

Air Quality in Sub-Saharan Africa: Measurement Issues and Empirical Challenges

Emilia Tjernström¹

¹Department of Economics, Macquarie Business School, Macquarie University, Sydney, New South Wales, Australia; email: emilia.tjernstrom@mq.edu.au

Annu. Rev. Resour. Econ. 2026. 18:1–22
Copyright © 2026 by the author(s).
All rights reserved

Author accepted manuscript, accepted for publication
in the Annual Review of Resource Economics, Vol. 18
(2026). Not copyedited or typeset by Annual Reviews;
not the version of record.

Keywords

sub-Saharan Africa, air pollution, measurement error, particulate matter

Abstract

Air pollution imposes large and uneven health and productivity costs in sub-Saharan Africa, yet evidence remains limited by severe data constraints. This review synthesizes what economists know about particulate matter exposure in the region, with a focus on measurement, sources, and implications for applied research. I review ground-based, satellite, reanalysis, and hybrid data products, emphasizing how spatially structured measurement error shapes empirical identification. I then assess evidence on health, agriculture, labor and human capital, highlighting both credible designs and remaining gaps. The review concludes by outlining research strategies that integrate exposure uncertainty into applied economic analysis.

1. INTRODUCTION

Air pollution represents one of the most significant unpriced externalities in sub-Saharan Africa (SSA). Exposure to fine particulate matter (PM_{2.5}) contributed to around 1.2 million deaths in Africa in 2021 (Health Effects Institute 2024).¹ Epidemiological evidence links long-term PM_{2.5} exposure to higher risks of cardiovascular disease, stroke, lung cancer, and chronic respiratory illness (Newell et al. 2017, Cohen et al. 2017, Fisher et al. 2021); more recent work adds adverse birth outcomes to this list (Murray et al. 2020).

The costs extend well beyond health: even moderate air pollution exposure lowers labor productivity, weakens cognitive performance, and distorts decision-making in ways that hinder economic development (Aguilar-Gomez et al. 2022, Hoffmann & Rud 2024, Duncan et al. 2025, Ma et al. 2025). The burden of air pollution falls disproportionately on the region's poor—of the 716 million people who both live on less than \$1.90 per day and face unsafe annual PM_{2.5}, nearly 60 percent live in sub-Saharan Africa (Rentschler & Leonova 2023).

The pollution sources that drive PM_{2.5} in sub-Saharan Africa differ from those in both industrialized countries and other high-pollution developing regions (McDuffie et al. 2021). Household combustion of solid fuels and kerosene is the dominant source across much of SSA (Fisher et al. 2021).² Saharan dust and crop burning add large but seasonally variable contributions, while vehicular exhaust, diesel generators, and open waste burning are growing sources in rapidly expanding cities (Chowdhury et al. 2023, Farquharson et al. 2018, Gordon et al. 2023, McDuffie et al. 2021). As a result, many areas without industrial activity still face high PM_{2.5} exposure. The resulting mix varies with rainfall, wind patterns, and agricultural cycles, which makes it harder to attribute pollution to specific sources and complicates both policy design and empirical identification.

The region was long viewed as relatively unaffected by air pollution—perhaps precisely because it lacks the heavy industry, coal-fired power, and dense urbanization associated with air quality crises in South and East Asia. As a result, neither domestic governments nor the global development community have invested heavily in monitoring infrastructure (Bililign et al. 2024, Health Effects Institute 2022), and very few countries in the region have any legislated ambient air quality standards (United Nations Environment Programme 2021). This has shaped the data environment in ways that constrain both research and policy (Amegah & Agyei-Mensah 2017, Mead et al. 2023). Monitoring infrastructure is not merely a measurement issue: Jha & La Nauze (2022) find that public data dissemination from the installation of US Embassy air-quality monitors in developing-country cities reduced PM_{2.5} by 2–4 μg/m³, likely by making local pollution levels salient and spurring policy responses.

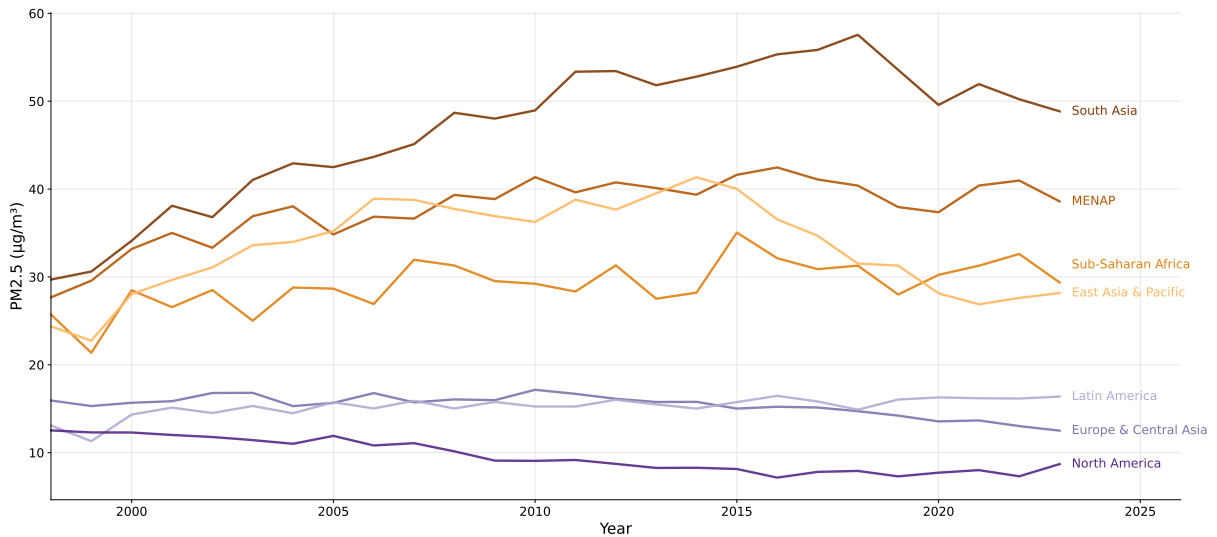
Sub-Saharan Africa's pollution trajectory diverges from other developing regions: while East Asia and South Asia have reduced PM_{2.5} levels in recent years, African population-weighted exposures are creeping upward. **Figure 1** shows population-weighted PM_{2.5} by World Bank region over 1998–2023, using data from Shen et al. (2024). Sub-Saharan Africa's PM_{2.5} concentrations have risen from around 25 μg/m³ in 1998 to around 30 μg/m³ in 2023, roughly six times the WHO annual guideline of 5 μg/m³ (World Health Organization 2021).

As **Figure 2** shows, the regional average hides substantial sub-regional and country-level heterogeneity. Western Africa faces the highest concentrations: several countries face pollution comparable to South Asia and the Middle East, with Niger's annual averages frequently exceeding 50 μg/m³. Southern Africa is much cleaner on average, averaging below 25 μg/m³, but Lesotho is an outlier with levels often in excess of 50 μg/m³. Eastern Africa's PM_{2.5} levels increased early in the period, but have since stabilized around 20 μg/m³, though the variation within the region is stark. Middle Africa tracks the regional average most closely.

¹Air pollution encompasses multiple pollutants—fine and coarse particulate matter (PM_{2.5} and PM₁₀), nitrogen dioxide (NO₂), sulfur dioxide (SO₂), ozone (O₃), and carbon monoxide (CO). In sub-Saharan Africa, PM_{2.5} accounts for most of the health burden and is easier to measure consistently across space and time. NO₂ concentrations tend to be highest near busy roads and dense urban cores, while SO₂ matters mainly near power plants and refineries; both receive even less systematic monitoring than PM_{2.5} because ambient levels are often low outside specific hot spots and health thresholds remain poorly established for African populations. Other pollutants also harm health, particularly among children, older adults, and people with pre-existing conditions, but this review focuses on PM_{2.5}.

²About one-fifth of the world's solid-fuel users live in SSA, a share that is projected to grow as other regions transition to cleaner fuels (Jeuland et al. 2015).

Figure 1: Population-weighted annual PM_{2.5} by World Bank region, 1998–2023



Source: Shen et al. (2024); population weights from the Gridded Population of the World v4 (GPWv4). Each line shows a regional population-weighted average, weighting each country's PM_{2.5} by its share of regional population.

The challenge of air pollution in sub-Saharan Africa is compounded by the scarcity and structure of the data available to measure it. While high-income countries use dense ground-based monitoring networks to inform policy, the vast majority of sub-Saharan African nations lack a single reference-grade monitor (see **Figure 3**). Africa averages just 0.03 PM_{2.5} monitors per million inhabitants, and the population-weighted distance to the nearest monitor is roughly 500 km (Martin et al. 2019). As a consequence, most evidence on PM_{2.5} levels and trends comes from satellite- and model-based hybrid products rather than dense ground networks of high-quality monitors. The reliance on these data sources introduces measurement error that is often non-classical and spatially structured in ways that can bias empirical estimates.

This review aims to provide guidance for applied economists with an interest in studying air quality in sub-Saharan Africa. I first outline the measurement processes that underlie the most commonly used data sources, their biases and limitations, and clarify the implications for empirical research designs. I then review existing empirical estimates of pollution's effects on health, agriculture, and human capital through the lens of this measurement environment, identifying where non-classical error structures interact with identification strategies in ways that may affect the interpretation of existing results.

2. DATA AND MEASUREMENT

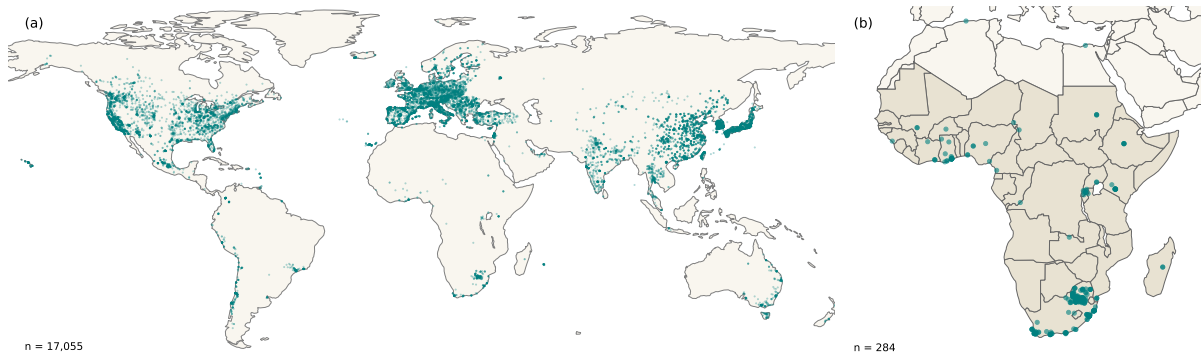
Air quality data in SSA fall into two broad categories. Ground-based instruments measure pollution directly but are sparse and unevenly distributed. Among these, reference-grade monitors provide the most reliable measurements, while low-cost sensors expand spatial coverage at the expense of accuracy. Model-based products fill the spatial gaps, but introduce measurement error whose structure depends on the product's design and the local environment. These error structures shape what empirical designs can credibly identify. Below, I present a simple framework to structure our discussion of how measurement error in different data types likely affects empirical estimates. I then review the most common data products through this lens.

