EPA Public Access
Author manuscript
Atmos Environ X. Author manuscript; available in PMC 2021 July 27.
About author manuscripts | Submit a manuscript

Published in final edited form as:
Atmos Environ X. 2019 April ; 2: . doi:10.1016/j.aeaoa.2019.100031.

Deliberating performance targets workshop: Potential paths for emerging $PM_{2.5}$ and O_3 air sensor progress

R. Williams[a], **R. Duvall**[a,*], **V. Kilaru**[a], **G. Hagler**[a], **L. Hassinger**[b], **K. Benedict**[c], **J. Rice**[c], **A. Kaufman**[c], **R. Judge**[d], **G. Pierce**[e], **G. Allen**[f], **M. Bergin**[g], **R.C. Cohen**[h], **P. Fransioli**[i], **M. Gerboles**[j], **R. Habre**[k], **M. Hannigan**[l], **D. Jack**[m], **P. Louie**[n], **N.A. Martin**[o], **M. Penza**[p,q], **A. Polidori**[r], **R. Subramanian**[s], **K. Ray**[t], **J. Schauer**[u], **E. Seto**[v], **G. Thurston**[w], **J. Turner**[x], **A.S. Wexler**[y], **Z. Ning**[z]

[a]U.S. Environmental Protection Agency, Office of Research and Development, Research Triangle Park, NC, USA

[b]Former Oak Ridge Institute for Science and Education (ORISE) staff assigned to the U.S. Environmental Protection Agency, Office of Research and Development, Research Triangle Park, NC, USA

[c]U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC, USA

[d]U.S. Environmental Protection Agency, Region 1, North Chelmsford, MA, USA

[e]Colorado Department of Public Health and the Environment, Denver, CO, USA

[f]Northeast States for Coordinated Air Use Management, Boston, MA, USA

[g]Pratt School of Engineering, Duke University, Durham, NC, USA

[h]College of Chemistry, University of California-Berkeley, Berkeley, CA, USA

[i]Clark County Department of Air Quality (Nevada), Las Vegas, NV, USA

[j]European Commission, Joint Research Centre, Ispra, Italy

[k]Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

[l]Mechanical Engineering Department, University of Colorado-Boulder, Boulder, CO, USA

[m]Mailman School of Public Health, Columbia University, New York, NY, USA

This is an open access article under the CC by-nc-nd license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

[*]Corresponding author. U.S. Environmental Protection Agency, 109 T.W. Alexander Drive, MD E343-02, Research Triangle Park, NC 27711, USA. duvall.rachelle@epa.gov (R. Duvall).

Disclaimer

The United States Environmental Protection Agency through its Office of Research and Development and Office of Air and Radiation conducted the research described here. It has been subjected to Agency review and approved for publication. Approval does not signify that the contents necessarily reflect the views and the policies of the Agency nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

Declaration of interest

There are no interests to declare.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aeaoa.2019.100031.

EPA Author Manuscript
EPA Author Manuscript
EPA Author Manuscript

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

[n]Hong Kong Environmental Protection Department, Hong Kong, China

[o]National Physical Laboratory, Teddington, Middlesex, United Kingdom

[p]Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA), Brindisi Research Center, Brindisi, Italy

[q]European Network on New Sensing Technologies for Air-Pollution Control and Environmental Sustainability (EuNetAir), Brindisi, Italy

[r]South Coast Air Quality Management District, Diamond Bar, CA, USA

[s]Center for Atmospheric Particle Studies, Carnegie Mellon University, Pittsburgh, PA, USA

[t]Confederated Tribes of the Colville Reservation, Nespelem, WAashington, USA

[u]College of Engineering, University of Wisconsin-Madison, Madison, WI, USA

[v]School of Public Health, University of Washington, Seattle, WA, USA

[w]School of Medicine, New York University, New York, NY, USA

[x]School of Engineering and Applied Sciences, Washington University, St. Louis, MO, USA

[y]Air Quality Research Center, University of California-Davis, Davis, CA, USA

[z]Hong Kong University of Science and Technology, Hong Kong, China

Abstract

The United States Environmental Protection Agency held an international two-day workshop in June 2018 to deliberate possible performance targets for non-regulatory fine particulate matter ($PM_{2.5}$) and ozone (O_3) air sensors. The need for a workshop arose from the lack of any market-wide manufacturer requirement for Ozone documented sensor performance evaluations, the lack of any independent third party or government-based sensor performance certification program, and uncertainty among all users as to the general usability of air sensor data. A multi-sector subject matter expert panel was assembled to facilitate an open discussion on these issues with multiple stakeholders. This summary provides an overview of the workshop purpose, key findings from the deliberations, and considerations for future actions specific to sensors. Important findings concerning $PM_{2.5}$ and O_3 sensors included the lack of consistent performance indicators and statistical metrics as well as highly variable data quality requirements depending on the intended use. While the workshop did not attempt to yield consensus on any topic, a key message was that a number of possible future actions would be beneficial to all stakeholders regarding sensor technologies. These included documentation of best practices, sharing quality assurance results along with sensor data, and the development of a common performance target lexicon, performance targets, and test protocols.

Keywords

Low-cost air quality sensors; Performance targets; $PM_{2.5}$

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

1. Introduction

Low-cost air quality sensors are being used by a diverse group of users to meet a wide variety of measurement needs. “Low-cost” generally represents devices costing under $2,500 USD although no formally recognized cost definition exists for such nomenclature. Measurement needs range from environmental awareness to purposeful data measurements for decision making (Clements et al., 2017). The data quality of these sensors is often poorly characterized and there appears to be confusion in how to effectively use the resulting data (Lewis et al., 2018). At the core of the confusion is the need to ensure sensors provide data quality meeting basic performance criteria (Lewis and Edwards, 2016). Currently, there is no standardized means of addressing performance for targeted end use (Woodall et al., 2017). As such, sensors are not yet appropriate for regulatory use in which specific performance and operation criteria are required.

Some of the many applications being reported for sensor technologies include:

- Real-time high-resolution mapping of air quality (spatio-temporal settings),
- Fenceline monitoring to detect industrial source air emissions,
- Community monitoring to assess hot spots,
- Personal, indoor or other microenvironmental monitoring to assess human exposures,
- Data collection in remote places or in locations that are not routinely monitored,
- Extreme events monitoring (e.g., wildfires), and
- Environmental awareness activities and citizen science.

As seen in the list above, data quality requirements may vary depending upon the intended end use. Given the rapid adoption and technological advances of new air sensor technologies, there are numerous questions about how well they perform and how lower-cost technologies can be used for certain non-regulatory applications. To facilitate discussion on these topics, the United States Environmental Protection Agency (US EPA) held the “Deliberating Performance Targets for Air Quality Sensors Workshop” from June 25–26, 2018 in Research Triangle Park, North Carolina (https://www.epa.gov/air-research/deliberating-performance-targets-air-quality-sensors-workshop). The purpose of the workshop was to solicit individual stakeholder views related to non-regulatory performance targets for sensors that measure fine particulate matter ($PM_{2.5}$) and ozone (O_3). $PM_{2.5}$ and O_3 were selected as the primary focus for all discussions because it was believed more technical information existed concerning low-cost sensors for these pollutants. Through on-site and webinar discussions, national and international participants totaling more than 700 registrants addressed a range of technical issues involved in establishing performance targets for air sensor technologies. Discussion topics included:

- The state of technology with respect to sensor performance for various measures (e.g., limits of detection, linearity, interference, calibration, etc.),

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

- Review of relevant government-based and peer-reviewed literature reports concerning performance requirements or technical findings,
- Important attributes of sensor technology to include in a performance target and why,
- Consideration of a single set of performance targets for all non-regulatory purposes or a tiered approach for different sensor applications,
- Manufacturers' perspectives on the need for and value of any performance certification process,
- Expert opinion from tribal/state/local air agencies on specific end-use categories including epidemiological, human exposure, and community-based monitoring, and
- Lessons learned from other countries about choices or trade-offs they have made or debated in establishing performance targets for sensor technologies.

2. Approach

The public workshop consisted of two days of structured presentations and panel discussions from subject matter experts (SEs) and general attendees using the US EPA-defined charge questions. A third day meeting (non-public) was convened among the SEs and the US EPA workshop organizers to review information from each session and begin the process of developing the workshop summary. The SEs identified in the sections below were selected to represent a variety of viewpoints (e.g., international, regulatory agencies, academia, etc.) to obtain a diverse perspective on sensor performance considerations. The names, affiliations, and expertise of the SEs are listed in Table S1. There was widespread interest in the workshop topics as suggested by the registered attendance statistics which included international governments/interests (137) from a total of 44 countries, US federal agencies (16), tribes (5), states (167), academics (119), private companies and nonprofits (47), sensor manufacturers and industry (121), and the general public and non-categorized attendees (29). Of over 700 registrants, the US EPA attendees represented 249.

The workshop consisted of 12 unique sessions with some commonality between various topics. In brief they are summarized as follows:

- $PM_{2.5}$ and O_3 peer-reviewed literature and/or government-based documentation on sensor use and certification efforts,
- International perspectives on ongoing or established sensor certification programs,
- Tribal/state/local air agency perspectives on sensor data quality benchmarks and application considerations,
- Needed performance targets for nonregulatory monitoring and other associated measurements applications,
- Case studies describing specific uses of $PM_{2.5}$ and O_3 sensor data,

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

- Perspective from $PM_{2.5}$ and O_3 sensor manufacturers on the value in and means by which sensor performance might be standardized,
- The value of either a binary (yes/no) or tiered (multiple performance conditions) approach to standardizing sensor technologies, and
- Miscellaneous: An open call from all attendees on topics of relevance pertaining to low-cost sensor performance targets.

