Taylor & Francis
Taylor & Francis Group

Journal of the Air & Waste Management Association

ISSN: 1096-2247 (Print) 2162-2906 (Online) Journal homepage: www.tandfonline.com/journals/uawm20

Performance evaluation of commercially available non-regulatory instruments and sensors in smoke for PM, CO, CO_2, NO_2, and SO_2

Matthew S. Landis, Russell W. Long, Jonathan D. Krug, Maribel Colón, Jenny Perth & Shawn P. Urbanski

To cite this article: Matthew S. Landis, Russell W. Long, Jonathan D. Krug, Maribel Colón, Jenny Perth & Shawn P. Urbanski (12 Mar 2026): Performance evaluation of commercially available non-regulatory instruments and sensors in smoke for PM, CO, CO_2, NO_2, and SO_2, Journal of the Air & Waste Management Association, DOI: 10.1080/10962247.2026.2629562

To link to this article: https://doi.org/10.1080/10962247.2026.2629562

JOURNAL OF THE AIR & WASTE MANAGEMENT ASSOCIATION
https://doi.org/10.1080/10962247.2026.2629562

Taylor & Francis
Taylor & Francis Group

TECHNICAL PAPER

Performance evaluation of commercially available non-regulatory instruments and sensors in smoke for PM, CO, CO_2, NO_2, and SO_2

Matthew S. Landis[a], Russell W. Long[a], Jonathan D. Krug[a], Maribel Colón[a], Jenny Perth[b], and Shawn P. Urbanski[c]

[a]Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, NC, USA; [b]National Technology and Development Program, U.S. Forest Service, Missoula, MT, USA; [c]Rocky Mountain Research Station, U.S. Forest Service, Missoula, MT, USA

ABSTRACT

Wildland fire smoke can severely degrade downwind air quality and impact responding firefighters and the public. Accurate and timely monitoring of particulate matter (PM) and gaseous pollutants around wildland fires are needed to support management activities like protecting responders and the public from the health and safety hazards of wildland fire smoke. This study systematically evaluates the performance of commercially available, non-regulatory PM and gas instruments/sensors under controlled wildland fire smoke conditions. We assessed accuracy, precision, and linearity by comparing instrument/sensor outputs to reference-grade samplers/instruments across a wide range of smoke concentrations. Results indicate that most non-regulatory PM instruments/sensors, particularly those utilizing optical particle sensing technology, often exhibit a high positive raw $PM_{2.5}$ measurement bias. However, applying smoke-specific calibration models substantially improved accuracy, with several models achieving post-calibration accuracy greater than 85%. Carbon monoxide (CO) and carbon dioxide (CO_2) gas instruments/sensors generally performed well after calibration, while the performance of electrochemical nitrogen dioxide (NO_2) and sulfur dioxide (SO_2) sensors was relatively poor. Non-regulatory $PM_{2.5}$ instrument/sensor response was primarily influenced by aerosol optical properties, particle size distribution, and effective density. We found that smoke-specific calibration and data quality assurance are key for reliable non-regulatory measurements during wildland fire events, especially for incident responder force protection and nearby communities. Integrating calibrated non-regulatory measurement data with regulatory networks can improve smoke exposure assessment and public health messaging. While non-regulatory instruments/sensors cannot fully replace reference instruments, their proper selection and calibration can provide valuable, actionable data for air quality management during wildfires.

Implications: The results presented in this paper provide federal, state, local, and tribal air monitoring and public health agencies with performance information on numerous commercially available non-regulatory instruments/sensors to address air quality monitoring data collection and interpretation challenges presented by wildland fire smoke. The presented accuracy, precision, and linearity performance metrics provide knowledge for the selection of new air quality monitors and insight on the deployment/use of existing monitors during wildland fire events to improve management oversight for professional practitioners/responders, and for communication of health messaging to the public.

PAPER HISTORY
Received August 11, 2025
Revised February 2, 2026
Accepted February 3, 2026

Introduction

In the western United States and Canada, wildfire burned area, fire severity, and smoke emissions have increased over the past 20 years (Westerling 2016; Holden et al. 2018; Parks and Abatzoglou 2020). Studies suggest the current fire activity levels will persist and may even increase in the coming decades due to a warming and drying climate (Kitzberger et al. 2007; Littell et al. 2009; Abatzoglou and Williams 2016; Westerling 2016). The smoke produced by these wildfires has resulted in severe air pollution across large areas—from regions proximate to fires to urban centers thousands of kilometers downwind (Mallia et al. 2015; Landis et al. 2018; Alonso-Blanco et al. 2018; Wu et al. 2018).

Smoke from wildland fires (wildfires and managed prescribed fires) contains particulate matter (PM) and hundreds of gaseous compounds, including many recognized as hazardous air pollutants (Urbanski et al. 2022). Fine particulate matter ($PM_{2.5}$; particles with a mass median aerodynamic diameter of less than 2.5 μm) and carbon monoxide (CO) are two primary pollutants of concern in wildfire smoke. Long-term

CONTACT Matthew S. Landis landis.matthew@epa.gov Office of Research and Development, U.S. Environmental Protection Agency, 109 T.W. Alexander Drive, Research Triangle Park, NC 27711, USA.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/10962247.2026.2629562

chronic exposure to $PM_{2.5}$ is associated with a wide range of negative health effects and is the basis for its regulation under the U.S. Clean Air Act (EPA 2019; Andersen et al. 2021; McCaffrey et al. 2022). Furthermore, evidence shows that short-term acute wildfire smoke exposure is associated with negative health impacts (Adetona et al. 2016; Reid et al. 2016; Jiao et al. 2024), and exposures to CO can impair oxygen transport, causing hypoxia and damage to organs and tissues (Omaye 2002). Wildland fire smoke can also directly threaten public safety by reducing visibility on roadways, contributing to injuries and fatalities from traffic accidents (Ashley et al. 2015; Bartolome et al. 2019), including massive multi-vehicle pileups in the United States (Associated Press 2022; Reuters 2023).

Real-time measurements of surface-level $PM_{2.5}$, CO, and other criteria pollutants are needed to protect the public from the health and safety hazards of wildland fire smoke. Critical activities include air quality and smoke forecasting, issuing public guidance for limiting smoke exposure, managing prescribed fires to mitigate smoke impacts, monitoring the exposure of wildland firefighters to $PM_{2.5}$ and CO, estimating population-level exposure to smoke in retrospective health studies, and developing and validating smoke modeling systems.

Health officials and air management agencies must provide timely guidance to the public on limiting smoke exposure. Guidance is based on observations and forecasts of $PM_{2.5}$ using the AirNow Fire and Smoke Map (https://fire.airnow.gov/), BlueSky smoke model (https://tools.airfire.org/websky/v2/#status), or various National Oceanic and Atmospheric Administration (NOAA) products, such as Air Quality Forecast Guidance Viewer (https://airquality.weather.gov/), Hazard Mapping System Fire and Smoke Product (https://www.ospo.noaa.gov/Products/land/hms.html#maps), and the HRRR-NCEP-smoke forecast (https://rapidrefresh.noaa.gov/hrrr/HRRRsmoke/). Wildfire smoke can be an extended and widespread threat to public health due to the relatively large acreages consumed, duration, and quantity of smoke emitted (Brey et al. 2018; Wilmot et al. 2022). However, prescribed fires, while generally small and short in duration, can have substantial impacts on nearby communities. Therefore, smoke monitoring is important for communicating with the public, and for managing prescribed burns to mitigate short-term acute and aggregate smoke exposures and impacts.

Wildland firefighters are subject to unique risks from wildland fire smoke. They can be exposed over long durations (often 16-hour shifts) to high levels of $PM_{2.5}$, CO, and other air pollutants while suppressing wildfires or conducting prescribed burns (Navarro 2020). Limited research reports firefighters' work-shift exposures at mean $PM_{2.5}$ concentrations as high as 1.2 mg m^{-3} and peak concentrations exceeding 10 mg m^{-3} (Navarro 2020). More data are available for firefighter CO exposures as a function of task, with one study reporting sawer/swamper (chain saw operator/brush clearing) having the highest geometric mean work shift exposures 6.8 ppm and maximum shift time weighted average (TWA) concentration of 64.5 ppm (Henn et al. 2019), and ~5% of firefighters exceeding the National Wildfire Coordinating Group 13-hr shift average concentration of 16 ppm (Semmens et al. 2021). Higher-resolution (1-min) data firefighter CO exposures have been reported to be as high as 339 ppm, with 17% of monitored participants exceeding the Occupational Safety and Health Administration (OSHA) ceiling limit of 200 ppm (Materna et al. 1992), and another study reporting 4% of monitored work shifts exceeding the American Conference of Governmental Industrial Hygienists (ACGIH) "not to be exceeded at any time" threshold limit level of 400 ppm for any 15-min period (Henn et al. 2019).

The potential health and safety hazard due to wildland fire smoke can vary spatially and temporally. Emission rate, plume behavior, terrain, and atmospheric conditions determine smoke transport and dispersion in the atmosphere and influence downwind surface air pollutant concentrations. Wildfire smoke can be widespread, impacting entire states, provinces, and regions thousands of kilometers downwind. Wildland fire smoke impacts can also be localized, heterogenous, and rapidly change. Regulatory air monitoring sites in the U.S. are limited in number, concentrated in urban areas, and often do not adequately characterize wildland fire smoke impacts. In the 11 western U.S. states, there are only 249 regulatory monitoring sites providing hourly observations of $PM_{2.5}$ and only 92 sites reporting CO. Most counties (~70%) in the United States and their rural smoke-susceptible communities have no regulatory monitors (U.S. GAO 2020).

Federal, state, local, and tribal agencies have initiated various efforts that provide non-regulatory monitoring of smoke impacts during the wildfire season and prescribed burn operations (Landis et al. 2021b). These efforts initially focused on professional-grade equipment, such as Met One Instruments (Grants Pass, OR, U.S.A.), Model E-BAM (^{14}C beta attenuation), and E-Sampler (optical) for $PM_{2.5}$, which require technical expertise to set up, calibrate, and run, with costs ranging from ~$15,000–$25,000 USD per instrument. Lower-cost PM sensors (<$500 USD) that require less technical expertise to set up and operate are widely available now, and are being used by the public, government agencies, and researchers. The PM sensors produced by PurpleAir

(Draper, Utah, U.S.A.) are a widely used and evaluated lower-cost PM monitor in the U.S. (Barkjohn K, Holder A, et al. 2021; Barkjohn KK, Gantt B, et al. 2021; Barkjohn et al. 2022; Malm et al. 2024). The interactive PurpleAir Map (https://www2.purpleair.com/) displays measurements from thousands of publicly visible PurpleAir sensors deployed across the U.S. and Canada. PurpleAir sensors use Plantower (Nanchang, Jiangxi, China) laser optical particle sensors (OPS) to derive PM mass concentrations. Other original equipment manufacturer (OEM) OPS manufactured by MetOne (Grants Pass, OR, U.S.A.), Plantower, Alphasense (Essex, UK), Senserion AG (Stäfa, Switzerland), ThermoFisher (Waltham, MA, U.S.A.), TSI (Shoreview, MN, U.S.A.), and Vaisala (Vantaa, Finland) are used in a variety of commercially available PM instruments/sensors (Zheng et al. 2018; Landis et al. 2021a; Barkjohn et al. 2024).

PurpleAir and other lower-cost sensors employ light-scattering optical methods for continuous (1-to-2-min) data acquisition rates. These optical methods convert measured light scattering into $PM_{2.5}$ mass concentrations using proprietary algorithms (Mehadi et al. 2020; He et al. 2020). The OPS light-scattering response depends on particle composition and size distribution (Hagan and Kroll 2020; He et al. 2020; Rueda et al. 2023). Lower-cost monitors require correction because they generally overestimate ambient $PM_{2.5}$ concentrations compared with federal reference method (FRM) and federal equivalent method (FEM) instruments (Barkjohn K, Holder A, et al. 2021; Barkjohn KK, Gantt B, et al. 2021; Liang 2021; Wallace et al. 2021, 2022; Barkjohn et al. 2022, 2024).

Evaluations show that PurpleAir sensors respond differently in smoke compared with ambient air (Holder et al. 2020; Wallace et al. 2021; Barkjohn et al. 2022). This finding is not surprising, given the differences in the composition and size distribution of particles produced by wildland fires compared with other major sources, such as fossil-fuel combustion, secondary aerosol formation, and fugitive dust from unpaved roads, agriculture, and construction activities. These studies have found that PurpleAir sensors respond linearly but require a polynomial calibration for $PM_{2.5}$ concentrations > ~300 $\mu g\ m^{-3}$ (Holder et al. 2020; Barkjohn et al. 2022). Lower-cost PM sensors are being used by government agencies and researchers to monitor and study the impacts of wildland fire smoke on ambient air quality. The AirNow Fire and Smoke Map, which was developed by the U.S. EPA and the U.S. Forest Service, provides the public with information on fire locations, smoke plumes, and near-real-time NowCast Air Quality Index (AQI) based on surface measurements (fire.airnow.gov). The Fire and Smoke map includes $PM_{2.5}$ data from nearly 15,000 PurpleAir monitors across the U.S. and Canada, currently corrected based on calibration equations detailed in K. Barkjohn, A. Holder, et al. (2021).

Several studies have combined publicly available non-regulatory (e.g., PurpleAir, E-Sampler, E-BAM) monitoring data with data from regulatory air monitoring networks (U.S. EPA Air Quality System; AQS) and/or satellite aerosol optical depth (AOD) observations to assess wildfire smoke (Gupta et al. 2018; Liang 2021; Vu et al. 2022; Yang et al. 2023; Carreras-Sospedra et al. 2024) and prescribed fire smoke (Sablan et al. 2024) impacts on air quality. The fusion of archived $PM_{2.5}$ observations from regulatory and non-regulatory instruments/sensors with satellite measured AOD has proven valuable for generating continuous surface interpolated $PM_{2.5}$ concentrations that can be used to evaluate smoke forecasting models (Chow et al. 2022). While lower-cost PM sensor data are now widely applied for the diagnosing and short-term forecasting of smoke impacts on air quality, estimating smoke exposure in health studies, and evaluating the performance of smoke forecasting systems, questions remain about the performance of this class of instruments in heavily smoke impacted environments.

Wildland firefighters on the fire line, in fire camps, and in communities or roadways near fires may experience $PM_{2.5}$ concentrations > 1000 $\mu g\ m^{-3}$. However, the performance of non-regulatory instrument/lower-cost sensors at these concentrations are uncertain, as previous evaluation efforts have focused primarily on environments with $PM_{2.5} < 1000\ \mu g\ m^{-3}$. Additionally, many companies market lower-cost PM and gas monitors that utilize several different OEM OPS, non-dispersive infrared (NDIR), and electrochemical sensors; however, most evaluation efforts have focused on the PurpleAir monitors, which use Plantower Model PMS5003 OPS. There is evidence that different low-cost PM monitors using the same Plantower OPS perform differently in smoke-impacted environments based on unique physical implantation and raw signal processing/calibration (Landis et al. 2021a). It is unknown whether the calibrations developed for PurpleAir sensors are applicable to other sensors. Finally, changes in the performance of OPS due to exposures to wildland fire smoke have not been studied. Thus, it is unknown how rapidly high concentrations of organic aerosol-rich PM in smoke may alter the response of these lower-cost monitors. These questions must be addressed so the public, health officials, air managers, air smoke forecasters, and researchers can confidently use data from low-cost PM monitors across the range of conditions experienced in smoke impacted environments.

This study expands on previous work by characterizing the ability of non-regulatory continuously reporting instruments/lower-cost sensors to meet PM and gaseous measurement needs at concentration ranges representative of heavily impacted wildland fire smoke environments. The primary study goals were to (1) test the performance (accuracy, precision, linearity) of non-regulatory instruments/sensors under controlled conditions at varying concentrations of biomass smoke concentrations typically observed near wildland fires; (2) create instrument/sensor-specific smoke calibrations and compare to previously published calibrations (if any) determined from ambient data sources; and (3) investigate primary factors leading to inaccuracies of raw $PM_{2.5}$ performance metrics (e.g., PM concentration, particle size distribution, black carbon concentration, effective aerosol density, previous use in smoke).

