performed there, the site operator is the ideal individual to perform initial verification of the collected data. With that in mind, the data review program developed by a monitoring organization should include the site operator.

Additional personnel are needed to review data after it has been initially verified by the site operator. A minimum of two additional reviewers are needed to further verify and then validate data after the operator's initial review. Sections 3.2.3 and 3.2.4 of this document detail these secondary and tertiary reviews. As a best practice, the additional data reviewers should be personnel independent from the monitoring organization's field operations, meaning they should not be individuals who generate air monitoring data. See Section 2.1.1 for more information on independence requirements.

A large monitoring organization will likely have numerous monitors and, therefore, several to many site operators. The amount of data generated by a large network of monitors will be substantial, especially if those monitors operate continuously. Therefore, the need for additional personnel to adequately validate the collected data increases. Where possible, monitoring organizations are encouraged to assemble a group (section) of personnel, consisting of multiple individuals, whose responsibilities include verification and validation of collected data. These could be the same individuals who perform QA activities for the organization, a separate section whose sole responsibility is data review and assessment, or a combination of both. Additional personnel may also be needed to process quality-assured data in preparation for AQS upload. An individual within the monitoring organization should be designated as a QA Manager or Officer (QAM or QAO, respectively), whose responsibilities include an independent, final review of the ambient monitoring data before it is released to the AQS database.

It is important to note that different stages of data review require different skill sets. Regardless of structure or number of personnel involved, a best practice is to establish a data review program for the ambient air monitoring network comprised of individuals who understand the data collection activities, the monitoring methodologies utilized, the fundamentals of quality assurance, and the monitoring objectives. At a minimum, the individual designated as the QAM or QAO should also have a keen understanding of the principles and rationale described in Section 1.2 of this document, as well as an understanding of how "big picture" decisions will be made with the collected data.

2.1.1 Independence

Independence in the monitoring program is an essential component of a monitoring program's quality system[23]. A monitoring program's QA management function must have sufficient technical expertise and management authority to conduct independent oversight and should be organizationally independent of environmental data generation activities (i.e., field operations). Likewise, data validation should be performed by individuals independent from the data collection activity. The independence of the data

[23] See 40 CFR 58, Appendix A, Section 2.2

validator and his or her review procedures is critical, to avoid any conflicts of interest or the appearance of such conflicts. As a result, an independent review is necessary for any environmental data to be used for regulatory purposes.

It is important to note that data and all supporting documentation are evidence to substantiate the decision that monitoring data are valid. An independent, third-party reviewer should concur with the original validity decision based upon objective, tangible evidence (documentation) and may assess the data set against additional benchmarks. Reproducibility is part of the scientific method. If an independent reviewer, with no stake in the data, can review the data and supporting documentation and come to similar conclusions, then data quality and defensibility are assured.

EPA acknowledges that smaller monitoring organizations may not have enough trained personnel to accomplish multiple levels of independent data review. However, those monitoring organizations can still find ways to fulfill the independent data review requirement. For example, it is possible that the secondary review of the monitoring data be performed by another site operator within the organization, but one who is independent from the sites/monitors under review. Under this scenario, though, the tertiary review of the data should not be performed by any site operator, in order to ensure adequate independence (i.e., separation from the data collection activity). Also, smaller organizations can work collectively, or with a qualified contractor, to achieve a comparable degree of independence. Likewise, separate programs within a larger organization could collaborate to complete data reviews (such as the QA staff in an environmental agency's Air and Water programs).

Forming or joining a PQAO with another monitoring organization(s) is another possible option to achieve independence of data reviews. PQAOs are responsible for a set of stations that monitors the same pollutant and for which data quality assessments will be pooled[24]. Many PQAOs across the country are established at the state level; and, in some cases, a state with local monitoring organizations may combine into a single PQAO. PQAOs can also be formed by noncontiguous monitoring organizations, which work together with the required degree of independence to conduct data validation. Examples to this approach may include the following:

- Tribal monitoring organizations that combine with other tribal or nearby state/local monitoring organizations within the same EPA Region that measure the same pollutant(s); or,
- Local monitoring organizations, separated by distance geographically within a large state, that form a PQAO for a single pollutant, such as lead, which may not be a pollutant monitored within the state network.

Under these circumstances, the monitoring organizations pooling resources must share the commonalities defined in the CFR (utilize a common QAPP, e.g.), and the PQAO formation must be approved by EPA. Please note the CFR states that each criteria pollutant sampler/monitor must be associated with only one PQAO. Other examples may be possible; when in doubt, the EPA Regional Office can be contacted for advice.

[24] See 40 CFR 58, Appendix A, Section 1.2

2.2 Tools

Personnel assembled to perform data review activities within a monitoring organization should be provided the tools necessary to ensure accurate, transparent, and consistent data validation procedures. The following identifies key tools and resources needed, at a minimum, to perform adequate data review. Other resources may be available.

2.2.1 EPA Data Validation Templates

The EPA QA/G-8 document provides in-depth discussion and specifications for data review, although the document is not ambient air monitoring-specific. As described in QA/G-8, the goals of data validation are to:

- Evaluate whether the data quality goals established during the project planning phase (i.e., the QAPP) have been achieved;
- Ensure that all project requirements are met;
- Determine the impact on data quality of those that are not met; and,
- Document the results.

The QA/G-8 document states, "**The main focus of data validation is determining data quality in terms of accomplishment of measurement quality objectives [MQOs].**" With that in mind, a primary goal for a monitoring organization should be to ensure data quality is evaluated in terms of accomplishment of the MQOs that were developed specifically for the National Ambient Air Quality Monitoring Program. Therefore, critical tools needed in a monitoring organization's data review program are the EPA Data Validation Templates, which can be found in Appendix D of the 2017 QA Handbook and also on the AMTIC website[25]. **The data validation templates contain the MQOs for the Ambient Air Quality Monitoring Program.**

The data validation templates (MQO tables) were initially developed in the late 1990s by a national QA workgroup consisting of stakeholders from SLT monitoring organizations, EPA Regional Offices, and OAQPS, among others. The preamble to Appendix D of the QA Handbook provides more details regarding this national collaboration and the resulting consensus-built templates. To date, the national QA workgroup remains active and weighs in on template revisions, although OAQPS is ultimately responsible for their upkeep. The templates are revised on a periodic basis, to stay current with changes in monitoring regulations, policies, other guidance, and advances in air monitoring technology. It is important to note that the templates can be revised outside of scheduled revisions of the QA Handbook, and for that reason, the templates are linked separately on the AMTIC website, where users can easily access the most current version at any time.

The data validation templates consolidate the MQOs for each pollutant and provide a tool that, when used as described in this document, promotes national consistency in the data quality decision-making process, fostering nationally comparable data sets. A best practice is to implement the acceptance criteria in the data validation templates as *control limits* (i.e., the thresholds at which defined acceptance criteria are considered acceptable, and above which associated data are considered "out of control" and should be invalidated (unless there is compelling evidence demonstrating otherwise), in the case of critical criteria, or investigated, mitigated, and/or justified, in the case of operational or systematic criteria). A significant

[25] https://www.epa.gov/amtic

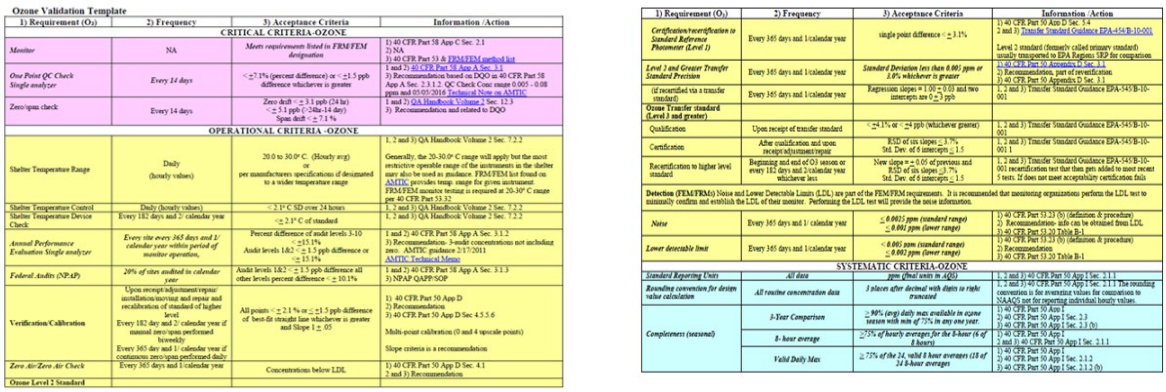
advantage to implementing the MQOs as control limits is that, during annual data certification and QAPP reconciliation, all data should statistically meet the quantitative DQOs established in CFR and the QAPP. Monitoring organizations are encouraged to adopt this approach.

Note: Monitoring organizations are also encouraged to establish and implement action (warning) limits that are more stringent than the MQOs (control limits) for their field operations. Being proactive in the field and performing instrument corrective actions prior to data control limits being exceeded will minimize data loss.

The following section discusses the design, structure, and intended implementation of the data validation templates, current as of the date of this publication.

2.2.1.1 Template Design and Utilization

The data validation templates for the gaseous pollutants are pollutant-based, meaning they are specific to the pollutant of interest and generally not developed for individual makes/models of instrumentation. The data validation templates for particulate pollutants, however, distinguish non-continuous (integrated sampling techniques with subsequent laboratory analyses) from those of continuous (concentrations generated *in situ*) monitors and provide limited technical distinctions based upon instrument-type. (For technical specifications important to individual instruments, the data reviewer should reference the FRM/FEM designation specifications discussed in Section 1.2 of this document, along with the instrument user manuals.) Figure 6 illustrates one of the data validation templates, which presents the MQOs using a color-coded format. **Understanding the structure, formatting, and coloration of the data validation templates is imperative for proper use.** The following subsections explain how to read and use the templates.



Ozone Validation Template

1) Requirement (O_3)	2) Frequency	3) Acceptance Criteria	Information /Action
CRITICAL CRITERIA-OZONE			
Monitor	NA	*Meets requirements listed in FRM/FEM designation*	1) 40 CFR Part 58 App C Sec. 2.1 2) NA 3) 40 CFR Part 53 & FEM/FRM method list
One Point QC Check Single analyzer	*Every 14 days*	≤ ±7.1% (percent difference) or ≤ ±1.5 ppb difference whichever is greater	1 and 3) 40 CFR Part 58 App A Sec. 3.1 3) Recommendation based on DQO in 40 CFR Part 58 App A Sec. 2.3.1.2. QC Check Conc range 0.005 - 0.08 ppm mid 05/05/2016 Technical Note on AMTIC
Zero/span check	Every 14 days	Zero drift +/- ± 3.1 ppb (24 hr) +/- ± 5.1 ppb (+/-3dia-14 day) Span drift < ± 7.1 %	1 and 3) QA Handbook Volume 2 Sec. 12.3 3) Recommendation not related to DQO
OPERATIONAL CRITERIA -OZONE			
Shelter Temperature Range	Daily (hourly values)	20.0 to 30.0° C (Hourly avg) or per manufacturers specifications if designated to a wider temperature range	1, 2 and 3) QA Handbook Volume 2 Sec. 7.2.2 Generally, the 20-30.0° C range will apply but the most restrictive operable range of the instruments in the shelter may also be used as guidance. FRM/FEM list found on AMTIC provides temp. range for given instrument. FRM/FEM monitor testing is required at 20-30° C range per 40 CFR Part 53.32.
Shelter Temperature Control	Daily (hourly values)	< 2.1° C SD over 24 hours	1, 2 and 3) QA Handbook Volume 2 Sec. 7.2.2
Shelter Temperature Device Check	Every 182 days and 2/ calendar year	<± 2.1° C of standard	1, 2 and 3) QA Handbook Volume 2 Sec. 7.2.2
Annual Performance Evaluation Single analyzer	*Every site every 365 days and 1/ calendar year within period of monitor operation,*	Percent difference of audit levels 3-10 ≤±15.1% Audit levels 1&2 ≤ ± 1.5 ppb difference or ≤±15.1%	1 and 2) 40 CFR Part 58 App A Sec. 3.1.2 3) Recommendation- 3-audit concentrations not including zero. AMTIC guidance 2/17/2011 AMTIC Technical Memo
Federal Audits (NPAP)	*20% of sites audited in calendar year*	Audit levels 1&2 ≤ ± 1.5 ppb difference all other levels percent difference < ± 10.1%	1 and 2) 40 CFR Part 58 App A Sec. 3.1.3 3) NPAP QAPP/SOP
Verification/Calibration	Upon receipt/adjustment/repair/ installation/moving and onsite and recalibration of standard of higher level Every 182 day and 2/ calendar year if manual zero/span performed biweekly Every 365 day and 1/ calendar year if continuous zero/span performed daily	All points < ± 2.1 % or ≤ ±1.5 ppb difference of best-fit straight line whichever is greater and Slope 1 ± .05	1) 40 CFR Part 50 App D 2) Recommendation 3) 40 CFR Part 50 App D Sec 4.5.5.6 Multi-point calibration (0 and 4 upscale points) Slope criteria is a recommendation
Zero Air/Zero Air Check	Every 365 days and 1/calendar year	Concentrations below LDL	1) 40 CFR Part 50 App D Sec. 4.1 2 and 3) Recommendation
Ozone Level 2 Standard			

1) Requirement (O_3)	2) Frequency	3) Acceptance Criteria	Information /Action
Certification/recertification to Standard Reference Photometer (Level 1)	Every 365 days and 1/calendar year	single point difference < ± 3.1%	1) 40 CFR Part 50 App D Sec. 5.4 2 and 3) Transfer Standard Guidance EPA-454/B-10-003. Level 2 standard (formerly called primary standard) usually transported to EPA Regions SRP for comparison.
Level 2 and Greater Transfer Standard Precision	Every 365 days and 1/calendar year	*Standard Deviation less than 0.005 ppm or 3.0% whichever is greater*	1) 40 CFR Part 50 Appendix D Sec. 3.1 2) Recommendation, part of recertification 3) 40 CFR Part 50 Appendix D Sec. 3.1
(if recertified via a transfer standard)	Every 365 days and 1/calendar year	Regression slopes = 1.00 ± 0.03 and two intercepts are 0 ± 3 ppb	1, 2 and 3) Transfer Standard Guidance EPA-454/B-10-003
Ozone Transfer standard (Level 3 and greater)			
Qualification	Upon receipt of transfer standard	< ±4.1% or < ±4 ppb (whichever greater)	1, 2 and 3) Transfer Standard Guidance EPA-454/B-10-003
Certification	After qualification and upon receipt/adjustment/repair	RSD of six slopes ≤ 3.7% Std. Dev. of 6 intercepts ≤ 1.5	1, 2 and 3) Transfer Standard Guidance EPA-454/B-10-003.1
Recertification to higher level standard	Beginning and end of O3 season or every 182 days and 2/calendar year whichever less	New slope = ± 0.05 of previous and RSD of six slopes ≤3.7% Std. Dev. of 6 intercepts ≤ 1.5	1, 2 and 3) Transfer Standard Guidance EPA-454/B-10-003 recertification test that then gets added to most recent 5 tests. If does not meet acceptability certification fails
Detection (FEM/FRM) Noise and Lower Detectable Limits (LDL) are part of the FEM/FRM requirements. It is recommended that monitoring organizations perform the LDL test to minimally confirm and establish the LDL of their monitor. Performing the LDL test will provide the noise information.			
Noise	Every 365 days and 1/ calendar year	*≤ 0.0025 ppm (standard range)* *≤ 0.001 ppm (lower range)*	1) 40 CFR Part 53.23 (b) (definition & procedure) 2) Recommendation- info can be obtained from LDL 3) 40 CFR Part 53.20 Table B-1
Lower detectable limit	Every 365 days and 1/calendar year	*≤ 0.005 ppm (standard range)* *≤ 0.002 ppm (lower range)*	1) 40 CFR Part 53.23 (b) (definition & procedure) 2) Recommendation 3) 40 CFR Part 53.20 Table B-1
SYSTEMATIC CRITERIA-OZONE			
Standard Reporting Units	*All data*	*ppm (final units in AQS)*	1, 2 and 3) 40 CFR Part 50 App I Sec. 2.1.1
Rounding convention for design value calculation	*All routine concentration data*	*3 places after decimal with digits to right truncated*	1, 2 and 3) 40 CFR Part 50 App I Sec. 2.1.1 The rounding convention is for truncation values for conversions to NAAQS not for reporting individual hourly values.
Completeness (seasonal)	*3-Year Comparison*	*≥ 90% (avg) daily max available in ozone season with min of 75% in any one year.*	1) 40 CFR Part 50 App I 2) 40 CFR Part 50 App I Sec. 2.3 3) 40 CFR Part 50 App I Sec. 2.3 (b)
	8- hour average	*≥75% of hourly averages for the 8-hour (6 of 8 hours)*	1) 40 CFR Part 50 App I 2 and 3) 40 CFR Part 50 App I Sec. 2.1.1
	Valid Daily Max	*≥ 75% of the 24, valid 8 hour averages (18 of 24 8-hour averages*	1) 40 CFR Part 50 App I 2) 40 CFR Part 50 App I Sec. 2.1.2 3) 40 CFR Part 50 App I Sec. 2.1.2 (b)

Figure 6: Ozone Data Validation Template

Format and Structure

Each **row** in the MQO table contains a specific line-item (QA/QC activity, sample, etc.) that is an important element (requirement) when monitoring for the pollutant of interest.

Each table has four columns. Figure 7 is an enlarged image to show this structure more clearly.

Each **column** has a header that is numbered and labeled, and provides the following significant information:

- Column # 1: Itemized element (**Requirement**)
- Column # 2: **Frequency** of the element
- Column # 3: **Acceptance criteria**
- Column # 4: Additional **Information/Action**, including citations noting where the element (requirement) originated. The column provides a source(s) for the itemized element, its frequency, and its acceptance criteria.

The use of ***Bold/Italics*** means that the specific information highlighted with this **font style** is identified as a requirement in the monitoring regulations (i.e., 40 CFR Parts 50, 53, or 58).

Ozone Validation Template

1) Requirement (O_3)	2) Frequency	3) Acceptance Criteria	Information /Action
CRITICAL CRITERIA-OZONE			
Monitor	NA	*Meets requirements listed in FRM/FEM designation*	1) 40 CFR Part 58 App C Sec. 2.1 2) NA 3) 40 CFR Part 53 & FRM/FEM method list
One Point QC Check Single analyzer	*Every 14 days*	< ±7.1% (percent difference) or < ±1.5 ppb difference whichever is greater	1 and 2) 40 CFR Part 58 App A Sec. 3.1 3) Recommendation based on DQO in 40 CFR Part 58 App A Sec. 2.3.1.2. QC Check Conc range 0.005 - 0.08 ppm and 05/05/2016 Technical Note on AMTIC

Figure 7: Close-Up Snapshot of Validation Template Line Items

Although the structure and formatting of the data validation templates is simplistic, the information presented in the tables is more complex than it appears. The importance of the information in Column 4 (Information/Action) cannot be overstated. Column 4 explains whether the requirement, frequency, and/or acceptance criteria are derived from the CFR, guidance, a specific methodology, or some other source. **It is critical that the data reviewer crosswalk the information in the table against the referenced source(s) to completely understand the specific line-item.** This cross-check should help clarify the coloration of the line-item in the template (discussed in the next section) as well as help the data reviewer gain a clearer understanding of the intent of the requirement.

The following is an example of how to read the templates, highlighting some of the complexities of the information summarized within their columns.

> See Figure 7, Column 1, second row (shaded pink). The line-item is ***One-Point QC Check Single Analyzer***, a QC activity for ozone monitoring shown in bold/italics, which alerts the data reviewer that this activity is found in the CFR. In Column 2, the frequency for the one-point QC check is ***every 14 days*** (again, bold/italics). In Column 3, the acceptance criteria for the ozone one-point QC check is stated as "< ±7.1% (percent difference) or < ±1.5 ppb difference, whichever is greater". However, the acceptance criteria for the QC check is not bold/italicized, which means it is not specified in the CFR. In Column 4, there are two main sources listed to clarify these specifications: for the requirement and frequency (Columns 1 & 2), the information

can be found in 40 CFR Part 58, Appendix A, Section 3.1; however, the acceptance criteria (Column 3) are recommendations of the QA workgroup, based on the DQO for ozone found in Section 2.3.1.2 to Appendix A, Part 58.

Cross-walking the template information, then, against the referenced CFR language, the following is observed:

*3.1.1 One-Point Quality Control (QC) Check for SO_2, NO_2, O_3, and CO. A **one-point QC check must be performed** at least **once every 2 weeks** on each automated monitor used to measure SO_2, NO_2, O_3 and CO.* (Hence, the information specified in Columns 1 and 2 highlighted as CFR requirements using bold/italics, where two weeks has been further defined as 14 days.)

2.3.1.2 Measurement Uncertainty for Automated O3 Methods. The goal for acceptable measurement uncertainty is defined for precision as an upper 90 percent confidence limit for the CV of 7 percent and for bias as an upper 95 percent confidence limit for the absolute bias of 7 percent. (The DQO for ozone, which is an **aggregate** statistic.)

The Section 2.3.1.2 citation does not speak to individual one-point QC checks for ozone or provide a percent difference acceptance criterion. Instead, it addresses coefficient of variation (CV), which is assessed annually (see 40 CFR Part 58, Appendix A, Section 4, as well as Section 4 of this document). However, MQOs are often established for individual phases of a measurement process and may be related to the DQO. If the results of individual ozone one-point QC checks (measurement phase) are held to more stringent limits (i.e., ±7% difference), then the aggregate measurement uncertainty, estimated annually using CV, should be controlled to the levels required by the DQO. (See Section 3.3 of the QA Handbook (2017) for additional information.) The recommended percent difference limit is a reasonable measurement-level acceptance criterion and, when utilized as a control limit, should ensure the ozone DQO of 7% CV and bias will be achieved.

Due to their formatting and structure, the data validation templates are, in essence, a data reviewer's summary sheet of the monitoring regulations, since they allow one to very quickly see which requirements are found in the CFR and, specifically, where to find them. However, it is in the best interest of the monitoring organization to ensure its data reviewers are proficient in the monitoring regulations, especially if one of the monitoring objectives for the organization is to generate data that are NAAQS-comparable. Data reviewers should not rely on the data validation templates alone as their sole source of regulatory information.

Coloration

As stated above, the data validation templates are designed to provide a tool that can yield consistent data validation procedures across the country. Towards that end, the pollutant MQOs are sorted and classified into three major criteria categories: critical, operational, and systematic, with each criteria category having a different degree of implication about data quality. Utilization of the templates, in part, is dictated by the criteria classification, which has specific instructions on how data reviewers are to judge data quality. The templates are color-coded to quickly highlight the three major criteria:

- Pink = Critical Criteria

- Yellow = Operational Criteria
- Blue = Systematic Criteria

Foremost, if data meet the MQOs, they can be deemed valid, unless other evidence demonstrates that they are invalid. When an MQO is not met, however, a judgment call must be made to determine the impact that deviation has had on associated data. The following describes the general protocol for making such judgment calls, which are based on the criteria classifications provided in the templates.

1) Criteria that are deemed critical to maintaining the integrity of *a sample or group of samples* are named **Critical Criteria**. As these criteria have the greatest implications on overall data quality, these items are placed first in the table. In most cases, the requirements classified as critical criteria are regulatory in nature. When performing data review, observations that do not meet each and every critical criterion identified in the MQO table should be **invalidated**, unless there is *compelling evidence* available to justify not doing so. In other words, when critical criteria are violated, the sample or group of samples is invalid until proven otherwise. The compelling evidence is needed to *prove* the data is valid. Typically, the EPA Regional Office will be in the best position to agree as to whether or not the evidence is compelling.

2) Criteria that are important for maintaining and evaluating the quality of the *data collection system* are named **Operational Criteria**. These criteria are placed second on the table. Violation of an operational criterion, or a number of operational criteria, may be cause for data invalidation, depending on the severity of the violation(s). However, the data reviewer should consider other QC information available that may or may not indicate the data are acceptable for the parameter being controlled. The sample or group of samples for which one or more operational criteria are not met are considered **suspect** unless additional QC information demonstrates otherwise and is documented. As a result, data may need to be **qualified (flagged)** to alert data users of the data quality issues.

3) The criteria important for correct data interpretation, but violation of which do not usually impact the validity of a sample or group of samples, are named **Systematic Criteria**. These criteria are placed last on the table. In some cases, violation of a systematic criterion may result in data **qualification**. (Invalidation may be recommended under egregious circumstances; please consult with the appropriate EPA Regional Office prior to invalidating data that violate systematic criteria.)

To summarize, in general, violations of criteria shaded pink in the data validation templates result in data invalidation, whereas violations of criteria shaded yellow or blue in the tables typically result in data qualification (flagging). However, a weight of evidence approach (see Section 2.2.1.3 below) should be taken when assessing the data and the number of violations observed. Generally speaking, when more than one violation of *any* criterion is identified, assurance of data quality decreases. Similarly, the application of more than two QA qualifier codes to any data point should be cause to question data quality; the weight of evidence should be more closely examined, as invalidation may be more appropriate depending on the data's end-use.

The designation of QA/QC activities as operational or systematic criteria does <u>not</u> imply that such activities are insignificant or need not be performed. **EPA notes that not performing an operational or systematic QA/QC check that is required in the CFR can be a basis for invalidation of all associated data.** Users of the templates are urged to notice the use of bold/italics in the yellow and blue sections of

the templates, as numerous elements designated as operational and systematic criteria are found in the CFR. Hence, reviewing the referenced sources in Column 4 of the templates is paramount to helping the data reviewer fully understand the import of each individual element and how to judge data quality against it.

Finally, it is important to note that, during the annual data certification process, EPA Regional Office staff may assess compelling evidence presented by monitoring organizations; this assessment may also occur during Technical Systems Audits (TSAs) and at other times throughout the year. Therefore, EPA Regional Office staff will be in the best position to determine whether there are compelling reasons and justification for retaining data as valid or invalidating data. The Regional Office evaluation will be informed by a weight of evidence approach, considering input from the monitoring organizations and OAQPS (when needed), and be documented. In accordance with CFR, EPA reserves the authority to use or not use monitoring data submitted by a monitoring organization when making regulatory decisions based on the EPA's assessment of the quality of the data.[26] With that in mind, when there are any doubts about data validity, the monitoring organization is encouraged to consult their respective EPA Regional Office for assistance.