Each of the individual sessions provided both in-person as well as web-based attendees the opportunity for open discussion on the workshop topics. During the workshop, except for the literature review discussions, the US EPA staff were not engaged in directly leading the discussions or presenting information pertaining to the key topics. The purpose of this was to recover a non-biased perspective on low-cost sensor performance targets and potential next steps to facilitate stakeholder needs. Presentations delivered by the subject experts are available on the US EPA Air Sensor Toolbox website (https://www.epa.gov/air-sensor-toolbox) and the workshop website.

3. Results

3.1. $PM_{2.5}$ and O_3 peer-reviewed literature and select government-based certification documentation

The US EPA sponsored a comprehensive review of select literature pertaining to both $PM_{2.5}$ and O_3 sensor-based measurements and government-based air quality instrumentation certification requirements. This work has been summarized in detail in an US EPA Technical Report (Williams et al., 2018). Excerpts from the report were presented during the workshop by Ronald Williams (US EPA) and are summarized here. As part of the literature review, automated database searches were conducted using Compendex, Scopus, Web of Science, Digital libraries of Theses and Dissertations, Open Grey, OpenAIRE, Worldcat, US Government Publications, Defense Technical Information Center, and UN Digital Library. Computer-based searches using relevant sensor-based words/phrases revealed more than 20,000 citations from Jan 1, 2007 to Dec 31, 2017. This time period was selected to ensure that only the timeliest literature would be examined relative to sensor performance targets. A systematic hand-curated review was conducted to reduce the quantity of citations to a meaningful number. Information from 56 titles was selected for this summary and included only the literature that provided the potential for informing the US EPA on 1) existing US and international performance standards, 2) identification of any non-regulatory technology performance standards, 3) discovery of the types of data identified in the reports, 4) data quality indicators (DQIs) and data quality statistics associated with low-cost sensor use, and 5) discovery of the air monitoring applications where specific sensor data quality indictors had been reported.

Four broad categories of performance requirements were defined from this selection including spatio-temporal variability, comparisons, trends analysis, and decision-making. Sixteen application categories were found including air quality forecasting, air quality index reporting, community near-source monitoring, control strategy effectiveness, data fusion, emergency response, epidemiological studies, personal exposure reduction, hot-spot

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

detection, model input, model verification, process study research, public education, public outreach, source identification, and supplemental monitoring. While many DQIs exist in the literature, a total of 10 were selected as part of the review process. Throughout this document both the correlation coefficient (r; measure of how well data points correlate) and the coefficient of determination (R^2; measure of variance for linear regression models) are referenced and discussed when referring to correlation and linearity as DQIs.

The technical report included a review of select government-based certification standards for ambient air monitoring instruments. These criteria provide information on attributes that characterize measurements. These included the US Federal Reference Method and Federal Equivalent Method Program (FRM/FEM) descriptions (US EPA, 40 CFR Part 50 and Part 53), the People's Republic of China National Environmental Monitoring Standards (Chinese Ministry of Environmental Protection, HJ 653–2013), the European Ambient Air Quality Directives [2008/50/EC; European Committee for Standardization (CEN) Technical Committee 264 for Air Quality, DS/EN 16450:2017], and the People's Republic of China Performance Standard for Air Sensors (Chinese Ministry of Environmental Protection, HJ 654–2013). A similar listing as above exists for government-based O_3 instrumentation certification. One specific example of a certification program that encompasses sensors was the aforementioned CEN Technical Committee 264. While the timing for enacting the certification program has not yet been determined, the CEN efforts represent some of the most mature certification concepts known including a three-tiered approach ranging from regulatory quality (meeting Air Quality Directives of 2008/50 EC), to those meeting simple ad hoc end user performance criteria.

The literature review showed differences in DQI requirements for regulatory monitoring across the US, EU, and China. For instance, with $PM_{2.5}$ measurement ranges, the US requires a range of 3–200 $\mu g/m^3$ (US EPA, 40 CFR Part 50), the EU requires a range of 0–1000 $\mu g/m^3$ per 24-h period (CEN, DS/EN 16450:2017), and China requires a range of 0–1000 $\mu g/m^3$ (Williams et al., 2018). Similar differences were found for regulatory requirements for O_3 monitoring. Examples for O_3 include lag and rise times ranging from 120 s (US) to up to 5 min (China). This information may suggest that even if low-cost sensors could meet regulatory requirements, DQIs would likely vary across the world.

In terms of purposes for using low-cost sensors, the literature review found that for both pollutants, sensor data were more often collected and reported for spatial-temporal understanding purposes (63% and 72%, respectively for $PM_{2.5}$ and O_3). Specifically, the need to use a low-cost sensor was primarily linked to the desire to collect a high spatial density of measurements. Only rarely (< 26%) were low-cost sensors used for decision support and with stated DQIs. It was interesting to note that the cost of the technology was reported as a deciding factor in its selection as much as 70% of the time under such circumstances. Even so, reported cost of purchase did not necessarily equate to good or poor sensor performance.

Another major outcome from the literature review was the disparate reporting of DQIs for both pollutants (summarized in Table 1). As an example, it was not uncommon to have as many as 9 different statistical attributes being reported for accuracy/uncertainty. Even when

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

a common term might be reported (e.g., drift), it was often reported either as a percentage (%) or as a given concentration value (e.g., ± X ppb) making direct comparison of the results difficult. This was also true for DQIs such as bias, completeness, measurement range, precision, and others. In short, conference attendees reported that because there are currently no data standards or even a common lexicon of statistical terms associated with reporting sensor data, it is difficult to clearly compare findings from one study to another. It was also noted that any effort to systematically define statistical terms and common data reporting criteria would benefit all sensor data users, regardless of the sponsoring organization. Examples of where data reporting confusion has been resolved as technologies matured, include organizations producing voluntary consensus standards such as ASTM International (ASTM D1356-17, 2017) where specific data reporting requirements exist. Similar standardization has been seen where professional organizations publish suggested lexicons on terminology for their members (Zartarian et al., 2006).

3.2. International perspectives on sensor use and certification

These sessions included discussions on $PM_{2.5}$ and O_3 sensor certification efforts, consideration of sensor performance issues, and how data from low-cost sensors were being used for different purposes internationally. The SEs included Nick Martin (UK National Physical Laboratory), Michel Gerboles (EU Joint Research Center), Peter Louie (Hong Kong Environmental Protection Department), Michele Penza (ENEA, EuNetAir), and Zhi Ning (The Hong Kong University of Science and Technology). Gerboles and Ning were unable to attend in person but provided material that was presented by others and is summarized here. The international SEs were asked to describe their organization and its purpose and efforts to establish or seek to establish sensor certification requirements. They were asked to define the performance standards and measurements practices to conduct certification and the rationale behind those choices. They provided examples of certification standards and their perspective on the future of low-cost sensors needing or benefiting from a certification requirement. Lastly, they provided their perspective on tiering or a binary (yes/no) consideration of sensor certification.

Martin, Gerboles, and Penza reported on ongoing work performed by their organizations in concert with the CEN Technical Committee 264 efforts (Gerboles, 2018; Martin, 2018; Penza, 2018). In brief, this multi-national organization led by air quality experts has the responsibility for defining low-cost air quality sensor certification procedures for the European governing bodies to consider. Working Group 42 has been devoted to examining current gas-phase and PM sensor performance and defining potential protocols and procedures needed to ensure harmonization of sensor performance benchmarks and relevant evaluation requirements. As seen in Table 2, specific consideration to key performance benchmarks are being considered. The listed benchmarks are proposed and have not been ratified. Developing the standard language has been ongoing for approximately 3 years. If, and when the CEN performance requirements are ratified, sensor manufacturers interested in obtaining certification would be required to submit their device (in replicate) to any European-nominated institution in charge of sensor evaluation according to the CEN Technical Committee Working Group 42’s protocol. Sensor manufacturers would be responsible for the cost of the certification activities that might include laboratory, field, or a

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

mixture of both test scenarios. Estimated costs of a single certification testing involving about 4 sensors (replicates) were valued in the range of $10,000-$100,000 (USD) and would remain valid in any European country.

Martin has been working extensively on $PM_{2.5}$ standards for the UK National Physical Laboratory where regulatory PM instrumentation are compared against UK standardized Monitoring Certification Scheme (MCERTS) protocols. Specific criteria have already been defined relative to any instrumentation desiring certification (Table 3). Martin noted that the UK considers language where the type of instrumentation in a certification process (e.g., sensor, monitor, analyzer, etc.) need not be identified by a given technology type, but rather as being “agnostic”. In other words, regardless of the cost, measurement technology, or other specifics of the instrument, it has no bearing on whether the instrumentation is eligible to be considered for certification. The only dependent variable is the ability of the device to meet the performance benchmark requirements. Such agnostic programs are believed to sometimes foster manufacturing innovation while having the potential to simplify the standard performance language needed for government-based approval of the certification program. Martin indicated that a simplified process in comparison to the MCERTS requirements was needed for low-cost sensor-based technologies. He provided examples of DQIs that might be meaningful for $PM_{2.5}$ and O_3 and how the distribution of a sensor-based network approach might have to be considered as part of any certification process.