Methods

FSL chamber testing

The U.S. Forest Service (USFS) Rocky Mountain Research Station Fire Sciences Laboratory (FSL) main combustion chamber is 12.4 m × 12.4 m × 19.6 m, with a total volume of 3000 m^3, and was described previously (Christian et al. 2004). Our "static chamber" burns simulate instrument/sensor exposure under field sampling conditions as described previously (Krug et al. 2021; Landis et al. 2021a; Whitehill et al. 2022; Long et al. 2023). Briefly, the instruments/sensors were placed in the combustion chamber; prior to each burn, the chamber was flushed with outdoor air, fuel beds were prepared and placed in the center of chamber and ignited, the chamber was sealed, and 10–15 mins were allotted for the consumption of the fuel bed and homogenization of the resulting smoke in the chamber prior to the start of the 1-hr test period (Appendix Figure A.1a-d). After the test period, the chamber was ventilated with outdoor air in preparation for the next test burn. The use of outdoor air limited the range of temperature and relative humidity (RH) during testing. Chamber temperatures ranged from 19–24°C in 2019, and 18–21°C in 2021, and relative humidity ranged from 21–53% in 2019 and 14–33% in 2021 (Appendix Tables B.1 & B.2).

A time series plot of 1-min $PM_{2.5}$ and CO concentrations throughout a typical burn day (April 19, 2021) is presented in Appendix Figure A.2, showing concentrations through the progression of burns 20–23 and highlighting fuel bed ignitions (red), start and stop times of official 1-hr test periods (vertical drop lines), and chamber ventilation (blue). The figure also demonstrates the dilution of measured concentrations over the course of each test due to outdoor air infiltration and sampler removal, which facilitates the evaluation of continuous measurements for response lag that can affect overall performance metrics. The fuels utilized for the studies were ponderosa pine (Pinus ponderosa) needles (PPN) and fine dead wood (PPW), and Douglas-fir (Pseudotsuga menziesii) boughs and cones, alone or mixed. Combustion efficiency under different combustion conditions (e.g., smoldering, flaming) of burns was managed by fuel moisture content, as summarized in Appendix Tables B.1 and B.2.

Reference measurements

Burn integrated (1-hr) FRM $PM_{2.5}$ mass measurements were performed in triplicate inside the burn chamber, according to 40 CFR Part 50 Appendix L (U.S. EPA) using Tisch Environmental (Cleves, OH, U.S.A.) Model TE-WILBUR2.5 filter-based samplers (Hall et al. 2012; EPA 2014a - RFPS-1014–219; Krug et al. 2021). During the 2021 study, a Model TE-WILBUR10 FRM sampler configured to sample PM_{10} was included (EPA 2014b - RFPS-0714–216). Each Tisch FRM sampler used a standard EPA PM_{10} louvered inlet, and $PM_{2.5}$ fractionation was performed by a BGI by Mesa Labs (Butler, NJ, U.S.A.) Model VSCCA very sharp cut cyclone (VSCC). Calibrations of sampler temperature, pressure, and flow rates were performed weekly using a certified BGI by Mesa Labs tetraCal® Air Flow Calibrator primary standard, as outlined in each respective sampler user manual. Measurement Technology Laboratories (MTL; Minneapolis, MN, U.S.A.) Model PT47P polytetrafluoroethylene (PTFE) 2 µm pore size pressure-drop equivalent membrane filters, with a hydro-inert support ring were used. Following ventilation of the chamber, sample filters were immediately retrieved, put into petri dishes, and frozen for shipment to the EPA weighing lab.

Filters were weighed by EPA in accordance with EPA Guidance Document 2.12 (EPA, 2016) for monitoring $PM_{2.5}$ in ambient air at the U.S. EPA National Reference Weighing Facility in Research Triangle Park, NC. Before and after sampling, filters were equilibrated for a minimum of 24 hr in a temperature- and humidity-controlled clean room prior to being weighed by an MTL Model AH1 automated robotic weighing system. The AH1 used a Mettler-Toledo (Columbus, OH, U.S.A.) Model XP2U micro balance with an accuracy of up to a tenth of a microgram. The response of the micro balance was checked every 10 filters by verification of Class 0 weight standards from Rice Lake Weighing Systems (Rice Lake, WI, U.S.A.). Filter

weights were performed in triplicate and were accepted as valid if the responses to the weight standards were ± 5 µg. Final weights were calculated as the mean of the triplicate readings.

Continuous (1-min) "reference" $PM_{2.5}$ concentrations were calculated for the purpose of visually evaluating instrument/sensor response time and stability during each burn. Continuous reference $PM_{2.5}$ concentrations were calculated by multiplying raw TSI DustTrak measurements by an integrated burn-specific correction factor defined as the quotient of FRM $PM_{2.5}$/DustTrak $PM_{2.5}$. The correction factor was calculated using data from each official 1-hr burn period but applied from fuel bed ignition through chamber ventilation separately for each burn. The DustTrak Model DRX8533 was factory serviced and calibrated prior to each study, the aerosol density was set to unity, and zero calibrations were conducted before each burn day. For the first nine 2021 burns when we did not have DustTrak data, "reference" $PM_{2.5}$ concentrations were calculated using FRM adjusted 2B Technologies (Broomfield, CO, U.S.A.) Photometer data. The DustTrak and Photometer both provided fast and linear response relative to the burn-integrated FRM concentrations over a wide dynamic range (see Raw and smoke calibrated PM performance), making these smoke-calibrated instruments a logical choice for which to visualize the performance of other instruments in smoke.

Gaseous reference instruments utilized in this testing included (1) Picarro Inc. (Santa Clara, CA, U.S.A.) Model G2401-m cavity ring-down spectroscopy (CRDS) CO/CO_2 gas analyzer, (2) ThermoScientific (Franklin, MA, U.S.A.) Model 43C pulsed fluorescence FEM sulfur dioxide (SO_2) analyzer (EPA 1986 - EQSA-0486–060,), and (3) Teledyne API (San Diego, CA, U.S.A.) Model T500U cavity attenuated phase shift (CAPS) FEM direct spectroscopic NO_2 instrument (EPA 2014c - EQNA-0514–212). Continuous-sampling gas reference instruments were placed in the observation room adjacent to the combustion chamber and connected to perfluoroalkoxy (PFA) Teflon® sampling manifolds that brought chamber air from the instrument/sensor testing zone (Appendix Figure A.1c). Except for the CRDS, all continuous gas analyzers were zeroed and span calibrated at the beginning and end of each chamber test day using Teledyne API Model T700U dynamic dilution calibration systems with certified ozone (O_3) photometers. EPA protocol certified gas standard cylinders diluted in ultra-scientific grade-zero air were used for CO, CO_2, nitrogen dioxide (NO_2), and SO_2 instruments. Gas phase titration (GPT) calibrations were utilized to calibrate the CAPS NO_2 instrument from the EPA nitric oxide (NO) protocol gas. Multi-point span calibrations were conducted at the beginning and end of each study week to ensure linearity. The CRDS response was stable over each set of burns (12 days in 2019 and 13 days in 2021), as confirmed with three-point calibrations using gas mixtures of CO_2 and CO in scientific-grade zero air prior to, midway through, and upon completion of each round of chamber burns (2019 and 2021). Gaseous reference instrument sampling and calibration procedures are detailed further in Long et al. (2021).

Aerosol size distributions were monitored during the 2021 study using a TSI Incorporated (Shoreview, MN, U.S.A.) Model 3321 Aerodynamic Particle Sizer (APS) and a TSI Model 3340 laser aerosol spectrometer (LAS). The APS and LAS were placed on a table within the designated smoke sampling area (see Appendix Figure A.1a). The TSI instruments were factory cleaned and calibrated prior to the study.

Continuous optical black carbon (BC) measurements were made with Magee Scientific (Berkley, CA, U.S.A.) Model AE33 seven wavelength spectrum Aethalometer equipped with a Model 5610 sample stream dryer and a BGI (Waltham, MA, U.S.A.) Model SCC-1.829 sharp cut $PM_{2.5}$ cyclone. The Magee AE33 measurement principle and real-time loading effect compensation algorithm based on two parallel spot measurements have been described previously by Drinovec et al. (2015). All reported BC concentrations represent the self-adsorption compensated BC6 (880 nm) channel. Performance of the AE33 light-emitting diode (LED) sources was verified prior to and after each study using the Magee Scientific Model 7662 neutral density optical filter validation kit. Results obtained from the AE33 are indicative of the optical properties of the smoke aerosols and, for the purposes of this paper, were averaged to the 1-hr filter-based FRM sample periods. Delta-C or ΔC (Allen et al. 2004) is an objective measure of the wavelength (λ) dependence of aerosol optical absorption between the AE33 self-adsorption compensated ultraviolet (UV; 370 nm) and infrared (IR; 880 nm) channels and was calculated using eq 1.

$$\Delta C = BC1(370\ nm) - BC6(880\ nm) \quad (1)$$

Instruments/sensors tested

The 17 PM instruments and sensor systems listed in Table 1 were tested *versus* FRM reference samplers in the FSL combustion chamber in burn-integrated mean $PM_{2.5}$ concentrations ranging from 11.2–3051.7 µg m^{-3}. The selected instruments/sensors were either pulled from existing EPA/USFS inventories (Used), purchased based on their current use for smoke applications (New), or

Table 1. $PM_{2.5}$ and PM_{10} instrument/sensors tested.

System manufacturer	Model	PM detector OEM	OEM model	Detection method	Status
2B	PAM	Plantower	PMS5003	Laser scattering	New & Used
2B	Black Carbon Photometer	2B	—	Light extinction	New
AmbiLabs	2WIN	AmbiLabs	Nephelometer	Light scattering	Demo
AQMesh	AQM-PM6	AQMesh	OPC	Laser sizing/scattering	Demo & New
Duke‡	—	Plantower	PMS3003	Laser scattering	Used
FTS	AQT400-GOES-PTBL	Vaisala	AQT420	Laser particle counter	New
Kunak	A10	Alphasense	OPC-R2	Laser sizing/scattering	Used$^{\delta}$
Kunak	Air Pro	Alphasense	OPC-N3	Laser sizing/scattering	New
MetOne	E-BAM	MetOne	Beta gauge	^{14}C Beta attenuation	Used
MetOne	E-Sampler	MetOne	Near-forward Light Scattering	Laser scattering	Used
PurpleAir	PA-II-SD	Plantower	PMS5003	Laser scattering	New & Used
Senserion	SPS30	Senserion	SPS30	Laser scattering	New
Thermo	pDR-1500	ThermoFisher	Laser Photometer	Light scattering	Used
Thingy	AQ	Senserion	SPS30	Laser scattering	New
TSI	APS 3321	TSI	Time-of Flight Dual Laser	Laser spectrometer	Used
TSI	DustTrak DRX 8533	TSI	Laser Photometer	Laser scattering	New & Used
Vaisala	AQT420	Vaisala	AQT420	Laser particle counter	New

Notes. ‡Zheng et al. (2018); $^{\delta}$Kunak A10 systems were previously evaluated (Landis et al. 2021b), and the OPC was replaced in system 1 prior to 2021 testing.

loaned from manufacturers wanting to evaluate their performance under typical wildland fire conditions (Demo). As detailed in Table 1, the methodologies utilized by the tested instruments/sensors for PM detection were laser or light scattering, laser spectroscopy, or ^{14}C beta attenuation systems.

The 2B (Broomfield, CO, U.S.A.), AQMesh (Stratford, UK), Duke (Durham, NC, U.S.A.), Kunak (Navarra, Spain), and Vaisala (Vantaa, Finland) multi-pollutant sensor systems also included one or more electrochemical (CO, CO_2, NO_2, SO_2) and/or NDIR gas sensors (CO_2) that were evaluated versus reference instruments along with Licor and Vaisala (NDIR) and Senserion single gas electrochemical sensors. Ranges of burn-integrated mean pollutant concentrations are summarized in Appendix Table B.3 for the 2019 burns and Appendix Table B.4 for the 2021 burns.

Statistical analysis

Data processing and all statistical analyses were performed using SAS v.9.4 (SAS Institute, Cary, NC, U.S.A.). Accuracy of nonregulatory instruments/sensors for each study were calculated using eq 2, and precision was calculated using the coefficient of variation or relative standard deviation using eq 3, where collocated instruments/sensors were tested.

$$\text{Accuracy}\,(\%) = 100 - \left[\frac{|\bar{X} - \bar{R}|}{\bar{R}}\right] \times 100 \qquad (2)$$

Where X is the reported instrument/sensor concentration and R is the reference concentration.

$$\text{Precision}\,(\%) = \frac{\sqrt{\sum \frac{(x_i - \bar{x})^2}{(n-1)}}}{\bar{x}} \times 100 \qquad (3)$$

Where $\bar{x}$ = mean of collocated sensor concentrations, $\sum (x_i - \bar{x})^2$ = the sum of the square of differences between individual collocated sensor concentrations and the mean, and n = the number of collocated sensor observations. All instruments and sensor data were temporally synchronized prior to calculating accuracy and precision.

Parametric statistics used in this analysis include polynomial least squares regression, Pearson product-moment correlation, and multilinear regression (MLR) analysis. The assumptions of the parametric procedures were examined using residual plots, skewness and kurtosis coefficients, the Shapiro-Wilk test, and the Brown-Forsythe test. Nonparametric statistics used in this analysis include Spearman's rank-order and Kendall's Tau-b correlation. A level of significance of $\alpha = 0.05$ was used for all statistical procedures unless otherwise stated. The SAS UNIVARIATE, CORR, and GLM procedures were used for testing data distributions for normality, equal variance, and hypothesis testing $Mu^0 = 0$ (population mean is not significantly different from zero); developing scatter plot matrices and calculating correlation coefficients; and conducting least square general polynomial model regressions and MLR analysis, respectively.

A separate MLR analysis model was run for each $PM_{2.5}$ instrument/sensor tested in 2021 to elucidate environmental conditions and aerosol physical/optical properties impacting its burn-integrated $PM_{2.5}$ concentration versus the FRM samplers. The MLR dependent variable for each model is the burn-integrated instrument/sensor $PM_{2.5}$ analytical artifact (both positive and negative), termed Δ_{FRM} and calculated using eq 4 as the difference between the mean FRM and each individual tested candidate instrument/sensor $PM_{2.5}$ mass concentration.

$$\Delta_{FRM} = \overline{FRM\ PM_{2.5}} - \text{instrument/sensor}\ PM_{2.5} \quad (4)$$

The MLR independent variables include temperature, relative humidity (RH), modified combustion efficiency (MCE; Urbanski 2013), BC, mass median aerodynamic diameter (MMAD), effective density, ΔC, and FRM $PM_{2.5}$ mass concentration. All significant independent variable interactions determined using parametric and non-parametric correlation methods were included in each model. Only MLR results that met the model residual normality assumption (Shapiro-Wilk test), with a mean not significantly different from zero, and constant variance are reported. Since the goal of the MLR analysis was to identify factors leading to $PM_{2.5}$ measurement artifacts, no model optimization attempts were made to remove non-significant independent variables and interactions.

Smoke-specific calibrations

Following the evaluation of raw instrument/sensor accuracy, precision, and linearity, study-specific integrated calibration models were developed and evaluated, including linear, linear piecewise, quadradic, and cubic (including those with intercepts and those forced through zero). Optimum calibration models were selected based on improved performance metrics and statistical significance of model parameters, and are provided in Appendix Table B.5 and Appendix Table B.6 for 2019 and 2021 studies, respectively.

Effective aerosol density estimates

Since many non-regulatory PM instruments and lower-cost sensors measure aerosol light scattering and estimate PM mass concentrations using proprietary algorithms that typically include parametrized aerosol density, we calculated an effective $PM_{2.5}$ aerosol density (ρ_{eff}, g cm^{-3}) for each FSL chamber burn in 2021 using a methodology adapted from DeCarlo et al. (2012) and presented in eq 5.

$$\rho_{eff} = \frac{m_p}{V_p} \quad (5)$$

Where m_p is the aggregate FRM $PM_{2.5}$ mass concentration and V_p is the aggregate partical volume $PM_{2.5}$ concentration, as exported from the LAS (aerodynamic diameter). The LAS particle diameters from the full $PM_{2.5}$ range (0.090 to 2.5 µm aerodynamic diameter) were used (Appendix Table B.4). Because the LAS has a low coincident particle counting threshold, the sample stream was diluted with a TSI Model 3302A Aerosol Diluter, possibly resulting in additional inlet losses and a subsequent overestimation of density. Regardless, these effective density estimations are comparable within the study period and are effective at showing relative shifts in effective density. Aerosols are assumed to be spherical, with no internal voids such that electrical mobility and aerodynamic diameters are equivalent. It should be noted that there are various definitions of effective density found in the literature that may result in numerically different results. Comparison of effective densities from studies using differing methodologies should be avoided.