2.2.1.2 Compelling Evidence

Compelling evidence is a term that is commonly used in air monitoring data validation that lacks a formal definition in the CFR. However, as defined in Section 1.1 of this document, compelling evidence is data (reason) that concretely establishes instrument performance or the validity of a QA/QC check; in other words, it's objective proof that data are usable despite a critical criterion violation. 40 CFR 58, Appendix A, Section 1.2.3 states that failure to conduct or pass a required check or procedure, or a series of required checks or procedures, does not by itself invalidate data. At quick glance, this regulatory statement may seem contradictory to the protocol in the QA Handbook that recommends data be invalidated that do not meet critical criteria. However, the statement is clarified when discussing it in terms of compelling evidence: there must be a reason(s) to invalidate the data; likewise, there must be a reason(s) to deem the data usable. The following will provide two examples.

1) Failure to conduct the check(s):

   See Figure 7 and the ozone QC check critical criterion requirement. Over a 2-month period, a newly hired operator performs QC checks on an ozone analyzer such that two QC checks are performed each month, but the spacing between checks is anywhere from 15 to 21 days. The data reviewer observes that the frequency does not meet the "biweekly" requirement (i.e., critical criterion), which is defined as "every 14 days" in the data validation templates and in the organization's QAPP. However, when examining the results of all the QC checks, the data reviewer also observes that each check is less than or equal to 2% difference (whereas, the acceptance criterion is $< \pm 7.1\%$ difference). Therefore, the analyzer itself was performing well within its established acceptance criterion (compelling evidence) when the operator conducted the QC checks. The data reviewer rationalizes that, although the operator failed to conduct checks in accordance with the template's 14-day requirement, the results of the tardy checks clearly showed the analyzer was producing acceptable data. Hence, in this example, the failure to conduct the QC checks on schedule did not result in immediate data invalidation. Instead, compelling evidence

[26] See 40 CFR 58, Appendix A, Section 1.2.3

supports the quality of the data. To be transparent about the procedural deviation, however, the data reviewer applied QA qualifier flags to the impacted data in AQS. (*Note: This scenario may have produced a different outcome had the operator not performed any QC checks during the 2-month time period, especially if a subsequent QC check or audit yielded poor results.*)

2) Failure to pass the check(s):

The ozone site is equipped to run automated QC checks. Upon review of the Daily Summary Report from the Central Office, the operator observes last night's automated ozone QC check results were 20% off; the acceptance criterion is < ±7.1% difference. Therefore, this check did not pass. The operator immediately travels to the site to determine the cause of failure. Upon arrival, the analyzer appears to be working normally; no warning or fault lights are observed. However, examination of the site calibrator reveals that it has malfunctioned. The operator hypothesizes that the poor QC results were likely the analyzer quantifying the concentration produced by the malfunctioning calibrator. However, to confirm this, the operator travels back to the office for a replacement calibrator. Returning to the site, the operator then performs a manual QC check. The results are within 3% difference, which confirms the analyzer is producing acceptable data and the poor QC results were caused by the failing calibrator. The operator documents all observations, troubleshooting techniques, and the manual QC results. Thus, in this example, the failure of the automated QC check does not result in immediate data invalidation. Instead, an investigation shows that the QC check itself was not valid due to a malfunctioning calibrator, and a subsequent manual QC check serves as compelling, quantitative evidence that the ozone analyzer continued to produce valid data during the time period in question.

As can be seen from these scenarios, compelling evidence (reason) can be data generated from independent audit point(s), multi-point verifications, and/or a prior zero/span check. Such data establishes whether the analyzer was operating within its acceptance limits. It also indicates whether a QC check itself is considered valid or invalid. Additional information on compelling evidence and how to qualify data in AQS can be found in the 2018 technical memorandum on the AMTIC website titled "*Steps to Qualify or Validate Data After an Exceedance of Critical Criteria Checks*."[27]

It is important to note that compelling evidence (reason) for justifying data validity is not limited to data from QA/QC checks. Compelling evidence can include data and documentation from a variety of other sources. For example, it can include: data from a collocated instrument; data from a nearby monitor (for regional pollutants like ozone and $PM_{2.5}$); biases and outliers identified in control charts; diagnostic data from an analyzer; an analyzer strip chart (i.e., minute data); data on certification records, such as the "as found" status being in or out of tolerance; among others. Be aware, these examples alone may not be "compelling", but rather, when evaluated in combination with other information, the cumulative effect may make the evidence compelling. All collected data and documentation considered compelling evidence in any data quality decision should be retained for data defensibility purposes, in accordance with the monitoring organization's QAPP record retention requirements.

[27] https://www.epa.gov/sites/production/files/2018-01/documents/critical_criteria_qualifier_memo_v1_0.pdf

2.2.1.3 Weight of Evidence Approach

40 CFR 58, Appendix A, Section 1.2.3 states the following:

> PQAOs and the EPA shall use the checks and procedures required in this appendix in combination with other data quality information, reports, and similar documentation that demonstrate overall compliance with Part 58. Accordingly, the EPA and PQAOs shall use a **"weight of evidence"** approach when determining the suitability of data for regulatory decisions. The EPA reserves the authority to use or not use monitoring data submitted by a monitoring organization when making regulatory decisions based on the EPA's assessment of the quality of the data. **Consensus built validation templates** or validation criteria already approved in QAPPs should be used as the basis for the weight of evidence approach. [Emphasis added]

"Weight of evidence" or the "weight of evidence approach" are expressions used when discussing data validation that currently lack formal definitions in the CFR. However, weight of evidence is an essential part of validation, and one the CFR specifically states that PQAOs and the EPA must use. The weight of evidence approach involves using all available supporting documentation, along with professional judgment, to make decisions about data validity and determining whether data meets the needs of the end user (i.e., intended use). For monitoring organizations, the data's end-use often includes NAAQS-decision making purposes, which implies data quality should be able to withstand public and legal scrutiny. The weight of evidence approach involves evaluating data and its supporting documentation and logically determining whether the number of deviations observed, combined with the implications of those deviations, impedes one's ability to defend the quality of the data. More simply put, it's whether the evidence that suggests the data cannot be used for its intended purpose **outweighs** the evidence available that suggests that it can, or vice versa. Although the weight of evidence decision is subjective, it is informed by objective evidence.

In reality, there are some occasions when validity is not a simple "yes or no" decision, but rather a complicated process based on varying types of evidence, layers of supporting documentation, and, quite simply, interpretation of regulatory and methodology requirements. The allowance for a weight of evidence approach affords monitoring organizations and EPA the opportunity to evaluate and analyze all available information, arrive at a validity decision, and then determine whether it can withstand various challenges. When doing this, pursuant to CFR, consensus-built templates and/or validation criteria already approved in QAPPs should be used as the foundation of the weight of evidence approach. The consensus-built templates referenced in the CFR are the QA Handbook's data validation templates. The weight of evidence approach is, therefore, informed by the data validation templates. However, as stated in the CFR, PQAOs and the EPA must use the checks and procedures required in 40 CFR Part 58, Appendix A, **in combination with other data quality information, reports, and similar documentation that demonstrate *overall compliance* with Part 58**. This distinction is important to emphasize. It means that data validation is not simply saying data are valid because required QA/QC checks were completed and passed. Instead, validation is going a step further and ensuring that, not only are the QA/QC checks in compliance, but also the other MQOs, summarized in the data validation templates – such as NIST-traceability, adherence to FRM/FEM specifications, and so forth – have been achieved.

The preamble to the data validation templates (Appendix D of the QA Handbook) recommends invalidation when data do not meet critical criteria, unless there is compelling evidence to justify not

doing so. In the context of weight of evidence, compelling evidence **informs** the weight of evidence decision. It is important to note this distinction, as the terms "compelling evidence" and "weight of evidence" are sometimes used interchangeably. Weight of evidence is often employed when multiple MQO violations have occurred, and most especially in situations where operational and/or systematic criteria have not been met. As stated earlier, where the line-item in the data validation templates originates plays an important role in informing the weight of evidence decision; the data reviewer should understand the intent of all requirements in the data validation templates, which makes review of the sources listed in the Information/Action column of the templates vitally important. When significant operational and/or systematic criteria deviations have occurred, data validity is compromised; invalidation may be warranted, depending on the data's end use. Therefore, the data reviewer must examine all the available evidence in order to inform the decision-making process as to whether **overall compliance with Part 58** has been achieved. The types of data and documentation available, collectively, help "build a case" for the validity decision. That decision should be scientifically sound and technically defensible, in line with the guiding principles described in Section 1.2 of this document.

Figure 8 is a generalized illustration to help visualize the weight of evidence concept. In this illustration, the operational criteria deviation observed is that shelter temperature exceeds 30 degrees Celsius. The data validator is charged with determining the impact of this deviation on overall data validity. As the illustration shows, additional evidence is available that demonstrates adherence to other QC requirements, such as a passing zero/span/precision check. Ultimately, the data validator must "weigh" all of this evidence in order to determine whether the impacted data should be retained, retained but qualified, or invalidated. Appendix C of this document provides several examples of using a weight of evidence approach for reconciling deviations identified in the data validation templates. The data scenarios in Appendix C range from straightforward to complex, and discuss the decision-making process (in other words, as Figure 8 suggests, "which way the scale tips") for each scenario, with suggestions on how the data should ultimately be reported to AQS.

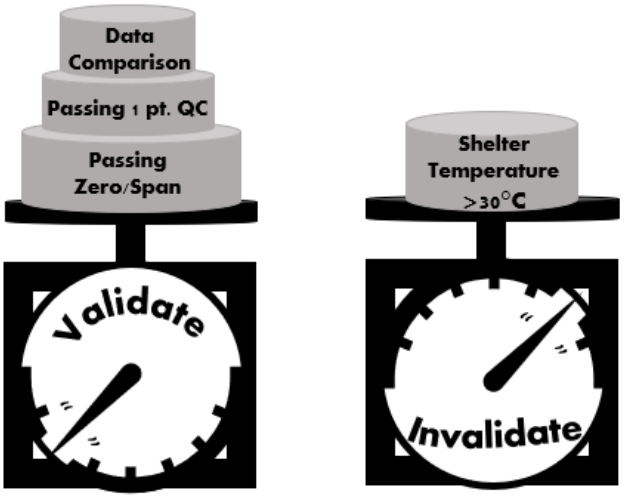


Figure 8: Illustration of Weight of Evidence Concept

It is important to note that the pollutant DQOs are listed as systematic criteria (shaded blue) in the data validation templates. If the DQOs are not met (as observed, for example, during annual data certification on an AQS AMP 600 report), this does not invalidate individual samples for that pollutant. Rather, it impacts the uncertainty associated with the attainment/non-attainment decision

made with that specific data. (Note, there is an inverse relationship between measurement uncertainty and decision-making confidence.) Generally speaking, not meeting DQOs indicates the need for quality system improvements at the monitoring organization level, so that measurement uncertainty is minimized going forward. See Section 15.4 of the QA Handbook (2017) for more information.

Finally, it is recommended that the monitoring organization's independent QAM or QAO be involved in the decision-making process for more complicated weight of evidence scenarios, as well as ones where the validity decision could impact a significant quantity of data. Moreover, in these situations, the monitoring organization is strongly encouraged to contact their EPA Regional Office for additional support. Monitoring organizations should avoid waiting until annual data certification or immediately prior to a TSA to discuss with EPA serious data validity concerns that could impact data completeness requirements or design values. The EPA Regional Office is typically the final decision-maker for these situations; under extreme circumstances, OAQPS may be consulted for additional support and guidance. With that in mind, frequent communication with the EPA Regional Office is strongly recommended as a proactive step in the monitoring organization's validation process.

2.2.2 Quality Assurance Project Plans

A QAPP is the monitoring organization's planning document for conducting a specific ambient air monitoring project. It is an overview of the organization's policies and QA/QC procedures and it formalizes how the monitoring organization plans to assure the quality of the project's data. EPA provides a graded approach to QAPP development (see the QA Handbook, Appendix C), which allows monitoring organizations some flexibility when writing QAPPs, dependent upon the monitoring objectives of the specific project. Monitoring projects that produce data comparable to the NAAQS require a Category 1 QAPP, which has the most stringent requirements. Elements required within a Category 1 QAPP include sections focused on data management, data usability, and verification and validation methods. Many of the data quality considerations described in Section 1.2 of this document are also discussed within the QAPP. A Category 1 QAPP, therefore, is designed to help the monitoring organization produce high quality, NAAQS-comparable data in a consistent manner, within a predetermined amount of measurement uncertainty based on the project's DQOs. **Once approved, the QAPP serves as a written contract between the monitoring organization and the EPA, and its requirements and specifications are expected to be implemented and followed.**

As a best practice, EPA strongly recommends the monitoring organization make efforts to organize its staff and resources in a manner that facilitates a tiered data review approach, such as the one shown in Figure 9, and formalize that structure in its QAPP. Figure 9 illustrates common data review levels, and sometimes overlapping, data review processes. Ideally, each level should include review of the work of prior reviewers, which helps ensure thorough validation. These levels are encouraged but not required; however, all data review essentially goes through stages like these. The tiered data review approach will be described in more detail in Sections 3 and 4 of this document, with particular emphasis on Levels 0 – 3, which are the verification/validation steps. Levels 4 and 5 are primarily the reconciliation steps with the project's DQOs that occur after data has been validated. By implementing a tiered data review approach, the monitoring organization sets itself in the best position to ensure the validity of data by maximizing peer review and independence in the validation process. Such a structure also maximizes the monitoring organization's ability to identify data reporting errors and anomalies, which in turn minimizes data reporting errors to AQS.

In addition to establishing a tiered data review structure in the QAPP, EPA also strongly recommends that the monitoring organization formally adopt the QA Handbook's data validation templates and include them, verbatim, in the QAPP. Towards that end, OAQPS issued a technical memo in July 2017 that specifically addresses the need for this adoption, especially as it relates to the adherence of the template's critical criteria[28]. The data validation templates are the premiere data validation tool for the monitoring organization and contain the MQOs and DQOs for a NAAQS-comparable monitoring network. Incorporating the templates into the QAPP promotes national consistency in ambient air monitoring validation and simplifies QAPP development/writing because the templates are already established, peer-reviewed, and accepted by EPA and the monitoring community. As stated previously, EPA encourages the monitoring organization to utilize the acceptance criteria in the templates as control limits for validation, and to ensure the QAPP clearly states that requirement.

It is important to note that the QAPP is an umbrella document, offering a broad overview of the monitoring organization's policies and procedures. The specific "how to" steps for conducting routine activities – such as data review – are captured in the SOPs the QAPP governs. With this in mind, SOPs must be included in the QAPP (see 40 CFR Part 58, Appendix A, Section 2.1.2). However, in cases where an SOP does not exist for a specific procedure, the QAPP should include the specific "how to" steps. More information about required QAPP elements can be found in the EPA documents *Requirements for Quality Assurance Project Plans* (EPA QA/R-5), *Guidance for Quality Assurance Project Plans* (EPA QA/G-5)[29] and the most recent *Guide to Writing QAPPs for Ambient Air Monitoring Networks* (EPA-454/B-18-006, August 2018)[30].

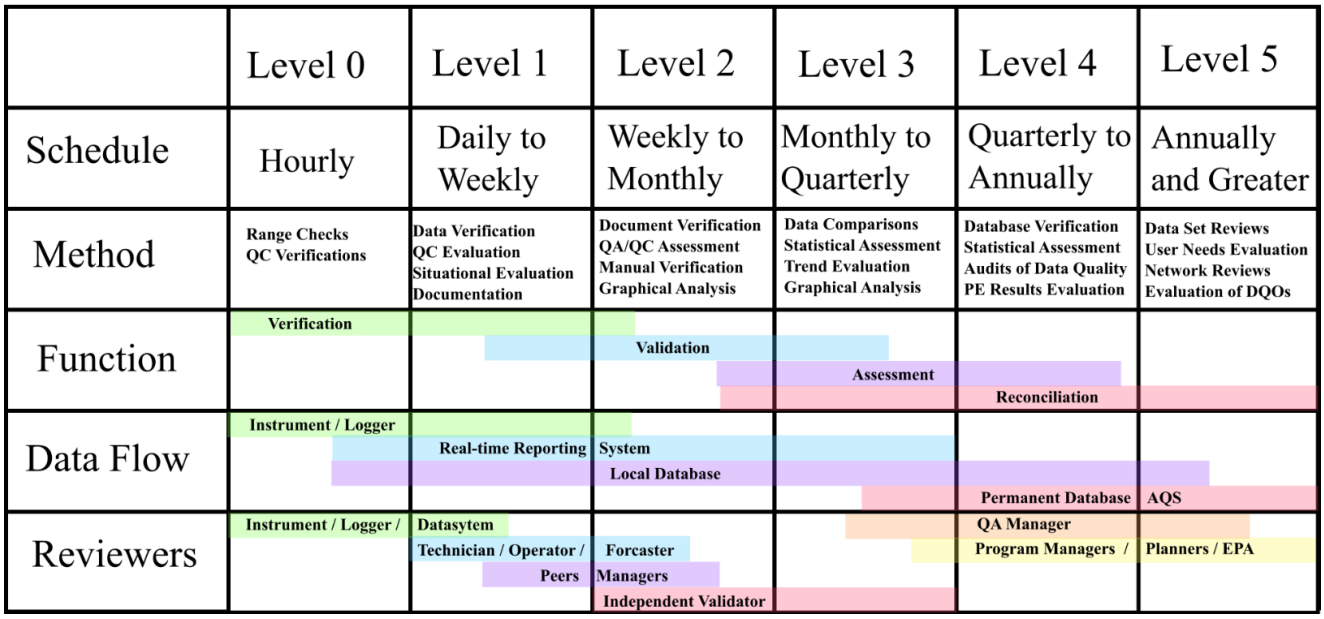


Figure 9: Tiered Data Review Structure for an Ambient Air Monitoring Program

[28] https://www.epa.gov/sites/production/files/2017-10/documents/qappmemo.pdf
[29] https://www.epa.gov/sites/production/files/2015-06/documents/g5-final.pdf
[30] https://www.epa.gov/sites/production/files/2020-10/documents/air_monitoring_qapp_guide_-_final.pdf

2.2.3 Data Review SOPs

An SOP is a "how to" document that provides prescriptive, step-by-step instructions on how to perform certain repetitive tasks. A data review SOP should implement the data review process discussed within its associated QAPP and provide sufficient detail that ensures monitoring organization staff validate data consistently over time. Although monitoring technology has advanced in recent years, automated datalogging and data management systems do not consider all quality indicators or prevent all recording errors. Therefore, in order to ensure data completeness and integrity, additional procedures are needed to complete and standardize the validation process. It is imperative that data reviewers understand their responsibilities as they relate to the data verification/validation process, as well as possess a general understanding of the data review process as a whole, in order to ensure accurate and timely dissemination of data.

All staff involved in data validation should follow the same procedures and utilize the same acceptance criteria. A data review SOP is, therefore, an essential tool for monitoring organization staff and is key to effectively validating data. The data review SOP should:

(1) define roles and responsibilities for data review;
(2) describe how to perform and document the completed reviews;
(3) provide acceptance criteria against which data should be evaluated;
(4) lay out how to address common data-related questions, including application of AQS null and qualifier codes; and
(5) establish timeframes/deadlines for completion of these activities to ensure regulatory reporting requirements are met.

A data review SOP ensures consistency and transparency, which increases confidence in validity decisions. Additionally, a data review SOP is useful for training data reviewers. The QA Handbook (2017) provides additional insight on the importance of SOPs, how they should be written, and what information they should contain. The EPA document, *Guidance for Preparing Standard Operating Procedures* (EPA QA/G-6)[31], also addresses SOPs.

With regards to the data review SOP goals outlined above, the review of supporting documentation is a critical part of validation (see Section 1.2 of this document). Logbooks, data forms, and other records must be maintained in order to justify data flagging or invalidation. Similarly, these records must be available to support that data are valid. The data review SOP should specify which records should be routinely reviewed, especially during the Levels 2 and 3 validation steps. Moreover, the SOP should specify the extent of documentation required by data reviewers to record their part of the review process. It is essential that the data review process be documented at the completion of each level of review, and all notes captured in a package that remains with the validated data set. As technology has advanced and monitoring organizations have moved more towards email and text messaging as a form of correspondence, it is important to note that these electronic conversations are considered records. As such, for those electronic conversations (emails, etc.) that contain the rationale for data validity decisions, or specific instructions to the data reviewer(s) on how to validate or AQS-code data, those emails should be converted to a PDF (or similar) and also maintained with the final data packages.

[31] https://www.epa.gov/sites/production/files/2015-06/documents/g6-final.pdf

The data review SOP should instruct users on how to utilize the data validation templates (i.e., the QAPP's MQO tables). As control limits, the MQOs with acceptance criteria should be considered the threshold at which invalidation will occur if exceeded. The data review SOP should also explain the weight of evidence approach and provide general guidelines for performing it. It is recommended that the SOP prescribe steps that include communicating with the EPA Regional Office when/if a substantial amount of data invalidation may be necessary or the weight of evidence decision is not straightforward.

It is important to note that the data review SOP cannot feasibly account for every scenario in which data will need to be qualified or invalidated. However, the SOP can provide examples of common scenarios and how to address them. Additionally, the data review SOP should prescribe the AQS codes to be applied for the common scenarios in order to facilitate consistent application of codes by all data reviewers. A way to accomplish this would be to include a table within the SOP that contains the AQS null and qualifier codes, defines them, and then provides a brief description for when to use them. (See Figure 10.) This is especially helpful because there is some redundancy in the AQS code list, and a table in the data review SOP could help clarify when to apply certain codes. For instance, the distinction between usage of "AT" (Calibration) and "BC" (Multi-point calibration) could be made in the SOP, as illustrated in Figure 10; in this case, the codes are distinguished such that "AT" means a single adjustment

Null Code	Code Description	When to Use
AH	Sample Flow Rate Out of Limits	Sample flow rate exceeds control limits (ex: failed flow check).
AN	Machine Malfunction	Machine/equipment malfunctions (ex: pump fails).
AQ	Collection Error	Data collection issues with continuous instruments (ex: less than 45 valid minutes collected).
AS	Poor Quality Assurance Results	Failed QC checks (ex: failed 1-point check and data invalidated back to last good check).
AT	Calibration	Continuous particulate matter calibrations (ex: calibrations on a TEOM).
AV	Power Failure	Power failure (ex: site loses power).
AY	Q C Control Points (Zero/Span)	Only a zero and span checks are performed (ex: equipment verification or troubleshooting).
AZ	Q C Audit	Internal quality control audit by agency (ex: agency does an official analysis of their procedures).
AX	Precision Check	Continuous particulate QC check (ex: flow checks).
BA	Maintenance/ Routine Repairs	Routine maintenance and repairs (ex: filter change).
BF	Precision/Zero/ Span	ZSP is performed (ex: prior to a filter change – to bracket the data).
BC	Multi-point Calibration	Gaseous calibrations (ex: multi-point calibrations are performed).

Figure 10: Example of SOP Table Defining When to Apply Common AQS Codes

(such as flow adjustment on a particulate monitor) as compared to multiple concentration points

associated with "BC", where adjustments are performed at multiple concentrations (zero/span). Moreover, the more ambiguous AQS codes should be clarified in the data review SOP. For instance, the "AM" null code (i.e., miscellaneous void) may be appropriate for numerous situations; however, the organization could highlight some specific instances in which the "AM" code will be applied. For example, the monitoring organization could define in the SOP that the "AM" code will be used for scenarios when invalidation is necessary due to water in the sample lines or when concentrations are diluted due to sample train leaks (i.e., sampling shelter air). Other uses for this code are acceptable, too. (More information about AQS codes will be provided in Section 3 and Appendices B and C of this document.)

An important way to further augment the data review SOP is to include screen shots of electronic strip charts that illustrate for data reviewers expected patterns and trends to look for in data sets. Foremost, the SOP should include screen shots of electronic charts illustrating proper operations, such as a quality calibration or QC check performed in accordance with field SOPs. For example, Figure 11 provides a six-hour view of ozone data on an electronic strip chart that illustrates an adjusted calibration, followed by a multi-point verification. In Figure 11, the red line is the analyzer output; the green line is the photometer output.

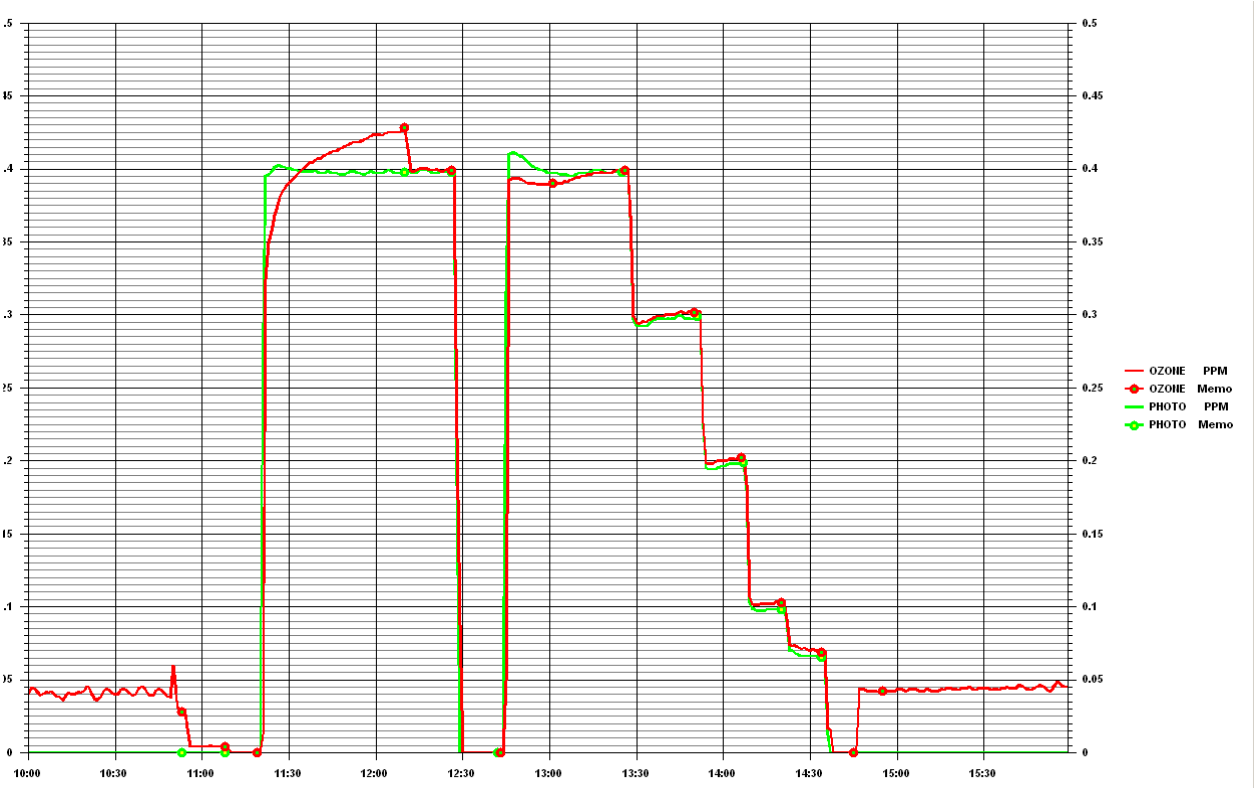


Figure 11: Six-hour View of Ozone Data that Illustrates an Adjusted Calibration

The data review SOP should also include examples of strip charts that highlight the expected behavior of pollutants, such as the diurnal pattern of ozone. For example, Figure 12 provides an example of a 24-hour view of ozone data on an electronic strip chart (i.e., time-series graph with hours as the x-axis, concentrations as the y-axis), illustrating both the diurnal pattern of ozone and an automated nightly zero/span QC check (at approximately *0100* hours). Moreover, the SOP should include illustrations of

strip charts captured during known analyzer issues. Figures 13 and 14 provide some examples of electronic strip chart images at 12- and 6-hour resolutions, respectively, that represent known instrument malfunctions. The images included here were captured when a data reviewer examined the minute data for a specific instrument in conjunction with the site operator's field records and notes. In some cases, the data reviewer also collaborated with the "shop" to confirm the root cause of the analyzer malfunction, after it had been investigated, diagnosed, and subsequently repaired.

