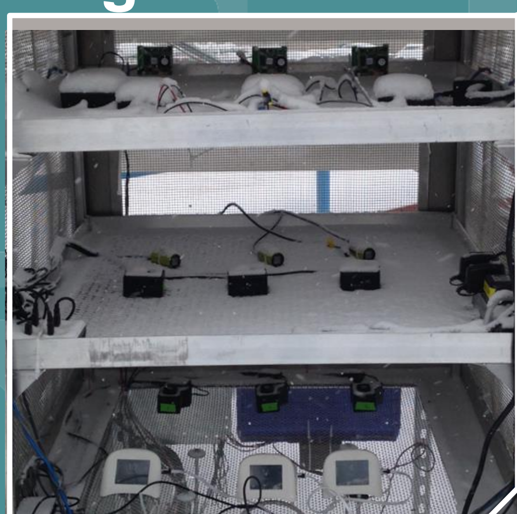
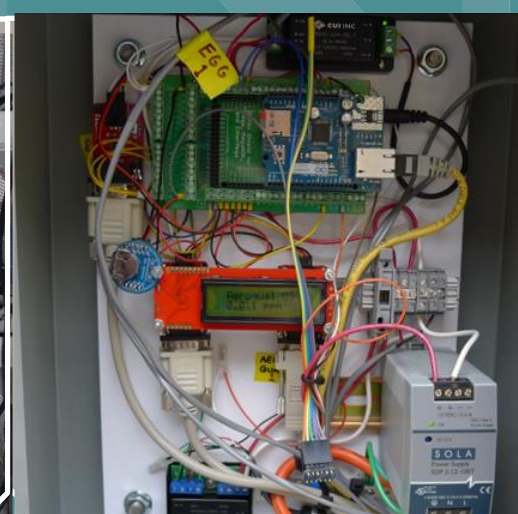
EPA
United States
Environmenta
Agency

EPA/600/R-18/324 | September 2018 | www.epa.gov/research

Peer Review and Supporting Literature Review of Air Sensor Technology Performance Targets





Office of Research and Development
National Exposure Research Laboratory

EPA 600/R-18/324
September 2018

Peer Review and Supporting Literature Review of Air Sensor Technology Performance Targets

by

Ron Williams, David Nash, Gayle Hagler, and Kristen Benedict
US Environmental Protection Agency
Research Triangle Park, NC 27711

Ian C. MacGregor, Brannon A. Seay, and Mitchell Lawrence
Battelle
Columbus, OH 43201-2696

and

Timothy Dye
TD Environmental Services
Petaluma, CA 94952

Disclaimer

This technical report presents work performed by the US Environmental Protection Agency's Office of Research and Development (ORD) with technical support provided through a Battelle (Columbus, OH) task order (Project # 100109904, EPA Contract # EP-C-16-014). The effort represents a collaboration between ORD and the US EPA's Office of Air Quality Planning and Standards with financial support from the Office of Enforcement and Compliance Assurance. Mention of trade names, commercial products, or various research institutions in the report does not constitute endorsement. The report has been internally and externally peer reviewed and approved by the US EPA for publication.

Acknowledgements

The authors thank the full technical staff associated with the Battelle task order for their original research efforts summarized here. The EPA 2018 Sensor Performance Deliberation Workshop team is acknowledged for their contributions in reviewing original data incorporated into this report and its summarization. Internal EPA peer reviewers and external peer reviewers who provided technical commentary are acknowledged. The staff of Jacobs Technology (EPA contract # EP-C-15-008) are thanked for their technical editing of the final document.

Executive Summary

Air quality monitoring is rapidly changing as miniaturized, lower-cost air sensors enable cities, community groups, businesses, and consumers to monitor local air quality conditions. Concurrently, air quality monitoring conducted by government agencies continues to use certified reference instruments that produce known, high-quality data necessary for regulatory applications. However, the lack of accepted performance specifications for air sensors limits understanding the quality of the data produced with this emerging technology.

Unlike more expensive instruments with comprehensive regulatory standards and processes for evaluation and certification, few standards and no certifications exist for low-cost air sensors. The lack of certification leads to confusion in the marketplace, as new buyers are uncertain of how well air sensors currently perform, how to operate (e.g., calibrate) them, and how well they need to perform to be fit for a given purpose.

To help improve data quality for sensors applied in a nonregulatory fashion, which is growing in prevalence, the United States Environmental Protection Agency (EPA) is considering development of a new voluntary sensor certification program for air sensors. The objective of this project is to evaluate peer-reviewed literature and other studies to identify performance attributes and metrics needed to obtain air monitoring data that are fit for a specific purpose or application. This work focused on ambient and near-source air monitoring for particulate matter ($PM_{2.5}$ and PM_{10}), carbon monoxide (CO), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), and ozone (O_3).

Substantial effort was invested to identify the information sources included in this literature review and synthesis. The process consisted of both automated and manual searches to identify relevant information sources from the peer-reviewed literature, technical reports, theses and dissertations, and regulatory air monitoring standards promulgated by government agencies, among others. Performance metrics for all potentially relevant air monitoring technologies were sought for inclusion in the literature review. However, given resource constraints, we included only those information sources published after 2007 and gave preference to the literature that provided quantitative performance characteristics of low-cost air sensors.

The quality of the information sources was assessed based on five different factors, with the primary focus on each source's applicability and utility to our study. Sources ranked highest for this factor when they contained quantitative, application-focused performance requirements for air monitoring instruments. A library of 257 potentially relevant information sources were identified, all of which were reviewed for applicability and utility. A total of 48 sources contained quantitative, and another 8 reported qualitative, performance requirements. Information about pollutants, applications, and performance results were extracted from these 56 different sources.

Quantitative performance requirements (data quality objectives [DQOs] and measurement quality objectives [MQOs]) were captured for ten different performance attributes/characteristics (also known as data quality indicators [DQIs]) that included accuracy/uncertainty, bias/trueness, completeness, detection limit, measurement duration, measurement frequency, measurement range, precision, response time, and selectivity. These ten data quality indicators are among the most common, were selected during the review of a variety of authoritative sources, and permit evaluation of performance requirements across a variety of applications and purposes but are not an exhaustive list of all possible performance

characteristics. The DQOs and MQOs were organized by 16 different air monitoring application types that were selected based on an initial literature review and in consultation with EPA, and were also binned into four broad categories (spatiotemporal variability, comparison, trend, and decision support), irrespective of the application. These categories describe the type of data analysis being performed with the measured pollutant concentrations and the decision sought, i.e., the purpose for the air monitoring. Stratification in this manner was performed to simplify the reported matrix of data (from 16 different applications to four broad data analysis types) and to facilitate the identification of potential qualitative trends in air sensor performance requirements. The performance requirements for regulatory air monitoring in the United States (US), European Union (EU), and China were captured under decision support. Also included in our review was information from the various extant and developing domestic and international air sensor performance evaluation and standards setting programs.

The results of the information review and synthesis are captured in the bullets below. In summary, more information, research, and resources are needed to determine fit-for-purpose air monitoring performance requirements.

- A total of 257 sources were located and assessed for applicability and utility; 48 (19%) contained quantitative performance information and 8 (3%) contained qualitative performance information. Thus, 56 (22%) of these information sources were included in the synthesis presented in this report.

- Performance requirements were found most frequently for spatiotemporal variation data analysis (40 to 72% of the time) and, more generally, quantitative air monitoring performance requirements detailing fitness for a given purpose were most abundant for O_3 (52%), followed by NO_2 (46%) and $PM_{2.5}$ (40%).

- Supplemental monitoring was most often cited as the purpose for collecting air pollution measurements, followed by community near-source monitoring, public education, and hot-spot detection.

- Across all data analysis types, high spatial density, cost, and accuracy/uncertainty are the main drivers for selection of air monitoring technologies. However, once the results are stratified, there is, in accord with expectations, the preference for regulatory monitoring is toward attainment of high accuracy, precision, and selectivity, whereas for non-regulatory monitoring purposes, accuracy remains important but high spatial density and low cost supersede precision and selectivity.

- Observations for pollutant concentration measurement range include that:
  - No information was found for the air quality forecasting and process study research applications
  - Supplemental monitoring typically requires measurements across the largest concentration range of all the different applications
  - The largest ranges were reported for CO and O_3.

- Among the 48 information sources containing quantitative performance requirements for air monitoring instruments, and excluding those sources pertaining to regulatory air monitoring (11), 70% (26 of 37) adjusted for measurement artifacts, 8% (3 of 37) intentionally chose not to perform adjustments for artifacts, and for 22% (8 of 37) of the studies such adjustments were not applicable. Some of the most frequent adjustments were made to account for cross-sensitivity and interference both from other airborne species and from changes in temperature and relative humidity.

- DQOs/MQOs for the various performance attributes (DQIs) were given most often for accuracy/uncertainty, followed by precision, measurement range, and detection limit.

- Among the 48 information sources containing quantitative performance requirements for air monitoring instruments, and excluding those sources pertaining to regulatory air monitoring (11), 68% (25 of 37) compared air pollution measurements to a reference instrument of some type, 11% (4 of 37) did not, and such information was not applicable for the remaining 21% (8 of 37). A wide range of performance against a reference method was reported in the literature for all the pollutants, indicating the presence of a range of data quality issues with the use of lower-cost sensors.

- Treatment of erroneous data was discussed in only 35% (13 of the 37) non-regulatory information sources with quantitative performance specifications. That most of the studies captured in this synthesis (which primarily included those using lower-cost sensors) did not explicitly discuss how to treat erroneous data suggests the need for more guidance on proper techniques and procedures on how such data should be managed. This is especially important given the many potential data quality issues encountered in lower-cost air sensor measurements.

- Two different international air sensor performance standards were identified. One set has already been developed and one is presently under development:

    - China's Ministry of Environmental Protection (MEP) and Environmental Protection Department of Hebei Province (China) have developed performance standards for sensors.

    - The European Committee for Standardization (CEN), Technical Committee 264-Air Quality, Working Group 42 is developing technical specifications for gas sensors. They have proposed three different classes (i.e., tiers) for sensor performance. Two of the three classes relate directly to the indicative and objective estimation targets in the CEN Air Quality Directive. The third class is for sensors that do not formally meet DQOs and can be used for research, educational purposes, and citizen information. Of particular importance is that Working Group 42 does not expect air sensors to be suitable for the purpose of fixed monitoring for regulatory compliance/decision support.

- In general, the *a priori* expectation was that the air monitoring performance requirements would increase in stringency (spatiotemporal < comparison < trend < decision support), where measurements performed for spatial or temporal analysis may in general be of lesser quality (e.g., they may have greater imprecision) than those measurements used for comparison to a threshold value, analysis of trends over longer periods, and for decision support. Major and cross-cutting findings include:

    - Decision support has the strictest performance requirements for precision, accuracy, completeness, and detection limit.

    - Required measurement durations are shorter for spatiotemporal, comparison, and trend data analyses, which is consistent with the conclusion that higher time resolution data are required for these applications.
    - In many instances, inconsistencies in the types of descriptors precluded the evaluation and detection of patterns in performance requirements. Due to resource limitations, no effort was undertaken to normalize the DQOs/MQOs for different descriptors for a given DQI. It is likely that such harmonization would have improved the ability to draw conclusions about trends in accuracy/uncertainty.
    - Due to a combination of inadequate, inconsistent, and limited information, non-regulatory air monitoring performance requirements cannot be stratified into tiers or categories of performance.
- Table A1 lists the DQOs/MQOs for four of the 10 performance attributes (DQIs) for regulatory monitoring in the US, the EU, and China. Numbers given in brackets, [ ], denote the citation number. Patterns in DQIs could not be ascertained for bias, completeness, measurement duration, measurement frequency, precision, or selectivity due to the lack of consistent and sufficient information on performance requirements.

Table A1. US, European Union, and Chinese Regulatory Monitoring Performance Requirements for Accuracy/Uncertainty, Detection Limit, Measurement Range, and Response Time for Measurements of $PM_{2.5}$, PM_{10}, CO, NO_2, SO_2, and O_3

Pollutant	Performance Attribute	US	EU	China
$PM_{2.5}$	Accuracy/uncertainty	R^2: 0.7225-0.9025 [1]		$R^2 \geq 0.8649$ [2]
	Measurement range	Measurement range: 3-200 µg/m^3 [1]	Measurement range: (0-1000$_{24h\text{-}avg}$, 0-10000$_{1h\text{-}avg}$ µg/m^3) [3]	Measurement range: 0-1000 µg/m^3 [2]
PM_{10}	Accuracy/uncertainty	$R^2 \geq 0.9409$ [1]		$R^2 \geq 0.9025$ [2]
	Measurement range	0-300 µg/m^3 [1]	(0-1000$_{24h_avg}$, 0-10,000$_{1hr_avg}$) µg/m^3 [3]	0 – 1000 µg/m^3 [2]
CO	Response time	Rise & Fall time: 120 sec [1],	Rise & Fall time: ≤180 sec [4]	Response time: ≤240 sec [5]
NO_2	Accuracy/uncertainty	12-hr zero drift: ±20 ppb [1] 24-hr zero drift: ±20 ppb [1] 24-hr 80% span drift: ±5.0 % [1] 24-hr 20% span drift: ±20.0% [1]	12-hr zero drift: ≤2.0 ppb [6] 12-hr span drift (ppb): ≤6.0 [6] Long-term zero drift: ≤5.0 ppb [6]	24-hr zero drift: ±5 ppb [5] 24-hr 80% span drift: ±10 ppb [5] 24-hr 20% span drift: ±5 ppb [5] Long-term zero drift: ±10 ppb [5]
	Detection limit	Detection limit: 10 ppb [1]		Detection limit: ≤2 ppb [5]
	Response time	Rise & Fall time: 15 min [1] Residence time: <2 min [7]	Rise & Fall time: ≤180 s [6] Residence time: ≤ 3.0 sec [6]	Response time: ≤5 min [5]
	Measurement range	Measurement range: 0-500 ppb [1]	Measurement range: ≤ 261 ppb [6]	Measurement range: 0-500 ppb [5]
SO_2	Accuracy/uncertainty	12-hr zero drift: ±4 ppb [1] 24-hr zero drift: ±4 ppb [1]	12-hr zero drift: ≤2.0 ppb [8]	24-hr zero drift: ±5 ppb [5]
	Response time	Rise & Fall time: 120 sec [1]	Rise & Fall time: ≤180 sec [8]	Response time: ≤5 min [5]
O_3	Accuracy/uncertainty	24-hr zero drift: ±4 ppb [1]		24-hr zero drift: ±5 ppb [5]
	Measurement range	Measurement range: 0-500 ppb [1]	Measurement range: ≤250 ppb [9]	Measurement range: 0-500 ppb [5]
	Detection limit	Detection limit: 5 ppb [1]		Detection limit: ≤2 ppb [5]
	Response time	Lag & Rise time: 120 sec [1]	Lag & Rise time: ≤180 sec [9]	Response time: ≤5 min [5]

Table of Contents

List of Figures

List of Tables

List of Appendices

Abbreviations and Acronyms

ADQ	audit of data quality
ANSI	American National Standard Institute
AQD	Air Quality Directive
AQI	Air Quality Index
AQ-SPEC	Air Quality Sensor Performance Evaluation Center
ASTM	American Society of Testing Materials
CEN	European Committee for Standardization
CFR	Code of Federal Regulations
CH_4	methane
CO	carbon monoxide
CO_2	carbon dioxide
CV	coefficient of variation
DQI	data quality indicator
DQO	data quality objective
EPA	US Environmental Protection Agency
ETV	Environmental Technology Verification
EU	European Union
FEM	Federal Equivalent Method
FRM	Federal Reference Method
H_2S	hydrogen sulfide
HCl	hydrogen chloride
HDMR	high-dimensional model representation
IUPAC	International Union of Pure and Applied Chemistry
ISO	International Organization for Standardization
LOD	Limit of Detection
MCERTS	Monitoring Certification Scheme (United Kingdom)
MDL	Method Detection Limit
MEP	Ministry of Environmental Protection (China)
MQO	measurement quality objective
NAAQS	National Ambient Air Quality Standards
NO	nitric oxide
NO_2	nitrogen dioxide
O_3	ozone
ORD	Office of Research and Development
Pb	lead
$PM_{2.5}$	particulate matter with aerodynamic diameter < 2.5 μm
PM_{10}	particulate matter with aerodynamic diameter < 10 μm
ppb	part per billion
ppbv	parts per billion by volume
ppm	parts per million
QA	quality assurance
QA/QC	quality assurance/quality control

QAPP	Quality Assurance Project Plan
r^2	coefficient of determination
RH	relative humidity
RIF	Reference Information File
RMSE	root mean squared error
RPD	relative percent difference
RSD	relative standard deviation
SCAQMD	South Coast Air Quality Management District
SO_2	sulfur dioxide
tVOC	Total volatile organic compounds
US	United States
VOC	volatile organic compound

1 Introduction

1.1 Background and Objectives

For decades, government agencies have deployed and operated expensive, complex reference instruments to measure air pollution for regulatory and research applications. These reference monitors are approved by organizations such as the US Environmental Protection Agency (EPA), the European Committee for Standardization (CEN), and China's Ministry of Environmental Protection (MEP). These organizations have established performance standards, which are documents with specific requirements that an instrument must meet to be acceptable for a given application or use. Performance standards typically include data quality objectives (DQOs) or measurement quality objectives (MQOs), data quality indicators (DQIs), testing methods, technical specifications, and operational criteria and may be based on the need to demonstrate attainment of air quality standards, adherence to laws, or achievement of specific requirements for a given application. A certification program is a process, typically with the force of law, to ensure that an instrument or measurement method meets the requirements of a given standard. These organizations and others develop, implement, and enforce programs that ensure the initial and ongoing quality of the data produced by manufacturers' instruments to ensure the measurements are fit for the required purpose, which is usually driven by compliance with a statute or regulation.

Recently, the rapid growth of miniaturized, lower-cost air sensors is changing the landscape of air pollution monitoring to enable cities, civil society, businesses, and consumers to monitor local air quality conditions, though with accuracy not on par with reference techniques. Unlike more expensive instruments with comprehensive, codified regulatory standards and performance certification processes, few standards or certification procedures exist for these new lower-cost air sensors. The lack of accepted performance specifications for air sensors is limiting the understanding of the quality of the data produced with this emerging technology and is leading to confusion in the marketplace, as new buyers are uncertain of how well air sensors currently perform, how to operate (e.g., calibrate) them, and how well sensors need to perform to be suitable for a given purpose. Yet interest in lower-cost sensors is proliferating because their price and size allow anyone to purchase and begin monitoring air pollution anywhere at any time. Furthermore, businesses, from small start-ups to large international companies, are promoting these devices for applications ranging from monitoring inside/outside homes, assessing urban air quality, measuring pollution near industry, and conducting school science programs.

Performance standard and certification programs for air quality sensor systems are therefore needed and would produce many benefits, including:

- Assuring that sensor systems produce data of sufficient quality and quantity to be fit for their intended purpose.
- Creating clear market incentives for manufacturers to improve the performance of their devices.
- Reducing confusion in the marketplace, as new buyers are more certain of the current performance of commercial air sensor systems compared to regulatory performance requirements.

EPA is interested in understanding whether there is evidence to support the definition of performance requirements for air sensor technology used in non-regulatory applications. If established and supported by sound, defensible science, these performance targets may facilitate the deployment of affordable air sensors that produce data of sufficient quality and quantity for non-regulatory air monitoring applications such as source identification, identification of spatiotemporal pollution gradients, and public awareness.