Figure 2: Population-weighted PM_{2.5} in sub-Saharan Africa by sub-region, 1998–2023



Source: Shen et al. (2024). Sub-Saharan African countries grouped by UN M49 Region Codes (United Nations Statistics Division 2024). Thin lines represent individual countries and bold lines show sub-regional population-weighted averages. Labels appear at each country’s final data point. Sudan is excluded because the UN M49 standard classifies it as Northern Africa.

Figure 3: Publicly accessible reference monitors. (a) World, (b) Sub-Saharan Africa.



Monitor locations compiled from three sources, each queried in April 2026: the OpenAQ API v3 supplies worldwide reference-grade monitors (OpenAQ 2026). The AirQo device API supplies beta-attenuation (BAM) reference monitors (AirQo 2026). The World Air Quality Index Project supplies stations within a sub-Saharan Africa bounding box (World Air Quality Index Project 2026). The map includes inactive sites and those that have previously shared data with OpenAQ but that no longer report. In panel b, shaded countries are those classified as sub-Saharan Africa by the World Bank.

2.1. A Framework for Exposure Measurement Error

Assume that we want to estimate the effect of pollution exposure on an outcome Y_{it} :

$$Y_{it} = \beta X_{it}^* + \mathbf{Z}'_{it}\gamma + \varepsilon_{it}, \quad 1.$$

where X_{it}^* is the true PM_{2.5} exposure for unit i at time t and \mathbf{Z}_{it} is a vector of controls. However, we do not observe true pollution exposure. Instead, our data measure exposure with error: $X_{it} = X_{it}^* + \eta_{it}$, where η_{it} is measurement error. Rewriting 1. in terms of observed exposure makes η_{it} part of the composite error, so OLS will generally be biased. The direction and magnitude of the resulting bias depends on how η relates to X^* , \mathbf{Z} , and ε .³

The simplest case is classical measurement error, where η is independent of both X^* and ε . OLS then yields:

$$\text{plim } \hat{\beta}_{\text{OLS}} = \beta \cdot \underbrace{\frac{\sigma_{X^*}^2}{\sigma_{X^*}^2 + \sigma_{\eta}^2}}_{\lambda}, \quad 2.$$

where our estimate is attenuated by the reliability ratio, λ (Fuller 1987). This benchmark is useful but, as we will see, the classical assumptions are unlikely to hold for pollution data in SSA.⁴

Maintaining the assumption that true exposure is exogenous (i.e., $\text{Cov}(X^*, \varepsilon) = 0$) but allowing η to correlate with true exposure or with the structural error, we instead get:

$$\begin{aligned} \text{plim } \hat{\beta}_{\text{OLS}} &= \beta \cdot \frac{\sigma_{X^*}^2 + \text{Cov}(\eta, X^*)}{\sigma_{X^*}^2 + \sigma_{\eta}^2 + 2 \text{Cov}(\eta, X^*)} + \frac{\text{Cov}(\eta, \varepsilon)}{\sigma_{X^*}^2 + \sigma_{\eta}^2 + 2 \text{Cov}(\eta, X^*)} \\ &= \beta \cdot \underbrace{\frac{\text{Cov}(X, X^*)}{\text{Var}(X)}}_{\lambda^*} + \frac{\text{Cov}(\eta, \varepsilon)}{\text{Var}(X)}. \end{aligned} \quad 3.$$

The bias decomposes into a multiplicative factor, λ^* , and an additive term driven by $\text{Cov}(\eta, \varepsilon)$. λ^* captures how variation in observed exposure maps into true exposure. Unlike the classical reliability ratio, λ^* depends on both noise variance and $\text{Cov}(\eta, X^*)$, and need not lie between zero and one. In practice, $\text{Cov}(\eta, X^*) \neq 0$ whenever measurement error varies with true pollution levels—for example, if satellite retrievals systematically overstate concentrations in clean areas and understate them in polluted areas.

The second term captures the endogeneity that arises from having η in the composite error term. If η correlates with unobserved determinants of the outcome, this term can reinforce or offset attenuation, independently of λ^* . For example, satellite-derived pollution estimates may systematically understate exposure in dense informal urban areas, where complex surface conditions degrade retrieval accuracy. If residents of these areas also face worse outcomes for reasons we cannot fully control for, then η and ε covary and this term biases our estimates.

For studies linking pollution to health, $\text{Cov}(\eta, \varepsilon)$ could be non-zero even if ambient concentrations were measured perfectly, because they are only a proxy for personal exposure. This review focuses on outdoor PM_{2.5}, but health outcomes depend on the total dose that a person inhales, which includes indoor exposure from cooking with solid fuels (Fisher et al. 2021). This can be a significant source of bias in SSA, since household combustion of solid fuels is an important source (see Jeuland et al. 2015, for a comprehensive treatment of household air pollution). When researchers use outdoor PM_{2.5} to proxy for total exposure, η includes the

³For simplicity, the bias expressions in this section suppress \mathbf{Z} and present the bivariate case. With controls, the same logic applies to partial regression coefficients, with population covariances replaced by their partialled-out analogues.

⁴The bias expressions in this section assume a linear model. In nonlinear specifications, measurement error introduces bias even under classical assumptions, with the direction depending on the curvature of the regression function (Chesher 1991). For broader surveys, see Bound et al. (2001) and Schennach (2016).

indoor-outdoor differential, which correlates with cooking fuel, housing quality, and income—all independent predictors of health outcomes.

The bivariate notation in Equation 3 obscures two important ways in which measurement error interacts with covariates and co-pollutants. First, converting satellite signals or model outputs to pollution estimates can introduce spatially varying bias—especially where ground data for calibration are scarce.⁵ If measurement error covaries with observables (i.e., $\text{Cov}(\eta_{it}, \mathbf{Z}_{it}) \neq 0$), then partialling out those controls—for example, by controlling for geographic or climatic variables that also predict measurement accuracy—can change both the effective noise-to-signal ratio and the residualized covariance structure, making it difficult to sign the direction of bias without further assumptions.⁶ Second, controlling for co-pollutants using noisy proxies creates a distinct problem: measurement error in one pollutant’s proxy can bias the coefficient on another—even one that is measured without error. Kluger et al. (2024) show that when pollutants are positively correlated and all have a non-positive effect on the outcome, measurement error in one proxy biases estimates of the others away from zero.⁷

Another error structure, called Berkson error (Berkson 1950, Carroll et al. 2006), arises when researchers assign gridded pollution values to survey respondents, for example when matching DHS clusters or LSMS households to gridded pollution measures. This approach assigns all individuals in grid cell c to a single pollution value, X_c , but the true individual exposure X_i^* varies within the cell:

$$X_i^* = X_c + u_i, \quad 4.$$

where u_i captures within-cell deviations. Unlike the error structures in Equations 2–3, the observed value X_c is now the conditional mean, and true exposure varies around it. If within-cell deviations are independent of the cell mean ($u \perp X_c$), Berkson error does not bias the OLS slope coefficient in a linear model, though it increases residual variance (Carroll et al. 2006).⁸ This no-bias result breaks down in nonlinear models: when the dose-response function is concave, Berkson error flattens the estimated curve, and under heterogeneous treatment effects, the bias depends on the joint distribution of exposure and effect heterogeneity (Carroll et al. 2006, Blundell et al. 2022). In practice, the error in gridded products is rarely pure Berkson. Szpiro et al. (2011) show that predicted exposures combine a Berkson component with a “classical-like” component, and that the latter can bias OLS (see also Sheppard et al. 2012).⁹

Beyond the structure of the error itself, estimation choices shape which components of it survive into the identifying variation. Location and time fixed effects, in particular, can either help or hurt, depending on which component of the error they absorb. We can decompose η_{it} into location-specific, period-specific, and residual components:

$$\eta_{it} = \bar{\eta}_i + \bar{\eta}_{\cdot t} + \tilde{\eta}_{it}, \quad 5.$$

where $\bar{\eta}_i$ and $\bar{\eta}_{\cdot t}$ are absorbed by location and time fixed effects, respectively. If the non-classical correlation between η and X^* is primarily spatial—for example, because satellite retrieval accuracy varies across locations but is stable over time—location fixed effects absorb $\bar{\eta}_i$ and eliminate the bias from that correlation. But fixed

⁵For example, Giordano et al. (2024) show that MERRA-2, a widely used reanalysis product described in Section 2.3.2, overestimates dust-driven $\text{PM}_{2.5}$ in West Africa, likely because dust scatters light differently than the aerosols that dominate where the model is calibrated.

⁶Correlation between η and \mathbf{Z} does not by itself imply bias in the partial coefficient; rather, it changes the partialled-out covariances that drive both bias channels in Equation 3. See Bound et al. (2001) for the multivariate case.

⁷Kluger et al. (2024)’s main results assume classical, uncorrelated measurement error across pollutants, though their simulations suggest that the direction of bias holds under weaker conditions. They also note that the nondifferential measurement error assumption underlying their framework may not hold for satellite- or model-based pollution proxies.

⁸If the gridded product estimates the cell mean with error, that error acts as a classical component and can bias OLS directly.

⁹The component is “classical-like” because parameter uncertainty in the exposure model increases the variability of X relative to the true conditional mean, introducing variation that is not informative for Y —analogous to standard classical error (Carroll et al. 2006). However, this component is neither independent of Y nor independent across locations, so standard correction techniques like regression calibration do not directly apply (Szpiro et al. 2011).

effects also remove the corresponding persistent components of true exposure, leaving identification to rest on within-unit, within-period variation in X_{it} , where the residual error $\tilde{\eta}_{it}$ can account for a large share of the remaining signal. As Griliches & Hausman (1986) show, this can worsen the noise-to-signal ratio, so the net effect of fixed effects on total bias is ambiguous. Temporal aggregation compounds the residual problem, because averaging over time smooths precisely the within-unit variation that identification relies on.