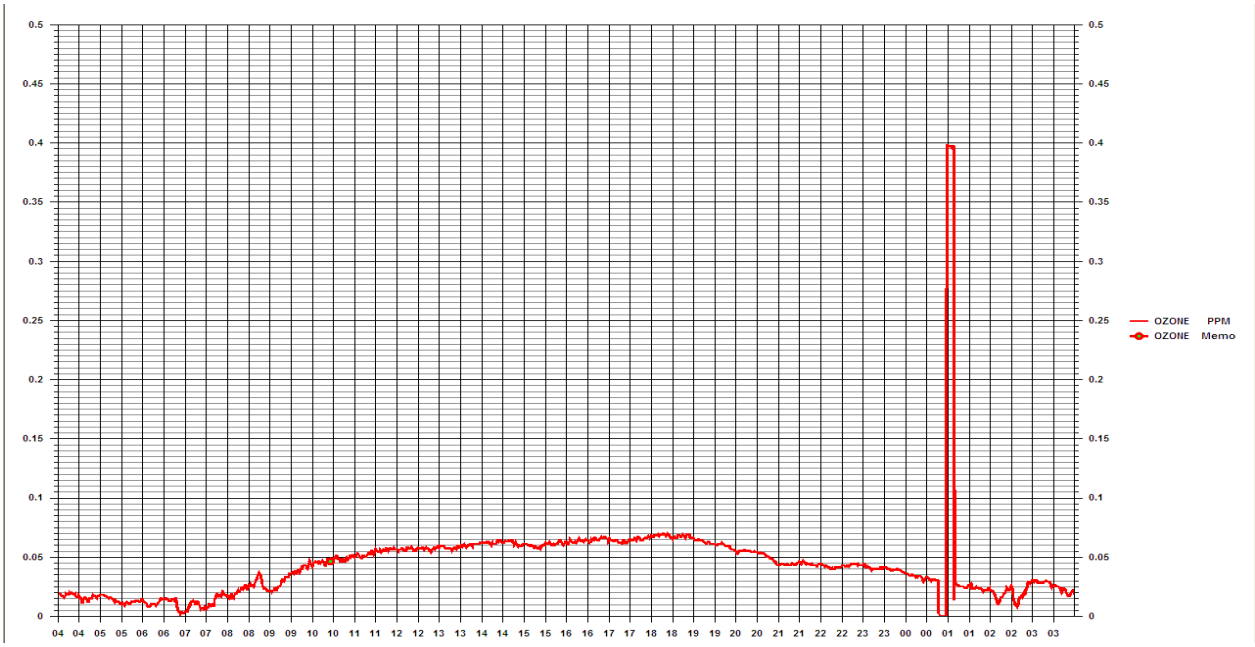


Figure 12: Diurnal Pattern of Ozone and an Automated Nightly Zero/Span Check

It is important to note that SOPs are dynamic and are intended to evolve over time, which is why an annual review and revision is the recommended best practice for document maintenance. With that in

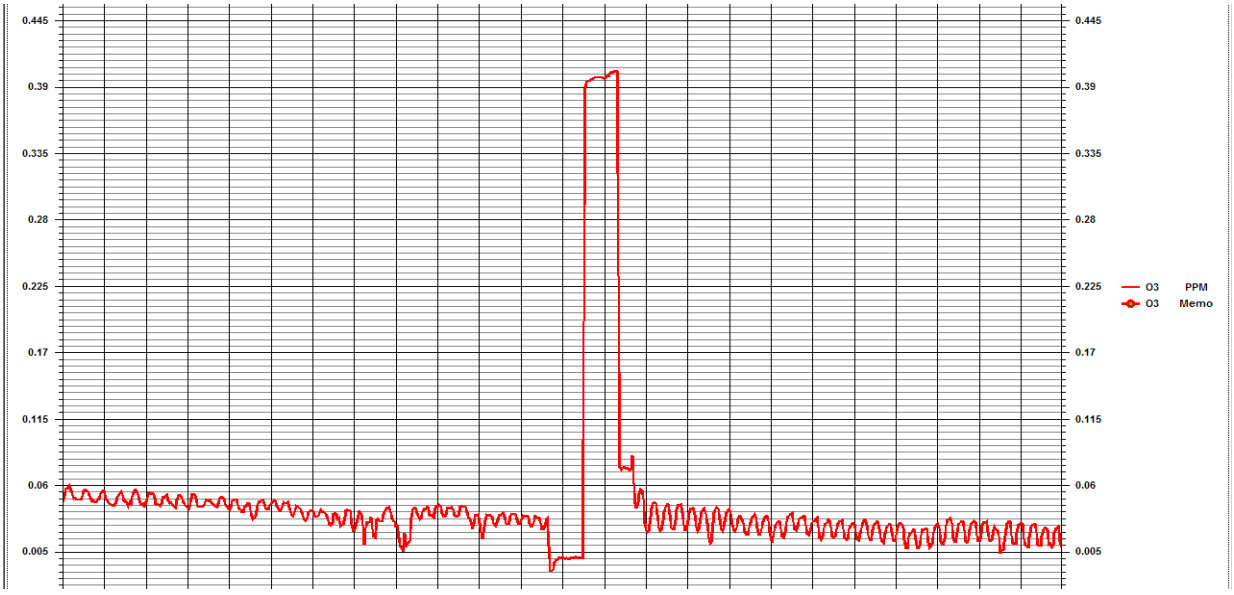


Figure 13: Electronic Chart Trace Illustrating an Ozone Analyzer with a Malfunctioning Detector

mind, screen shots of various analyzer issues can be taken throughout the year and then added to the SOP during routine revision. It is further recommended that the screen captures be taken of the electronic chart

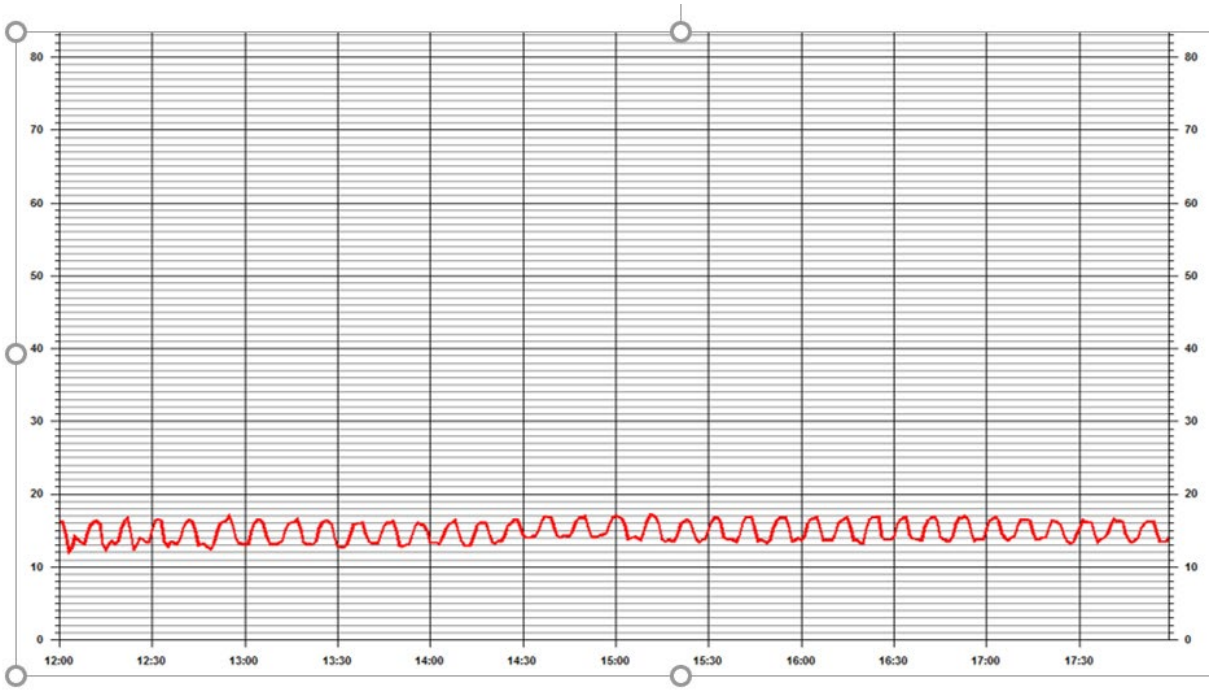


Figure 14: Electronic Chart Trace Illustrating an Ozone Analyzer Impacted by Water in the Sample Line

trace both before and during a known, diagnosed malfunction so that the analyzer response (visual pattern) can be retained for future reference. In this manner, over time, the data review SOP will serve not only as a thorough data review tool, but also as an excellent information repository that will help site operators and data reviewers alike more easily identify monitoring issues, which will also help prevent and minimize data loss.

2.2.4 Data Management Systems

Much of the data collected by a monitoring organization will be collected through the use of automated systems. These systems must be effectively managed and documented by using a set of guidelines and principles by which adherence will ensure data integrity. Discussions of data management activities and requirements can be found in Sections 14 and 17 of the QA Handbook (2017). The monitoring organization's QAPP must detail its data management framework.

Data management systems are an integral piece of the data review process and, thus, are an essential tool. Systems should be configured to:

- Collect and organize 1-minute, 5-minute, and hourly averages of pollutant concentrations;
- Apply pre-programmed flags to data that meet specified conditions;
- Track all changes to data and who they were made by while retaining the original, unedited, data set;
- Provide a platform for adding qualification, comments related to data quality, and/or links to additional data quality documentation (e.g., corrective action reports);
- Provide a means to analyze and visualize data (e.g., charts and tables);

- Provide a means to retrieve and archive data; and,
- Provide a mechanism to output validated data for submittal to EPA's AQS database.

A variety of data management systems are currently available to air monitoring organizations. Some of these systems have sophisticated data verification abilities. The monitoring organizations are encouraged to explore and utilize these capabilities in order to streamline and enhance their data verification processes. At a minimum, basic software can be programmed to scan data for extreme values, rates of change, and other outliers (see Section 3.2.1 below for examples). The automated review can be further refined to account for time of day, time of week, and other cyclic conditions. Utilizing these capabilities as part of a Level 0 review, questionable data values can be automatically flagged to indicate a possible error. This application of initial qualifiers by the data management systems immediately notifies operators of potential data quality issues so they can be corrected quickly. This feature is invaluable, especially when a monitoring organization has a sizeable monitoring network that heavily utilizes continuous monitors.

If the data management software provides the monitoring organization the option of adding user-defined flags, then the monitoring organization is encouraged to define the flags such that they align with AQS codes as much as possible. The monitoring organization should ensure data management system or logger-applied flags are defined in the organization's data review SOP, to ensure proper translation, especially for any instances where the flags do not match those used in AQS.

An additional feature that is a strongly recommended component to a monitoring organization's overall data management system is that of the electronic strip chart. Monitoring organizations are strongly encouraged to invest in this feature. Electronic strip charts should be utilized in conjunction with the continuous analyzers at field sites and documented by site operators during routine operations and site visits. Data reviewers in the monitoring organization's central office should be able to access and review these charts as well. The graphical display of data in an electronic strip chart – particularly data at the 1-minute resolution – is an invaluable tool to assist monitoring staff in determining data quality, as well as assessing the quality and stability of QA/QC procedures performed in the field, including calibrations, QC checks, and audits. The visualization of data on a time-series graph allows data reviewers to more easily identify instrument and site-level problems that might go undetected if only reviewed in a numerical table. Therefore, the importance of its use during data review cannot be overstated. Depending on the averaging time of the data management system in use by the monitoring organization, the graph of the electronic strip chart may vary. **EPA strongly recommends 1-minute data be collected and used for this purpose.** Figures 11-14 above provide some examples of 1-minute data collected and displayed on an electronic strip chart. Section 10 of the QA Handbook discusses electronic strip charts and their review in more detail.

3.0 Data Review Process

Data is influenced by many processes, events, and people, and as such, has a chain-of-custody. There are multiple layers of processing as data travels from the time it is initially collected until it is reported to AQS, all of which can have an impact on the final product. Additionally, the activities of individuals involved in data collection and review have an impact on its overall quality, integrity, and legal defensibility. Therefore, when validating data, the reviewer must examine this chain-of-custody, taking into consideration the many elements that have influenced the data, and determine if it is usable for its intended purposes.

This section of the document is designed to assist monitoring organization staff whose responsibilities include data review, including the site operator. This section will discuss Levels 0-3 data review, which are the verification and validation stages that, in essence, ready data for upload into AQS. Levels 4-5, illustrated in Figure 9, will be discussed in Section 4 of this document.

Data verification and validation often overlap in what stages of data review they occur; AQS codes can be added at any of these stages. Initial data review can begin as early as the data are logged. However, verification and validation must occur before data are entered into AQS and prior to performing final data quality assessments, such as annual data certification. (See Figures 1 and 25.) The monitoring organization's data review SOP should prescribe the order of operations and specify reporting timeframes and deadlines.

Any editing of data, including adding AQS null and QA qualifier codes, should be documented as to why, by whom, and when the edits were made. This information should be retained in data packages that attest to the final validation of the monitoring data. The documentation must be retained in accordance with the records management requirements stipulated in the monitoring organization's QAPP and defined in 2 CFR 1500 and 2 CFR 200.334.

3.1 Application of AQS Codes

Pursuant to the CFR, monitoring organizations must submit ambient air monitoring data to the AQS database in accordance with AQS reporting conventions. AQS codes are an indicator of the reason that a data value:

(1) did not produce a numeric result;
(2) produced a numeric result but it is qualified in some respect relating to the type or validity of the result; or
(3) produced a numeric result but for administrative reasons is not to be reported outside the monitoring organization.

Qualifier codes are used in AQS to provide additional information to a data point (sample). There are four main types of AQS codes: null data qualifier, QA qualifier, request exclusion, and informational only. These codes should be applied as follows:

- Null data qualifiers are required when submitting a null (i.e., nothing was collected) value for the sample measurement. Null codes are also used to represent data (including QC data) that have been invalidated for a specific reason.
- QA qualifiers are used when the sample measurement is available and valid, but the monitoring organization needs to identify (flag) issues with the data to alert end-data users of known limitations with its use.
- Request Exclusion is required when submitting data that is affected by an Exceptional Event and for which an exclusion will be requested from EPA.
- Information Only is optional and can be used in place of a Request Exclusion flag when an exclusion of data will not be requested from EPA or to simply provide additional context to the data. These codes are also useful to provide transparency and a more complete story regarding local impacts or other possible issues associated with a data point.

Adding AQS codes provides more useful information than just reporting data as valid or invalid. AQS allows up to 10 qualifier codes to a single record, although the monitoring organization is strongly encouraged to apply codes judiciously. Valid, flagged data may be usable for some objectives and not others. Flagging data can help ensure that data are legally defensible, because the codes demonstrate awareness of issues and transparency in data reporting. For example, flagging data for exceptional events makes clear that data are undergoing exceptional event review and processing by EPA. Available AQS codes and descriptions can be found on the AQS website[32].

The AQS code used to flag or invalidate data needs to correspond to the specific issue/activity that impacted the data value. For example, if a site operator takes a gaseous analyzer offline and performs a multi-point calibration, the hour(s) affected by that specific activity should be coded "BC", the AQS code for multi-point calibration. Similarly, if the operator performs a stand-alone one-point QC check (1 hour), which exceeds acceptance criteria, then performs instrument maintenance/repair (1 hour) followed by an adjusted calibration (1 hour) to return the analyzer to good working order, the three hours affected should be coded as "AX" (precision check), "BA" (maintenance), and "BC" (multi-point calibration). Forty-five minutes (i.e., 75% of an hour) are needed to have a valid hour in AQS. If multiple activities are performed in one hour, it is recommended that the AQS code that reflects the activity that consumed the majority of the hour be utilized. For the example described above with the failed QC check followed by maintenance/recalibration, if the maintenance event (such as changing a filter) only took 5 minutes of the hour, with the remaining 55 minutes of the hour being the calibration event, then the AQS coding sequence would be "AX" followed by "BC". The AQS codes utilized by the monitoring organization should be defined in the organization's QAPP. Additionally, the monitoring organization's data review SOP should contain a table that lists common AQS codes and how they will be applied (see Section 2.2.3). **The monitoring organization should have and retain supporting documentation to justify the use of specific AQS codes.** Appendix B of this document provides examples of AQS coding for different scenarios.

It is important to note that the AQS AMP 350 (Raw Data) report provides the concentration (hourly or daily) values for the pollutant monitors, and "tells a story" to external users of the data. The AMP 350 can be "read" by viewing the null value codes or QA qualifiers added to the data set. (*Note: The AMP 350 will only display 1 qualifier code per concentration value; if multiple qualifiers have been applied, an AQS AMP 501 report would be needed to view them.*) For example, when an EPA auditor preparing for a TSA reviews the AMP 350 and sees a 3-hour sequence of null codes as "AX, BA, BC" for a gaseous analyzer, it tells the auditor the site operator followed best practices when addressing an instrument issue. However, if the auditor sees hourly coding such as "AX" followed immediately by "AN" (i.e., malfunction), followed by valid hourly concentrations, it raises a red flag to the auditor because a malfunction followed by valid data without evidence of maintenance, repair, or recalibration would not be the best practice in the field. Similarly, a "BC" code followed by "BL" (i.e., QA Audit) would be another example where coding implies best practices may not have been followed, because calibrations should not be performed immediately prior to a performance audit. With this is mind, accurate code selection is important – and the **AMP 350 report should be reviewed routinely by the monitoring organization to ensure the coding reflects the true activities at the monitor/site**. Additionally, the coding on the AQS AMP 350 report should match the hourly display of the electronic chart when comparing them.

[32] https://www.epa.gov/aqs/aqs-code-list

Additional AQS data coding best practices:

- Always code missing data. There should be no "gaps" on an AMP 350 report for a continuous analyzer.
- Apply null codes for scheduled, but missed, intermittent (physical) samples, such as $PM_{2.5}$ FRM or TSP Pb samples.
- Select either a null value code or a QA qualifier code(s). The data point should not contain a combination of both a null code and QA qualifier to describe the scenario.
- Limit use of the Miscellaneous Void (AM) null data code – or, define specific applications of the code's usage in the data review SOP.
- Limit the use of the "1" (i.e., Critical Criterion Not Met) QA qualifier flag. This code is not intended for widespread use and should only be applied under specific circumstances (for an example, see Appendix B of this document). **Most importantly, the "1" flag is not intended to "save" weeks of data that should be otherwise invalidated.** When the "1" flag is applied, EPA will expect to see compelling evidence and documentation to justify the validity of the data.
- Apply null codes and QA qualifiers **consistently.**

Note: AQS codes are updated periodically by the EPA AQS Team. The monitoring organization is encouraged to visit the AQS website on a routine basis to review the current data coding options available.

3.1.1. Data Bracketing

When valid zero, span, or one-point QC checks exceed acceptance limits, ambient measurements should be invalidated back to the most recent point in time where such measurements are known to be valid. Similarly, data following such QC check exceedances that result in invalidated data, or data following an analyzer malfunction or period of non-operation, should be regarded as invalid until the next subsequent acceptable QC check or calibration – in other words, data is invalidated forward until the point of time when measurements are again known to be valid[33]. These validity markers, so to speak, are often referred to in the air monitoring QA community as "data brackets" (see Section 17 of the 2017 QA Handbook). An important concept that is utilized during data verification/validation activities includes appropriately "bracketing" data with AQS codes.

When a calibration, which is a type of QC activity, is performed, the calibration serves as the beginning of data collection – in other words, it's a beginning data bracket. When the next QC check is performed – such as an automated or manual one-point QC check – that QC activity verifies the quality of data that has been collected since the initial calibration. The as-found results of the QC check, then, serve as an "ending bracket". When thinking of these QC checks then, from a data validation standpoint, the data reviewer can quantitatively judge the quality of data between these two known points. When the next QC check is performed, the last QC check serves as the "beginning bracket" and the newer QC check is then an "ending bracket". The cycle repeats itself, with each subsequent QC check serving as both a beginning and ending bracket, depending on which time period of data is being validated. See Figure 15, which is a

[33] https://www.epa.gov/sites/production/files/2018-01/documents/critical_criteria_qualifier_memo_v1_0.pdf

visual representation of QC checks as seen on an electronic strip chart. The first check is an as-found QC check with poor results (ending bracket for the previous 2 weeks of data), which prompts subsequent instrument adjustment (i.e., recalibration – a beginning bracket for a new period of data). To assess the stability of QA/QC activities completed in the field, site operators and data reviewers are encouraged to review the electronic strip chart (1-minute resolution) as the best practice both when performing QA/QC activities in the field and when reviewing data.

Looking more closely at Figure 15, it shows the as-found QC check (i.e., first circled area on the graph) as having multiple concentrations, including zero, precision, and span; therefore, "BF" (i.e., Zero, Precision, Span Check) would be the recommended AQS code for the hour. In response to the poor QC results, the operator initiates a recalibration, which is a lengthy process shown in the second, larger circled area in the figure. Within the larger circle, the electronic chart clearly shows the instrument adjustment at the span concentration, followed by re-spanning the instrument at the same concentration to ensure the adjustment was successful. Afterwards, the operator performs a multi-point verification to ensure linearity of the calibration as a whole. As this entire process (adjustment, followed by multi-point verification)[34] is considered a "multi-point calibration", the recommended AQS coding for these hours would be "BC" (i.e., Multi-Point Calibration).

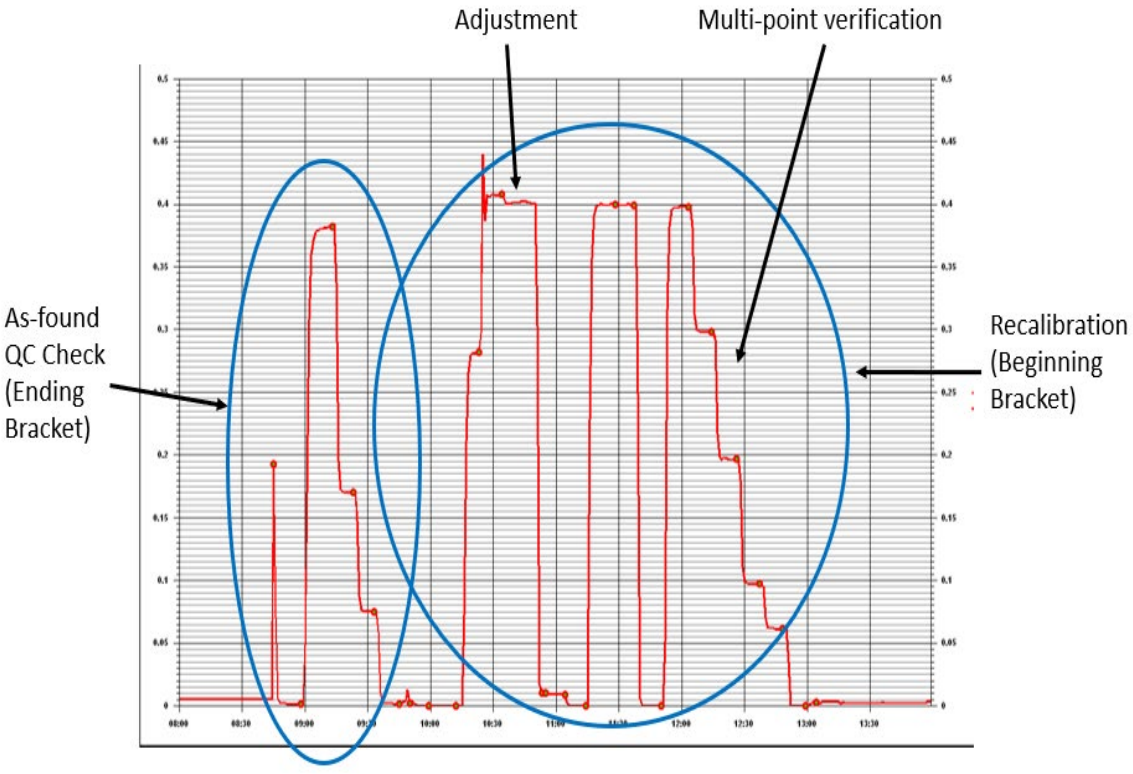


Figure 15: Data Bracketing QC Checks Observed on an Electronic Strip Chart

[34] See Sections 12.2 and 12.3, QA Handbook (2017)

Figure 16 provides another visualization of data bracketing, but this time on a monthly concentration report generated by a monitoring organization's data management software. The example report shows five QC checks evenly spaced during the month (the checks are highlighted in blue with the AQS code "BF"). The checks serve as beginning and ending brackets for the four weeks of data shown on the report. For example, the QC checks on January 1 and 8 (Bracket #1) confirm the quality of data collected between the two checks (i.e., from approximately *1900* hours on January 1 until *1500* hours on January 8). The next data bracket starts with the January 8th QC check and ends with the January 15th QC check, and so on.