To inform the development of potential air sensor performance targets, this project aimed to:

- Review the recent peer-reviewed literature and other studies/programs to identify the most important performance attributes, or DQIs, that characterize the performance required for instruments suitable for monitoring pollutants in ambient and near-source air.
- Identify quantitative performance metrics (DQOs/MQOs) needed for each performance attribute (DQI) such that the results obtained are fit for the given purpose or desired use of the pollutant measurements.

This work focused on informing the development of performance requirements for air monitoring instruments that measure particulate matter ($PM_{2.5}$ and PM_{10}), carbon monoxide (CO), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), and ozone (O_3). To place the outcomes of this literature review and synthesis into context, the next section describes existing air monitoring performance certification programs.

1.2 Overview of Existing Performance Standards and Certification Programs

This section provides background information on established performance programs focused on reference instruments and covers emerging programs for air quality sensor systems. Various government agencies develop and implement these performance standard/certification programs. As shown in Figure 1, several stages of development are involved in setting standards, evaluating equipment, and certifying equipment:

- **Need**. Definition of the purpose for the air monitoring with a desire for measurements of known and sufficient quality and quantity.
- **Performance setting**. A consensus-building process to establish the technical and credible standard.
- **Publishing a standard**. There are several different types of standards: 1) performance-based standards that specify acceptance criteria that must be achieved for fitness for purpose but do not stipulate the instruments that must be employed, and 2) method-based standards that designate the instruments, operating conditions, and performance requirements to be fit for a given purpose.
- **Evaluation**. Evaluating instruments/sensor systems against the standard may be performed by organizations (public and private) and results are published as evidence of meeting the standard.
- **Certification**. When an accredited organization validates and certifies the results of the evaluation against the standard.

As shown in Table 1, the US, the European CEN, and the People's Republic of China have established standards and certification programs for reference instruments as part of their regulatory and compliance monitoring programs. Numbers given in brackets, [ ], denote the citation number. These programs are described in more detail below, along with voluntary performance evaluation programs.

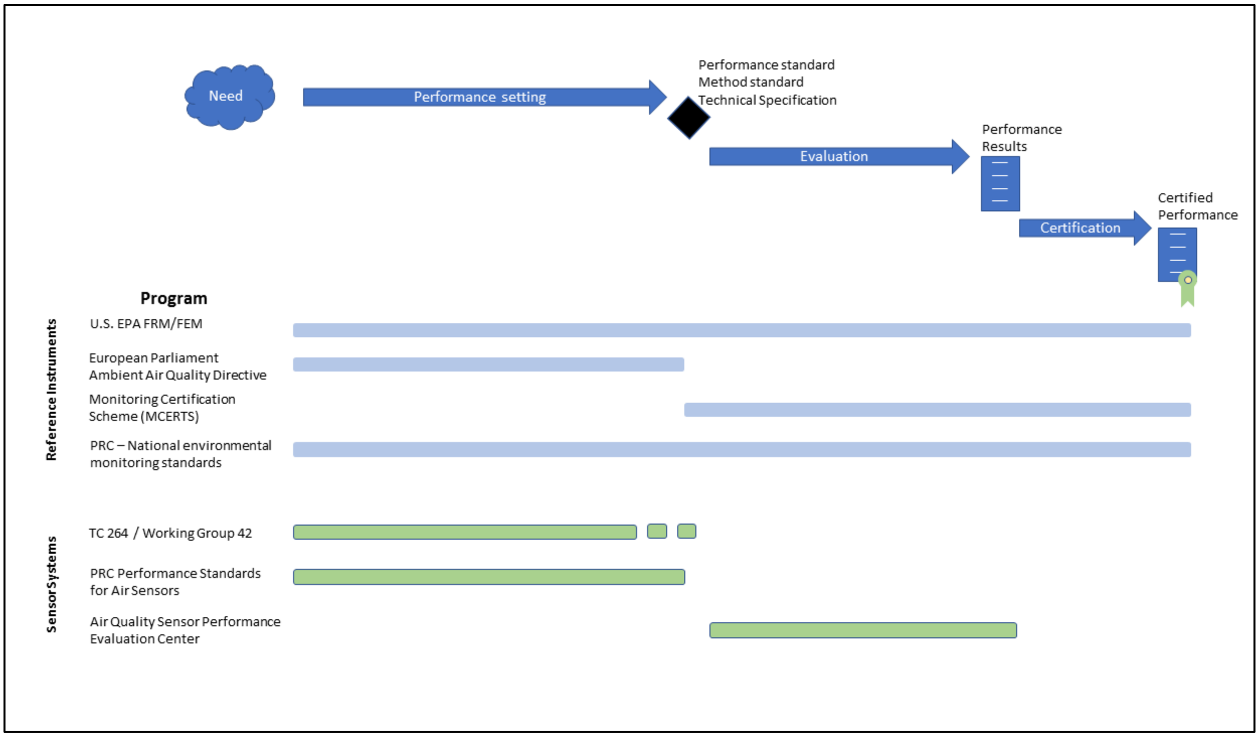


Figure 1. Stages of Development Involved in Setting Standards, Evaluating Instruments, and Certifying Instruments for Those Programs Identified in Table 1

Table 1. Characteristics of Different Evaluations and Certification Programs for Reference Instruments and Air Sensor Measurement Instruments and Systems

Program	US EPA FRM/FEM Program	European Parliament and of the Council of Ambient Air Quality Directive (2008/50/EC)[1]	Monitoring Certification Scheme (MCERTS)	People's Republic of China National environmental monitoring standards	US EPA Performance Standard 18	European Committee for Standardization (CEN) Technical Committee 264 (Air Quality) Working Group 42 (Gas sensors)	People's Republic of China Performance Standards for Air Sensors	Air Quality Sensor Performance Evaluation Center
Organization	US EPA	European Committee for Standardization	Environment Agency (UK)	Chinese Ministry of Environmental Protection (MEP)	US EPA	European Committee for Standardization	Chinese Ministry of Environmental Protection (MEP)	South Coast Air Quality Management District (SCAQMD)
Type	Performance Standards Certification (instruments)	Performance Standards (instruments)	Certification (instruments)	Performance Standards Certification (instruments)	Performance Standards (instruments)	Technical Specifications (air sensors)	Performance Standards (air sensors)	Performance Evaluation (air sensors)
Pollutants	Ambient O_3, NO_2, CO, SO_2, $PM_{2.5}$, PM_{10}, and Pb	Ambient $PM_{2.5}$, PM_{10}, CO, NO_2, SO_2, and O_3, NO_3, $PM_{2.5}$, PM_{10}	Ambient $PM_{2.5}$, PM_{10}, CO, nitric oxide (NO), NO_2, SO_2, O_3, benzene, and benzene-like VOCs	Ambient $PM_{2.5}$, PM_{10}, CO, NO_2, SO_2, and O_3	Source Hydrogen Chloride (HCl)	Ambient O_3, NO, NO_2, CO, SO_2, and carbon dioxide (CO_2)	Ambient $PM_{2.5}$, PM_{10}, CO, NO_2, SO_2, O_3, and total VOCs (tVOC)	Ambient $PM_{2.5}$, PM_{10}, CO, NO_2, NO_x, SO_2, O_3, VOCs, hydrogen sulfide (H_2S), and methane (CH_4)
Application Tiers	Single Tier Designated reference or equivalent method for use in regulatory monitoring for the NAAQS	Three Tiers 1. Fixed measurements (highest quality) 2. Indicative measurements 3. Objective estimation	Two tiers 1. Fixed measurements (highest quality) 2. Indicative measurements	Single Tier	Single Tier Any instrumental technology that can meet performance criteria may be used	Three tiers Class 1 - meets the DQOs of indicative measurements set in the Air Quality Directive (2008/50/EC) Class 2: meets the DQOs of objective estimation Class 3: measuring device delivering measurements that are not formally associated with any mandatory target measurement uncertainty	Single Tier	Single Tier

Table 1. Characteristics of Different Evaluation and Certification Programs for Reference Instruments and Air Sensor Measurement Instruments and Systems (continued)

Program	US EPA FRM/FEM Program	European Parliament and of the Council of Ambient Air Quality Directive (2008/50/EC)[1]	Monitoring Certification Scheme (MCERTS)	People's Republic of China National environmental monitoring standards	US EPA Performance Standard 18	European Committee for Standardization (CEN) Technical Committee 264 (Air Quality) Working Group 42 (Gas sensors)	People's Republic of China Performance Standards for Air Sensors	Air Quality Sensor Performance Evaluation Center
Test Locations	Laboratory and Field	Laboratory and Field	Laboratory and Field	Field	Field	Laboratory and Field	Field	Laboratory and Field
Outcomes	Designated reference or equivalent method by US EPA	Stamp of approval for the use of specific analyzers (in their tested configuration) in national monitoring networks	Product Conformity Certificate issued for an instrument and concentration range	Unknown	Any instrumental technology that can meet performance criteria may be used	Unknown	Unknown	Evaluation report posted on AQ-SPEC website
References	Title 40, Parts 50 and 53 of the Code of Federal Regulations [10]	Ambient Air Quality Directive (2008/50/EC) and in the amending Directive (EU 2015/1480) [11]	Environment Agency [12-16]	National environmental standards People's Republic of China (HJ 653-2013 and HJ 654-2013) [2, 5]	US EPA PS18 [17]	TD Environmental personal communication with M. Gerbolis [18]	Compiled from MEP documents and Hebei documents. [2, 5, 19-21]	Papapostolou et al. [22]

[1]DQOs stated in Ambient Air Quality Directive (2008/50/EC)

Note that test duration varies for each evaluation and certification program and is not discussed here. The reader is referred to the referenced documents for more information.

1.2.1 US Regulatory Air Monitoring Requirements and Certification Programs

EPA has a program to evaluate instruments suitable for use in determining compliance with the National Ambient Air Quality Standards (NAAQS). As part of this program, candidate instruments measuring $PM_{2.5}$, PM_{10}, CO, NO_2, SO_2, O_3, and lead (Pb) are evaluated against requirements codified in Title 40, Parts 50 and 53 of the Code of Federal Regulations (CFR) [7,10]. Instruments that attain the applicable performance specifications are designated as either a Federal Reference Method (FRM) or a Federal Equivalent Method (FEM). This program currently categorizes instruments into one tier – either designated as FRM/FEM, which supports use in regulatory monitoring, or not FRM/FEM designated (i.e., non-regulatory). Attainment of both technical specifications in 40 CFR Parts 50 and 53 and the national air monitoring program DQOs given in 40 CFR Part 58 Appendix A ensure that errors in NAAQS attainment/nonattainment decision-making are controlled to acceptable levels.

1.2.2 European Regulatory Air Monitoring Requirements and Certification Programs

Similarly, the European Commission, acting through the CEN, has produced a series of standard methods [3, 4, 6, 8, 9, 23] for monitoring air pollutants applicable to air monitoring in the European Union (EU). These standards outline minimum performance requirements to ensure instruments meet the DQOs established in the Ambient Air Quality Directive (2008/50/EC) [11] and in the amending Directive (EU 2015/1480) [24]. These DQOs are divided into three performance tiers:

- Fixed measurements – highest quality, used for trends and compliance.
- Indicative measurements – DQOs that are less strict than those required for fixed measurements.
- Objective estimation – supplemental information with which pollution levels below the lower assessment threshold may be measured.

The United Kingdom’s Environment Agency operates the Monitoring Certification Scheme (MCERTS), which certifies that instruments, personnel, and organizations comply with European Directives. Certification is based on the CEN standard methods ($PM_{2.5}$, PM_{10}, CO, NO_2, NO, SO_2, and O_3, benzene, benzene-like volatile organic compounds [VOCs]) and indicative dust monitoring (PM_{10} only) [12-16]. Both laboratory and field evaluations are conducted, and certification is performed by accredited third-party organizations. Certified instruments are issued a Product Conformity Certificate for a specified concentration range. TUV Rheinland [25] provides certification of compliance with the EU standards as well.

1.2.3 Chinese Regulatory Air Monitoring Requirements and Certification Programs

The People’s Republic of China created instrument performance standards (HJ 654-2013 and HJ 653-2013) [2, 5] to support its Prevention and Control of Air Pollution and Ambient Air Quality Standards (GB 3095-2012) [26]. These standards cover specifications and test procedures for continuous automated monitoring of $PM_{2.5}$, PM_{10}, CO, NO_2, SO_2, and O_3. These standards include technical requirements, performance indexes, and test procedures. China also has a certification program; however, no program-specific details could be found during this review.

1.2.4 Performance Specification and Evaluation Programs for Non-Regulatory Monitoring

Performance evaluation programs for sensor systems were recently established in California, China, and the European Union. In 2014, the South Coast Air Quality Management District (SCAQMD) established the Air Quality Sensor Performance Evaluation Center (AQ-SPEC) [22] to inform the public about the initial performance of commercially available, low-cost air quality sensors in ambient air at fixed sites. Until the creation of the AQ-SPEC program, there had not been an objective way to systematically evaluate the performance of air sensors and sensor data. The center evaluates the performance of air sensors in both field and laboratory settings; provides guidance on sensor technology; and seeks to catalyze the successful evolution, development, and use of sensor technology.

In 2017, the People's Republic of China's MEP developed performance standards for particle and gas sensor systems [20, 21]. These performance standards include criteria for laboratory and field evaluations and describe the methods to compare data measured by sensors to data collected from reference instruments. Guidance also includes information on network design, technical requirements and testing methods, monitoring system quality assurance/quality control (QA/QC) and operation, and network installation and acceptance.

The CEN Technical Committee 264, Working Group 42 is currently developing technical specifications for performance requirements and test methods for low-cost sensors under prescribed laboratory and field conditions [18]. These technical specifications remain under development and will describe the general principles, including testing procedures and requirements, for the evaluation of the performance of low-cost air sensor systems for the monitoring of gaseous compounds in ambient air at fixed sites. It is likely that the evaluation of sensor systems will include tests performed under prescribed laboratory and/or field conditions that are collocated at reference stations.

The Working Group 42 protocols specify the methods to evaluate the sensitivity, selectivity, and stability of air sensor measurements. Working Group 42 is anticipating three classification regimes and test procedures:

- Class 1 represents the highest accuracy reachable with sensor systems; it meets the DQOs of indicative measurements established in the European Air Quality Directive (AQD).
- Class 2 sensor systems meet the DQOs of objective estimation techniques in the ADQ.
- Class 3 sensor systems are those that do not formally meet the AQD's DQOs but can be used for research, educational purposes, and citizen information.

The EPA, the CEN, and many other organizations (American Society for Testing and Materials [ASTM], American National Standards Institute [ANSI], etc.) have also developed other instrument performance evaluation programs. For example, EPA Performance Specification 18 [17] applies for measuring gaseous concentrations of hydrogen chloride (HCl). It allows the use of different sampling and analytical technologies as long as the required performance criteria are met. In 1995, EPA's Office of Research and Development (ORD) created and administered the Environmental Technology Verification (ETV) Program [27] to perform credible, third-party testing and evaluation of innovative environmental technologies. More than 400 technologies were verified under the program before it concluded in 2014. In November 2016, the International Organization for Standardization (ISO) promulgated standard 14034 [28], which provides an approach to technology evaluation based on a standardized procedure that encourages the sharing of verification results across multiple jurisdictions. The ISO 14034 process is a quality-assured approach for the identification of credible performance parameters and permits independent verification of the actual performance of technologies, enabling informed and effective decisions on technology selection and use. VerifiGlobal [29] is a member-based program that performs third-party testing according to the ISO standard.

The remainder of this report describes the approach to gathering information on performance requirements and present and discuss the applications, performance attributes, and performance specifications gleaned from the literature search. We summarize our results, the limitations of this work, and provide recommendations for future work.

2 Approach

2.1 Information Source Identification

The overall objective of this work was to inform the selection of required performance specifications that air monitoring instruments and low-cost air sensors must meet to measure the criteria pollutants $PM_{2.5}$, PM_{10}, CO, NO_2, SO_2, and O_3 in ambient and near-source air for a variety of applications, such as trends analysis, decision making, research, and citizen science. A comprehensive, but not exhaustive, literature review was performed to identify information sources that contained such performance requirements. The term “information source” refers to a specific document that was considered for incorporation into the information review. The search was limited to documents published between 2007 and 2017, and, given resource constraints, focused on and gave preference to the literature that provided quantitative performance characteristics of lower-cost air sensors. The following types of information sources were targeted:

1. Existing ambient, personal exposure, and near-source regulatory air quality measurement technology standards (e.g., CFR), including initial and ongoing quality assurance (QA) requirements and DQOs;
2. Existing or draft non-regulatory air quality measurement technology standards (e.g., European Union);
3. Peer-reviewed science journal articles and technical reports, as well as websites and other sources describing the use of measurement technology (1) to characterize air pollution trends in different environments and in different modes of use (e.g., stationary, portable), including near-roadway or other near-source air quality, general outdoor air quality trends, indoor air quality, personal exposure, health studies, and citizen science, and (2) to evaluate or describe air sensor technology;
4. QA documentation supporting non-regulatory measurements; and
5. Several EPA reports provided by the Task Order Contracting Officer’s Representative.

Two different types of searches were performed. The first consisted of an automated search of reference databases such as Compendex, Scopus, and Web of Science, which permitted identification of relevant peer-reviewed literature; the Networked Digital Library of Theses and Dissertations, OpenGrey, OpenAIRE, and WorldCat for identification of relevant information sources available in the “grey literature;” and the Catalog of US Government Publications, the Defense Technical Information Center, and the United Nations Digital Library for applicable US and international government documents. The automated searches identified information sources with relevant metadata in the source’s title, abstract, and/or keywords. For example, sources were selected if their database citations contained metadata that matched a relevant pollutant (particulate matter, CO, or NO_2, etc.), air type (ambient, near-source, or near-road, etc.), activity type (assess*, measure*, or monitor*, etc., where the asterisk indicates that any word containing the letters previous to the asterisk would be selected), and application (research, citizen science, or emergency response, etc.). Various similar search strategies were performed and approximately 20,000 potentially applicable information sources were identified. The Battelle Team, in consultation with EPA, determined that further down selection of relevant sources would require manual inspection and review of all candidate sources, and that insufficient resources were available to perform such a review. Thus, the automated search process was discontinued per EPA's technical direction.

The final list of potentially relevant information sources was determined instead by a hand-curated approach based on the Battelle Team's subject matter expertise. The literature was surveyed to select those sources expected to contain air measurement performance requirements, international subject matter experts were contacted to provide recommendations for literature to include, and reference sections of various information sources were inspected to identify additional relevant information sources. Regulatory air monitoring requirements for the US, EU, and China were intentionally sought out and captured. This approach enabled the Battelle Team to focus on finding highly relevant information sources. A master list of information sources was compiled as an Endnote library. Sources were selected such that the questions listed below could be answered.

1. In the review of existing performance standards:
    a. How do current regulatory technology performance standards for criteria pollutants in the US compare with those internationally (with a focus on ambient and near-source)?
    b. Are there any non-regulatory technology performance standards for criteria pollutants internationally? What is the justification for how these standards were set, and to what applications/pollutants do they apply?
2. For the review of research studies and information sources containing data called out below:
    a. What are the various purposes of applying the measurement technology (applications such as control strategy effectiveness, source identification, near- source monitoring, emergency response, public outreach, etc.)?
    b. What appear to be the drivers affecting the air measurement technology employed for specific monitoring purposes (such as cost, performance [accuracy, precision, bias], portability, reliability, etc.)?
    c. What were the expected concentrations and actual measured concentration ranges for specific measurement applications and environments?
    d. How are measurement artifacts addressed, such as impacts on measurement performance related to environmental conditions (adjustment, no adjustment; explanation)?[a]
    e. What, if any, in-use DQIs or other automated data quality checks were employed to flag and/or adjust data (precision, bias, accuracy, completeness, etc.)?
    f. If applicable, were the selected measurement techniques compared to FRM/FEM or other regulatory/reference instruments, and if so, what were the outcomes of these comparison(s) (compared to FRM/FEM or other reference standard, yes or no; if yes, indicate degree of agreement as bias range)?