Penza reiterated many of the same points provided by both Gerboles and Martin. He provided a number of examples where low-cost $PM_{2.5}$ sensor technologies were deployed in multiple European cities with varying degrees of comparability versus regulatory monitors. He noted open questions of air sensors in particular lower accuracy compared to reference monitors, cross-sensitivity, low stability, more periodic calibration needs, and regular maintenance in the field, among others. Penza also noted advantages and benefits of air quality sensors for uses including deployment in cities at high spatial-temporal resolution, personal exposure studies, gaining information on emission sources, outdoor and indoor monitoring, and combining sensors with modeling for micro-scale analysis.

Louie presented on the development of marine-based pollutant standards in Hong Kong. While no $PM_{2.5}$ or O_3 data were shared relative to low-cost sensor use, Louie demonstrated successful field and laboratory calibration of numerous gas-phase sensors. His tests included examination of precision and linearity, range of measurement, long term drift examination, temperature and relative humidity response. Following calibration, the plume sniffer technology his team developed (portable and aerial platforms) has been used to enforce the Domestic Emission Control Areas ambient air quality with respect to marine-based combustion product emissions.

Ning’s presentation (delivered by Louie) provided field-based and laboratory examples of where low-cost air quality sensors had performance issues (response) related to changing temperatures, relative humidity, and sensor drift over time (Fig. 1). Not only did the low-cost O_3 sensor shown yield only a modest agreement with a collocated reference monitor over a 0–60 ppb challenge range, it had multiple groupings of residuals around the regression line where vastly different sensor response was observed depending on temperature and humidity

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

conditions. In lieu of any certification requirements or even if such a requirement was ever developed, Ning proposed a process for examining sensor performance over time which included:

- Collocated testing with a reference monitoring for a 2-day period using a 1-min time resolution,
- Collocation calibration with a reference monitor at 1-min time resolution for another 24-h period, immediately after the initial “training” period to ensure the training algorithm was sufficient,
- Immediate deployment of sensors to the area of interest,
- In field zero audits and if possible, use of temperature and humidity stabilization, hardware with the sensors to minimize unwanted response features, and
- Collocation testing with a reference monitor every 2 months of operation to ensure basic performance characteristics and correction algorithm review.

It must be stated that some of the experts felt that collocation periods longer than 2 days as suggested above were needed. The shorter collocation periods might not be sufficient in potentially training a given sensor to quantify a wide enough range of ambient O_3 concentrations.

As a general summary, the international SEs indicated that the current state of air quality sensors warrant localized calibration, and if possible, verification/certification of manufacturer’s performance claims be performed. This is needed because of concerns about variations between sensor components, potentially inaccurate data processing algorithms, detection to output conversion processes, response variability due to localized pollutant mixtures, and sensor drift issues, among others. At this time, the greatest progress in certification discussions have taken place in Europe and Asia. Considering that the CEN Technical Committee 264 has been engaged for approximately 3 years in developing consensus-based standards for member states to consider, the process is complex and many of the technical details remain to be resolved within that committee’s work. The Europe model of a comprehensive laboratory and/or field-based approach for sensor certification was very well received by those in attendance. Many requested the US EPA workshop organizers attempt to stay abreast of the EU’s efforts. Sensor manufacturers expect the EU to be successful in implementing Working Group 42 efforts and hope that any future US or independent third-party approach to sensor certification, if it was ever attempted, would heavily rely upon the performance benchmark values of the CEN. This sentiment was voiced because of the manufacturer’s desire for common standards to promote global-scale marketing of sensor products without the conflict of different country-specific performance requirements.

3.3. Tribal/state/local air agency panel: perspectives on sensor use and certification

This session included a panel dedicated to Tribal/State/Local air agency perspectives on sensor applications. The panelists included Kris Ray (Confederated Tribes of the Colville Reservation), Paul Fransioli (Clark County Department of Air Quality, Nevada), Gordon Pierce (Colorado Department of Public Health and the Environment), and Andrea Polidori

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

(South Coast Air Quality Management District). Four charge questions were provided for each panelist to address as follows:

- How would your organization use non-regulatory air quality measures?
- Please identify sensor applications and needs.
- Are there specific data quality objectives (DQOs) or measurement criteria that are most important in non-regulatory data for them to be useful?
- What sampling time intervals are important for your end-use of non-regulatory sensor data, and why?

During the discussions, it was evident that all four air agencies were dealing with sensors on one level or another, from simply working with other groups to performing full laboratory testing. For non-regulatory air quality measures, a range of sensor types were being used, with costs ranging from a few hundred dollars to a few thousand (USD). Non-regulatory uses were varied and included assessing pollutant concentrations in areas without monitors or areas that have experienced growth, assessing the impact of regulatory measures, screening risk assessments, defining hot spots, evaluating facilities, supplementing regulatory networks, event sampling, validating models, responding to community/environmental justice concerns, and education/outreach. While these non-regulatory uses were typically performed with non-regulatory monitors and as such cannot be used for determining compliance with air quality standards, the panelists noted that sensor data could be very useful to determine if air quality issues do exist. Based on findings, decisions could be made if new or additional air monitoring is needed, if higher quality sensors need to be employed, or if additional regulations need to be implemented.

Specific applications that were noted by the air agencies include monitoring wildfires and prescribed burns, residential wood-burning, specific industries, and oil and gas development. Other applications included forecasting, air toxics studies, and community studies and grants. It was noted that low-cost sensors are needed and very useful, but they also need to have at least a certain level of quality for the data to be useable. While PM sensors are currently developed enough to provide a level of confidence in the data, sensors for other pollutants such as volatile organic compounds (VOCs) still need significant work. PM sensors still need further development for higher concentration applications. The panelists also discussed that handling the large data sets generated by sensor networks is a problem and how data are displayed and interpreted is a challenge to the general user. As pointed out, the initial cost of a given sensor might not reflect its cost-effectiveness when deployment and data recovery/data reduction requirements are factored into the full cost of use. Education is needed to ensure that users can collect the highest quality data.

Panelists stated that DQOs that are the most important can be application specific. However, accuracy, repeatability, and good sensitivity are critical for all sensors. Different tiers to quickly decide what sensors might be appropriate for an application was discussed. One option noted was a 5-tier system such as regulatory monitoring, fenceline monitoring, community monitoring, health research, and environmental education. Air agencies noted a

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

desire for secure data transfer communications, non-proprietary software, standardized communication protocols, and standardized data outputs in common units.

For sampling time intervals, agencies recognized that longer-term averages (e.g., 1-h, 8-h, 24-h) are generally the primary interval needed for comparisons to criteria air quality standards, but shorter timeframes of 1-min are very useful for seeing short term spikes that can be related to specific sources. It was discussed that public education is very important so that people recognize timeframes that relate to potential health concerns, and that a short spike in concentration values does not necessarily translate into an immediate health concern.

Overall, this panel session provided some useful discussion as Tribal/State/Local air agencies are often the first contacts for the general public. The public is looking for guidance on how to use sensors for air monitoring. There is often hesitation from air agencies on using sensors due to lack of knowledge on how sensors work, the perceived or actual accuracy of sensors, and a lack of resources to take on additional work. Air agencies may need to consider addressing these and other issues as sensor networks become more widespread and data quality improves.

3.4. Manufacturer's perspectives on sensor performance targets

During the workshop, two open-ended sessions were held for those involved in building sensors and integrated sensor systems to provide their perspectives on sensor performance targets. Approximately 10 individuals provided commentary, with the majority representing systems integrators – those buying original equipment manufacturer (OEM) sensor components and combining these with other ancillary electronics and algorithms to have a fully working monitoring system. Some of the perspectives shared were from organizations developing sensor components and advancing the fundamental measurement technology over time. In addition to sharing comments, about half of the participants opted in to sit on an impromptu panel session and receive questions from the audience.

An area of commonality among all the participants was expressed support for sensor performance targets. Additionally, that the performance targets should be at the "integrated device" level rather than at the OEM sensor level, given the system design including its packaging and data handling/processing algorithms, may modify the performance of OEM sensors. Performance targets were viewed as driving technology development, identifying well-performing devices from substandard technology, and aiding customers in their purchasing decisions. Several representatives noted their desire to see harmonization in sensor performance targets that may be developed in the US versus by international bodies (e.g., CEN Technical Committee 264 Working Group 42). There was not uniformity among the representatives toward a tiered versus binary approach for sensor performance targets. Various speakers noted challenges in setting performance targets given the wide variety of potential use cases and environments of measurement, and they also noted that users are almost always pushing the limits of sensor technology in unique, and sometimes unexpected, applications.

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

In terms of testing protocols, a variety of perspectives were shared. Collocation with reference monitors was noted as an attractive approach; however, questions were raised about whether performance in one geographic location would translate to other geographic areas when meteorological and other environmental conditions might be highly variable. Additionally, cost was a concern if a performance target required testing in multiple geographic areas. Laboratory testing was also noted as attractive due to its reproducibility; however, experience has shown that sensors may perform very differently in real-world scenarios and thus the potential need for field calibrations to be part of any testing protocol. A “gold standard” reference monitor was proposed, that could be used in both laboratory and field environments to support comparability in performance tests. Several representatives noted the need for simplicity and urged a rapid drafting of performance standards, even if interim in nature or without a full-fledged certification program, to help guide the fast-paced technology development.

