Check for updates

ATMOSPHERIC SCIENCE

Hyperlocal sensing and inverse modeling reveal community impacts of urban air pollutant emissions

Chirag Manchanda[1], Ronald C. Cohen[2,3], Ramon A. Alvarez[4], Tammy M. Thompson[4], Maria H. Harris[4], Alexander J. Turner[5], Julian D. Marshall[6], Robert A. Harley[1], Joshua S. Apte[1,7]*

Copyright © 2026 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).

Addressing urban air pollution requires identifying, quantifying, and mitigating emission sources, yet bottom-up inventories often lack fine spatiotemporal precision and rarely capture unpermitted or informally operated sources. We combine dense mobile and fixed-site measurements in a Bayesian inverse modeling framework to produce hourly, hyperlocal black carbon (BC) emission maps (150 meters by 150 meters) that reveal how and where key sources are underrepresented. In West Oakland, CA—an overburdened community near major freight infrastructure—our approach identifies underestimated and unaccounted emission sources, increasing estimated BC emissions by ~33%. These corrections quadruple the contribution of neighborhood and port-related sources to population-weighted exposures (from 5 to 20%), disproportionately affecting low-income neighborhoods adjacent to freight corridors. Our results demonstrate that combining dense observational data with inverse modeling enables detection and quantification of overlooked emissions, supporting targeted intervention. Sensitivity analyses indicate similar performance with a substantially reduced fixed-site network (~10 monitors; ~1 site per square kilometer, compared with 6.7 sites per square kilometer in the full campaign), paired with time-averaged mobile maps, supporting applicability in data-limited urban settings.

INTRODUCTION

Concentrations of urban air pollutants vary sharply in space, owing to a diverse array of unevenly distributed sources that often contribute to inequitable exposures and health outcomes (*1*). Bottom-up emission inventories—the foundation for most regulatory action—rarely capture this fine-scale variability, often missing unpermitted, informally operated, or transient sources, and offering little insight into who is most affected and why. To mitigate human exposure to air pollution, it is essential to understand which pollution sources degrade air quality and by how much, whom these sources impact, and the underlying reciprocal linkages between emissions and exposures. While findings from mobile monitoring studies and dense networks of stationary low-cost pollution sensors increasingly highlight sharp gradients and localized exposure hotspots (*2*), translating these observations into quantitative, source-specific emissions estimates has remained elusive. Here, we leverage these dense hyperlocal air pollution observations (i.e., subneighborhood scale), in a measurement-modeling framework to quantify source-specific emission fluxes and their subsequent population exposure impacts at a 150-m and hourly resolution (*3*–*8*). Using this approach, we reveal ubiquitous neighborhood-scale pollution sources that have a substantial but previously poorly quantified impact on population exposures to air pollution, offering opportunities for precise, targeted mitigation.

These gaps in fine-scale emissions understanding are becoming increasingly consequential as the urban emissions landscape rapidly evolves. Economy-wide shifts—including transitions in energy use in the transport and building sectors, the rise of e-commerce, and changing patterns of goods movement—are reshaping the sources that influence community-scale exposure. Yet, emission inventories often lag behind these changes. High-resolution observations are essential to detect and quantify these emerging, high-impact contributors to urban emissions (*3*, *9*–*13*). Here, we use "hyperlocal" to refer to subneighborhood spatial scales, within the broader fine-scale monitoring range of roughly tens of meters to 1 km.

Traditional air quality management uses "bottom-up" emission inventories that are based on estimates of the spatial distribution of activities and their emission factors. These inventories then drive atmospheric dispersion or chemical transport models (CTMs), which simulate pollutant dynamics and infer exposures (*14*, *15*). CTMs are indispensable for exploring atmospheric processes, testing emission scenarios, and evaluating chemical drivers of observed patterns. Their predictive accuracy, however, often depends heavily on the quality of input emission inventories, which frequently disagree with in situ concentration measurements, overlook key sources, and lack the spatial resolution required for effective local-scale interventions (*4*, *16*). Bayesian inverse modeling offers a complementary "top-down" approach to constrain and update emissions using atmospheric observations. Recent urban inversions have estimated emissions of short-lived and combustion-related pollutants across urban to regional domains and at increasingly fine spatial and temporal scales (*17*–*22*). These studies demonstrate the value of top-down methods for evaluating bottom-up inventories and resolving spatiotemporal variability in emissions. However, estimating emissions at scales relevant to intra-urban exposure gradients remains challenging, particularly where pollutant concentrations vary sharply near roads, ports, rail corridors, and other urban combustion sources. While inversions can reduce dependence on prior inventories, they too have limitations, most notably their reliance on sufficient observational coverage, which has historically limited their application to coarse spatial scales due to the high cost of acquiring dense measurements (*23*–*30*). In the past decade, hyperlocal air pollution mapping—for example, using dense networks of fixed sensors and/or mobile

[1]Department of Civil and Environmental Engineering, University of California, Berkeley, Berkeley, CA 94720, USA. [2]Department of Earth and Planetary Science, University of California, Berkeley, Berkeley, CA 94720, USA. [3]College of Chemistry, University of California, Berkeley, Berkeley, CA 94720, USA. [4]Environmental Defense Fund, Washington, DC 20004, USA. [5]Department of Atmospheric and Climate Science, University of Washington, Seattle, WA 98195, USA. [6]Department of Civil and Environmental Engineering, University of Washington, Seattle, WA 98195, USA. [7]School of Public Health, University of California, Berkeley, Berkeley, CA 94720, USA. *Corresponding author. Email: apte@berkeley.edu

Downloaded from https://www.science.org on October 29, 2026

monitoring fleets—has emerged as a powerful tool to measure air pollutant concentrations at very fine spatial scales, typically from tens of meters to ~1 km. Yet, these techniques typically stop at mapping patterns: They identify where pollution is high, but not why (*2*). The framework we present here integrates these threads—hyperlocal observations, CTM-based transport modeling, and Bayesian inference to quantify source-specific emissions at 150 m, hourly resolution—directly linking community-level concentration patterns to the emission sources driving them. This innovation enables the detection and quantification of overlooked urban sources with the precision needed for targeted mitigation.

This study focuses on West Oakland, CA, a historically and predominantly African-American community (population ~ 40,000, 10 km^2) that has long faced a high cumulative air pollution exposure burden (*31*). Like many overburdened communities, West Oakland has a diverse array of pollution sources, including three interstate highways, a large port complex, and multiple rail yards and intermodal freight-handling facilities. Industries, warehouses, and major commercial activities are located adjacent to and interspersed within residential areas of the community. Diesel emissions are a major contributor to elevated pollution, and black carbon (BC), a major component of diesel exhaust, imposes a substantial exposure burden on the community. Despite more than a decade of intensive air quality studies led by regulatory agencies, research scientists, and community groups to inform mitigation efforts (*15*, *31*–*35*), much of West Oakland's pollution burden remains poorly resolved at the source level. Prior work has produced rich datasets, including hyperlocal fixed-site and mobile monitoring, detailed inventories, and validated transport models, yet these have not been fully integrated to pinpoint and quantify the sources most responsible for community exposures. This wealth of data, combined with the community's complex source mix and history of environmental injustice, makes West Oakland an ideal setting to test our high-resolution inversion framework and demonstrate how it can uncover overlooked emissions that matter most for population exposures.

We build on prior Bayesian inversion studies by combining dense mobile and fixed-site air pollution monitoring datasets into an assimilated observational constraint that allows emissions to be estimated at hyperlocal resolution (*5*, *6*, *36*, *37*). We term the method $INFE^2R$ (INversion for Fine-scale Emission and Exposure Refinement) (see section S1 and fig. S1 for details). Compared with bottom-up inventories, which smooth emissions across large grid cells and depend on activity-based estimates, $INFE^2R$ produces observation-constrained BC emissions at 150 m resolution—capturing sharp spatial gradients and transient emission plumes that inventories miss. This framework enables three advances: (i) precise BC emission estimates tied directly to measured concentrations rather than proxy activity data; (ii) identification of sources that were previously unaccounted for in emission inventories; and (iii) a more actionable understanding of the link between emissions and population exposures that comes from reintroducing corrected emissions into forward models. Even in a well-studied area like West Oakland, where disparities and sources are broadly known, $INFE^2R$ changes the picture from "we know there is a problem" to "we know exactly where, when, and by how much." While the broad patterns may align with local expectations, our results quantify the magnitude and temporal dynamics of these impacts, enabling a shift from generalized assumptions to targeted, data-informed mitigation. Because the approach relies on detailed observational data, posterior estimates remain stable under large systematic and randomized perturbations to prior emissions estimate, while still benefiting from minimally informative spatial structure in the prior emissions for fine-scale detail.