The residual $\tilde{\eta}_{it}$ creates another problem when the identifying variation itself changes the particle mix that sensors or models are measuring. Many pollution data products measure different particle types with different biases (Sections 2.2.2–2.3.2). If the shock or instrument that a researcher exploits also changes the particle mix, η_{it} shifts at the same time as X_{it}^* , generating $\text{Cov}(\tilde{\eta}_{it}, \tilde{X}_{it}^*) \neq 0$ in precisely the residual variation that fixed effects leave for identification. For example, a researcher using COVID-19 lockdowns as a temporal shock to estimate the effect of pollution on a health outcome would face this problem: lockdowns reduce traffic emissions while leaving household cooking emissions largely unchanged, shifting aerosol composition and with it the measurement error properties of the exposure data.

The same logic applies to instrumental-variable strategies. Deryugina et al. (2019) use wind direction to instrument for daily $\text{PM}_{2.5}$ and estimate its effect on elderly mortality in the US. Adapting this approach to SSA is complicated by the fact that wind direction also determines aerosol composition: northerly Harmattan winds carry mineral dust, while other wind patterns leave locally generated traffic and biomass aerosols dominant. If wind direction predicts measurement error, IV no longer solves the problem it was designed to fix.

The next section reviews common data products, considering the error structures that each introduces and the implications for estimation and interpretation.

2.2. Ground-Based Measurements

Ground-based instruments sample air at a fixed location and measure pollutant concentrations on site. They span two classes that trade accuracy against cost and coverage: reference-grade monitors and low-cost sensors.

2.2.1. Reference Monitors. Reference-grade air quality monitors provide direct measurements of air pollutants using methods that meet regulatory standards for accuracy and precision. These instruments are used to establish air quality standards for compliance purposes and serve as ground truth for validating other measurement methods. However, they require stable power, technical expertise for calibration and maintenance, and a reliable supply chain for replacement parts.

These infrastructure barriers may help explain why sub-Saharan Africa has an extremely small fraction of the world's reference monitors. **Figure 3** maps the global distribution of reference-grade monitors compiled from three publicly accessible sources: the [OpenAQ platform](#), [AirQo](#), and the [World Air Quality Index Project \(AQICN\)](#). Of over 17,000 reference monitors worldwide, only 285 (1.7%) are in sub-Saharan Africa, and South Africa accounts for around 230 of these. These counts include inactive sites, so they represent an upper bound on functional monitoring capacity. Data availability has deteriorated in recent years due to two main events: in March 2024, South Africa's national monitoring network stopped sharing data through OpenAQ. A year later, in March 2025, the U.S. State Department's embassy monitoring network ceased public data sharing ([Associated Press 2025](#)). Embassy monitors had been among the few sources of openly accessible, reference-grade air quality data in SSA. After these two events, only 15 monitors on OpenAQ have shared data, with only one monitor on OpenAQ reporting data during the first quarter of 2026.

Among the operational $\text{PM}_{2.5}$ monitors in SSA on OpenAQ, the data that do exist are sparse. I define completeness as the share of potential hours with a recorded observation, measured from a monitor's first observation through its last, over a ten-year window (April 2016 to April 2026).¹⁰ Median completeness across

¹⁰The denominator begins at the start of this window, or from a monitor's first recorded observation, whichever comes later. The completeness measure therefore captures the fraction of hours during this window that have data, of the hours that a monitor could have reported, so that a monitor

SSA $PM_{2.5}$ monitors is just 37%, meaning that the typical monitor has publicly accessible data for roughly one-third of its potential hours. Fewer than a third of monitors reach 50% completeness, and 30% have data for less than 25% of potential hours. Additional reference-grade monitoring efforts operate in the region, but data availability varies and most are not shared through centralized platforms.¹¹ The [EPIC Air Quality Fund](#) is funding new reference monitors and expects their grantees to share data, which may improve coverage going forward.

This scarcity leaves SSA without both the regulatory foundation to set air quality standards and without the calibration baseline to validate alternative measurements. Absent a reference monitor network, researchers studying air quality in SSA have to rely on the alternatives discussed below: low-cost sensors for direct measurement, or gridded products derived from some combination of models and satellites. Each involves trade-offs between measurement quality, spatial coverage, and the potential for systematic bias, all of which are amplified by the lack of reference monitors for calibration and validation.

2.2.2. Low-Cost Sensors (LCS). Low-cost sensors (LCS) have emerged as a potential response to these data gaps, providing hyper-local and real-time particulate matter data at a fraction of the cost of reference-grade monitors. These sensors infer $PM_{2.5}$ by shining light through a small chamber, measuring how strongly particles scatter that light, and converting this into a particulate matter estimate using an internal, often proprietary, formula.¹² The optical approach comes with inherent limitations: sensor performance is sensitive to humidity as well as the size distribution and composition of the sampled aerosols ([Ouimette et al. 2024](#)). In addition, data quality can degrade over time ([deSouza et al. 2023](#)) and accuracy depends on whether the local particle mix resembles the conditions under which the manufacturer calibrated the device ([Molina Rueda et al. 2023](#)). For example, if a LCS is calibrated in a city where vehicle exhaust is the primary aerosol but deployed in an area with significant windblown dust or biomass smoke, the manufacturer’s correction factors may fail ([Malings et al. 2020](#), [Molina Rueda et al. 2023](#)).

The most relevant biases for SSA are those related to particle size and humidity. Field evaluations of common LCSs have found that they severely underestimate $PM_{2.5}$ during windblown dust episodes, which is particularly problematic for SSA, where dust is often a major pollution source ([Molina Rueda et al. 2023](#), [Ouimette et al. 2024](#)).¹³ Humidity is also problematic, since many particles absorb water in humid air, swell, and therefore scatter more light, which leads to overestimation of particulate matter concentrations.¹⁴ A collocation study in Accra quantified both biases in an SSA setting: [Nimo et al. \(2025\)](#) find that mean absolute errors during the Harmattan dry season are roughly five times larger than in the wet season (MAE of 20–29 vs. 2–10 $\mu\text{g}/\text{m}^3$), with dust-driven underestimation dominating dry-season errors and humidity-driven overestimation dominating during the wet seasons.

Despite these limitations, LCSs can provide useful data if researchers follow careful calibration protocols. Properly calibrated LCSs can achieve reasonably high correlations with reference instruments, with R^2 of 0.8–0.9 quite common (see e.g., [Raheja et al. 2023](#)). Ideally, LCSs should be collocated with a reference-grade monitor before they are deployed to study sites, using collocation data to evaluate sensor performance and to develop calibration models (see e.g., [Giordano et al. 2021](#)).¹⁵ In the absence of a reference monitor, [Bagkis et al.](#)

that came online in 2023 is measured against its three active years, not the full decade. The completeness analysis covers only OpenAQ monitors, as the AirQo and WAQI APIs do not expose hourly observation counts.

¹¹The [Afri-SET](#) network in Ghana and the GeoHealth Hub ([Atuyambe et al. 2025](#)) have both deployed reference monitors.

¹²These proprietary algorithms mean that even “raw” $PM_{2.5}$ values typically embed opaque assumptions about particle size and density that researchers cannot observe or adjust for in post-processing.

¹³The performance for PM_{10} is so poor that [Molina Rueda et al. \(2023\)](#) recommend crowdsourced sensor networks to stop reporting PM_{10} concentrations from LCSs altogether, as they detect almost none of the particle mass above 2.5 μm and apparently instead report a mathematical allocation by the sensor’s firmware.

¹⁴At very high relative humidity, the relationship can reverse as the swollen particles move into size ranges in which the sensor undersamples ([Ouimette et al. 2024](#)).

¹⁵Most sensors log temperature and humidity alongside their main measurements, providing data that researchers can use to correct for environmental

(2025) describe two alternative approaches. The first, called proxy-based field calibration, treats one or more better-tested sensors as benchmarks and calibrates remaining sensors to match them using collocation periods and corrections. The second, transfer-based field calibration, applies a calibration relationship from a different site or time, with adjustments as needed for the current context.

These calibration requirements raise the actual cost of deploying LCSs beyond the purchase price of the devices. Reference collocation requires installation labor and a nearby reference monitor, and the alternative approaches demand additional sensors, field time, or rely on the assumption that calibration relationships transfer across sites. LCSs also have short operational lifespans compared to reference monitors (see e.g., [deSouza et al. 2023](#)), which adds to the cost of any longer-term study.

Calibration introduces a secondary data-generating process: the corrected measurements depend on model specification, training period selection, and algorithm choice, so researchers should test the sensitivity of their results to these calibration decisions. Imperfect humidity correction is particularly concerning for identification and residual calibration error will likely be largest in humid periods and locations. If humidity also independently affects the outcome—for example, through respiratory function, disease transmission, or crop growth—then the measurement error correlates with unobserved determinants of the outcome ($\text{Cov}(\eta, \varepsilon) \neq 0$ in Equation 3), and the resulting bias can push estimates in either direction.

Additionally, the measurement limitations of LCSs can introduce threats to identification for randomized controlled trials (RCTs) that use LCSs to collect outcome data. If treatment affects an area's aerosol levels or particle composition (e.g., dust versus biomass smoke or traffic-related soot), it can also change how the LCS responds to a given true PM concentration. Treatment could also change the rate at which sensors degrade, for example through differential dust deposition or smoke residue. These issues all lead to non-classical measurement error (Equation 3), with sensor drift and bias both potentially differing by treatment status.

Periodic re-collocation with a reference instrument can detect drift before it contaminates results. If study resources prohibit this, another approach is to rotate sensors across treatment and control sites on a fixed schedule. Device fixed effects then absorb each sensor's average bias, while the rotation prevents the time-varying component of drift from correlating with treatment status. These steps do not eliminate measurement error, but they reduce the chances that it will be correlated with treatment status.

2.3. Model-Based Data Products

Model-based data products estimate $\text{PM}_{2.5}$ concentrations using models of the atmosphere to represent how pollution is emitted, transported, and removed from the air. They underpin most global $\text{PM}_{2.5}$ datasets that economists use. Chemical transport models (CTMs) use emissions inventories and weather data to simulate pollution on a 3D grid. Reanalysis systems combine a forecasting model with available observations to produce a gridded dataset that is consistent across time and space. Hybrid models use combinations of CTM simulation and reanalysis outputs and satellite-based estimates as inputs into a statistical or, increasingly, a machine-learning model that is trained on ground monitor data. They then use the fitted relationships to extrapolate pollution estimates globally and over time.