Results and discussion

Raw and smoke calibrated PM performance

We tested over 23 days totaling 71 static chamber burns with 31 burns completed in 2019 and 40 burns completed in 2021. The results demonstrate smoke testing over a range of 11.2–3051.7 µg m^{-3} ($PM_{2.5}$) and 14.0–1535.8 µg m^{-3} (PM_{10}; 2021 only). The burn-integrated MCEs ranged from 0.85–0.97, representing both flaming and smoldering combustion conditions (Akagi et al. 2011; Urbanski et al. 2022). Smoke homogeneity was verified using the precision of three filter-based $PM_{2.5}$ FRM samplers located within the chamber and triangulating the testing zone for all 71 chamber burns; it was determined to be 2.6 ± 2.5% (mean ± standard deviation), with a median of 1.7%.

The raw $PM_{2.5}$ and PM_{10} testing accuracy and collocated precision results are summarized for all tested instruments/sensors in Tables 2 and 3 for 2019 and 2021 studies, respectively. The 2021 PM_{10} performance metrics were included for completeness, but the chamber smoke $PM_{2.5}$ aerosol size distributions had MMADs ranging from 0.182–0.756 µm (Appendix Table B.4), showing relatively minor contributions of coarse $PM_{10-2.5}$ mass (1.7 ± 3.9%). We observed a wide range of mean $PM_{2.5}$ accuracies from −260 to 87%, collocated precision was generally within ±15%. Study-specific smoke (SSS) calibration equations were developed, evaluated, and applied to all the tested instruments and sensors to potentially improve their performance over wide concentration ranges. All the selected calibration equations are presented in Appendix Table B.5 and Appendix Table B.6 for the 2019 and 2021 studies, respectively. The updated SSS calibrated PM instrument/sensor performance results are summarized in Table 4 (2019) and Table 5 (2021) and demonstrate that most performance metrics significantly improved versus their raw metrics (Tables 2 and 3). As expected, sensors with well fit calibration regression equations produced substantial performance improvements.

Table 2. 2019 burn-integrated raw $PM_{2.5}$, CO, CO_2, and NO_2 testing accuracy (%) and collocated precision (%).

				Accuracy (mean ± standard deviation)				Collocated precision (mean ± standard deviation)			
Vendor	Model	Unit	n‡	$PM_{2.5}$	CO	CO_2	NO_2	$PM_{2.5}$	CO	CO_2	NO_2
2B	PAM	2015	12	76.5 ± 15.4	93.9 ± 5.4	48.6 ± 17.8	–	6.3 ± 3.5	2.7 ± 1.2	46.7 ± 25.8	–
2B	PAM	2016	20	79.0 ± 11.8	94.3 ± 6.0	−133.7 ± 67.8	–				
AmbiLabs	Nephelometer		31	−254.2 ± 70.2	–	–	–	–	–	–	–
AQMesh		#1	17	56.7 ± 13.6	–	–	62.6 ± 21.8	–	–	–	–
Duke		#14	31	79.9 ± 13.6	–	–	–	3.3 ± 2.0	–	–	–
Duke		#23	30	76.6 ± 15.8	–	–	–				
FTS	AQT420		31	67.3 ± 25.6	81.6 ± 5.4	–	42.7 ± 52.3	–	–	–	–
Kunak	A10	1	23	73.2 ± 22.7	87.6 ± 3.3	–	73.9 ± 17.4	1.8 ± 1.3	14.4 ± 1.0	–	14.5 ± 7.8
Kunak	A10	2	26	74.2 ± 21.8	92.4 ± 3.6	–	65.9 ± 17.0				
Licor	850		31	–	–	94.0 ± 1.3	–	–	–	–	–
Met-One	E-BAM	1	26	78.1 ± 11.1	–	–	–	4.7 ± 6.6	–	–	–
Met-One	E-BAM	2	26	82.0 ± 9.0	–	–	–				
Met-One	E-Sampler	1	24	77.8 ± 17.3	–	–	–	5.0 ± 1.9	–	–	–
Met-One	E-Sampler	2	26	72.5 ± 20.3	–	–	–				
PurpleAir	PA-II-SD	#1	30	56.7 ± 25.3	–	–	–	1.6 ± 0.9	–	–	–
PurpleAir	PA-II-SD	#2	30	57.0 ± 24.6	–	–	–				
Thermo	PDR	1	26	73.4 ± 19.5	–	–	–	13.4 ± 1.9	–	–	–
Thermo	PDR	2	20	50.9 ± 27.4	–	–	–				
TSI	APS	APS	29	43.6 ± 18.0	–	–	–	–	–	–	–
TSI	DustTrak	1	31	−26.1 ± 27.9	–	–	–	2.7 ± 1.7	–	–	–
TSI	DustTrak	2	26	−31.2 ± 24.4	–	–	–				

Note. ‡Number of burns with valid data.

Table 3. 2021 burn-integrated raw $PM_{2.5}$ and PM_{10} testing accuracy (%) and collocated precision (%).

				Accuracy (mean ± standard deviation)		Collocated precision (mean ± standard deviation)	
Vendor	Model	Unit	n	$PM_{2.5}$	PM_{10}	$PM_{2.5}$	PM_{10}
2B	PAM	1022	40	75.1 ± 19.0	52.9 ± 27.0	–	–
2B	Photometer		40	83.9 ± 13.3	–	–	–
AQMesh	2,450,480		40	76.2 ± 17.6	77.6 ± 19.2	–	–
Kunak	A10	1	–	–	–	–	–
Kunak	A10	2	40	74.6 ± 20.8	76.1 ± 22.2		
Kunak	Pro	1	40	49.0 ± 16.4	46.8 ± 14.8	9.7 ± 4.1	9.9 ± 4.3
Kunak	Pro	2	40	55.9 ± 18.8	53.8 ± 17.6		
Met-One	E-Sampler	171	37	58.9 ± 19.4	–	0.9 ± 0.5	–
Met-One	E-Sampler	594	37	59.2 ± 19.1	–		–
PurpleAir	PA-II-SD	746D – New	40	35.2 ± 30.6	30.7 ± 29.8	3.2 ± 1.5	5.0 ± 2.6
PurpleAir	PA-II-SD	8171 – New	40	49.1 ± 27.7	44.4 ± 27.4	9.5 ± 0.8	13.7 ± 2.3
PurpleAir	PA-II-SD	BE2E – New	40	50.9 ± 29.7	45.0 ± 28.1	5.2 ± 2.4	4.6 ± 2.6
PurpleAir	PA-II-SD	CBE8 – New	40	49.6 ± 28.0	46.5 ± 27.3	4.2 ± 2.5	5.6 ± 2.5
PurpleAir	PA-II-SD	8BDE – Used	40	42.6 ± 29.9	34.9 ± 29.4	5.8 ± 4.2	5.9 ± 4.0
PurpleAir	PA-II-SD	9349 – Used	40	43.0 ± 28.2	40.7 ± 28.1	3.3 ± 3.1	3.0 ± 2.3
PurpleAir	PA-II-SD	C0BA – Used	40	36.2 ± 30.5	32.0 ± 30.0	6.2 ± 6.2	5.0 ± 5.4
PurpleAir	PA-II-SD	EACF – Used	40	45.5 ± 30.6	41.6 ± 28.4	3.8 ± 3.5	6.5 ± 3.5
PurpleAir	PA-II-SD	8BED – Used	40	57.4 ± 28.5	45.2 ± 27.7	–	–
PurpleAir	PA-II-SD	09DE – Used	40	41.2 ± 32.6	8.3 ± 32.6	11.0 ± 6.0	9.3 ± 7.0
PurpleAir	PA-II-SD	39C4 – Used	40	47.0 ± 31.1	43.4 ± 29.1	–	–
PurpleAir	PA-II-SD	CCC3 – Used	40	51.7 ± 25.9	44.7 ± 26.8	14.6 ± 1.9	13.5 ± 3.0
PurpleAir	PA-II-SD	1CCA – Used	40	48.3 ± 27.1	42.6 ± 26.7	7.5 ± 3.8	4.7 ± 2.4
PurpleAir	PA-II-SD	8E2D – Used	40	40.8 ± 28.2	36.0 ± 28.8	7.0 ± 2.2	6.3 ± 2.5
PurpleAir	PA-II-SD	8F98 – Used	40	34.0 ± 34.3	30.7 ± 32.0	5.8 ± 7.9	5.9 ± 7.1
PurpleAir	PA-II-SD	39C8 – Used	40	35.9 ± 41.2	33.3 ± 38.7	6.8 ± 6.4	5.6 ± 5.3
Senserion	SPS30	Used	39	86.7 ± 11.1	84.6 ± 13.8	–	–
Thingy	AQ	2005		81.2 ± 11.8	83.9 ± 12.5		
Thingy	AQ	2007		86.3 ± 10.0	85.4 ± 12.7		
Thingy	AQ	2008	40	82.6 ± 12.0	80.0 ± 14.3	6.7 ± 1.1	6.7 ± 0.9
Thingy	AQ	2016		85.9 ± 12.0	83.4 ± 14.2		
Thingy	AQ	2019		86.0 ± 12.5	83.0 ± 15.0		
TSI	DustTrak	1	31	−10.7 ± 28.9	9.9 ± 28.3	–	–
TSI	APS		31	37.3 ± 13.7	44.9 ± 15.0	–	–
Vaisala	AQT420		37	5.4 ± 3.3	7.1 ± 4.4	–	–

Table 4. 2019 burn-integrated study-specific smoke calibrated $PM_{2.5}$, CO, CO_2, and NO_2 testing accuracy (%) and collocated precision (%).

				Accuracy (mean ± standard deviation)				Collocated precision (mean ± standard deviation)			
Vendor	Model	Unit	n	$PM_{2.5}$	CO	CO_2	NO_2	$PM_{2.5}$	CO	CO_2	NO_2
2B	PAM	2015	12	94.1 ± 4.5	97.3 ± 2.3	98.8 ± 1.1	–	6.2 ± 3.7	0.9 ± 0.6	0.8 ± 0.7	–
2B	PAM	2016	20	86.1 ± 8.7	96.5 ± 2.6‡	98.8 ± 0.9	–				
AmbiLabs	Nephelometer		31	91.5 ± 9.5	–	–	–	–	–	–	–
AQMesh		1	17	88.9 ± 8.7	–	–	65.4 ± 42.6	–	–	–	–
Duke		14	31	83.3 ± 16.3	–	–	–	5.9 ± 4.8	–	–	–
Duke		23	30	89.6 ± 8.0	–	–	–				
FTS	AQT420		31	73.2 ± 16.8	95.7 ± 3.0	–	⁂	–	–	–	–
Kunak	A10	1	23	δ	96.9 ± 2.4	–	75.6 ± 26.1	δ	0.8 ± 0.9	–	1.4 ± 1.5
Kunak	A10	2	26	δ	97.3 ± 1.4	–	79.2 ± 18.8				
Licor	850		31	–	–	99.1 ± 0.8	–	–	–	–	–
Met-One	E-BAM	1	26	93.4 ± 5.6	–	–	–	4.1 ± 5.9	–	–	–
Met-One	E-BAM	2	26	92.3 ± 7.2	–	–	–				
Met-One	E-Sampler	1	24	81.7 ± 17.6	–	–	–	2.7 ± 1.9	–	–	–
Met-One	E-Sampler	2	26	80.5 ± 17.4	–	–	–				
PurpleAir	PA-II-SD	1	30	89.5 ± 8.6	–	–	–	1.2 ± 1.3	–	–	–
PurpleAir	PA-II-SD	2	30	89.9 ± 7.5	–	–	–				
Thermo	PDR	1	26	87.2 ± 12.8	–	–	–	1.2 ± 0.8	–	–	–
Thermo	PDR	2	20	85.7 ± 14.3	–	–	–				
TSI	APS	A	29	72.5 ± 22.6	–	–	–	–	–	–	–
TSI	DustTrak	1	31	91.1 ± 8.7	–	–	–	2.1 ± 1.2	–	–	–
TSI	DustTrak	2	26	92.2 ± 7.9	–	–	–				

Notes. ‡Burn 31 was not included; ⁂Sensor not significantly correlated to reference; δCalibration did not significantly improve accuracy.

Table 5. 2021 burn-integrated study-specific smoke calibrated $PM_{2.5}$ and PM_{10} testing accuracy (%) and collocated precision (%).

				Accuracy (mean ± standard deviation)		Collocated precision (mean ± standard deviation)	
Vendor	Model	Unit	n	$PM_{2.5}$	PM_{10}	$PM_{2.5}$	PM_{10}
2B	PAM	1022	40	89.2 ± 10.6	88.2 ± 11.7	–	–
2B	BCP™		40	90.2 ± 8.8	–	–	–
AQMesh	2,450,480		40	78.1 ± 15.9	δ	–	–
Kunak	A10	1	–	–	–	–	–
Kunak	A10	2	40	δ	δ	–	–
Kunak	Pro	1	40	73.6 ± 18.9	74.5 ± 17.7	2.0 ± 1.7	15.0 ± 2.9
Kunak	Pro	2		74.1 ± 19.8	75.4 ± 20.8		
Met-One	E-Sampler	171	37	80.4 ± 17.2	–	0.9 ± 0.5	–
Met-One	E-Sampler	594	37	80.1 ± 17.2	–		–
PurpleAir	PA-II-SD	746D – New	40	88.7 ± 8.5	87.6 ± 9.0	1.9 ± 1.1	2.1 ± 1.6
PurpleAir	PA-II-SD	8171– New		88.6 ± 8.9	86.9 ± 9.8		
PurpleAir	PA-II-SD	BE2E – New		89.4 ± 7.9	87.3 ± 9.7		
PurpleAir	PA-II-SD	CBE8 – New		88.7 ± 8.9	87.2 ± 9.6		
PurpleAir	PA-II-SD	8BDE – Used		88.5 ± 9.4	87.1 ± 10.0		
PurpleAir	PA-II-SD	9349 – Used		89.0 ± 8.5	87.2 ± 9.5		
PurpleAir	PA-II-SD	C0BA – Used		88.2 ± 9.3	87.6 ± 9.6		
PurpleAir	PA-II-SD	EACF – Used		88.3 ± 8.6	87.5 ± 9.0		
PurpleAir	PA-II-SD	8BED – Used		88.5 ± 9.2	86.5 ± 10.6		
PurpleAir	PA-II-SD	09DE – Used		89.3 ± 8.4	87.1 ± 9.6		
PurpleAir	PA-II-SD	39C4 – Used		89.0 ± 8.5	87.5 ± 9.5		
PurpleAir	PA-II-SD	CCC3 – Used		88.8 ± 9.0	87.5 ± 9.9		
PurpleAir	PA-II-SD	1CCA – Used		88.5 ± 8.7	87.8 ± 8.7		
PurpleAir	PA-II-SD	8E2D – Used		88.5 ± 9.4	86.4 ± 10.2		
PurpleAir	PA-II-SD	8F98 – Used		87.1 ± 9.4	86.7 ± 8.9		
PurpleAir	PA-II-SD	39C8 – Used		86.8 ± 13.0	85.4 ± 13.9		
Senserion	SPS30		39	86.4 ± 12.3	85.4 ± 13.3	–	–
Thingy	AQ	2005	40	85.7 ± 12.9	84.5 ± 13.5	2.3 ± 1.3	3.3 ± 2.0
Thingy	AQ	2007		86.8 ± 11.5	85.4 ± 12.6		
Thingy	AQ	2008		83.9 ± 13.9	81.0 ± 15.8		
Thingy	AQ	2016		85.5 ± 12.4	83.9 ± 13.7		
Thingy	AQ	2019		86.0 ± 12.5	84.2 ± 13.8		
TSI	DustTrak	1	31	91.2 ± 6.5	88.9 ± 8.9	–	–
TSI	APS		31	74.3 ± 22.8	79.3 ± 18.5	–	–
Vaisala	AQT420		37	65.0 ± 30.9	60.8 ± 38.6	–	–

Note. δCalibration did not significantly improve accuracy.