Like any Bayesian approach, $INFE^2R$ combines three elements of a Bayesian inverse problem: a prior emissions estimate, a forward transport model that maps emissions to concentrations, and observational constraints, each represented with associated uncertainties. Here, we leverage an existing, previously validated BC emission inventory for Oakland (*32*), which incorporates detailed measurements of the local diesel vehicle fleet (*38*–*40*). We applied a combination of a meteorological model and a Lagrangian transport model [Weather Research and Forecasting–Stochastic Time-Inverted Lagrangian Transport (WRF-STILT); see the Supplementary Materials] (*41*, *42*) to generate a source-receptor atmospheric transport matrix **H** that links BC concentrations at receptors to BC emissions at the source (resolution: 150 m, hourly). Together, these steps provide prior predictions of BC concentrations across the domain (see section S1.1 and figs. S2 and S3 for details). For the second input—observational constraints—we use a published data-assimilation technique (*43*) that combined hourly measurements from a 100-day hyperlocal monitoring campaign (May to August 2017; 97 sites in 15 km^2, or 6.7 monitors km^{-2}) with repeated mobile monitoring on every street (output: time-average pollutant maps at 30 m) (*44*, *45*). The data assimilation provides a spatiotemporally complete BC observational dataset (150 m resolution, ~ 450 assimilated hourly observations in domain). This uniquely rich dataset blending dense fixed-site and mobile coverage enables us to demonstrate an inversion at hourly, neighborhood scale, where observational constraints are typically too sparse for robust source attribution.

Next, we solve an inverse problem, reconciling the mismatch between modeled BC concentrations and assimilated observations to derive spatiotemporally resolved "posterior" BC emission flux estimates. When reintroduced into the forward model, these corrected emissions reduce (by ~60%) out-of-sample mismatches between predicted and observed concentrations, and enable precise, data-driven quantification of how specific sources expose West Oakland's population to BC. Crucially, we show that this inversion approach remains robust even under large uncertainties or systematic errors in the prior inventory, underscoring its applicability across various urban settings.

The advance here is not hourly inversion alone, which has been demonstrated in prior top-down studies, but the combination of hourly temporal resolution with 150 m spatial resolution, source-resolved emission estimation, and exposure attribution. Whereas recent time-resolved urban inversion studies resolve emissions down to ~1 km^2 (*3*, *12*, *26*, *27*, *30*, *37*, *46*), here we provide robust emissions estimates for pixels with about 50× smaller area (150 m × 150 m ≈ 0.02 km^2). We demonstrate how data assimilation allows even a limited set of fixed site monitors—when coupled with mobile monitoring—to constrain emissions precisely and with high spatiotemporal resolution. While many of the individual building blocks used here have been demonstrated separately, $INFE^2R$ integrates them into a fully observation-constrained inversion at subneighborhood scale, showing how emerging hyperlocal data streams can be translated into actionable insights for air pollution control.

RESULTS AND DISCUSSION

Inverse model reveals sharp BC corrections

Inverse modeling yields substantial observationally constrained corrections to the prior emission inventory, with notable spatiotemporal

Downloaded from https://www.science.org on October 29, 2026

nuance and structure. Figure 1A contrasts the prior and posterior estimates for BC emissions. Aggregated over space and time, domain-wide BC emissions increase by ~33%—from ~2.7 kg hour^{-1} in the prior to ~3.6 kg hour^{-1} in the posterior, but substantial differences emerge at fine scales. Corrections at the 150-m grid scale align closely with land uses (Fig. 1B), with the largest corrections at and near port facilities, at locations within residential neighborhoods, and along nearly all highways and major roads. Time-averaged posterior emissions increase nearly everywhere across the domain, except on one stretch of highway (I-880) dominated by drayage trucks. BC emission fluxes are highly variable in space, with adjacent 150-m pixels often differing by more than 5×, suggesting the presence of emission hotspots.

Emissions spatial patterns show a marked weekend-to-weekday difference, with particularly large weekend reductions in areas near major physical infrastructure (ports, railyards, and highways). Domain-wide BC emissions are 2× to 3× higher on weekdays than on weekends, with marked day-to-day variation (~3.5 to 5.5 kg hour^{-1}, compared with a weekday prior of nearly constant ~3.0 kg hour^{-1}, Fig. 1C). A demonstration of the model's efficacy is its detection of a sharp emissions drop on July 4—a public holiday with reduced freight and industrial activity that is entirely absent from the prior inventory. This capability shows that the framework can dynamically correct for the impact within the domain of unanticipated events or holidays, so long as these deviations are captured in the observational record. Conversely, emissions corrections show

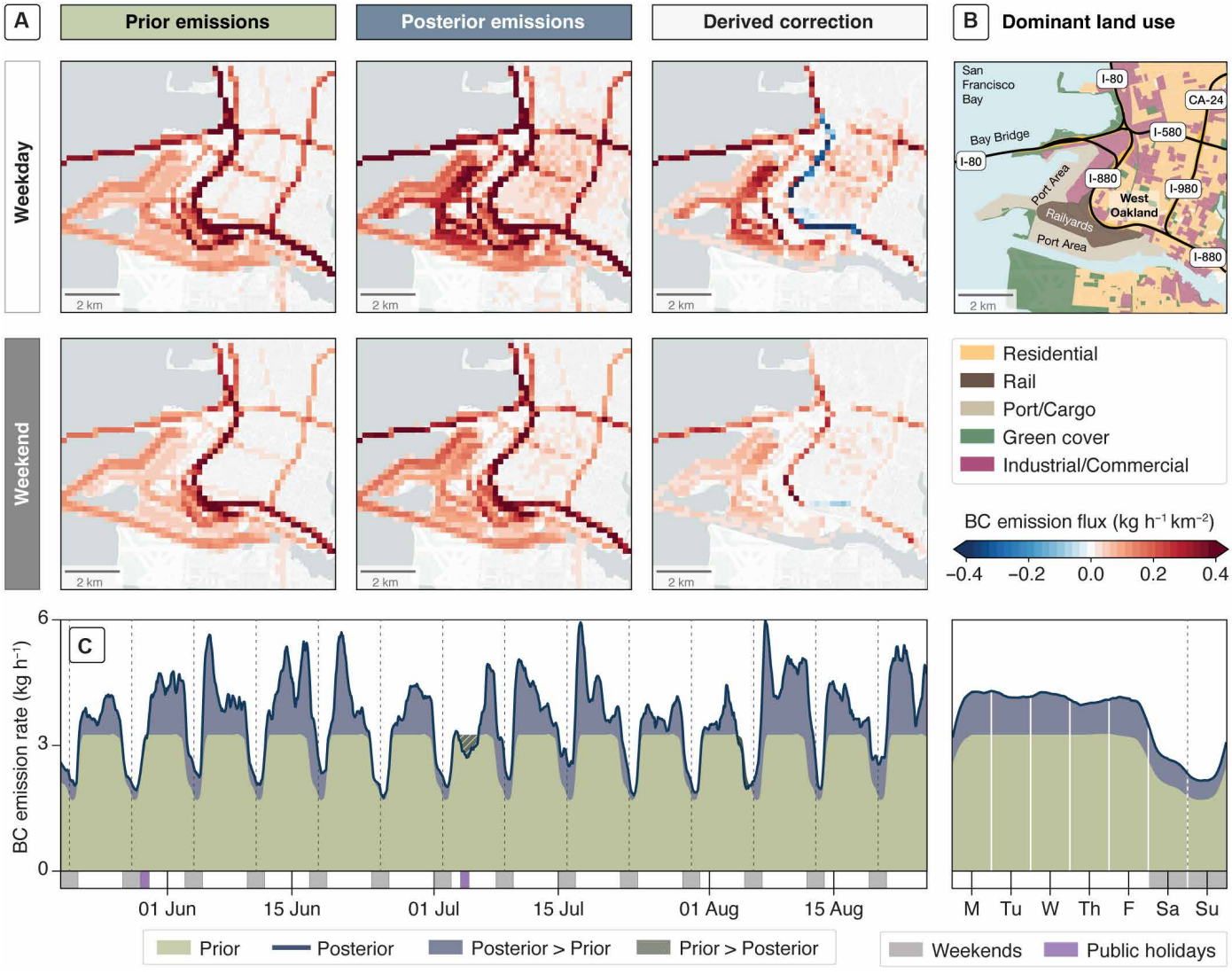


Fig. 1. Spatiotemporal variability of inverse-derived emission corrections. (**A**) Time-averaged emission correction maps illustrate fine-scale (150 m) spatial variability in BC emissions for weekdays and weekends. The prior inventory generally underestimates emissions. The inversion ("posterior") results in large upward correction in weekday emissions across the domain. The derived corrections vary sharply in space, with a reduction in emissions over part of the truck-dominated I-880 highway. Notably large additions to emissions in the posterior are in residential neighborhoods, which were almost entirely absent in the prior inventory, as well as in the area surrounding the Port of Oakland. (**B**) Map of land-use patterns across the domain. (**C**) Hourly trends in BC emissions compare the posterior (bold blue line) and prior (light green shaded area) estimates. The posterior emissions consistently exceed the prior (blue shaded area). A notable exception occurs on a public holiday (July 4) not accounted for in the prior inventory (dark green hatched area). Vertical dashed lines in (C) indicate the beginning of each week (0h00 Sunday).

Downloaded from https://www.science.org on October 29, 2026

negligible temporal association with domain-average wind speed and boundary layer height (fig. S4), suggesting that day-to-day correction magnitudes are not primarily controlled by broad meteorological variability. More generally, this approach can capture episodic or short-lived changes in emissions—such as special events, temporary shutdowns, or traffic anomalies—if these events are observed and fall within the footprint window of the inversion. The framework is thus sensitive to real-world variability, but not omniscient: It cannot detect events that fall outside the observational record or spatial footprint of the modeling domain.