2.3.1. Chemical-Transport Models (CTMs). CTMs simulate pollution on a 3D grid at fine time scales. They combine emissions inventories (databases of how much pollution each source type emits, where, and when) for different sources, such as traffic, industry, household fuel use, biomass burning, dust, and sea-salt, with weather data such as wind, temperature, and rainfall. The model then applies a large system of physics and chemistry equations that describe how pollution moves, changes, and leaves the atmosphere. Where ground monitors exist, CTMs mostly use them to evaluate or tune the model, rather than as direct inputs. GEOS-Chem, a widely used CTM, simulates how pollutants move with the wind, mix in the atmosphere, and get removed ([Bey et al.](#)

effects, with relative humidity being the most important correction.

2001). As we will see below, GEOS-Chem also feeds into several global PM_{2.5} products that economists and public health researchers use.

One common use of CTMs is to support counterfactual simulations. For example, [Park et al. \(2024\)](#) feed fire emissions estimates into GEOS-Chem to simulate yearly surface PM_{2.5} concentrations under both observed historical climate and counterfactual “no-climate-change” scenarios, attributing estimated mortality differences to climate-change-driven fire activity.

For applied microeconomics research, it is important to remember that CTM outputs are modelled concentrations, not measurements. In sub-Saharan Africa, CTMs face important constraints: emissions inventories often rely heavily on proxies, such as fuel use or population counts, instead of direct data on activities like burning or industrial activity. Because these proxies are themselves economic variables, CTM errors are likely correlated with economic activity, generating non-classical measurement error (Equation 3). Sparse ground monitoring compounds the problem, since model validation may be unable to detect or correct these biases. A recent global study of GEOS-Chem ([Zhang et al. 2023](#)) confirms that coarse global simulations understate population exposures and misstate which sectors and locations drive the highest concentrations, with especially large discrepancies in the Global South.

2.3.2. Reanalysis Datasets. Reanalysis datasets combine a numerical weather model with many different observations using what is called “data assimilation” to create a best-guess estimate of the atmosphere over time and space. At each time step, these models start from a recent forecast and adjust it, nudging it toward existing data, including surface stations and satellite measurements. When ground-based observations are sparse, like most of SSA, these models rely more heavily on the model and on satellites, so the uncertainty and potential biases are greater.

Most reanalyses focus on weather variables such as wind, temperature, and humidity, but some also include pollution concentrations. The two most relevant are NASA’s MERRA-2 ([Randles et al. 2017](#), [Gelaro et al. 2017](#)) and ECMWF’s CAMS-RA ([Inness et al. 2019](#)). MERRA-2 provides estimates of particulate matter, broken down into five components (dust, black carbon, organic carbon, sea salt, and sulfate), at $0.5^\circ \times 0.625^\circ$, or around 50 km, resolution from 1980 onward. CAMS-RA covers a shorter period, January 2003 to December 2024 at the time of writing, at $0.75^\circ \times 0.75^\circ$ (roughly 80 km) with three-hourly temporal resolution, but instead tracks a wider range of pollutants: in addition to particulate matter, it includes ozone, NO₂, SO₂, and CO. CAMS also provides forecasts for the same variables at a finer spatial resolution ($0.4^\circ \times 0.4^\circ$, or roughly 45 km). MERRA-2 was designed primarily to represent dust and smoke aerosols and their role in the climate system, while CAMS-RA places more emphasis on urban and industrial pollutants of the kind that European regulatory networks track.

Validation studies suggest that reanalysis products track pollution variation reasonably well on average, but this average masks substantial region-specific biases. Using data from Indian cities, [Navinya et al. \(2020\)](#) show that MERRA-2’s monthly averages are correlated with regulatory monitors ($\rho = 0.96$), but underestimate annual average PM_{2.5} by nearly $30 \mu\text{g}/\text{m}^3$ (34%). Further, while regulatory monitors recorded pollution above India’s national standard on 34% of days, MERRA-2 only flagged 11% of those days. In SSA, performance is worse: [Giordano et al. \(2024\)](#) compare MERRA-2 to reference monitor readings and find that the product explains very little of the PM_{2.5} variation in Kenya and South Africa ($R^2 < 0.1$). In Ghana, it attributes most particulate matter to dust in all seasons, even though Saharan dust is only a major source during parts of the year. In Niamey, [Gualtieri et al. \(2024\)](#) compare calibrated low-cost sensors against CAMS forecasts and find that they overestimate PM_{2.5} and underestimate PM₁₀, with substantial seasonal heterogeneity in the amount of measurement error.

Mapping these findings back to Equation 3, reanalysis products can have different biases in different regions and over time. This means that researchers cannot rely on validation results from one setting to sign the bias in another, and that panel designs should pay careful attention to how fixed effects interact with the error

structure. A stable level shift ($\bar{\eta}_i$ in Equation 5) would be absorbed by location fixed effects. The Ghana and Niamey evidence, however, suggests that reanalysis errors also vary systematically with seasons; in a location fixed effects design, this bias would enter through $\tilde{\eta}_{it}$ and remain in the error term, potentially confounding estimates if it correlates with outcomes or treatments under study.

2.3.3. Satellite-Based Geophysical Products. Satellite instruments retrieve column aerosol optical depth (AOD), which captures how much sunlight is scattered and absorbed by particles in the atmospheric column, i.e., between the top of the atmosphere and the Earth’s surface. The most commonly used AOD retrievals come from NASA’s MODIS instruments, available since 2000. MAIAC, a widely-used retrieval algorithm applied to MODIS data, produces AOD estimates at 1 km resolution. Geophysical products convert AOD into surface PM_{2.5} estimates based on models (often CTMs) that specify how much of the column mass sits near the ground and how particle size and humidity affect the AOD-PM relationship.

Most AOD retrievals share two systematic problems: non-random missingness and retrieval bias. Satellite AOD retrieval requires a clear view of the surface, so the algorithms discard observations where clouds or snow interfere. This removes a large share of daily observations—about 90% in MAIAC—and the missing days tend to be higher-pollution, because aerosols absorb water and scatter more light under the humid conditions that generate clouds (Bi et al. 2019). Among retrievals that do succeed, MAIAC tends to overestimate AOD in clean conditions and underestimate it during dusty conditions, because its assumptions break down when mineral dust dominates (Rogozovsky et al. 2023).

Both of these issues matter for empirical designs that rely on day-to-day variation as the non-random missingness removes the high-pollution tail of the exposure distribution and thus reduces the available identifying variation. The retrieval bias introduces a non-classical component (as in Equation 3), since the measurement error varies systematically with the true pollution level—overestimating in clean conditions and underestimating in dusty ones implies $\text{Cov}(\eta, X^*) < 0$. When this negative covariance is large enough, λ^* in Equation 3 can exceed one, meaning that the multiplicative bias amplifies rather than attenuates the true effect. If, in addition, cloudiness affects the outcome independently—through channels such as rainfall, temperature, or activity patterns—dropping cloudy days may introduce sample selection that could bias estimates in either direction. While few economists work with raw AOD directly, the hybrid products we discuss next use satellite data as inputs and the biases can persist even in corrected estimates, especially in monitor-sparse regions.

2.3.4. Hybrid Products. Hybrid products combine some or all of these ingredients—ground monitor data, CTM or reanalysis output, and satellite-derived estimates—to predict PM_{2.5} concentrations on a global grid using statistical or machine-learning models. The van Donkelaar et al. (2016) study is a leading example and the resulting dataset one of the most commonly used for global PM_{2.5} studies. First, they combine satellite AOD with GEOS-Chem output to build a global prior of annual PM_{2.5}. They then use a geographically weighted regression of the bias (ground monitor PM_{2.5} minus the prior), using data from around 10,000 monitors worldwide, on a small set of covariates, including indicators of aerosol type and land use, with location-varying coefficients.¹⁶

The resulting global PM_{2.5} surface agrees closely with monitors where they exist, but the correction is likely weaker far from calibration data. In SSA, where van Donkelaar et al. (2016) report zero or very few monitors contributing to the GWR, the correction relies heavily on the aerosol-type and land-use covariates, which in turn inherit uncertainty from CTMs (Section 2.3.1).

More recent work extends this approach using deep learning methods. Shen et al. (2024) build a convolutional neural network that takes existing estimates as a prior (instead of learning PM_{2.5} from scratch), and

¹⁶The GWR coefficients are effectively at $1^\circ \times 1^\circ$, weighted by inverse squared distance to ground monitors, and matching to monitors based on the greater of 100 km or the third-nearest monitor. Predictors are then interpolated to a 0.01° grid, so the output resolution is fine-scale but the coefficient variation is coarse—and in monitor-sparse regions like SSA, the effective smoothing radius can span many hundreds of kilometers.

learns their local bias relative to monitors. The authors then use a loss function that penalizes large deviations from this prior, so in places with few or no monitors it stays close to the physical prior instead of extrapolating freely from monitors in other regions. This design reduces the risk of unrealistic values far from monitors, which has been a common issue with earlier approaches. Their $PM_{2.5}$ estimates, released as dataset version [V6.GL.02.04](#), are available monthly from 1998–2023 at $0.01^\circ \times 0.01^\circ$ (equivalent to about 1.1 kilometer at the equator) and show substantially better out-of-sample fit in cross-validation than earlier approaches, especially in monitor-sparse regions.

A recent application of these methods in West Africa illustrates both their promise and their constraints for applied work ([Westervelt et al. 2025](#)). They construct a daily, 1×1 km gridded $PM_{2.5}$ dataset spanning 2005–2024 for West Africa by training several machine-learning models using data from ground-based reference monitors located on 10 US embassies, satellite measures of gases that form or travel with $PM_{2.5}$, as well as land meteorology from ERA5, a global atmospheric reanalysis produced by the European Centre for Medium-Range Weather Forecasts. Their preferred model, based on out-of-sample performance metrics, fits their urban training sites well, with an R^2 of around 0.9. The resulting data reveal very high exposures and strong Harmattan-season peaks across West Africa. Since the model is primarily training on data from ten urban centers, measurement error likely grows in rural areas without local training data and in places with very different emissions composition that differ from these sites. Even within these urban areas, US Embassy compounds may not be representative of the exposure patterns in other parts of the city.