Finally, manufacturing representatives noted the unique challenge of setting performance targets for sensors that are utilizing sophisticated algorithms and may process data from a group of sensor devices (nodes) in a network. A variety of perspectives were shared on whether manufacturers would be willing to disclose data processing algorithms to organizations evaluating the technology, be entirely open source with algorithms, and whether performance targets should consider a calibrated network approach. It was noted that the network calibration approach was so new, that peer reviewed research was needed, but conducting this research would be challenging if sensor raw data and proprietary algorithms are not shared with users conducting independent research. Even with this caveat, some panelists suggested that encouraging research in network calibrations is being reported in the literature and at international conferences.

3.5. Non-regulatory monitoring and associated measurement performance targets

This session covered non-regulatory monitoring needs associated with use of low-cost sensors for measurement of $PM_{2.5}$ and O_3. These discussions featured real-world applications or case studies where sensor technologies were being used for different purposes. Viewpoints associated with personal exposure monitoring, epidemiological data needs, air quality modeling considerations, and atmospheric exposure assessment, among other monitoring scenarios were discussed. We report summaries of the case study presentations given during the two-day public workshop as well as summaries from the third day follow-on discussions (charge questions shown in Table S2). These follow-on discussions resulted in non-consensus-based identification of suggested sensor performance attributes (e.g., monitoring focus/use, key data quality indicators, value or range of suggested performance benchmark). A variety of data quality terminologies and definitions were discussed by the SEs and are reported here without bias or endorsement. While discussions concerning $PM_{2.5}$ and O_3 findings on performance were often quite similar, they are presented separately here as an aid to the reader.

3.5.1. $PM_{2.5}$ focus—For $PM_{2.5}$, the SEs included Mike Hannigan (University of Colorado-Boulder), Rima Habre (University of Southern California), and James Schauer (University of Wisconsin-Madison). As described by Hannigan and others during this

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

portion of the workshop, low-cost $PM_{2.5}$ sensors have been used in a wide range of applications. These include investigations of spatio-temporal variability, indoor/outdoor residential air quality monitoring, emission and other source characterizations, personal exposure monitoring, and sensor network applications to provide supplemental or complimentary ambient air quality monitoring.

Hannigan provided an example where he investigated the appropriate averaging time of collocated low-cost $PM_{2.5}$ sensors (HAPEx Nano Monitor; Fig. 2) to achieve the best results. In particular, 1-min averaging times resulted in poorer agreement (median $r = 0.65$) as compared to 1-h time intervals ($r = 0.93$). Even so, this sensor yielded very poor accuracy as compared to direct gravimetric analysis comparisons (adjusted $R^2 = 0.47$). This highlighted one of the re-occurring points made by many of the SEs, namely, the ability of a low-cost $PM_{2.5}$ sensor to have acceptable linearity and precision might not equate to it producing a concentration value ($\mu g/m^3$) of benefit for any intended application. Care must be taken by the end users in correcting raw sensor response to a meaningful value. Collocated sensor calibration with reference monitors was often reported during the workshop as key to developing and applying such correction factors.

Habre provided an overview of the Los Angeles PRISMS Center (https://www.mii.ucla.edu/research/projects/prisms/) BREATHE informatics platform where a low-cost PM sensor (Habitat Map AirBeam) was integrated for real-time personal exposure monitoring, in conjunction with sensors collecting human activity data and biomonitoring sensors for epidemiological assessments of pediatric asthma. Desirable performance features of the PM sensor are summarized in Table 4. In brief, these features included the need for accuracy and precision errors < 15%, low participant burden (e.g., safety, weight and form factor of sensor), and other data integration and ease-of-use functions (e.g., internal storage, secure wireless data transmission, 12^{+} h battery capacity). She reiterated that sensor performance features of greatest need are highly dependent upon the intended use and exposure or health research question. For example, if an epidemiological study was focusing upon inter-personal (between-person) exposure comparisons within a cohort, sensor accuracy might be the most important performance feature. Conversely, sensor precision and bias (e.g., drift) performance attributes might be more important for an epidemiological investigation involving within-person monitoring over a longitudinal period. Habre reported that the successful integration of a low-cost PM sensor as part of a system or informatics platform is providing definitive and highly spatially and temporally resolved information on potential human exposure characterization and its relation to acute health outcomes.

Schauer noted many pros and cons associated with low-cost PM sensor use. Positive features included low-cost and ease-of-use, portability that allows for wide spatial coverage across multiple monitoring scenarios, and sensor failure is often easy to diagnose (obvious poor data quality). He shared negative performance features of low-cost PM sensors such as their susceptibility to systematic bias, time dependent drift, and initial as well as “aging” sensor accuracy. He described the difficulty in conducting field-based calibrations and the sensor’s susceptibility to negative influencing factors (e.g., relative humidity, temperature extremes, variable response to differing aerosol composition). Even so, Schauer suggested a number of ideal applications for the use of low-cost PM sensors including supplemental monitoring to

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

augment true gravimetric-based assessments, spatio-temporal investigations, investigations involving “hot spots”, source tracking and characterization, and assessing the value of emission source interventions and control measures. Schauer suggested specific precision and accuracy values of benefit to a wide range of application scenarios (Table 5). The suggested precision and accuracy errors ranged from 10 to 50%. Source tracking and scaling filter-based measurement scenarios were judged as being able to support higher error rates (50%), while micro-environmental monitoring, intervention assessments and spatio-temporal monitoring required significantly lower precision and accuracy errors (< 25%).

There was also discussion that most of the low-cost sensors do not measure $PM_{2.5}$, rather they measure particle number in a specific size range and then use a conversion factor (user derived, or manufacturer supplied, based on reference data) to convert the measured number of particles to $PM_{2.5}$. There may be conversion errors because the size range of particles in the environment is different than the size range for the reference data. For example, some of the sensors only measure particles larger than 1 μm. Particles in that size range could reasonably comprise 75% of the $PM_{2.5}$ mass or 50% of the mass. In such circumstances, the sensor could be performing perfectly but our ability to interpret the measurement correctly is limited by incomplete knowledge of the particles in the environment. In addition, the low-cost PM sensors measure of particle number is affected by particle composition. Differences between the reference particle composition and the observed particles could cause a bias. This type of error is not associated with poor performance of the sensors but can result in misinterpretation of the relationship between the measured quantity and $PM_{2.5}$ mass. Additionally, for optical sensors (bulk or single particle), size is usually specified as physical size. When referring to aerodynamic particle size, that number becomes larger.

The third day discussions were focused on the issue of performance targets for sensors. While no formal survey information on performance attributes were obtained, the responses from the SEs were summarized as general findings (Table 6). The listed sensor performance attributes should not be considered suggestive about any specific US EPA (recommended) means of establishing such a value. Rather, they are meant to be indicative of the general workshop discussions. The summary provided in Table 6 sometimes required aggregating different performance attributes in an attempt to define a common measure. For example, precision was discussed in terms of percent error, ± a given concentration value, root mean square error, or coefficient of variability among others. Aggregation was performed only when a common metric was described (e.g., a type of %) even if the means of calculating that measure might be variable. The range of performance attributes, and associated values, voluntarily shared by the SEs as necessary attributes were widely variable. Accuracy was considered a key performance attribute with acceptable ranges suggested to be 10–100% depending upon the purpose of the data collection. Median acceptable accuracy error values were calculated to be 25%. Precision was another performance attribute generating much discussion and in particular, the statistical term of merit. Regardless of the statistical term (e.g., root mean square error, coefficient of variation) most of the volunteered terms could be designated as percent error. We captured a range of 10–50% precision error with the median approaching 25%.

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

Significant concerns with bias (and its various definitions) were expressed throughout the workshop. Both mass concentration as well percent error value terms were often shared. In general, a concentration range of acceptable bias from 1 to 5 $\mu g/m^3$ or 2.5–50% over the full concentration measurement range were reported. There was general agreement that a highly linear response need ($r \geq 0.8$) exists for PM sensor-based measurements over the full measurement range. The median acceptable correlation was estimated to be $r = 0.89$. Only a few of the SEs shared their opinions regarding a needed limit of detection (1–4 $\mu g/m^3$). Even so, many provided input on the needed performance range. Many SEs from the US suggested a range of approximately 0–200 $\mu g/m^3$ while SEs from Europe and Asia suggested a wider response range (0–1000 $\mu g/m^3$). Lastly, of the performance attributes discussed, acceptable relative humidity ranges of performance varied between 0 and 100%. The upper range suggested by some (100%) would indeed be a very hard performance target for low-cost sensors to achieve considering the well-known relative humidity bias issues for the popular nephelometer-based PM sensors.

3.5.2. O_3 focus—Workshop O_3 discussions were led by George Thurston (New York University), Ronald Cohen (University of California-Berkeley) and Anthony Wexler (University of California-Davis). Thurston shared information where spatial consideration of O_3 measurements was a determining factor in observed measurements, and where such measurements using sensors would be highly valuable for epidemiological assessments. For example, 1-h based outdoor O_3 concentrations were reported to vary significantly (up to a factor of 3) depending upon the height at which the measurement was taken in one urban environment (Restrepo et al., 2004). Other examples he provided where low-cost O_3 sensor measurements would be useful included greater spatio-temporal coverage to supplement regulatory monitoring data, and deployment in proximity to near road locations, where O_3 "quenching" takes place, as O_3 reacts locally with fresh emissions of nitric oxide (NO). Thurston indicated that one of the most significant benefits of additional personal sensor measurements would be in support of future epidemiological assessments of the adverse health effects at the individual rather than overall population level.