The observationally constrained posterior exhibits spatiotemporal fidelity in reconstructing out-of-sample BC observations using cross-validation (section S1.4). Specifically, comparison of 150-m, 1-hour BC predictions against held-out observations yielded Pearson's $R \sim 0.81$ to 0.84 and a normalized mean absolute error (NMAE) of 17-20% (table S1 and figs. S5 to S7). As an additional independent check, we compared prior and posterior predicted BC concentrations against measurements from an EPA monitoring site that was not included in the mobile/fixed-site network or inversion and uses a different measurement approach. Posterior improved agreement relative to the prior at this independent site (Pearson's $R = 0.87$ versus 0.45; fig. S8), providing validation beyond held-out inputs to the assimilated product. Additional fixed-site out-of-sample validation and reduced-observation sensitivity tests are presented in sections S5 and S6. In contrast, the prior bottom-up emission inventory performs poorly in predicting observations (Pearson's $R \sim 0.35$ to 0.40, NMAE of 35 to 50%). Using an ensemble approach (fig. S9), we estimate high precision for the posterior estimates, with temporal and spatial coefficients of variation (SD/mean) of 8 and 14%, respectively, based on heterogeneity in the observational constraints. Because explicitly forming the full averaging kernel is computationally intractable for this high-dimensional state vector (see section S2), we interpret the posterior emission field at 150 m spatial resolution as a covariance-regularized, observation-constrained estimate rather than an independent retrieval for every grid cell and hour. Confidence is highest where posterior corrections are supported by footprint sensitivity, low ensemble variability, and stability across prior-perturbation tests (figs. S9 and S18 to S21). These results demonstrate that INFE^{2}R not only corrects large-scale biases in the inventory but also resolves fine-scale spatiotemporal variability with high fidelity. This capability makes it broadly applicable to other cities, including those with fewer fixed-site monitors or less detailed prior inventories. Together, the held-out validation, EPA-site comparison, ensemble analysis, and prior-perturbation tests provide converging evidence that the major fine-scale posterior features are not driven by isolated observation or transport errors, while also showing that pixel-scale corrections are less certain for grid cells where receptor footprints provide limited observational sensitivity or where emission signals are very small.

Fine-scale insights into sectoral emissions

We derive emissions estimates at a fine spatial scale that reflects land-use-based distinctions among activity sectors. We group sources into five sectors that generate distinct temporal and spatial patterns of BC emissions: (i) major roads, which include diesel vehicles on highways and designated truck routes; (ii) cargo handling equipment at the Port of Oakland (on land) (e.g., diesel-fueled on-shore equipment such as cranes and loaders; trucks on port premises); (iii) rail (line-haul and switching locomotives; trucks and diesel equipment at intermodal facilities); (iv) marine (active and berthed ships); and (v) "neighborhood" emissions, representing diverse in-community sources such as warehouses, industrial or commercial activity, and vehicle emissions on smaller roads not classified as truck routes. Although such sources are often included in bottom-up inventories, aggregation across broad categories can mask important differences in magnitude, timing, and mitigation potential. While diesel trucks can appear in multiple sectors, their operational contexts and spatial clustering differ, enabling differentiated interpretation and targeted intervention. To allow flexibility, grid cells with zero prior emissions were initialized with a negligible placeholder value (1×10^{-8} kg hour^{-1}), ensuring numerical stability and permitting the inversion to allocate and scale up "missing" or underrepresented sources where supported by observations.

To enable sector-specific interpretation, we apply a single-sector assignment per 150-m grid cell (Fig. 2A; see section S4.1 for details)—a practical simplification that works well in this setting, where source types are relatively spatially segregated at this resolution. Sectoral masks are used to structure the prior error covariance and support post hoc interpretation; they do not hard-constrain posterior emissions or prevent corrections in any grid cell. The covariance regularization encourages nearby pixels within the same source context to covary, while posterior emissions remain driven by the mismatch between observations and forward-model predictions (see section S2). Thus, adjacent highway pixels may be temporally correlated, while an adjacent highway and neighborhood pixel are not. Emissions can be assigned to any grid cell regardless of prior attribution or mask. This spatially explicit segmentation enables us to evaluate how different emission-generating land uses contribute to overall emissions patterns and how they differ in both magnitude and temporal dynamics. This step allows direct comparison of the prior and posterior inventories by sector, making it possible to identify where bottom-up methods systematically under- or overestimate emissions.

Each sector exhibits distinct inventory corrections (Fig. 2B). Domain-wide emissions from rail change modestly (+4%), while port (+270%) and neighborhood sources (+100%) undergo larger and more complex corrections. Temporal signatures of posterior estimates likewise differ among sectors. Whereas rail and marine emissions exhibit only moderate temporal variability, we find that major roads, port, and neighborhood sectors show sharp diurnal trends with 2× to 3× lower emissions on weekends (Fig. 2B). These corrections are not simply changes in total magnitude. They also revise the temporal allocation of emissions. To isolate changes in timing from changes in total emissions, we compare peak-normalized prior and posterior diurnal profiles by sector (fig. S10). This comparison shows that the inversion modifies the shape of daily activity patterns, especially for major roads, port, and marine emissions. This suggests that inventory errors arise not only from missing or underestimated sources, but also from temporal allocation assumptions that smooth real-world activity patterns, particularly for freight-related diesel activity.

Our results indicate that diesel emissions on major roads were already reasonably captured in the prior inventory (Fig. 2B). However, this likely does not reflect the overall impact of diesel truck emissions, as sectoral attribution in our framework is based on spatial land use and activity patterns rather than discrete emission sources. Some of this underestimation likely stems from truck activity off designated routes and within facilities. Because the prior inventory spatially

Downloaded from https://www.science.org on October 29, 2026

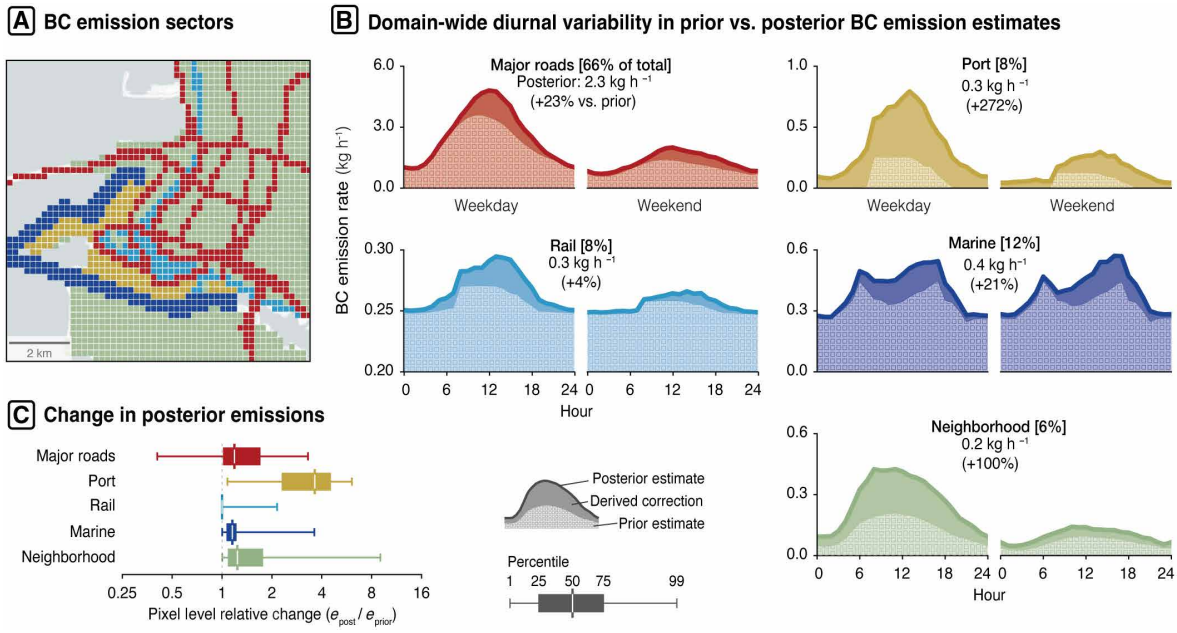


Fig. 2. Sector-resolved variability in emission corrections. (**A**) Emissions in the prior and posterior inventories (150 m resolution) are categorized into five major activity sectors based on their spatial distribution. (**B**) Domain-wide diurnal emissions rate profiles (kg BC hour^{-1}) for each sector. Note the substantial variation among sectors in diurnal and weekday/weekend temporal signatures. Sectors vary markedly in the degree to which emissions increase in our observationally constrained posterior estimate. (**C**) The magnitude of emissions corrections in the posterior varies from sector to sector, with the distribution of the pixel-level posterior-to-prior emissions ratios indicated in the box-and-whisker plots. At the pixel level, median emissions in the port increase by +300% in the posterior, while neighborhood sources and major roads have median corrections closer to +50%. Sector labels indicate dominant source context at 150 m resolution. The major-roads category reflects freeways and designated truck routes in the prior inventory, not all large or multilane roads; individual pixels may also contain adjacent land uses.

allocates on-road emissions using activity proxies such as traffic counts, road density, truck-route designations, and facility activity data, these posterior corrections help identify where such proxies underallocate diesel activity. This pattern reflects a structural limitation of bottom-up inventories that may be generalizable to cities where truck-intensive activities are interspersed among other land uses, with warehouse clusters and small industrial operations contributing emissions that are poorly tracked or underreported. While diesel trucks likely contribute across sectors, each sector has a distinct regulatory and mitigation context, shaping the policy relevance of the corrections. Consequently, the biases revealed here may be generalizable beyond the specific place and time of our study (West Oakland in 2017).