---

[a] In the context of this project the term artifact captures the potential impact of co-collected pollutants and/or temperature/relative humidity (RH) changes on reported concentrations. An artifact may be manifested as imprecision, bias, change in sensitivity, etc.

g. How were erroneous data handled (not flagged and used; not flagged and not used [discarded/null coded]; flagged and used)?

h. What are the commonalities or differences among measurement DQOs within similar studies conducting non-regulatory air quality measurements (e.g., multiple near-road outdoor air quality studies) and between differing purposes of non-regulatory monitoring (e.g., indoor versus outdoor monitoring)?

Once potentially relevant information sources had been identified, they were further screened to ensure that they were in fact applicable and useful for the given purpose. The criteria for selecting the information sources were based on EPA's Assessment Factor process [30] as described in Section A7 of the Quality Assurance Project Plan (QAPP) [31]. The process permitted a qualitative and semi-quantitative evaluation of the fitness of the information source for inclusion in this literature review. Primary focus was on the applicability and utility, that is, on whether an information source contained quantitative performance requirements (i.e., DQOs) that describe how well air-monitoring instruments must perform to be fit for a given purpose. The information sources that contained such information were given a score of 2 (on a scale ranging from -1 to 2) and were down selected for inclusion in the present data synthesis. Sources that contained qualitative information on air monitoring instrument performance (but lacked quantitative, numerical DQOs/MQOs were scored a 1 on applicability/utility and were also down selected and included to permit a better understanding of, for example, the most important performance characteristics of air-monitoring instruments. Sources that scored -1 or 0 on applicability/utility were excluded. During the extraction and capture of the performance requirements from the down selected information sources, a score of -1 to 2 was assigned to each of the other four assessment factors (soundness, clarity and completeness, uncertainty and variability, and evaluation and review).

More details of the process by which the information sources were selected for inclusion in the literature review can be found in the QAPP in Sections A6 and B9 [31].

2.2 Air Monitoring Applications

The following applications were selected as the most relevant for inclusion in this work based on our discussions with EPA following an initial review of the literature and on our subject matter expertise in air pollution monitoring. The applications capture the typical anticipated purposes for which low-cost air sensors could be deployed to measure pollutants in ambient air, including near-source ambient air monitoring. These applications included (in alphabetical order):

- Air quality forecasting
- Air quality index (AQI) reporting
- Community near-source monitoring
- Control strategy effectiveness
- Data fusion
- Emergency response
- Epidemiological studies
- Exposure reduction (personal)
- Hot-spot detection

- Model input
- Model verification
- Process study research
- Public education
- Public outreach
- Source identification
- Supplemental monitoring

The selection of the various air monitoring applications, although important and relevant, was of secondary importance to the work presented herein because, as described in the section below, the performance requirements identified in the down selected information sources were categorized not by application, but by the type of data analysis that was to be performed with the air monitoring results.

2.3 Organization of Performance Requirements

Air monitoring instrument performance requirements found in the down selected information sources were binned into four broad categories, irrespective of the application. These categories describe the type of data analysis being performed with the measured pollutant concentrations and the decision sought and purpose of the air monitoring. The categorization scheme is based on the work of Lewis et al. [32], in which spatial and temporal variability are combined and the decision support category is an added feature to capture the regulatory monitoring applications.

The performance requirements were stratified in this manner to simplify the reported matrix of data (from 16 different applications to four broad categories of decision sought) to facilitate the identification of potential qualitative trends in air sensor performance requirements and in acknowledgement of the expectation that relatively few applicable sources (those containing relevant quantitative information describing the fitness of air monitors for a given purpose) would be located. The four categories are given below along with examples of decisions sought.

- Spatiotemporal variability – Characterizing a pollutant's concentration over geographic area and/or time
    - "Is pollution higher in the morning at location A or B?"
- Comparison – Analysis of differences and/or similarities in air pollution characteristics against a threshold value or between different networks, locations, regions, time periods, etc.
    - "Does a location show high pollution levels, but other locations do not?"
- Long-term trend – The change in a pollutant's concentration over a period of (typically) years
    - "How did $PM_{2.5}$ concentrations change at a location over a 5-year period?"
- Decision Support – Includes all regulatory monitoring applications monitoring
    - "What percentage of all ozone exceedances in a city are caused by motor vehicle emissions?"

2.4 Performance Characteristics, Descriptors, and DQOs/MQOs

Ten performance characteristics (also referred to as performance attributes and DQIs) were selected, taking into consideration (1) those likely to be of most importance to air quality measurements with air sensors; (2) the guidance promulgated by a variety of authoritative domestic and international government agencies and consensus standards organizations such as EPA [10, 33, 34], ASTM [35-37], the International Union of Pure and Applied Chemistry (IUPAC) [38, 39], and ISO [40, 41]; (3) those most likely to have quantitative DQOs/MQOs reported in the literature; and (4) those likely to permit evaluation of performance across a variety of air monitoring applications. The final 10 performance attributes selected were (in alphabetical order):

- Accuracy/uncertainty
- Bias/trueness
- Completeness
- Detection limit
- Measurement duration
- Measurement frequency
- Measurement range
- Precision
- Response time
- Selectivity

It is important to note that this list is not exhaustive or inclusive of the many performance parameters often used in FRM/FEM certifications or evaluations of research-grade instruments.

Each of the various attributes may be described in any multitude of ways, and each such "descriptor" may have DQOs/MQOs with different units. For example, the descriptor for precision may be a standard deviation with units of part per billion (ppb) or μg/m^3; a coefficient of variation (CV) or relative standard deviation (RSD) with units of percent (%); or a relative percent difference (RPD) with units of %. To the extent feasible, similar descriptors (e.g., standard deviation, CV, RPD) for each performance attribute (e.g., precision) were selected and captured to enable comparison of the various DQOs/MQOs for the different decisions sought and among the regulatory monitoring requirements in the US, EU, and China.

Definitions of each of the performance attributes and many of the descriptors are given in Appendix A.

2.5 Information Extraction

To capture a consistent set of information from each information source selected for inclusion in our synthesis, a template (a Reference Information File [RIF]) was developed as described in QAPP Section A6 [31]. The down-selected information sources were read and reviewed, and the relevant information was extracted into individual RIFs.

In many cases, the assignment of a measurement application to a data analysis/decision sought required a subjective judgment on the part of the reviewer of the information source. To the extent feasible, the three reviewers on the team sought to be consistent in their selection methodology.

2.6 Technical and Quality Reviews

Where applicable and as required in Sections C and D of the QAPP [31], peer, programmatic, QA, and management reviews were performed on work products, including this report. The outcomes of the audit of data quality (ADQ) for this report were provided separately to EPA.

In summary, Mr. Zachary Willenberg, Battelle's STREAMS III Contract Quality Assurance Manager, conducted the ADQ on April 30 and May 1, 2018. He assessed the accuracy of 100% of the information contained in six (6) RIFs (out of a total of 56 [11%]). Each RIF was compared to its source document to verify the accuracy of the extracted information, to determine if transcription or other data entry errors were made, and to review the completeness of the information captured. The intent of the ADQ was to verify that the information captured in each RIF was supported by the source documentation. The contents of each of the six (6) RIFs were also compared to the information presented in Tables B1 to B6 to verify the correct transcription of information. Overall, minor comments/observations were noted, with only six (6) minor findings called out (incorrect transcriptions or data in report tables not matching RIFs). Following receipt of the ADQ report, technical staff corrected all errors and inconsistencies found during the ADQ and subsequently performed a 100% review of all remaining RIFs to determine if any of the auditor's comments also applied to other RIFs. Technical staff identified and corrected all other similar transcription errors and inconsistencies in decisions regarding the inclusion/exclusion of reported air monitoring performance requirements. The results of the 100% data review are reflected in the data presented in the draft and in this final report.

3 Results and Discussion

3.1 Information Source Selection Summary

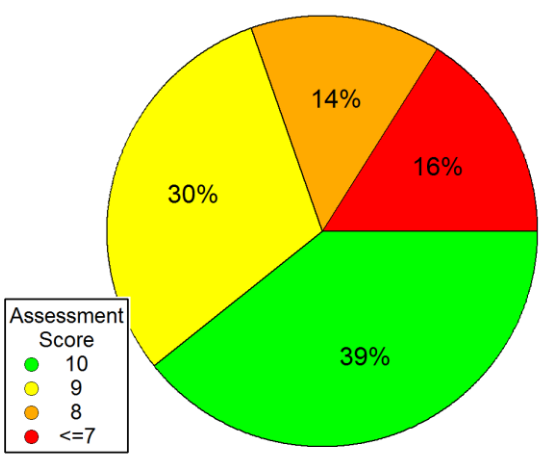
Shown in Table 2 is a summary of the outcomes of information source selection process.

Table 2. Breakdown on Information Sources Evaluated and Those Containing Quantitative and Qualitative Air Monitoring Performance Requirements

Number of Sources	Description
257	Information sources assessed for applicability and utility
201	Sources scored as a "-1" or "0", i.e., those containing neither qualitative nor quantitative performance requirements for air monitors
8	Sources scored as a "1", i.e., those containing qualitative performance requirements and ancillary, contextual information
48	Sources scored as a "2", i.e., those containing quantitative performance requirements for air monitors to be fit for a given purpose
56	Sources scored as "1" or "2" for which RIFs will be completed and with information that will be included in the information synthesis

A total of 257 sources were located and assessed for applicability and utility; 48 (19%) contained quantitative performance information (were scored a 2 on applicability/utility), and 8 (3%) contained qualitative performance information (were scored a 1 on applicability/utility). A total of 48 + 8 = 56 (22%) of all the hand-selected information sources were included in the information synthesis and form the basis of the results presented in this and the following sections.

A breakdown of the total assessment factor scores is shown in Figure 2. As can be seen, most of the information sources (39 of 56, or 69%) scored a 9 or a 10 out of a possible score of 10.



Assessment Score	Number of Info. Sources
10	22
9	17
8	8
7	4
6	3
5	2

Figure 2. Distribution of Total Assessment Factor Score for Information Sources in which Qualitative and Quantitative Performance Requirements Were Found

Table 3 shows the frequency and number of times that DQOs/MQOs were found, by pollutant and stratified by data analysis type, in the 48 information sources containing quantitative performance requirements. For a given combination of data analysis type and pollutant, the frequency is calculated as the number of times that, for example, $PM_{2.5}$ comparison studies were performed (6) out of all the times any information source presented performance requirements for any type of measurement for $PM_{2.5}$ (19); 6/19 = 32%. Shown in the right-most column is the frequency and number of times with which DQOs/MQOs were found for a pollutant regardless of the data analysis being performed. For $PM_{2.5}$, 19 of the 48 information sources contained DQOs/MQOs; 19/48 = 40%.

Table 3. Frequency and Number of Times Quantitative Performance Requirements (DQOs/MQOs) Were Found by Pollutant for the Various Data Analyses Performed

Pollutant[a]	Comparison	Spatio-temporal Variation	Trend	Decision Support	Other	% All Info. Sources
$PM_{2.5}$	32% (6)	63% (12)	5% (1)	26% (5)	5% (1)	40% (19)
PM_{10}	23% (3)	46% (6)	15% (2)	38% (5)	0% (0)	27% (13)
Carbon Monoxide (CO)	35% (6)	65% (11)	18% (3)	24% (4)	0% (0)	35% (17)
Nitrogen Dioxide (NO_2)	32% (7)	68% (15)	18% (4)	27% (6)	0% (0)	46% (22)
Sulfur Dioxide (SO_2)	20% (1)	40% (2)	20% (1)	60% (3)	0% (0)	10% (5)
Ozone (O_3)	20% (5)	72% (18)	20% (5)	20% (5)	0% (0)	52% (25)

[a] Totals across all of the data analyses are always greater than the figures in the right-most column because a single information source may contain performance requirements for more than one pollutant and/or data analysis type.

As can be seen in Table 3, performance requirements were found most frequently for the spatiotemporal variation data analysis type (40% to 72% of the time) and, more generally, quantitative air monitoring performance requirements detailing fitness for a given purpose were most abundant for O_3 (52%), followed by NO_2 (46%), then $PM_{2.5}$ (40%).

3.2 Purposes for Applying the Air Measurement Technology

Table 4 shows the frequency and number of times, by pollutant, that the 16 different air monitoring applications were discussed in the 48 information sources that contained quantitative performance information. The last row shows the frequency and number of times that quantitative performance requirements were found for a pollutant regardless of the application (this is information identical to that in the right-most column of Table 3). The frequencies are calculated as described in Section 3.1.

Supplemental monitoring was most often cited as the purpose for collecting air pollution measurements, followed by community near-source monitoring, public education, and hot-spot detection. Monitoring of SO_2 appears to be frequently cited for a range of applications but is an artifact resulting from the relative infrequency with which performance requirements for SO_2 were found.

Table 4. Frequency with Which and Number of Times Air Monitoring was Performed, by Pollutant and Application

Application[a]	$PM_{2.5}$	PM_{10}	Carbon Monoxide (CO)	Nitrogen Dioxide (NO_2)	Sulfur Dioxide (SO_2)	Ozone (O_3)
Air Quality Forecasting	16% (3)	23% (3)	12% (2)	14% (3)	40% (2)	8% (2)
Air Quality Index Reporting	26% (5)	31% (4)	24% (4)	23% (5)	40% (2)	16% (4)
Community Near-Source Monitoring	42% (8)	38% (5)	35% (6)	36% (8)	60% (3)	48% (12)
Control Strategy	32% (6)	46% (6)	18% (3)	18% (4)	40% (2)	24% (6)
Data Fusion	16% (3)	23% (3)	12% (2)	18% (4)	40% (2)	8% (2)
Emergency Response	21% (4)	31% (4)	18% (3)	14% (3)	40% (2)	8% (2)
Epidemiological Studies	42% (8)	46% (6)	24% (4)	27% (6)	40% (2)	28% (7)
Exposure Reduction	16% (3)	15% (2)	35% (6)	23% (5)	40% (2)	20% (5)
Hot-Spot Detection	42% (8)	38% (5)	18% (3)	23% (5)	60% (3)	20% (5)
Model Input	16% (3)	23% (3)	12% (2)	18% (4)	40% (2)	8% (2)
Model Verification	21% (4)	31% (4)	18% (3)	18% (4)	40% (2)	16% (4)
Process Study Research	16% (3)	23% (3)	12% (2)	14% (3)	40% (2)	8% (2)
Public Education	37% (7)	38% (5)	29% (5)	32% (7)	60% (3)	16% (4)
Source Identification	16% (3)	23% (3)	35% (6)	32% (7)	40% (2)	20% (5)
Supplemental Monitoring	68% (13)	62% (8)	47% (8)	50% (11)	80% (4)	56% (14)
Other[b]	11% (2)	8% (1)	12% (2)	23% (5)	20% (1)	12% (3)
% All Information Sources	40% (19)	27% (13)	35% (17)	46% (22)	10% (5)	52% (25)

[a] Totals across all of the air monitoring applications for a given pollutant are always greater than the figures shown in the last row because a single information source may contain performance requirements for more than one pollutant and/or application.

[b] The "Other" category captures all applications not among the 16 shown.

3.3 Factors Driving the Selection of Air Measurement Technology

Figure 3 shows, in order of decreasing frequency, the various factors identified in the 56 information sources (those containing both qualitative [42-49] and quantitative performance metrics) as driving the selection of air monitoring technologies for all applications and data analysis types taken together.

Overall, high spatial density, cost, and accuracy/uncertainty are the main drivers across all data analysis types. However, once the results are stratified, there is the preference for regulatory monitoring toward attainment of high accuracy, precision, and excellent selectivity; whereas for non-regulatory monitoring purposes, accuracy remains important but high spatial density and cost supersede precision and selectivity.

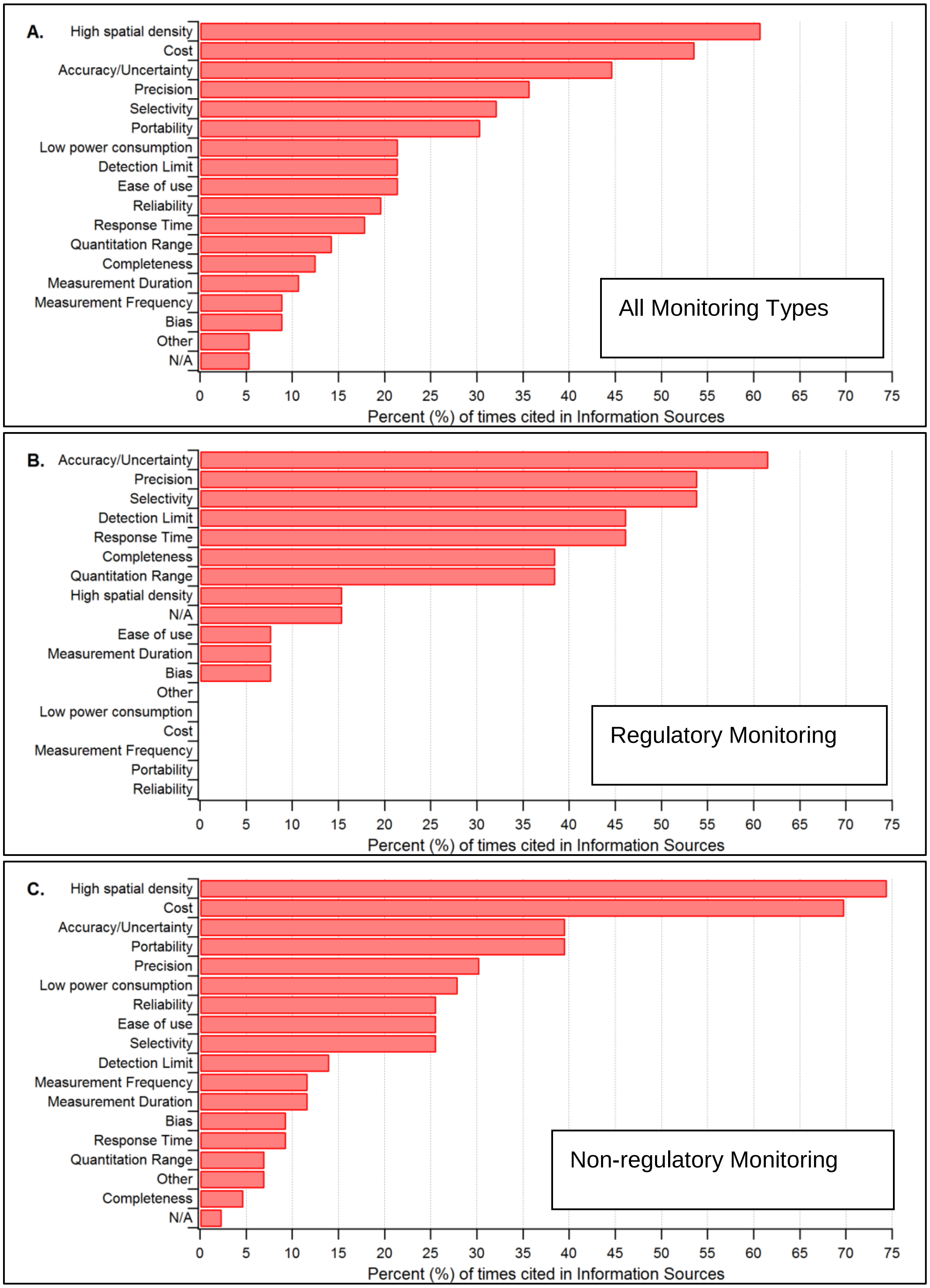


Figure 3. Frequency with Which Different Factors were Identified as Driving the Selection of Air Monitoring Technologies for: (A) All Data Analysis Types; (B) Decision Support (Regulatory Monitoring); and (C) Spatiotemporal, Comparison, and Trends Work (Non-regulatory Monitoring)

3.4 Concentration Ranges for Different Applications

Table 5 gives the concentration ranges for ambient and near-source measurements as reported in the 56 information sources (those containing both qualitative and quantitative performance metrics, stratified by pollutant and non-regulatory air monitoring application). References to the specific information sources reporting the concentration ranges are shown in brackets.