Monthly Report

Parameter: NO 42601 January 2017 Avg Interval: 1 hour Units: PPB 008 Method: 200

Hours

Day	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	Summary Max	Avg	RDS
01	.5	2.6	3.2	8.5	5.4	5.0	5.7	6.2	7.6	12.2	12.5	14.2	14.9	14.5	21.6	19.1	BF	BF	BF	13.9	18.9	20.9	10.3	12.9	21.6	10.9	21
02	13.6	9.4	6.7	9.8	7.4	24.4	29.6	25.3	23.2	26.4	16.2	24.5	43.8	33.3	34.3	40.9	24.5	35.3	34.6	35.1	38.4	21.1	18.4	22.2	43.8	24.9	24
03	.5	3.0	17.7	42.4	41.5	50.5	58.4	7.6	15.9	9.6	3.9	2.5	2.6	2.9	2.4	1.9	3.1	.9	.4	2.2	1.7	.7	1.6	1.1	58.4	11.4	24
04	1.2	2.9	.6	5.2	6.8	16.6	22.9	27.2	27.7	30.5	26.4	24.9	23.3	26.7	29.9	19.0	20.0	21.2	20.8	29.1	18.5	20.0	22.8	21.9	30.5	19.4	24
05	13.9	29.8	36.0	15.3	53.3	40.1	118.0	121.4	105.4	88.9	47.3	15.8	5.4	2.6	5.0	7.3	7.3	4.9	23.2	19.1	13.0	29.8	16.5	11.9	121.4	34.6	24
06	13.6	12.5	10.0	18.0	21.9	40.1	36.5	41.7	47.3	36.6	37.8	48.1	36.7	24.8	40.8	39.4	30.0	28.6	25.0	23.9	13.8	13.6	11.5	7.8	48.1	27.5	24
07	4.4	4.2	2.3	2.2	1.4	.9	1.8	2.5	3.5	6.6	7.7	8.4	8.2	8.8	7.5	9.8	7.2	7.3	7.4	8.0	4.6	5.2	4.6	4.4	9.8	5.2	24
08	4.0	3.2	2.4	1.7	2.7	3.5	4.7	5.1	7.2	11.0	10.7	12.1	11.1	11.3	7.7	12.2	BF	BF	BF	25.1	21.3	18.0	21.9	23.4	25.1	10.4	21
09	14.9	20.1	21.0	13.7	45.3	45.4	5.2	7.5	31.7	6.1	1.9	1.9	1.7	1.2	2.0	2.4	1.4	.5	10.8	34.5	70.4	82.6	37.8	21.2	82.6	20.0	24
10	20.0	44.7	41.1	34.7	20.2	36.9	25.3	77.6	30.4	4.8	4.3	2.0	1.6	1.7	1.5	.8	.5	1.3	.6	.8	.7	.2	.2	.4	77.6	14.6	24
11	.2	.1	.1	.1	.1	.2	.5	4.0	7.5	1.2	.7	.9	.7	.5	.4	1.0	.5	.3	1.3	7.0	5.7	.5	.3	.2	7.5	1.4	24
12	.2	.1	.1	17.2	60.3	95.2	110.6	119.8	32.4	8.2	4.6	1.4	1.0	.7	.7	.7	.5	.2	.2	78.2	89.2	108.5	59.4	99.8	119.8	37.0	24
13	88.6	76.1	73.0	83.1	101.3	119.7	167.5	181.1	176.3	215.9	54.5	26.0	10.4	3.7	9.1	14.7	24.1	35.9	49.1	80.0	75.0	74.2	77.7	112.3	215.9	79.5	24
14	92.7	96.2	82.2	90.3	84.1	94.0	93.6	115.0	135.0	163.2	94.7	34.0	13.7	3.3	1.4	.7	10.9	32.6	28.5	40.1	49.2	31.8	20.5	37.2	163.2	60.2	24
15	55.3	36.1	41.2	53.9	60.8	73.9	71.3	75.2	97.2	86.0	57.2	17.3	3.0	2.5	1.6	1.0	BF	BF	BF	44.5	48.1	41.2	76.2	59.3	97.2	47.7	21
16	61.4	56.1	77.3	65.3	52.4	59.6	78.3	91.2	86.2	106.9	63.9	5.6	2.7	2.0	4.3	2.3	.6	1.1	11.1	27.6	34.5	33.9	29.2	41.7	106.9	41.4	24
17	42.7	34.7	35.1	38.3	11.8	43.5	49.9	64.8	63.5	24.6	10.8	3.1	.5	.5	.4	.6	.4	.2	.1	17.4	54.0	45.8	2.6	.2	64.8	22.7	24
18	.1	5.0	40.9	13.2	.4	16.4	7.3	14.5	17.3	33.4	17.9	12.6	8.7	10.6	8.4	2.8	3.3	1.4	3.0	16.2	14.5	24.0	43.4	31.1	43.4	14.4	24
19	55.3	75.7	62.9	77.0	64.8	106.9	99.1	225.1	123.2	39.8	13.3	2.6	1.6	1.4	1.0	.5	.3	.4	27.4	25.5	10.9	7.0	14.6	9.7	225.1	43.5	24
20	9.7	24.9	.1	.0	.0	.3	.2	.3	1.4	2.3	4.7	3.1	2.7	3.9	1.9	1.3	1.1	.4	.3	1.2	35.8	47.3	55.4	62.4	62.4	10.8	24
21	64.7	53.9	38.4	10.5	7.7	28.9	28.4	11.0	29.2	26.6	25.5	18.5	10.7	10.0	9.2	3.5	.6	9.5	21.7	38.3	37.9	19.7	59.1	40.7	64.7	25.0	24
22	4.7	2.8	3.8	5.7	5.2	.2	.2	.9	1.0	4.0	1.5	8.0	3.3	12.2	7.4	7.4	BF	BF	BF	.5	.0	.2	.0	.0	12.2	3.2	21
23	.0	.0	.0	.0	9.3	32.8	33.8	17.6	9.0	9.2	17.5	15.1	15.8	15.2	14.9	14.2	13.4	14.5	9.8	8.6	8.3	7.3	7.1	7.1	33.8	11.6	24
24	5.5	5.3	4.4	5.4	7.9	16.1	19.5	17.6	23.7	29.6	22.2	17.4	15.7	19.0	13.0	17.4	11.2	14.5	56.7	36.7	39.8	76.6	95.0	73.6	95.0	26.8	24
25	76.2	77.3	102.1	113.0	96.3	111.9	299.1	353.0	238.8	144.7	71.8	18.7	1.7	1.4	1.6	1.3	.7	.3	.2	.2	.7	49.3	55.8	18.3	353.0	76.4	24
26	.2	.1	.2	.8	3.9	5.8	6.6	9.8	13.5	15.3	11.5	9.4	7.9	12.8	8.8	10.8	10.9	9.7	16.5	14.2	15.4	14.4	8.6	10.6	16.5	9.0	24
27	5.0	.5	.4	1.0	3.4	12.0	12.1	12.8	11.8	17.6	12.4	14.9	11.6	12.8	8.7	9.3	9.4	6.5	9.1	13.3	5.1	1.5	1.2	1.7	17.6	8.0	24
28	1.6	3.6	2.3	4.8	9.6	2.0	8.7	6.1	3.2	6.0	10.1	6.4	5.9	3.0	4.1	1.5	1.3	.5	.3	.5	.3	.4	.4	.5	10.1	3.4	24
29	.3	.3	.4	.7	2.9	.4	.7	.9	3.4	6.3	5.0	2.7	1.8	1.4	1.2	.9	BF	BF	BF	7.3	5.8	4.4	5.1	3.7	7.3	2.6	21
30	2.9	2.8	2.5	5.6	3.5	2.8	7.5	37.5	10.6	24.2	19.3	14.9	11.1	10.3	11.2	7.5	5.7	4.1	6.8	64.6	59.0	71.8	63.7	108.8	108.8	23.2	24
31	94.4	88.4	153.3	143.0	135.5	162.9	196.2	196.4	110.0	14.6	7.6	6.6	3.4	3.2	2.2	1.7	1.7	.4	.2	.2	.3	73.0	61.4	117.9	196.4	65.6	24
Max	94.4	96.2	153.3	143.0	135.5	162.9	299.1	353.0	238.8	215.9	94.7	48.1	43.8	33.3	40.8	40.9	30.0	35.9	56.7	78.2	89.2	108.5	95.0	117.9	353.0		
Avg	24.1	24.9	27.7	28.4	29.9	40.2	51.5	60.5	48.1	39.1	22.4	12.6	9.1	8.3	8.5	8.1	7.3	8.9	14.0	22.3	25.5	30.4	28.4	31.0		25.8	
Count	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	26	26	26	31	31	31	31	31			729

Bracket #1 Jan 1 - 8

Bracket #2 Jan 8 - 15

Bracket #3 Jan 15 - 22

Bracket #4 Jan 22 - 29

Figure 16: Monthly Data Report with Data-Bracketing QC Checks Highlighted

Similar principles can be applied to non-continuous, or intermittent, sampler data. The following are two examples of scenarios discovered during data review, where the data validator must invalidate data for an intermittent sampler and needs to appropriately bracket the data using QC checks.

Example 1:

A $PM_{2.5}$ FRM flow check on November 18 exceeds acceptance criteria at 4.5% difference (d). The site operator recognizes this value is outside the SOP control limit and immediately performs the necessary maintenance and flow recalibration (same day). At this site, the operator performs flow rate verifications only once per month. The last passing flow rate verification was on October 23 at 3.7% d. The data validator, using the QC checks as brackets, invalidates all samples that were collected between October 23 and November 18. The AQS code chosen by the data validator to invalidate these weeks of data is "AS" (i.e., "Poor Quality Assurance Results).

Example 2:

A $PM_{2.5}$ FRM flow check on November 18 exceeds acceptance criteria at 4.5% difference (d). The site operator does not recognize the value is outside the SOP control limit and does not perform any maintenance. At this site, the operator performs flow rate verifications approximately once per month. A semi-annual flow rate audit is performed on December 21 with results of 3.9% d. The auditor reminds the operator that the acceptance criterion is 4% and suggests that a recalibration be performed before the acceptance criterion is exceeded. The site operator does as suggested that same day. The next passing flow rate check following recalibration is on December 30 at 1.6% d.

The data validator reviews this information and determines that samples before and after the failed November flow check must be invalidated. Upon review of documentation, the last passing QC check prior to the failure was on October 23 at 3.7% d. Going forward after the November check, the first passing QC check is the semi-annual flow rate audit on December 21, which is followed by a recalibration on that same day. Therefore, the data validator invalidates all samples between October 23 and December 21. The AQS code chosen to invalidate these weeks of data is "AS" (i.e., "Poor Quality Assurance Results).

3.2 Tiered Data Review Approach

The procedures by which ambient air quality data are obtained, processed, and reduced to the various reporting formats in a monitoring organization is a complex undertaking, involving the coordinated work of multiple staff. A QAPP describes the monitoring organization's data management framework and

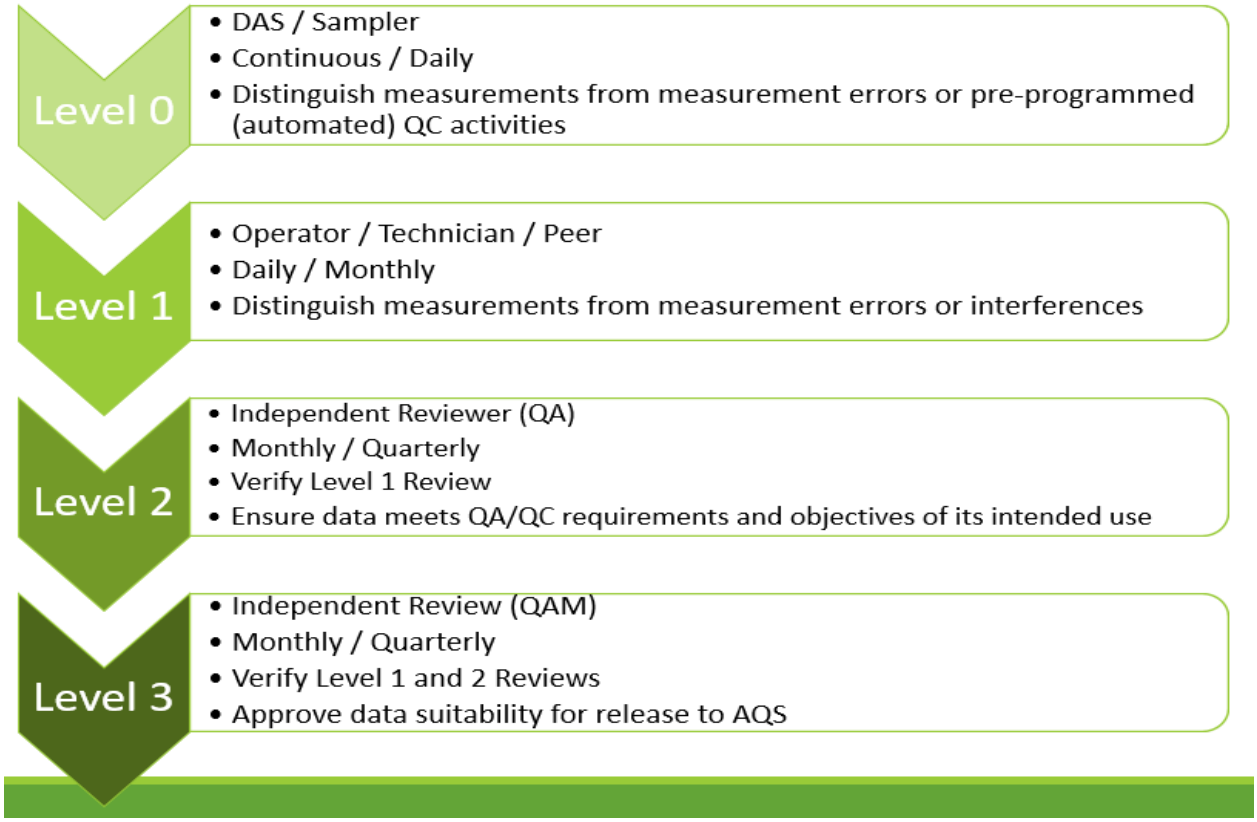


Figure 17: Summary of Levels 0 – 3 Data Review Activities

review requirements. Figure 9 illustrates a manageable, organized framework for performing effective data review. As stated above, EPA encourages monitoring organizations to adopt this approach, or construct one similar (resource-dependent). The data review SOP, on the other hand, should provide specific, detailed instructions on how to complete the Levels 0-3 reviews, in particular, which are primarily the verification and validation stages. A summary of the Levels 0-3 stages and their primary goals are shown in Figure 17.

It is important to note that monitoring organizations have different data handling procedures, acquisition systems, and staffing levels. This section provides general principles and examples for reviewing ambient air monitoring data that apply across all agencies, regardless of these differences. **Appendix A of this document provides a data review tool that can be used in conjunction with the procedures described herein.**

3.2.1 Level 0 Data Review

Data acquisition systems display continuous/near real-time concentrations from air monitoring instruments. The first, or Level 0 phase, of data review utilizes automated systems and occurs as the monitoring data are originally acquired. It includes automatic flagging of data by an instrument, datalogger, and/or management system, which has been pre-programmed with specific acceptance criteria. This is a continuous, daily process. The Level 0 data review can help distinguish valid measurements from measurement errors, as well as distinguish actual measurements from automated QC activities, such as nightly zero/span/precision checks. For example, if automated nightly QC checks are scheduled, the data associated with those checks can be automatically flagged by the automated system with a user-defined flag that alerts the data reviewer of the specific check. Similarly, some monitoring equipment, such as particulate samplers, have this ability to flag data as they are acquired. Other management systems, such as AirNow-Tech, screen data prior to reporting real-time to a public interface. System codes for flow rate, filter loading, or any out-of-range parameters can be pre-programmed in the automated system / software and are very useful in diagnosing problems.

Some examples of pre-programmed factors that commonly are applied in automated flagging for Level 0 review include:

(1) Out-of-range parameters (e.g., identifying data that have some parameter outside of an expected range, such as exceedances of shelter temperature designated to fall between 20-30 degrees Celsius);
(2) Values that exceed an established low or high ceiling (such as values that exceed the NAAQS standard or values that exceed the calibration range of the monitor);
(3) "Stuck" or repeating identical values for more than a few hours/days that can be flagged as suspect and require further investigation;
(4) Data that change by more than preset limits from one hour to the next (e.g., ozone rate of change) can be flagged for further investigation;
(5) Power failures of more than a certain number of seconds/minutes
(6) Hours with less than 45 minutes of data;
(7) Automated QC checks / maintenance activities that are controlled by the data acquisition system; and,
(8) Results of QC activities that exceed defined thresholds.

Additionally, different flagging can be set up for different seasons, as the expected range and behavior of pollutants change. Data sets polled/downloaded will display the flags applied by the instruments/data acquisition systems and, in some cases, field operations staff will be notified by text or email on instrument status so they can begin the next level of review.

Given the amount of data that is collected in a monitoring network, especially one with a high percentage of continuous instrumentation, an automated (Level 0) review process ultimately increases the likelihood that erroneous data will be identified and appropriately addressed, while simultaneously reducing the amount of staff hours needed to manually evaluate data for certain criteria. EPA encourages monitoring organizations to explore and utilize automated data verification capabilities in order to streamline and enhance their Level 0 – 1 data verification processes.

3.2.2 Level 1 Data Review

Data should be reviewed as soon as reasonable after it is gathered. Ideally, the first step in the monitoring organization's data review process includes an automated Level 0 review stage that evaluates data on a near real-time basis for criteria such as that listed in Section 3.2.1 above. A more thorough verification, which includes the review of additional records and supporting information, should follow soon afterwards and be documented. This next review stage is referred to as Level 1 data review.

Level 1 data review should occur on a daily basis, with the data reviewer verifying the previous 24-hour's worth of data. If problems are readily identified during this review stage, they can be fixed more quickly, documented, and the system can resume gathering valid data sooner, minimizing data loss. Timely review also ensures that data quality issues, including any local impacts near the site/monitor, are consistently and accurately documented. The **goals** of Level 1 data review include:

- To distinguish measurements from measurement errors, interferences, or contamination; and,
- To document events that impact data quality clearly when they first occur, so they don't have to be reconstructed weeks or months later.

The site operator is the most knowledgeable about the site, instrument(s), procedures, and surrounding environment, including local activities that can affect the data, such as nearby prescribed burns and construction activity, among others. Therefore, the site operator is best positioned to make site-level decisions and document them. With that in mind, Level 1 data review should be performed, ideally, by the site operator. During the Level 1 review, the site operator should document observations in the data set, so that subsequent reviewers can understand and build upon the site operator's experiences and technical expertise. In the event the site operator is not delegated this responsibility, the monitoring organization should ensure another technician or peer with knowledge of the monitoring equipment and requirements is available to perform the Level 1 review.

Level 1 data review should evaluate 100% of the data collected. Although this may sound challenging, when reviewing data on a daily basis, it equates to small data sets. The workload is even more manageable when it is distributed amongst the monitoring organization's site operators. The automated review performed during the Level 0 stage will have already highlighted areas of concerns within the data set, which expedites the review process. In addition, daily review offers the most efficient strategy for reviewing the accompanying documentation (e.g., data forms, logbook entries, etc.) because

the number of records available for daily review tends to be limited. The monitoring organization's data review SOP should include the specific how-to steps to instruct the data reviewer (site operator) on how to access the monitoring data and navigate through the data management system. Likely, the data management system (software) should provide some type of "daily summary report" (or similar) that will allow the reviewer to see the hourly averages from the previous 24 hours. The data review SOP should prescribe how the Level 1 reviewer is to document findings and observations in the data set; this may be done electronically or manually, depending on the monitoring organization's resources and capabilities. It is important to note that the Level 1 reviewer (site operator) can make recommendations on how data should be null coded or qualified in AQS. The data review SOP should clearly instruct the Level 1 reviewer on how to communicate and document coding recommendations.

To perform daily Level 1 review, access and view the previous 24 hours of data. Access to the electronic strip charts that correspond to these hours of data should be available to the Level 1 reviewer as well. Follow the steps below, which serve as a thorough guide to evaluating the collected data and supporting documentation, in order to achieve the Level 1 goals stated above. Note that variations in the review approach are acceptable; for example, some monitoring organizations may perform some of these steps during Level 2 review, depending on their resources and capabilities. Appendix A of this document contains a Data Verification Checklist (tool) that can assist the Level 1 reviewer when performing the review.

Recommended Daily Level 1 Review Approach:

1. Look for gaps in data collection (i.e., missing data). See Figure 18 for an example.
    a. If identified, determine root cause of data loss and document it.
    b. Re-poll datalogger or instrument, if possible, to see if missing data can be restored.
2. Review all status flags applied by the data management system (datalogger, sampler, etc.) during the Level 0 review. Some software packages may color-code this data. ***Note: If an automated Level 0 review is not performed, then the reviewer will need to verify the data for criteria such as that listed in Section 3.2.1 above.***
    a. Determine if the status flags are expected and accurate.
        - For example, if a nightly, automated QC check is programmed to occur during the *0100-0200* hours, does the daily/hourly summary report show a QC check flag for those specific hours?
    b. If unexpected, investigate the data points further to determine root cause(s) and document findings.
        - For example, if a user-defined flag indicates a power failure occurred, the reviewer should look at the associated minute data to see precisely when the power failure occurred and how many minutes of data were lost. (Some software packages may apply a power-loss flag when mere seconds of data are impacted.) If 45 minutes or more of ambient data are available in the hour, the hour is likely valid (barring other issues). Observe whether other instruments at the site experienced power loss during that same hour. It is likely that a significant power surge would impact most or all the instruments at the site. Also, it is important to note that power failures can cause continuous instruments to "spike" or otherwise show an erratic strip chart for a few minutes or longer upon powering back up. The Level 1 reviewer should look for these scenarios in the data.

3. For each pollutant monitored, verify the maximum and minimum hourly concentrations, and document any errors.
    a. Do the values make sense? (Site operators should be familiar with what pollutant concentrations are considered normal for the site, at different times of day and year, and be on the lookout for unexpected results.)
    b. Are the values real, the result of an automated QC procedure, or an anomaly? Compare the value to the electronic strip chart and any available logbook / QC data forms.
    NOTE: Be sure to review the 1-hour concentration maximum values for <u>all</u> pollutants, including ozone and CO. This is especially important because, if the hourly maximum concentration is erroneous (calibration gases reported as ambient, e.g.), then the 8-hour averages that encompass that hour will be incorrect.

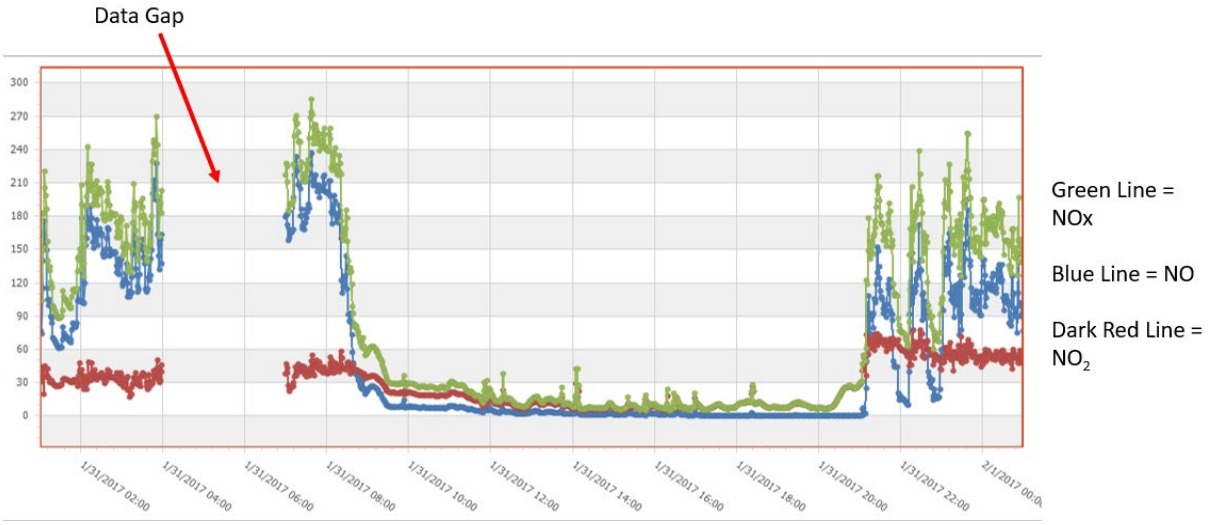


Figure 18: Strip Chart of NO-NO2-NOx Data that Illustrates a Gap (i.e., missing data) during the 0400-0700 time period

4. Look for the expected behavior of the pollutant. If anomalies are identified, investigate why and document. Some common examples to look for may include, but are not limited to:
    - Is the diurnal pattern of ozone present? (View the strip chart to confirm the presence of the expected curve; see Figure 12 for an example.) If not, investigate why. This may include reviewing the weather conditions for the specific location.
    - Do NO-NO_2-NOx values rise and fall as expected during rush hour? If plotting ozone and the oxides of nitrogen together on an electronic chart, is titration visible when expected?
    - Does the NOx concentration minus the NO concentration equal (approximately) the NO_2 concentration? (This calculation can be verified manually and can also be easily observed on the electronic strip chart. See Figure 20 for an example.)
    - Are PM_{10} concentrations higher than $PM_{2.5}$ concentrations at a collocated site?
5. Verify data values against FRM/FEM designation specifications, such as shelter temperature requirements for the instrument.
    a. If identified, document the impacted hour(s). Determine if a site visit is warranted to perform corrective actions.
6. Verify the data against instrument diagnostics specifications (e.g., lamp intensities, flow rates, monitor slope/offset, etc.).

a. Site operators need to be aware of the acceptance ranges for various instrument diagnostics, as they can fluctuate.
b. Document any excursions from the user manual/SOP specifications and determine if a site visit is warranted to perform corrective actions.
NOTE: Availability of diagnostic data is often resource and equipment-dependent. Some diagnostic data may be polled electronically by the DAS. At a minimum, critical instrument diagnostics should be manually recorded in logbooks or on data forms by the operator during routine site visits.

7. Look for negative readings.
   a. If observed, do the negative readings exceed AQS reporting limits? (See EPA Technical Note titled "Reporting Negative Values for Criteria Pollutant Gaseous Monitors to AQS."[35])
   b. Investigate cause(s) and document. Determine if a site visit is warranted to perform corrective actions. *Note: If the monitor has been consistently producing negative readings for some time, it may indicate zero drift and the need to adjust the monitor's calibration baseline.*
8. Look for constantly repeating values. Figure 19 illustrates "stuck" (i.e., repeating) concentration values, which appear as "stair steps" on the electronic strip chart. Two periods of missing data are also visible.
   a. If identified, investigate to determine root cause(s) and document. Determine if a site visit is warranted to perform corrective actions.