$PM_{2.5}$ performance results by instrument/sensor

The Senserion SPS30 OPS native OEM sensor logged with manufacturer devkit software demonstrated fast response to changing concentrations (Appendix Figure A.3) and reported the best raw burn integrated accuracy of 86.7 ± 11.1% (mean ± standard deviation; Table 3) for all systems evaluated. The SPS30 has a visually linear response relative to the FRM through ~1 mg m^{-3} (Appendix Figure A.4a) that is slightly better fit to a quadratic model forced through zero (FTZ; Appendix Figure A.4b), with the coefficient of determination (r^2) improving from 0.976 to 0.985 due to a single high concentration. The 1-min SPS30 data show a similar relationship to the $PM_{2.5}$ reference, with linear FTZ (r^2 = 0.951), quadratic FTZ (r^2 = 0.982), and linear piecewise FTZ (r^2 = 0.985) models with clustering of responses appearing to show burn-specific differences (Appendix Figures A.5a-c). To highlight the influence of burn-specific optical and physical properties of the smoke aerosols on the SPS30 native sensor response relative to the reference $PM_{2.5}$, a subset of 12 2021 burns were plotted in Appendix Figure A.6, with data points from each burn identified along with unique linear regression model slopes presented (ranging from 0.673–1.512), similar to previously reported OPS behavior (Landis et al. 2021b) that will be investigated in $PM_{2.5}$ MLR analysis. The highest $PM_{2.5}$ FRM integrated burn (#33) concentration (1.39 mg m^{-3}) that appears to be a SPS30 response outlier in Appendix Figure A.4a, is bound within response slopes of other lower concentration burns (e.g., burn #3), suggesting that in the absence of other burn concentrations in this elevated range that the burn 33 response may not be a concentration-driven outlier. Therefore, the Senserion SPS30 was calibrated with a SSS linear FTZ regression model (Appendix Table B.6) that did not improve its overall reported accuracy 86.4% ± 12.3% (Table 5).

The Thingy AQ implementation of the Senserion SPS30 sensor reported similar raw accuracies to the previously discussed native OEM sensor, ranging from 81.2–86.3%, with a precision of 6.7 ± 1.1%. The Thingy AQ linear response relative to the reference $PM_{2.5}$ concentrations for each of the five tested units are presented in Appendix Figures A.7a–e and A.8a–e for burn integrated and 1-min data, respectively. The Thingy AQ units were calibrated with SSS linear FTZ marginally improving their reported accuracies to 83.9–86.8% and precision to 2.3 ± 1.3% (Table 5).

The 2B technologies dual wavelength black carbon photometer (BCP™) measures particulate light extinction at 405 nm (near UV) and 880 nm (IR) without filter preconcentration and is primarily marketed as a black carbon instrument. However, it also reports a PM mass concentration calculated using a wavelength-dependent mass extinction coefficient. The 2B BCP demonstrated fast response to changing concentrations (Appendix Figure A.9) and reported a raw accuracy of 83.9 ± 13.3%. Burn-integrated (Appendix Figure A.10) and 1-min data (Appendix Figure A.11) versus the reference $PM_{2.5}$ are presented, and good linear FTZ model fits of r^2 = 0.969 and 0.983 were observed, respectively. Similar to the burn-specific responses previously discussed, the BCP response to burn 33 was bounded by other lower-concentration response slopes, so a SSS linear FTZ calibration model was applied that improved the reported accuracy to 90.2 ± 8.8%.

The AmbiLabs (Warren, RI, U.S.A.) 2WIN Nephelometer consistently overestimated $PM_{2.5}$, resulting in large negative raw accuracies −254 ± 70% (Table 2). The nephelometer demonstrated fast response to changing concentrations (Appendix Figure A.12), and a linear response through ~1.5 mg m^{-3} for hourly integrated burns (Appendix Figure A.13a) and a linear response through ~1.6 mg m^{-3} for 1-min data (Appendix Figure A.13b), as determined using two-segment piecewise linear regression models. The nephelometer response above the identified inflection points are not supported by response characteristics at lower concentrations, suggesting a true response change. When data with reference $PM_{2.5}$ concentrations > 1.6 mg m^{-3} are removed, an excellent linear response (r^2 = 0.993) is observed, with some minor burn-specific clustering (Appendix Figure A.14). When a linear calibration model was applied (Appendix Table B.5), the instrument accuracy improved to 91.5 ± 9.5% (Table 4). After our 2019 testing, the AmbiLabs 2WIN nephelometer was designated by EPA for $PM_{2.5}$ (EQPM-0922–260) on October 28, 2022, with the addition of an omnidirectional sampling inlet, a $PM_{2.5}$ sharp cut cyclone, and firmware changes (version 1.36 or later), incorporating a new proprietary $PM_{2.5}$ algorithm (EPA 2022). When Ambilabs reprocessed the raw 2WIN chamber testing data using their FEM PM algorithm, the instrument's raw accuracy improved to 62.2 ± 21.6%.

The TSI DustTrak instruments tested also consistently overestimated $PM_{2.5}$ concentrations, with unit #1 accuracies reported −26.1 ± 27.9% in 2019 and −10.7 ± 28.9% in 2021, respectively (with aerosol density set to 1). The overall combined 2019 and 2021 DustTrak #1 accuracy was −18.4 ± 29.2% (n = 62). DustTrak precision was 2.7 ± 1.7% in 2019 when collocated instruments were tested. The DustTrak instrument provided fast (Appendix Figure A.15) and linear response over the full range of combined 2019 and 2021 burn-integrated FRM $PM_{2.5}$ concentrations (Appendix Figure A.16), with an r^2 = 0.990. Calibration using SSS regression models (Appendix Tables B.5, B.6)

dramatically improved the reported accuracy to 91.1–92.2% (Tables 4 and 5).

The well-documented inefficiency of counting small diameter aerosols (e.g., <40% for 0.5 µm aerosols) impacted the accuracy of the TSI APS derived $PM_{2.5}$ concentrations (Armendariz and Leith 2002; Peters and Leith 2003; Volckens and Peters 2005) and systematically underestimated concentrations by greater than a factor of two, with accuracies of 43.6 ± 18.0% in 2019 and 37.3 ± 13.7% in 2021, respectively. The overall combined 2019 and 2021 APS raw accuracy was 40.3 ± 16.1% ($n = 60$). APS response time was adequate (Appendix Figure A.17), but non-linear and scattered (Appendix Figure A.18) compared with the DustTrak. APS accuracy improved (72.5–74.3%; $r^2 = 0.882$) after calibrating with SSS regression models.

The Kunak A10 (OPC-R2), and Kunak Air Pro (OPC-N3) multi-sensor pod systems both incorporated Alphasense OPCs. When reporting 1-min data, these sensors showed PM measurement gaps for 5–7 mins for the OPC-R2 and 2–3 mins for the OPC-N3 every hour, and the first reported measurement after the pause was lower, requiring the screening and invalidation of those results. The response time and visible gaps in reported data are presented in Appendix Figure A.19. The OPC-R2 provided a linear response through ~1.5 mg m^{-3} during both the 2019 and 2021 studies, albeit with scatter (Appendix Figures A.20a-b). The collocated measurements in 2019 demonstrate that the scatter is reproducible (Appendix Figure A.20a), with a calculated precision of 1.8 ± 1.3% (Table 2), suggesting that the differences with the FRM may be related to optical aerosol properties. The raw accuracies of the OPC-R2 (Kunak A10) ranged between 73.2–74.2%, and attempts to implement SSS calibration models did not appreciably improve their performance due to burn-specific scatter. The OPC-N3-equipped Kunak Air Pro sensors provided non-linear $PM_{2.5}$ responses relative to the FRM (Appendix Figure A.21a), with raw accuracies ranging from 49.0–55.9% and a mean precision of 9.7 ± 4.1% in 2021 testing (Table 3). When burns with FRM $PM_{2.5} > 1$ mg m^{-3} are removed, the response is linear, with burn-specific scatter that is reproducibly reported by the collocated unit (Appendix Figure A.21b). When calibrated with SSS quadratic FTZ smoke regression models (Appendix Table B.6), their reported accuracies improve to 73.6–74.1% (Table 5).

The AQMesh sensor systems use their manufactured OPC units. The demo unit we tested in 2019 was configured to report 1-min resolution PM data and 15-min resolution gas data, and the PM response on a typical test day is presented in Appendix Figure A.22. The newly purchased unit tested in 2021 came configured to measure and report 15-min resolution integrated PM data and 1-min resolution gas data. Appendix Figure A.23 shows the lower resolution PM data for a typical test day. The raw accuracy of the AQMesh system was 56.7 ± 13.6% in the 2019 configuration and 76.2 ± 17.6% in the 2021 configuration. The different AQMesh units provided unique $PM_{2.5}$ responses relative to the FRM (Appendix Figure A.24). The 2021 data were best fit to a quadratic FTZ calibration model that only marginally improved the overall accuracy to 78.1 ± 15.9%. While the raw accuracy of the AQMesh was lower in 2019, the data were conducive to a higher-order cubic FTZ calibration model that increased the accuracy to 88.9 ± 8.7%.

The raw accuracy of the ThermoScientific Personal DataRAM pDR-1500 was 73.4 ± 19.5% and 50.9 ± 27.4% for units 1 and 2, respectively, with a precision of 13.4 ± 1.9%. The non-linear burn-integrated response versus the FRM concentrations for both units is presented in Appendix Figure A.25 and shows unit-specific response differences. Plots of 1-min $PM_{2.5}$ versus the reference concentrations are presented in Appendix Figures A.26a-b, showing linear clusters of differing responses relative to the reference. Appendix Figure A.27 presents a subset of 10 2019 burns plotted with data points from each burn identified along with unique linear regression model slopes (ranging from 0.917–1.644), highlighting that the clusters are associated with burn-specific responses. Application of quadratic calibration models to the data improved the pDR accuracy to 87.2 ± 12.8% (unit 1) and 85.7 ± 14.3% (unit 2), with precision improving to 1.2 ± 0.8%.

The new integrated FTS Inc. (Victoria, BC, Canada) remote satellite data telemetry sensor system tested in 2019 was equipped with a Vaisala AQT420 sensor that reported a raw $PM_{2.5}$ accuracy of 67.3 ± 25.6%. The FTS system PM sensor reported a concentration every 10 mins. The PM response on a typical test day is presented in Appendix Figure A.28, showing both a slow ramp up at fuel ignition and a delayed response during chamber ventilation. It is not clear whether this observed response pattern is a function of the OPC ventilation rate or the timing of actual measurement signal acquisition. A linear SSS calibration improved the accuracy to 73.2 ± 16.8%. A new Vaisala AQT420 was purchased for the 2021 testing, and raw ASCII data output was streamed directly to a laptop to investigate the native OEM performance of the stand-alone sensor versus the implementation of the sensor into the FTS system and its power management configuration as tested in 2019. The Vaisala AQT420 sensor reported PM readings every minute, but with the same repeated concentrations for 10-min before the value refreshed.

However, the new AQT420 responded in a similar delayed fashion as the FTS implemented AQT420 in 2019 and inexplicably would periodically not report elevated concentrations during some burns (affecting at least one burn each day), and the overall response factor was approximately an order of magnitude lower than the FRM. Appendix Figure A.29 shows the typical OEM AQT420 response pattern for burns 20 and 22, but no response to burn 21. Appendix Figures A.30a–b and A31a–b show the burn-integrated and 1-min AQT420 versus FRM $PM_{2.5}$, respectively. The impact of the instrument's delayed response can be seen in the 1-min resolution data plots, and the clustering of non-responsive burns are clearly observable in the 2021 study plots. The 2021 raw accuracy was 5.4 ± 3.3% and was improved to 65.0 ± 30.9% after calibration using a quadratic FTZ SSS regression model.

The E-BAM and E-Sampler instruments tested were pulled from the USFS Region 1 air monitor cache. The USFS Interagency Wildland Fire Air Quality Response Program (IWFAQRP; https://www.wildlandfiresmoke.net/) maintains a cache of Met-One E-BAM and E-Sampler units, which are dispatched nationwide with Air Resource Advisors (ARAs; USFS, 2020) to measure and disseminate smoke impact information to the public during large wildfire incidents. The E-BAM is a ^{14}C beta attenuation-based system that reports hourly (_HR) and 10-min estimated (_RT) data. Since the burns were not timed exactly to the top of an hour, E-BAM burn-integrated concentrations were calculated using the 10-min data stream. The raw accuracy of the E-BAMS ranged between 78.1–82.0%, with a collocated precision of 4.7 ± 6.6%. The near linear response and excellent precision is depicted in Appendix Figure A.32 and is consistent with previously reported Met-One BAM1022 (^{14}C beta attenuation-based FEM) results in smoke reported in Long et al. (2023). The SSS calibrated E-BAM accuracies improved to 92.3–93.4%, with a precision of 4.1 ± 5.9%. The E-Sampler light-scattering instrument reported raw accuracy range from 72.5–77.8% in 2019 and 58.9–59.2% in 2021, with precision of 5.0 ± 1.9% and 0.9 ± 0.5%, respectively. Appendix Figures A.33a (2019) and A.33b (2021) present the burn-integrated E-Sampler versus FRM $PM_{2.5}$. After quadratic FTZ SSS model calibration, the accuracies improved to 80.5–81.7%, with a precision of 2.7 ± 1.9% in 2019, and 80.1–80.4%, with a precision of 0.9 ± 0.5% in 2021.

The 2B PAM, Duke, and PurpleAir sensor systems tested all utilized Plantower OEM laser scattering sensors (Table 1). Since there are thousands of PurpleAir sensors in use, the sensors have been widely evaluated (Barkjohn K, Holder A, et al. 2021; Barkjohn KK, Gantt B, et al. 2021; Barkjohn et al. 2022; Malm et al. 2024), and are currently being utilized by the AirNow Fire and Smoke Map (https://fire.airnow.gov/), the next section is dedicated to presenting our results on these units. The 2B PAM units had intermittent issues initializing connection to local cellular service in 2019, which reduced the overall data completeness but provided complete data in 2021. The PAM raw accuracy ranged from 76.5–79.0% with precision of 6.3 ± 3.5% in 2019 and accuracy was 75.1 ± 19.0% in 2021. Appendix Figures A.34a (2019) and A.34b (2021) show the burn-integrated PAM versus FRM $PM_{2.5}$, with cubic FTZ regression models. Plots of 1-min PAM $PM_{2.5}$ versus the reference concentrations are presented Appendix Figures A.35a (2019) and A.35b (2021) to highlight the non-linear response and some burn-specific clustering. After SSS calibration, the accuracies ranged from 86.1–94.1% in 2019 and was 89.2 ± 10.6% in 2021. The used Duke sensors tested were the same units as previously described and evaluated by Zheng et al. (2018). The Duke sensors raw accuracy ranged from 76.6–79.9%, with precision of 3.3 ± 2.0%. Appendix Figure A.36a shows the burn-integrated Duke sensors versus FRM $PM_{2.5}$ with cubic FTZ regression models, and Appendix Figure A.36b shows the 1-min Duke sensors versus Reference $PM_{2.5}$ with cubic FTZ regression models. After calibration, the Duke accuracies ranged from 83.3–89.6%, with a precision of 5.9 ± 4.8%. The PAM and Duke units performed similarly, with reported concentration overestimating the FRM when $PM_{2.5} < 500\ \mu g\ m^{-3}$ and then underestimating higher concentrations between a factor of 1 – 2 at concentrations $> 1000\ \mu g\ m^{-3}$.