While the latest hybrid models perform well in cross-validation, they degrade away from their training data. Evidence from the US illustrates the problem: [Sullivan & Krupnick \(2018\)](#) find substantial bias in globally calibrated $PM_{2.5}$ models even where regional calibration performs well, and [Fowlie et al. \(2019\)](#) show that satellite estimates diverge sharply from co-located monitors, with prediction errors that are spatially heterogeneous and biased downward at higher concentrations. In SSA, where monitoring is concentrated in very few urban sites, the problem is likely worse, especially in rural areas where the model falls back on priors. If these rural areas also have systematically different true $PM_{2.5}$ levels, the prediction error correlates with true exposure, generating $\text{Cov}(\eta, X^*) \neq 0$ in Equation 3.

2.3.5. Model Uncertainty. Model-based data products offer broad spatial and temporal coverage in SSA, filling an important gap where ground-based data remain scarce. However, they all embed significant uncertainty: CTMs and reanalyses inherit uncertainty from the proxies used in emissions inventories (fuel sales, population, land cover) when direct activity data are absent. Where surface monitors are sparse, models also lean heavily on satellite retrievals that struggle in ways that may distort the AOD-emissions relationship in spatially-patterned ways. Hybrid products reduce bias where monitors are dense but remain model-dependent in monitor-sparse regions. The most recent models improve fit by learning bias corrections, but estimates over SSA and other monitor-sparse regions still draw heavily on the model-based priors.

Model-based pollution products often report uncertainty estimates, but these usually come from internal consistency checks and sensitivity analyses, not from direct validation against external ground-based monitors. For example, [Van Donkelaar et al. \(2021\)](#) quantify uncertainty at the pixel level as having two components: satellite noise (how much do different satellite products affect the results) and calibration noise (what happens to their corrections when they change the weights that different monitors receive in the model). They then take the maximum of the two types of noise as the reported uncertainty. In SSA, this uncertainty primarily reflects disagreements across satellite and model inputs, given the scarcity of reference monitors, and the gap between reported and actual uncertainty is likely wide.

3. FROM CONCENTRATIONS TO OUTCOMES

The data constraints described above help explain why few causal estimates exist for how pollution affects health, productivity, or human capital in SSA. The studies reviewed below are building this evidence base,

but comparing their magnitudes to estimates from other regions requires care. SSA's pollution environment differs from better-studied settings, and the direction of the resulting bias is ambiguous. Lower baseline health, limited healthcare access, and low avoidance capacity (few air-conditioned buildings, frequent outdoor work) all suggest that the marginal health impact of pollution is larger in SSA than in the settings that generated most existing estimates. On the other hand, mineral dust accounts for a large share of SSA's $PM_{2.5}$, and dust aerosols may be less toxic per unit mass than combustion-generated particles (Fussell & Kelly 2021). This could attenuate effects relative to settings where traffic and industrial emissions dominate. The net direction is difficult to sign without local studies.

3.1. Health and Mortality

Most global mortality estimates for $PM_{2.5}$ rely on exposure-response functions in the style of the Global Burden of Disease (GBD), the global epidemiological study that quantifies deaths attributable to major risk factors (Burnett et al. 2014, 2018).¹⁷ These approaches attribute deaths to $PM_{2.5}$ exposure using a single concentration-response curve applied across all populations. This effectively assumes that a given inhaled dose of air pollution has similar toxicity across regions and countries. The studies underlying these curves come predominantly from North America, Europe, and China, so their applicability to SSA depends on how closely local populations, exposure ranges, and particulate mixtures match those in the estimation samples. Burnett et al. (2018) illustrate the sensitivity very clearly: when they replace the standard GBD risk function with a model that draws on broader evidence, the estimated $PM_{2.5}$ -attributable mortality burden for Africa more than doubles.

Recent economics research estimates the health impact of pollution exposure in developing countries (Jayachandran 2009, Greenstone & Hanna 2014, Arceo et al. 2016). Evidence from SSA remains limited, but two influential studies use Saharan dust to estimate mortality effects using quasi-experimental designs (Heft-Neal et al. 2020, Adhvaryu et al. 2024). Both papers find large mortality responses, comparable to or larger than prominent quasi-experimental estimates from high-income countries. Heft-Neal et al. (2020) instrument annual mean $PM_{2.5}$ with dust originating in the Bodélé Depression, isolating variation that is plausibly unrelated to local economic activity. They estimate that a $10 \mu\text{g}/\text{m}^3$ increase in $PM_{2.5}$ raises infant mortality by 24%. Adhvaryu et al. (2024) focus on prenatal exposure, use Harmattan-driven variation with rich fixed effects, and show that impacts decline over time, which they take as evidence of adaptation.

While these designs credibly address confounding from local economic activity, the exposure data introduce their own problems. Both studies use gridded $PM_{2.5}$ products matched to DHS clusters, which introduces Berkson-like error (Equation 4), and both data sources smooth true exposure peaks through spatial and temporal averaging.¹⁸ How these errors affect the estimates depends on the empirical design. Adhvaryu et al. (2024) use a fixed-effects specification, so the bias expressions in Equations 2–3 apply directly; monthly pollution measures likely smooth short Harmattan spikes, increasing the noise-to-signal ratio in the residual variation that fixed effects leave for identification (Griliches & Hausman 1986) and pushing estimates toward zero.

Heft-Neal et al. (2020) use instrumental variables, which addresses classical measurement error but introduces a different concern. Dust shifts the aerosol composition, and satellite retrievals measure dust- $PM_{2.5}$ with different accuracy than combustion- $PM_{2.5}$ (Section 2.1), so the instrument may be correlated with the proxy's measurement error. In either design, smoothing also distorts the estimated dose-response shape: compressing the right tail of the observed exposure distribution flattens the estimated curve at high concentrations, understating the health impact of the short, intense spikes that characterize e.g., Harmattan episodes.¹⁹

¹⁷For a detailed review of the public health literature on air pollution in Africa, see Abera et al. (2021).

¹⁸Heft-Neal et al. (2020) use a hybrid satellite product at 0.01° resolution, while Adhvaryu et al. (2024) use dust-specific $PM_{2.5}$ from a CTM at 1.25° resolution.

¹⁹This point extends beyond the linear framework of Equation 3, which characterizes bias in a slope coefficient but is silent on curvature. Carroll et al. (2006) shows that in nonlinear models, measurement error distorts the estimated shape of the relationship, not just its average slope.

Beyond measurement error, both studies use a single $PM_{2.5}$ measure and omit co-pollutants that may covary with $PM_{2.5}$ and independently harm health. As [Kluger et al. \(2024\)](#) show, this omission biases the coefficient away from zero, and the common assumption that measurement error makes estimates “conservative” can fail when this OVB dominates attenuation. Particle composition raises a related concern: because the identifying variation comes from dust, the estimates reflect the health effects of dust-driven $PM_{2.5}$ specifically. As we’ve seen, dust particles differ in size, composition, and toxicity from combustion aerosols ([Fussell & Kelly 2021](#)), so these estimates may not generalize to $PM_{2.5}$ from other sources. In addition, if dust composition independently affects health through channels other than $PM_{2.5}$ mass—for example, as a transport vector for microorganisms and toxic biogenic allergens ([Lwin et al. 2023](#))—the estimates conflate the effect of $PM_{2.5}$ mass with dust-specific toxicity.

These studies therefore speak more to the sign, timing, and heterogeneity of pollution damages and do not necessarily reflect the per-unit causal effect of $PM_{2.5}$. They nonetheless provide some of the clearest causal evidence on air pollution and early-life mortality in sub-Saharan Africa. [Pullabhotla et al. \(2023\)](#) extend this evidence to biomass fires, matching DHS births across 54 countries to satellite-derived burned area and using wind direction in an attempt to isolate the air-quality channel. They estimate that upwind burning raises infant mortality by roughly 2% per km^2 , with Africa accounting for about 75% of fire-attributable infant deaths globally.²⁰

By focusing on medium- to long-run average exposure, these studies have less to say about how short, intense spikes translate into exposure and health outcomes. [Berkouwer & Dean \(2026\)](#) address this gap by following charcoal-using households in an urban setting (Nairobi) for 3.5 years and randomizing subsidies for an energy-efficient stove. They then combine 48-hour personal $PM_{2.5}$ and CO exposure readings with hourly time-use data to study how reducing routine cooking spikes affects average exposure and a rich set of clinical and self-reported health outcomes. They find that the large reductions in spikes lead to improved self-reported respiratory symptoms, but they do not translate into measurable clinical gains.

To mitigate concerns about low-cost sensors, the authors collocate devices with each other, calibrate following [Giordano et al. \(2021\)](#), and include device fixed effects. Combined with sensor rotation across respondents, this makes it unlikely that the main spike result is due to static calibration artifacts. However, device fixed effects do not fully address time-varying shifts or drift, especially if the amount of drift correlates with when or where the device is deployed. This could be resolved in future work by explicitly randomizing device rotation across treatment arms.

3.2. Agricultural Productivity

There is evidence from other regions that air pollution lowers agricultural productivity. In India, [Burney & Ramanathan \(2014\)](#) estimate that greenhouse gases and short-lived climate pollutants reduced wheat and rice yields by 36% and 20%, respectively. In the United States, [Burney \(2020\)](#) show that decommissioning coal-fired power plants lowered local $PM_{2.5}$ concentrations and raised nearby corn yields by roughly one percent.

Evidence from sub-Saharan Africa remains limited but points in the same direction. [Aragón & Rud \(2016\)](#) exploit the rapid expansion of gold mining in southwestern Ghana between 1997 and 2005 to study impacts on nearby smallholder farms. They find large declines in agricultural productivity, around 40%, and substantial increases in local poverty in mining areas relative to comparable non-mining regions. Although the pollution channel rests on indirect proxies and partial tests, the paper presents a compelling case that air emissions link mining activity to farm productivity losses.

[Nwokolo \(2026\)](#) provides closely related evidence using satellite-based measures of gas flaring in Nigeria, combined with three waves of farm surveys from the World Bank’s Living Standards Measurement Study—

²⁰Their $PM_{2.5}$ -equivalent effect sizes have wide confidence intervals: 95% CI: 26,000–237,000 deaths per year, in part because the burned-area-to- $PM_{2.5}$ rescaling relies on ground monitors that are sparse in Africa.

Integrated Surveys on Agriculture (LSMS-ISA), spanning 2012–2019. The paper shows that a one-standard-deviation increase in gas flare volume during the harvest season reduces crop yields by roughly 18% and labor productivity by roughly 19% using OLS with household fixed effects. The paper also uses a shift-share instrument that interacts lagged changes in global oil prices with each local government area’s (LGA) initial (1998) gas flaring intensity, exploiting the fact that global oil price changes affect flares more in areas with historically high flaring levels. The IV estimates are much larger than OLS—roughly 3.5 times the OLS coefficients—suggesting that measurement error in satellite-derived flare volumes severely attenuates the fixed-effects estimates.