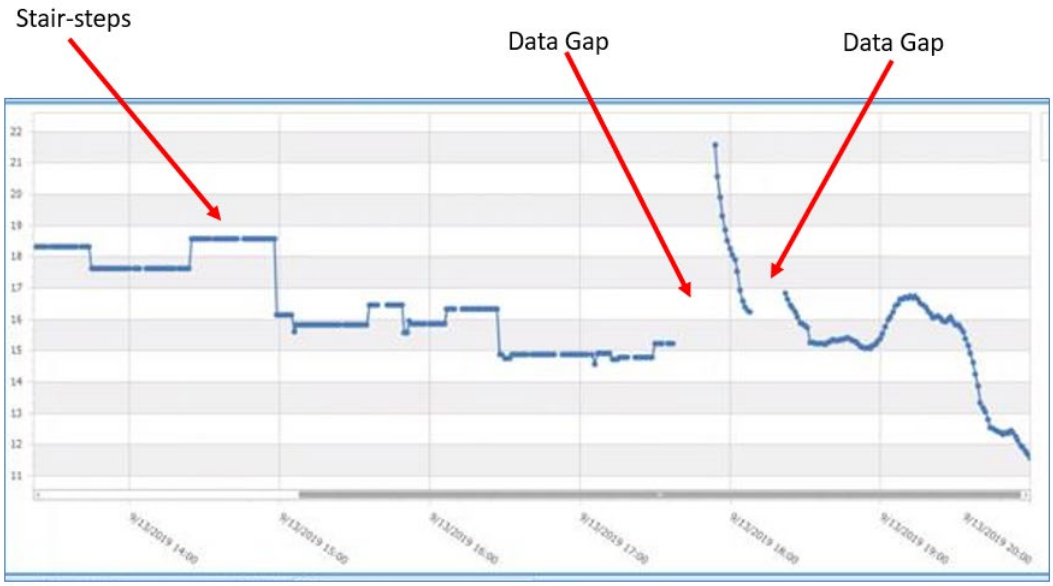


Figure 19: Strip Chart for a Continuous $PM_{2.5}$ Sampler that Shows "Stuck" Concentration Values

[35] https://www.epa.gov/sites/production/files/2017-02/documents/negative_values_reporting_to_aqs_10_6_16.pdf

9. Look for outliers, such as values that appear anomalously high, or those the DAS may have highlighted as exceeding defined thresholds (such as values greater than 2 to 3 times the standard deviation of the historical average concentrations of the monitor, etc.).
    a. If identified, investigate to determine root cause(s) and document.
10. Compare results of any instrument calibrations, QA/QC checks, and maintenance activities to applicable specifications to look for anomalies or failures. Document, if identified, and determine if corrective actions are warranted.
11. Select a random hour and compare the pollutant concentration on the summary report to the analyzer's strip chart (analog or digital) to check for DAS accuracy. Do the values match? *Note: The data review SOP should prescribe an allowable ppm/ppb difference between the strip chart and the DAS; corrective action would be warranted if that allowable difference is exceeded.*
12. Review documentation associated with the 24-hour data set to ensure records and commentary are complete, accurate, descriptive, legible (if handwritten), and, where appropriate, signed and dated.

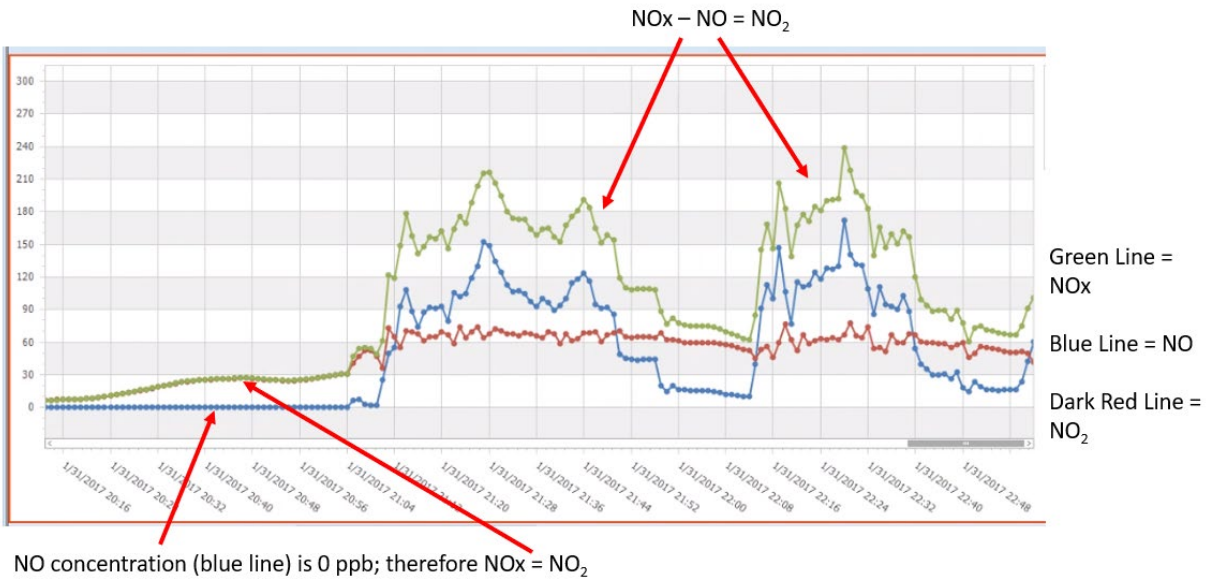


Figure 20: Electronic Strip Chart of NO-NO2-NOx, where the Three Pollutants Traces Demonstrate the Expected Pollutant Behavior

<u>Additional Review for Intermittent Samplers</u>:

The Level 1 review process for intermittent data, such as $PM_{2.5}$ FRM or lead (Pb) samples collected on a 1-in-3, 1-in-6, or 1-in-12 day schedule, follows the same basic concepts as described above. During Level 1 review for intermittent data, the data and records readily available to the reviewer would be those associated primarily with field operations, which includes pre- and post-sample collection activities. The Level 1 review should occur as soon as the data is available, but at least on a weekly or monthly basis. An example Level 1 review approach for intermittent samples includes:

1. Download and verify data collected by the sampler (if available) to look for errors.
    a. Some models of intermittent samplers contain dataloggers (or similar) that are pre-programmed to identify exceedances of critical performance specifications or other outliers. If identified, investigate root cause(s) and document.
    b. Some samplers will also throw status flags in the event of certain mechanical failures. If identified, investigate root cause(s) and document.
    c. Some samplers collect 5-minute data and provide summary files of the 24-hour sampling event. Review these files for anomalies or errors; document any issues identified and investigate root cause(s).
    d. Perform corrective actions, as needed, and document them.
2. Review sampler performance specifications and diagnostics that were recorded manually during the site visit (e.g., flow rate, temperature and barometric pressure readings, leak rate, sampler clock/timer).
    a. Ensure supporting records, such as logbooks and data forms, are complete and accurate.
    b. Earmark noted exceedances of acceptance criteria. Perform and document corrective actions.
3. Review sampler/station conditions at the time of sample set-up, during the sample run, and at the time of sample collection. For example, a notable sampler/station condition may be power loss and/or sampler damage found upon arrival due to a recent storm.
    a. Ensure supporting records, such as logbooks and data forms, are complete and accurate, and operator commentary is descriptive.
    b. Perform and document corrective actions, if needed.
    c. If observations impact data, ensure they are earmarked for the next level reviewer.
4. Review documentation regarding atmospheric conditions at the time of sample set-up, during the sample run, and at the time of sample collection. For example, a heavy rain event on a sample collection day may result in an extremely low particulate concentration; therefore, this known weather condition would be important information for the Level 2 reviewer.
    a. Ensure supporting records, such as logbooks and data forms, are complete and accurate, and operator commentary is descriptive.
    b. If observations potentially impact data, ensure they are marked (highlighted) for the next level reviewer.
5. Visual inspection of sample media.
    a. If the sample filter is received from the laboratory with visible damage or imperfection, this should be immediately documented and a decision made regarding its use. The laboratory may need to be contacted to request a replacement filter.
    b. If the sample filter is damaged during transport or upon collection in the field, it should be documented, along with a description of how the damage occurred (if known). Any necessary corrective actions should be documented. *Note: Photographs of the damaged sample filters should be taken as a best practice.*
    c. Similarly, the use of make-up samples should be thoroughly documented so the next level reviewer understands what transpired in the field. The Level 1 reviewer should provide a suggested null data code that best fits the reason why the sample was not collected on the scheduled run day (per the EPA sampling calendar). For example, if the sample media was damaged and a replacement unavailable prior to the scheduled run date, the Level 1 reviewer could suggest the missed run be coded with “AJ” (i.e., filter damage) or “AF” (scheduled but not collected).

6. Review documentation to ensure all activities and observations which could impact sample integrity are detailed and descriptive. Events occurring at or near the monitoring site, such as construction, prescribed burns, or source-facility maintenance, are important details that should be captured and marked for the next level data reviewer. *Note: The Level 1 reviewer (site operator) can recommend a sample be "void" based on known issues that bias the sample results*.
7. Review chain-of-custody documentation for completeness and accuracy.

Monthly Report

Parameter: O3 44201 October 2018 Avg Interval: 1 hour Units: PPM 007 Method: 047

Hours

Day	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	Summary		
																									Max	Avg	RDS
01	.003	.003	.002	.001	BF	.001	.009	.017	.023	.031	.041	.043	.046	.045	.046	.044	.044	.040	.020	.012	.007	.000	.000	.000	.046	.020	23
02	-.001	-.001	-.001	-.001	-.001	-.001	-.001	.000	.010	.026	.038	.043	.044	.041	.040	.036	.028	.022	.010	.002	-.001	-.001	-.001	-.001	.044	.013	24
03	-.001	-.001	-.001	-.001	-.001	-.001	-.001	.000	.009	.025	.039	.043	.049	.054	.057	.060	.058	.044	.024	.015	.004	.001	.000	-.001	.060	.019	24
04	-.001	-.001	-.001	-.001	-.001	-.001	-.001	.000	.001	.019	.034	.042	.049	.051	.050	.050	.050	.040	.023	.006	.001	.000	-.001	-.001	.051	.016	24
05	-.001	-.001	-.001	-.001	-.001	-.001	-.001	.000	.008	.021	.048	.064	.063	.064	.065	.072	.067	.056	.031	.017	.007	.002	.000	.006	.072	.024	24
06	.012	.009	.005	.001	.000	.000	-.001	.000	.008	.024	.042	.050	.052	.047	.043	.038	.034	.025	.013	.012	.008	.005	.010	.009	.052	.018	24
07	.008	.010	.008	.004	.001	.000	-.001	.002	.015	.030	.040	.041	.042	.040	.038	.036	.037	.036	.029	.018	.012	.021	.024	.023	.042	.021	24
08	.024	.019	.013	.015	BF	.015	.015	.017	.022	.028	.029	.030	.031	.028	.032	.030	.021	.016	.003	.004	.007	.016	.020	.024	.032	.019	23
09	.026	.028	.028	.026	.024	.023	.021	.014	.019	.026	.030	.033	.036	.036	.031	.024	.025	.021	.016	.016	.020	.019	.020	.021	.036	.024	24
10	.020	.018	.015	.013	.011	.008	.006	.007	.009	.011	.015	.017	.018	.017	.019	.019	.015	.013	.014	.013	.013	.015	.017	.017	.020	.014	24
11	.017	.016	.016	AV	AV	AV	AV	AV	AV	AV	AV	AV	AV	AV	AV	.025	.028	.029	.031	.032	.031	.031	.030	.027	.032	.026	12
12	.025	.023	.021	.017	.013	.011	.007	.009	.013	.023	.028	.032	.037	.039	.040	.038	.034	.026	.008	.000	-.001	-.001	-.001	-.001	.040	.018	24
13	.000	.009	.007	.009	.007	.007	.019	.021	.025	.031	.035	.038	.039	.040	.041	.042	.042	.035	.015	.007	.001	.000	.000	.001	.042	.019	24
14	.003	.004	.006	.005	.008	.008	.006	.010	.018	.025	.035	.043	.043	.045	.045	.042	.042	.036	.015	.008	.003	.001	.001	.001	.045	.018	24
15	.001	.000	-.001	-.001	BF	-.001	-.001	.000	.009	.012	.035	.037	.039	.039	.040	.038	.035	.022	.011	.004	.003	.008	.012	.007	.040	.015	23
16	.003	.000	.000	.000	.000	.005	.009	.003	.008	.018	BF	BF	.035	.039	.038	.038	.040	.030	AV	AV	.000	.000	.018	.015	.040	.014	20
17	.015	.011	.005	.003	.001	.000	.000	.005	.013	.014	.021	.025	.030	.030	.030	.031	.030	.023	.016	.017	.021	.025	.023	.021	.031	.017	24
18	.027	.029	.028	.026	.027	.027	.026	.025	.028	.031	.035	.038	.043	.046	.046	.047	.046	.043	.035	.031	.021	.025	.025	.016	.047	.032	24
19	.015	.013	.014	.012	.010	.009	.004	.015	.023	.027	.033	.039	.041	.043	.042	.036	.028	.025	.018	.021	.018	.019	.018	.013	.043	.022	24
20	.015	.012	.021	.023	.024	.024	.022	.020	.021	.022	.024	.021	.021	.024	.023	.024	.024	.024	.025	.023	.024	.027	.026	.026	.026	.022	24
21	.029	.027	.025	.023	.029	.028	.029	.032	.034	.034	.034	.034	.036	.036	.036	.036	.035	.030	.013	.007	.001	.000	.000	.000	.036	.024	24
22	.000	.000	.000	.000	BF	.000	.000	.000	.005	.019	.027	.034	.037	.039	.040	.041	.040	.024	.006	.000	.000	.000	.000	.000	.041	.013	23
23	.000	.000	.000	.000	.000	.000	.000	.001	.004	.011	.020	.028	.032	.042	.042	.038	.034	.020	.007	.007	.001	.000	.000	.000	.042	.011	24
24	.000	.000	.000	.004	.010	.006	.002	.002	.013	.028	.036	.042	.046	.050	.052	.052	.051	.042	.012	.012	.014	.016	.014	.012	.052	.021	24
25	.014	.013	.017	.016	.011	.017	.017	.015	.021	.026	.032	.033	.031	.031	.030	.029	.027	.023	.023	.023	.024	.023	.023	.023	.033	.022	24
26	.021	.020	.019	.020	.018	.018	.016	.014	.015	.016	.015	.013	.016	.017	.017	.017	.018	.020	.020	.019	.018	.015	.017	.017	.021	.017	24
27	.017	.013	.012	.010	.006	.008	.007	.008	.010	.012	.016	.020	.022	.028	.031	.036	.033	.030	.027	.022	.019	.025	.023	.021	.036	.019	24
28	.020	.019	.019	.017	.018	.017	.015	.012	.022	.030	.034	.037	.037	.038	.037	.037	.036	.030	.018	.011	.033	.034	.034	.033	.038	.026	24
29	.033	.032	.032	.030	BF	.023	.019	.020	.024	.030	.036	.039	.040	.040	.040	.040	.038	.021	.004	.000	.001	.000	.000	.000	.040	.023	23
30	.000	.000	.000	.000	.000	.000	.000	.002	.012	.028	.034	.036	.038	.040	.043	.043	.043	.023	.006	.000	.000	.000	.000	.000	.043	.014	24
31	.000	.000	.000	.000	.000	.000	.000	.000	.003	.011	.024	.040	.043	.048	.051	.054	.044	.020	.006	.001	.000	.000	.000	.000	.054	.014	24
Max	.033	.032	.032	.030	.029	.028	.029	.032	.034	.034	.048	.064	.063	.064	.065	.072	.067	.056	.035	.032	.033	.034	.034	.033	.072		
Avg	.011	.010	.009	.009	.008	.008	.008	.009	.014	.022	.031	.035	.037	.039	.039	.038	.036	.028	.016	.012	.010	.010	.011	.010		.019	
Count	31	31	31	30	25	30	30	30	30	30	29	29	30	30	30	31	31	31	30	30	31	31	31	31			723

Figure 21: Example Monthly Report for Ozone

Level 1 Review of Data Trends

In addition to daily data review, data should also be reviewed **weekly to monthly** to look for trends and patterns in the data that may not have been obvious when only assessing a single day's worth of data. Knowledge of expected data patterns helps reviewers distinguish actual measurements from problem data. The monitoring organization's data management system (software) should provide some type of summary report that will allow the reviewer to see the hourly averages from the previous week or month. Figures 16 and 21 provide examples of monthly data reports for a continuous monitor. To perform the Level 1 weekly/monthly review for data trends, access and view the time period of interest for the specific monitor per the instructions in the data review SOP. Evaluate the data and records with the following in mind:

- Scan the data set for any missing data, outliers, or anomalies that may not have been identified during the daily review.

- Re-review minute data (in strip chart format) to watch for trends or shifts in analyzer response that were not apparent when reviewing one day's worth of values.
- Review control charts to see if any new trends or patterns are revealed (see Figure 22 for an example).
- Review logbook notations in conjunction with the data set, to ensure accuracy of data coding, and to see if annotations reveal new issues.
- Verify documentation on all spreadsheets, QA/QC data forms, and/or supporting data reports.
    - Is documentation complete and accurate?
    - Does it convey everything the next level data validator needs to know?

Note: Control charts enhance both site operations and data verification activities. The visualization of data on a control chart illustrates trends in analyzer performance, such as slow drift or consistent bias in one direction, that may not be apparent when reviewing data in a table. The charts also quickly identify outliers, which illustrate data points that require closer review. Control charts can be easily created by the site operator (data reviewer) using the DAS or other software (such as Microsoft Excel™, etc.).

It is imperative that results from the Level 1 data review be documented. The monitoring organization's QAPP and data review SOP should detail these requirements. Level 1 daily review may be documented in a variety of ways, including annotations in a logbook or within the DMS software, digitally on the daily summary reports (saved to PDF), or manually on a hard copy of the daily printout, to name a few. The documentation should attest that Level 1 review was completed, with a signature/initials of the reviewer and date, along with any questions or comments for the next-level reviewer. Monthly reviews can be summarized in more formal reports, if preferred, but in all cases the documentation should include what data were reviewed, who did the review, when, and any details, especially of corrective action.

Finally, it is important to emphasize that all data review staff are reliant on the effective communication and documentation of prior reviewers. With that in mind, the documented Level 1 review begins this critical communication chain. The end product of Level 1 verification goes to the designated Level 2 reviewer (and so on); comments and concerns identified during the review need to be clearly explained. Documentation should be sufficient such that the next-level reviewer can "reproduce" how decisions about the data were made. After all Level 1 data review activities have been completed, the data and associated reports/documentation should be organized and transferred to the designated next-level reviewer.

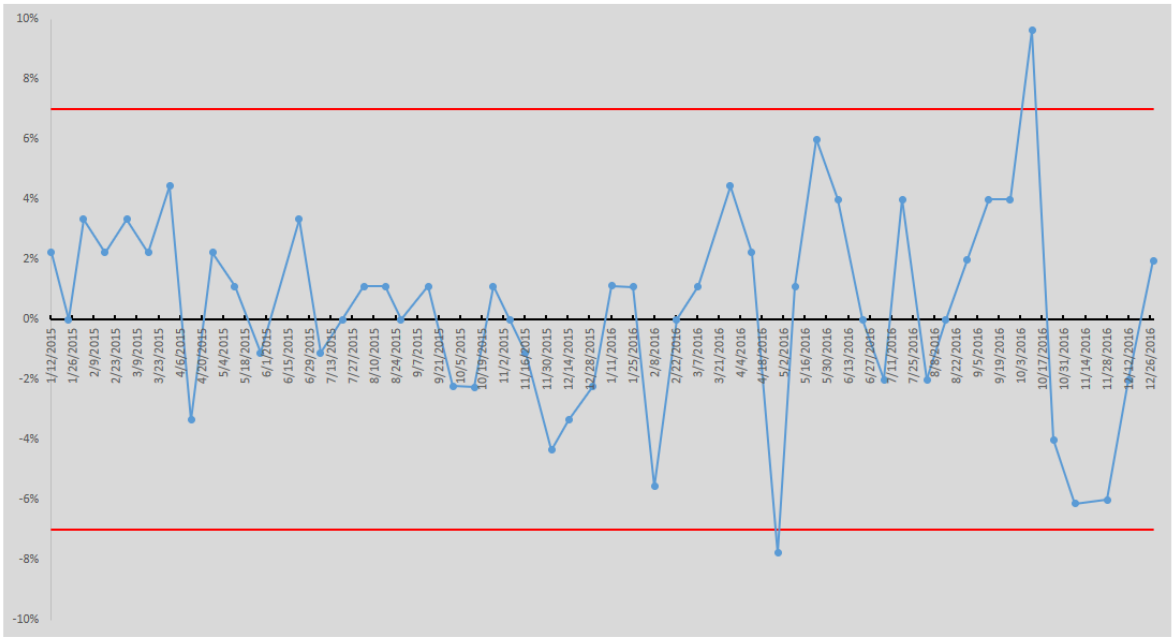


Figure 22: Example Control Chart Plotting Results of Biweekly QC Checks

3.2.3 Level 2 Data Review

As stated earlier, verification and validation often overlap in what stages of data review they cover. Level 2 review typically begins the validation stage of the review process, as more QA/QC data is available to the data reviewer; verification is also performed during Level 2 review. To ensure data defensibility, procedures for data validation should be handled completely separately and independently from data collection. **Therefore, the most important distinction between the Level 1 and Level 2 review is the independence of the Level 2 reviewer.** Level 2 reviews should be performed by someone other than the site operator (technician). (See Figure 9.) An independent reviewer can bring an unbiased perspective and potentially find issues that were missed in the previous review. See Section 2.1.1 of this document for more information, including example scenarios to achieve independence in data review when there are staffing limitations.

Ideally, Level 2 reviews should be conducted monthly and quarterly. The goals of Level 2 data review include:

- Verifying the Level 1 review occurred properly and there is sufficient documentation to support decision making; and,
- Ensuring data meets QA/QC requirements and the objectives of their intended use (**validation**).

Level 2 review initially mimics components of the Level 1 review in that certain data verification steps are completed, but these are done primarily to confirm the accuracy and completeness of the Level 1 review. Afterwards, the Level 2 review goes beyond the verification steps and also validates the monitoring data. It is important to note that validation starts at the monitor-level and determines if data from the individual monitor, at a specific moment in time, produced usable results. Towards that end, it is

important to note that data validation includes looking at the **"fitness for use"** (i.e., data usability) of the data collected and ensuring that it can be used in the ways intended (see Section 1.2 of this document). "Intended use" refers to the monitoring objectives found in the QAPP. The question of "fitness for use" of the data should always be kept in mind by the Levels 2-3 data reviewers. The QAPP includes specifications on data collection elements including sampling design; sample collection; sample handling and custody; QC procedures; calibration procedures; analytical procedures; and data processing procedures. For each of these, the data reviewer must ask questions and make judgment calls regarding the usability of the collected data. For example, if a particulate monitoring station is located adjacent to (and downwind of) a temporary building construction project, and atypical concentrations are observed after construction start-up, would these concentrations be representative of ambient concentrations (neighborhood scale exposure) or would they be demonstrating the localized impact of the construction project? Similarly, if a Pb sampler is intended to collect data at a microscale level outside a facility fence-line to determine ambient Pb concentrations near a local elementary school, but the data reviewer learns that the sampler is positioned such that it is upwind of the facility, does the collected data meet the monitoring objectives? **Keep in mind that some data may be useful for some objectives, but not for others.**

The Level 2 reviewer compares data to QA/QC requirements prescribed in the QAPP and SOPs, and flags or invalidates data that do not meet these criteria. To inform the Level 2 review, the data validation templates (i.e., MQO tables / control limits) should be utilized, and the requirements compared to the actual QA/QC records. Strip charts can identify the frequency, correctness, and stability of QA/QC activities. Because data from a longer time period is reviewed, trends and patterns in the data should be more easily identifiable. If the Level 1 data reviewer had any concerns, the Level 2 reviewer needs to confirm that those concerns were answered and documented. It is imperative that the data flagging and validity decisions by the Level 2 reviewer be thoroughly documented. **All data reviewer judgments about the data must be supported by evidence and <u>not</u> assumptions.** Level 2 data review needs to be as consistent and objective as possible, so that data from different years can be compared, even when data reviewers change.

Another notable distinction between the Level 1 and Level 2 reviews is that only a **percentage** of data are reviewed by the Level 2 reviewers, whereas the Level 1 reviewer evaluates 100% of the data. Generally speaking, the Level 1 reviewer evaluates data on a daily basis, which makes the workload more manageable. However, the Level 2 reviewer is likely reviewing the data packages from multiple operators – and considerably larger data sets (i.e., months and/or quarters). Therefore, the Level 2 reviewer can only be expected to review a reasonable percentage of the data and its supporting records. The amount established will likely be determined based on the number of staff available for Level 2 activities, the sophistication (complexity) of the Level 1 review, and/or the number of instruments in the monitoring network. With this in mind, the percentage of data reviewed during Level 2 activities may vary across monitoring organizations.

In order to ensure consistency in process, however, a clear description of what the Level 2 review specifically entails and the targeted amount of data to be evaluated should be prescribed in the monitoring organization's data review SOP. At a minimum, the Level 2 reviewer should evaluate 100% of the data for completeness (i.e., missing data) and perform an in-depth review of 100% of the data earmarked by the Level 1 reviewer as needing additional review in order to make a judgment call on data validity. Similarly, the Level 2 reviewer should evaluate all QA/QC data forms (or similar) for which a verification signature is required. After these initial steps have been completed, the Level 2 reviewer should perform

an in-depth review of the supporting documentation, records, and underlying minute data for a lesser percentage of the collected data. There are a variety of strategies that could be used to designate the targeted percentage; they will differ based on whether it is continuous or intermittent data under review. Ultimately, the monitoring organization has the flexibility to determine the amount selected and how it is implemented, given their capabilities and resources. For illustration purposes, the strategy employed for continuous monitoring data could be to review the data and records for approximately 7-8 days out of the month (i.e., ~25% of the total hours in the month), focusing on the highest concentration days. Another approach could be to perform an in-depth review on **all** the days for which the Level 1 reviewer has suggested AQS null codes or qualifiers be applied. This latter approach would yield a variable amount of data reviewed monthly, but would serve to ensure correct decisions have been made and documentation exists to support decision-making. Under this second scenario, for months when little to no data is flagged by the Level 1 reviewer, the Level 2 reviewer would need to select additional, random days to augment the review. Other scenarios are possible and can be discussed with the appropriate EPA Regional Office, if needed. Many monitoring organizations target ~25% data review for Level 2 activities, which EPA encourages as a best practice.