Cohen shared his perspective on the applicability of O_3 sensors for non-regulatory data collections from his experience with the BEACO$_2$N study. He presented measurement findings associated with their low-cost sensor pod (Kim et al., 2018). While the BEACO$_2$N study often had a primary focus on CO_2 observations, many of the same considerations could be applied to O_3-based measurements. Published findings (Turner et al., 2016) reported that use of a large number of low-cost sensors had the potential of outperforming a smaller number of more expensive monitors with respect to estimating near-road line source emission rates. In other words, a network-based approach using low-cost O_3 sensors might achieve significant benefit to those trying to measure and model spatial variability/chemical mechanisms of gaseous pollutants. He further extrapolated on the value of a network-based approach relative to performance considerations which might best be described as "the sum is greater than any single measurement". Weather forecasting was cited as an example where it was the cumulative information obtained from the synthesis of many measurements provided significantly more useful information as compared to any one measurement itself. Precision and accuracy of an O_3 network approach would be expected to improve with

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

averaging over integration time as well as potentially being dependent upon the square root of the number of sensors in the network.

Wexler shared his perspective on the potential value (benefits) of low-cost O_3 sensors. He indicated that their low-cost provided the means to deploy them in large numbers. Such deployments would provide the opportunity to capture spatial-temporal variability here-to-fore unable to be achieved with more traditional monitoring approaches. This is especially important when one considers fast reaction rates for select gas pollutants (e.g., NO, O_3, NO_2). Such situations might be expected to occur in near road locations and therefore more spatial coverage of O_3 monitoring would be of benefit in assessing O_3 gradients over any given geographical location. Low-cost O_3 sensors have the potential to gather information on the spatial variability of ever decreasing ambient concentrations, now often well below the 70 ppb, 8-h primary standard (US National Ambient Air Quality Standards). He pointed out that it would be inappropriate to expect low-cost O_3 sensors to be as reliable or as accurate as federal reference monitors, although these monitors also have uncertainties due to light scattering and absorption by other pollutants such as PM and VOCs. Their performance attributes (e.g., accuracy, precision) must be judged specifically with their intended use and therefore no single set of criteria best serves the scientific community at large relative to establishing any performance standards for such technology.

The third day discussions captured general viewpoints from the SEs on suggested performance attributes for O_3 sensors. Similar to the $PM_{2.5}$ discussions, a wide range of views were shared concerning O_3 sensor performance. There was general agreement that this technology already provides often good to excellent agreement when compared to federal reference monitors. The most frequently heard negative opinion concerning the current state of the technology was co-responsiveness to other gas species (e.g., oxides of nitrogen or NOx) and the uncertainty of its useful lifespan (sensor aging). Table 7 provides a general summary of suggested O_3 sensor performance attributes. The SEs covered a wide range of applications (e.g., epidemiology, atmospheric modeling, exposure assessment). Even so, the suggested performance needs for O_3 sensors were significantly smaller in overall scale than previously discussed for $PM_{2.5}$. For example, a tighter range of accuracy error was suggested (7–20%) with an equally smaller degree of acceptable bias (5–25% or 2–5 ppb). Precision error of 7–20% was deemed necessary and currently achievable. Linearity needs from $r = 0.70$ to $r = 0.95$ were expressed. SEs indicated that the current state of O_3 technology should be able to achieve a 0–200 ppb working range. Other performance attributes were discussed (e.g., response time, cross-interferences, need for limits on relative humidity and temperature effects) but to such a limited degree it cannot be successfully categorized here.

3.6. Perspectives on a testing and certification program: scope and structure

This session involved discussion on the scope and structure of a testing and/or certification program for low-cost sensors. In general, a testing program [e.g., Air Quality Sensor Performance Evaluation Center (AQ-SPEC), EU JRC, US EPA, etc.] provides objective data on the technical performance of sensor devices either in the laboratory and/or in the field. No judgements are rendered, and no recommendations are given as to how well a device

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

performs for a given application. The US EPA's former Environmental Technology Verification Program (ETV) represents an example of this nature (https://archive.epa.gov/nrmrl/archive-etv/web/html/). A certification program commonly goes beyond testing and evaluation to provide judgement on the data quality and fit-for-purpose of a device for a given application or set of applications. In addition, a certification program would provide a stamp of approval from an authoritative body [e.g., Underwriters Laboratories (UL) for electrical devices, Energy Star Program].

In this session, SEs including Edmund Seto (University of Washington), George Allen (NESCAUM), and Stephan Sylvan (US EPA), provided presentations on considerations for a tiered and/or binary sensor certification program. Seto provided an example of Tier 1 (apples to apples) versus Tier 2 (apples to oranges) certification testing (Seto, 2018). Tier 1 was based on standardized testing protocols and recommended performance measures that aim to address the need for highly reproducible testing results. Examples provided for PM included use of reference aerosol composition (e.g., Arizona Road dust), temperature, humidity, concentration ranges, reference instruments and methods, and specific performance metrics (e.g., correlation, bias, sensor to reference plots; Seto, 2018). He suggested that Tier 1 testing might be more easily adopted by manufacturers as a standardized testing regime because it could be performed in a laboratory under controlled conditions to provide reproducible results and allow for apples to apples performance comparisons among different sensors. Tier 1 results may be misleading because they may not represent realworld use cases (e.g., indoor monitoring, mobile monitoring, personal exposure assessment, etc.). Tier 2 was based on developing a protocol or process that is use-specific, acknowledging that use cases have different requirements. The first step would be defining a sensor use case, followed by specific DQOs/DQIs for that use, a Quality Assurance Project Plan (QAPP), and systematic data reporting/sharing of results. Tier 2 testing would ideally be conducted in real-world field settings to understand how a sensor performs outside of controlled laboratory conditions or other specific settings. As part of Tier 2 testing, sensors collocated with a reference instrument in the field can help identify bias, noise, interferences, and degradation over time. Seto noted that consumers want a device that is trusted and would like recommendations on which device to use and how good it will perform for their use case.

Allen acknowledged that certification, while attractive, is complex, expensive, and difficult to implement, but would still be worth considering as a long-term goal (Allen, 2018). He suggested constructing a more formalized testing and evaluation program that includes effective communication of the results to a wide range of end users with differing data quality needs. Allen noted that current testing programs, like AQ-SPEC, were at the "high end" of the spectrum, requiring large resources and providing results that are more suited to a technical audience. He proposed a list of parameters as an example of a testing protocol which included precision, bias, linearity, goodness of fit using either R^2 or root mean square deviation (RMSD), sensitivity, level of detection (LOD), and interferences such as relative humidity for PM sensors and NO_2 for electrochemical O_3 sensors (Allen, 2018). Allen also provided an example of a five-tiered system for sensor performance targets that describe a sensor's suitability for a given use case (Table 8). He noted that the errors were a function of averaging time and needed to be specified.

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

Sylvan provided a perspective on several ecolabel and certification programs, including the Energy Star Program. These programs commonly promote trust, provide unique value, and offer simplicity. Sylvan noted that ecolabel/certification make sense when there are many buyers (or a market) that cannot easily distinguish product performance, a significant demand, misleading/confusing claims, market confusion leading to under investment, and existence of at least one credible "actor" that would issue the ecolabels with a market-based approach as opposed to a legal or technical approach (Sylvan, 2018). In the case of ratings/multi-tier approaches, Sylvan stated that these approaches are often complex and only make sense in special cases because they require the same elements of an ecolabel/certification program along with sophisticated buyers, a deliberative purchasing process, and credible actors with sufficient, ongoing resources to develop/maintain/enforce multiple standards. In general, multi-tier certifications have not worked however a few success stories include Leadership in Energy and Environmental Design (LEED) rating system for commercial buildings, Electronic Product Environmental Assessment Tool (EPEAT) for certain electronic devices, and level (developed by a trade association) for office/institutional furniture. Sylvan noted that even the most professional buyers want simple and trusted choices.

In summary, opinions on the scope and structure of a testing and/or certification program varied among the SEs. Some were against a certification program given the complexities, nuances, and long-time frames involved. As an alternative, they recommended sensor performance targets, voluntary standards, and best practices. Others preferred pursuing a testing type structure like AQ-SPEC to achieve progress in the short/medium term, while continuing to explore how to structure a viable certification program. Most agreed that the target audience for any effort would be the general consumer, mass market sector. Opinions varied on whether tiering should be an element of any program. Some SEs preferred presenting objective testing results with no tiering or judgement involved, while others were partial to a binary tiering strategy that defines a lower threshold to separate out very poorly performing devices from the remainder. Several SEs believed that a multi-tiered approach, harmonious with the EU efforts, should be pursued. Lastly, the SEs indicated that regardless of the scope and structure of a program, it needs to provide flexibility to handle innovation and technological development. Additionally, the data and metadata used to summarize the technical performance of any device should be widely available and transparent. They emphasized that a portal is needed to effectively communicate and share findings, information, and data on low-cost sensor evaluations.