Two additional exercises support [Nwokolo \(2026\)](#)’s interpretation that air pollution drives the productivity losses. First, LGA-level regressions of satellite-measured pollutant concentrations on gas flare volume show that flaring LGAs have significantly higher black carbon, SO₂, and tropospheric NO₂ during the harvest season, though PM_{2.5} differences are not statistically significant. These regressions rely on gridded satellite products—NO₂ from the Ozone Monitoring Instrument (OMI), a spectrometer aboard NASA’s Aura satellite, at 0.2 × 0.2 degrees, and black carbon, SO₂, and PM_{2.5} from MERRA-2 at 0.5 × 0.625 degrees. Since a single MERRA-2 cell spans roughly 55 × 70 km, it can contain both flaring and non-flaring areas, which likely compresses the pollution contrast these regressions are trying to detect (see Equation 4). Second, a wind-direction check using ERA5 data shows that upwind flares reduce crop yields and labor productivity during the harvest season, while downwind flares have no significant effect, consistent with an airborne pollution channel.

[Nwokolo \(2026\)](#)’s gas flare setting provides a concrete illustration of the interaction between fixed effects and measurement error (see Section 2.1, Equation 5) in the context of model-based pollution data. Satellites detect large, hot flares reliably but underestimate or entirely miss small, cooler ones ([Seymour et al. 2025](#)). This creates nonclassical measurement error in the main explanatory variable, where areas dominated by small flares can appear in the data as having zero flaring ($\eta \approx -X^*$), while large-flare areas are measured accurately ($\eta \approx 0$).

Household fixed effects amplify the resulting attenuation by absorbing the persistent cross-sectional variation in true exposure—whether a household lives near a flare site—which is the main source of variation. What remains for identification, then, is within-household temporal variation across time—in this case, three survey waves—meaning that satellite detection noise likely accounts for a large share of the residual signal. Only 9% of households are ever exposed to measured flaring, and fewer than 1% are exposed in all three waves. As [Griliches & Hausman \(1986\)](#) show, when the signal is more persistent than the noise, fixed effects worsen the noise-to-signal ratio, and the resulting attenuation can be severe. This makes the large observed gap between OLS and IV estimates unsurprising in this setting.

The shift-share IV helps precisely because it reintroduces cross-sectional variation through the 1998 baseline flaring share, where the noise-to-signal ratio is lower than in the thin within-household temporal variation that the fixed-effects specification relies on. The sign, direction, and broad pattern of the results appear well-identified, though the magnitudes may embed some residual uncertainty due to the fact that the instrument’s 1998 baseline flaring-share was itself satellite-measured and subject to similar detection limitations. Despite these measurement challenges, the paper advances the frontier on pollution and agriculture in SSA by credibly linking a major industrial pollutant to farm-level productivity losses and demonstrating how extractive industries reshape local economies through environmental channels.

3.3. Labor and Human Capital

Air pollution affects labor and human capital through several interacting margins in SSA. Firms’ locational choices influence workers’ pollution exposure, labor supply responses determine how cumulative exposure changes on high-pollution days, and short-run cognitive impacts influence productivity conditional on working. Recent work from East Africa and Latin America illustrates each of these channels and shows that adjustment capacity and damages vary sharply across firms, workers, and tasks.

Along the first margin, job location, [Bassi et al. \(2022\)](#) show that exposure risk is already embedded in firm sorting within African cities. Using a combination of stationary and mobile PM_{2.5} monitoring in Ugandan

urban areas, they document sharp within-neighborhood pollution gradients tied to major roads. They match these data to a geo-coded survey of roughly 1,000 small manufacturing firms and their workers. Entrepreneurs sort onto the busiest, most polluted roads because traffic brings customers and is associated with around 15 percent higher profits for each step up in road size. Wages, however, rise by only about 2.5 percent over the same gradient, implying that workers capture little of the profit–pollution premium, all while bearing most of the added exposure.

The main profit and wage estimates are robust to sensor error, since they use road size as the regressor. The pollution gradient that motivates the paper’s welfare accounting, however, is more sensitive to measurement accuracy. [Bassi et al. \(2022\)](#) rely on AirQo’s low-cost sensors, validating the data mainly through internal consistency checks and comparisons to national averages. In the regressions that document the gradient, measured $PM_{2.5}$ is the dependent variable, and classical measurement error in a dependent variable does not bias coefficients. If sensor bias varies systematically with road size—larger roads may generate more heat, humidity, or traffic-specific aerosols that affect LCS readings differently—the measured gradient and the welfare accounting built on it would both be distorted.

A second margin operates through labor supply responses to day-to-day pollution. Using high-frequency data from Mexico City, [Hoffmann & Rud \(2024\)](#) show that workers reduce hours worked on high $PM_{2.5}$ days and that responses differ sharply by income. Combining hourly $PM_{2.5}$ readings from a dense ground-monitoring network with detailed labor market data, they estimate that workers in the top income decile cut hours by roughly two to three times more than workers in the bottom decile on very polluted days. This behavioral margin widens health and welfare gaps even in the absence of differential healthcare access, a pattern that shows how adjustment capacity mediates pollution damages. No comparable high-frequency evidence exists for SSA cities, where outdoor work is more prevalent and avoidance options are more limited, suggesting that labor supply responses may be smaller even as health costs are larger.

Cognitive performance creates a third margin where identical ambient shocks generate different impacts across workers and tasks. [Ma et al. \(2025\)](#) measure real-time $PM_{2.5}$ exposure during cognitive testing in the Kenya Life Panel Survey. They show that moderate, short-lived pollution shocks impair general cognition, with pronounced effects on memory, executive function, and visuospatial skills. Cognitive losses are larger among more educated respondents, suggesting that pollution can erode high-skill human capital at precisely the margins most relevant for complex tasks and decision-making. These findings suggest that pollution can reduce productivity even when workers remain on the job.

While [Ma et al. \(2025\)](#) sits at the frontier of what LCSs can offer in SSA, the design also illustrates some of the measurement challenges highlighted in Section 2.2.2. The devices they use have been documented to achieve good laboratory and field correlations with reference monitors, but they still rely on optical measurements that are sensitive to humidity, temperature, and aerosol composition. The authors control linearly for mean humidity and temperature and trim extreme values, but do not report formal collocation, calibration, or device fixed effects. This matters because the humidity-induced bias in optical sensors is nonlinear (Section 2.2.2). Particles absorb water and swell, causing overestimation that grows with both humidity and true concentration. A linear humidity control cannot fully absorb a bias that scales with both the regressor and the confounder. The result is scope for nonclassical measurement error (Equation 3): humidity both worsens cognitive performance directly and causes sensors to overestimate $PM_{2.5}$ conditional on true concentrations, so the estimated pollution effect captures both channels. These caveats notwithstanding, the study is one of the first to measure real-time personal pollution exposure during cognitive testing anywhere in sub-Saharan Africa, and the finding that short-lived shocks impair cognition is consistent with evidence from other settings.

The studies reviewed in this section illustrate both the promise and the difficulty of estimating pollution impacts in SSA. Credible identification strategies exist—quasi-experimental dust variation, shift-share instruments for gas flaring, within-neighborhood pollution gradients, real-time personal exposure during testing—and several produce large, policy-relevant effect sizes. At the same time, every study relies on pollution measures

that introduce some form of the measurement error characterized in Section 2.1, and in each case the sign and pattern of effects are typically more credible than the precise magnitudes.

4. IMPLICATIONS FOR FUTURE RESEARCH

Many studies reviewed in this article share a common challenge: estimates typically depend on pollution measures that introduce one or more of the measurement error problems discussed in Section 2.1. Future work can address this by (i) rethinking *what* to estimate, i.e., what questions to ask, and (ii) by treating exposure measurement as part of the identification strategy rather than a background assumption.

On the question of what to estimate, three gaps in the current literature stand out. First, researchers can better match the time scale of estimation to the time scale of pollution variation. Most studies reviewed here use annual or seasonal exposure averages, but much of the pollution variation that characterizes SSA—daily dust transport during the Harmattan, biomass burning episodes, traffic peaks on unpaved roads—unfolds over much smaller time scales. Temporal averaging compresses these spikes, worsening the noise-to-signal ratio under fixed effects (Section 2.1), potentially flattening estimated dose-response curves. Estimating daily impacts on productivity, labor supply, or health outcomes would preserve variation that annual averages discard, but doing so requires exposure data at matching temporal resolution.

Second, clearly separating physiological harm from behavioral avoidance would sharpen both measurement and interpretation. Current estimates typically capture a composite: if workers reduce hours or move indoors on polluted days, the estimated effect of pollution reflects both direct health damage and costly avoidance. Disentangling these channels matters for policy, since they call for different types of interventions. Doing so requires pairing data on individual responses—time use, mobility, indoor-outdoor transitions—with exogenous variation that shifts exposure separately from avoidance opportunities.

Third, developing locally estimated exposure-response functions would address one of the most consequential assumptions in the literature. As Section 3.1 showed, changing the risk function can matter tremendously for the magnitude of estimated PM_{2.5}-attributable mortality in Africa (Burnett et al. 2018). The functions currently used in global burden estimates come primarily from North American, European, and Chinese populations, where pollution mixtures, baseline health, and healthcare access differ substantially.

All three directions require exposure data that current infrastructure in SSA cannot easily supply. For studies that rely on satellite or reanalysis products, a promising approach is to report results across multiple products rather than treating any single one as ground truth. Different products embed different assumptions about emissions and aerosol properties, and they tend to disagree most in precisely the monitor-sparse regions where SSA research takes place. Where estimates agree across products, researchers can be more confident that results are robust to measurement assumptions and where they diverge, the range may provide informative bounds on the true effect. Partial-identification frameworks can formalize this, letting researchers report credible intervals that reflect measurement uncertainty alongside sampling uncertainty.

For studies that deploy low-cost sensors, the measurement challenge is different but just as serious. One practical approach combines LCS data with reanalysis products that provide information on meteorology and aerosol composition. Giordano et al. (2024) show that MERRA-2 data can help correct humidity-related bias in optical PM sensors, improving accuracy even when collocation with a reference monitor is infeasible. More broadly, economists using LCS data should treat calibration and validation as part of the research design, not a post-processing step. Pre-specifying calibration protocols and reporting performance metrics, placebo tests, and sensitivity to model choice would strengthen credibility. So would validating against whatever anchors are available, through a combination of collocation with portable reference units, rotations through regulatory sites, and cross-device consistency checks that bound sensor drift.