At the scale of individual 150-m pixels (Fig. 2C), the distribution of relative changes is highly skewed. While most pixels show substantial increases (posterior-to-prior ratios of 1.9 and 1.25 for the mean and median, respectively), the top 1% of pixels in the neighborhood, marine, and port sectors experience marked increases of 6× to 8×. These corrections not only are large in relative terms, but also correspond to meaningfully large absolute fluxes in many cases, disproportionately affecting population-weighted exposure due to proximity to residential areas. Where the prior inventory was near-zero, the absolute mass may be small, but localized increments can still create peaks of concern for sensitive receptors. These substantial corrections indicate that the inversion framework can identify and substantially revise emissions for sources that are absent or nearly absent from the prior inventory—particularly smaller facilities or informal land uses that are often difficult to characterize using traditional bottom-up methods.

Although we leverage a localized, high-resolution prior in this study, INFE^{2}R does not require such a detailed inventory to function effectively. As demonstrated by our sensitivity analyses (see the "Broader applicability and limitations" section and sections S5 and S6), the framework can still recover location- and time-specific emission patterns even when the prior lacks spatial detail, including when replaced with a uniform field. This highlights that dense and representative observational constraints can reduce sensitivity to detailed prior magnitudes, but do not eliminate the value of spatial prior information. In particular, the flat-prior tests show that temporal dynamics can be recovered with little prior structure, whereas fine-scale spatial patterns improve substantially when land-use information is included through the covariance.

This approach enables quantitative attribution of emissions intensity and temporal behavior to specific locations—moving from anecdotal local knowledge to observation-constrained, policy-relevant evidence. Figure 3A highlights the 10 grid cells with the largest average increase in neighborhood emissions (posterior-minus-prior). An aerial view (Fig. 3B) reveals specific facilities within these grid cells, including a smog check station, trucking companies, and warehouses (table S2). The facility-level spatial resolution, combined with distinct diurnal profiles observed in adjacent grid cells (Fig. 3, C and D), demonstrates INFE^{2}R's ability to resolve both the location and timing of localized emission activity. Several high-correction pixels originated from uniform or near-zero prior values, and we observe both reallocation within sectors (from identified to previously unidentified locations) and the emergence of spatially missing sources. These locations therefore provide diagnostic evidence of spatial allocation errors in the prior, especially for facility-associated truck activity, warehouse

Downloaded from https://www.science.org on October 29, 2026

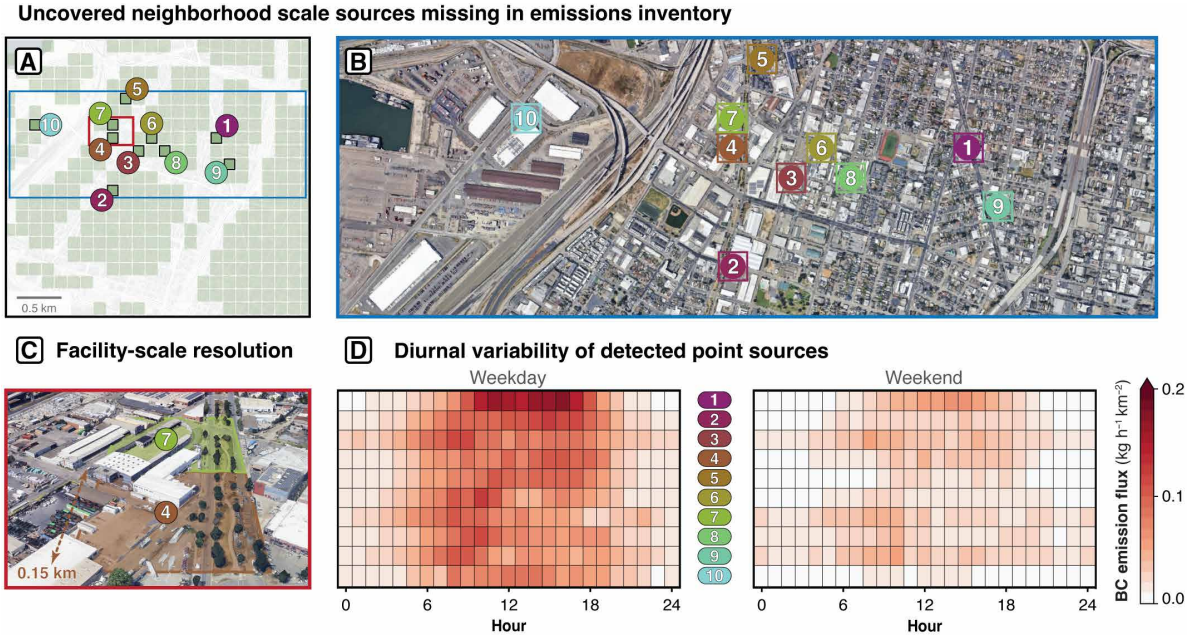


Fig. 3. Neighborhood scale emissions hotspots. We illustrate here the capability of the inversion framework to identify and quantify fine-scale local BC emission hotspots that are missing in the bottom-up inventory. (**A**) Here, we highlight the top 10 pixels (0.02 km^2 each) with the largest increases in the "neighborhood" sector. (**B**) Aerial imagery shows how individual pixels correspond to discrete facilities (table S2). (**C**) A zoomed-in aerial view shows the grid resolution, capturing emissions from two adjacent facilities with clear spatial delineation. (**D**) Distinct diurnal emission patterns for all 10 highlighted facilities reveal diverse activity profiles, including pronounced weekday/weekend variability. Notably, the two adjacent facilities in (C) exhibit distinct and contrasting temporal signals, further underscoring the framework's ability to resolve both spatial and temporal variations in emissions. Aerial imagery in (B) and (C) is from Google Earth (image source: Google 2023).

clusters, smaller roads, and port-adjacent interfaces that are difficult to represent using conventional activity-based proxies. In the core case (Fig. 3), multiple high-impact posterior locations were spatially absent or heavily muted in the prior, confirming that "previously unaccounted" includes missing or misallocated sources, not merely rescaled patterns. This finding is robust to a wide range of sensitivity scenarios (table S1). Many of these sources are located close to sensitive receptors such as residences and schools. By converting qualitative awareness into quantitative, time-resolved evidence, these fine-scale, source-linked emission corrections provide a direct pathway from hyperlocal observation to identifying context-specific mitigation pathways.

Connecting emissions to exposures

A key component of our approach is a source-receptor ("footprint") matrix **H** that quantifies how emissions from any grid cell influence concentrations at receptor locations across the domain. We leverage **H** and our observationally constrained emission estimates to understand how distinct emission sources contribute to BC concentrations—and thus, population exposures—in West Oakland. To do so, we combine grid cell concentration predictions (Fig. 4) with block-level US Census demographic data (see section S4.2).

Source-oriented predictions of BC concentrations differ substantially for the prior versus posterior inventories (Fig. 4). In the highest concentration areas, corrections shift predicted concentrations by up to 80% (Fig. 4, A and B), driven by revised emissions near key population centers. Posterior estimates indicate a more distributed source profile: exposures from major roads decreases, from 84% (prior) down to 68% (posterior), while the contribution from the other four source categories doubles from 16% (prior) to 32% (posterior) (Fig. 4C). Of those four, neighborhood and port-related sources together account for ~20% of BC exposures (posterior), or ~4× the prior value (~5%). These changes are not spatially uniform: concentrations from the port and neighborhood sources increase up to eightfold the prior level in certain areas (fig. S11), underscoring the effect of localized emission hotspots overlooked in prior estimates. Furthermore, for the most exposed population decile, the prior inventory underestimates population-weighted BC exposures by 50% (fig. S12), emphasizing the importance of the refined posterior estimates for accurately capturing exposure disparities. While the campaign-averaged change in domain-wide population-weighted exposure appears modest (~0.2 μg/m^3), this masks larger, localized differences at the native 150-m, hourly scale. Across space-time, the mean absolute posterior-prior difference is ~1 μg/m^3, with extremes up to 13 μg/m^3, concentrated in the most burdened neighborhoods. INFE^{2}R enables attribution of these changes to specific source categories, allowing previously obscured emissions patterns to be identified and traced to their origins in a way that aligns with observed concentrations. This redistribution in exposure shares does not assert that nonroad sources dominate emissions; rather, it shows that off-highway, facility-associated, and neighborhood-scale activities contribute more to exposure per unit emissions than previously captured, consistent with proximity effects embedded in **H**. Although some of the largest emissions corrections occur in port container yards with few residents, substantial concentration increases also occur in adjacent residential neighborhoods, with average BC levels rising by ~0.4 μg/m^3 (~50%) (Fig. 4). These areas, while often underrepresented in bottom-up inventories, are directly exposed to