Table 5. Reported Concentration Ranges by Non-Regulatory Application and Pollutant

Application	$PM_{2.5}$ ($\mu g/m^3$)	PM_{10} ($\mu g/m^3$)	Carbon Monoxide (CO) (ppb)	Nitrogen Dioxide (NO_2) (ppb)	Sulfur Dioxide (SO_2) (ppb)	Ozone (O_3) (ppb)
Air Quality Forecasting	--	--	--	--	--	--
Air Quality Index Reporting	0-60 [50, 51]	5-25 [52]	350-1000 [51]	10-100 [51]	--	0-45 [51]
Community Near-Source Monitoring	8-400 [53, 54]	--	84-1706 [55]	0-140 [55, 56]	--	0-500 [55-60]
Control Strategy	--	--	--	--	--	0-500 [57, 58]
Data Fusion	--	--	--	5-95 [61]	--	--
Emergency Response	--	50-150 [62]	--	--	--	--
Epidemiological Studies	0-150 [54, 63-65]	0-150 [64, 65]	0-1706 [55, 64]	0-95 [55, 64, 66]	0.8-4.2[a] [64]	0-99 [55] [64, 66, 67]
Exposure Reduction	--	--	150-6000 [68, 69]	20-250 [66, 69]	--	0-45 [66, 67]
Hot-Spot Detection	0-400 [53, 54, 63]	50-150 [62]	--	25-95 [66]	--	0-500 [58, 66]
Model Input	--	--	--	5-95 [61]	--	--
Model Verification	0-81 [64]	0-113 [64]	0-1360 [64]	2-50 [64]	0.8-4.2[a] [64]	0-45 [64, 67]
Process Study Research	--	--	--	--	--	--
Public Education	0-100 [63, 64]	0-113 [64]	0-1360 [64]	2-50 [64]	0.8-4.2[a] [64]	0-44 [64]
Source Identification	--	--	150-6000 [68, 69]	0-250 [56, 69]	--	0-140 [56]
Supplemental Monitoring	0-400 [53, 54, 63-65]	0-150 [62, 64, 65]	0-6000 [64, 69]	0-250 [56, 64, 69]	0.8-4.2[a] [64]	0-10000 [56, 58-60, 64, 67, 70, 71]
Other[b]	--	--	200-1000 [72]	5-50 [72]	--	0-100 [60]

[a] One-hour average concentration range determined via the EU standard method for measuring SO_2 concentration (EN 14212) [8] over a 2-week field campaign. Note that this is a lower and more narrow range than expected in ambient environments, that the unusual measurement range was not discussed in the information source, and that reported here is the measurement range from the reference instrument that was collocated with two low-cost sensor packages that reported more a more realistic concentration range of ~0-55 ppb SO_2.

[b] The "Other" category captures all applications not among the 16 shown.

Observations for the concentration measurement ranges include:

- No information was found for the air quality forecasting and process study research applications
- Supplemental monitoring typically requires measurements across the largest concentration range of all the different applications
- The largest ranges were reported for CO and O_3.

3.5 Handling of Measurement Artifacts

Among the 48 information sources containing quantitative performance requirements for air monitoring instruments, and excluding those sources pertaining to regulatory air monitoring (11), 70% (26 of 37) adjusted for measurement artifacts, 8% (3 of 37) intentionally chose not to perform adjustments for artifacts, and for 22% (8 of 37) of the studies, such adjustments were not applicable.

Examples of the adjustments that were performed are given in Table 6, sorted by the general type of measurement artifact. Some of the most frequent adjustments were made to account for cross-sensitivity and interference both from other airborne species and from changes to temperature and relative humidity.

Table 6. Types of Measurement Artifacts and Typical Adjustments Performed

Type of Artifact	Adjustment Type
Calibration	Adjusted for baseline drift *"Sensor baseline drift is corrected, for every sensor for every measurement, via a linear (time-dependent) correction function."* [59] *"O_3 sensors were periodically calibrated in the field against the nearby reference instrument between 1-4 AM when O_3 concentrations are relatively homogeneous."* [66]
	Adjustments made during calibration against reference instrument *Calibration equation derived using sensor signal (corrected for variations in temperature, relative humidity, and signal drift) compared to the FEM concentration.* [71]
	Model adjustments *"Interferences from variable ambient gas concentration mix, sensor flow-cell temperature changes, and relative humidity changes were corrected with a high-dimensional model representation (HDMR)."* [55] *"Sensor is sensitive to temperature and humidity, so correction equations which implement temperature and humidity measurements were used."* [53] *"Meteorological filter was applied to remove local influences."* [68]
Calibration, meteorological	Cross interferences *"Temperature, relative humidity, and cross-interference corrections were applied following procedures described in reference [69]"* [64] *"The data were post-processed by the manufacturer with the aim to correct cross-interferences as well as the effect of temperature and relative humidity. Platform includes an O_3-filtered NO_2 sensor from Alphasense, designed to reject O_3 and hence eliminate cross-sensitivity issues."* [61]

Table 6. Types of Measurement Artifacts and Typical Adjustments Performed (continued)

Type of Artifact	Adjustment Type
	Temperature and relative humidity adjustments *"Algorithms derived from laboratory tests under various temperature and relative humidity conditions were applied to compensate for the impact of variations in reported concentrations."* [51] "*Lab experiment tests the instruments to see how their performance are impacted by relative humidity."* [73] *"Meteorological adjustments made in the model. Can include sinusoidal seasonal adjustments or allow for different slopes and/or intercepts for each quarter/season*." [50] *Response of all three sensors (CO, NO, NO_2) were adjusted for variation in ambient temperature (which also accounted for variation in relative humidity); adjustment needed (but not made) for O_3 on NO_2 sensor ([NO_2]sensor = [NO_2]ambient + [O_3]ambient)."* [69] *"Temperature-dependent baseline changes to concentration measurements are corrected for with a temperature-dependent equation."* [72] *"Sensor responses were adjusted for temperature/relative humidity (per manufacturer's built-in algorithms); but high bias and inter-sensor variability were nonetheless observed."* [52]
Miscellaneous	Comments and insights *"No adjustments mentioned due to meteorological conditions, but they do acknowledge "Changes in ambient water vapor and temperature have long been known to affect sensor performance, but there is also potential interference due to exposure and response to other co-pollutants. Our aim was to establish the selectivity of these sensors to their target compounds, and quantitatively characterize chemical interference to other pollutants. We then evaluated the scale of impacts of co-pollutants through an inter-comparison exercise alongside reference measurements of the same pollutants in ambient air."* [67] *"It is noted that abrupt changes in temperature and relative humidity can affect the performance of the sensors, and that is why both parameters are measured concurrently with concentration data. It does not detail how to account for these impacts, however."* [74] *"No data adjustments, but instruments were housed in heated environment to protect from water and extreme temperatures."* [54] *"No adjustments made but recognized that both meteorological conditions and aerosol conditions can influence measurements and should be accounted for when sensor measurements are calibrated with established standards."* [62] Quality control *"Instrument failure can be detected via remote diagnostics. For pump (air flow) degradation: measure heater current vs potential difference across it. For sensor failure (semiconductor structure variation): measure zero-ozone resistance of sensor versus time."* [60]

3.6 Important Performance Attributes

The frequency and number of times that quantitative DQOs/MQOs were available for the various performance characteristics (DQIs) are shown in Table 7. Given in the last row is the frequency and number of times (out of 48) that quantitative performance requirements were found for a pollutant regardless of the specific performance characteristic (this is information identical to that in the right-most column of Table 3 and last row of Table 4). The frequencies are calculated as described in Section 3.1. Cited most often were DQOs/MQOs for accuracy/uncertainty, followed by precision, measurement range, and detection limit.

Table 7. Frequency and Number of Times Information Sources Contained DQOs/MQOs for Different Performance Attributes

Performance Characteristic/DQI[a]	$PM_{2.5}$	PM_{10}	Carbon Monoxide (CO)	Nitrogen Dioxide (NO_2)	Sulfur Dioxide (SO_2)	Ozone (O_3)
Accuracy/Uncertainty	84% (16)	77% (10)	65% (11)	68% (15)	80% (4)	76% (19)
Bias	5% (1)	8% (1)	18% (3)	9% (2)	40% (2)	16% (4)
Completeness	26% (5)	31% (4)	12% (2)	14% (3)	40% (2)	16% (4)
Detection Limit	26% (5)	8% (1)	47% (8)	32% (7)	80% (4)	24% (6)
Measurement Duration	26% (5)	8% (1)	18% (3)	14% (3)	0% (0)	20% (5)
Measurement Frequency	26% (5)	15% (2)	35% (6)	23% (5)	0% (0)	32% (8)
Measurement Range	47% (9)	46% (6)	35% (6)	32% (7)	80% (4)	40% (10)
Precision	42% (8)	31% (4)	29% (5)	36% (8)	80% (4)	32% (8)
Response Time	0% (0)	0% (0)	29% (5)	32% (7)	80% (4)	20% (5)
Selectivity	11% (2)	8% (1)	24% (4)	23% (5)	80% (4)	16% (4)
Other[b]	5% (1)	8% (1)	0% (0)	0% (0)	0% (0)	8% (2)
% All Information Sources	40% (19)	27% (13)	35% (17)	46% (22)	10% (5)	52% (25)

[a] Totals across all performance characteristics for a given pollutant are always greater than the figures shown in the last row because a single information source may contain performance requirements for more than one pollutant and/or performance characteristic.

[b] The “Other” category captures all performance characteristics not among the 10 shown.

3.7 Comparison to Reference Instruments

Among the 48 information sources containing quantitative performance requirements for air monitoring instruments, and excluding those sources pertaining to regulatory air monitoring (11), 68% (25 of 37) compared air pollution measurements to a reference instrument of some type, 11% (4 of 37) did not, and such a comparison was not applicable for the remaining 21% (8 of 37). Table 8 provides the outcomes of these comparisons to relevant reference instruments.

Table 8. Outcomes of Comparisons as both Ranges and Medians (if comparisons were provided from at least three information sources) of the Reported Coefficients of Determination (r^2), of Air Monitors to Reference Instruments, by Pollutant

$PM_{2.5}$	PM_{10}	Carbon Monoxide (CO)	Nitrogen Dioxide (NO_2)	Sulfur Dioxide (SO_2)	Ozone (O_3)
0.07-0.91; 0.78 [50, 53, 63-65]	0.13-0.91; 0.36 [52, 64, 65]	0.53-0.87 [64]	0.02-0.96; 0.89 [61, 64, 66]	0.09-0.20 [64]	0.12-0.98; 0.9 [64, 66, 67, 70, 71]

A coefficient of determination ranges from 0 to 1 and, in the case of a comparison (based on a linear regression) of a "lower quality" air monitoring instrument to a reference instrument, indicates to a first approximation the extent of agreement between the two. Values closer to 1 indicate closer agreement and therefore greater accuracy and lower uncertainty. A wide range of performance against a reference method was reported in the literature for all the pollutants, which highlights the presence of a range of data quality issues with the use of lower-cost sensors.

3.8 Handling of Erroneous Data

As shown in Figure 4, among the 48 information sources containing quantitative performance requirements for air monitoring instruments, and excluding those sources pertaining to regulatory air monitoring (11), 2 of 37 (5%) reported and qualified erroneous data, 2 of 37 (5%) reported and did not qualify such data, 7 of 37 (19%) invalidated and did not report erroneous data, and 2 of 37 (5%) took other action. This information was not discussed for 24 of the 37 (65%) remaining information sources. Table 9 gives two different examples of how erroneous data were treated when they were invalidated and not reported.

Treatment of erroneous data was discussed in only 35% (13 of the 37) of the non-regulatory information sources with quantitative performance specifications. That most of the studies captured in this synthesis (which primarily included those using lower-cost sensors) did not explicitly discuss how to treat erroneous data indicates the need for more guidance on proper techniques and procedures for managing such data. This is especially important given the many possible data quality issues encountered with lower-cost air sensor measurements.

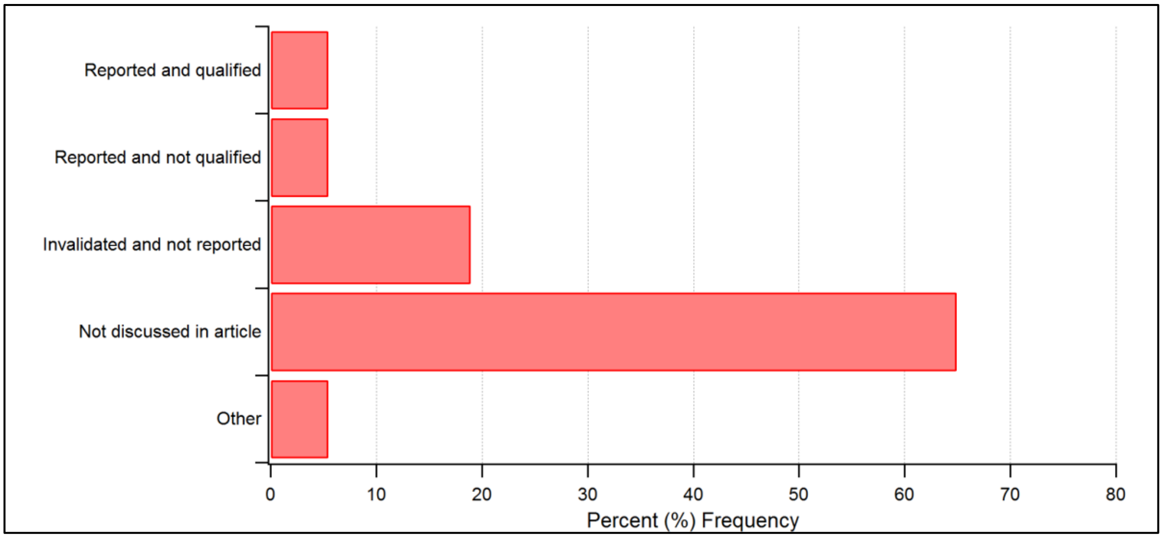


Figure 4. Frequency of Various Treatments of Erroneous Data

Table 9. Examples of Treatment of Erroneous Data

How Erroneous Data Were Handled	Specific Procedure
Invalidated and not reported	Data discarded due to temperature effects and electromagnetic interferences from every two-hour data transmission. [68]
Invalidated and not reported	Data from the deployment period were filtered to eliminate points that had temperature and relative humidity values out of the ranges recorded during calibration. Logged data were collected into minute medians to reduce the influence of outliers within each minute. Based on reference instrument data during deployment, 171 parts per billion by volume (ppbv) was set as a maximum level of ozone, with any sensor concentration above this threshold removed. Lastly, data were omitted when they fell more than 8 standard deviations away from the mean consecutive difference in values. [71]

3.9 International Non-regulatory Technology Performance Standards for Criteria Pollutants, Justification for Setting the Performance Requirements, and Pollutants to Which the Standards Apply

Two sets of international air sensor performance standards were identified. China's MEP and Environmental Protection Department of Hebei Province have developed generic performance standards for sensors, and Technical Committee 264-Air Quality, Working Group 42 in the EU is developing technical specifications for gas sensors.

The MEP Sensor Performance Standard and Hebei Local Standard contain DQOs for air sensors for $PM_{2.5}$, PM_{10}, CO, NO_2, SO_2, O_3, and total VOCs, but do not provide DQOs for different applications, and as no justification is available for how the DQOs were determined, are of limited usefulness. However, these standards do provide guidance for locating sensor networks (referred to as grids) for various applications: [19-21]

- Ambient Air Quality Monitoring Grid
- Pollution Source Area Monitoring Grid (includes: Road Traffic Grid, Dust Grid on Construction Site, Gas-related Enterprise Grid, Industrial Park Grid, Life Source Grid)
- Gradient Station Selection (vertical deployment from 10 to 300 meters above ground level)

Working Group 42 is creating technical specifications for sensors that measure gaseous pollutants such as O_3, NO and NO_2, CO, SO_2 and benzene [18] with a focus on fixed sensor systems and not on mobile devices, networks of sensor nodes, or indoor air monitoring. The rationale underpinning the selection of the relevant technical specifications is to map air sensor performance to the DQOs defined in the EU Ambient Air Quality Directive. Working Group 42 has proposed three different classes (i.e., tiers) for sensor performance based on the EU Ambient Air Quality Directive that provides some indication of the applications, as shown in Table 10. Of particular importance is that Working Group 42 does not expect air sensors to be fit for the purpose of fixed monitoring for regulatory compliance/decision support.

Table 10. Comparison of the Different Performance Tiers for the European Ambient Air Quality Directive and Working Group 42's Draft Specifications for Air Sensors

Example DQOs for $PM_{2.5}$ and O_3 Monitoring	**CEN Air Quality Directive (2008/50/EC)**	**CEN, Technical Committee 264-Air Quality, Working Group 42**
Uncertainty: **$PM_{2.5}$: 25%** **O_3: 15%** **Minimum data capture:** **$PM_{2.5}$: 90%** **O_3: 90% during summer; 75% during winter**	**Fixed measurements** *Highest quality, used for trends and compliance (decision support)*	**Not applicable**
Uncertainty: **$PM_{2.5}$: 50%** **O_3: 30%** **Minimum data capture:** **$PM_{2.5}$: 90%** **O_3: 90%** **Minimum time coverage:** **$PM_{2.5}$: 14%** **O_3: >10% during summer**	**Indicative measurements** *DQOs that are less strict than those required for fixed measurements*	**Class 1 Sensor System** *Consistent with the DQOs of indicative measurements set in the AQD*
Uncertainty: **$PM_{2.5}$: 100%** **O_3: 75%**	**Objective estimation** *Supplemental information that can be used to assess ambient air quality at levels below the lower assessment threshold*	**Class 2 Sensor System** *Consistent with the DQOs of objective estimation techniques set in the AQD*
Not applicable	**Not applicable**	**Class 3 Sensor System** Can be used for research, educational purposes, and citizen information

Performance requirements for air monitoring instruments nominally decrease in stringency from top to bottom.

3.10 Commonalities and Differences among Air Measurement Performance Requirements for Different Purposes

Tables B1 through B6 (found in Appendix B) give the descriptors and DQOs/MQOs for 10 different performance attributes for four different data analysis types for $PM_{2.5}$, PM_{10}, CO, NO_2, SO_2, and O_3. This section summarizes what qualitative patterns exist, if any, in the performance requirements, by attribute, both across the data analysis types and decisions sought. Where applicable, it is indicated where inconsistent data precluded an assessment of the presence of such patterns.

In general, the *a priori* expectation is that the performance requirements will increase in stringency from left to right across the columns of Tables B1 through B6, where measurements performed for spatial or temporal analysis must in general be of lesser quality (e.g., they may have greater imprecision) than those measurements used for comparison to a threshold value, analysis of trends over longer periods, and for decision support. The decision support column includes only regulatory air monitoring performance requirements. Performance requirements without data are not discussed. DQOs/MQOs were found for some performance characteristics, but in many cases, there were too few DQOs/MQOs available to discern a pattern.