The question is no longer whether pollution data in SSA are good enough—they are not, and will not be for some time—but whether the studies that use them make this limitation visible and informative. The identifying variation exists, and the stakes for the hundreds of millions of people breathing unsafe air, as well as

for the governments allocating scarce resources, are too high to wait for perfect data.

SUMMARY POINTS

1. PM_{2.5} exposure imposes large health and productivity costs in sub-Saharan Africa, and the burden falls disproportionately on poor populations.
2. SSA's dominant pollution sources differ from those in the settings where most measurement tools and exposure-response functions are developed, which undermines the reliability of model-based data products and damage estimates.
3. The scarcity of publicly accessible reference-grade monitors forces empirical work to rely on satellite, reanalysis, and hybrid products, which embed substantial and spatially structured measurement error.
4. Measurement error in African pollution data rarely behaves like classical noise; rather, it varies with many factors that may correlate with development, urban form, and economic activity.
5. Applied work can improve credibility by treating exposure measurement as part of the identification strategy, reporting how measurement error interacts with the empirical design, and where it could change the conclusions.

FUTURE ISSUES

1. Most estimates use annual or seasonal averages, which hides important variation. It would be valuable to study the extent to which temporal averaging flattens dose-response curves.
2. How much of the estimated cost of pollution in SSA reflects behavioral avoidance rather than direct physiological harm? Disentangling these channels could provide important insights for policy-makers.
3. The exposure-response functions used in global burden estimates come from non-African populations and data environments with denser monitoring and different pollution profiles. Locally estimated functions that reflect SSA pollution mixtures, baseline health, and healthcare access could substantially revise current damage estimates in either direction.
4. Exposure measurement is often treated as a background assumption rather than part of identification. Exploiting disagreement across satellite products to bound effect sizes, and pre-specifying calibration and validation protocols for low-cost sensors, could change that.
5. Studies rarely make explicit which measurement assumptions drive the results and where they may break down. Doing so in a structured way would help readers assess what the estimates do and do not identify.

DISCLOSURE STATEMENT

The author is a co-founder of AfriqAir, a low-cost sensor air quality monitoring network operating across sub-Saharan Africa. This relationship involves no financial interest.

ACKNOWLEDGMENTS

I thank R. Subramanian and an anonymous referee for helpful comments. All errors are my own.

LITERATURE CITED

Abera A, Friberg J, Isaxon C, Jerrett M, Malmqvist E, et al. 2021. Air quality in Africa: Public health implications. *Annual Review of Public Health* 42(1):193–210

- Adhvaryu A, Bharadwaj P, Fenske J, Nyshadham A, Stanley R. 2024. Dust and death: Evidence from the West African Harmattan. *The Economic Journal* 134(659):885–912
- Aguilar-Gomez S, Dwyer H, Graff Zivin J, Neidell M. 2022. This is air: The “nonhealth” effects of air pollution. *Annual Review of Resource Economics* 14(1):403–425
- AirQo. 2026. AirQo air quality monitoring network. Accessed April 2026, <https://airqo.net>
- Amegah AK, Agyei-Mensah S. 2017. Urban air pollution in Sub-Saharan Africa: Time for action. *Environmental Pollution* 220:738–743
- Aragón FM, Rud JP. 2016. Polluting industries and agricultural productivity: Evidence from mining in Ghana. *The Economic Journal* 126(597):1980–2011
- Arceo E, Hanna R, Oliva P. 2016. Does the effect of pollution on infant mortality differ between developing and developed countries? Evidence from Mexico City. *The Economic Journal* 126(591):257–280
- Associated Press. 2025. US stops sharing air quality data from embassies worldwide. scientists say that cuts out a vital resource for global health. CNN. <https://edition.cnn.com/2025/03/05/world/us-air-quality-pollution-intl-hnk>
- Atuyambe LM, Etajak S, Walyawula F, Kasasa S, Nyabigambo A, et al. 2025. Air quality and attributable mortality among city dwellers in Kampala, Uganda: Results from 4 years of continuous PM_{2.5} concentration monitoring using BAM 1022 reference instrument. *Journal of Exposure Science & Environmental Epidemiology* 35(2):288–293
- Bagkis E, Hassani A, Schneider P, DeSouza P, Shetty S, et al. 2025. Evolving trends in application of low-cost air quality sensor networks: Challenges and future directions. *npj Climate and Atmospheric Science* 8(1):335
- Bassi V, Kahn ME, Lozano Gracia N, Porzio T, Sorin J. 2022. Jobs in the smog: Firm location and workers’ exposure to pollution in African cities. 30536, National Bureau of Economic Research. <https://www.nber.org/papers/w30536>
- Berkouwer SB, Dean JT. 2026. Cooking, health, and daily exposure to pollution spikes. *American Economic Journal: Economic Policy* 18(2):511–547
- Berkson J. 1950. Are there two regressions? *Journal of the American Statistical Association* 45(250):164–180
- Bey I, Jacob DJ, Yantosca RM, Logan JA, Field BD, et al. 2001. Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. *Journal of Geophysical Research: Atmospheres* 106(D19):23073–23095
- Bi J, Belle JH, Wang Y, Lyapustin AI, Wildani A, Liu Y. 2019. Impacts of snow and cloud covers on satellite-derived PM_{2.5} levels. *Remote Sensing of Environment* 221:665–674
- Bililign S, Brown SS, Westervelt DM, Kumar R, Tang W, et al. 2024. East African megacity air quality: Rationale and framework for a measurement and modeling program. *Bulletin of the American Meteorological Society*
- Blundell R, Horowitz J, Patey M. 2022. Estimation of a Heterogeneous Demand Function with Berkson Errors. *The Review of Economics and Statistics* 104(5):877–889
- Bound J, Brown C, Mathiowetz N. 2001. Measurement error in survey data. In *Handbook of Econometrics*, vol. 5. Elsevier, 3705–3843
- Burnett R, Chen H, Szyszkowicz M, Fann N, Hubbell B, et al. 2018. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *Proceedings of the National Academy of Sciences* 115(38):9592–9597
- Burnett RT, Pope CA, Ezzati M, Olives C, Lim SS, et al. 2014. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. *Environmental Health Perspectives* 122(4):397–403
- Burney J, Ramanathan V. 2014. Recent climate and air pollution impacts on Indian agriculture. *Proceedings of the National Academy of Sciences* 111(46):16319–16324
- Burney JA. 2020. The downstream air pollution impacts of the transition from coal to natural gas in the United States. *Nature Sustainability* 3(2):152–160
- Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM. 2006. Measurement error in nonlinear models: A modern perspective. New York: Chapman and Hall/CRC, 2nd ed.
- Chesher A. 1991. The effect of measurement error. *Biometrika* 78(3):451–462
- Chowdhury S, Pillarisetti A, Oberholzer A, Jetter J, Mitchell J, et al. 2023. A global review of the state of the evidence of household air pollution’s contribution to ambient fine particulate matter and their related health impacts. *Environment International* 173:107835
- Cohen AJ, Brauer M, Burnett R, Anderson HR, Frostad J, et al. 2017. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: An analysis of data from the Global Burden of Diseases Study 2015.

- The Lancet* 389(10082):1907–1918
- Deryugina T, Heutel G, Miller NH, Molitor D, Reif J. 2019. The mortality and medical costs of air pollution: Evidence from changes in wind direction. *American Economic Review* 109(12):4178–4219
- deSouza P, Barkjohn K, Clements A, Lee J, Kahn R, et al. 2023. An analysis of degradation in low-cost particulate matter sensors. *Environmental Science: Atmospheres* 3(3):521–536
- Duncan SD, Adeniyi CY, Ayoola FI, Giurcanu M, Ibigbami T, et al. 2025. Impact of fine particulate matter air pollution (PM_{2.5}) on Nigerian children’s performance on tests of cognitive and neurobehavioral development at age seven years. *Environment International* 205:109843
- Farquharson D, Jaramillo P, Samaras C. 2018. Sustainability implications of electricity outages in sub-Saharan Africa. *Nature Sustainability* 1(10):589–597
- Fisher S, Bellinger DC, Cropper ML, Kumar P, Binagwaho A, et al. 2021. Air pollution and development in Africa: Impacts on health, the economy, and human capital. *The Lancet Planetary Health* 5(10):e681–e688
- Fowle M, Rubin E, Walker R. 2019. Bringing satellite-based air quality estimates down to earth. *AEA Papers and Proceedings* 109:283–288
- Fuller WA. 1987. Measurement error models. *Wiley Series in Probability and Mathematical Statistics*
- Fussell JC, Kelly FJ. 2021. Mechanisms underlying the health effects of desert sand dust. *Environment International* 157:106790
- Gelaro R, McCarty W, Suárez MJ, Todling R, Molod A, et al. 2017. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate* 30(14):5419–5454
- Giordano MR, Beekmann M, Appoh E, Hughes A, Gatari MJ, et al. 2024. Utilizing reanalysis datasets to improve the performance of low-cost air sensors in the global south. ChemRxiv preprint, <https://doi.org/10.26434/chemrxiv-2024-j5dcn>
- Giordano MR, Malings C, Pandis SN, Presto AA, McNeill VF, et al. 2021. From low-cost sensors to high-quality data: A summary of challenges and best practices for effectively calibrating low-cost particulate matter mass sensors. *Journal of Aerosol Science* 158:105833
- Gordon JND, Bilsback KR, Fiddler MN, Pokhrel RP, Fischer EV, et al. 2023. The Effects of Trash, Residential Biofuel, and Open Biomass Burning Emissions on Local and Transported PM_{2.5} and Its Attributed Mortality in Africa. *GeoHealth* 7(2):e2022GH000673
- Greenstone M, Hanna R. 2014. Environmental regulations, air and water pollution, and infant mortality in India. *American Economic Review* 104(10):3038–3072
- Griliches Z, Hausman JA. 1986. Errors in variables in panel data. *Journal of Econometrics* 31(1):93–118
- Gualtieri G, Ahbil K, Brill L, Carotenuto F, Cavaliere A, et al. 2024. Potential of low-cost PM monitoring sensors to fill monitoring gaps in areas of Sub-Saharan Africa. *Atmospheric Pollution Research* 15(7):102158
- Health Effects Institute. 2022. The state of air quality and health impacts in Africa: A report from the state of global air initiative. , Health Effects Institute, Boston, MA. <https://www.stateofglobalair.org/resources/report/state-air-quality-and-health-impacts-africa>
- Health Effects Institute. 2024. The state of global air 2024: A special report on global exposure to air pollution and its health impacts, with a focus on children’s health. , Health Effects Institute, Boston, MA. <https://www.stateofglobalair.org/resources/report/state-global-air-report-2024>
- Heft-Neal S, Burney J, Bendavid E, Voss KK, Burke M. 2020. Dust pollution from the Sahara and African infant mortality. *Nature Sustainability* 3(10):863–871
- Hoffmann B, Rud JP. 2024. The unequal effects of pollution on labor supply. *Econometrica* 92(4):1063–1096
- Inness A, Ades M, Agustí-Panareda A, Barré J, Benedictow A, et al. 2019. The CAMS reanalysis of atmospheric composition. *Atmospheric Chemistry and Physics* 19(6):3515–3556
- Jayachandran S. 2009. Air quality and early-life mortality: Evidence from Indonesia’s wildfires. *Journal of Human Resources* 44(4):916–954
- Jeuland M, Pattanayak SK, Bluffstone R. 2015. The economics of household air pollution. *Annual Review of Resource Economics* 7(Volume 7, 2015):81–108
- Jha A, La Nauze A. 2022. US Embassy air-quality tweets led to global health benefits. *Proceedings of the National Academy of Sciences* 119(44):e2201092119
- Kluger DM, Lobell DB, Owen AB. 2024. Biases in estimates of air pollution impacts: The role of omitted variables and