4. Conclusions

This workshop was attended by national and international participants from multiple sectors, which helped to obtain a diverse set of perspectives and feedback on sensor performance targets. It provided participants with a valuable opportunity to discuss: the air sensor performance state-of-science; how sensors can and are being used for non-regulatory purposes; perspectives on sensor performance targets of greatest importance; and, lessons learned from organizations that have established or ongoing plans for implementing sensor certification programs.

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

There are many emerging issues that could drive the need for sensor performance targets. For example, California's Assembly Bill (AB) 617 passed in June 2017, requires community air monitoring systems in locations that are adversely impacted by air pollution in order to provide information to help reduce exposures. It is anticipated that sensor technologies will play an important role in implementing AB 617 and ensuring data quality will be imperative. Another issue is the increased prevalence of wildfires which brings a need to rapidly obtain and communicate air quality data to protect public health. For this application, sensors are advantageous as they are easy to deploy (individually or in networks), portable, obtain real-time data that can be quickly communicated, and allow for monitoring in dangerous conditions. The quality of data is critical not only to inform immediate emergency response efforts but also in exposure and health studies aimed at detecting and quantifying public health impacts on communities for a range of health outcomes. Public confusion can result when such individual and short-term personal sensor concentrations are inappropriately compared to ambient air quality standards with different measurement methods and averaging times. Matters such as these will continue to inform the dialogue about sensor performance targets and appropriate public health messaging.

The topic of sensor performance targets also introduces many other issues and questions to consider. For example, if implemented, what would the structure of a certification-type program look like (Pass/Fail? Tiering? Ecolabel approach?); given the varied backgrounds of users ranging from technical to non-technical, would guidance or tools need to be provided for proper use and deployment of sensors, QA/QC, and interpreting and visualizing data; and cost to both manufacturers for "certifying" their products and users for purchasing, operating, maintaining, processing and storing data, among other activities.

A significant number of challenges exist in advancing the purposeful use of sensor technologies. These include how data are being collected, processed, and reported. Should any determination of sensor performance be specific to the instrument itself or should it be performed at the final stage of reporting where vendor-specific data processing (including machine learning and other data treatments) might be performed? A real need for end users is greater confidence in the accuracy of sensor data. Continued discussions that bring together a variety of expertise, from manufacturers to users, will be beneficial in shaping the future of sensor technologies and potential performance targets or certification-type programs.

5. Next steps

Innovative measurement and information technologies are providing a tremendous amount of new air quality data. However, as discussed during the workshop, uncertainties exist regarding the characterization of these measurements and the end use of non-regulatory data. While opinions differed on the best approach, there was a basic desire to understand air sensor data in a more systematic way. Recognizing that sensors are developing rapidly, the US EPA plans to take an agile approach in setting priorities, developing outputs or programs, and conducting research to ensure that all current and future efforts support the continued need. Our workshop discussions with regulatory officials from Europe and Asia indicated that ultimately government organizations worldwide will choose their own best path on any

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

certification or suggested air sensor performance targets, and coordination or at a minimum enhanced dialogue on such efforts from all parties will be vital in ensuring the best outcomes for advancing air sensor technologies. Sensor manufacturers communicated during the workshop that consistent performance targets or certification requirements from governing institutions should be pursued.

The efforts described above are intended to provide governments and other stakeholders alike with streamlined, unbiased assessments of sensor performance, both initially and over time. Moving forward, the US EPA will consider the various opinions expressed during the workshop on the attributes required to characterize sensor performance, the need for various tiers of performance, the development of a common vocabulary when defining performance attributes, the desire to coordinate with other international programs and colleagues, and the market demand for certified products.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors acknowledge the contributions of US EPA workshop planning members (Stacy Katz, Gail Robarge, Jenia McBrian, David Nash) who contributed to establishing the goals of the workshop and worked to ensure its success. Stephan Sylvan of US EPA is acknowledged for his presentation on historical certification considerations. We wish to thank the participants of the 2018 Air Sensor Performance Deliberation Workshop and the subject matter experts who shared their perspectives on performance targets. Kelly Poole and the Environmental Council of States (ECOS) are recognized for supporting Tribal and State air agency participation in the workshop. US EPA staff including Andrea Clements, Amara Holder, Stephen Reece, Stephen Feinberg, and Maribel Colon are thanked for their efforts in supporting the workshop webinar question/answer periods that contributed to the overall discussions. Angie Shatas and Laura Bunte of US EPA are recognized for facilitating the workshop and collecting the extensive notes used in development of this article. The staff of ICF (US EPA contract number EP-C-14–001) are recognized for their selection of subject experts who attended the workshop and in handling logistics for conducting the workshop event. Ian MacGregor and the staff of Battelle (US EPA contract number EP-C-16–014) are acknowledged for their efforts in conducting the peer review literature report summarized in this article.

References

Allen G, 2018. The role of PM and ozone sensor. Testing/Certification Programs [presentation slides]. Retrieved from. https://www.epa.gov/sites/production/files/2018-08/documents/session_07_b_allen.pdf.

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

Chinese Ministry of Environmental Protection, 2013a. HJ 653–2013 Specifications And Test Procedures For Ambient Air Quality Continuous Automated Monitoring Systems For PM10 and PM2.5.

Chinese Ministry of Environmental Protection, 2013b. HJ 654–2013 Specifications and Test Procedures for Ambient Air Quality Continuous Automated Monitoring Systems for SO2, NO2, O3 and CO.

Clements AL, Griswold WG, RS A, Johnston JE, Herting MM, Thorson J, Collier-Oxandale A, Hannigan M, 2017. Low-cost air quality monitoring tools: from research to practice (A workshop summary). Sensors 17, 2478. 10.3390/a17112478.

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.

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

Gerboles M, 2018. The road to developing performance standards for low-cost sensors in Europe [PowerPoint slides]. Retrieved from. https://www.epa.gov/sites/production/files/2018-09/documents/session_01_b_gerboles.pdf.

Habre R, 2018. Performance targets for wearable PM2.5 sensors in epidemiologic studies (of pediatric asthma) using real-time enabled informatics platforms [PowerPoint slides]. Retrieved from. https://www.epa.gov/sites/production/files/2018-08/documents/session_04_b_habre.pdf.

Hannigan M, 2018. Using low-cost PM sensing to assess emissions and apportion exposure in Ghana [PowerPoint slides]. Retrieved from. https://www.epa.gov/sites/production/files/2018-08/documents/session_04_a_hannigan.pdf.

Kim J, Shusterman AA, Lieschke KJ, Newman C, Cohen RC, 2018. The Berkley Atmospheric CO2 Observation Network: field calibration and evaluation of low-cost air quality sensors. Atmos. Meas. Tech. 11, 1937–1946. 10.5194/atm-11-1937-2018.

Lewis A, Edwards P, 07 7 2016. Validate personal air-pollution sensors. Nature 535, 29–31. 10.1038/535029a. [PubMed: 27383969]

Lewis A, von Schneidemesser E, Peltier R, 2018. Low-cost Sensors for the Measurement of Atmospheric Composition: Overview of Topic and Future Applications. World Meteorological Organization, WMO-No.1215 Available at: https://www.wmo.int/pages/prog/arep/gaw/documents/Low_cost_sensors_post_review_final.pdf.

Martin NA, 2018. The road to developing performance standards in Europe for low-cost sensors-Part 1: example of implementation of an existing reference instrument standardization method [PowerPoint slides]. Retrieved from. https://www.epa.gov/sites/production/files/2018-08/documents/session_01_a_martin.pdf.

Ning Z, 2018. Application of UAV based sensor technology for ship emission monitoring and high sulfur fuel screening in Hong Kong [PowerPoint slides]. Retrieved from. https://www.epa.gov/sites/production/files/2018-08/documents/session_01_c_louie.pdf.

Penza M, 2018. Developing air quality sensors: perspectives and challenges for real applications [PowerPoint slides]. Retrieved from. https://www.epa.gov/sites/production/files/2018-08/documents/session_01_d_penza.pdf.

Restrepo C, Chen LC, Thurston G, Clemente J, Gorczynski J, Zhong M, Blaustein M, Zimmerman R, 2004. A Comparison of Ground-Level Air Quality Data with New York State Department of Environmental Conservation Monitoring Stations Data in South Bronx. (R827351). Atmospheric Environment 10.1016/j.atmosenv.2004.06.004.

Schauer J, 2018. Deployment of low-cost PM sensors: aligning sensor performance and data quality needs [PowerPoint slides]. Retrieved from. https://www.epa.gov/sites/production/files/2018-08/documents/session_04_c_schauer.pdf.

Seto E, 2018. Apples to apples vs. Apples and oranges: a perspective on tiered lab vs. Field testing of PM instruments [PowerPoint slides]. Retrieved from. https://www.epa.gov/sites/production/files/2018-08/documents/session_07_a_seto.pdf.

Sylvan S, 2018. Ecolabels/certifications: when they make sense & how to ensure success [PowerPoint slides]. Retrieved from. https://www.epa.gov/sites/production/files/2018-08/documents/session_07_c_sylvan.pdf.

Turner AJ, Shusterman AA, McDonald BC, Teige V, Harley RA, Cohen RC, 2016. Network design for quantifying urban CO_2 emissions: assessing trade-offs between precision and network density. Atmos. Chem. Phys. 16, 13465–13475. 10.5194/acp-16-13465-2016.

US EPA, 40 CFR Part 50 – National Primary and Secondary Ambient Air Quality Standards.

US EPA, 40 CFR Part 53 – Ambient Air Monitoring Reference and Equivalent Methods.