Major and cross-cutting findings include:

- Decision support has the strictest performance requirements for accuracy, completeness, detection limit, and precision.
- Required measurement durations are shorter for spatiotemporal, comparison, and trend data analyses, which is consistent with the conclusion that higher time resolution data are required for these applications.
- In many instances, inconsistencies in the types of descriptors (e.g., root mean squared error [RMSE] and coefficient of determination [r^2]) for a given DQI (e.g., accuracy/uncertainty) and units for the DQOs/MQOs (µg/m^3 for RMSE, unitless for r^2) precluded the evaluation and detection of patterns in performance requirements. Due to resource limitations, no effort was undertaken to normalize the DQOs/MQOs for different descriptors for a given DQI. Such harmonization would likely have improved the ability to draw conclusions about trends in accuracy/uncertainty.
- Due to a combination of inadequate and inconsistent information, non-regulatory air monitoring performance requirements cannot be stratified into tiers or categories of performance.

The following qualitative patterns in the air measurement performance attributes were observed for each pollutant. The reported numerical DQOs/MQOs may be found in Tables B1 to B6.

Particulate Matter ($PM_{2.5}$)

- Accuracy/uncertainty – higher r^2 for decision support compared to spatiotemporal
- Completeness – higher requirements for decision support
- Detection limit – lower detection limit for decision support
- Measurement duration – shorter measurement duration for comparison and spatiotemporal data analyses

- Measurement range – smaller concentration range (0-200 µg/m^3) for comparison and spatiotemporal compared to larger ranges (0-1000 µg/m^3) for European Union and China Standards under decision support.
- Precision – lower CV for concentration and flow for decision support

Particulate Matter (PM_{10})

- Accuracy/uncertainty – higher r^2 for decision support
- Completeness – higher completeness requirements for decision support
- Measurement duration – shorter measurement duration for comparison data analyses as compared to decision support
- Precision – lower CV for concentration and flow for decision support

Carbon Monoxide (CO)

- Accuracy/uncertainty – inconsistent information
- Completeness – highest completeness requirements for decision support; however, data capture is limited to two information sources
- Measurement range – higher measurement ranges for non-regulatory air monitoring work (all but decision support-related applications)

Nitrogen Dioxide (NO_2)

- Accuracy/uncertainty – performance requirements (based on the descriptor of %difference [%Diff]) increase from left to right across the table; however, data capture is limited to two information sources
- Completeness – higher completeness requirements for decision support
- Detection limit – inconsistent information
- Measurement frequency – no pattern present
- Measurement range – higher measurement ranges for non-regulatory air monitoring work (all but decision support-related applications)
- Precision – no pattern present
- Response time – appears to be a pattern in that faster response times are needed for spatiotemporal air monitoring as compared to decision support applications

Sulfur Dioxide (SO_2)

- Accuracy/uncertainty – inconsistent information
- Completeness – highest requirements for decision support; however, data capture is limited to two information sources

Ozone (O_3)

- Accuracy/uncertainty – inconsistent information

- Completeness – highest requirements for decision support
- Measurement duration – monitoring to discern spatiotemporal variations requires shorter measurement durations as compared to longer-term trends monitoring, in accord with expectations
- Measurement frequency – similar across comparison, spatiotemporal, and trends monitoring applications
- Measurement range – higher measurement ranges are required for non-regulatory air monitoring work (all but decision support-related applications)
- Response time – faster response times are needed for non-regulatory purposes such as spatiotemporal trends monitoring; note that data are limited (one spatiotemporal study, three regulatory monitoring methods)
- Precision – no pattern present

3.11 Commonalities and Differences in the Regulatory Monitoring Requirements for Criteria Pollutants in the US Compared to International Requirements

The regulatory monitoring requirements in the U.S, EU, and China shown under Decision Support were extracted from Tables B1 to B6 and are summarized in Table 11 below. The DQOs/MQOs for 4 of the 10 performance attributes (DQIs) are given; patterns in DQIs could not be ascertained for bias, completeness, measurement duration, measurement frequency, precision, and selectivity due to the lack of consistent and sufficient information on performance requirements.

Table 11. US, European Union and Chinese Regulatory Monitoring Performance Requirements for Accuracy/Uncertainty, Detection Limit, Measurement Range, and Response Time for Measurements of $PM_{2.5}$, PM_{10}, CO, NO_2, SO_2, and O_3

Pollutant	Performance Attribute	US	EU	China
$PM_{2.5}$	Accuracy/uncertainty	R^2: 0.7225-0.9025 [1]		$R^2 \geq 0.8649$ [2]
	Measurement range	Measurement range: 3-200 µg/m³ [1]	Measurement range: (0-1000$_{24h\text{-}avg}$, 0-10000$_{1h\text{-}avg}$ µg/m³) [3]	Measurement range: 0-1000 µg/m³ [2]
PM_{10}	Accuracy/uncertainty	$R^2 \geq 0.9409$ [1]		$R^2 \geq 0.9025$ [2]
	Measurement range	0-300 µg/m³ [1]	(0-1000$_{24h_avg}$, 0-10,000$_{1hr_avg}$) µg/m³ [3]	0 – 1000 µg/m³ [2]
CO	Response time	Rise & Fall time: 120 sec [1],	Rise & Fall time: ≤180 sec [4]	Response time: ≤240 sec [5]
NO_2	Accuracy/uncertainty	12-hr zero drift: ±20 ppb [1] 24-hr zero drift: ±20 ppb [1] 24-hr 80% span drift: ±5.0 % [1] 24-hr 20% span drift: ±20.0% [1]	12-hr zero drift: ≤2.0 ppb [6] 12-hr span drift (ppb): ≤6.0 [6] Long-term zero drift: ≤5.0 ppb [6]	24-hr zero drift: ±5 ppb [5] 24-hr 80% span drift: ±10 ppb [5] 24-hr 20% span drift: ±5 ppb [5] Long-term zero drift: ±10 ppb [5]
	Detection limit	Detection limit: 10 ppb [1]		Detection limit: ≤2 ppb [5]
	Response time	Rise & Fall time: 15 min [1] Residence time: <2 min [7]	Rise & Fall time: ≤180 s [6] Residence time: ≤ 3.0 sec [6]	Response time: ≤5 min [5]
	Measurement range	Measurement range: 0-500 ppb [1]	Measurement range: ≤ 261 ppb [6]	Measurement range: 0-500 ppb [5]
SO_2	Accuracy/uncertainty	12-hr zero drift: ±4 ppb [1] 24-hr zero drift: ±4 ppb [1]	12-hr zero drift: ≤2.0 ppb [8]	24-hr zero drift: ±5 ppb [5]
	Response time	Rise & Fall time: 120 sec [1]	Rise & Fall time: ≤180 sec [8]	Response time: ≤5 min [5]
O_3	Accuracy/uncertainty	24-hr zero drift: ±4 ppb [1]		24-hr zero drift: ±5 ppb [5]
	Measurement range	Measurement range: 0-500 ppb [1]	Measurement range: ≤250 ppb [9]	Measurement range: 0-500 ppb [5]
	Detection limit	Detection limit: 5 ppb [1]		Detection limit: ≤2 ppb [5]
	Response time	Lag & Rise time: 120 sec [1]	Lag & Rise time: ≤180 sec [9]	Response time: ≤5 min [5]

4 Summary

Substantial effort was invested to identify the information sources included in this literature review and synthesis. The process consisted of both automated and manual searches to identify relevant information sources from among those in the peer-reviewed literature, technical reports, theses and dissertations, and regulatory air monitoring standards promulgated by government agencies, among others. Quantitative performance requirements (DQOs/MQOs) were captured for ten different performance attributes/characteristics, and the DQOs/MQOs were organized by 16 different air monitoring application types and binned into four broad categories, irrespective of the application. The performance requirements for regulatory air monitoring in the US, EU, and China were also captured along with information from the various extant and developing domestic and international air sensor performance evaluation and standard-setting programs.

The results of the information review and synthesis are captured in the bullets below. In summary, more information and research is needed to determine the fit-for-purpose air monitoring performance requirements.

- A total of 257 sources were located and assessed for applicability and utility; 48 (19%) contained quantitative performance information and 8 (3%) contained qualitative performance information. Thus, 56 (22%) of these information sources were included in the information synthesis presented in this report.
- Performance requirements were found most frequently for the spatiotemporal variation data analysis type (40 to 72% of the time) and, more generally, quantitative air monitoring performance requirements detailing fitness for a given purpose were most abundant for O_3 (52%), followed by NO_2 (46%), then $PM_{2.5}$ (40%).
- Supplemental monitoring was most often cited as the purpose for collecting air pollution measurements, followed by community near-source monitoring, public education, and hot-spot detection.
- Across all data analysis types, high spatial density, cost, and accuracy/uncertainty are the main drivers for selection of air monitoring technologies. However, once the results are stratified, there is, in accord with expectations, the preference for regulatory monitoring toward attainment of high accuracy, precision, and excellent selectivity, whereas for non-regulatory monitoring purposes, accuracy remains important but high spatial density and cost supersede precision and selectivity.
- Observations for pollutant concentration measurement range include that:
  - No information was found for the air quality forecasting and process study research applications
  - Supplemental monitoring typically requires measurements across the largest concentration range of all the different applications
  - The largest ranges were reported for CO and O_3.

  Among the 48 information sources containing quantitative performance requirements for air monitoring instruments, and excluding those sources pertaining to regulatory air monitoring (11), 70% (26 of 37) adjusted for measurement artifacts, 8% (3 of 37) intentionally chose not to perform adjustments for artifacts, and for 22% (8 of 37) of the studies, such adjustments were not applicable. Some of the most frequent adjustments were made to account for cross-sensitivity and interference both from other airborne species and from changes in temperature and relative humidity.

- DQOs/MQOs for the various performance attributes (DQIs) were given most often for accuracy/uncertainty, followed by precision, measurement range, and detection limit.
- Among the 48 information sources containing quantitative performance requirements for air monitoring instruments, and excluding those sources pertaining to regulatory air monitoring (11), 68% (25 of 37) compared air pollution measurements to a reference instrument of some type, 11% (4 of 37) did not, and such a comparison was not applicable for the remaining 21% (8 of 37). A wide range of performance against a reference method was reported in the literature for all the pollutants, demonstrating the presence of a range of data quality issues with the use of lower-cost sensors.
- Treatment of erroneous data was discussed in only 35% (13 of the 37) of the non-regulatory information sources with quantitative performance specifications. That most of the studies captured in this synthesis (which primarily included those using lower-cost sensors) did not explicitly discuss how to treat erroneous data indicates the need for more guidance on proper techniques and procedures on how to treat such data. This is especially important given the many potential data quality issues encountered in lower-cost air sensor measurements.
- Two different international air sensor performance standards were identified. China’s MEP and Environmental Protection Department of Hebei Province have developed generic performance standards for sensors, and CEN, Technical Committee 264-Air Quality, Working Group 42 is developing technical specifications for gas sensors.
- In general, the *a priori* expectation was that the air monitoring performance requirements will increase in stringency (spatiotemporal < comparison < trend < decision support) where measurements performed for spatial or temporal analysis must in general be of lesser quality (e.g., they may have greater imprecision) than those measurements used for comparison to a threshold value, analysis of trends over longer periods, and for decision support. Major and cross-cutting findings include:
    - Decision support has the strictest performance requirements for accuracy, completeness, detection limit, and precision.
    - Required measurement durations are shorter for spatiotemporal, comparison, and trend data analyses, which is consistent with the conclusion that higher time resolution data are required for these applications.
    - In many instances inconsistencies in the types of descriptors precluded the evaluation and detection of patterns in performance requirements. Due to resource limitations, no effort was undertaken to normalize the DQOs/MQOs for different descriptors for a given DQI. Such harmonization would likely have improved the ability to draw conclusions about trends in accuracy/uncertainty.
    - Due to a combination of inadequate and inconsistent information, non-regulatory air monitoring performance requirements cannot be stratified into tiers or categories of performance.

- DQOs/MQOs for the DQIs accuracy/uncertainty, detection limit, measurement range, and response time for regulatory monitoring in the US, EU and China. Patterns in DQIs could not be ascertained for bias, completeness, measurement duration, measurement frequency, precision and selectivity due to the lack of consistent and sufficient information on performance requirements.

5 References

Table 12 lists the references that correspond to those (1) cited for background information; (2) that are included in the 48 that contained quantitative performance metrics; and (3) that are included in the eight that contained qualitative performance information.

Table 12. Specific References that Belong to Different Groups of Information Sources

Type of Information Source	Specific References
Background information	[11-22, 24-41, 75-85]
Quantitative Performance Requirements	[1-10, 23, 50-66, 68-74, 86-98]
Qualitative Performance information	[42-49]

1. US EPA, *CFR Part 53: Ambient Air Monitoring Reference and Equivalent Methods.* 2018.
2. Chinese Ministry of Environmental Protection, *HJ 653-2013 Specifications and Test Procedures for Ambient Air Quality Continuous Automated Monitoring Systems for PM10 and PM2.5*. 2013.
3. European Committee for Standardization (CEN), *DS/EN 16450:2017 Ambient Air - Automated measuring systems for the measurement of the concentration of particulate matter (PM10; PM2.5)*. 2017: Brussels, Belgium.
4. European Committee for Standardization (CEN), *DS/EN 14626:2012 Ambient air - Standard method for the measurement of the concentration of carbon monoxide (CO) by non-dispersive infrared spectroscopy*. 2012: Brussels, Belgium.
5. Chinese Ministry of Environmental Protection, *HJ 654-2013 Specifications and Test Procedures for Ambient Air Quality Continuous Automated Monitoring Systems for SO2, NO2, O3 and CO*. 2013.
6. European Committee for Standardization (CEN), *DS/EN 14211:2012 Ambient air - Standard method for the measurement of the concentration of nitrogen dioxide (NO2) and nitrogen monoxide (NO) by chemiluminescence*. 2012: Brussels, Belgium.
7. US EPA, *CFR Part 50 - National Primary and Secondary Ambient Air Quality Standards.* 2018.
8. European Committee for Standardization (CEN), *DS/EN 14212:2012 Ambient air - Standard method for the measurement of the concentration of sulphur dioxide by ultraviolet fluorescence*. 2012: Brussels, Belgium.
9. European Committee for Standardization (CEN), *DS/EN 14625:2012 Ambient air - Standard method for the measurement of the concentration of ozone (O3) by ultraviolet photometry*. 2012: Brussels, Belgium.
10. US EPA, Office of Research and Development, National Exposure Research Laboratory, *EPA 600/R-14/159 Air Sensor Guidebook*. 2014: Research Triangle Park, NC, USA.
11. The European Parliament and the Council of the European Union, *Directive 2008/50/ED of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe*. 2008.
12. Environment Agency, *MCERTS: Performance Standards for Indicative Ambient Particulate Monitors*. 2017.

13. Environment Agency, *MCERTS: Performance Standards for Continuous Ambient Air Quality Monitoring Systems*. 2016.

14. Environment Agency, *MCERTS: Performance Standards for Open Path Ambient Air Quality Monitoring Systems*. 2017.

15. Environment Agency, *MCERTS: Performance Standards for Portable Emission Monitoring Systems*. 2014.

16. Environment Agency, *MCERTS: Performance Standards and Test Procedures for Continuous Emission Monitoring Systems*. 2016.

17. US EPA, *Performance Specification 18 - Performance Specifications and Test Procedures for Gaseous Hydrogen Chloride (HCl) Continuous Emission Monitoring Systems at Stationary Sources*. 2017.

18. Dye, T., *Personal communication with Michel Gerbolis*. 2018.

19. Environmental Protection Department of Hebei Province (China), *DB13/T2544-2017: Specifications and Test Procedures for Air Pollution Control Gridded Monitoring System*. 2017.

20. Environmental Protection Department of Hebei Province (China), *DB13/T2545-2017: Technical Regulation for Selecting the Location of Air Pollution Control Gridded Monitoring System*. 2017.

21. Environmental Protection Department of Hebei Province (China), *DB13/T2546-2017: Technical Specification for Installation Acceptance and Operating of Air Pollution Control Gridded Monitoring System*. 2017.

22. Papapostolou, V., et al., *Development of an environmental chamber for evaluating the performance of low-cost air quality sensors under controlled conditions.* Atmospheric Environment, 2017. **171**: p. 82-90.

23. European Committee for Standardization (CEN), *DS/EN 12341:2014 Ambient air - Standard gravimetric measurement method for the determination of the PM10 or PM2.5 mass concentration of suspended particulate matter*. 2014: Brussels, Belgium.

24. Union, T.E.P.a.t.C.o.t.E., *Commission Directive (EU) 2015/1480 of 28 August 2015 amending several annexes to Directives 2004/107/EC and 2008/50/EC of the European Parliament and of the Council laying down the rules concerning reference methods, data validation and location of sampling points for the asseessment of ambient air quality*. 2015. p. 8.

25. *TUV Rheinland (R) Precisely Right.* [cited 2018 June 4]; Available from: https://www.tuv.com/usa/en/.

26. Chinese Ministry of Environmental Protection, *GB 3095-2012 Ambient air quality standards.* 2012.

27. US EPA. *Environmental Technology Verification (ETV) Progam.* [Accessed May 7, 2018]; Available from: www.epa.gov/etv.

28. International Organization for Standardization (ISO). *14034:2016 - Environmental management -- Environmental technology verification (ETV)*. [Accessed May 7, 2018]; Available from: https://www.iso.org/obp/ui/#iso:std:iso:14034:ed-1:v1:en.

29. VerifiGlobal. *VerifiGlobal Alliance*. [May 7, 2018]; Available from: http://www.verifiglobal.com/en/.

30. US EPA, *EPA/100/B-03/001 A Summary of General Assessment Factors for Evaluating the Quality of Scientific and Technical Information*. 2003, Office of Research and Development, Science Policy Council: Washington, DC.

31. Battelle, *Quality Assurance Project Plan for "Comprehensive technical information review to inform air sensor technology performance targets"*. 2017: Columbus, OH.

32. Lewis, A.C., et al., *Technical advice note on lower cost air pollution sensors*. 2017, World Meteorlogical Organization and Global Atmosphere Watch.

33. US EPA, *EPA QA/G-5, Guidance for Quality Assurance Project Plans*. 2002: Washington, DC.

34. US EPA, *FEM Document Number 2005-01 Validation and Peer Review of US Environmental Protection Agency Chemical Methods of Analysis*. 2016.

35. ASTM, *D3249-95(2011): Standard Practice for General Ambient Air Analyzer Procedures*. 2011: West Conshohocken, PA.

36. ASTM, *D3670-91(2014): Standard Guide for Determination of Precision and Bias of Methods of Committee D22*. 2014, ASTM International: West Conshohocken, PA.

37. ASTM, *E2655-14: Standard Guide for Reporting Uncertainty of Test Results and Use of the Term Measurement Uncertainty in ASTM Test Methods*. 2014, ASTM International: West Conshohocken, PA.

38. Vessman, J., et al., *Selectivity in analytical chemistry (IUPAC Recommendations 2001)*, in *Pure and Applied Chemistry*. 2001. p. 1381-1386.

39. den Boef, G. and A. Hulanicki, *Recommendations for the usage of selective, selectivity and related terms in analytical chemistry*, in *Pure and Applied Chemistry*. 1983. p. 553-556.

40. International Organization for Standardization (ISO), *ISO 5725-1: 1994(E): Accuracy (trueness and precision) of measurement methods and results - Part 1: General principles and definitions*. 1994: Geneva, Switzerland.

41. International Organization for Standardization (ISO), *International vocabulary of metrology — Basic and general concepts and associated terms (VIM), 3rd Ed.* 2008, Geneva, Switzerland.

42. District, S.C.A.Q.M. *Air Quality Sensor Performance Evaluation Reports*. [cited 2018 March 22]; Available from: http://www.aqmd.gov/aq-spec/evaluations/summary.