Downloaded from https://www.science.org on October 29, 2026

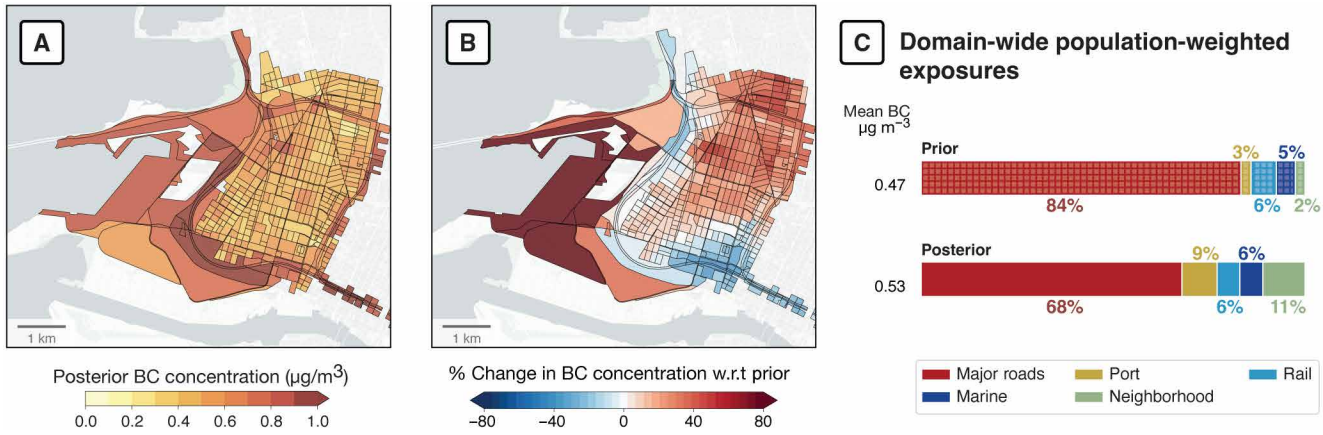


Fig. 4. Substantial revision to the spatial distribution and magnitude of community-scale BC exposures. (**A**) Census-block level BC concentration estimates simulated using inverse-modeled emissions show the sharp concentration gradients across West Oakland. (**B**) Percent change in BC concentrations compared to prior-based predictions, emphasizing how emission corrections reshape exposure landscapes. (**C**) Fractional contribution of each source sector to population-weighted mean BC exposures. Notably, the contribution of the four smallest source categories doubles from 16% to 32%, with an especially large increase for the "neighborhood" source category. Overall, while the domain-wide average exposure increases by ~13%, both the spatial distribution of exposure and the source contributions to these exposures are sharply revised.

facility-associated and off-road diesel activity. Moreover, the corrected sources are coemitters of fine particulate matter ($PM_{2.5}$), underscoring the relevance of these findings beyond BC alone. Crucially, the ability to resolve emission origins—rather than just magnitudes—enables spatially targeted mitigation strategies that better protect overburdened communities.

Thus far, we have demonstrated how use of hyperlocal observations provides sharp revisions to our understanding of the geography of emissions (i.e., Fig. 1A) and exposures (Fig. 4, B and C) in West Oakland. What connects these two endpoints, however, is a crucial question for effective air quality policy. Translating these improved estimates into actionable policy guidance requires identifying not just the largest emissions sources but those that disproportionately influence exposure due to proximity and local transport conditions (*47*, *48*). This distinction aligns with the policy concept of "indirect sources," wherein a facility's operations (e.g., shipping/receiving, warehousing, queueing, and idling) induce emissions from mobile equipment and trucks that are not stack emissions of the facility itself but nonetheless generate substantial nearby exposure (*9*–*11*). Our framework captures these effects by assigning emissions and exposure influence to the grid cells (and land-use sectors) where activity occurs, irrespective of legal ownership of the emissions.

We further decompose population exposure into the product of two terms (Fig. 5): domain-wide BC emissions E from a source category (kg hour^{-1}) and an exposure factor XF (μg m^{-3} per kg hour^{-1}) that reflects the efficiency with which an emitting sector produces exposures for a specific population (*49*). XF is computed as the population-weighted receptor concentration per unit emissions and incorporates the source-receptor influence (**H**), receptor population distribution, and spatial transport. It captures not only atmospheric dispersion and dilution, but also proximity to densely populated areas—factors that make some sources disproportionately impactful. Here, we compute XF as a campaign-average metric to emphasize persistent exposure drivers, although absolute values and sectoral rankings of XF vary across time-of-day and weekday/weekend subsets (fig. S13). This reflects genuine differences in how emissions translate into exposure depending on population presence and meteorological regimes, underscoring the importance of temporally resolved analyses for targeted intervention. To illustrate how this varies spatially, we split West Oakland into western and eastern zones, each home to approximately 20,000 people. Distinct sources within each zone differ markedly in their influence on exposure. In the eastern zone, neighborhood emission sources cause ~5× more exposure per unit emissions (i.e., XF) than do rail or ship emissions. XF is not simply the **H** matrix or a population-weighted concentration; it quantifies exposure per unit emissions and allows direct comparison among sectors with very different emission magnitudes. Beyond emission magnitude corrections, the posterior reveals a reordering in XF across sectors—highlighting sources that may be smaller in emissions but disproportionately high in exposure impact due to proximity and transport. In the prior model, roads not only account for the largest fraction of total exposure (82 and 85% in the W and E zones, respectively), but also have the highest XF. The posterior shifts our understanding of relative exposure drivers, revealing how other source categories (especially neighborhood sources)—though less dominant in the prior—contribute more strongly to population exposure than previously estimated.

Sources that cause high exposures typically have high emission rates (E) and/or close proximity to populations (high XF). Analogously to "super-emitters" in infrastructure systems, we find "super-exposer" sources with disproportionate impact on population exposure (*50*, *51*). For example, for one 0.6-km^2 neighborhood within West Oakland (fig. S14), 20% of the emissions (and 7% of pixels) in the domain account for 60% of neighborhood exposures. These high-impact sources are sprinkled across West Oakland, pointing toward the possible benefit of spatially targeted emissions mitigation efforts to protect specific especially sensitive receptors (*52*, *53*). Crucially, a single facility like a port or neighborhood industry may be more amenable to localized control measures than, for example, on-road vehicles, for which emission reductions may be limited by slow annual vehicle turnover rates and by limitations in local government

Downloaded from https://www.science.org on October 29, 2026

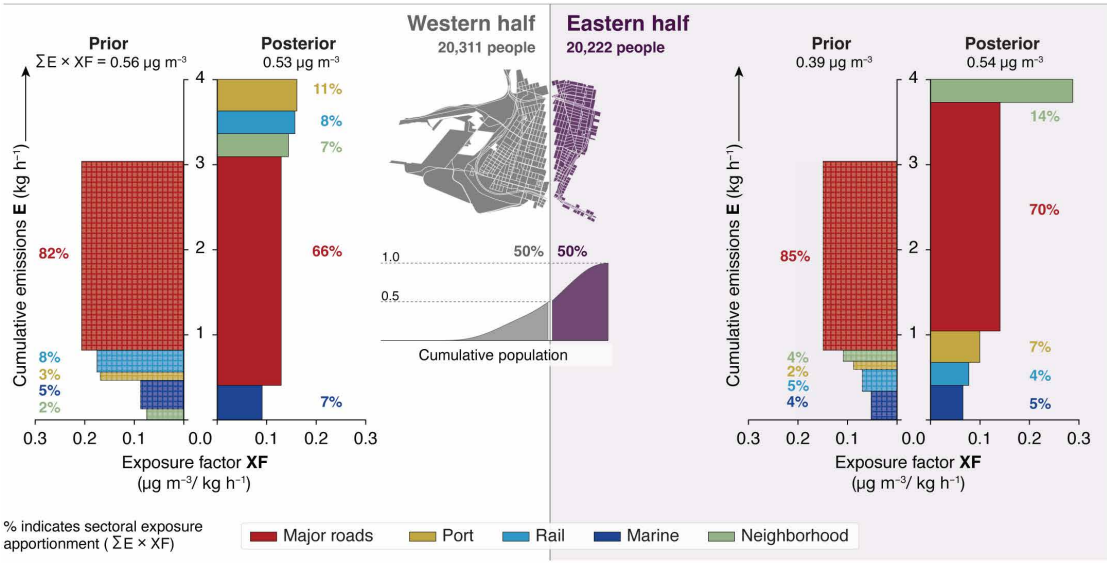


Fig. 5. Reordering of which sources have the highest marginal contribution to exposure. We divide West Oakland into two equal population halves (west and east) to illustrate how observationally constrained emissions estimates (posterior) update the understanding of which sources have the greatest impact on exposures. Exposures resulting from each source type are decomposed into two factors: emissions rate (kg hour^{-1}) and (ii) exposure factor XF, which indicates the population-weighted BC concentrations (and thus exposure mitigation potential) per unit emissions. The area of each plotted element represents its contribution to the BC exposure, with total concentrations indicated in the title of each plot. Sectors are ordered by XF from highest to lowest, highlighting which sectors generate the most exposure per unit emissions. Observationally constrained posterior estimates present a marked reordering of XF in addition to general increases in emissions estimates. The details vary substantially between west and east, suggesting the potential value of more locally targeted interventions. Overall, the relative importance of mitigating the nonroad source types increases markedly in our revised estimate. Indicated percentages quantify the sector-specific attribution of population-weighted mean BC concentrations. Note the sharp increases in the percentage contribution for the neighborhood and port sources, and the 39% increase in total BC exposure in the eastern zone.

authority (*54*). Thus, INFE^{2}R provides a scalable pathway to generate observation-driven emissions corrections and exposure insights—providing policy-relevant information to support locally tailored air quality interventions, even in settings with limited or outdated inventories—and doing so in a way that regular bottom-up inventories cannot. Without the fine-scale observational constraints embedded in INFE^{2}R, the influence of colocated and indirect sources, facility-induced truck activity, and non-truck diesel equipment on local exposures would remain obscured, leaving gaps in both source attribution and targeted mitigation planning.