The steps that follow focus on continuous monitoring data and serve as a guide to achieve the Level 2 data review goals stated above. The monitoring organization's data management system (software) should provide some type of summary report that will allow the Level 2 reviewer to see the hourly averages for the month/quarter. To perform the Level 2 review, access and view the time period of interest for the specific monitor per the instructions in the data review SOP. Appendix A of this document contains a Data Verification Checklist (tool) that can also assist the Level 2 reviewer when completing the review.

Recommended Level 2 Review Approach:

Part 1: Verify the accuracy and completeness of the Level 1 review. For example:

1. Look for any gaps in data collection (i.e., missing data) during the month.
    a. If gaps are found, determine the cause of data loss and select the appropriate AQS null code to reflect the reason for data loss.
    b. If data can be restored, ensure the concentration reported is accurate and document the reason for data validity.
2. Review the appropriateness of AQS null and qualifier codes recommended by the Level 1 reviewer. Ensure the documentation and records available, including strip charts and corrective action reports, support the application of the codes. *Note: The data management software may add color to data that meet certain specifications or highlight data that have been modified in the database for any reason. Figure 21 shows an example of a monthly report for ozone with coloration: the AQS null codes are color-coded dependent on the specific code; very low concentrations and negatives in the month are shaded pale yellow; and the highest concentration of the month is shaded gray.*
    a. If necessary, select another code(s) that is more appropriate and document the reason why. Earmark this data for concurrence by the Level 3 reviewer.
    b. Communicate coding changes with the Level 1 reviewer. Further discussion may be warranted, especially if the Level 2 reviewer does not concur that the data needs to be qualified or invalidated.
3. Verify the maximum hourly/daily concentrations. Compare the value to the strip chart to ensure values are not the result of a QC activity or instrument malfunction.

4. Look for negative readings or constantly repeating values during the month. Did the Level 1 reviewer provide explanations for their cause? Is data jeopardized because of an underlying issue?
5. Ensure data has not been post-processed to correct for failing QC or zero/span drift.[36]
6. Look for outliers, or unusual and undocumented patterns.
    a. If identified, does the documentation from the Level 1 review (or other available information) discern whether the outlier is a real data point or one that should be invalidated?
    b. If needed, compare concentrations to nearby and downwind sites. Are other monitors reading similarly?
    c. If needed, compare concentrations to historical averages for the specific monitor and time period under review. Do the concentrations reasonably compare or are they markedly different?

If errors are found in the Level 1 data review, the Level 2 reviewer can either review a higher percentage of data (e.g., more than 25%) and/or return the package to the Level 1 reviewer for a second review.

Part 2: Ensure data meets QA/QC requirements and intended use. For example:

1. Review any data marked by the Level 1 reviewer as questionable/suspect, as well as any data for which the Level 1 reviewer indicated additional review is warranted. Make a judgment call(s) on the affected data's validity and select the appropriate AQS codes, if needed.
2. For the selected percentage of data to be reviewed in-depth, ask: Is the data comparable and representative of ambient conditions?
    a. Was the instrument(s) operated in accordance with its SOP (for example, see Step #5, below).
    b. Would any issues identified in the field impact the data's "fitness for use"? For example, if documentation indicated that the sample line to the instrument had been disconnected inside the shelter and the reported concentrations during the affected time period were abnormally low (confirming dilution), then the instrument was sampling shelter air instead of ambient air. The impacted data would need to be invalidated.
    c. Compare the results of collocated instruments.
        i. If there is a significant difference in the concentration values of the two instruments, investigate the supporting records and documentation for the data pair to further assess data validity. In some cases, one of the two instruments may need to be recalibrated or may have malfunctioned. For intermittent samplers, one of the samples may be impacted by contamination, for example, or could be a field blank inadvertently reported as a sample filter. The data review SOP should define a targeted threshold value between collocated sampling results (e.g., percent difference, absolute unit difference, etc.), whereupon exceedance triggers in-depth review and investigation.
        ii. Document the outcome of the investigation and determine which AQS codes are needed, if applicable.

[36] See the QA Handbook (2017), Sections 10.4, 12.2. and 12.5

d. Compare the results of nearby sites to determine the reasonableness of data. See Figure 23 for an example of data tracking demonstrated when plotting nearby monitors on a time-series graph. Figure 23 illustrates two monitors tracking concentrations similarly; the two pollutant traces generally display the same trends. However, from June 5-8, there is a large discrepancy between the monitors. Upon further review, the monitor with the red chart trace was sampling shelter air, which explains the visible anomaly.

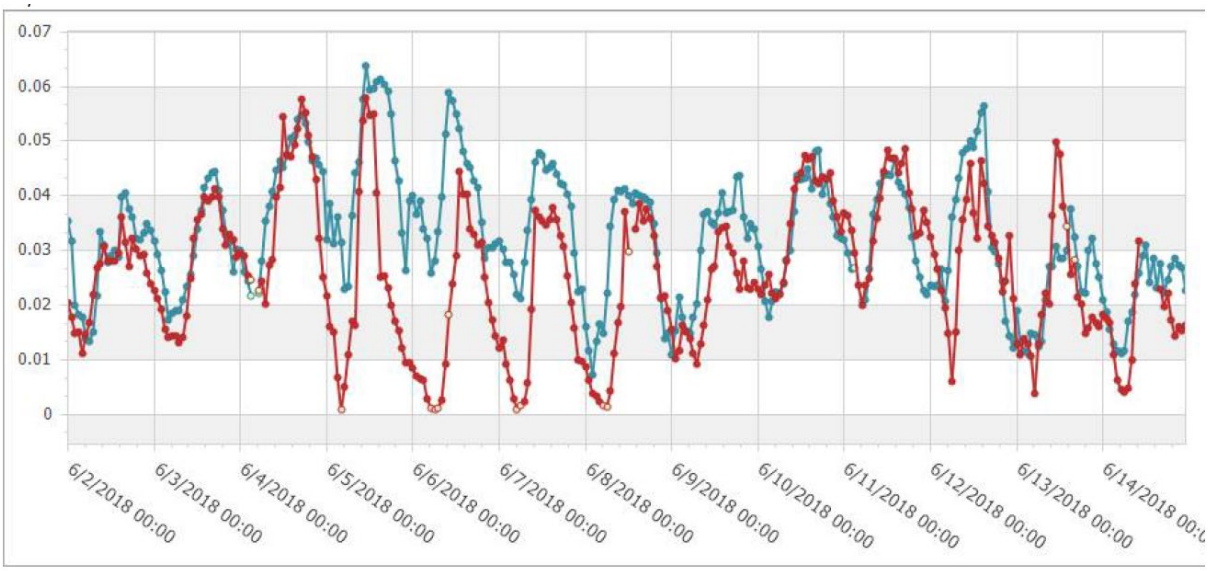


Figure 23: Time-Series Graph that Compares Concentrations of Nearby Monitors

3. Ensure FRM/FEM designation specifications have been met for NAAQS-comparable instrumentation.
4. Ensure calibrations, QC checks, and performance evaluations (audits) were performed using NIST-traceable equipment that was "in certification" (i.e., not expired).
    a. If expired standards were used, then the audits or QC checks performed with that equipment would likely be considered invalid. The procedural error would need to be further investigated to determine its impact, if any, on the concentration data.
        i. The "1C" code would need to replace an invalid QC check of a gaseous analyzer when reported to AQS. *(Note: The "1C" null code should not be used to invalidate concentration data. Rather, the "1C" code is used to report an invalid QC check in an AQS QA transaction.)* The data reviewer may need to qualify associated ambient concentration data to be transparent about the procedural error if multiple QC checks were determined to be invalid.
        ii. If the error caused a significant amount of QC data to be questionable (e.g., a quarter or more of QC checks were jeopardized), the monitoring organization should consult the EPA Regional Office.
    b. If calibrations were performed using expired standards, it is recommended that the monitoring organization reach out to its EPA Regional Office for consultation.
5. Determine whether the operator adhered to the SOP for the pollutant under review.
    a. Access the electronic strip chart to confirm that calibrations and QC checks were performed correctly and concentrations were allowed appropriate time to stabilize. *Note: For gaseous monitors, QA/QC stability will appear as "walkable stair steps" on the*

electronic strip chart. (See Figures 11 and 15 for examples.) However, if walkable "stair steps" are seen in the routine ambient concentration data (as opposed to QA/QC data), it could indicate an instrument issue in the field. (See Figure 19 for an example.)

b. Confirm that the correct concentrations were generated during the gaseous QC activities (i.e., as stated in the SOP and required by the QAPP and CFR), or that the appropriate flow rates were tested for particulate monitors.
c. Confirm that the correct number of QC checks were performed during the time period under review and that they were spaced appropriately. For example, if the SOP requires manual QC checks be performed every 7 days, does the supporting documentation demonstrate that the checks were performed ~7 days apart?
d. Determine whether the "order of activities" performed by the operator adheres to the SOP. For example, if the SOP specifies that an as-found QC check is performed on site prior to any maintenance activity, and/or that an as-left QC check is performed following any maintenance activity, does the documentation support that these activities occurred? (*Note: Code the hours such that they represent the order of operations performed. If more than one activity occurs in a single hour, determine which activity took the greatest number of minutes to complete. See Section 3.1 of this document.*)
e. For any procedural deviation that is observed during Steps A-D above, determine whether the non-conformance has impacted the associated concentration data. If so, select the appropriate AQS null or QA qualifier flag and apply the code such that it appropriately brackets the data.
    i. For example, the Level 2 reviewer may determine that the site operator did not adhere to the SOP requirement of performing checks every 7 days; however, an automated ZPS is performed nightly, which is more frequent than requirements in the CFR. To be transparent about the SOP deviation, the Level 2 reviewer adds a QA qualifier code of "6" to the data from the time period when the QC check was expected until it was completed.
f. Confirm that instrument maintenance and other SOP-required procedures were performed as directed. Review maintenance logs and other records, as applicable.

6. Verify calculated computations (such as percent differences, linear regression, etc.) are correct on QA/QC forms (or similar). Then, compare the results of the instrument calibrations, QC checks, and QA audits to QA/QC specifications in the QAPP/SOPs to look for anomalies or failures.
    a. If the results of QC checks exceeded the established acceptance criteria, does documentation show whether an investigation was performed to determine if the QC check itself was valid or invalid? If not, investigate the QC check validity, which may involve communicating with the site operator (Level 1 reviewer).
    b. If the check was invalid, flag (replace) the QC check with the "1C" code. If the check was valid, determine the impact to the associated concentration data, null-code/qualify in accordance with the OAQPS directive[37], and bracket the data appropriately.
7. If available, review the results of any external performance audits (such as NPAP). Were outcomes acceptable? If not, why? Is data impacted?
8. If available, review the results of any external/internal systems audits. Do non-conformances identified impact data integrity? If so, determine whether data should be qualified or invalidated.

[37] https://www.epa.gov/sites/production/files/2018-01/documents/critical_criteria_qualifier_memo_v1_0.pdf

9. Review documentation associated with the specific data set under review to ensure records and commentary are complete, accurate, descriptive, legible (if handwritten), and, where appropriate, signed and dated.

As a final step, after data have been examined as discussed in Steps 1 – 9 above, the Level 2 data reviewer should do a final cross-walk of the data against the remaining line items in the data validation templates. This final step ensures the selected data under review satisfies the necessary requirements and that its supporting documentation justifies data validity decisions and associated AQS coding (if applicable). *Note: The DQOs, many of which are in the systematic criteria (blue section) on the templates, are based on annual summary statistics of validated data; comparison to the DQOs would occur during Level 4-5 review, when data assessments, such as annual certification, are performed.*

Additional Review for Intermittent Samplers:

The Level 2 review process for intermittent data, such as $PM_{2.5}$ FRM or Pb samples, follows the same basic concepts as described above, except that Level 2 review also includes data and records received from the analytical laboratory. That said, whether the analytical laboratory is an in-house laboratory operated by the monitoring organization directly or by one of its governmental partners, or is an outside laboratory that has been contracted to perform the analyses, the monitoring organization should receive analytical QA/QC data along with the analytical results, in order to perform the necessary validation steps. Contract laboratories will be discussed more in Section 3.2.4.1 of this document.

When validating intermittent sampling data, the data and records from the pre- and post-sampling activities must be combined for each individual sample to determine overall validity; however, this information is available to the data reviewer at different time periods. Level 1 review would include evaluating the data and records as they first become available (primarily from the field), whereas Level 2 review would include looking at the pre- and post-field records and the laboratory information simultaneously. As stated previously, the in-depth review of documentation, records, and data during Level 2 data review should occur on only a percentage of the collected data. Therefore, for intermittent sampler data, as a best practice, EPA suggests the highest concentration and a random concentration for each sampler/month be minimally selected for in-depth review.

The Level 2 review should occur on a monthly, or at least, quarterly basis. An example Level 2 review approach for intermittent samples includes repeating the Part 1, Steps 1-5, and Part 2, Steps 1-9, above, assessing both the field and laboratory data/records to the best extent possible, along with the following additional steps:

1. Compare final sample concentrations to the national EPA sampling schedule to ensure there is either a concentration reported or an AQS null value code for each scheduled run date.
2. Verify the calculated final concentration for the sample is correct. For example, for $PM_{2.5}$ FRM, this calculation would require obtaining the initial and final filter weights (laboratory), as well as the volume of air sampled (field). *Note: EPA guidance recommends at least 7% of computations be verified manually in order to ensure correctness*[38].
    a. If errors in calculations are observed, the Level 2 reviewer should verify the math for a higher percentage of samples in the batch. If gross reporting errors are observed, the Level 2 reviewer should determine the source of the computational error (if possible) and

[38] https://www.epa.gov/sites/production/files/2021-03/documents/p100oi8x.pdf

take corrective action, which may include contacting the laboratory and/or field staff to address the source of error.

b. Determine whether concentrations need to be re-reported. Document all outcomes.

3. Review any available sampler summary files and/or minute data (5-minute interval files, e.g.) available from the sampler for the 24-hour sampling event. Ensure field requirements have been met. If not, select the appropriate AQS codes.
4. Review documentation of sampler/station conditions, weather conditions, and other observations from the time of sample set-up, during the sample run, and at the time of sample collection. Determine whether any of the noted observations impact data quality. Select the appropriate AQS codes, if needed.
5. Review documentation regarding sample integrity from the Level 1 reviewer and/or from the laboratory. Request photographs of damaged sample media from the site operator or the laboratory, if available. Make a judgment call(s) on sample integrity and select the appropriate AQS codes, if needed.
6. Review chain-of-custody documentation, including field and laboratory information. Determine whether any of the noted observations impact data quality. Similarly, observe custody seal documentation (if required by the monitoring organization as a best practice) and note whether the seal was intact or broken. Select the appropriate AQS codes, if needed.
7. Review the results of field, trip, and lab blanks, as well as Pb audit strip data, all of which should be provided by the laboratory. Do **trends** indicate potential contamination or other issues in the field or laboratory (e.g., problems with static control, etc.)? Control charts can help with this evaluation. If so, determine if associated concentration data should be qualified in AQS. *Note: If trends or issues are observed with the blank data, only the associated concentration data needs to be qualified in AQS. The blank data reported to AQS should not be null-coded or qualified.*
8. Review the QA/QC data from the laboratory. Ensure critical analytical method requirements were met for the sample batch containing the specific sample under review.
   a. Using the data validation templates as a tool, cross-walk the laboratory QA/QC data against the line items specified in the data validation templates as **laboratory critical criteria** for the pollutant under review, **at a minimum**.
   b. Compare the operational and systematic criteria as well, as available, which should include access to NIST traceability certificates for laboratory standards and equipment.

Some laboratories will provide the monitoring organization with an AQS-ready text file that contains the monthly/quarterly results for all sites in the network. **Under no circumstances should the monitoring organizations upload this file to AQS without first completing Level 2 and 3 reviews of the intermittent data.** Although the laboratory providing the AQS-ready text file to the monitoring organization is an acceptable practice, it does not substitute for completing the necessary validation of the intermittent data. The monitoring organization remains accountable for ensuring that the laboratory analysis meets regulatory and method requirements. Therefore, it is critical that the data reviewer perform Steps 7-8 above each month/quarter to ensure the analytical MQOs were met, at a minimum. After completion of this review, changes to the AQS-ready text file provided by the laboratory may be necessary. If changes are made, they should be documented.

During Level 2 (and/or Level 3) data review activities, the data reviewer should determine the most descriptive and appropriate null code or qualifier(s) to apply to the intermittent data, when validation results indicate such is warranted. This means looking at the overall results of the data/documentation, which includes combining the field and laboratory information. The data reviewer may be tempted to

immediately invalidate or qualify data with AQS flags upon observing lab-applied codes in the analytical data package, especially those that may translate to "lab error". However, given that the analytical laboratory is charged with analyzing all samples received regardless of status, there will be times when samples received at the laboratory are "void" upon arrival but still analyzed. With this in mind, the data reviewer must consider the site operator's comments and field documentation and select an AQS code that best represents the primary reason why the sample is questionable or invalid. For example, if the site operator documents that the field sampler did not maintain the appropriate flow rate throughout the 24-hour sample run, the sample would still be analyzed at the laboratory. Then, for example, if a lab qualifier is added to this same sample (such as one that indicates the analyst observed an imperfection (pinhole, spot, or blemish) on the sample filter), the data reviewer would need to take both issues into account and select the one that is more significant with regards to overall validity. In this example, the data reviewer should invalidate the sample using a code that represents the known field issue (i.e., AH null code, meaning "Sample Flow Rate or CV Out of Limits"), as opposed to applying a qualifier that would indicate that filter damage was the reason for invalidation. This specific sample was technically invalid upon arrival at the laboratory, but it was not the lab analyst's responsibility to make that determination. That responsibility falls to the monitoring organization.

Finally, it is imperative that results from the Level 2 data review be documented. The monitoring organization's QAPP and data review SOP should detail this requirement. Moreover, the data review SOP should prescribe how the Level 2 data review will be documented by the reviewer, including the format and required contents of the data package that will be prepared for and transferred to the designated Level 3 reviewer. The documentation should attest that Level 2 review was completed, with a signature/initials of the reviewer and date, along with any questions or comments for the designated next-level reviewer. Any email communications that discussed data validity concerns or was used to justify validity decisions for specific data under review should be included in the package. Documentation should be sufficient such that the next-level reviewer can "reproduce" how decisions about the data were made. After all Level 2 data review activities have been completed, the data and associated reports/documentation should be organized and transferred to the designated next-level reviewer.

3.2.4 Level 3 Data Review

The Level 3 review concludes the validation phase of the data review process. The Level 3 review should be performed by someone independent from the data collection activities, such as another independent reviewer, the monitoring organization's QAM, or in some cases, a program manager (see Figure 9). It is important to note that, being completely independent, a Level 3 reviewer may not have a comprehensive, technical understanding of each individual monitoring method and associated instrumentation. However, the Level 3 reviewer should be proficient in understanding how decisions are made with the data by the monitoring organization and its external partners, such as EPA (see Section 1.2). That said, the Level 3 reviewer may view the collected monitoring data through a lens that is significantly different from that of the previous reviewers.

Validation must occur before data are uploaded to AQS (see 40 CFR 58.16). Therefore, to meet federal data reporting requirements, Level 3 review must occur on a quarterly basis, at a minimum. More frequent review (e.g., monthly) is strongly encouraged. The goals of Level 3 data review include:

- Verifying the Level 1 and 2 reviews and supporting documentation;

- Ensuring data are accurate, complete, comparable, representative, and defensible, given the supporting documentation (validation); and,
- Approving data suitability for release to AQS.

Generally speaking, more information, including QA/QC data, is available to the Level 3 reviewer, especially when evaluating data quarterly. Therefore, additional outliers, trends, and patterns may be apparent during the Level 3 review. The Level 3 review includes evaluating supporting documentation and flags added during prior reviews. It is the final chance for errors to be discovered and fixed, and to ensure that there is documentation to justify decisions on data validity, prior to AQS data submittal. As the final validation stage, the Level 3 review focuses heavily on the data's fitness for use, corresponding to the objectives stated in the monitoring organization's QAPP. Review activities generally include data comparisons, trends evaluations, and graphical analyses (such as those described in the Level 2 discussion, among others). During Level 3 review, data should also be assessed in terms of the DQIs (i.e., precision, bias, completeness, comparability, representativeness, sensitivity); with that in mind, this review stage may overlap with some assessment activities discussed in Section 4. As the final validator, the Level 3 reviewer should ensure validity decisions can withstand public and legal challenges, and that final data packages contain the necessary documentation and records to support those decisions, should such challenges arise. At its conclusion, the Level 3 review provides final confirmation that data are valid, based on available information at the time of the review, and can be used in decision making. The Level 3 reviewer then approves the release of the data for subsequent upload into the AQS database.

The Level 3 review follows the same general approach as the Level 2 review, except that a smaller percentage of data and records are evaluated. For example, if the Level 2 review included an in-depth review of supporting records and information for ~25% of the collected data, then the Level 3 review may include a similar in-depth review for only 10% of the data. The percentage of data reviewed during Level 3 activities – and precisely how that percent review is implemented – may vary across monitoring organizations; it is often dependent upon the size of the monitoring network and the resources available to perform the previous levels of review. Therefore, it is important that the percentage established for Level 3 review be clearly defined in the monitoring organization's QAPP and data review SOP. Moreover, a clear description of what the in-depth review specifically entails, and steps on exactly how to perform it, should also be included in the data review SOP.

The Level 3 reviewer should utilize the data validation templates and the QAPP when performing the review. Other tools/resources (e.g., strip charts, control charts, the DASC tool, other spreadsheets, etc.) may be used to graphically analyze the data and evaluate it for trends. Appendix A of this document contains a Data Validation Checklist (tool) that can assist the Level 3 reviewer when completing the review as well.

To perform the Level 3 review, access the monitoring data and supporting documentation as prescribed in the monitoring organization's data review SOP. Then, in general, **repeat the steps for the Level 2 review detailed above**. The first part of the review should verify the accuracy and completeness of the Level 1 and 2 reviews. **If significant errors or inconsistencies are found in the Level 2 data review, the Level 3 reviewer can either review a higher percentage of data (e.g., more than 10%) and/or return the package to the Level 2 reviewer for a second review.** The second part of the Level 3 review should include an in-depth review of a percentage of data/records. Lastly, the data is reviewed to ensure it has the appropriate AQS null codes and qualifiers and is suitable for release to AQS.

Recommended Level 3 Review Approach:

1. Evaluate 100% of the data for completeness (i.e., missing data). There should be no gaps in the data submittal to AQS.
2. Spot-check the AQS null codes and qualifiers that have been applied to the data by the previous reviewers for accuracy and consistency.
3. Consider the "story" that the AQS coding presents when looking at the data from a monthly to quarterly perspective. Investigate any apparent anomalies in coding (e.g., recommendations of unexpected or uncommon codes).
4. Validate 100% of the data points marked by the Level 2 reviewer as questionable or needing further evaluation (such as complex situations where a weight of evidence judgment call is needed). Ensure outcomes are appropriately documented.
5. Ensure data has not been post-processed to correct for failing QC or zero/span drift.
6. Conduct an in-depth review of the supporting documentation, records, and underlying minute data for a selected percentage of the concentration data. Cross-walk the records and information against the MQOs in the data validation templates.

Throughout these six steps, the Level 3 reviewer should keep in mind these questions:

- Has any new objective evidence (additional documentation, QA/QC data, TSA reports, etc.) been collected since the Level 2 review that impacts decisions made by prior reviewer(s)? If so, document the justification for changing the validity decision(s).
- Is the data usable for its intended purpose?
- Is the audit trail for the data complete?
- Is the data (validity decision) defensible, given the documentation and records included in the data package (objective evidence)?

Finally, with regards to AQS-release, the Level 3 reviewer should keep in mind these additional data handling questions as well when reviewing the data:

- Was the same set of rules followed by the different Level 2 reviewers for the different sets of data?
- Is data coding consistent? (Have reviewers utilized the same codes for similar situations, in accordance with the data review SOP?)
- Is there continuity in coding? (Meaning, if an issue that spans months has impacted data, does the appropriate AQS coding continue from one month to the next?)
- Are validity decisions consistent? (Meaning, were similar situations with similar outcomes handled in the same manner, with any resulting AQS codes applied consistently?)

Because data submitted to AQS can be immediately used by EPA, other external entities, researchers, and the public, it is imperative that the data uploaded to AQS be thoroughly reviewed and validated prior to AQS submittal. The data review SOP should prescribe how the completed Level 3 data review will be documented, including signatures/initials of the Level 3 reviewer, along with the date of the review. The SOP should also explain how notification will be communicated and documented to inform the monitoring organization's AQS data submitter that data is ready for upload into AQS. The monitoring organization is encouraged to retain the AQS upload files for future reference.

3.2.4.1 Data Validation and Analytical Laboratories

The data validation process for intermittent data, such as $PM_{2.5}$ FRM or Pb samples, should include a review of field and laboratory records in tandem, in order to ensure CFR and method requirements have been fully met. The data validation templates for the intermittent pollutant methods include individual critical criteria sections for both field and laboratory activities, so a lot of information is needed to confirm the collected samples meet all regulatory and method requirements. With that in mind, the Level 2-3 review frequency will likely be dictated, to some degree, by the laboratory's data reporting schedule and subsequent receipt of analytical data packages by the monitoring organization.

With regards to the analytical data, one of the most important validation steps that should be performed by the monitoring organization is ensuring that the laboratory is indeed utilizing an FRM or FEM for its analytical method. Although this may sound like an unnecessary step, there have been multiple findings during EPA TSAs in recent years where an analytical laboratory has been found to not be utilizing an FRM or FEM for the analysis of ambient air samples. These findings have resulted in either invalidation of large data sets or costly reanalysis of samples (when possible given filter holding time requirements). For example, some laboratories have utilized water methods to analyze criteria Pb samples; although similar in technique, the water methods ultimately do not meet the regulatory requirements prescribed for ambient air analysis. Therefore, to avoid this serious issue, confirmation of the analytical method employed by the laboratory is critical. Ideally, this first step should be performed prior to establishing any agreements (contracts) for analytical services; or, if utilizing an in-house laboratory, prior to beginning analysis. The monitoring organization should obtain a copy of the laboratory's QA manual (e.g., QAPP or similar), as well as the analytical SOP for the specific method, to verify the procedures. In particular, the monitoring organization is strongly encouraged to crosswalk the laboratory SOP against the FRM/FEM to verify compliance (for example, compare the SOP procedures to those in 40 CFR Part 50, Appendix G, if the laboratory is utilizing the Pb FRM for its analytical method). Similarly, if the laboratory is utilizing an FEM – such as one of the numerous Pb FEMs available – then the monitoring organization should obtain a copy of the FEM (likely available from EPA OAQPS) and then crosswalk the SOP against the FEM. **When utilizing an FEM, in order to be in compliance, the laboratory must follow the FEM verbatim and not make any modifications to the analytical procedure.** (The monitoring organization is encouraged to consult with their respective EPA Regional Office if there are any concerns about the specific analytical method.) **Subsequently, during the Level 2-3 reviews of the intermittent data, the data reviewer should spot-check data packages and the supporting laboratory documentation to ensure continued compliance with the analytical FRM/FEM method requirements.** The monitoring organization is encouraged to develop a data review checklist (or similar) for intermittent data that includes confirmation that the analytical method utilized remains an FRM/FEM throughout the duration of the monitoring effort.