Comparison with previous studies

Trent (2003) evaluated the performance of one new E-Sampler and four new E-BAM instruments in pine needle smoke (30–1288 $\mu g\ m^{-3}$) versus a MesaLabs (Lakewood, CO, U.S.A.) Model BGI PQ200 FRM sampler during 30 test burns at the FSL. Ambient conditions for the testing were 18–27°C and RH ~ 30%. All four E-BAM units were highly correlated with the FRM ($r^2 > 0.994$) and overestimated $PM_{2.5}$, with E-BAM versus FRM burn-integrated linear regression slopes between 1.09 and 1.21 (intercepts ranged from −23.8 to −11.5). The two USFS cache E-BAM units we tested in 2019 also overestimated $PM_{2.5}$, with slopes of 1.13 and 1.16 ($r^2 \geq 0.99$) and intercepts of 26.9 and 6.6, respectively. The Trent (2003) E-Sampler unit performed similarly to E-BAMs, with a linear regression model versus the FRM of $1.13x + 3.41$ ($r^2 = 0.963$). The two USFS cache E-Sampler units we tested in 2019 overestimated $PM_{2.5}$, with slopes of 1.26 and 1.44 ($r^2 \geq 0.964$) and intercepts of −61.5 and −85.1,

respectively. The E-Sampler units we tested in 2021 overestimated $PM_{2.5}$, with the identical slope of 1.57 ($r^2 \geq 0.954$) and intercepts of −73.1 and −71.5. Trent (2003) did not report accuracy and precision information or attempt to evaluate the E-BAM or E-Sampler performance in terms of burn conditions.

Mehadi et al. (2020) assessed the performance of several instruments for measuring $PM_{2.5}$ in smoke, including the MetOne E-BAM, TSI Model 8520 DustTrak, Thermo pDR-1500, and Purple Air PA-II sensors in a chamber and at ambient sites. Pairs of each monitor type were collocated with Met-One Model BAM 1020 FEM (EPA 2015 - EQPM-0715–266), to measure smoke PM in laboratory testing (188–883 μg m^{-3}; RH = 24–68%). All four instrument/sensors overestimated $PM_{2.5}$, with the E-BAM and pDR-1500 agreeing best with the FEM reference. The PA-II units had a large positive offset and were poorly correlated with reference (linear regression line versus reference = 0.67x + 622; $r^2 = 0.33$). The DustTrak linear regression slope versus the FEM (y = 2.16x − 7.0; $r^2 = 0.84$) was about twice that of the E-BAM (y = 1.17x − 67.2; $r^2 = 0.95$) and pDR-1500 (y = 1.07x − 40.3; $r^2 = 0.80$). Overall, the Trent (2003), Mehadi et al. (2020), and our testing report consistent E-BAM performance versus an FRM/FEM reference with mean high bias of ~10–20%.

Delp and Singer (2020) evaluated the performance of several monitor types, including a Thermo Model pDR-1500 (PDR) and TSI Model DustTrak II-8533 (DT), during the sampling of smoke-impacted ambient air in Berkley, CA, over 13 days during the 2018 Camp Fire (~225 km away). The testing took place inside a lab with "high infiltration" of outdoor air and occasionally with the doors to the outside opened. The monitors were evaluated versus a Thermo Tapered Element Oscillating Microbalance (TEOM) Model 1045-DF FEM (EPA 2009 - EQPM-0609–182). Delp and Singer (2020) calculated adjustment factors (AF), the slope of a linear FTZ regression of candidate monitors versus the FEM, for 4-hour intervals of data. The PDR and DustTrak were only available for a "few days." The AF for the PDR was 1.89 (1.23, 2.21) (median and 10th & 90th percentiles), and the AF for the DT was 3.99 (3.57, 4.26). The DustTrak AF is almost double the 2.16 slope reported by Mehadi et al. (2020) and the slopes from our 2019 (2.34–2.45) and 2021 (2.37) testing. The Delp and Singer (2020) pDR-1500 AF of 1.89 is consistent with the slopes of our 2019 testing results (1.44–1.82), but higher than the 1.07 reported by Mehadi et al. (2020). The authors suggested the different responses between the pDR-1500 and DustTrak is related to the factory calibrations—"the DustTrak uses A1 Ultrafine dust, with mass median diameter (mmd) in the range of 3 to 5 μm. The PDR is calibrated with Arizona test dust A2 Fine, with mmd of 8 to 10 μm."

PurpleAir sensors raw and calibrated performance

The PurpleAir outdoor sensors tested include two Plantower PMS5003 OPS units that can be used to characterize some performance metrics and provide for ambient data to be aggregated and publicly displayed on the manufacturer's website (https://community.purpleair.com/c/data/7) and the AirNow Fire and Smoke Map (https://fire.airnow.gov/). Eighteen PurpleAir units were tested, two in 2019 and 16 in 2021. The exact use history of the USFS units tested in 2019 is uncertain, and therefore, we limit this discussion to the 16 units from the 2021 testing. Four of the 16 units were brand new (Tables 3, 5 and 6), while the remainder had been previously deployed for ~1 yr of ambient air sampling at sites in Boise, ID; Missoula, MT; and Reno, NV. The raw burn integrated data (PurpleAir data stream cf = 1) were screened using a quality assurance procedure adapted from the Barkjohn et al. (2022) criteria but based on our unique testing data, whereby if the two sensors in a single unit (A/B) had an hourly integrated concentration absolute relative percent difference (RPD) >50% then the data were flagged. None of the new PurpleAir units exceeded this criterion for any of the hourly integrated burns. However, we noted that two of the previously used units (8BED, 39C4) reported hourly concentrations that exceeded the paired sensor screening criteria for every test burn, and two other units (8F98, 39C8) reported six and eight paired hourly burn concentrations that exceeded the criteria, respectively. To preserve as much data as possible in these limited comparisons, we used the FRM reference concentration to identify which sensor in each pair was reporting erroneous readings and then invalidated that value and used the remaining sensor results for calculating burn integrated $PM_{2.5}$. However, in typical ambient smoke monitoring applications where A/B sensor precision screening criteria are applied (e.g., AirNow Fire and Smoke Map), it would not be possible to identify which sensor is erroneous, and it would be necessary to invalidate the flagged data points. Automated monitoring of A/B sensor precision would provide a mechanism for tracking performance deterioration and notifications for sensor replacement.

The study burn integrated PurpleAir $PM_{2.5}$ raw accuracy ranged from 34.0–57.4%, with a mean of 44.2% (Table 3). There was no discernible difference in study mean raw accuracy between the new and used units following data screening, and on a burn-by-burn basis, the

difference in mean accuracy between new and used units was <5.2% for 38 of 40 burns. Our test results are consistent with previous studies in smoke impacted environments; $PM_{2.5}$ mass concentrations reported by PurpleAir sensors are biased high, particularly for concentrations <1 mg m^{-3}, and require calibration (Holder et al. 2020; Barkjohn K, Holder A, et al. 2021; Barkjohn KK, Gantt B, et al. 2021). Quadratic FTZ SSS calibration models were fit to all PurpleAir sensors (Appendix Table B.6). Figure 1a presents an example PurpleAir #9349 raw $PM_{2.5}$ nonlinear response versus the FRM fit with a quadratic FTZ model, with an inset showing the raw accuracy distribution, and Figure 1b shows the calibrated linear response and improved accuracy distribution. Some of the PurpleAir units could be fit to a cubic FTZ model (Figure 1c) that potentially would improve accuracy in the 100–500 μg m^{-3} concentration range, but the improvement to accuracy over the full test range was not substantially different 89.0 ± 8.5% versus 89.9 ± 9.7% for unit #9349 (Figure 1d). Therefore, for consistency, the quadratic FTZ SSS calibration was applied to all the PurpleAir units. Appendix Figure A.37a shows the burn-integrated

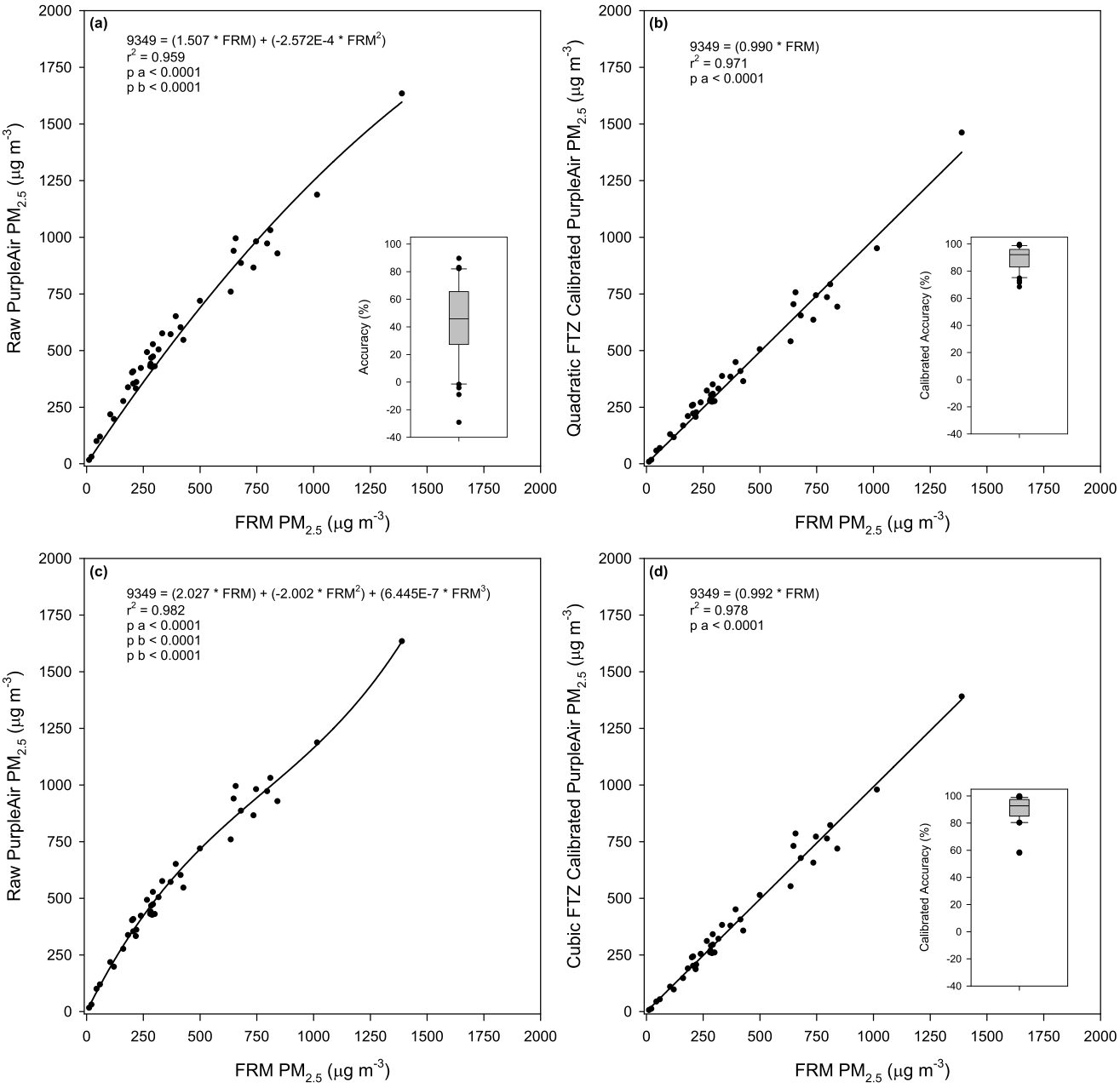

Figure 1. Example 2021 PurpleAir sensor raw $PM_{2.5}$ versus FRM fit with quadratic FTZ model with raw accuracy distribution whisker plot inset (a); quadratic calibrated $PM_{2.5}$ versus FRM with calibrated accuracy distribution inset (b); raw $PM_{2.5}$ versus FRM fit with cubic FTZ model (c); and cubic calibrated $PM_{2.5}$ versus FRM with calibrated accuracy distribution inset.

responses for an additional two PurpleAir units, and Appendix Figure A.37b shows the continuous 1-min comparison to the reference. The SSS quadratic FTZ regression calibration models based on the burn integrated FRM $PM_{2.5}$ dramatically improved accuracy to 86.8–89.4% (Table 5), roughly doubling the study mean raw accuracy for new and used sensors, and reduced the analytical artifact (Δ_{FRM}) across the range of $PM_{2.5}$ tested (Appendix Figures A.38a–b). We also applied the Fire and Smoke Map (FSM; Barkjohn K, Holder A, et al. 2021) and EPA (EPA; Barkjohn et al. 2022) calibration protocols to our raw PurpleAir data. Both the FSM calibration and the EPA calibration provided study mean burn integrated $PM_{2.5}$ accuracies comparable to the SSS calibrations (Table 6).

We examined the SSS, FSM, and EPA calibrated PurpleAir versus the FRM $PM_{2.5}$ accuracies to gauge how the calibrations perform across a wide range of $PM_{2.5}$ concentrations that may be encountered in smoke-impacted environments. The Δ_{FRM} for the three calibration models are plotted versus the FRM $PM_{2.5}$ in Appendix Figures A.39a–b. The accuracies of the three calibration models are comparable over most of the tested range; however, the SSS calibration is somewhat less accurate than the other models for FRM reference $\leq 200\ \mu g\ m^{-3}$. Given that PurpleAir sensors are widely used for smoke monitoring, confirmation of their accuracy in heavily smoke-impacted environments is an important aim of this study. Defining heavily smoke-impacted environments using the U.S. EPA Air Quality Index (AQI) threshold for the PM category "hazardous," $PM_{2.5} \geq 225.5\ \mu g\ m^{-3}$, we find the mean accuracies were 88.8% for the FSM calibration, 86.5% for the EPA calibration, and 90.9% for the SSS calibration. These results indicate that using the FSM or EPA calibration model will provide accurate estimates of $PM_{2.5}$ in heavily smoke-impacted environments. While occurrences of smoke $PM_{2.5} > 225\ \mu g\ m^{-3}$ have been somewhat limited at U.S. regulatory monitoring sites, (our query of the U.S. EPA Air Quality System (AQS; EPA 2025), which contains ambient air monitoring data collected by federal, state, and local agencies nationwide at mostly urban sites, found that across the western 11 states between 2010–2024 there were 13,099 site-hours (202 sites) for which FEM $PM_{2.5}$ exceeded $225\ \mu g\ m^{-3}$ during periods of wildland fire smoke impacts), they are likely far more common on fire lines, in fire camps, and in rural communities nearby active wildfires (Navarro 2020).

$PM_{2.5}$ MLR analysis

The MLR analysis results for the 2021 FSL chamber burns are presented in Appendix Table B.7, and the independent variables and included interaction terms used were typically able to explain >90% of the variability in instrument/sensor $PM_{2.5}$ measurement artifacts (Δ_{FRM}) in smoke, with coefficients of multiple determination (R^2) ranging from 0.806–1.000. The MLR analysis found that $PM_{2.5}$ measurement artifacts in instrument/sensor systems were mainly a result of changes in smoke aerosol optical properties (BC, Δ_C), effective aerosol density (ρ_{eff}), and aerosol size distribution (MMAD; see significant type III sum of squares F values in Appendix Table B.7). The dominance in the physical and optical aerosol properties affecting $PM_{2.5}$ measurement artifacts is not surprising due to the limited temperature (18.4–21.3°C) and RH (14–33%) range during the FSL chamber testing. However, the AQMesh

Table 6. 2021 PurpleAir sensor burn-integrated $PM_{2.5}$ testing accuracy (%) for raw and several adjustment regimens.