Although our analysis focuses on BC, many highlighted sources also contribute to $PM_{2.5}$ mass through coemitted primary particles and precursors. BC is a major constituent of diesel PM (classified in California as a carcinogenic toxic air contaminant), and hyperlocal variation in BC has been associated with adverse cardiovascular outcomes in older adults (*35*, *55*). Prior intraurban studies further indicate that much of the spatial variation in PM mass is driven by BC and coemitted primary organic aerosol (*56*). Accordingly, BC-resolved corrections provide a high-resolution indicator of $PM_{2.5}$ relevant activity in urban combustion settings and can inform $PM_{2.5}$ exposure-reduction strategies where bottom-up inventories undercount facility activity, off-route truck traffic, or material handling. Linking BC-based corrections with $PM_{2.5}$ mass via source profiles or reduced-form modeling is a natural next step.

Broader applicability and limitations

INFE^{2}R provides a framework for generating high-resolution, observationally constrained emissions estimates that can inform community-scale air quality strategies. While demonstrated here for BC in West Oakland, the approach is transferable to a wide range of urban contexts, especially in communities that are overburdened by local pollution sources but lack data-driven emissions estimates. In such settings, coarser emissions data or proxy inventories (e.g., land use, traffic models, or emission factor-based surrogates) may serve as useful starting points. Practical implementation requires interdisciplinary expertise across hyperlocal measurements, atmospheric modeling, data assimilation, and community-engaged air quality management, as well as the availability of a high-resolution (150 m) meteorological simulation (i.e., WRF) to estimate **H**. While these prerequisites are nontrivial, our sensitivity analyses (figs. S15 to S20) indicate that the modeling approach is robust even when observational data are limited or prior emissions data contain substantial uncertainties, enhancing the potential feasibility of applying this framework in resource-constrained or data-scarce urban settings.

Practically, INFE^{2}R does not require a permanently dense network: A one-time, representative time-averaged mobile map can supply spatial detail, while an ongoing network of ~10 fixed monitors can provide temporal continuity. In down-selection tests, assimilating such a mobile map with ~10 fixed sites reproduced near-core posterior performance (spatial $R \geq 0.93$; figs. S16 and S17). These results suggest that future applications may not require the same level of permanent fixed-site monitoring used in the original West Oakland campaign. A representative mobile map plus a reduced fixed-site network can still be assimilated into a dense receptor field, here 549 receptor locations, while preserving hyperlocal spatial constraints. Computationally, the receptor-oriented STILT calculation scales with the number of assimilated receptors, modeled hours, and footprint duration, while daily inversions are independent and readily parallelized. This

Downloaded from https://www.science.org on October 29, 2026

reduced-monitor result should be interpreted for primary BC under the transport and observational conditions studied here. Applications to shorter-lived or chemically reactive pollutants would require source-receptor operators that capture the relevant chemistry, lifetime, and transport at appropriate scales. Nonetheless, the ability to generate robust fine-scale corrections despite imperfect priors is particularly valuable in contexts where structural inventory biases—such as underrepresentation of off-highway freight activity, small industrial sites, or indirect sources—are likely to recur across cities and years.

These strengths should be interpreted in light of the assumptions required for large Bayesian inversions. Our implementation assumes Gaussian prior and observational errors, uses a diagonal observational covariance for tractability, and solves for the posterior mean rather than the full posterior distribution. These assumptions do not imply that emissions or observations themselves are normally distributed; rather, they define an error model for the mismatch between prior emissions, forward predictions, and observations. Unresolved transport, representation, and covariance errors may remain, and future applications should revisit these assumptions where richer uncertainty information is available.

Because our framework attributes emissions to "source contexts" based on spatial land use rather than direct observations of emission activities, colocated sources—such as diesel trucks and yard equipment operating within the same facility—are pooled within the same grid cell. As such, our results reveal both misestimated activity levels and fundamental limitations in how conventional inventories disaggregate emissions. For example, truck activity occurring off designated freight routes or within facilities may go uncounted in bottom-up estimates that assume spatially fixed source profiles. Regular bottom-up inventories, without such observational constraints, cannot reliably detect these colocated or facility-induced emissions, leaving them hidden from policy action.

While the observational dataset used in this study is from 2017, its unique spatiotemporal resolution makes it well-suited for developing and validating new inversion techniques. The methodological insights gained here—particularly regarding data assimilation, sensitivity to priors, and integration of mobile and fixed-site measurements—can inform future deployments of INFE^{2}R in other cities and contexts. At the same time, the framework itself is not limited to historical analysis: It can be applied in updated monitoring campaigns to track sector-specific emission reductions and evaluate policy effectiveness over time. Since 2017, there have been substantial reductions in ambient BC levels in the region, presenting an opportunity to use new data to assess the real-world impact of regulatory interventions.

A crucial insight is that efficacy is driven by the combination of mobile and fixed-site data through data assimilation resulting in spatiotemporally rich observational constraints (see sections S5 and S6). The fixed-site monitors provide temporal continuity, while mobile measurements contribute fine spatial detail. The enhanced observational density provided by this approach increases the model's sensitivity to spatially resolved emissions, enabling precise, nuanced emission corrections (see section S3 and fig. S21 for details). We show explicitly that removing mobile monitoring data—thus relying solely on fixed-site observations—markedly reduces the inversion's ability to resolve local emission hotspots (figs. S15 and S16). Conversely, preserving mobile data and sharply reducing fixed-site inputs (as few as 10 randomly-selected sites) still yields robust inversions nearly identical to our core analysis (Pearson's $R \sim 0.93$, NMAE <10%, see figs. S16 and S17). These results suggest that by using data assimilation to incorporate the high spatial density information contained in mobile monitoring data, relying on a sparser fixed site network (<1 site km^{-2}) could still permit a robust inversion at our 150-m (~0.02 km^2) resolution.

We next explored the sensitivity of our results to the fidelity of the a priori emission inventory. We hypothesized that the very high density of the assimilated dataset (>50 observations km^{-2}) might reduce the dependence of our results on the accuracy of the prior emission inventory. Therefore, in a series of sensitivity analyses, we systematically degraded the prior inventory using both sector-wide and randomized grid-level perturbations (sections S5 and S6 and fig. S18). Specifically, we applied sector-specific scaling factors (0.25× to 4×) to emulate systematic directional bias and randomized grid-level perturbations within each sector to emulate spatially heterogeneous inventory errors, including local over- and underallocation of emissions. These randomized tests provide a complementary assessment of whether localized prior errors can project into the posterior. Notably, doubling or halving the assumed prior emissions from a given sector had virtually no impact on resulting posterior emissions estimates across the domain (temporal Pearson's $R \sim 0.96$ to 0.98; spatial $R > 0.99$). In a more extreme case, where the prior was eliminated altogether, the high density of observational constraints still allowed reconstruction of a temporally reliable emissions estimate ($R \sim 0.84$), though spatial precision was lower ($R \sim 0.45$, fig. S19). When the same flat prior mean was paired with a land-use–structured covariance (fig. S20), spatial agreement with the core posterior improved substantially ($R \sim 0.83$) while temporal skill was retained. This outcome shows that meaningful spatial patterns can be recovered when observations are paired with simple land-use covariance structure, even without detailed prior magnitudes.

Together, these analyses illustrate a key difference between inversions and CTMs: While CTMs are essential for process understanding and scenario testing, inversions like INFE^{2}R leverage observations to directly constrain emissions, reducing dependence on potentially biased priors. The two approaches are complementary—CTMs diagnose why emissions produce certain outcomes, while inversions diagnose what the emissions actually are. These findings imply that INFE^{2}R can be deployed in regions with less detailed or incomplete inventories. A localized prior is helpful, and some spatial structure in either the prior mean or covariance remains important for fine-scale spatial fidelity. Even when initialized with coarse or uniform priors, the framework identified spatially and temporally resolved emissions—including sources absent from inventory data—especially when land-use–structured covariance was applied (figs. S19 and S20). However, accuracy depends on measurement coverage and representativeness; in settings with sparse measurement data, missing inventory detail poses greater challenges and uncertainty.