It is also important to note that a multi-step (i.e., secondary and tertiary) data review process should occur at the analytical laboratory prior to the laboratory releasing data packages to its customers. The tiered review structure is often a requirement of various laboratory accreditations; however, this is not always the case. The monitoring organization should be aware of the data review strategy employed by the analytical laboratory and know specifically what their data review entails. Frequent and routine communication with the analytical laboratory is essential. Although the laboratory will likely review analytical sample batches to ensure the QA/QC requirements of the analytical method are met, it is highly unlikely that the laboratory will evaluate the data further for compliance with ambient air monitoring regulations. Additionally, the analytical laboratory will likely not have access to the supporting records

and documentation for the field activities associated with the samples, such as the results of monthly flow rate verifications and so forth, which are vital to successful review. **Therefore, the final validation of the ambient air monitoring data – incorporating records and data from both the field and laboratory operations, reviewing them in tandem – must be performed by the monitoring organization.**

To facilitate successful validation of data where laboratory analysis has been performed, the monitoring organization is encouraged to do the following:

- Establish an upfront agreement that specifies data packages received from the analytical laboratory will include the QA/QC data from the analytical batches from which the samples were run/analyzed.
    - Make sure the QA/QC data received from the laboratory includes all of the elements that are defined in the data validation tables as **critical criteria for laboratory activities**. If the packages do not contain all the necessary information for Level 2-3 reviewers to confirm MQOs have been met, contact the laboratory and request the missing information be added to the data package. *Note: In some instances, the request for additional information in data packages will incur additional fees.*
    - Request to receive the QA/QC sample data considered operational criteria in the data validation templates, such as the results of laboratory blanks.
    - Ensure that copies of NIST traceability certificates are accessible, or hardcopies provided at least annually, so that the monitoring organization can confirm laboratory standards and equipment are in good order during sample analysis.
    - Ensure data packages provide explanations of any laboratory-applied codes or flags. Laboratory-applied qualifiers may or may not have the same meaning/implications as AQS codes. Therefore, it is important the Level 2-3 reviewers have a clear understanding as to what lab-applied qualifiers mean, and be able to translate those flags to the appropriate AQS codes when necessary. ***Note: The application of lab-applied qualifiers does not always necessitate the application of an AQS code. When in doubt, confer with the laboratory QA liaison to discuss why the flag(s) was applied and how it impacts data quality. Or, consult with the EPA Regional Office.***

- Regularly communicate with the laboratory, especially the QA liaison, and ask questions when anomalies or issues are observed in the data packages.

It is important to note that it is **not** the responsibility of the analytical laboratory (or laboratory analyst) to make determinations on data validity for the monitoring organization. The analytical laboratory will analyze all samples received from the monitoring organization, unless the samples are damaged to the point where analysis is physically impossible. The laboratory is accountable for documenting observations of damaged samples, or those that were received with known issues; however, it is the monitoring organization's responsibility to invalidate or qualify samples upon receipt of the analytical results.

3.2.4.2 Post-AQS Data Verification

After the Level 3 review has been completed – and data validity, accuracy, and completeness confirmed based on the available information – the data is approved for release to AQS. After the AQS upload is completed, however, verification should not stop. Instead, as a final review step, various AQS reports

should be generated to verify the success of the data upload. This final review should include spot-checking that:

1) all data submitted, including AQS null and qualifier codes, were successfully and accurately entered (i.e., transmitted);
2) all QA/QC data reported were successfully and accurately entered;
3) units of measure and method/instrument codes are correct for all data reported, including QA/QC data; and,
4) typographical errors are not present in any data manually entered, especially QA/QC data.

Section 4 provides a listing of several helpful AQS reports, along with recommendations on how they can be used during post-AQS data verification.

EPA recommends, at a minimum, that the AQS AMP 350, "Raw Data Report", which shows the hourly/daily concentration values for the sites/monitors, and the AQS AMP 251, "QA Raw Assessment Report", which shows the results of all QA/QC checks, be generated following AQS data entry to confirm the accuracy of the upload. The monitoring organization is encouraged to retain these AQS reports for future reference, and sign/initial/date them, to attest to the final review and accuracy of the reported data in AQS. The monitoring organization should determine which data reviewer(s) will be responsible for completing this final check and include the responsibility in the data review SOP.

4.0 Overall Assessment of Data Quality

After ambient air monitoring data have been validated and uploaded to AQS, the next stages of data review include performing various statistical assessments and data quality audits. As discussed in Section 1.2 of this document, assessments are evaluation processes used to measure the performance or effectiveness of a system and its elements. With regards to data quality, assessment is the process of evaluating the *aggregated* data set's ability to meet the intended objectives (i.e., DQOs). Assessments can occur on a quarterly, annual, or multi-year basis, when larger sets of data are available for evaluation. With regards to achievement of criteria pollutant DQOs, assessment statistics are typically calculated for annual and 3-year time periods. For some pollutants, however, EPA assessments can include evaluating data from a 5-, 6-, or 10-year perspective. QA/QC data can be statistically assessed at various levels of aggregation (e.g., monitor level, PQAO level, nationally). A noteworthy difference between assessments and validation is assessments can be performed by persons external to the monitoring organization.

Statistical assessments of monitoring data can begin as early as the Level 3 review stage, although such evaluations are limited to a quarterly level of aggregation. Figure 9 illustrates an expanded data review strategy that goes beyond the Levels 0 – 3 structure discussed in Section 3 of this document. During the Levels 4-5 data review stages, quarters to years of data are assessed. To implement this ideal review structure, additional resources and personnel may be needed. However, many monitoring organizations do not have the resources to staff additional data review tiers within their organizational structure. Therefore, the Level 3 reviewer (e.g., the QAM or Program Manager) is often the individual charged with performing the tasks that are associated with these expanded review levels (see Figure 9), including activities such as database verification and audits of data quality. In some cases, the tasks are performed by external data users, such as the EPA. Because this document is primarily designed to assist monitoring organizations in performing data verification and validation, in-depth guidance and instruction on performing assessments will not be provided here. Instead, this section will offer only a general overview, including brief discussions of annual data certification and audits of data quality (ADQ). This section will

also highlight some readily-accessible tools the monitoring organization can use. Other EPA documents and training courses, such as APTI SI-470, are available to provide additional guidance on performing assessments.

Once validated data are uploaded to AQS, the monitoring organization is immediately provided with a number of additional, helpful reports (tools) that can be utilized to further assess data quality. The various AQS AMP reports can help identify issues or trends that may have not been observed when performing Levels 0-3 data review procedures. The following list suggests common AQS reports that may be helpful to the QAM or data review staff, and provides a description of the report's purpose and potential use. It is important to note that this list is not all-inclusive, nor does it suggest that each report listed must be generated. Other AQS reports are available for review which may be beneficial to the QAM or data review staff.

- AMP251 QA Raw Assessment Report

  This report lists the results (i.e., % difference) of each individual QC check performed for the pollutant of interest. It also includes the results of performance evaluations (presented to reflect the 10 audit concentration range levels), NPAP/PEP audits, lead audit strip analyses, and collocation assessments. Because many monitoring organizations prepare QA/QC results manually for AQS submission, this report is helpful to review to ensure no typographical errors were made when processing the data. The report can be reviewed to ensure the correct reporting units were used for the listed parameters. Additionally, the report will calculate the percent differences for the data values entered into AQS, and so can be used to cross-check calculations on the monitoring organization's QA/QC data forms. When reviewing the AMP 251 results, any anomalously high percent differences should be further investigated. For instance, any extremely high percent differences noted in the collocation assessments should be reviewed. Also, collocation is used to calculate aggregate particulate precision; high percent differences between data pairs could be an indicator of field / sampler issues, and warrants a closer review.
- AMP256 QA Data Quality Indicator Report

  This report calculates summary statistics at both the monitor and PQAO levels for the QC checks performed by the organization. It is used to determine whether or not the organization is meeting the established DQOs for the criteria pollutants. A companion document that explains the AMP 256 statistics can be found on AMTIC.[39] This report is most useful when reviewing a full calendar year's worth of data, although it can be reviewed anytime.
- AMP350 Raw Data Report

  This report shows hourly ambient concentrations for the continuous analyzers and samplers in the monitoring network, as well as concentration results for the intermittent particulate (PM and Pb) samples (i.e., 24-hour samples) collected by the organization. This report can be used to verify reported AQS codes are correct and to see if the data, reading the codes, "tells the correct story" (as discussed in Section 3 of this document). It is important to note that the AMP 350 will only show one AQS qualifier flag per individual hourly or 24-hour concentration. If multiple qualifier flags have been added, the data reviewer will need to generate an AMP 501 report (see below) in order to verify that all applicable flags have been appropriately added. For intermittent data, the data

[39] https://www.epa.gov/sites/production/files/2016-09/documents/boxplots_companion-generic_v2_9_9_16.pdf

reviewer will be able to see the "pattern" the 1-in-3, 1-in-6, or 1-in-12 day sampling schedule present in the data and, from that, should be able to spot if data have been reported on the wrong sampling schedule or were not reported (i.e., missing samples). Also, the report will make more readily visible particulate concentrations reported that are less than 1.0 µg/m^3. Any $PM_{2.5}$ sample that is less than the federal LDL of the sampler (i.e., 2.0 µg/m^3, per 40 CFR 50, Appendix L) should be further reviewed to ensure it is a valid sample concentration.

- AMP350MX Raw Data Max Values Report

  This report provides the highest concentration value for each day for the pollutant of interest. This report is helpful for reviewing 5-minute SO_2 data, particularly for those organizations that report only the highest 5-minute average from each hour. (Compare to the AMP 501 report below.) For hours where SO_2 concentrations have been invalidated in AQS, the corresponding 5-minute data would also need to be invalidated. The data reviewer could compare an AMP 350 to the AMP350MX to cross-check the SO_2 data to ensure this has occurred. Also, the data reviewer could use this report to identify the highest concentrations in the data sets and then further review those specific hours/days to ensure they are real, representative ambient concentrations.

- AMP430 Data Completeness Report

  This report calculates monthly statistics, and provides quarterly or annual data completeness statistics for each monitor in the specific pollutant network, depending on the data range criteria selected. If unexpected data completeness statistics are computed for a monitor that either began or discontinued data reporting, the monitor "Begin" and/or "End" date(s) should be reviewed for accuracy in the AQS monitor metadata by querying an AMP390 Monitor Description Report.

- AMP480 Design Value Report

  This report calculates the design value for each site in the network for most criteria pollutants. From that, the data reviewer can determine the overall highest concentration sites in the network. Monitoring organizations may want to consider performing ADQs or other reviews of these specific sites, to ensure data reported is accurate and defensible.

- AMP501 Extract Raw Data

  This extract produces a text file that contains the organization's ambient concentration data in raw format. This report will show all AQS qualifier flags that have been applied to the data. This extract may be helpful for reviewing SO_2 5-minute data, particularly for those organizations that report 12, 5-minute blocks of data for each hour.

- AMP503 Extract Sample Blank Data

  This extract produces a text file that contains the organization's $PM_{2.5}$ field and/or trip blank data, if the monitoring organization operates a manual (FRM) $PM_{2.5}$ network. $PM_{2.5}$ field blank data is required to be reported to AQS in accordance with 40 CFR 58.16; however, the submittal of trip blank data to AQS is optional. The data reviewer can review the text file as is, or import it into Excel so it can be more easily assessed. The best practice is to control-chart the field blank data in order to quickly identify trends which could point to data collection issues, such as sampler contamination in the field. The data reviewer should look for trends and investigate the results of any blanks that exceed the $PM_{2.5}$ acceptance criterion (i.e., ±30 µg). It is not uncommon for field blank samples and actual field samples to be inadvertently mislabeled (or "swapped") when handling samples in the field. Data reviewers should pay close attention to high field blank results and compare those values to the actual field samples collected during the

same time period; if the corresponding sample concentrations are extremely low (such as values less than 1.0 ug/m^3), a handling error could have occurred.

- AMP504 Extract QA Data

  This extract produces a text file that contains the organization's QA/QC data in raw format. The text can be imported into Excel, in order to more easily sort and review the data; or, the data can be imported into a similar data assessment program that will sort the data, such as the 504 QA Data Assessment Tool available on AMTIC.

- AMP600 Certification Evaluation and Concurrence Report

  This is the primary report generated by organizations during annual data certification. The report shows whether DQOs have been met for the criteria pollutants, provides dates for QAPP approvals, and illustrates other QA considerations.

In addition to these AQS reports, other tools that can be helpful in assessing data quality include control charts, box-and-whisker plots, the EPA DASC tool, and the aforementioned 504 QA Data Assessment Tool. Examples of these quality indicator assessment reports can be found on the AMTIC website[40]. It is important to note that control charts and the DASC tool may have been utilized during Level 2-3 data review activities to identify outliers or other issues in small data sets. The distinction to keep in mind here is, for big picture assessments, these same tools would be used to evaluate validated data over the long-term (quarters-to-years), looking for marked biases in the quality system (PQAO / network) and calculating statistics such as coefficient of variation (CV), which is <u>not</u> used as an acceptance criterion at the analyzer level.

Data reviewers can also access the EPA Air Data webpage[41], where more EPA tools are readily available to help visualize data trends. These tools use the data that have been uploaded to AQS. The following (Figure 24) is an example of a box-and-whisker plot generated online using one such tool, the Single Point Precision and Bias Report[42]. The box-and-whisker plots[43] in Figure 24 help the data reviewer quickly see there are outliers at several monitoring sites which should be further reviewed (particularly, Sites 0035 and 0004). For a number of years, EPA produced annual box-and-whisker plots of the gaseous pollutants using this specific tool and posted them to AMTIC. OAQPS's goal is to perform data quality assessments for the pollutants of the Ambient Air Quality Monitoring Network at a yearly frequency for data reports and at a 3-year frequency for more interpretative reports.

40 CFR Part 58, Appendix A, Section 4, *Calculations for Data Quality Assessment*, details the specific calculations that are used to statistically assess the monitoring data QA/QC data. The regulation stipulates quarterly, annual, and triannual aggregation when performing some of the calculations. Section 4 to Appendix A further states:

> *Calculations of measurement uncertainty are carried out **by the EPA** according to the following procedures. The PQAOs must report the data to AQS for all measurement quality checks as specified in this appendix even though they may elect to perform some or all of the calculations in this section on their own.* [Emphasis added]

[40] https://www.epa.gov/amtic/ambient-air-monitoring-quality-assurance
[41] https://www.epa.gov/outdoor-air-quality-data
[42] https://www.epa.gov/outdoor-air-quality-data/single-point-precision-and-bias-report
[43] https://www.epa.gov/sites/production/files/2016-09/documents/boxplots_companion-generic_v2_9_9_16.pdf

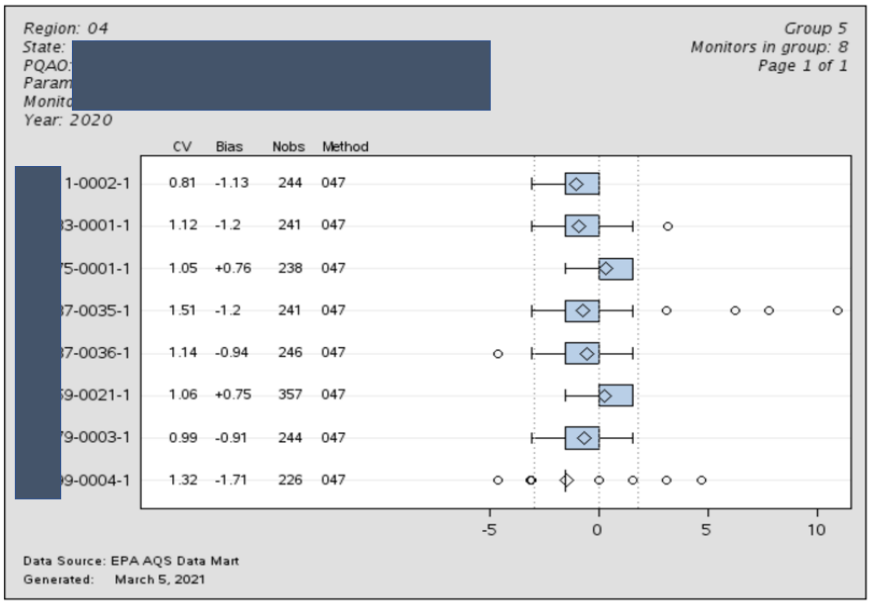


Figure 24: Example Box-and-Whisker Plot Generated Using EPA Online Reports

Section 4 to Appendix A further specifies that EPA will provide annual assessments of data quality aggregated by site and PQAO for SO_2, NO_2, O_3 and CO, and by PQAO for PM_{10}, $PM_{2.5}$, and Pb. The AQS reports listed above, particularly the AMP 256 and AMP 600, can quickly and easily calculate the statistics that the CFR requires. (The DASC tool also has the ability to perform these calculations.) The results presented on these reports can then be compared to the DQOs that are codified in 40 CFR 58, Appendix A, Section 2.3.1. It is important to emphasize that the DQOs are data quality **goals**, and stated as such in the CFR. The DQOs were set by the EPA, in collaboration with the SLT monitoring organizations, as a result of the systematic planning process for each pollutant (see Figures 1 and 25). If the calculated CV for the aggregate data set does not meet the CFR requirement, that does not mean automatic invalidation of data. Instead, **it serves as an indicator of a quality system issue(s)** that should be further investigated and remedied. However, depending on the egregiousness of the results, and/or the quality system issue(s) those results illuminate, EPA reserves the right to not use a monitoring organization's data, based on EPA's assessment of the data (see 40 CFR 58, Appendix A, Section 1.2.3). As a best practice, monitoring organizations should discuss elevated CVs with their EPA Regional Office before data certification.

For clarification, the precision (i.e., CV) and bias estimators for the gaseous pollutants are based on the aggregation of the results of single-point QC checks performed throughout the year. For the particulate pollutants (i.e., PM and Pb), precision (or CV) is estimated using collocated sampler data pairs greater than the minimum concentration specified in 40 CFR 58, Appendix A, Section 4, whereas bias is primarily estimated using PEP and Pb audit strip analyses results (although flow rate verification and/or audit bias may be estimated, as well). It is important to note that the AMP reports and DASC tool can be used at any time, but if they are generated when there is less than one quarter's worth of QC data available, the results will be misleading. Precision and bias calculations are best performed when there are more data available – such as quarters to years. The precision and bias calculations should **not** be used to determine whether individual QC checks are valid; moreover, the precision and bias calculations are **not** designed to affirm the validity of the concentration data during the intervals between QC checks. The

precision and bias calculations are designed to estimate measurement uncertainty for the *aggregate* data set **at the project level** (e.g., all QC checks for a single monitor over the course of a year; all QC checks for a single pollutant over the course of a year for the entire PQAO). (See Figure 5.) Also, for clarification, the results of annual PEs and NPAP/PEP audits are evaluated as statistical averages and should be performed when there are sufficient data available to perform the required calculations. However, the results of individual performance audits and NPAP/PEP audits *can* be used to inform decisions at the monitor-level at the time of the audit, and therefore, should also be reviewed on an individual basis during Level 2 or 3 review activities.

The final level of assessment (see Figure 9) includes such tasks as network reviews, user-needs evaluations, and reconciliation of DQOs, with these latter two often occurring in tandem. These assessments evaluate large data sets (e.g., annual, tri-annual, 5-year) and are usually performed by the monitoring organization's QAM and program planners.

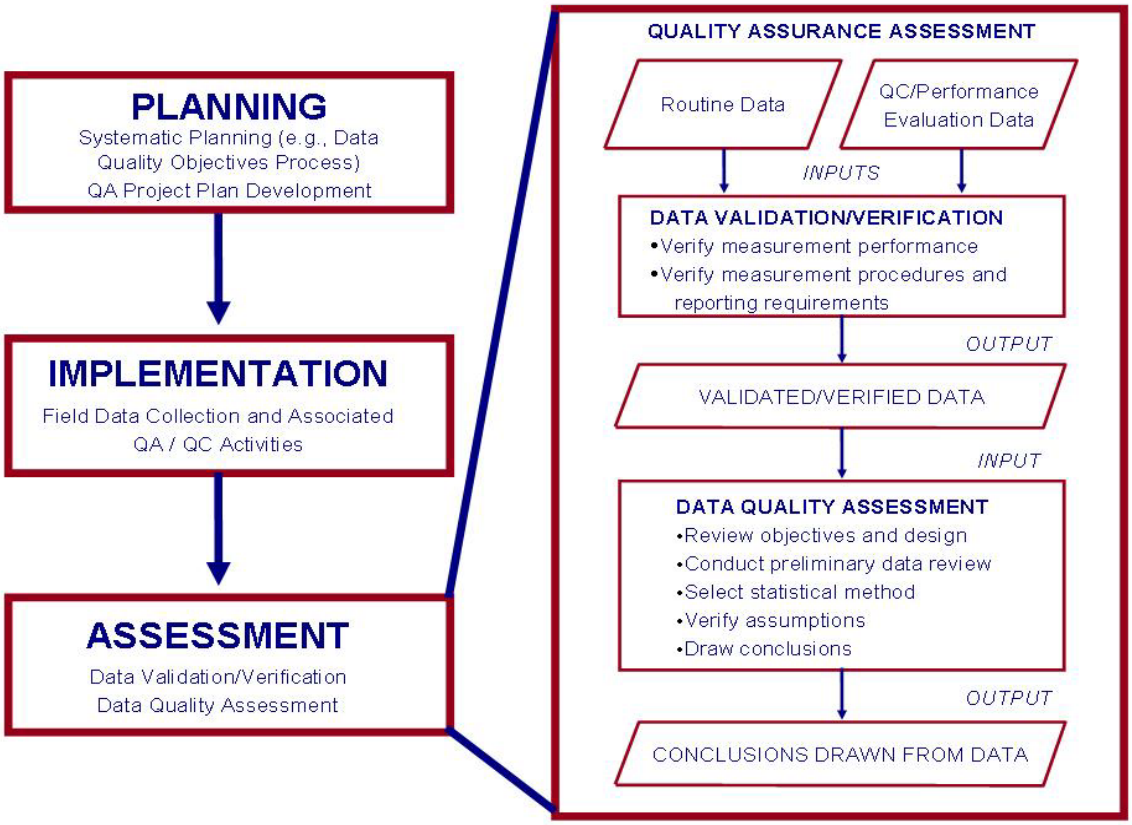


Figure 25: Data Quality Assessment in Context of the Data Life Cycle

Reconciliation represents the completion of the quality cycle, and is where quality system improvements are considered and recommendations made to update DQOs. With that in mind, reconciliation includes an evaluation of the aggregated data set's ability, in combination with the specified objectives of the project's ability, to meet the needs of the end data user. Reconciliation of DQOs is a required element in Category I air monitoring QAPPs (see Appendix C of the QA Handbook (2017)); all monitoring organizations operating NAAQS-compliant monitoring networks must have a mechanism in place that serves this function. The EPA guidance document *Data Quality Assessment: A Reviewers Guide* (EPA

QA/G-9R)[44] and the QA Handbook (Section 18.1) discuss a 5-step DQA process that is helpful when reconciling DQOs at the end of a project, or at defined intervals (such as annually). This formal DQA includes the scientific and statistical evaluation of environmental data to determine if they meet the planning objectives of the project, and as such, are of the right type, quality, and quantity to support their intended use. DQA is built on a fundamental premise: *data quality* is meaningful only when it relates to the *intended use* of the data, which in many cases stems from the DQOs. DQAs can be used to determine whether modifications to the DQOs are necessary, or "tighter" quality control is required. **QAPPs are often revised as a result of reconciliation** (see Figure 1). Figure 25 illustrates DQA in context of the data life cycle. The 5-steps of the DQA are included as bullets in the figure. The figure illustrates: 1) verification and validation are part of QA assessment of data; 2) validated data are critical inputs for DQAs to help data users evaluate whether and how data can be used for decision-making purposes; and 3) the outputs of assessments are conclusions drawn from the validated data. EPA is responsible for setting the DQOs that are codified in 40 CFR Part 58, Appendix A, and is also charged with performing DQAs. Although enhancements to AQS over the years have provided reports, such as the AMP 600, which assist monitoring organizations in quickly performing annual assessments of monitoring data, the monitoring organizations are encouraged to perform DQAs as well.

Ultimately, from data verification to data validation to DQA, each step in the data review process benefits from and builds on the previous one. Together, they assure achievement of the ultimate goals of the ambient air monitoring program: credible data and sound, defensible decisions.

4.1 Audits of Data Quality (ADQ)

The QAM or designated data reviewer should be tasked with performing periodic ADQs. An ADQ includes reviewing supporting documentation and records, in order to ensure the data reported to EPA is accurate, traceable, and defensible. It is usually performed in conjunction with an internal systems audit, but can be performed separately. An ADQ can be a time-consuming process, but is designed to ensure a solid "audit trail" exists for the data evaluated. To perform the audit, the QAM (or other data reviewer delegated this responsibility) selects a limited number of data points to scrutinize. Through this process, the QAM will evaluate whether or not the monitoring organization is validating its data in accordance with its QAPP/SOPs and EPA requirements. These few data points, then, will be used to generally surmise the quality of the overall data sets, through confirmation of the effectiveness of the organization's quality system and its documentation and recordkeeping practices.

The EPA is charged with performing an ADQ during TSAs. For more information about how the EPA performs an ADQ, see the EPA *Quality Assurance Guidance Document (QAGD) Conducting TSAs of Ambient Air Monitoring Programs* (November 2017). The monitoring organization can use the techniques described in the TSA QAGD as a guide to enhance its data review strategies.

4.2 Annual Data Certification

Annual data certification is a process performed by the monitoring organization that is required pursuant to 40 CFR 58.15. The data is certified at the PQAO level. The monitoring organization identified as the PQAO is responsible for the oversight of the quality of data of all monitoring organizations within the

[44] https://www.epa.gov/quality/guidance-data-quality-assessment

PQAO, pursuant to 40 CFR Part 58, Appendix A, Section 1.2.1. Therefore, the PQAO will submit the required annual data certification package to EPA on behalf of all organizations operating under the PQAO.