		Accuracy (mean ± standard deviation)			
PA-II-SD Unit	Study	Raw	Fire & Smoke Map[δ]	EPA smoke[‡]	Study sensor specific
#1	2019	56.7 ± 25.3	90.6 ± 8.1	87.6 ± 9.0	89.5 ± 8.6
#2	2019	57.0 ± 24.6	91.0 ± 7.9	88.4 ± 9.2	89.9 ± 7.5
746D – New	2021	35.2 ± 30.6	84.1 ± 11.0	87.6 ± 9.2	88.7 ± 8.5
8171 – New	2021	49.1 ± 27.7	89.7 ± 7.3	86.8 ± 8.2	88.6 ± 8.9
BE2E – New	2021	50.9 ± 29.7	90.0 ± 6.8	85.9 ± 8.8	89.4 ± 7.9
CBE8 – New	2021	49.6 ± 28.0	90.0 ± 7.2	86.5 ± 8.1	88.7 ± 8.9
8BDE – Used	2021	42.6 ± 29.9	87.7 ± 8.9	87.8 ± 8.5	88.5 ± 9.4
9349 – Used	2021	43.0 ± 28.2	88.0 ± 8.4	88.1 ± 8.1	89.0 ± 8.5
C0BA – Used	2021	36.2 ± 30.5	83.9 ± 11.9	86.7 ± 10.1	88.2 ± 9.3
EACF – Used	2021	45.5 ± 30.6	88.9 ± 8.0	87.7 ± 8.3	88.3 ± 8.6
8BED – Used	2021	57.4 ± 28.5	88.4 ± 7.3	82.7 ± 9.3	88.5 ± 9.2
09DE – Used	2021	41.2 ± 32.6	88.2 ± 9.8	87.1 ± 10.3	89.3 ± 8.4
39C4 – Used	2021	47.0 ± 31.1	89.8 ± 7.5	86.9 ± 8.2	89.0 ± 8.5
CCC3 – Used	2021	51.7 ± 25.9	89.9 ± 7.1	86.4 ± 8.9	88.8 ± 9.0
1CCA – Used	2021	48.3 ± 27.1	88.4 ± 7.6	86.4 ± 8.4	88.5 ± 8.7
8E2D – Used	2021	40.8 ± 28.2	86.4 ± 10.3	87.5 ± 9.3	88.5 ± 9.4
8F98 – Used	2021	34.0 ± 34.3	82.5 ± 12.9	87.0 ± 11.0	87.1 ± 9.4
39C8 – Used	2021	35.9 ± 41.2	85.7 ± 14.3	87.8 ± 13.1	86.8 ± 13.0

Notes. [δ]AirNow Fire & Smoke Map (Barkjohn K, Holder A, et al. 2021); [‡]Barkjohn et al. (2022).

(temp & RH), Thingy (RH * temp), and Vaisala AQT420 (RH) measurement artifacts were still found to be significantly impacted by environmental conditions. The 2B photometer was found to be impacted by the fewest independent variables and interactions, with only Δ_C, MMAD*$PM_{2.5}$, and Δ_C*MCE explaining 94% of the variability in $PM_{2.5}$ measurement artifacts.

The non-regulatory instrument/sensor $PM_{2.5}$ measurement artifact MLR model independent variable and interaction drivers determined in this analysis are mostly consistent with those reported by Long et al. (2023) for understanding $PM_{2.5}$ measurement artifacts for the FEM-designed Teledyne T640 scattered light spectroscopy $PM_{2.5}$ regulatory instrument. One difference in the non-regulatory instrument/sensor reported results is the significance of aerosol size distribution (MMAD) determined in this study, derived from the LAS instrument, and the APS-derived MMAD utilized in the Long et al. (2023) analysis, not a significant independent variable or interaction. Either the Teledyne T640 measurement and/or $PM_{2.5}$ mass concentration algorithm are better at controlling for changes in aerosol size distribution, or the APS-derived MMAD used in the Long et al. (2023) analysis was not able to adequately resolve the aerosol size distribution of the smoke in the chamber due to known inefficiencies in counting particle sizes below 0.8 microns (Volckens and Peters 2005; Peters et al. 2006). Our MMAD measurement comparisons between the APS versus $PM_{2.5}$ LAS are summarized in Appendix Table B.8 and show the large differences between MMAD estimates based on the different measurement technologies.

Biomass burning smoke aerosols are predominantly (60–95%) organic aerosol (Urbanski et al. 2022), with optical properties that are highly variable (Kirchstetter et al. 2004; Saleh et al. 2013; Laskin et al. 2015). The significant MLR model interactions between combustion conditions (MCE) and the amount of BC (MCE*BC) and the wavelength (λ) dependence of aerosol optical absorption (MCE*ΔC) observed for most tested instruments/sensors highlight the potential raw measurement artifacts as a function of fire activity (e.g., flaming, smoldering) and, therefore, specific job tasks that firefighters are assigned. The non-linear SSS and the FSM/EPA (PurpleAir) calibrations are shown to do a reasonable correction for these changes in aerosol physical and optical properties, with accuracy improvements to >80% enabling reliable public health messaging for firefighters and impacted nearby communities.

Gas sensor performance

The raw burn integrated gaseous testing accuracy and collocated precision results are summarized for all tested instruments/sensors in Tables 2 and 7 for 2019 and 2021 studies, respectively. The 2019 burn-integrated gaseous reference concentrations (Appendix Table B.3) ranged from 0.69–6.81 ppm (CO), 453–530 ppm (CO_2), 15.9–135.1 ppb (NO), 7.5–48.1 (NO_2), and 1.4–77.6 ppb (SO_2); and the 2021 burn-integrated gaseous reference concentrations (Appendix Table B.4) ranged from 0.31–4.25 ppm (CO), 439–517 ppm (CO_2), 8.6–144.9 ppb (NO), 3.1–39.3 (NO_2), and 1.5–35.1 ppb (SO_2). We observed generally good accuracies for CO and CO_2 sensors (75–99%), disparate moderate-good performance for NO_2 (45–80%), and generally poor performance for SO_2 (−889 to 57%). SSS calibration equations were developed, evaluated, and applied to all the gas sensors to evaluate the potential for improving their performance in smoke. All the selected calibration equations are presented in Appendix Tables B.5 and B.6 for the 2019 and 2021 studies, respectively. The updated regression calibrated gas instrument/sensor performance results are summarized in Table 4 (2019) and Table 8 (2021) and demonstrate that most performance metrics significantly improved versus their raw metrics. As expected, sensors with well-fit SSS calibration regression equations produced the most significant performance improvements.

Table 7. 2021 burn-integrated raw gas species testing accuracy (%) and collocated precision (%).

				Accuracy (mean ± standard deviation)				Collocated precision (mean ± standard deviation)			
Vendor	Model	Unit	n	CO	CO_2	NO_2	SO_2	CO	CO_2	NO_2	SO_2
2B	PAM	1022	40	73.5 ± 10.7‡	94.6 ± 1.7	–	–	–	–	–	–
AQMesh		2450480	40	69.6 ± 12.6	89.9 ± 1.5	44.7 ± 19.4	56.9 ± 22.1	–	–	–	–
Kunak	A10	1	40	87.8 ± 11.8	–	56.7 ± 25.2	–	7.9 ± 3.6	–	22.0 ± 20.3	–
Kunak	A10	2		91.9 ± 12.0	–	52.7 ± 18.6	–				
Kunak	Pro	1	40	88.6 ± 4.1	93.2 ± 2.5	65.7 ± 17.5	−889 ± 737	1.6 ± 1.3	1.7 ± 0.7	39.2 ± 31.9	10.8 ± 6.8
Kunak	Pro	2		90.1 ± 3.3	91.0 ± 2.3	43.1 ± 20.2	−730 ± 634				
Licor	850		40	–	94.7 ± 1.0	–	–	–	–	–	–
Senserion	SCD30		38	–	94.5 ± 2.3	–	–	–	–	–	–
Senserion	SCD41		6	–	93.3 ± 0.2	–	–	–	–	–	–
Vaisala	AQT420		37	82.8 ± 9.4	–	49.9 ± 66.2	−687 ± 613	–	–	–	–
Vaisala	GMP343		40	–	94.0 ± 2.4	–	–	–	–	–	–

Notes. ‡Six burns (7–8; 14–16; 22) where CO sensor response became very noisy and erratic were excluded; if included, the accuracy would have been 65.3 ± 25.9.

Table 8. 2021 burn-integrated calibrated gas species testing accuracy (%) and collocated precision (%).

				Accuracy (mean ± standard deviation)				Collocated precision (mean ± standard deviation)			
Vendor	Model	Unit	n	CO	CO_2	NO_2	SO_2	CO	CO_2	NO_2	SO_2
2B	PAM	1022	40	92.3 ± 8.0‡	98.9 ± 0.9	–	–	–	–	–	–
AQMesh		2450480	40	92.5 ± 12.3	99.1 ± 1.1	61.2 ± 61.8	δ	–	–	–	–
Kunak	A10	1	40	90.8 ± 11.3	–	63.2 ± 41.6	–	4.1 ± 5.3	–	13.2 ± 9.2	–
Kunak	A10	2		91.8 ± 10.3	–	80.0 ± 21.8	–				
Kunak	Pro	1	40	95.0 ± 5.3	98.2 ± 1.2	79.3 ± 27.8	38.9 ± 62.3‡	1.5 ± 1.2	0.4 ± 0.3	7.2 ± 7.4	4.4 ± 5.4‡
Kunak	Pro	2		96.0 ± 4.7	98.1 ± 1.2	75.5 ± 30.2	41.4 ± 62.3‡				
Licor	850		40	–	99.3 ± 0.6	–	–	–	–	–	–
Senserion	SCD30		38	–	98.4 ± 1.4	–	–	–	–	–	–
Senserion	SCD41		6	–	99.9 ± 0.9	–	–	–	–	–	–
Vaisala	AQT420		37	90.5 ± 14.2	–	δ	44.9 ± 60.0‡	–	–	–	–
Vaisala	GMP343		40	–	98.6 ± 0.9	–	–	–	–	–	–

Notes. ‡Six burns (7–8; 14–16; 22) where CO electronic sensor response became very noisy and erratic were excluded; ‡Chemical interference impeded calibration of sensors; δCalibration did not improve accuracy.

Two dedicated NDIR-based CO_2 instruments, the Licor Model 850 (used) and the Vaisala Model GMP343 (new), were tested. Unlike many of the other systems tested, these units can be user calibrated using standard air quality calibration systems; therefore, both systems were zero and span-calibrated using an EPA protocol gas diluted into ultra scientific grade zero air prior to each study period and were evaluated versus the USFS CRDS CO_2 reference instrument calibrated using a separate system. The Licor 850 raw accuracy was 94.0 ± 1.3% in 2019 (Appendix Figure A.40a) and 94.7 ± 1.0% in 2021 (Appendix Figure A.40b) and were linear through the entire dynamic test range (r^2 = 0.958–0.992). Study calibration to the reference analyzer resulted in slightly improved accuracy to 99.1 ± 0.8% and 99.3 ± 0.6% in 2019 and 2021, respectively. The Vaisala GMP343 was only tested in 2021, and its raw accuracy was 94.0 ± 2.4% and was linear through the entire test range (r^2 = 0.963; Appendix Figure A.41). Study calibration slightly improved accuracy to 98.6 ± 0.9%.

The 2B PAM units tested in 2019 (#2015, #2016), which employed Alphasense Model CO-A4 amperometric sensors, were new and had excellent raw accuracy for CO (93.9–94.3%), precision (2.7 ± 1.2%), and linearity (r^2 > 0.995–0.998; Appendix Figure A.42a). The Amphenol (Wallingford, CT, U.S.A.) Model TelAire T6713 NDIR CO_2 response from the two PAM units were linear (r^2 = 0.855–0.998; Appendix Figure A.42b), but units had large offsets versus the reference concentrations and different slopes (1.562–2.125) that negatively affected raw accuracy (−133.7 to 48.6%) and precision (46.7 ± 25.8%). Following SSS linear calibration, the CO_2 accuracies (99 ± 1%) and precision (0.8 ± 0.7%) were dramatically improved, and the CO accuracies were slightly improved to 96.5–97.3%, with corresponding precision improvements to 0.9 ± 0.6%. We attempted to test two other 2B PAM units in 2021 after two years of ambient use (#1022, #1036), but unit #1036 sensors were all non-responsive to changing concentrations in the chamber, and the results were invalidated. The CO sensor on unit #1022 intermittently provided extremely noisy signals, and so results from six burns were invalidated (Burns 7–8; 14–16; 22). The raw CO accuracy for the remaining 34 burns was 73.5 ± 10.7%, with good linearity (r^2 = 0.874; Appendix Figure A.43a). The raw accuracy for CO_2 including all 40 burns, was 94.6 ± 1.7%, with good linearity (r^2 = 0.846; Appendix Figure A.43b). SSS linear calibrations substantially improved the CO accuracy to 92.3 ± 8.0% and slightly improved the CO_2 accuracy to 98.9 ± 0.9%.

The Kunak A10 units, which employed Alphasense OEM sensors for all gases, were a year old when tested in 2019, after having been tested in 2018 as part of the EPA Wildland Fire Sensor Challenge (Landis et al. 2021a). Following the 2019 chamber testing, they were deployed to ambient monitoring sites until FSL testing again in 2021. The 2019 raw accuracies for Alphasense Model CO-B4 CO ranged from 87.6–92.4%, with performance remaining virtually unchanged after two years of continuous ambient use, with 2021 raw accuracies ranging between 87.8–91.9%. The CO precision was 14.4 ± 1.0%, with excellent linearity (r^2 > 0.996; Appendix Figure A.44a) in 2019 and precision of 7.9 ± 3.6%, with excellent linearity (r^2 > 0.980; Appendix Figure A.45a) in 2021. As the CO sensors aged, we observed increased variability, with accuracy standard deviations increasing by approximately a factor of three and decreased study calibrated accuracies. The Alphasense Model NO_2-B43F sensor performance did degrade over this two-year period, with raw accuracies ranging from 65.9–73.9% in 2019 and 52.7–56.7% in 2021. The NO_2 sensors had substantially more scatter *versus* the FEM in 2019 (r^2 = 0.445–0.600; Appendix Figure A.44b) compared with 2021 (r^2 = 0.659–0.868; Appendix Figure A.45b), which was also true of the other NO_2 sensors tested.

The Kunak Air Pro units were new in 2021 when tested and were configured with PM and manufacturer branded CO, CO_2, NO_2, and SO_2 sensors. The raw accuracies for CO ranged from 88.6–90.1%, with a precision of 1.6 ± 1.3%, and were linear over the full dynamic test range (r^2 ≥ 0.994; Appendix Figure A.46a). The CO_2 sensor raw accuracies ranged from 91.0–93.2%, with a precision of 1.7 ± 0.7%, and were linear over the full dynamic test range with more scatter (r^2 ≥ 0.611; Appendix Figure A.46b). The NO_2 sensor raw accuracies ranged from 43.1–65.7%, with a precision of 39.2 ± 31.9%, and were nonlinear over the dynamic test range. The NO_2 response was well fit to a quadratic regression model ($r^2 > 0.867$; Appendix Figure A.46c). The SO_2 sensor response appeared to be affected by an electrochemical interference, resulting in over reporting concentrations in some cases by an order of magnitude resulting in raw accuracies ranging from −730 to −889% and precision of 10.8 ± 6.8% (Appendix Figure A.46d). The Kunak Air Pro study calibrated mean accuracy results improved to 95.0–96.0% for CO, 98.1–98.2% for CO_2, and 75.5–79.3% for NO_2. The calibrated SO_2 accuracy improved to 38.9–41.4%, with a precision of 4.4 ± 5.4%, but was impeded by a suspected chemical interferent(s).

The AQMesh demo unit tested in 2019 was configured for PM and NO_2. The NO_2 raw accuracy was 62.6 ± 21.8%. The NO_2 performance was negatively affected by delayed sensor response, variable baseline, and differential response sensitivity during tests. Appendix Figure A.47 demonstrates the delayed response on a typical test day and the similar reported integrated NO_2 concentrations during burns 27 (16.8 ppb) and 28 (15.2 ppb), while the FEM instrument reported NO_2 concentrations for these burns (19.2 and 37.8 ppb) that were almost a factor of two different. The AQMesh limited NO_2 response range (12.4–26.7 ppb) *versus* the FEM NO_2 range of response is (7.5–38.3 ppb) depicted in Appendix Figure A.48. The sensor reported NO_2 was not significantly correlated with the FEM reference concentration using parametric (Pearson) or non-parametric (Spearman; Kendall Tau b) tests; therefore, a calibration was not attempted.

The newly purchased AQMesh unit tested in 2021 was configured with PM, CO, CO_2, NO_2, and SO_2 sensors. This unit reported a raw accuracy for CO was 69.6 ± 12.6% and was linear over the full dynamic test range ($r^2 = 0.988$; Appendix Figure A.49a). The CO_2 sensor raw accuracy was 89.9 ± 1.5% and had a linear response ($r^2 = 0.914$; Appendix Figure A.49b). The NO_2 and SO_2 sensors generally underestimated smoke concentrations by a factor of two, but showed no evidence of gross interferences. The NO_2 raw accuracy was 44.7 ± 19.4% and was linear over the dynamic test range ($r^2 = 0.686$; Appendix Figure A.49c), and the SO_2 sensor raw accuracy was 56.9 ± 22.1% ($r^2 = 0.632$; Appendix Figure A.49d). The AQMesh SSS calibrated mean accuracy results improved to 92.5 ± 12.3% for CO, 99.1 ± 1.1% for CO_2, and 61.2 ± 61.8% for NO_2. Attempts to calibrate the SO_2 sensor results using standard polynomial regression models did not result in meaningful improvement.