Together, these sensitivity analyses support the applicability of our approach to other urban areas, including those with more limited monitoring networks and more uncertain priors. That said, certain prerequisites must still be met. Temporal resolution of observational constraints must be adequate to resolve hourly emission dynamics; this may not be feasible with heavily time-averaged data. While this study focused on BC, the method could, in principle, be extended to other pollutants such as $PM_{2.5}$ or NO_x, provided transport and chemistry can be reasonably linearized (e.g., via CTM adjoints or reduced-form models). Similarly, transient pollution events (e.g., wildfires, structural fires, or major disruptions such as

Downloaded from https://www.science.org on October 29, 2026

COVID-19 lockdowns) can be captured in our modeling approach, provided they are adequately represented in the observational dataset. These capabilities mean that INFE^{2}R is not locked to a single moment in time: repeating such analyses over time using updated observations could help track changes in urban emissions and evaluate the real-world impacts of mitigation policies, including those that have driven recent BC reductions in West Oakland since the time of data collection. More broadly, the framework is designed to be transferable: It requires only (i) a source-receptor operator resolved at spatial and temporal scales appropriate to local urban transport and (ii) observational constraints with fine spatial structure (from periodic mobile mapping) and modest temporal continuity (from a sparse fixed-site network). As long as the source-receptor operator can be reliably derived at the relevant scales, the framework remains robust even when prior inventories are coarse or uncertain. This supports applicability across diverse cities with differing land-use patterns, emission sources, and monitoring infrastructures.

While BC is not directly regulated, the same source contexts typically contribute to $PM_{2.5}$ mass. BC-based inversions can therefore guide $PM_{2.5}$ mitigation by identifying where primary particles and coemitted species are underestimated and by localizing the most impactful source contexts. For context, regional analyses have associated average BC levels of ~0.3 μg/m^3 with ~2 deaths per 100,000 and ~20 cardiovascular hospitalizations, underscoring the potential public-health value of the source-targeted reductions identified here (*55*).

Computationally, the framework is tractable and scalable: Each day of hourly emissions is solved independently and can be parallelized across cores. In our implementation, with precomputed STILT footprints, inversions required ~6 to 8 min/day on 16 CPU cores, enabling full 100-day campaigns to run within a few hours. Multipollutant applications are similarly feasible, particularly when footprints can be reused across pollutants. As such, city-scale or longitudinal deployments can be supported on modest high-performance computing resources.

Overall, our findings demonstrate how combining mobile and fixed-site monitoring data, atmospheric transport modeling, and inversion can yield observationally constrained, fine-scale emissions estimates that meaningfully revise our understanding of who is exposed to what—and from where. In a historically overburdened community such as West Oakland, this capability is itself environmental-justice relevant: Measurement-constrained emissions at 150 m and hourly resolution can support facility- and corridor-scale identification of exposure-relevant sources, even when the domain is too small to reliably quantify race/ethnicity-stratified disparities within the neighborhood. Our key finding, that smaller, distributed urban emission sources have a substantial and underappreciated impact on air pollution exposures, warrants further validation through studies in other settings. While our analysis used data from a past campaign, the INFE^{2}R framework is well suited for longitudinal application. If applied repeatedly with new data, this approach could help quantify emissions mitigation progress over time and assess the effectiveness of evolving regulatory policies. In doing so, it offers a replicable, observation-driven pathway for cities worldwide to uncover and address emissions sources that traditional bottom-up inventories alone cannot reliably detect.

MATERIALS AND METHODS

INFE^{2}R builds on the Bayesian inverse modeling paradigm, using measured pollutant concentration data to adjust and refine prior emission inventories. Its distinguishing feature—and primary policy relevance—lies in its ability to produce emissions estimates at a granular 150-m, hourly resolution. This capability is driven by two innovations: (i) the use of a dense observational constraint field (~50 data points per km^2) derived from assimilated fixed-site and mobile measurements, and (ii) a source-receptor transport matrix constructed at matching resolution, enabling fine-scale corrections by back-propagating mismatches between predicted and observed concentrations.

Such observational density is infeasible through conventional fixed-site monitoring alone. Instead, we leverage a previously published data assimilation technique that fuses mobile and fixed-site measurements into a spatiotemporally complete BC dataset. This assimilated dataset is equivalent to a uniform network of 450 sensors covering the 9-km^2 study domain at hourly resolution. The transport matrix—derived using the WRF–STILT model (*57*)—connects emissions at source to concentrations at receptor, preserving the 150-m resolution across space and time.

The following sections mirror the structure of the Bayesian inverse problem and its implementation. We begin with the forward model, which defines the likelihood by mapping emissions to predicted concentrations. Next, we describe the assimilated observational constraints. Last, we detail how the prior emissions estimate, forward model, observations, and error covariances are combined to compute posterior emission estimates.

Forward model

The forward model, as illustrated in fig. S1, uses a priori emissions data $(\boldsymbol{x_a})$ alongside a source-receptor transfer matrix or footprint matrix ($\mathbf{H}$) to estimate emissions-derived pollutant concentrations $(\boldsymbol{y_a})$ at designated receptor locations following Eq. 1. The forward model calculations were run at hourly resolution over a 100-day study period from 19 May to 26 August 2017, corresponding to the availability of observational data.

$$\boldsymbol{y_a} = \mathbf{H}\boldsymbol{x_a} \qquad \mathbf{H} \in \mathbb{R}^{\mathbf{m}\times\mathbf{n}}; \boldsymbol{x_a} \in \mathbb{R}^{\mathbf{n}\times\mathbf{1}} \tag{1}$$

Source receptor transfer, or footprints matrix ($\boldsymbol{H}$)

The forward model serves as the linkage between pollutant emissions at their sources and pollutant concentrations at receptor locations. In this study, we derive a "footprints" matrix, representing the transfer function between emission sources and receptors. This matrix is generated using the STILT model (*42*), which uses high-resolution meteorological inputs from a WRF model run (*41*). The matrix quantifies the time-varying sensitivity of pollutant concentrations at receptors to emissions from upwind grid cells. Footprint sensitivities, initially expressed in units of ppm/μmol m^2 s^{-1}, were converted to μg m^{-3}/g m^{-2} hour^{-1} (or observed BC concentration per unit BC emission flux) using an appropriate conversion factor, similar to Feng *et al.* (*58*).

WRF model setup

The WRF model is used to generate fine-scale meteorological fields, which serve as inputs for the STILT model to calculate footprints. The physics parameterizations and boundary conditions are consistent with those used and validated by Hamilton and Harley (*32*) for the same spatial domain and time period. WRF was run in a nested configuration (fig. S2) with three domains:

1) Domain D01: Covers a 400 km by 400 km area in Northern California (99 × 99 grid cells at a 4.05-km horizontal resolution).

Downloaded from https://www.science.org on October 29, 2026

2) Domain D02: Focuses on the greater Bay area, covering a 131 km by 131 km region (99 × 99 grid cells at a 1.35-km horizontal resolution).

3) Domain D03: Refined to focus on West Oakland, the primary study area, covering 9.3 km by 9.3 km (72 × 72 grid cells at a 150 m horizontal resolution).

The meteorological fields produced using these physics parameterizations and domain configurations were validated in prior work by Hamilton and Harley (*32*), ensuring their reliability for the study region and time frame. The West Oakland domain is dominated by low-rise residential, industrial, port, and freight-corridor land uses rather than dense high-rise street-canyon environments. Thus, while WRF-STILT does not explicitly resolve building-scale turbulence or street-canyon flow, prior high-resolution WRF-STILT applications in this same domain have demonstrated reliable performance for neighborhood-scale transport (*4*, *12*, *13*, *32*, *59*). Dense observational constraints can reduce sensitivity to random or site-specific transport errors in inverse estimates (*4*), although they do not eliminate coherent transport biases in the source-receptor operator. Remaining subgrid dispersion effects are not separately modeled but are implicitly reflected in the empirical spatial and temporal uncertainty terms used in the observational error covariance.

STILT model setup

The STILT model was configured to run within the fine-scale innermost domain (D03). Although WRF outputs covered all 72 × 72 grid cells, STILT simulations were restricted to a 62 × 62 grid, trimmed by five cells on each side to mitigate boundary effects. STILT simulations were conducted for each receptor location with observational constraints, linking these receptors back to the 62 × 62 grid used for the inverse modeling calculations.

Concentration sensitivities were computed using an ensemble of 1000 particles released from each receptor and transported backward in time. Tests with higher particle ensembles (5000 and 10,000 particles per receptor) changed footprint sensitivities by <2% on average, confirming that 1000 particles is sufficient for convergence in this domain. The residence time of these particles within emission grid cells quantified the influence of emissions on receptor concentrations. Each simulation ran for 6 hours, at which point more than 98% of particles had exited the innermost D03 domain. To capture contributions from sources outside this domain, we estimated background concentrations at each receptor and hour using prior emissions from the surrounding region (D02 excluding D03), combined with 24-hour STILT back trajectories. These regional background concentrations were incorporated into the observational error covariance matrix at each receptor to account for "background error" (see section S2 and fig. S22). On average, 99% of particles exited D02 within 24 hours, enabling us to capture regional influence while isolating local emissions during inversion. Because receptor observations can capture localized plumes sharper than a 150-m transport grid can resolve, some high-concentration events may contribute to representativeness error. This effect is partly accounted for in the observational error covariance and partly mitigated by the least-squares inverse formulation and sector-structured prior covariance, which discourage isolated single-pixel corrections and promote locally coherent adjustments among neighboring pixels of the same source context (see section S2 for more details).

STILT incorporates meteorological influences such as relative humidity, precipitation scavenging, and boundary layer mixing by directly ingesting WRF meteorological outputs at hourly resolution. Precipitation and wet deposition processes are modeled by probabilistically removing particles when precipitation occurs along the trajectory, while turbulent mixing and entrainment are treated via stochastic perturbations to particle velocities informed by modeled turbulence parameters (*42*). This coupling allows STILT to capture the impacts of changing atmospheric stability, wind shear, and precipitation on particle trajectories and sensitivities.