Williams R, Nash D, Hagler G, Benedict K, MacGregor I, Seay B, Lawrence M, Dye T, 9 2018. Peer Review and Supporting Literature Review of Air Sensor Technology Performance Targets. EPA Technical Report Undergoing Final External Peer Review. EPA/600/R-18/324.

Woodall G, Hoover M, Williams R, Benedict K, Harper M, Soo Jhy-Charm, Jarabek A, Stewart M, Brown J, Hulla J, Caudill M, Clements A, Kaufman A, Parker A, Keating M, Balshaw D, Garrahan K, Burton L, Batksa S, Limaye V, Hakkinen P, Thompson R, 2017. Interpreting mobile and handheld air quality sensor readings in relation to air quality standards and health effect

reference values: tackling the challenges. Atmosphere 8, 182. 10.3390/atmos8100182. [PubMed: 29093969]

Zartarian V, Duan N, Ott W, 2006. Basic Concepts and Definitions of Exposure and Dose. pp. 33–63. 10.1201/9781420012637.ch2.

Web References

Deliberating performance targets for air quality sensors workshop. https://www.epa.gov/air-research/deliberating-performance-targets-air-quality-sensors-workshop.

Air quality sensor performance evaluation center (AQ-SPEC) South Coast air quality management district. http://www.aqmd.gov/aq-spec.

Evaluation of emerging air pollution sensor performance" air sensor Toolbox, US EPA https://www.epa.gov/air-sensor-toolbox/evaluation-emerging-air-pollution-sensor-performance.

Underwriters Laboratories (UL). https://www.ul.com.

Energy star program. https://www.energystar.gov/.

US EPA air sensor Toolbox for citizen scientists, researchers and developers. https://www.epa.gov/air-sensor-toolbox.

European Commission, Joint Research Centre https://ec.europa.eu/info/departments/joint-research-centre_en.

Pediatric research using integrated sensor monitoring systems (PRISMS). https://www.nibib.nih.gov/research-funding/prisms.

US EPA environmental technology verification program (ETV). https://archive.epa.gov/nrmrl/archive-etv/web/html/.

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

HIGHLIGHTS

- Certification standards do not exist for low-cost air sensors.
- Some sensor performance targets are being considered by various organizations.
- Lack of standards impacts data quality and confidence in using sensor data.
- In literature, reported sensor performance attributes are highly variable.
- Possible next steps include recording best practices and quality assurance results.

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

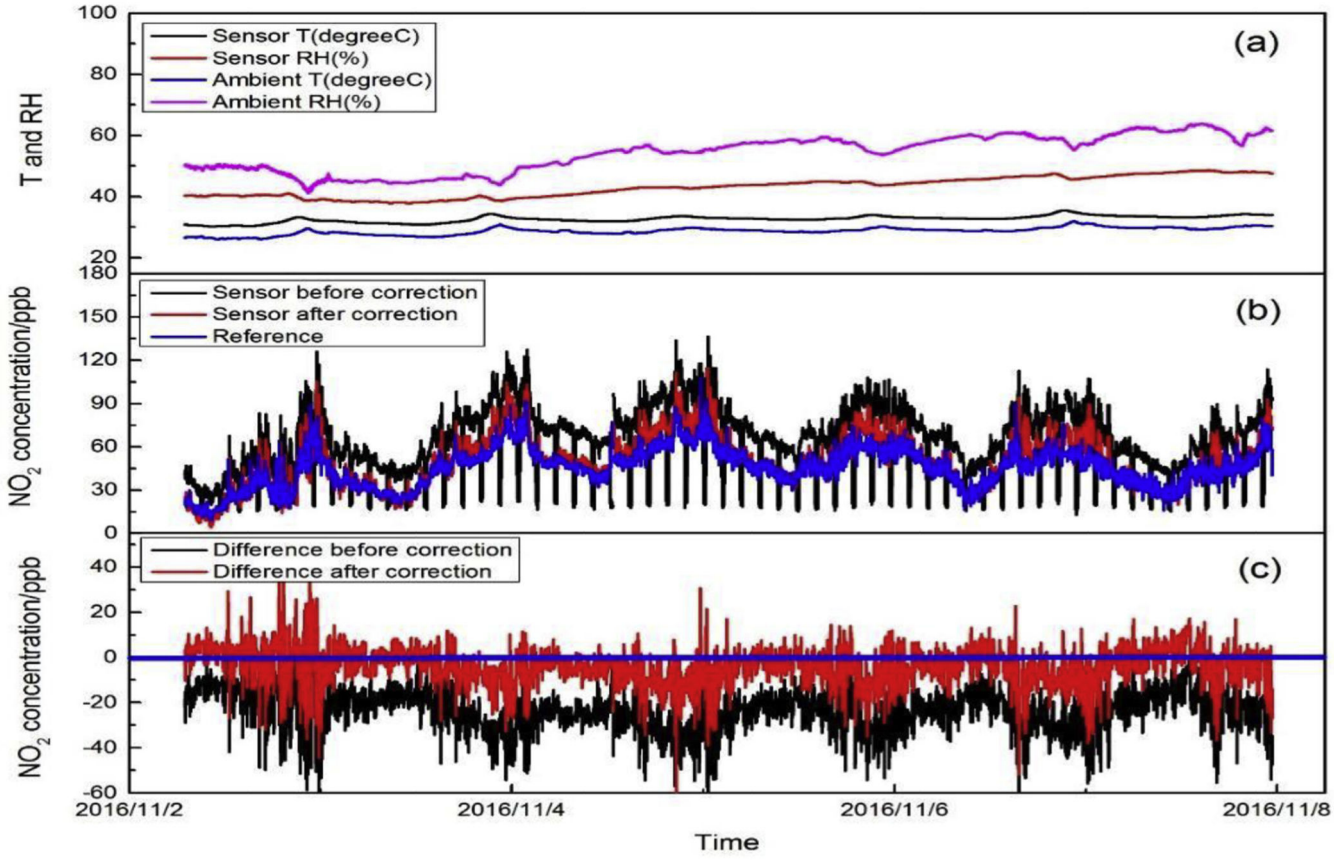

Fig. 1.

Sensor performance relating to changes in temperature, RH, and drift. Example of a low-cost NO_2 electrochemical sensor with an O_3 filter. Test is displaying responses to changes in temperature, relative humidity, and drift over time (Ning, 2018).

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

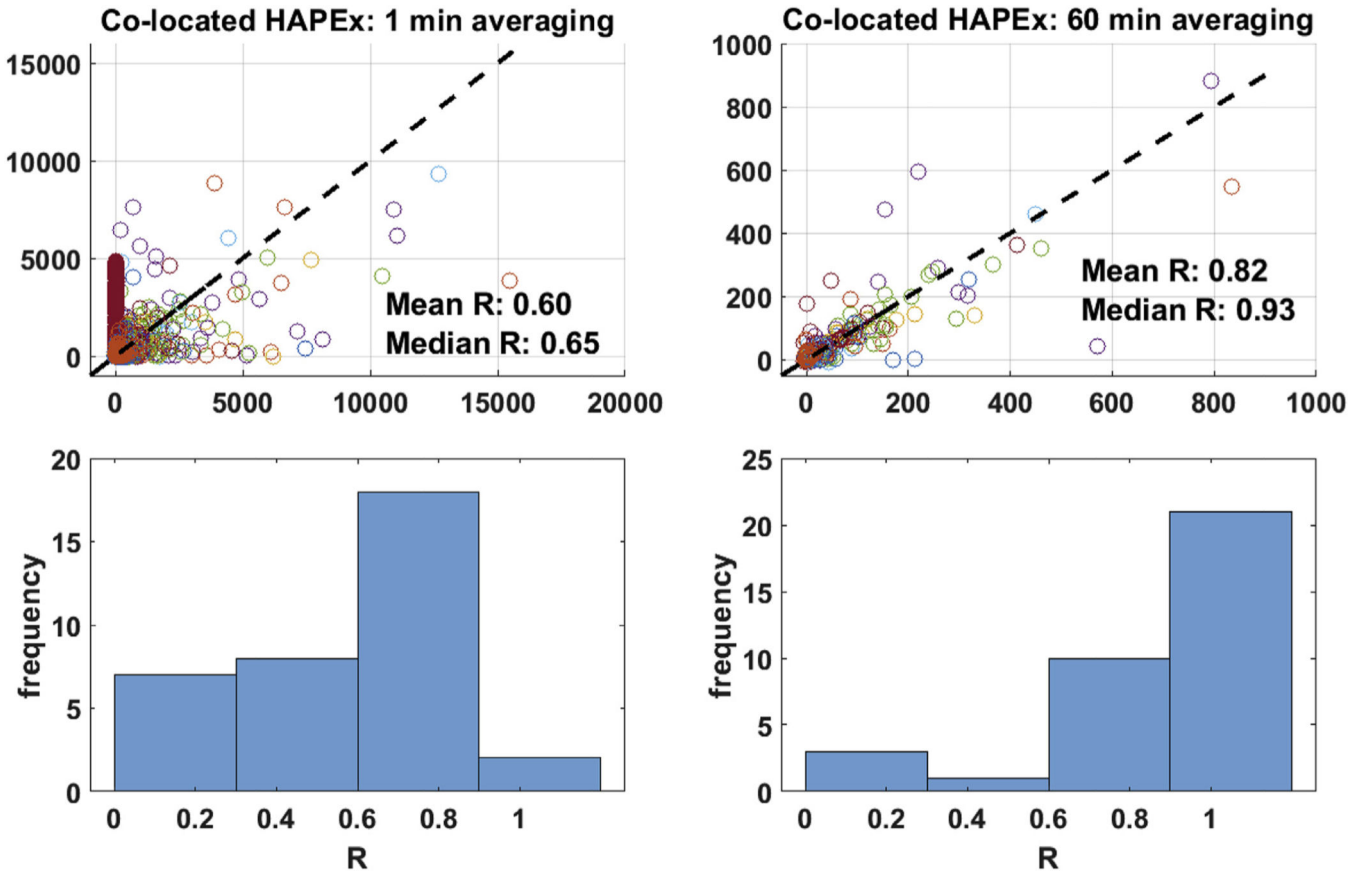

Fig. 2.