43. Baldauf, R.W., et al., *Air quality variability near a highway in a complex urban environment.* Atmospheric Environment, 2013. **64**: p. 169-178.

44. Kaufman, A., et al., *A Citizen Science and Government Collaboration: Developing Tools to Facilitate Community Air Monitoring.* Environmental Justice, 2017. **10**(2): p. 51-61.

45. Kleffmann, J., et al., *NO2 Measurement Techniques: Pitfalls and New Developments*. 2013, Springer Netherlands: Dordrecht. p. 15-28.

46. Jovašević-Stojanović, M., et al., *On the use of small and cheaper sensors and devices for indicative citizen-based monitoring of respirable particulate matter.* Environmental Pollution, 2015. **206**: p. 696-704.

47. Castell, N., et al., *Real-world application of new sensor technologies for air quality monitoring.* 2013.

48. Kumar, P., et al., *The rise of low-cost sensing for managing air pollution in cities.* Environment International, 2015. **75**: p. 199-205.

49. Yi, W.Y., et al., *A Survey of Wireless Sensor Network Based Air Pollution Monitoring Systems.* Sensors (Basel), 2015. **15**(12): p. 31392-427.

50. US EPA, Office of Air Quality Planning and Standards,, *EPA-454/B-02-002 Data Quality Objectives (DQOs) for relating Federal Reference Method (FRM) and Continuous PM2.5 Measurements to Report an Air Quality Objective (AQI)*. 2002: Research Triangle Park, NC, USA.

51. Sun, L., et al., *Development and Application of a Next Generation Air Sensor Network for the Hong Kong Marathon 2015 Air Quality Monitoring.* Sensors, 2016. **16**(2): p. 211.

52. Castell, N., et al., *Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates?* Environment International, 2017. **99**: p. 293-302.

53. Gao, M., J. Cao, and E. Seto, *A distributed network of low-cost continuous reading sensors to measure spatiotemporal variations of PM2.5 in Xi'an, China.* Environmental Pollution, 2015. **199**: p. 56-65.

54. Zikova, N., et al., *Estimating Hourly Concentrations of PM2.5 across a Metropolitan Area Using Low-Cost Particle Monitors.* Sensors, 2017. **17**(8).

55. Cross, E.S., et al., *Use of electrochemical sensors for measurement of air pollution: Correcting interference response and validating measurements.* Atmospheric Measurement Techniques, 2017. **10**(9): p. 3575-3588.

56. Duvall, R., et al., *Performance Evaluation and Community Application of Low-Cost Sensors for Ozone and Nitrogen Dioxide.* Sensors, 2016. **16**(10): p. 1698.

57. Bart, M., et al., *High Density Ozone Monitoring Using Gas Sensitive Semi-Conductor Sensors in the Lower Fraser Valley, British Columbia.* Environmental Science & Technology, 2014. **48**(7): p. 3970-3977.

58. David, E.W., et al., *Validation of low-cost ozone measurement instruments suitable for use in an air-quality monitoring network.* Measurement Science and Technology, 2013. **24**(6).

59. MacDonald, C., et al., *Ozone Concentrations In and Around the City of Arvin, California*. 2014, Sonoma Technology, Inc.

60. Williams, D., et al., *Development of Low-Cost Ozone Measurement Instruments Suitable for Use in an Air Quality Monitoring Network.* Chemistry in New Zealand, 2009. **73**: p. 27-33.

61. Schneider, P., et al., *Mapping urban air quality in near real-time using observations from low-cost sensors and model information.* Environment International, 2017. **106**: p. 234-247.

62. Mukherjee, A., et al., *Assessing the Utility of Low-Cost Particulate Matter Sensors over a 12-Week Period in the Cuyama Valley of California.* Sensors (Basel), 2017. **17**(8).

63. Austin, E., et al., *Laboratory Evaluation of the Shinyei PPD42NS Low-Cost Particulate Matter Sensor.* Plos One, 2015. **10**(9).

64. Borrego, C., et al., *Assessment of air quality microsensors versus reference methods: The EuNetAir joint exercise.* Atmospheric Environment, 2016. **147**: p. 246-263.

65. Crilley, L.R., et al., *Evaluation of a low-cost optical particle counter (Alphasense OPC-N2) for ambient air monitoring.* Atmospheric Measurement Techniques, 2018. **11**(2): p. 709-720.

66. Moltchanov, S., et al., *On the feasibility of measuring urban air pollution by wireless distributed sensor networks.* Science of The Total Environment, 2015. **502**: p. 537-547.

67. Lewis, A.C., et al., *Evaluating the performance of low cost chemical sensors for air pollution research.* Faraday Discussions, 2016. **189**: p. 85-103.

68. Heimann, I., et al., *Source attribution of air pollution by spatial scale separation using high spatial density networks of low cost air quality sensors.* Atmospheric Environment, 2015. **113**: p. 10-19.

69. Mead, M.I., et al., *The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks.* Atmospheric Environment, 2013. **70**: p. 186-203.

70. Pang, X., et al., *Electrochemical ozone sensors: A miniaturised alternative for ozone measurements in laboratory experiments and air-quality monitoring.* Sensors and Actuators B: Chemical, 2017. **240**: p. 829-837.

71. Sadighi, K., et al., *Intra-urban spatial variability of surface ozone in Riverside, CA: viability and validation of low-cost sensors.* Atmospheric Measurement Techniques, 2018. **11**(3): p. 1777-1792.

72. Popoola, O.A.M., et al., *Development of a baseline-temperature correction methodology for electrochemical sensors and its implications for long-term stability.* Atmospheric Environment, 2016. **147**: p. 330-343.

73. Sonoma Technology Inc. (STI), *Characterization of low-cost NO2 sensors.* Draft final report, Prepared for US Environmental Protection Agency, 2010.

74. US EPA, Office of Research and Development,, *EPA 600/R-15/008 Citizen Science Air Monitor (CSAM) Quality Assurance Guidelines*. 2015.

75. US EPA, Office of Air Quality Planning and Standards,, *Measurement Quality Objectives and Validation Templates: Appendix D, QA Handbook VII*. 2017: Research Triangle Park, NC, USA.

76. US EPA, Office of Air Quality Planning and Standards,, *EPA-454/B-17-001 Quality Assurance Handbook for Air Pollution Measurement Systems Volume II*. 2017: Research Triangle Park, NC, USA.

77. Verbic, T., Z. Dorko, and G. Horvai, *Selectivity in analytical chemistry.* Revue Roumaine de Chimie, 2013. **58**: p. 569-575.

78. US EPA, Office of Water,, *EPA 821-R-16-006 Definition and Procedure for the Determination of the Method Detection Limit, Revision 2*. 2016: Washington, DC, USA.

79. Kimothi, S.K., *The Uncertainty of Measurements: Physical and Chemical Metrology Impact and Analysis*. 2002, Milwaukee, WI: ASQ Press.

80. ASTM, *D1356-17: Standard Terminology Relating to Sampling and Analysis of Atmospheres*. 2017, ASTM International: West Conshohocken, PA.

81. ASTM, *E3080-17: Standard Practice for Regression Analysis*. 2017, ASTM International: West Conshohocken, PA.

82. ASTM, *E2586-16: Standard Practice for Calculating and Using Basic Statistics*. 2016, ASTM International: West Conshohocken, PA, 2016.

83. ASTM, *E456-13A(2017)e1: Standard Terminology Relating to Quality and Statistics*. 2017, ASTM International: West Conshohocken, PA.

84. Spinelle, L., M. Gerboles, and M. Aleixandre, *Report of Laboratory and in-Situ Validation of Micro-Sensor for Monitoring Ambient Air Pollution-NO9: CairClipNO2 of CAIRPOL (F).* Joint Research Centre Technical Report, 2013.

85. US EPA, Office of Air Quality Planning and Standards,, *EPA-454/B-16-001 Quality Assurance Guidance Document 2.12 - Monitoring PM2.5 in Ambient Air Using Designated Reference or Class I Equivalent Methods*. 2016: Research Triangle Park, NC, USA.

86. Moore, K., et al., *Ambient Ozone Concentrations Cause Increased Hospitalizations for Asthma in Children: An 18-Year Study in Southern California.* Environmental Health Perspectives, 2008. **16**(8): p. 1063-1070.

87. Snyder, E.G., et al., *The Changing Paradigm of Air Pollution Monitoring.* Environmental Science & Technology, 2013. **47**(20): p. 11369-11377.

88. McKercher, G.R., J.A. Salmond, and J.K. Vanos, *Characteristics and applications of small, portable gaseous air pollution monitors.* Environmental Pollution, 2017. **223**: p. 102-110.

89. Jiao, W., et al., *Community Air Sensor Network (CAIRSENSE) project: evaluation of low-cost sensor performance in a suburban environment in the southeastern United States.* Atmospheric Measurement Techniques, 2016. **9**(11): p. 5281-5292.

90. European Environment Agency (EEA), *EEA Report No 7/2013 Air Implementation Pilot - Lessons learnt from the implementation of air quality legislation at urban level.* 2013: Copenhagen, Denmark.

91. European Environment Agency (EEA), *EEA Technical report No 1/2009 Spatial assessment of PM10 and ozone concentrations in Europe (2005)*. 2009: Copenhagen, Denmark.

92. European Environment Agency (EEA), *EEA Technical report No 7/2009 Assessment of ground-level ozone in EEA member countries, with a focus on long-term trends*. 2009: Copenhagen, Denmark.

93. Fishbain, B., et al., *An evaluation tool kit of air quality micro-sensing units.* Science of The Total Environment, 2017. **575**: p. 639-648.

94. Brienza, S., et al., *A low-cost sensing system for cooperative air quality monitoring in urban areas.* Sensors, 2015. **15**(6): p. 12242-59.

95. Michel, S.L., A. Manuel, and Gerboles, *Protocol of evaluation and calibration of low-cost gas sensors for the monitoring of air pollution*. 2013.

96. US EPA, *Quality Assurance Guidance Document 2.12 - Monitoring PM2.5 in Ambient Air Using Designated Reference or Class I Equivalent Methods.* Office of Air Quality Planning and Standards, 2016: p. 174.

97. Wang, A. and M. Brauer, *Review of Next Generation Air Monitors for Air Pollution*. 2014, The University of British Columbia: Vancouver, Canada.

98. White, R., et al., *Sensors and 'apps' for community-based: Atmospheric monitoring.* 2012: p. 36-40.

Appendix A. Definitions

Accepted Reference Value

A value that serves as an agreed-upon reference for comparison, and which is derived as: (1) a theoretical or established value, based on scientific principles, (2) an assigned or certified value, based on experimental work of some national or international organization, or (3) a consensus or certified value, based on collaborative experimental work under the auspices of a scientific or engineering group. [37]

Accuracy

- Accuracy is a measure of the overall agreement of a measurement with a known value (an accepted reference value). Accuracy includes a combination of systematic error (bias) and random error (precision). [33]
- Accuracy is sometimes confused with bias; the terms are used interchangeably. EPA recommends using the terms “precision” and “bias”, rather than “accuracy”, to convey the information usually associated with accuracy. [10]
- The degree of conformity of a value generated by a specific procedure to the assumed or accepted true value that includes both precision and bias. [80]
- The meaning of the term “accuracy” has changed over the years, and accuracy should be viewed as a qualitative concept rather than a synonym for bias. [34, 40, 79]
- A process is considered accurate only if it precise as well as unbiased. [79]

Bias

- The systematic or persistent distortion of a measurement process that causes error in one direction. [33]
    - Bias will be determined by estimating the positive and negative deviation from the true value as a percentage of the true (or accepted reference) value. [76]
    - The presence of systematic errors can only be determined by comparison of the average of many results with a reliable, accepted reference value. [34]
- An error in the measurement that is repeatable, which can be determined by taking multiple measurements with the sensor and comparing these data with the “true” concentration (or accepted reference value). The true concentration can be established by a reference monitor located in close proximity to the sensor. Bias means an average systematic or persistent distortion of a measurement process that causes errors in one direction. [10]
- A systematic (non-random) deviation of the method’s average value or the measured value from an accepted value. [80]
- The term bias has been in use for statistical matters for a very long time, but because it caused certain philosophical objections among members of some professions (such as medical and legal practitioners), the positive aspect has been emphasized by the invention of the term trueness. [40]

Comparability

- A measure of the confidence with which one data set or method can be compared to another, considering the units of measurement and applicability to standard statistical techniques. [76]

- A qualitative term that expresses the measure of confidence that one data set can be compared to another and can be combined for the decision(s) to be made. [33]

Completeness

- Describes the amount of valid data obtained from a measurement system compared to the amount that was expected to be obtained under correct, normal conditions. [76]

- A measure of the amount of valid data needed to be obtained from a measurement system. [33]

Coefficient of Determination

- The square of the correlation coefficient, r. [81]

- The coefficient of determination (r^2) varies from 0 to 1 and measures the proportion of the variance removed from the raw Y data by the regression model. [81]

Coefficient of Variation

- For a non-negative characteristic, the ratio of the standard deviation to the mean for a population or sample. [82]
    - Also known as the relative standard deviation (RSD)

Correlation Coefficient

- For a population, the correlation coefficient is a dimensionless measure of association between two variables X and Y, equal to the covariance divided by the product of σ_X and times σ_Y. [81]

Detection Limit / Limit of Detection (LOD) / Method Detection Limit (MDL)

- The lowest concentration or amount of the target analyte that can be determined to be different from zero by a single measurement at a stated level of probability. [76]

- The lowest amount of an analyte that is detectable with a given confidence level. For normal distributions, the limit of detection can be calculated as 3 times the standard deviation of blank measurements. The limit of detection can be used as a threshold value to assert the presence of a substance with a known confidence. [80]

- MDL – the minimum concentration of an analyte that can be reported with a 99% confidence that the value is above zero, based on a standard deviation of greater than seven replicate measurements of the analyte in the matrix of concern at a concentration near the low standard. [80]

- The MDL is defined as the minimum measured concentration of a substance that can be reported with 99% confidence that the measured concentration is distinguishable from method blank results. [78]

- The term “detection limit” is used to describe the lowest analyte level that can be confidently identified. [34]

- Lower detectable limit: The minimum pollutant concentration that produces a measurement or measurement output signal of at least twice the noise level. [10]
- Noise: Spontaneous, short duration deviations in measurements or measurement signal output, about the mean output, that are not caused by input concentration changes. Measurement noise is determined as the standard deviation of a series of measurements of a constant concentration about the mean and is expressed in concentration units. [10]

Drift

- A gradual change in instrument response to a constant, quantitative characteristic, i.e., a standard concentration or zero air. [10]
- Span drift – the change in analyzer output over a stated time period, usually 24 hours of unadjusted continuous operation, when the input concentration is at a constant, stated upscale value. Span drift is usually expressed as a percentage change of full scale over a 24-hour operational period. [35]
- Zero drift – the change in analyzer output over a stated time period of unadjusted continuous operation when the input concentration is zero; usually expressed as a percentage change of full scale over a 24-hour operational period. [35]

Instrument Calibration

- Procedures used for correlating instrument response to an amount of analyte (concentration or other quantity). [34]

Intercept

- Of a regression model, the value of the response variable when the predictor variable is zero. [81]

Interferences

- Factors that hinder, obstruct, or impede the ability of a sensor to make accurate measurements. May include pollutants or other chemical compounds, weather conditions, radio frequencies, power fluctuations, vibration, dirt, dust, and insects. [10]
- An interfering substance for an analytical procedure is one that causes a predeterminate systematic error in the analytical result. [39]
    - The observation of the interference depends on the amount of the interferent and of the analyte in the sample. In the case of a quantitative method of analysis, the allowable magnitude of the systematic error should be fixed beforehand in terms of the standard deviation of an individual determination of the analyte. Again, whether or not a substance interferes depends on the amount of the interferent and that of the analyte in the sample. Furthermore, it should be noted that the extent of an interference is not necessarily proportional to the concentration or the content of the interferent in a sample and that the effect of the presence of several interferents is not always additive. Synergistic as well as compensating effects may occur. [39]
- An undesired output caused by a substance or substances other than the one being measured. The effect of the interfering substance(s) on the measurement of interest, shall be expressed as: percentage change of measurement compared with the molar amount of the interferent. If the

interference is nonlinear, an algebraic expression should be developed (or curve plotted) to show this varying effect. [35]

Linearity

- The maximum deviation between an actual analyzer reading and the reading predicted by a straight line drawn between the upper and lower calibration points. This deviation is expressed as a percentage of full scale. [35]
- Linear dynamic range: the range of concentrations over which the calibration curve for an analyte is linear. It extends from the detection limit to the onset of calibration curvature. [80]

Measurement Duration

- The length of time over which a measurement is [performed or] collected (e.g., 1 minute, 1 hour). [10]

Measurement Frequency

- Describes the number of measurements collected per unit of time. [10]
- Sampling rate: the rate at which data collection occurs, usually presented in samples per second. [80]

Measurement Range

- An instrument's dynamic range is the concentration range from minimum to maximum values that the instrument is capable of measuring. [10]
- The concentration region between the minimum and maximum measurable limits. [80]
- Full scale: The maximum measuring limit for a given range of an analyzer. [35]
- Measurement range is distinct from "quantitation range." The term quantitation range describes the span of analyte levels, as contained in a sample matrix, for which the method's performance has been tested, and data quality is deemed acceptable for its intended use. [34]
- Range: The nominal minimum and maximum concentrations that a method is capable of measuring. [10]

Precision

- A measure of agreement among repeated measurements of the same property under identical, or substantially similar, conditions; calculated as either the range or as the standard deviation. May also be expressed as a percentage of the mean of the measurements, such as relative range or relative standard deviation (coefficient of variation). [33]
- The random component of error. Precision is estimated by various statistical techniques typically using some derivation of the standard deviation. [76]
- Precision measures the agreement among repeated measurements of the same property under identical or substantially similar conditions. The more frequently data are collected over a given period, the more confidence one has in the concentration estimate. Precision can be expressed in terms of standard deviation39. Precision can be thought of as the scatter introduced into data by

random (indeterminate) errors when an instrument attempts to measure the same concentration of a pollutant multiple times. [10]

- The degree of agreement of repeated measurements of the same property, expressed in terms of dispersion of test results about the mean result obtained by repetitive testing of a homogenous sample under specified conditions. The terms *repeatability* and *reproducibility* are not standardized, but have generally become to mean single-laboratory-operator-material precision and multi-laboratory, multi-operator, single-material precision, respectively. [80]
- The general term "precision" is used to describe the magnitude of random (indeterminate) errors associated with the use of an analytical method. [34]

Repeatability

- A measure of the precision of the analyzer to repeat its results on independent introductions of the same sample at different time intervals. [80]
- Closeness of agreement between the results of successive measurements of the same measure carried out under the same conditions of measurement. These conditions are called repeatability conditions. Repeatability conditions include the same measurement procedure, observer, measuring instrument (used under the same conditions), location, (and) repetition over a short period of time. Repeatability may be expressed quantitatively in terms of the dispersion characteristics of the results. [34, 41]

Representativeness

- Refers to the degree to which data accurately and precisely represent a characteristic of a population, a parameter variation at a sampling point, a process condition, or a condition. Population uncertainty, the spatial and temporal components of error, can affect representativeness. [76]
- A qualitative term that expresses "the degree to which data accurately and precisely represent a characteristic of a population, parameter variations at a sampling point, a process condition, or an environmental condition." [33]

Reproducibility

- A measure of the precision of different analyzers to repeat results on the same sample. [80]
- Closeness of agreement between the results of measurements of the same measure carried out under changed conditions of measurement. A valid statement of reproducibility requires specification of the conditions changed. The changed conditions may include principle of measurement, method of measurement, observer, measuring instrument, reference standard, location, conditions of use, (and) time. Reproducibility may be expressed quantitatively in terms of the dispersion characteristics of the results. [34, 41]

Response Time

- The amount of time required for a sensor to respond to a change in concentration. [10]
- The time interval from a step change in the input concentration at the analyzer inlet to an output reading of 90% of the ultimate reading. [35]
- Rise time: response time minus lag time. [35]

- Lag time: the time interval from a step change in the input concentration at the analyzer inlet to the first corresponding change in the analyzer signal readout. [35]

Ruggedness

- The extent to which an analytical method remains unaffected by minor variations in operating conditions. [34]
- Insensitivity of a test method to departures from specified test or environmental conditions. [83]

Selectivity

- The ability to correctly identify the analyte(s) of interest in the presence of expected chemical/physical interferences. [34, 77]
    - Selectivity is typically expressed qualitatively. A qualitative selectivity statement includes a description of known interferences, interference effects, and the nature of the analytical data and information that substantiates the identity of the analyte(s) in the matrix of concern. [34]
- The ability of a sensor to respond to a particular pollutant and not to other pollutants. [10]
- Selectivity is the recommended term in analytical chemistry to express the extent to which a particular method can be used to determine analytes under given conditions in the presence of other components of similar behavior. [38]
- Interference equivalent: Positive or negative measurement response caused by a substance other than the one being measured. [10]

Sensitivity

- The capability of a method or instrument to discriminate between measurement responses representing different levels of the variable of interest. [33]
- This term is often confused with and used to describe the detection limit.