The new Vaisala AQT420 sensor that was integrated into the FTS system tested in 2019 demonstrated good temporal (Appendix Figure A.50a) and linear response versus the FEM reference instrument for CO (Appendix Figure A.51a), erratic response for NO_2 (Appendix Figure A.50b and Appendix Figure A.51b), and generally delayed response with severe hysteresis for SO_2 (Appendix Figure A.50c and Appendix Figure A.51c). The raw accuracy results were 81.6 ± 5.4% for CO, 42.7 ± 52.3% for NO_2, and −83.3 ± 250.4% for SO_2. The linear SSS calibrated CO results improved to 95.7 ± 3.0%. Attempts to calibrate the SO_2 sensor did not appreciably improve its accuracy due to the observed temporal response and hysteresis issues. The sensor reported NO_2 was not significantly correlated with the FEM reference concentration using either parametric or non-parametric tests; therefore, a calibration was not attempted.

A new Vaisala AQT420 sensor was tested in 2021, independent of the FTS system. Similarly good linear CO sensor response was observed in 2021 (Appendix Figure A.52a), with raw accuracies of 82.8 ± 9.4% and SSS calibrated accuracies of 90.5 ± 14.2%. The raw NO_2 performance was slightly improved at 49.9 ± 66.2% and was significantly correlated with the FEM reference concentrations (Appendix Figure A.52b), but attempts to calibrate the sensor did not result in improved accuracy due to the high degree of scatter observed. The SO_2 sensor reported concentrations that, on average, were almost an order of magnitude higher than the FEM reference (Appendix Figure A.52c), resulting in high negative raw accuracies of −687 ± 613%. SSS Calibration improved accuracies to 44.9 ± 60.0%.

The native Senserion OEM SCD30 and SCD41 NDIR CO_2 sensors logged with manufacturer evaluation kit hardware and software performed well, with raw accuracies of 94.5 ± 2.3% (SCD30; $n = 38$) and 93.3 ± 0.2% (SCD41; $n = 6$). The SCD30 provided a linear response (Appendix Figure A.53; $r^2 = 0.678$) with some scatter. The SCD30 SSS-calibrated accuracy improved slightly to 98.4 ± 1.4%. Unfortunately, the SCD41 was received late during the study and was only available for the last two days of testing. Nonetheless, we included the results for the limited comparisons due to its excellent linearity (Appendix Figure A.54; $r^2 = 0.994$) and calibrated accuracy 99.9 ± 0.9%.

Conclusion

This study comprehensively evaluated commercially available non-regulatory PM and gas instruments/sensors under simulated wildland fire smoke conditions. The findings can aid in the selection and implementation of non-regulatory air quality instruments/sensors during prescribed fires and wildfire response operations to improve workforce protection and public health messaging. While non-regulatory instruments/sensors offer rapid deployment and broad spatial coverage, their performance varied. Selection and configuration of instruments/sensors is critical to quantifying firefighter exposure during the extreme dynamic conditions experienced on the fire line. Most lower-cost PM sensors, particularly those using OPS technology, exhibited a high bias in raw $PM_{2.5}$ measurements in smoke (Figure 2). However, applying SSS calibration models significantly improved their accuracy, with many instruments/sensors achieving post-calibration accuracies >85% (Figure 2). Notably, native OEM/Thingy AQ implemented Senserion SPS30 sensor and 2B Tech BCP instrument demonstrated the best raw and SSS-calibrated performance, while the widely used PurpleAir sensors demonstrated substantial performance improvements after SSS/FSM/EPA calibration. The study also found that instrument/sensor $PM_{2.5}$ measurement artifacts are mainly driven by changes in aerosol optical properties, effective density, and size distribution and that non-linear heavy smoke-specific calibrations could account for most of these confounding effects on reported concentrations. Electrochemical and NDIR-based gas sensors for CO and CO_2 generally achieved accuracies >90% after SSS calibration. However, sensors for NO_2 and especially SO_2 showed variable and often moderate to poor performance, with interferences and non-linearities observed in several models.

This study highlights the importance of smoke-specific calibration and periodic reference checks for non-regulatory instruments/sensors, especially when deployed in high-concentration smoke environments or after extended field use. Data quality assurance protocols, such as collocated sensor comparisons and outlier screening, ensure reliable measurements. Recent studies indicate that integration of non-regulatory instrument/sensor data with regulatory monitoring networks and satellite

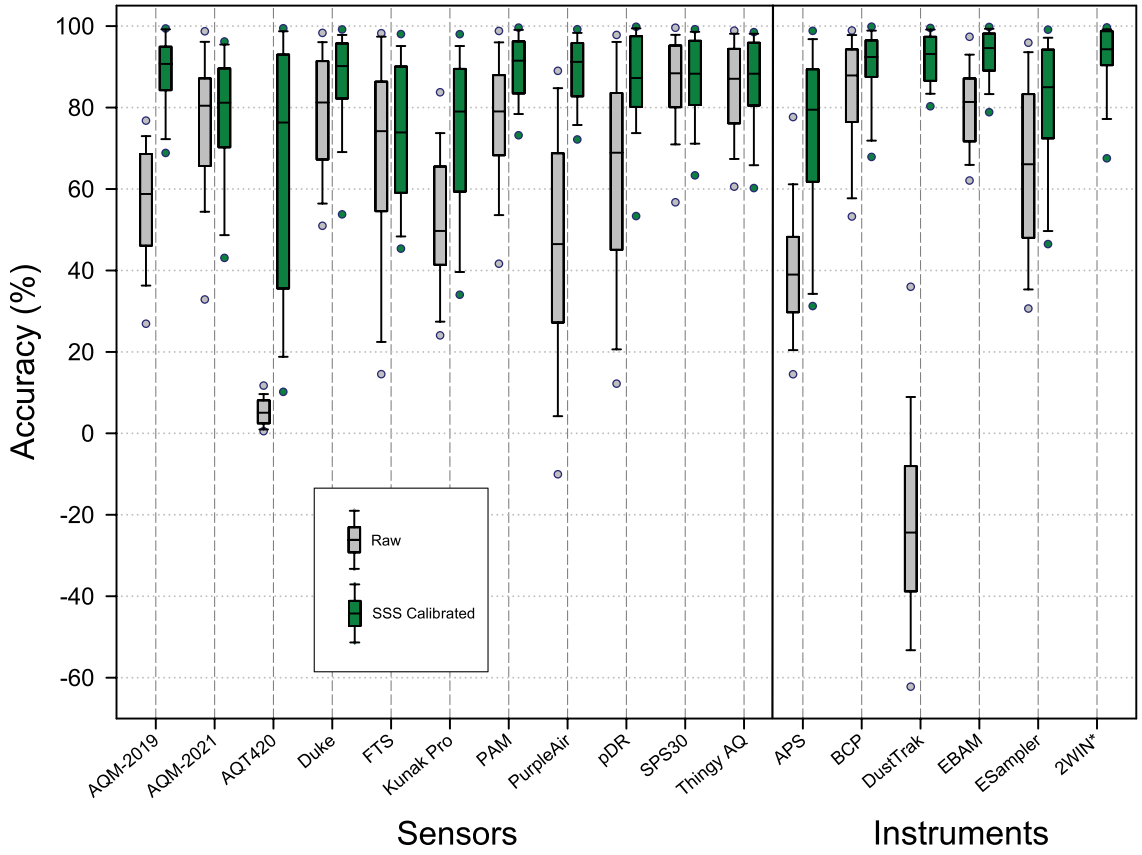

Figure 2. Summary boxplot of the combined 2019 and 2021 raw and SSS calibrated $PM_{2.5}$ accuracy aggregated by instrument/sensor model (Table 1). The box shows the interquartile range, the line inside box is the median, the error bars are the 10th and 90th percentiles, and the dots are the largest outliers. The solid vertical line separates sensors (left) and instruments (right). *2WIN nephelometer raw accuracy is off scale (−254.2 ± 70.2).

observations can enhance smoke exposure assessment, support public health advisories, and improve smoke modeling and forecasting. However, our study shows that non-regulatory instruments/sensors are limited and require ongoing calibration near active wildland fires producing high $PM_{2.5}$ concentrations.

Acknowledgment

The authors thank Kyle Digby (Amentum), Andrew Habel (Amentum), Ali Kamal (EPA), Emily Lincoln (USFS), and Brittany Thomas (Amentum) for their assistance with chamber measurements and Bob Vanderpool (EPA) for technical advice.

Author contributions

Matthew S. Landis served as EPA principal investigator and Shawn P. Urbanski served as USFS principal investigator and prepared the manuscript with contributions from all coauthors. Matthew S. Landis; Shawn P. Urbanski; Russell W. Long, Jonathan Krug, and Jenny Perth, performed the Missoula FSL chamber testing and data quality assurance review. Maribel Colón prepared and weighed all the FRM filters. Matthew S. Landis was the study data manager and conducted all the statistical data analysis and modeling.

Disclosure statement

The EPA through its Office of Research and Development (ORD) funded and conducted this research in collaboration with the USFS Rocky Mountain Research Station under Memorandum of Understanding 20-MU-112216737-003. The views expressed in this paper are those of the authors and do not necessarily reflect the views or policies of EPA or USFS. It has been subjected to Agency review and approved for publication. Mention of trade names or commercial products do not constitute an endorsement or recommendation for use.

Funding

The work was supported by the U.S. Environmental Protection Agency [Intramural Funding]; U.S. Forest Service [In Kind Support].

About the authors

Matthew S. Landis, Ph.D., corresponding author, is a senior research environmental health scientist with the U.S. EPA, Office of Research and Development, based in Research Triangle Park, NC.

Russell W. Long, Ph.D., is a senior research chemist with the U.S. EPA, Office of Research and Development, based in Research Triangle Park, NC.

Jonathan D. Krug is an engineer with the U.S. EPA, Office of Research and Development, based in Research Triangle Park, NC.

Maribel Colón was a research chemist with the U.S. EPA, Office of Research and Development, based in Research Triangle Park, NC and is now retired.

Jenny Perth was a project leader with US Forest Service, National Technology and Development Program, based in Missoula, MT and is now retired.

Shawn P. Urbanski, Ph.D., is a research physical scientist with the USDA Forest Service, Rocky Mountain Research Station, based in Fort Collins, CO.

ORCID

Matthew S. Landis http://orcid.org/0000-0002-8742-496X
Russell W. Long http://orcid.org/0000-0001-9051-7246
Jonathan D. Krug http://orcid.org/0000-0002-7045-7650
Shawn P. Urbanski http://orcid.org/0000-0002-9669-5120

Data availability statement

The EPA data system has assigned the data associated with the paper as https://doi.org/10.23719/1532438.

References

Abatzoglou JT, Williams AP. 2016. Impact of anthropogenic climate change on wildfire across Western U.S. forests. Proc Natl Acad Sci USA. 113(42):11770–11775. https://doi.org/10.1073/pnas.1607171113

Adetona O et al. 2016. Review of the health effects of wildland fire smoke on wildland firefighters and the public. Inhalation Toxicol. 28(3):95–139. https://doi.org/10.3109/08958378.2016.1145771

Akagi SK et al. 2011. Emission factors for open and domestic biomass burning for use in atmospheric models. Atmos Chem Phys. 11(9):4039–4072. https://doi.org/10.5194/acp-11-4039-2011

Allen G, Babich P, Poirot R. 2004. Evaluation of a new approach for real time assessment of wood smoke PM. 2004. In: Proceedings of air & waste management association visibility specialty conference on regional and global perspectives on haze: causes, Consequences and Controversies; Asheville, NC. Air & Waste Management Association. https://www-f.nescaum.org/documents/2004-10-25-allen-realtime_woodsmoke_indicator_awma.pdf/view

Alonso-Blanco E et al. 2018. Wildfire smoke plumes transport under a subsidence inversion: climate and health implications in a distant urban area. Sci Total Environ. 619-620:988–1002. https://doi.org/10.1016/j.scitotenv.2017.11.142

Andersen A et al. 2021. Chapter 6: health and ecological effects of wildland fire smoke exposure. Comparative Assessment of The Impacts of Prescribed Fire Versus Wildfire (CAIF): A Case Study in the Western U.S., U.S. EPA, EPA/600/R-21/197. https://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=352824

Armendariz AJ, Leith D. 2002. Concentration measurement and counting efficiency for the aerodynamic particle sizer 3320. J Aerosol Sci. 33(1):133–148. https://doi.org/10.1016/S0021-8502(01)00152-5

Ashley WS, Strader S, Dziubla DC, Haberlie A. 2015. Driving blind weather-related vision hazards and fatal motor vehicle crashes. Bull Amer Meteor Soc. 96(5):755–778. https://doi.org/10.1175/BAMS-D-14-00026.1

Associated Press. 2022. 3 Dead in crashes amid fog and smoke on I-95 in Florida. [accessed 2024 May 22]. https://apnews.com/article/florida-daytona-beach-florida-highway-patrol-a79077468133cfbddfb2091700dbc36b

Barkjohn K et al. 2024. Air quality experts convene: current air quality assurance considerations for credible data. ACS EST Air. 1(10):1203–1214. https://doi.org/10.1021/acsestair.4c00125

Barkjohn K, Holder A, Clements A, Fredrick S, Evans R. 2021. Sensor data cleaning and correction: application on the AirNow fire and smoke map. Presented at the American Association for Aerosol Research Conference; Albuquerque, NM, Oct 18-22, 2021. https://cfpub.epa.gov/si/si_public_record_report.cfm?dirEntryId=353088&Lab=CEMM

Barkjohn KK, Gantt B, Clements AL. 2021. Development and application of a United States-wide correction for $PM_{2.5}$ data collected with the PurpleAir sensor. Atmos Meas Tech. 14(6):4617–4637. https://doi.org/10.5194/amt-14-4617-2021

Barkjohn KK, Holder AL, Frederick SG, Clements AL. 2022. Correction and accuracy of PurpleAir $PM_{2.5}$ measurements for extreme wildfire smoke. Sensors. 22:9669. https://doi.org/10.3390/s22249669

Bartolome C et al. 2019. Laboratory and numerical modeling of the formation of superfog from wildland fires. Fire Saf J. 106:94–104. https://doi.org/10.1016/j.firesaf.2019.04.009

Brey SJ: Ruminski M, Atwood SA, Fischer EV. 2018. Connecting smoke plumes to sources using hazard mapping system (HMS) smoke and fire location data over North America. Atmos Chem Phys. 18(3):1745–1761. https://doi.org/10.5194/acp-18-1745-2018

Carreras-Sospedra M et al. 2024. Air quality and health impacts of the 2020 wildfires in California. Fire Ecol. 20(1):6. https://doi.org/10.1186/s42408-023-00234-y

Chow FK et al. 2022. High-resolution smoke forecasting for the 2018 Camp Fire in California. Bull Amer Meteor Soc. 103(6):E1531–E1552. https://doi.org/10.1175/BAMS-D-20-0329.1

Christian TJ et al. 2004. Comprehensive laboratory measurements of biomass-burning emissions: 2. First intercomparison of open-path FTIR, PTR-MS, and GC- MS/FID/ECD. J Geophys Res Atmos. 109(D2):D02311. https://doi.org/10.1029/2003JD003874

DeCarlo PF, Slowik JG, Worsnop DR, Davidovits P, Jimenez JL. 2012. Particle morphology and density characterization by combined mobility and aerodynamic diameter measurements. Part 1: theory. Aerosol Sci Tech. 38(12):1185–1205. https://doi.org/10.1080/027868290903907

Delp WW, Singer BC. 2020. Wildfire smoke adjustment factors for low-cost and professional $PM_{2.5}$ monitors with optical sensors. Sensors. 20(13):3683. https://doi.org/10.3390/s20133683

Drinovec L et al. 2015. The "dual-spot" aethalometer: an improved measurement of aerosol black carbon with real-time loading compensation. Atmos Meas Tech. 8(5):1965–1979. https://doi.org/10.5194/amt-8-1965-2015

Gupta P et al. 2018. Impact of California fires on local and regional air quality: the role of a low-cost sensor network and satellite observations. GeoHealth. 2(6):172–181. https://doi.org/10.1029/2018GH000136

Hagan DH, Kroll JH. 2020. Assessing the accuracy of low-cost optical particle sensors using a physics-based approach. Atmos Meas Tech. 13(11):6343–6355. https://doi.org/10.5194/amt-13-6343-2020

Hall ES, Beaver MR, Long RW, Vanderpool RW. 2012. EPA's reference and equivalent methods research program: supporting NAAQS implementation through research, development, and analysis. EM Air Waste Manag Assoc. 5:8–12.