Plots demonstrating the 1 and 60-min averaging times of 35 collocated low-cost (~$100 USD) Sharp GP2Y10 light scattering sensors during a 48-h deployment (Hannigan, 2018).

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

Table 1

Frequency and number of times information sources contained DQOs/MQOs[a] for different performance attributes (Williams et al., 2018).

Performance Characteristic/DQI[b]	$PM_{2.5}$	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[c]	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] MQO =Measurement Quality Objective.

[b] 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.

[c] "Other" category captures all performance characteristics not among the 10 listed.

Atmos Environ X. Author manuscript; available in PMC 2021 July 27.

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

Table 2

Measurement parameters for air sensor performance (Penza, 2018).

Transducer	Sensitivity	Selectivity	Stability	Limit of Detection	Open Questions
Electrochemical	High	Variable	Improved	ppb	Interference, calibration, signal processing
Spectroscopic and NDIR	High	Variable	Low	ppm	Interference, calibration, signal processing
Photo-Ionization Detector	High	Low	Improved	ppb	Interference, calibration, signal processing
Optical Particulate Counter	High	Improved	Improved	$\mu g/m^3$	Interference, calibration, signal processing
Metal Oxides	High	Variable	Low	ppb	Interference, calibration, signal processing
Pellistors	High	Low	Improved	ppm	Interference, calibration, signal processing

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

Table 3

$PM_{2.5}$ criteria/DQIs defined relative to any instrumentation for regulatory air monitoring requirements in the U.S., EU, and *China* (Williams et al., 2018; Table B1). Note: A description of the terms and abbreviations are listed in Table S3.

Performance Attributes/DQIs	Decision Support
Accuracy/Uncertainty	RPD_{flow}: ≤2% %$Diff_{specifiedflow}$: ± 5%, ± 5% %$Diff_{onepointflow}$: ± 4% %$Diff_{multipointflow}$: ± 2% T_{amb} (°C): ± **2**, ± *2,* ± 2 P_{amb} (mm Hg): ± 10, ≤ *7.5*, ± 7.5 RH_{amb}: ± 5% Clock/timer (sec): ± **60**, ± *20* D_{50}: *2.5 ± 0.2* μm Collection efficiency: σ_g = 1.2 ± 0.1 Average flow indication error: ≤ *2*% Slope: *1 ± 0.15*, **1 ± 0.10** Intercept (μg/m^3): 0 ± 10, 0 ± 2 Aerosol transmission efficiency: ≥ 97% Expanded uncertainty: < 25% in 24-hr averages Zero level: < 2.0 μg/m^3 Zero check: 0 ± 3 μg/m^3 Maintenance interval: < 14 days
Bias	None
Completeness	Completeness (%): 85, ≥90
Detection Limit	Detection limit (μg/m^3): < 2.0, 2 T_{amb} resolution: 0.1°C P_{amb} resolution: 5mm Hg
Measurement Duration	Measurement duration: 60min
Measurement Frequency	Flow rate measurement intervals: ≤30s
Measurement Range	Concentration range: 0–1000 μg/m^3, (0–1000 24-hr avg, 0–10000 1-hr avg μg/m^3), 3–200μg/m^3
Precision	CV_{conc}: ≤5%, ≤ 15% CV_{flow}: < 2%,≤ 2%, (Avg: ≤2%, Inst.: ≤5%) σ:≤2μg/m^3 Precision: < 2.5 μg/m^3

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

Performance Attributes/DQIs	Decision Support
	RMS: 15%
Response Time	None
Selectivity	Temperature influence: zero temperature dependence under 2.0 $\mu g/m^3$, < 5.0% change in min and max temperature conditions Voltage influence: ≤ 5% change in min and max voltage conditions Humidity influence: ≤ 2.0 $\mu g/m^3$ in zero air

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

Table 4

Desirable performance features of wearable PM sensors established as part of the Los Angeles PRISMS Center BREATHE informatics platform and pediatric asthma panel study (Habre, 2018).

Parameter	Selection Criteria
Accuracy and precision	As close as possible to equivalent FRM/FEM±15% or less
Interferences	Minimal
Data collection, storage, and retrieval	Internal storage, wireless, secure and real-time communication
Energy consumption	Minimal: Battery life ~8–12 h and/or simple charging requirements
Participant burden	Low: Low weight, low noise, unobtrusive form factor, “wearable”, flexible wear options
Durability, known performance	Consistent and proven performance, across microenvironments and mobility levels, low drift over time

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

Table 5

Suggested precision and accuracy targets for low-cost sensors to benefit a wide range of application scenarios (Schauer, 2018).

Application	Precision	Accuracy
Comparison to Standards	± 10%	± 10%
Scaling Filter Based Measurements	± 50%	± 50%
Spatial Gradients	± 10%	± 25%
Microenvironmental Monitoring	± 25%	± 25%
Meteorological Drives	± 10%	± 25%
Source Tracking	± 50%	± 50%
Intervention and Control Measures	± 25%	± 25%

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

Table 6

Summary of $PM_{2.5}$ sensor performance attributes from the subject matter expert discussion on Day 3.

Technology Attribute	Minimum Acceptable Value/Range (count)[a]	Estimated Minimum Acceptable Value/Range
Accuracy	10% (2)	Range: 10%–100%
	15% (2) 20% (1) 30% (1) 50% (1) 20–50% (1) 20–30% (2) Factor of 2, 100% (1)	Median: 25%
Bias	2.5% (1)	Range: 1 $\mu g/m^3$ – 5 $\mu g/m^3$
	10% (2)	or 2.5%–50%
	20% (1) 35% (1) 50% (1) 1 $\mu g/m^3$ (1) 3 $\mu g/m^3$ (1) 5 $\mu g/m^3$ (1)	Median: 2 $\mu g/m^3$ or 15%
Correlation	r = 0.84 (1)	Range: r = 0.84–0.95
	r = 0.87 (1) r = 0.89 (2) r = 0.95 (1)	Median: r = 0.89
Detection Limit	1 $\mu g/m^3$ (1)	Range: 2–4 $\mu g/m^3$
	2 $\mu g/m^3$ (1) 4 $\mu g/m^3$ (1)	Median: 2 $\mu g/m^3$
Precision	10% (1)	Range: 10%–50%
	20% (2) 25% (1) 30% (1) 50% (1)	Median: 23%

[a] Numbers (X) represent the count of SEs who suggested each metric.

(Note: The listed sensor performance values should not be considered suggestive about any specific US EPA recommended means of establishing such a value).

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

Table 7

Summary of O_3 sensor performance attributes from the subject matter expert discussion on Day 3. (Note: The listed sensor performance values should not be considered suggestive about any specific US EPA recommended means of establishing such a value).

Technology Attribute	Minimum Acceptable Value/Range (count)[a]	Estimated Minimum Acceptable Value/Range
Accuracy	7% (1)	Range: 7%–20%
	10% (2) 15% (2) 20% (1)	Median: 13%
Bias	5% (1)	Range: 5%–25%
	10% (1)	or 2 ppb-5 ppb
	25% (1) 10–20% (1) 2 ppb (1) 5 ppb (1) 3–5 ppb (1)	Median: 12.5% or 4 ppb
Correlation	r = 0.74 (1)	Range: r = 0.74–0.95
	r = 0.89 (1) r = 0.95 (1)	Median: r = 0.86
Detection Limit	2 ppb (1)	Range: 2ppb-10 ppb
	10 ppb(1) 2–5 ppb (1)	Median: 5 ppb
Precision	7% (1)	Range: 7%–20% or 5 ppb-200 ppb
	15% (2) 20% (2) 5 ppb (1) 5–10 ppb (1) 30–200 ppb (1)	Median: 15% or 102 ppb
Range	0–180 ppb (1)	Range: 0 ppb-200 ppb
	0–200 ppb (1) 10–200 ppb (1) 200 ppb (1)	Median: 100 ppb

[a]Numbers (X) represent the count of SEs who suggested each metric.

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

Table 8

Breakdown of an example tiering system for sensor performance targets (Allen, 2018).

Tier	Performance targets
0 – Just Don't Use It	R^2 < 0.25 or RMSD > 100%
1 – Qualitative	R^2 0.25 to 0.50, RMSD < 100%
2 – Semi-Quantitative	R^2 0.50 to 0.75, RMSD < 50%, bias < 50%
3 – Reasonably Quantitative	R^2 0.75 to 0.90, RMSD < 20%, bias < 30%
4 – Almost Regulatory	R^2 > 0.90, RMSD < 10%, bias < 15%