Slope

- Of a regression model, the incremental change in the response variable due to a unit change in the predictor variable. [81]

Specificity

- Specificity is considered to be the ultimate of selectivity, mean[ing] that no interferences are supposed to occur. [39]
- To avoid confusion, the use of the term specificity for the [concept of selectivity] is to be discouraged, as it is incorrect. A method is either specific, or it is not. Few, if any, methods, are specific. [38]

Standard Deviation

- Of a population, σ, the square root of the average or expected value of the squared deviation of a variable from its mean; of a sample, s, the square root of the sum of the squared deviations of the observed values in the sample from their mean divided by the sample size minus 1. [82]

Uncertainty

- An indication of the magnitude of error associated with a value that takes into account both systematic errors and random errors associated with the measurement or test process. [37]
    - Uncertainty is a closely related but not identical concept to precision and bias. The primary difference between concepts of precision and bias and of uncertainty is the object that they address. Precision (repeatability and reproducibility) and bias are attributes of the test method. They are estimates of statistical variability of test results for a test method applied to a given material. Repeatability and intermediate precision measure variation within a laboratory. Reproducibility refers to interlaboratory variation. Uncertainty is an attribute of the particular test result for a test material. It is an estimate of the quality of that particular test result. [37]
    - In the case of a quantity with a definition that does not depend on the measurement or test method (for example, concentration, pH, modulus, heat content), uncertainty measures how close it is believed the measured value comes to the quantity. [37]
- Standard uncertainty is reported as the standard deviation of the estimated value of the quantity subject to measurement. [37]
    - The uncertainty is reported as the standard deviation of the reported value. The report $x \pm u$ implies that the value should be between $x - u$ and $x + u$ with an approximate probability of two-thirds, where x is the test result. [37]
    - Expanded uncertainty is reported as a multiple of the standard uncertainty. [37]
    - Relative Standard Uncertainty—The uncertainty is reported as a fraction of the reported value. For a measured value and a standard uncertainty, $x \pm u$, the relative standard uncertainty is u/x. This method of expressing uncertainty may be useful when standard uncertainty is proportional to the value over a wide range. [37]

Appendix B. Air Monitoring Performance Requirements by Data Analysis Type/Decision Sought for $PM_{2.5}$, PM_{10}, CO, NO_2, SO_2, and O_3

In Tables B1 through B6, the information provided under the decision support column relates to regulatory air monitoring requirements in the US, EU, and China and is shown in **bold**, underline, and *italics*, respectively. Numbers shown in brackets are citations to the references given in Section 5. Empty cells indicate that no quantitative performance information was found for that given combination of pollutant, performance attribute, and data analysis type.

Table B1. $PM_{2.5}$ DQOs/MQOs for Various Performance Characteristics/Attributes/DQIs by Data Analysis Type/Decision Sought

Performance Attributes/DQIs	Spatiotemporal Variation	Comparison	Trend	Decision Support	Overview
Accuracy/ Uncertainty	R^2: (0.4225-0.4356, 0.3969-0.4489) [89], 0.62-0.71 [51], 0.91 [65]	R^2: ≥0.73-0.76 [50]		R^2: ≥*0.8649 [2]*, **(0.7225-0.9025) [1]**	Higher r^2 for decision support compared to spatiotemporal
	$\%Diff_{flow}$: ±10% [74]	$\%Diff_{flow}$: ±10% [74]			
	$\%Diff_{zerodrift}$: <20% [74]	$\%Diff_{zerodrift}$: <20% [74]			
	σ: 1-10 μg/m³ [53] %Diff: 9% [63] Relative expanded uncertainty: 50% at 25 μg/m³ with an averaging period of 1 year [84] Short-term drift: <0.5%/24 hours [97] Long-term drift: <5%/month [97] $RMSE/\sigma_{reference}$ ≤1 [64]			RPD_{flow}: ≤2% [3] $\%Diff_{specifiedflow}$: **±5% [7],** ±5% [2] $\%Diff_{onepointflow}$: **±4% [7]** $\%Diff_{multipointflow}$: **±2% [7]** T_{amb} (°C): **±2 [85]**, *±2 [2]*, ±2 [3] P_{amb} (mm Hg): **±10 [7]**, ≤ *7.5 [2]*, ±7.5 [3] RH_{amb}: ±5% [3] Clock/timer (sec): **±60 [7]**, *±20 [2]* *D_{50}: 2.5±0.2 μm[2]* Collection efficiency: *σ_g = 1.2±0.1 [2]* Average flow indication error: *≤2% [2]* Slope: *1±0.15 [2]*, **1±0.10 [1]** Intercept (μg/m³): *0±10 [2]*, **0±2 [1]** Aerosol transmission efficiency: *≥97% [2]* Expanded uncertainty: <25% in 24-h averages [3] Zero level: <2.0 μg/m³ [3] Zero check: 0±3 μg/m³ [3] Maintenance interval: <14 days [3]	

Table B1 (continued). $PM_{2.5}$ DQOs/MQOs for various performance characteristics/attributes/DQIs by data analysis type/decision sought

Performance Attributes/DQIs	Spatiotemporal Variation	Comparison	Trend	Decision Support	Overview
Bias	Bias (%): (<20, <50)[10]	Bias (%): (<30, <30, <50) [10]	Bias (%): <50 [10]		
Completeness	Completeness (%): (≥50, ≥80) [10], 75 [54]	Completeness (%): (≥50, ≥75, ≥80) [10], ≥75% [50]	Completeness (%): ≥50 [10]	Completeness (%): *85 [2]*, ≥90 [3]	higher required completeness for decision support
Detection Limit	Detection limit: 10 µg/m^3 [54], 5 µg/m^3 [97]			Detection limit (µg/m^3): ≤2.0 [3], **2 [7]**	lower detection limit for decision support
				T_{amb} resolution: **0.1 °C [7]** P_{amb} resolution: **5 mm Hg [7]**	
Measurement Duration	Measurement duration: 30 sec [53], 1 hour [54]	Measurement duration = 1 min [51], 1 hour [50]		Measurement duration: **60 min [7]**	shorter measurement duration for comparison and spatiotemporal
Measurement Frequency	Reporting interval: 1 second raw sensor output interval [63] Minimum measurement frequency: 10 s [65], 12 h [89] Averaging time: >4 times the sensor response time [84]			Flow rate measurement intervals: **≤30 sec [7]**	
Measurement Range	Concentration range: <100 µg/m^3 [63], 0.1-200 µg/m^3 [74], 0-250 µg/m^3 [97]	Concentration range: 0.1-200 µg/m^3 [74]		Concentration range: *0-1000 µg/m^3 [2]*, (0-$1000_{24h\text{-}avg}$, 0-$10000_{1h\text{-}avg}$ µg/m^3) [3], **3-200 µg/m^3 [1]**	smaller concentration range (0-200 µg/m^3) for comparison and spatiotemporal compared to larger ranges (0-1000 µg/m^3) for European Union and China Standards under decision support
Precision	CV (%): (<20, <50)[10]	CV (%): (<30, <30, <50)[10]	CV (%): <50 [10]	CV_{conc}: **≤5%[1]**, *≤15% [2]*	lower CV for concentration and flow for decision support

Table B1 (continued). $PM_{2.5}$ DQOs/MQOs for various performance characteristics/attributes/DQIs by data analysis type/decision sought

Performance Attributes/DQIs	Spatiotemporal Variation	Comparison	Trend	Decision Support	Overview
	CV_{flow}: ±10% [74]	CV_{flow}: ±10% [74]		CV_{flow}: **<2% [7],** *≤2% [2]*, (Avg: ≤2%, Inst.: ≤5%) [3]	
	$CV_{zerodrift}$: ±10% [74]	$CV_{zerodrift}$: ±10% [74]			
	R^2: 0.95-0.99 [51], 0.9801 [89] Unbiased variance estimate: 12% [54],			**σ: ≤2 µg/m³ [1]** Precision: <2.5 µg/m³ [3] RMS: **15% [1]**	
Response Time					
Selectivity	Temperature impact on sensor sensitivity: <0.3% from -10 to 50 °C [97]			Temperature influence: zero temperature dependence under 2.0 µg/m³ [3], <5.0% change in min and max temperature conditions [3]	
				Voltage influence: <5% change in min and max voltage conditions [3] Humidity influence: <2.0 µg/m³ in zero air [3]	

Table B2. PM_{10} DQOs/MQOs for Various Performance Characteristics/Attributes/DQIs by Data Analysis Type/Decision Sought

Performance Attributes/DQIs	Spatiotemporal Variation	Comparison	Trend	Decision Support	Overview
Accuracy/ Uncertainty	R^2: 0.53-0.81 [51], 0.91 [65]			R^2: ≥*0.9025 [2]*, **≥0.9409 [1, 10]**	slightly higher r^2 for decision support
	Relative expanded uncertainty: 50% at 50 μg/m³ with an averaging period of 1 hour [84] RMSE/$\sigma_{reference}$: ≤1 [64]	Average match score = 0.91 [52]		D_{50} (μm): *10±0.5 [2]*, **10±0.5 [7]** Collection efficiency: σ_g = *1.5±0.1 [2]* Clock/timer (sec): *±20 [2]*, **±900 [7]** D_{50}: **10±0.5 μm [7]** T_{amb} (℃): *±2[2]* P_{amb} (mm Hg): *≤7.5 [2]* $\%Diff_{flow}$: ±10% [2], **±5% [75]** $\%Diff_{oneptflow}$: **7% [75]** $\%Diff_{multiptflow}$: **10% [75]** RPD_{flow}: ≤ 2%[3] CV_{flow}: *±5% [2]*, <u>(Avg.: <2%, Inst.: <5%)[3]</u> Slope: *1±0.15[2]*, **1±0.10[1]**, <u>1±0.02[3]</u> Intercept (μg/m³): *0±10[2]*, **0±5[1]**, <u>0±1[3]</u> $\%Diff_{conc}$: **±10% [7]**	

Table B2 (continued). PM_{10} DQOs/MQOs for various performance characteristics/attributes/DQIs by data analysis type/decision sought.

Performance Attributes/DQIs	Spatiotemporal Variation	Comparison	Trend	Decision Support	Overview
				Zero level: <2.0 µg/m³ [3] Zero check: 0±3 µg/m³ [3] Expanded uncertainty: <25% [3] Maintenance interval: <14 days [3]	
Bias	Bias (%): (<20, <50)[10]	Bias (%): (<30, <30, < 50)[10]	Bias (%): <50 [10]		
Completeness	Completeness (%): (≥50, ≥80) [10]	Completeness (%): (≥50, ≥75, ≥80) [10]	Completeness (%): ≥50 [10], 75 [90]	Completeness (%): ≥*85 [2]*, >90 [3]	higher required percent completeness for decision support
Detection Limit				Detection limit: <2.0 µg/m³ [3]	
Measurement Duration		Measurement duration: 1 min [51]		Measurement duration: **60 min [7]**	shorter measurement duration for comparison as compared to decision support
Measurement Frequency	Minimum time frequency: 10 seconds [65] Averaging time: >4 times the sensor response time [84]				
Measurement Range				Concentration range (µg/m³): *0 - 1000[2]*, **0 -300 [1]**, (0-1000$_{24h\ avg}$, 0-10,000$_{1hr\ avg}$) [3]	

Table B2 (continued). PM_{10} DQOs/MQOs for various performance characteristics/attributes/DQIs by data analysis type/decision sought.

Performance Attributes/DQIs	Spatiotemporal Variation	Comparison	Trend	Decision Support	Overview
Precision	CV (%): (<20, <50)[10]	CV (%): (<30, <30, <50)[10]	CV (%): <50 [10]	CV: **(7% for concs >80 µg/m^3) [7]**, ≤ *10% [2]*	lower CV for concentration and flow for decision support
	R^2: 0.79-0.91 [51]			σ: **(5 µg/m^3 for concs <80 µg/m^3) [7]** MSE: <2.5 µg/m^3 [3] RMS: **10% [1]**	
Response Time					
Selectivity				Temperature dependence of zero on temperature: under 2.0 µg/m^3 [3] Voltage influence: <5% [3] Humidity influence: <2.0 µg/m^3 in zero air [3]	

Table B3. Carbon Monoxide (CO) DQOs/MQOs for Various Performance Characteristics/Attributes/DQIs by Data Analysis Type/Decision Sought

Performance Attributes/DQIs	Spatiotemporal Variation	Comparison	Trend	Decision Support	Overview
Accuracy/ Uncertainty	Accuracy: ±10% [97]	Accuracy (±20-30%) [87]			inconsistent information
	Relative expanded uncertainty: 25% at 10 mg/m^3 at an averaging period of 8 hours [84] R^2 = 0.6241-0.6724 [89], 0.87 [64], 0.9996 [69] RMSE/$\sigma_{reference}$: ≤1 [64]			12-hr zero drift parts per million (ppm): **±0.5[1]**, ≤0.10 [4] 24-hr zero drift (ppm): **±0.5[1],** *±1 [5]* 12-hr span drift (ppm): 0.60 [4] 24-hr 20% Span drift: ± *1 ppm [5]* 24-hr 80% span drift: **±2.0% [1]**, *±1 ppm [5]* Long-term zero drift (ppm): ≤0.50 [4], *±2 [5]* Long-term span drift: ≤5.0% [4], *±2 ppm [5]* Period of unattended operation: ≥2 weeks and ≤3 months [4], *≥7 days [5]* %Diff$_{flow}$: *±10% [5]* %Diff$_{FullScale}$: *±2% [5]*	
Bias	Bias: (<20%, <50%)[10]	Bias: (<30%, <30%, <50%)[10]	Bias: <50% [10]		
				Maximum linear fit residuals: 4% [4] Linear fit residuals at zero: 0.5 ppm [4]	
Completeness	Completeness (%): (≥50, ≥80)[10]	Completeness (%): (≥50, ≥75, ≥80)[10]	Completeness (%): ≥50 [10]	Completeness (%): >90 [4]	highest required completeness for air monitoring for decision support; however, data capture is limited to two information sources
Detection Limit	Detection Limit (ppb): 4 [68], <4 [69], 1000 [97]			Detection Limit (ppm): **0.4[1]**, *≤0.5 [5]*	
				Noise, σ (ppm): **0.2[1],** *(≤0.25$_{zero}$, ≤1$_{range}$) [5]*	

Table B3 (continued). Carbon Monoxide (CO) DQOs/MQOs for various performance characteristics/attributes/DQIs by data analysis type/decision sought.

Performance Attributes/DQIs	Spatiotemporal Variation	Comparison	Trend	Decision Support	Overview
Measurement Duration				Averaging of short-term fluctuations: <= 7.0% [4]	
Measurement Frequency	Sample Time (min): (1, 1) [88]	Sample Time (min): (1, 1) [88]	Sample Time (min): (1, 1) [88]		
	Temporal Resolution: 5 sec [72]	Temporal Resolution: 5 sec [72]	Temporal Resolution: 5 sec [72]		
	Time: 1 measurement every 10 seconds [68] Measurement Frequency (min): 5[55], 30[94], 60[89] Averaging time: > 4 times the sensor response time [84]				
Measurement Range	Concentration range (ppm): (0.1-8, 0-25, 0-100, 1-1000)[88], 0-1000[97]	Concentration range (ppm): (0-25, 0-100)[88]	Concentration range (ppm): (0.1-8, 0-25, 0-100, 1-1000)[88]	Concentration range (ppm): **0-50 [1]**, *0-50[5]*	higher measurement ranges are required for non-regulatory air monitoring work (all but decision support-related applications)
Precision	Precision (ppb): (50, 200) [88]	Precision: 50 ppb [88]	Precision (ppb): (50, 200)[88]		
	CV: (<20%, <50%) [10]	CV: (<30%, <30%, <50%) [10]	CV: <50% [10]		

Table B3 (continued). Carbon Monoxide (CO) DQOs/MQOs for various performance characteristics/attributes/DQIs by data analysis type/decision sought.

Performance Attributes/DQIs	Spatiotemporal Variation	Comparison	Trend	Decision Support	Overview
	Reproducibility (R^2 = 0.95) [69] Repeatability: (σ: <1.1 ppb) [69]			σ: ≤5.0% of the average of a 3 month period [4] Repeatability standard deviation at zero: 0.3 ppm [4] Repeatability standard deviation: 0.4 ppm [4] Repeatability standard deviation at zero: 0.3 ppm [4] Repeatability standard deviation: 0.4 ppm [4] $\sigma_{20\%URL}$: **1.0%[1]**, *≤0.5 ppm [5]* $\sigma_{80\%URL}$: **1.0%[1]**, *≤0.5 ppm [5]* $\%Diff_{Sampler/CaibrationPort}$: ≤1.0%[4], *±1%[5]* $\sigma_{20\%URL}$: **1.0%[1]**, *≤0.5 ppm [5]* $\sigma_{80\%URL}$: **1.0%[1],** *≤0.5 ppm [5]* $\%Diff_{Sampler/CalibrationPort}$: ≤1.0%[4], *±1%[5]*	
Response Time	Response time: <150 sec [97]			Response time: *≤4 min [5]*	
				Lag time (sec): **120 [1]** Rise time (sec): **120[1]**, ≤180 [4] Fall time (sec): **120[1]**, 180 [4] Rise time - Fall time: ≤ 10 sec [4]	
Selectivity	temperature dependence (ppb/°C): (6.6, 10.3)[72]	temperature dependence (ppb/°C): (6.6, 10.3)[72]	temperature dependence (ppb/°C): (6.6, 10.3)[72]	Temperature Interference: *≤ 0.3 ppm/K [5]*, 0.30 ppm/K [4]	

Table B3 (continued). Carbon Monoxide (CO) DQOs/MQOs for various performance characteristics/attributes/DQIs by data analysis type/decision sought.