He ML, Kuerbanjiang N, Dhaniyala S. 2020. Performance characteristics of the low-cost Plantower PMS optical sensor. Aerosol Sci Tech. 54(2):232–241. https://doi.org/10.1080/02786826.2019.1696015

Henn SA et al. 2019. Carbon monoxide exposures among W. S. wildland fire fighters by work, fire, and environmental characteristics and conditions. J Occup Environ Hyg. 16(12):793–803. https://doi.org/10.1080/15459624.2019.1670833

Holden ZA et al. 2018. Decreasing fire season precipitation increased recent western U.S. forest wildfire activity. Proc Natl Acad Sci USA. 115(36):E8349–E8357. https://doi.org/10.1073/pnas.1802316115

Holder AL et al. 2020. Field evaluation of low-cost particulate matter sensors for measuring wildfire smoke. Sensors. 20(17):4796. https://doi.org/10.3390/s20174796

Jiao A, Headon K, Han T, Umer W, Wu J. 2024. Associations between short-term exposure to wildfire particulate matter and respiratory outcomes: a systematic review. Sci Total Environ. 907:168134. https://doi.org/10.1016/j.scitotenv.2023.168134

Kirchstetter TW, Novakov T, Hobbs PV. 2004. Evidence that the spectral dependence of light absorption by aerosols is affected by organic carbon. J Geophys Res. 109(D21):D21208. https://doi.org/10.1029/2004JD004999

Kitzberger T, Brown PM, Heyerdahl EK, Swetnam TW, Veblen TT. 2007. Contingent Pacific-Atlantic Ocean influence on multicentury wildfire synchrony over western North America. Proc Natl Acad Sci USA. 104(2):543–548. https://doi.org/10.1073/pnas.0606078104

Krug J et al. 2021. Evaluation of small form factor, filter-based $PM_{2.5}$ samplers for temporary non-regulatory monitoring during wildland fire smoke events. Atmos Environ. 265:118718. https://doi.org/10.1016/j.atmosenv.2021.118718

Landis M et al. 2021a. The U.S. EPA wildland fire sensor challenge: Performance and evaluation of solver submitted multi-pollutant sensor systems. Atmos Environ. 247:118165. https://doi.org/10.1016/j.atmosenv.2020.118165

Landis M et al. 2021b. Chapter 4: air quality monitoring of wildland fire smoke in air quality monitoring of wildland fire smoke in comparative assessment of the impacts of prescribed fire versus wildfire (CAIF): a case study in the western U.S., U.S. EPA, EPA/600/R-21/197. U.S. Environmental Protection Agency. https://cfpub.epa.gov/si/si_public_record_Report.cfm?Lab=CPHEA&dirEntryId=352824

Landis MS et al. 2018. The impact of the 2016 Fort McMurray Horse River wildfire on ambient air pollution levels in the Athabasca Oil Sands Region, Alberta, Canada. Sci Total Environ. 618:1665–1676. https://doi.org/10.1016/j.scitotenv.2017.10.008

Laskin A, Laskin J, Nizkorodov SA. 2015. Chemistry of atmospheric brown carbon. Chem Rev. 115(10):4335–4382. https://doi.org/10.1021/cr5006167

Liang L. 2021. Calibrating low-cost sensors for ambient air monitoring: techniques, trends, and challenges. Environ Res. 197:111163. https://doi.org/10.1016/j.envres.2021,111163

Littell JS, McKenzie D, Peterson DL, Westerling AL. 2009. Climate and wildfire area burned in Western U.S. ecoprovinces, 1916-2003. Ecol Appl. 19(4):1003–1021. https://doi.org/10.1890/07-1183.1

Long R et al. 2021. Comparison of ozone measurement methods in biomass burning smoke: an evaluation under field and laboratory conditions. Atmos Meas Tech. 14 (3):1783–1800. https://doi.org/10.5194/amt-14-1783-2021

Long RW et al. 2023. Summary of $PM_{2.5}$ measurement artifacts associated with the Teledyne T640 PM mass monitor under controlled chamber experimental conditions using polydisperse ammonium sulfate aerosols and biomass smoke. J Air Waste Manag Assoc. 73(4):295–312. https://doi.org/10.1080/10962247.2023.2171156

Mallia DV, Lin JC, Urbanski S, Ehleringer J, Nehrkorn T. 2015. Impacts of upwind wildfire emissions on CO, CO_2, and $PM_{2.5}$ concentrations in Salt Lake City, Utah. J Geophys Res. 120(1):147–166. https://doi.org/10.1002/2014JD022472

Malm WC et al. 2024. The dependence of PurpleAir sensors on particle size distribution and meteorology and their relationship to integrating nephelometers. Atmos Environ. 337:120757. https://doi.org/10.1016/j.atmosenv.2024.120757

Materna BL, Jones JR, Sutton PM, Rothman N, Harrison RJ. 1992. Occupational exposures in California wildland fire fighting. Am Ind Hyg Assoc J. 53(1):69–76. https://doi.org/10.1080/15298669291359311

McCaffrey SM et al. 2022. Social considerations: Health, economics, and risk communication. In: Peterson D, McCaffrey S, Patel-Weynand T, editors. Wildland fire smoke in the United States: a scientific assessment. Springer Nature Switzerland AG p 199–237. Chapter 7. https://doi.org/10.1007/978-3-030-87045-4_7

Mehadi A et al. 2020. Laboratory and field evaluation of real-time and near real-time $PM_{2.5}$ smoke monitors. J Air Waste Manag Assoc. 70(2):158–179. https://doi.org/10.1080/10962247.2019.1654036

Navarro K. 2020. Working in smoke: Wildfire impacts on the health of firefighters and outdoor workers and mitigation strategies. Clin Chest Med. 41(4):763–769. https://doi.org/10.1016/j.ccm.2020.08.017

Omaye ST. 2002. Metabolic modulation of carbon monoxide toxicity. Toxicology. 180(2):139–150. https://doi.org/10.1016/S0300-483X(02)00387-6

Parks SA, Abatzoglou JT. 2020. Warmer and drier fire seasons contribute to increases in area burned at high severity in western U.S. forests from 1985 to 2017. Geophys Res Lett. 47(22):e2020Gl089858. https://doi.org/10.1029/2020Gl089858

Peters TM, Leith D. 2003. Concentration measurement and counting efficiency of the aerodynamic particle sizer 3321. J Aerosol Sci. 34(5):627–634. https://doi.org/10.1016/S0021-8502(03)00030-2

Peters TM, Ott D, O'Shaughnessy PT. 2006. Comparison of the GRIMM 1.108 and 1.109 portable aerosol spectrometer to the TSI 3321 aerodynamic particle sizer for dry particles. Ann Occup Hyg. 50:843–850. https://doi.org/10.1093/annhyg/mel067

Reid CE et al. 2016. Critical review of health impacts of wildfire smoke exposure. Environ Health Persp. 124 (9):1334–1343. https://doi.org/10.1289/ehp.1409277

Reuters. 2023. Louisiana 'super fog' kills 7 in massive car pile-up. [accessed 2024 May 22]. https://www.reuters.com/world/us/louisiana-super-fog-kills-7-massive-car-pile-up-2023-10-24/

Rueda EM, Carter E, L'Orange C, Quinn C, Volckens J. 2023. Size-resolved field performance of low-cost sensors for particulate matter air pollution. Environ Sci Technol Lett. 10(3):447–253. https://doi.org/10.1021/acs.estlett.3c00030

Sablan O et al. 2024. Quantifying prescribed fire smoke exposure using low-cost sensors and satellites: springtime burning in Eastern Kansas. GeoHealth. 8(4):e2023GH000982. https://doi.org/10.1029/2023GH000982

Saleh RC et al. 2013. Absorptivity of brown carbon in fresh and photo-chemically aged biomass-burning emissions. Atmos Chem Phys. 13(15):7683–7693. https://doi.org/10.5194/acp-13-7683-2013

Semmens EO et al. 2021. Carbon monoxide exposures in wildland firefighters in the United States and targets for exposure reduction. J Expo Sci Environ Epidemiol. 31(5):923–929. https://doi.org/10.1038/s41370-021-00371-z

Trent A. 2003. Laboratory evaluation of new real-time smoke particulate monitors. Tech. Rep. 0325–2834–MTDC. U.S. Department of Agriculture, Forest Service, Missoula Technology and Development Center. 14. [accessed 2025 Apr 18]. https://www.fs.usda.gov/t-d/pubs/pdfpubs/pdf03252834/pdf03252834dpi300.pdf

Urbanski SP. 2013. Combustion efficiency and emission factors for wildfire-season fires in mixed conifer forests of the Northern Rocky Mountains, U.S. Atmos Chem Phys. 13(14):7241–7262. https://doi.org/10.5194/acp-13-7241-2013

Urbanski SP, O'Neill SM, Holder AL, Green SA, Graw RL. 2022. Emissions. In: Peterson D, McCaffrey S, Patel-Weynand T, editors. Wildland fire smoke in the United States: a scientific assessment. Springer Nature Switzerland AG; pp 121–165. Chapter 5. https://doi.org/10.1007/978-3-030-87045-4_5

[U.S. EPA] U.S. Environmental Protection Agency. 1986. Thermo Environmental Instruments, Inc./Thermo Electron, Thermo Scientific models 43A, 43B, 43C, 43C-TLE, 43i, 43i-TLE, 43iQ, 43iQTL SO2 analyzer automated equivalent method: EQSA-0486-060. Federal Register: Vol. 51. 12390. [accessed 1986 Apr 10]. https://www.epa.gov/amtic/air-monitoring-methods-criteria-pollutants

[U.S. EPA] U.S. Environmental Protection Agency. 2009. Thermo Scientific TEOM® 1405-DF dichotomous ambient particulate monitor with FDMS®. Automated equivalent method: EQPM-0609-181. Federal Register: Vol. 74. 28696. [accessed 2009 Jun 17]. https://www.epa.gov/amtic/air-monitoring-methods-criteria-pollutants

[U.S. EPA] U.S. Environmental Protection Agency. 2014a. Teledyne advanced pollution instrumentation model T500U nitrogen dioxide analyzer automated equivalent method: EQNA-0514-212. Federal register: Vol.79. 34734–34735. [accessed 2014 Jun 18]. https://www.epa.gov/amtic/air-monitoring-methods-criteria-pollutants

[U.S. EPA] U.S. Environmental Protection Agency. 2014b. Tisch environmental model TE-Wilbur10 particulate

sampler – PM_{10} manual reference method: RFPS-0714-216. Federal Register: Vol. 79. 49307. [accessed 2014 Sep 20]. https://www.epa.gov/amtic/air-monitoring-methods-criteria-pollutants

[U.S. EPA] U.S. Environmental Protection Agency. 2014c. Tisch environmental model TE-Wilbur2.5 particulate sampler – $PM_{2.5}$ manual reference method: RFPS-1014-219. Federal register: Vol. 79. 65392. [accessed 2014 Nov 4]. https://www.epa.gov/amtic/air-monitoring-methods-criteria-pollutants

[U.S. EPA] U.S. Environmental Protection Agency. 2015. MetOne instruments BAM-1020 beta attenuation mass monitor – $PM_{2.5}$ FEM configuration. Automated equivalent method: EQPM-0715-266. Federal register: Vol. 80. 51802. [accessed 2015 Aug 26]. https://www.epa.gov/amtic/air-monitoring-methods-criteria-pollutants

[U.S. EPA] U.S. Environmental Protection Agency. 2016. Quality assurance guidance document 2.12: monitoring PM2.5 in ambient Air using designated reference or class I equivalent methods. Office Air Qual Plann Stand Air Qual Anal Div Publ No. EPA-454/B-16-001. https://www3.epa.gov/ttnamti1/files/ambient/pm25/qa/m212.pdf

[U.S. EPA] U.S. Environmental Protection Agency. 2019. Integrated science assessment for particulate matter. Office of Research and Development, Center for Public Health & Environmental Assessment, Research Triangle Park. EPA/600/R–19/188. https://assessments.epa.gov/risk/document/&deid=347534

[U.S. EPA] U.S. Environmental Protection Agency. 2022. Ambilabs model 2WIN $PM_{2.5}$ FEM monitor. Automatic equivalent method: EQPM-0922-260. Federal register: Vol. 87. 65203–65204. [accessed 2022 Oct 28]. https://www.epa.gov/amtic/air-monitoring-methods-criteria-pollutants

[U.S. EPA] U.S. Environmental Protection Agency. 2025. Air data: air quality data collected at outdoor monitors across the US. https://www.epa.gov/outdoor-air-quality-data

[U.S. GAO] U.S. Government Accountability Office. 2020. Report to congressional requestors, air pollution: opportunities to better sustain and modernize the National Air Quality Monitoring System. GAO-21-38. https://www.gao.gov/assets/720/710612.pdf

[USFS] United States Forest Service. 2020. Interag wildland fire Air qual response program. Air Resource Advisors; [accessed 2026 Feb 1]. https://www.wildlandfiresmoke.net/ara

Volckens J, Peters TM. 2005. Counting and particle transmission efficiency of the aerodynamic particle sizer. J Aerosol Sci. 36(12):1400–1408. https://doi.org/10.1016/j.jaerosci.2005.03.009

Vu BN et al. 2022. Application of geostationary satellite and high-resolution meteorology data in estimating hourly $PM_{2.5}$ levels during the Camp Fire episode in California. Remote Sens Environ. 271(112890):112890–114257. https://doi.org/10.1016/j.rse.2022.112890

Wallace L, Bi JZ, Ott WR, Sarnat J, Liu Y. 2021. Calibration of low-cost PurpleAir outdoor monitors using an improved method of calculating $PM_{2.5}$. Atmos Environ. 256:118432. https://doi.org/10.1016/j.atmosenv.2021.118432

Wallace L, Zhao TK, Klepeis NE. 2022. Calibration of PurpleAir PA-I and PA-II monitors using daily mean $PM_{2.5}$ concentrations measured in California, Washington, and Oregon from 2017 to 2021. Sensors. 22(13):4741. https://doi.org/10.3390/s22134741

Westerling AL. 2016. Increasing western U.S. forest wildfire activity: sensitivity to changes in the timing of spring. Philos Trans R Soc. 371(1707):B371. https://doi.org/10.1098/rstb.2016.0373

Whitehill A, Long R, Habel A, Landis MS. 2022. Evaluation of Cairpol and Aeroqual sensors in wildland fire plumes. Atmosphere. 13(6):877. https://doi.org/10.3390/atmos13060877

Wilmot TY, Mallia DV, Hallar AG, Lin JC. 2022. Wildfire activity is driving summertime air quality degradation across the western U.S.: a model-based attribution to smoke source regions. Environ Res Lett. 17(11):114014. https://doi.org/10.1088/1748-9326/ac9a5d

Wu Y, Arapi A, Huang J, Gross B, Moshary F. 2018. Intra-continental wildfire smoke transport and impact on local air quality observed by ground-based and satellite remote sensing in New York City. Atmos Environ. 187:266–281. https://doi.org/10.1016/j.atmosenv.2018.06.006

Yang H, Ruiz-Suarez S, Reich BJ, Guan Y, Rappold AG. 2023. A data-fusion approach to assessing the contribution of wildland fire smoke to fine particulate matter in California. Remote Sens. 15(17):4246. https://doi.org/10.3390/rs15174246

Zheng T et al. 2018. Field evaluation of low-cost particulate matter sensors in high- and low-concentration environments. Atmos Meas Tech. 11(8):4823–4846. https://doi.org/10.5194/amt-11-4823-2018