- measurement errors. <https://arxiv.org/abs/2310.08831>
- Lwin KS, Tobias A, Chua PL, Yuan L, Thawonmas R, et al. 2023. Effects of desert dust and sandstorms on human health: A scoping review. *GeoHealth* 7(3):e2022GH000728
- Ma X, Layvant MN, Miguel E, Ochieng E, Pillarisetti A, Walker MW. 2025. Impacts of contemporaneous air pollution exposure on cognitive performance in Kenya. Working paper 34557, National Bureau of Economic Research. <https://www.nber.org/papers/w34557>
- Malings C, Tanzer R, Haurlyuk A, Saha PK, Robinson AL, et al. 2020. Fine particle mass monitoring with low-cost sensors: Corrections and long-term performance evaluation. *Aerosol Science and Technology* 54(2):160–174
- Martin RV, Brauer M, van Donkelaar A, Shaddick G, Narain U, Dey S. 2019. No one knows which city has the highest concentration of fine particulate matter. *Atmospheric Environment: X* 3:100040
- McDuffie E, Martin R, Yin H, Brauer M. 2021. Global burden of disease from major air pollution sources (GBD MAPS): A global approach. , Health Effects Institute, Boston, MA. <https://www.healtheffects.org/publication/global-burden-disease-major-air-pollution-sources-gbd-maps-global-approach>
- Mead MI, Okello G, Mbandi AM, Pope FD. 2023. Spotlight on air pollution in Africa. *Nature Geoscience* 16(11):930–931
- Molina Rueda E, Carter E, L'Orange C, Quinn C, Volckens J. 2023. Size-Resolved Field Performance of Low-Cost Sensors for Particulate Matter Air Pollution. *Environmental Science & Technology Letters* 10(3):247–253
- Murray CJL, Aravkin AY, Zheng P, Abbafati C, Abbas KM, et al. 2020. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: A systematic analysis for the Global Burden of Disease Study 2019. *The Lancet* 396(10258):1223–1249
- Navinya CD, Vinoj V, Pandey SK. 2020. Evaluation of PM_{2.5} surface concentrations simulated by NASA's MERRA version 2 aerosol reanalysis over India and its relation to the air quality index. *Aerosol and Air Quality Research* 20(6):1329–1339
- Newell K, Kartsonaki C, Lam KBH, Kurmi OP. 2017. Cardiorespiratory health effects of particulate ambient air pollution exposure in low-income and middle-income countries: A systematic review and meta-analysis. *The Lancet Planetary Health* 1(9):e368–e380
- Nimo J, Borketey MA, Appoh EKE, Morrison AK, Ibrahim-Anyass Y, et al. 2025. Low-cost PM_{2.5} sensor performance characteristics against meteorological influence in Sub-Saharan Africa: Evidence from the air sensor evaluation and training facility for the West Africa project. *Environmental Science & Technology* 59(13):6623–6635
- Nwokolo A. 2026. Gas flaring, agricultural productivity and labor reallocation: Evidence from Nigeria. Working paper, Growth Research Platform, Centre for Economic Policy Research. <https://grp.cepr.org/publications/steg-working-paper/gas-flaring-agricultural-productivity-and-labor-reallocation>
- OpenAQ. 2026. OpenAQ air quality data platform. Accessed April 2026, <https://openaq.org>
- Quimette J, Arnott WP, Laven P, Whitwell R, Radhakrishnan N, et al. 2024. Fundamentals of low-cost aerosol sensor design and operation. *Aerosol Science and Technology* 58(1):1–15
- Park CY, Takahashi K, Fujimori S, Jansakoo T, Burton C, et al. 2024. Attributing human mortality from fire PM_{2.5} to climate change. *Nature Climate Change* 14(11):1193–1200
- Pullabhotla HK, Zahid M, Heft-Neal S, Rathi V, Burke M. 2023. Global biomass fires and infant mortality. *Proceedings of the National Academy of Sciences* 120(23):e2218210120
- Raheja G, Nimo J, Appoh EKE, Essien B, Sunu M, et al. 2023. Low-cost sensor performance intercomparison, correction factor development, and 2+ years of ambient pm_{2.5} monitoring in Accra, Ghana. *Environmental Science & Technology* 57(29):10708–10720
- Randles CA, Da Silva AM, Buchard V, Colarco PR, Darmenov A, et al. 2017. The MERRA-2 aerosol reanalysis, 1980 onward. Part I: system description and data assimilation evaluation. *Journal of Climate* 30(17):6823–6850
- Rentschler J, Leonova N. 2023. Global air pollution exposure and poverty. *Nature Communications* 14(1):4432
- Rogozovsky I, Ohneiser K, Lyapustin A, Ansmann A, Chudnovsky A. 2023. The impact of different aerosol layering conditions on the high-resolution MODIS/MAIAC AOD retrieval bias: The uncertainty analysis. *Atmospheric Environment* 309:119930
- Schennach SM. 2016. Recent advances in the measurement error literature. *Annual Review of Economics* 8:341–377
- Seymour SP, Aylward B, Galvin KP, Kang M, Xie D. 2025. Gas flaring is likely underestimated by satellites due to undetected small flares. *Environmental Science & Technology* 59(40):21465–21475
- Shen S, Li C, van Donkelaar A, Jacobs N, Wang C, Martin RV. 2024. Enhancing global estimation of fine particulate

- matter concentrations by including geophysical a priori information in deep learning. *ACS ES&T Air* 1(5):332–345
- Sheppard L, Burnett RT, Szpiro AA, Kim SY, Jerrett M, et al. 2012. Confounding and exposure measurement error in air pollution epidemiology. *Air Quality, Atmosphere & Health* 5(2):203–216
- Sullivan DM, Krupnick A. 2018. Using satellite data to fill the gaps in the US air pollution monitoring network. Working paper 18-21, Resources for the Future. <https://www.rff.org/publications/working-papers/using-satellite-data-to-fill-the-gaps-in-the-us-air-pollution-monitoring-network/>
- Szpiro AA, Sheppard L, Lumley T. 2011. Efficient measurement error correction with spatially misaligned data. *Biostatistics* 12(4):610–623
- United Nations Environment Programme. 2021. Regulating air quality: The first global assessment of air pollution legislation. Nairobi: UNEP. <https://www.unep.org/resources/report/regulating-air-quality-first-global-assessment-air-pollution-legislation>
- United Nations Statistics Division. 2024. Standard Country or Area Codes for Statistical Use (M49). Accessed December 2025, <https://unstats.un.org/unsd/methodology/m49>
- Van Donkelaar A, Hammer MS, Bindle L, Brauer M, Brook JR, et al. 2021. Monthly global estimates of fine particulate matter and their uncertainty. *Environmental Science & Technology* 55(22):15287–15300
- van Donkelaar A, Martin RV, Brauer M, Hsu NC, Kahn RA, et al. 2016. Global estimates of fine particulate matter using a combined geophysical-statistical method with information from satellites, models, and monitors. *Environmental Science & Technology* 50(7):3762–3772
- Westervelt DM, Amooli JA, Anand A. 2025. Twenty years of high spatiotemporal resolution estimates of daily PM_{2.5} in West Africa using satellite data, surface monitors, and machine learning. *ACS ES&T Air* 2(8):1468–1477
- World Air Quality Index Project. 2026. AQICN air quality data platform. Accessed April 2026, <https://aqicn.org>
- World Health Organization. 2021. WHO global air quality guidelines: Particulate matter (PM_{2.5} and PM₁₀), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. Geneva: World Health Organization, 1st ed. <https://www.who.int/publications/i/item/9789240034228>
- Zhang D, Martin RV, Bindle L, Li C, Eastham SD, et al. 2023. Advances in simulating the global spatial heterogeneity of air quality and source sector contributions: Insights into the global south. *Environmental Science & Technology* 57(17):6955–6964

4.1. Terms & Definitions

Particulate matter: Tiny solid particles and liquid droplets suspended in the atmosphere, varying in size, shape, and chemical composition.

PM_{2.5}: Particles smaller than 2.5 micrometers (roughly 30 times thinner than a human hair), reported in $\mu\text{g}/\text{m}^3$. Often produced by combustion.

PM₁₀: Particles smaller than 10 micrometers, reported in $\mu\text{g}/\text{m}^3$. Includes mineral dust and other coarse particles.

Aerosol: A suspension of particles in air, including dust, soot, smoke, sulfates, and sea salt, that scatters and absorbs sunlight.

Reference-grade monitor: An air quality instrument meeting regulatory accuracy standards, used as ground truth for validating other measurements.

Low-cost sensor (LCS): A portable device that estimates PM concentrations, typically by measuring how particles scatter light. Requires calibration against reference instruments.

Collocation: Deploying a low-cost sensor alongside a reference monitor to develop site-specific calibration corrections.

Chemical transport model (CTM): A numerical model that simulates pollution concentrations using emissions inventories and weather data as inputs, rather than direct measurements.

Reanalysis: A gridded atmospheric dataset produced by blending a numerical model with observations through data assimilation to fill spatial and temporal gaps.

Aerosol optical depth (AOD): A satellite-retrieved measure of how much sunlight is scattered and absorbed

by particles in the atmospheric column.