Performance Attributes/DQIs	Spatiotemporal Variation	Comparison	Trend	Decision Support	Overview
	Interference: 0.24 ± 0.05 %CO/%NO; 0.20 ± 0.08 %CO/%NO_2 [69]			Interference equivalent: **(concentration change: ±1 ppm).[1]** 19 mmol/mol H_20 : ≤ 1.0 ppm CO [4] *2.5% H_20 Interference: ±5% FS [5]* 500 ppm CO_2: ≤ 0.5 ppm CO [4] *1000 ppm CO_2 interference: ±5% FS[5]* 1 ppm NO: ≤ 0.5 ppm CO [4] 50 ppb N_2O ≤ 0.5 ppm CO [4] *Voltage Stability: ±1% FS [5]* Sensitivity coefficient of electrical voltage: ≤ 0.30 ppm/V [4] Sensitivity coefficient of sample gas pressure: ≤ 0.70 ppm/kPa % [4] Sensitivity coefficient of sample gas temperature: ≤ 0.30 ppm/K [4]	

Table B4. Nitrogen Dioxide (NO_2) DQOs/MQOs for Various Performance Characteristics/Attributes/DQIs by Data Analysis Type/Decision Sought

Performance Attributes/DQIs	Spatiotemporal Variation	Comparison	Trend	Decision Support	Overview
Accuracy/ Uncertainty	%Diff: ±20% [74]	%Diff: ±20% [74]		$\%Diff_{FullScale}$: *±2% [5]* $\%Diff_{flow}$: *±10% [5]*	performance requirements increase from left to right across the table; however, data capture is limited to two information sources
	Relative expanded uncertainty at 106 [21] ppb: 25% [25%] [84] RMSE: 9 ppb [61] Accuracy: 5 ppb [97] R^2: 0.6084-0.9604 [66], 0.89 [64], 0.9823-0.9962 [73], >0.9996 [69] Short-term drift: <5 ppb/24 hr [97] Long-term drift: <10 ppb/month [97] $RMSE/\sigma_{reference}$: ≤ 1 [64]			12-hr zero drift (ppb): **±20[1]**, ≤2.0 [6] 12-hr span drift (ppb): ≤6.0 [6] 24-hr zero drift (ppb): **±20[1]**, *±5 [5]* 24-hr 80% span drift: **±5.0 %[1]**, *±10 ppb [5]* 24-hr 20% span drift: **±20.0%[1]**, *±5 ppb [5]* Long-term zero drift (ppb): *±10 [5]*, ≤5.0 [6] Long-term span drift: *±20 ppb [5]*, ≤5.0% of maximum of certification range [6] Converter efficiency: **>96% [7]**, *>96% [5]*, ≥98% [6] Residuals from linear fit at conc. = [>] 0 (<= 5.0 ppb [4.0]) [6] Period of unattended operation: 3 months [6], ≥*7 days [5]*	
Bias	Bias: (<20%, <50%) [10]	Bias: (<30%, <30%, <50%) [10]	Bias: <50% [10]		
Completeness	Completeness (%): (≥50, ≥80) [10]	Completeness (%): (≥50, ≥75, ≥80) [10]	Completeness (%): 75 [90], ≥50 [10]	Completeness (%): ≥90 [6]	higher required percent completeness for decision support

Table B4 (continued). Nitrogen Dioxide (NO_2) DQOs/MQOs for various performance characteristics/attributes/DQIs by data analysis type/decision sought

Performance Attributes/DQIs	Spatiotemporal Variation	Comparison	Trend	Decision Support	Overview
Detection Limit	Detection limit (ppb): <1 [69], (10, 20) [97]			Detection limit (ppb): **10[1]**, *≤2 [5]*	inconsistent information
	Resolution: 5 ppb [66]			Noise, σ (ppb): **5 [1]**, *(≤1_{zero}, ≤2_{range}) [5]*	
Measurement Duration	Measurement duration: 20 s [66], 1 hour [61]				
Measurement Frequency	Measurement Frequency: 5 sec [72], 1 min [88], 5 min [55], 30 min [94], 1 hr [89]	Measurement frequency: 5 sec [72], 1 min [88]	Measurement frequency: 5 sec [72], 1 min [88]		no pattern observed
	Averaging time: >4 times the sensor response time [84]				
Measurement Range	Measurement range (ppb): (10-250, 10-1000, 0-1000, 50-5000, 50-5000) [88], (10-2000) [66], (20-200) [74], (0-250) [97], (10-2000) [97]	Measurement range (ppb): (0-1000, 50-5000) [88], (20-200) [74]	Measurement range (ppb): (10-250, 10-1000, 0-1000, 50-5000, 50-5000) [88]	Measurement range (ppb): **0-500[1]**, *0-500 [5]*, ≤ 261 [6]	higher measurement ranges are required for non-regulatory air monitoring work (all but decision support-related applications)
Precision	Precision (ppb): (1.2, 3.0, 10, 20)[88]	Precision: 20 ppb [88]	Precision (ppb): (1.2, 3.0, 10, 20) [88]		
	CV: (<20%, <50%) [10], ±20% [74]	CV: (<30%, <30%, <50%) [10], ±20% [74]	CV: <50% [10]		

Table B4 (continued). Nitrogen Dioxide (NO_2) DQOs/MQOs for various performance characteristics/attributes/DQIs by data analysis type/decision sought

Performance Attributes/DQIs	Spatiotemporal Variation	Comparison	Trend	Decision Support	Overview
	Repeatability: (σ<0.32 ppb) [69] Reproducibility: (R^2 = 0.94) [69]			Repeatability standard deviation at zero [concentration] (≤ 1.0 [3.0] ppb) [6] σ: (≤ 5.0% of 3-month avg.) [6] $\sigma_{20\%URL}$: **2%[1]**, *≤5 ppb [5]* $\sigma_{80\%URL}$: **6%[1]**, *≤10 ppb [5]* $\%Diff_{SamplerCalibrationPort}$: ≤1.0% [6], *±1% [5]*	
Response Time	Response time: 0.2 sec [97] t_{90} = 21 s [69]			Response time: *≤5 min [5]* Lag time: **20 min [1]** Rise time: **15 min[1]**, ≤180 s [6] Fall time: **15 min[1]**, ≤180 s [6] Difference in rise and fall time (≤ 10 s) [6] Residence time: **<2 min [7]**, ≤ 3.0 s) [6]	appears to be a pattern in that faster response times are needed for spatiotemporal air monitoring compared to decision support applications
Selectivity	Temperature impact on sensor sensitivity: <0.5% from -20 to 40 °C [97]			Temperature Interference: ≤ *3 ppb/°C [5]*	
	Interference: -0.02 ± 0.03 $\%NO_2/\%CO$; 1.2 ± 0.11 $\%NO_2/\%NO$ [69]			Individual [total] interference equivalent: **concentration change: ±0.02 [0.04] ppm [1]** Voltage Stability: *±1% FS [5]* *2.5% H_20 Interference: ± 4% FS [5]* Interferent from 19 mmol/mol of H_2O: ≤5 ppb [6] *1 ppm NH_3 interference: ± 4% FS [5]* Interferent from 200 ppb NH_3: ≤5.0 ppb [6] *200 ppb O_3 interference: ± 4% FS [5]* *500 ppb SO_2 interference: ± 4% FS [5]* Interferent from 500 ppm of CO_2: ≤5.0 ppb [6] Sensitivity coefficient of sample gas pressure: ≤8.0 ppb/kPa [6]	

Table B4 (continued). Nitrogen Dioxide (NO_2) DQOs/MQOs for various performance characteristics/attributes/DQIs by data analysis type/decision sought

Performance Attributes/DQIs	Spatiotemporal Variation	Comparison	Trend	Decision Support	Overview
				Sensitivity coefficient of sample gas temperature: ≤3.0 ppb/K [6] Sensitivity coefficient of surrounding temperature: ≤3.0 ppb/K [6] Sensitivity coefficient of electrical voltage (≤ 0.3 ppb/V) [6]	

Table B5. Sulfur Dioxide (SO_2) DQOs/MQOs for Various Performance Characteristics/Attributes/DQIs by Data Analysis Type/Decision Sought

Performance Attributes/DQIs	Spatiotemporal Variation	Comparison	Trend	Decision Support	Overview
Accuracy/ Uncertainty	Short term drift: <2 ppb/24 hours [97] Long term drift: 10 ppb/month [97] Accuracy: ±0.5 ppm [97]			12-hr zero drift (ppb): **±4[1]**, ≤2.0 [8] 12-hr span drift (ppb): ≤6.0 [8] 24-hr zero drift (ppb): **±4[1]**, *±5 [5]* 24-hr 80% span drift: **±3.0%[1]**, *±10 ppb [5]* *24-hr 20% span drift: ± 5 ppb [5]* $\%Diff_{flow}$: *±10% [5]* $\%Diff_{FullScale}$: *± 2% [5]* Long-term zero drift (ppb): *±10 [5]*, ≤4.0 [8] Long-term span drift: *±20 ppb [5]*, ≤5.0% of maximum of certification range [8] Residuals from linear fit at conc. = [>] 0 (≤ 5.0 ppb [4.0]%) [8]	inconsistent information
Bias	Bias: (<20%, <50%) [10]	Bias: (<30%, <30%, <50%) [10]	Bias: <50% [10]		
Completeness	Completeness (%): (≥50, ≥80) [10]	Completeness (%): (≥50, ≥75, ≥80) [10]	Completeness (%): ≥50 [10]	Completeness (%): >90 [8]	highest requirements for air monitoring for decision support; however, data capture is limited to two information sources
Detection Limit	Detection limit (ppb): (50, 200) [97]			Detection limit (ppb): **2[1]**, *≤2 [5]*	

Table B5 (continued). Sulfur Dioxide (SO_2) DQOs/MQOs for various performance characteristics/attributes/DQIs by data analysis type/decision sought

Performance Attributes/DQIs	Spatiotemporal Variation	Comparison	Trend	Decision Support	Overview
				Noise, σ (ppb): **1[1]**, *(≤1_{zero}, ≤5_{range}) [5]*	
Measurement Duration					
Measurement Frequency					
Measurement Range	Measurement range (ppb): 0-1000 [97], 0-10000 [97]			Measurement range (ppb): **0-500[1]**, *0-500 [5]*, ≤376 ppb [8]	
Precision	CV: (<20%, <50%) [10]	CV: (<30%, <30%, <50%) [10]	CV: <50% [10]		
				$\sigma_{20\%URL}$: **2%[1]**, *≤5 ppb [5]* $\sigma_{80\%URL}$: **2%[1]**, *≤10 ppb [5]* Repeatability standard deviation at zero [concentration] (≤ 1.0 [3.0] ppb) [8] σ: (≤ 5.0% of 3-month avg.) [8] $\%Diff_{SamplerCalibrationPort}$: ≤1.0%[8], *±1% [5]*	
Response Time	Response time: <60 sec [97]			Response time: *≤5 min [5]*	
				Lag time (sec): **120 [1]** Rise time (sec): **120[1],** ≤180 [8] Fall time (sec): **120[1]**, ≤180 [8] Difference in rise and fall time: ≤10 sec [8]	

Table B5 (continued). Sulfur Dioxide (SO_2) DQOs/MQOs for various performance characteristics/attributes/DQIs by data analysis type/decision sought

Performance Attributes/DQIs	Spatiotemporal Variation	Comparison	Trend	Decision Support	Overview
Selectivity	Temperature impact on sensor sensitivity: <0.2% from -20 to 40 °C[97]			Temperature Interference: *≤1 ppb/°C [5]* Individual interference equivalent: **±0.005 ppm [1]** Voltage Stability: *±1% FS [5]* 2% H_20 Interference: *±4% FS [5]* Interference from 19 mmol/mol of H_2O: ≤10 ppb [8] *Interference from 0.1 ppm Toluene: ±4% FS [5]* *Interference from 3000 ppm CH_4: ±4% FS [5]*	

Table B5 (continued). Sulfur Dioxide (SO_2) DQOs/MQOs for various performance characteristics/attributes/DQIs by data analysis type/decision sought

Performance Attributes/DQIs	Spatiotemporal Variation	Comparison	Trend	Decision Support	Overview
				Interference from 200 ppb H_2S: ≤5.0 ppb [8] Interference from 200 ppb NH_3: ≤5.0 ppb [8] Interference from 500 ppb NO: ≤5.0 ppb [8] Interference from 200 ppb NO_2: ≤5.0 ppb [8] Interference from 1 ppm m-xylene: ≤10.0 ppb [8] Sensitivity coefficient of sample gas pressure: ≤2.0 ppb/kPa [8] Sensitivity coefficient of sample gas temperature: ≤1.0 ppb/°C [8] Sensitivity coefficient of surrounding temperature: ≤1.0 ppb/°C [8] Sensitivity coefficient of electrical voltage: ≤0.3 ppb/V [8]	

Table B6. Ozone (O_3) DQOs/MQOs for Various Performance Characteristics/Attributes/DQIs by Data Analysis Type/Decision Sought

Performance Attributes/DQIs	Spatiotemporal Variation	Comparison	Trend	Decision Support	Overview
Accuracy/ Uncertainty	Standard error (ppb): 3 [59], 5 [58] Estimation Error, 2σ: ±4 ppb [60] Long-term drift: <4 ppb [58], <10 ppb/month [97] Short-term drift: <5 ppb/24 hr [97] Stability over time: yearly average offset < factor of 2 [91] Mean difference: 2.0 ± 1.6 ppb [58] Relative expanded uncertainty: 30% at 120 µg/m^3 over an 8 hour averaging period [84] R^2: 0.95-0.97 [71], 0.8464-0.9801[66], (0.82-0.94, 0.8281-0.9409)[89], 0.84 [70], >0.9 [67], 0.77 [64] Accuracy: 6.5 ppb [97] $RMSE/\sigma_{reference}$: ≤ 1 [64]			12-hr zero drift (ppb): **±4 [1]** 24-hr zero drift (ppb): **±4[1]**, *±5 [5]* 24-hr 80% span drift: **±3.0%[1]**, *±10 ppb [5]* 24-hr 20% span drift: ± *5 ppb [5]* Long-term zero drift (ppb): *±10 [5]*, ≤5.0 [9] Long-term span drift: *±20 ppb [5]*, ≤5.0% of max certification range [9] $\%Diff_{flow}$: *±10% [2]* $\%Diff_{FullScale}$: *±4% [5]* Residuals of linear fit at conc. = [>] 0 (<= 5.0 ppb [4.0]) [9] Period of unattended operation: 3 months [9], *≥7 days [5]*	inconsistent information
Bias	Bias (%): (<20, <50) [10]	Bias (%): (<30, <30, <50) [10]	Bias (%): <50 [10]		
	Standard error (ppb): (3±2, 6) [57], (<5, 5) [58] Mean bias (ppb): -1 [57], 0 [58]				
Completeness	Completeness (%): (≥50, ≥80) [10]	Completeness (%): (≥50, ≥75, ≥80) [10]	Completeness (%): ≥50 [10], ≥75 [90]	Completeness (%): >90 [9]	highest requirements for air monitoring for decision support

Table B6 (continued). Ozone (O_3) DQOs/MQOs for Various Performance Characteristics/Attributes/DQIs by Data Analysis Type/Decision Sought

Performance Attributes/DQIs	Spatiotemporal Variation	Comparison	Trend	Decision Support	Overview
	Sample frequency: >75% of available hourly data collected [92] Time: 8 years in a 10 year period [92]				
Detection Limit	Detection limit (ppb): 5 [70], (1, 20) [97]			Detection limit (ppb): **5[1]**, *≤2 [2]*	
	Resolution: 1 ppb [66]			Noise, σ (ppb): **2.5[1]**, *(≤1_{zero}, ≤5_{range}) [2]*	
Measurement Duration	Measurement duration: 1 min [60], 1 min [71], 1 min [66]		1-hr daily maximum values averaged quarterly [86]		spatiotemporal variations require shorter measurement durations as compared to longer-term trends monitoring, in accord with expectations
Measurement Frequency	Sample time: 10 s [88], 1 min [59], (1 min, 1 min) [88], 1 min [57], 1 min [58], hourly [89], 5 minutes [70], 5 min [55], 30 min [94]	Sample Time: (10 s, 1 min, 1 min) [88]	Sample Time: (10 s, 1 min, 1 min) [88]		similar across comparison, spatiotemporal, and trends monitoring applications
	Averaging time: >4 times the sensor response time [84]				
Measurement Range	Measurement range (ppb): (2-10000, 10-250, 0-500, 0-150, 10-1000) [88], 0-100 ppb [60], 0-150 [66], (0-250, 0-500) [97]	Measurement range (ppb): (2-10000, 0-500, 0-150) [88]	Measurement range (ppb): (2-10000, 10-250, 0-500, 0-150, 10-1000) [88]	Measurement range (ppb**): 0-500[1]**, *0-500 [5]*, <u>≤250 [9]</u>	higher measurement ranges are required for non-regulatory air monitoring work (all but decision support-related applications)
Precision	Precision (ppb): (0.5, 0.6, 2.0, 5.0, 6.0, 10, 10.3) [88]	Precision (ppb): (2.0, 5.0, 6.0) [88]	Precision (ppb): (2.0, 5.0, 6.0, 10, 10.3) [88]		no pattern present
	CV: (<20%, <50%) [10]	CV: (<30%, <30%, <50%) [10]	CV: <50% [10]		

Table B6 (continued). Ozone (O_3) DQOs/MQOs for Various Performance Characteristics/Attributes/DQIs by Data Analysis Type/Decision Sought

Performance Attributes/DQIs	Spatiotemporal Variation	Comparison	Trend	Decision Support	Overview
	Precision: 4% at 95% confidence level [59] Mean absolute deviation: 1.3 [0.6-3.1] ppb [66] R^2 = 0.9±0.06 [67], 0.9995 [70]			$\sigma_{20\%URL}$: **2%[1]**, *≤5 ppb [5]* $\sigma_{80\%URL}$: **2%[1]**, *≤10 ppb [5]* Repeatability standard deviation at zero [concentration] (≤ 1.0 [3.0] ppb) [9] σ: (≤ 5.0% of 3-month avg) [9] $\%Diff_{SampleCalibrationPort}$: ≤ 1.0% [9], *±1% [5]*	
Response Time	Response time: 65 sec [97]			Response time: *≤5 min [5]*	faster response times are needed for non-regulatory purposes such as spatiotemporal trends monitoring; note that data are limited (one spatiotemporal study, three regulatory monitoring methods)
				Lag time (sec): **120 [1]** Rise time (sec): **120[1]**, ≤180 [9] Fall time (sec): **120[1]**, ≤180 [9] Difference in rise and fall time: ≤10 sec [9] Residence time inside analyzer: ≤3.0 sec [9]	
Selectivity	Temperature impact on sensor sensitivity: <0.5% from -20 to 40 °C [97]			T_{amb} Interference: *≤1 ppb/°C [5]*	

Table B6 (continued). Ozone (O_3) DQOs/MQOs for Various Performance Characteristics/Attributes/DQIs by Data Analysis Type/Decision Sought

Performance Attributes/DQIs	Spatiotemporal Variation	Comparison	Trend	Decision Support	Overview
				Individual interference equivalent: **±0.005 ppm [1]** Voltage Stability: ±*1% FS [5]* *2% $H_2$0 Interference: ±4% FS [5]* 19 mmol/mol H_2O interference: ≤10 ppb [9] *1 ppm Toluene interference: ±4% FS [5]* 0.5 ppm Toluene interference: ≤ 5.0 ppb [9] *0.2 ppm SO_2 interference: ±4% FS [5]* *0.5 ppm NO/NO_2 interference: ±6% FS [5]* 0.5 ppm m-xylene interference: ≤ 5.0 ppb [9] Sensitivity coefficient of sample gas pressure: ≤2.0 ppb/kPa [9] Sensitivity coefficient of sample gas temperature: ≤1.0 ppb/K [9] Sensitivity coefficient of surrounding temperature: ≤1.0 ppb/K [9] Sensitivity coefficient of electrical voltage: ≤0.3 ppb/V [9]	

EPA
United States
Environmental Protection
Agency

Office of Research and
Development (8101R)
Washington, DC 20460