In some monitoring organizations, the Level 3 data reviewer may be the same individual who completes annual data certification procedures for the PQAO. Because of this dual role, it is not uncommon for the Level 3 reviewer to look at data from a "certification" perspective when completing quarterly Level 3 reviews. Once data has been entered into AQS, any monitoring organization staff member can pull a variety of AQS reports to further check and assess the data, as described in Section 4 above. The AQS reports are additional tools that can help reviewers spot issues in the data that may not have been identified previously; with this in mind, the monitoring organization is strongly encouraged to routinely utilize these reports. However, the AQS reports are not intended to, nor can they replace, the validation techniques that are described in Section 3 of this document.

During the data certification process, the monitoring organization assesses **validated** data. It is a confirmation that the data for the previous calendar year has been reviewed and deemed acceptable by the monitoring organization. All data collected by FRM, FEM, and ARM monitors at SLAMS or SPM monitoring stations that are required to meet 40 CFR Part 58, Appendix A, must be certified. That includes monitoring data for CO, NO_2, SO_2 (hourly and 5-minute averaged data), O_3, Pb, PM_{10}, $PM_{2.5}$, and $PM_{10\text{-}2.5.}$ If the monitoring organization's validation process has been robust and thorough throughout the year (utilizing the review strategies and best practices provided in Section 3 of this document), then data certification should be a quick and easy process for the monitoring organization.

When the head official in a monitoring organization, or the official's designee, submits a formal data certification letter along with other necessary material described below, the official certifies that the previous year's ambient concentration data and all of the QA/QC data that were collected, have been completed and passed the monitoring organization's data validation process and have been submitted to AQS. With this letter, the official also confirms that the ambient concentration data are accurate to the best of his or her knowledge, taking into consideration the QA findings. This means, the official has considered the results of periodic QC checks and has determined that any other relevant performance reviews meet regulatory requirements and data quality requirements specified in the monitoring organization's QAPP(s). This formal letter attesting to ambient data completeness and accuracy must be submitted by May 1st of each year for data collected the previous calendar year.

Along with the submittal of the signed data certification letter, an agency is also required to provide the AMP 600 data certification report, which is a summary report of all the ambient air quality data collected by FRM, FEM, and ARM monitors at SLAMS and SPM sites. This report serves as the record of the specific data that is the object of the certification letter, and it contains a summary of precision and bias data, as well as a summary of data completeness, for all ambient air quality data to be certified. This report also assesses the data that has been certified and identifies if there are quality assurance or data completeness issues associated with the certified data. Also required as part of the data certification submittal is the AMP450NC for $PM_{10\text{-}2.5}$ and 5-minute SO_2 data.

Following submittal of this data certification package, EPA Regional Office staff will review all submitted materials to assure completeness and adherence to CFR requirements. EPA will review the assessments made as part of the AMP600 described above and apply "yes" or "no" flags to the data in AQS to indicate that EPA has evaluated the certified data and has (or has not) identified data

completeness or QA issues with the certified data. However, EPA evaluation of that certified data does not mean that EPA has completed any validation of the agency's data. **Data validation is the sole responsibility of the monitoring organization.** Furthermore, the application of a "Y" or a "N" to this data has no effect on it being "certified." Data is certified when the head official in the monitoring organization signs a letter pursuant to 40 CFR 58.15 saying that data is certified. It is important to note that before the submission of these materials on May 1st, EPA presumes that monitoring organizations may still be reviewing and validating data, but after this deadline, EPA may move ahead and use the most current, three complete years of data available to propose and make designations or findings of attainment. EPA does not typically use AQS data in broadly distributed publications until the deadline for certification has passed. Ultimately, annual certification gives EPA (and the public) formal permission to use the data for a variety of purposes, including determinations of attainment/nonattainment relative to the NAAQS.

Additional guidance on data certification and the setting of the "certification evaluation flags" is available on the AMTIC at this address: https://www.epa.gov/amtic/data-certificationvalidation. There is also a module dedicated to annual data certification in the APTI SI-470 training course.

5.0 References

40 CFR Parts 35, 50, 53, and 58

2 CFR Part 1500

48 CFR Part 46

EPA, *EPA Quality Policy CIO 2105.1*, 03/31/21

EPA, *EPA 2185 – Good Automated Laboratory Practices*, USEPA, 8/10/95

EPA*, Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity, of Information Disseminated by the Environmental Protection Agency*, EPA/260R-02-008, October 2002

EPA*, Quality Assurance Handbook for Air Pollution Measurement Systems, Volume II, Ambient Air Quality Monitoring Program*, EPA-454/B-17-001, January 2017

National Environmental Laboratory Accreditation Conference (NELAC), *2003 NELAC Standard*, EPA/600/R-04/003, June 5, 2003

The NELAC Institute (TNI), 2016 Laboratory Standards

EPA, *Guidance on Environmental Data Verification and Data Validation, EPA QA/G-8*, EPA/240/R-02/004, November 2002

EPA, *Guidance on Data Quality Assessment, Practical Methods for Data Analysis, EPA QA/G-9*, EPA/600/R-96/084, July 2000

EPA, *EPA Requirements for Quality Management Plans, EPA QA/R-2*, EPA/240/B-01/002, March 2001

EPA, *EPA Requirements for Quality Assurance Project Plans, EPA QA/R-5*, EPA/240/B-01/003, March 2001

EPA, *Guide to Writing Quality Assurance Project Plans for Ambient Air Monitoring Networks,* EPA-454/B-18-006, August 2018

EPA. SW-846 Chapter Nine, *Sampling Plan*, September 1986

Taylor, John K., NBSIR, *Principles of Quality Assurance of Chemical Measurements*, NBSIR 85-3105, US Department of Commerce, National Bureau of Standards, February 1985

ISO/IEC, *General Requirements for the Competence of Testing and Calibration Laboratories*, ISO/IEC 17025, May 15, 2005

ASQ/ANSI, *E4: Quality Management Systems for Environmental Information and Technology Programs (E-Standard)*, 2014

EPA, *NEIC Policies and Procedures Manual*, EPA-330/2-9-78-001R, May 1978 (Revised August 1991)

EPA, *NEIC Procedures Manual for Contract Evidence Audit and Litigation Support for EPA Enforcement Case Development,* EPA-330/9-89-002, February 1989

Appendix A: Data Verification and Validation Checklists

This appendix provides example checklists that can be used by Level 1 – 3 data reviewers as a comprehensive guide to verify and validate ambient air monitoring data on a monthly or quarterly basis. The checklists included in this appendix have been developed for the gaseous pollutants only and reflect the guidance available at the time of this publication. Checklists for particulate pollutants may be developed and posted to AMTIC at a later time.

The checklists generally mimic the data validation templates found in Appendix D of the *QA Handbook* (2017); questions are associated with specific critical, operational, and systematic criteria. The checklists also list several other questions of significance for the data reviewers to consider. For monitoring organizations in the early stages of building a data review program, the checklists offer a ready tool to facilitate and document a tiered data review approach. The checklists could also be utilized as a training tool for monitoring organizations, offering a consistent, step-by-step guide to teach data reviewers the numerous monitoring requirements that must be evaluated. Likewise, the checklists could assist QA staff when performing ADQs. If needed, users can further customize these example checklists to add more information, such as unique requirements that may be emphasized in an agency's QAPP/SOPs.

General Instructions for Use:

The "*Data Verification Checklist*" is intended to be used by the Level 1-2 data reviewers when verifying the ambient air monitoring data. Level 1 reviewers will first document their verification of the monitoring data. After completing the appropriate information in the checklist's header, the data reviewer will read the instructions and proceed with completing the checklist. All questions assigned to the Level 1 reviewer should be answered. The column second from the right of the checklist lists recommended response actions to certain criteria that are determined to have not been met. The column furthest right lists hyperlinked references associated with the questions. Once the checklist is complete, the data reviewer's name should be inserted at the end, with the checklist signed/initialed and dated. The checklist – along with supporting documentation – will then advance to the Level 2 data reviewer, who will follow these same steps in completing the checklist. Once the checklist is complete, the Level 2 data reviewer's name should be entered, with the checklist signed/initialed and dated. The checklist – along with supporting documentation – will then advance to the Level 3 data reviewer. Both the Level 1 and 2 data reviewers are encouraged to retain a copy of the completed checklist that is "locked" to any future edits for their records.

The Level 3 data reviewer will complete the "*Data Validation Checklist*", which is intended to be used when validating the ambient air monitoring data. The checklist is similar in appearance and structure to the "Data Verification Checklist", and the Level 3 data reviewer should answer each question on the checklist. Once the checklist is complete, the Level 3 data reviewer's name should be entered, with the checklist signed/initialed and dated. A copy of the completed verification and validation checklists – along with supporting documentation – should then be retained as a record in a designated location, to ensure it is easily accessible at a later date and is protected from edits, damage, or loss. These records are intended to provide supporting documentation for the data quality review that was completed.

Note: Although the checklists are useful tools for documenting data quality review, each monitoring organization should weigh the potential benefit and burden associated with completing these checklists. For smaller organizations, it may be feasible to routinely complete these checklists for every monitor operating in the network. For organizations with larger monitoring networks, completing the checklists for every monitor may be too burdensome; such organizations may alternatively consider completing the checklists for their network monitors on a rotational basis, or may complete the checklists for a select number of high priority monitors (e.g., monitors with design values near or exceeding the NAAQS). As noted earlier, the data verification and validation checklists, at a minimum, may be useful as tools for developing a tiered data review process, training staff, and/or for completing periodic ADQs.

Click the "PaperClip" to access the workbook.

Appendix B: Data Coding Examples

The purpose of this appendix is to help data reviewers select the most applicable AQS null or QA qualifier codes, using real-world monitoring scenarios as examples. EPA recommends data be coded in a manner that best reflects the actual event, or series of events, that occurred at a monitoring site (analyzer/sampler) and impacted the resulting data. For each example, multiple AQS coding options are presented as a means to illustrate the different types of coding choices AQS offers. Each scenario will discuss the coding choice EPA recommends as the "best answer", given the amount of information available. The example scenarios range from easy to more complex.

The scenarios presented in this appendix are the "as found" information provided to the data reviewer. In some of the examples, corrective actions are warranted in the field or laboratory in order to prevent recurrence of the issue(s). It is important to emphasize that when an issue that warrants corrective action is left unaddressed, data qualification or invalidation must continue until such time as the situation is successfully remedied.

It is also important to note that the AQS QA qualifier code of "1" (i.e., Deviation from a CFR / Critical Criteria Requirement) should be applied sparingly and only when compelling evidence is available. The "1" flag is not intended to "save" data that should be otherwise invalidated. Frequent data reviews will identify critical problems quickly, which should prevent larger data sets from developing problems that would require "1" flags or invalidation. Monitoring organizations are encouraged to contact their EPA Regional Office to discuss data scenarios that may result in the application of the "1" QA code. One example of such data coding is offered in this appendix to explain a situation in which the use of this specific code would be deemed appropriate.

Click the "PaperClip" to access the presentation.

Appendix C: Weight of Evidence Examples

The purpose of this appendix is to help data reviewers better understand the "weight of evidence" concept, using real-world monitoring scenarios as illustrations. A variety of scenarios are included here, ranging from straightforward to complex. Although this appendix does provide AQS data coding "solutions" for each scenario, the examples are more geared towards illustrating the "thought process" data reviewers should walk through to arrive at the final validity determination. Data reviewers are encouraged to consider all possible outcomes, eliminating choices based on implications or other reasons. The determination of data usability (i.e., which way the scale "tips" when weighing evidence, per Figure 8 in Section 2.2.1.3 of this document) includes an evaluation of whether data is technically sound and legally defensible, in addition to its adherence to CFR.

Monitoring organizations are encouraged to contact their EPA Regional Offices when weight of evidence decisions are not straightforward and/or could impact data completeness requirements. Likewise, monitoring organizations are also strongly encouraged to contact their EPA Regional Offices when impacted data includes potential exceedance(s) of any NAAQS standard.

Example 1:

The particulate laboratory received a **PM_{10} high volume** filter from a 1-in-3 day sampling site with a recommendation from the site operator that the filter be invalidated due to a fingerprint on the exposed filter. The filter was heavily loaded and the laboratory determined the PM_{10} concentration to be 165 $\mu g/m^3$ (the 24-hour standard for PM_{10} is 150 $\mu g/m^3$). The validator was provided a picture of the filter by the laboratory, which showed a small imperfection in one of the corners. The validator rationalized that, although a human fingerprint has a mass, its impact on the resulting concentration of an 8x10 inch high-volume PM_{10} filter would be significantly less than it would be on a 47mm $PM_{2.5}$ filter. The validator also noted that the site had passed all QC checks; a continuous PM_{10} monitor ~10 miles away recorded a concentration of 187 $\mu g/m^3$ on the same day; and the samples 3 days before and after the sampling event in question were <50$\mu g/m^3$. Weight of evidence options include:

1. Accept the filter as valid;
2. Accept the filter as valid, but apply the QA qualifier 'FX' (filter integrity issue) to be transparent about the fingerprint; or
3. Invalidate the filter.

Based on weight of evidence, the validator reversed the recommendation of the site operator and reported the PM_{10} data as valid, but with the qualifier 'FX', denoting a filter integrity issue.

Example 2:

A rural ozone monitor reported an unexpected exceedance of the 8-hour ozone standard. The closest urban site (~20 miles south) recorded similar ozone trends, but slightly lower concentrations. An internal TSA of the monitoring site the week prior to the exceedance noted that the instrument was being operated

in an office space where temperature was being controlled with the office thermostat, but was not being monitored using a NIST-traceable, certified thermometer. The instrument was also being operated with its cover removed. Because instrument issues were suspected, a performance audit of the equipment was also conducted. The audit passed at all concentrations with an average 3% difference from the audit system. Weight of evidence options include:

1. Accept data as valid;
2. Accept data as valid, but apply the QA qualifier '2' (Operational Deviation) to the data to denote the concerns with the traceability of the shelter temperature-monitoring device;
3. Accept data as valid, but apply two QA qualifiers to the data – the "2" flag and a "3" flag (Field Issue), to denote the observation of the analyzer operating without its lid, which is poor form;
4. Accept data as valid, but apply three QA qualifiers to the data – the "2" flag, the "3" flag, and a "1" (Deviation from CFR/Critical Criteria Deviation): because without its lid, the internal temperature regulation of the ozone analyzer may have been compromised, which could impact its FEM status; or
5. Invalidate the data.

As these options were considered, the validator considers that the office space temperature was controlled at ~72 degrees Fahrenheit and the ozone monitor's FEM allowed for the instrument to be operated in a shelter where the temperature range is between 5 – 40 degrees Celsius (i.e., 41 – 104 degrees Fahrenheit). The validator rationalizes it is unlikely that the office space exceeded this temperature range. Moreover, the passing performance audit results and the similar concentrations trends observed at the closest ozone site are compelling evidence that demonstrate the analyzer to be properly functioning. The audit results also demonstrate that the instrument is able to analyze ozone concentrations within acceptable limits, despite its lid being temporarily removed. The validator further rationalizes that, had the internal temperature of the ozone analyzer been out of range, the instrument would have thrown a diagnostic warning flag, which the operator and/or the auditor should have noted. None were identified. Therefore, the validator decides to retain this data (based on the weight of evidence) but, in order to be conservative and transparent, qualifies it in AQS with two QA qualifier flags (Option 3).

Example 3:

On January 1, two additional continuous $PM_{2.5}$ BAM FEM monitors were officially added to a monitoring network with 8 BAM FEM instruments, making 10 FEMs total, without the addition of an FEM/FEM collocated monitoring site to supplement the existing FEM/FRM collocation. This was noticed during the quarterly Level 3 data review prior to AQS upload in late April. The PQAO CV for the FEM/FRM collocation for the designated method was 13.6% for the previous year (i.e., aggregate, annual statistic) and the trend appeared to continue into the first quarter of the current year. In March of the current quarter, one month of data from the FEM at the existing collocated site was suspect due to holes punched in the filter tape. This was discovered during the monthly flow check on March 29th (the previous flow check was on March 1) and, despite holes being punched in the tape, the BAM in question passed an as-found leak and flow check. Weight of evidence options include:

1) Accept all data as valid;
2) Invalidate one month of data (March 1 to March 29) from the collocated FEM due to holes in the filter tape;

3) Invalidate data for the two new FEMs that were added without an additional collocated site established; or,
4) Invalidate all $PM_{2.5}$ FEM data due to insufficient collocation and poor CV.

Although this situation was complicated, the data validator weighed each option closely against regulatory and scientific/technical requirements. When *weighing* the evidence, professional judgment combined with technical understanding of the instrument led the data reviewer to determine Option 2 (i.e., invalidate one month of data from the collocated FEM) was the best data validation approach. The BAM user manual states that pinholes punched through the filter tape can cause erroneous beta ray measurements. This implies that pinholes can impact the detection system, which would be considered a major issue with the instrument. With that in mind, although the monthly flow check passed, the presence of the holes clearly indicated the presence of a technical issue / instrument malfunction. Moreover, although collocation and the precision (CV) are operational and systematic criteria in the data validation templates, respectively, the QA Handbook clearly states that not meeting the DQOs does not necessarily invalidate data. The validator rationalizes that the collocated FEM/FRMs are different methodologies, which in itself can result in an elevated CV. Upon making this decision, the validator recognizes that invalidation of the collocated monitor's data leaves the primary sampler without collocated data for a month. To be transparent, the data validator also applies an AQS QA qualifier flag of "2" (i.e., Operational Criteria Not Met) to the FRM data during this time period and documents the rationale for these FEM/FRM validation decisions in the associated data package.

Example 4:

The data validator confirms that critical, operational, and systematic criteria were met for the organization's $PM_{2.5}$ samples for all field parameters. However, a TSA identifies multiple non-conformances in the monitoring organization's recently relocated in-house $PM_{2.5}$ gravimetric laboratory. The audit occurred within 2 months of start-up. The non-conformances identified are all considered "operational criteria". The TSA findings include:

- The laboratory's aged microbalance has no known calibration or certification (traceability) documentation;
- The microbalance is not properly grounded;
- Laboratory blanks (QC samples) are out of specification (acceptance criterion is $\leq \pm 15\mu g$; blank results range from 98 µg to -477 µg);
- Field blanks (QC samples) are also significantly out of specification; and,
- The newly purchased relative humidity (RH)/temperature datalogger doesn't meet accuracy specifications.

Weight of evidence options include:

1. Accept all samples weighed in this laboratory as valid;
2. Accept all samples as valid, but apply the QA qualifier '2' (Operational Deviation) to be transparent about the multiple deviations observed in the laboratory that are considered "Operational Criteria" in the data validation templates;

3. Accept all samples as valid, but apply three qualifiers to the data to be even more transparent about the operational deviations: the "2", "LB" (Lab Blank Value Above Acceptable Limit), and "FB" (Field Blank Value Above Acceptable Limit) qualifiers; or
4. Invalidate all the samples.

The data validator weighs each option closely against regulatory and scientific/technical requirements. Understanding the gravimetric method for $PM_{2.5}$, the validator recognizes that the severe swings in the laboratory blank data – which is also apparent in the field blank data – indicates that static electricity is significantly impacting the ungrounded microbalance. Static electricity will cause the microbalance to incorrectly weigh filters. Therefore, given that the microbalance lacks NIST-traceability documentation and the QC data for the laboratory supports that the microbalance's readings are unreliable, the $PM_{2.5}$ samples weighed in the laboratory since start-up cannot be trusted. Moreover, there is doubt in the accuracy of the RH and temperature readings in the laboratory, which means the laboratory climate control (i.e., filter conditioning) data is also suspect. Therefore, based on the weight of evidence, the data validator invalidates the data for the laboratory back to the date of start-up two months prior. This would include samples analyzed during pre-sampling weigh sessions as well as exposed samples returned for final weigh.

Example 5:

A BAM instrument began operating in June. In November, a back-up site operator performed the required monthly flow rate verification and observed that it did not pass the acceptance criterion; the flow rate measured 6.8% difference. The back-up operator recorded this information in the logbook and sent an email to the primary operator about the issue. The primary operator, upon returning to the site, began an extensive investigation into the cause of the flow exceedance. Upon review of the data captured by the sampler on the day of the unsatisfactory flow rate verification, a filter temperature exceedance was also observed. It was determined that the filter temperature sensor malfunction had caused the instrument's mass flow controller to produce the incorrect flow. Upon further review of the sampler meta data, looking specifically at the sampler temperature, the primary operator observed that there had been an intermittent issue with temperature throughout the time period since instrument installation. At times, the temperature was greater than 10 degrees inaccurate (based on the ideal gas law and average ambient temperatures, a 10 to 12 degree inaccuracy would result in a flow error of greater than 4 percent). The meta data showed there were often multiple temperature malfunctions in most hours.

After this review, the operator recommended to the data validator that the data from this instrument be invalidated back to instrument start-up in June. The data validator, upon receipt of this documentation from the site operator, noted that a performance audit conducted on the sampler in August passed, and the monthly flow rate checks conducted in June through October passed. Weight of evidence options include:

1) Accept all data as valid;
2) Scrutinize each hour of temperature data for the entire time period in question (i.e., June – November) by comparing results to any available, certified temperature data, in order to validate individual hours of sampler operation;
3) Invalidate all data from the time of the failed flow rate verification in November back to the last passing flow rate verification the month prior;

4) Invalidate data from the time of the failed flow rate verification, back to the last passing flow rate verification, and forward until a successful repair of the instrument and recalibration is performed; or,
5) Invalidate all data back to the time of sampler installation, per the written recommendation of the site operator, and forward until instrument repair/recalibration or replacement is completed.

Based on the scientific and technical principles upon which this instrument operates, the data validator determines that the temperature fluctuations are so frequent and severe that the validity of the collected data set cannot be defended, despite the passing QC checks. The validator also determines that Option 2 would require extensive resources and, despite the outcome, would likely not change the defensibility of the data set; the instrument was, ultimately, in a state of malfunction. The validator also recognizes that for NAAQS-compliance, 40 CFR Part 50, Appendix N specifies that 18 or more hours of valid data are needed on a given day for the day to be considered valid; it may not benefit the program much to save a few hours if complete days of data cannot be saved. Therefore, the decision is made to invalidate the data back to the time of sampler installation in June and forward until appropriate corrective actions were completed (Option 5).

Example 6:

A large PQAO performs an internal systems audit of one of the local monitoring organizations within its jurisdiction. When reviewing the documentation and records for a specific sulfur dioxide (SO_2) analyzer, the following issues are identified, which prompts the PQAO to re-validate the specific data set.

- The analyzer's sample flow rate (730 – 750 ccm) does not meet instrument manual specifications (650 ccm ± 10%) over the course of a six-month time period;
- The analyzer's slope (1.46 – 1.68) does not meet instrument manual specifications (1.0 ± 0.3) during this same time period;
- One-point QC checks, although passing, show a negative drift in the monitor over the past six months;
- Slope values are observed to change in the instrument logbook, indicating span adjustments are being made to the monitor. During audit interviews, the operator verbally acknowledges span adjustments were made during site visits as a "quick fix" to resolve sample flow issues and avoid multi-point calibrations;
- Zero adjustments are also observed in the instrument logbook, to which the operator also verbally acknowledges;
- Results of a performance evaluation exceed acceptance criteria at 2 out of 4 test concentrations, with the greatest deviation (40% difference) at the Level 3 concentration;
- Pump replacement after the PE immediately brings the analyzer's sample flow rate and slope back into compliance.

When reviewing this information, the QA auditors note that the zero and span adjustments were not documented on the PQAO calibration forms and proper calibration procedures, per the QAPP and SOP, were not followed. In fact, the QAPP specifically prohibits span adjustments exclusive of a multi-point calibration. The high slope value observed in the documentation further indicates the operator adjusted the span numerous times without following protocol. The auditor also observes that other SOP requirements were not met; for example, the operator did not perform monthly flow checks (with a flow

meter) to verify the analyzer flow rate, and accurate flow rate is important to achieve accurate concentrations for an SO_2 analyzer. Moreover, the SOP states that the instrument user manual must be followed; the manual, in turn, states that a PMT hardware calibration is required when the instrument slope is greater than 1.3. The instrument manual further states that the slope should be verified following calibration procedures in order to ensure linearity, which is an indicator of data quality. The replacement of the pump, followed by an immediate positive response of the analyzer, confirms the analyzer had an underlying equipment issue and was in a state of malfunction. However, it passed QC checks during the time period under evaluation. Calibration and audit criteria are listed as *operational criteria* in the data validation templates. The FEM designation of the SO_2 analyzer states that the instrument manual must be followed, but the slope and flow rate specifications are not directly included in the designations specifications listed in the *List of Designated Federal Reference and Equivalent Methods*. Operation of the instrument as an FRM/FEM, however, is considered a critical criterion.

Under this scenario, the PQAO's data reviewers must *weigh* the evidence to make a validity determination. Weight of evidence options include:

1) Accept all data as valid;
2) Accept all data as valid, but apply the QA qualifier '2' (Operational Deviation) to be transparent about the multiple deviations observed that are considered "Operational Criteria" in the data validation templates for SO_2 analyzers;
3) Accept data as valid, but apply two QA qualifiers to the data – the "2" flag and a "6" flag (QAPP Issue), to denote the deviations from the organization's QAPP/SOP requirements;
4) Accept data as valid, but apply three QA qualifiers to the data – the "2" flag, the "6" flag, and a "1" (Deviation from CFR/Critical Criteria Deviation): because the user manual (flow rate and slope) exceedances imply the SO_2 analyzer may have been operating outside of its FEM specifications;
5) Invalidate data for the entire 6-month time period where documentation demonstrates the instrument flow rate and slope are out of specification, and forward until the time of the pump replacement and subsequent recalibration of the SO_2 analyzer.

This complex situation is one in which it would be prudent for the monitoring organization to outreach to its EPA Regional Office for consultation. At a minimum, the data described in this situation would need to be qualified in AQS. However, understanding the scientific/technical requirements for the SO_2 analyzer, the more conservative approach would be to invalidate the data set; its quality would be difficult to defend, given the numerous deviations identified. As stated earlier in this document, understanding the intent of each criterion in the data validation templates (i.e., Information/Action column) is incredibly important, and the designation of items as operational or systematic does not negate their significance. Although QC checks passed during this time period, the out-of-specification slopes and flow rates – combined with the resolution brought on by the pump replacement – indicate the analyzer was in a significant state of decline throughout the 6-month time period. This, combined with the more egregious issues of the operator's lack of adherence to the program's QAPP/SOPs, specifically in regard to undocumented zero and span adjustments – adds a significant amount of uncertainty to the quality of the data.

United States Environmental Protection Agency	Office of Air Quality Planning and Standards Air Quality Assessment Division Research Triangle Park, NC	Publication No. EPA-454/B-21-007 August 2021