STILT has been widely applied and validated in a range of studies to estimate source contributions to pollutant concentrations (*57*, *60*, *61*). These applications demonstrate that STILT effectively reproduces observed concentration patterns in urban and regional contexts when driven by high-resolution meteorological fields, supporting its suitability for fine-scale inversion studies such as this work.

While this approach effectively captures regional background using modeled prior emissions, it does not account for episodic or uninventoried background events (e.g., wildfires) originating outside the domain. However, no major wildfire activity affected the Bay Area during the study period, and as shown in fig. S22, regional background concentrations remained a small fraction of observed values—about one-third of the 5th percentile across the domain. Even if wildfire smoke had doubled or tripled those values, the overall concentration patterns would still be dominated by local emissions, which drive the spatial gradients that our inversion is designed to resolve.

Prior emission inventory (x_a)

The prior inventory was adapted from a bottom-up BC emission inventory for West Oakland by Hamilton and Harley (*32*), previously used for WRF simulations at 150 m resolution (*32*, *41*) and to compare these predictions with measured BC concentrations. The inventory is structured across three nested WRF domains (D01 to D03, see fig. S2). While emission corrections are derived for the innermost domain (D03), prior emissions from D02 (excluding D03) are used to estimate background concentrations error estimates in D03 (see section S2).

The bottom-up inventory was constructed using a combination of measured and parameterized data, including traffic counts, road density, and on-road measurements of emission factors, supplemented by reported data on port operations and locomotive emissions. Emission factors for major sources, such as drayage and over-the-road trucks contributing to on-road vehicle emissions, were derived from in situ measurements conducted over the past decade, tracking fleet-specific variations within the study area (*38*–*40*). For further details, see Hamilton and Harley (*32*).

The original inventory includes seven emission sectors. For this study, we merged two on-road vehicle categories (light-duty and heavy-duty vehicles) and two port-related marine vessel categories (ocean-going vessels and harbor craft) into consolidated groups. This reorganization results in five emission source categories: major roads; rail; the Port of Oakland ("port"), located on land; marine, consisting of ships and harbor craft; and a large assortment of mostly smaller-scale sources distributed throughout the community ("neighborhood sources"). This grouping reflects the dominant land use or activity in each grid cell rather than the precise technological origin of emissions. For instance, diesel truck activity may contribute to multiple categories, directly or indirectly, but each represents a distinct regulatory and emissions control context. This categorization enables spatial segmentation and comparison of prior and posterior emissions by sector and location (see section S4.1 for details).

Downloaded from https://www.science.org on October 29, 2026

Observational constraints (y_o)

In this study, observational constraints are derived from the West Oakland Community Air Quality Study, an intensive 100-day measurement campaign conducted from 19 May to 27 August 2017 in West Oakland, CA. This campaign included a fixed-site measurement network deployed by Caubel *et al.* (*44*), consisting of approximately low-cost 100 BC sensors (*62*) strategically distributed across residential, industrial, and high-traffic locations. The network used the Aerosol Black Carbon Detector (ABCD), a custom low-cost sensor platform previously evaluated for neighborhood-scale BC mapping. Although not a commercial instrument, the ABCD provided high precision for this application (MAE: 0.02 μg/m^3) and with valid observations available for 84% of hours on average [see section S1; Caubel *et al.* (*44*) and Manchanda *et al.* (*43*)]. Complementing the fixed-site network, two Google Street View vehicles, custom-equipped with fast-response gas and particle instrumentation (*63*), repeatedly surveyed every street in the study area. Fixed-site data completeness and QA/QC procedures are documented in Caubel *et al.* (*44*), while the paired mobile and fixed-site measurements entering the assimilation framework are described in Chambliss *et al.* (*45*). The assimilated observational product used here, including its validation and uncertainty characterization, is described in Manchanda *et al.* (*43*). Further details on measurement approach and instrumentation are provided in section S1.

To transform these measurements into observational constraints suitable for inverse modeling, we used a recently published data assimilation technique (*43*) (see fig. S1). Unlike traditional data assimilation methods that require predefined model states, this technique leverages inherent cross-dimensional relationships between pollutant concentrations driven by shared emission sources and meteorological conditions. Specifically, spatially dense but temporally sparse mobile multipollutant data fill spatial gaps in the temporally continuous, but spatially sparse, fixed-site BC observations. This process yields a spatiotemporally complete BC concentration dataset. Validation experiments previously conducted by Manchanda *et al.* (*43*) demonstrated that this assimilation method reproduced out-of-sample fixed-site observations with high fidelity (Pearson's $R \sim 0.95$; NMAE ~ 0.20).

For the present analysis, we adjusted the spatiotemporal resolution of the assimilation outputs to align with the meteorological and emission inputs required by our inversion framework. Whereas Manchanda *et al.* (*43*) originally applied this assimilation at 30-m spatial and 15-min temporal resolution, here we aggregated data to a coarser 150-m spatial resolution (150-m point buffers spaced along roads; fig. S3) and hourly time steps. While distance-decay relationships for BC have been shown to be steep at scales below 150 m (*63*), our assimilation product is designed to reflect spatial structure at scales resolvable by the transport model. As such, sub–150-m gradients are not explicitly resolved in the inversion but are implicitly captured in the form of smoothed observational constraints that are internally consistent with the resolution of the inversion framework. This resolution strikes a balance between computational feasibility and representational accuracy, resulting in assimilated observations for 549 receptor locations over the 100-day study period (2400 hourly time steps). Revalidation at this 150-m (coarser) resolution confirmed similar high performance, reproducing out-of-sample fixed-site observations with Pearson's $R \sim 0.96$ and NMAE ~ 0.18.

As shown in Manchanda *et al.* (*43*), the mobile component need not be continuous and coinciding with the fixed site measurements: A representative, time-averaged mobile map generated over a finite campaign can be reused to supply spatial detail, with a periodic targeted refresh if source patterns change materially. In this study, we treat the assimilated dataset as the "truth" for inversion purposes, and we evaluate inversion performance not only by its ability to reproduce out-of-sample assimilated observations, but also via comparison against fully independent fixed-site sensors that were entirely withheld from the data assimilation step (i.e., sensors not contributing data to any assimilated product; see section S1.4). In addition, we performed sensitivity analyses to evaluate the impact of observational density on emission corrections (table S1 and figs. S14 to S16). These included scenarios where observational constraints relied solely on the underlying fixed-site measurements, without mobile data assimilation. We further explored the scalability and accuracy trade-off by assessing how performance varies when assimilation is conducted with progressively fewer fixed-site sensors, following prior demonstrations that similar observational density can be achieved with fewer inputs (*43*) (see table S1 and section S5). Together, these analyses demonstrate the robustness of the inversion approach under varied observational scenarios, including those with limited or spatially sparse measurements.

Bayesian estimation of emission corrections

The Bayesian inversion framework updates prior emissions estimates (x_a) using the mismatch between observational constraints (y_o) and forward model predictions (Hx_a). This mismatch forms the basis of an inference problem to computing posterior emission estimates. Assuming Gaussian-distributed errors, a closed-form solution for the posterior (x_c) can be derived as shown below in Eq. 2 [cf. Rodgers (*64*)]

$$x_c = x_a + (HS_a)^T (HS_a H^T + S_o)^{-1} (y_o - Hx_a) \qquad S_o \in \mathbb{R}^{n\times n}; S_a \in \mathbb{R}^{m\times m} \tag{2}$$

Here, x_c is the vector of posterior emission estimates, S_o is the observational error covariance matrix, and S_a is the prior error covariance matrix. We solve the inverse problem separately for each day of emissions. The H matrix is constructed from STILT-generated footprints that trace concentrations to emissions up to 6 hours back in time. This enables a given day's emissions (24 hours) to influence up to 30 hours of observed concentrations. The likelihood is defined by the mismatch between observed concentrations and forward-model predictions, weighted by S_o. Thus, the model enters the Bayesian update through H, while the prior and observational uncertainties enter through S_a and S_o, respectively.

The observational error covariance (S_o) is treated as a diagonal matrix, assuming observational errors are uncorrelated spatiotemporally. In line with previous studies, each diagonal element represents a total error σ_m, as a composite of measurements uncertainty and background variability. Here, σ_m aggregates instrument noise, assimilation uncertainty, and background/representation error, ensuring posterior updates are uncertainty weighted. Constructing a fully populated prior error covariance matrix (S_a) would result in more than 10^{12} elements for a single day's inversion. To maintain computational tractability, and in line with previous literature (*3*, *4*, *12*, *65*), S_a was modeled as a Kronecker product of a temporal (D) and a spatial covariance (E) matrix (see section S2 for more details).

Downloaded from https://www.science.org on October 29, 2026

In this formulation, spatial correlations are segmented by emission sector. For example, grid cells corresponding to the "roads" sector should be correlated with other road sector grid cells but not with rail or marine sectors. An exponential decay in correlation is imposed in both space and time, ensuring nearby grid cells have higher correlations. This means that if one road pixel is over- or underestimated for a given time step, adjacent road grid cells and time steps are more likely to experience similar biases. However, this effect should not extend to nearby grid cells from other sectors. For instance, if a traffic jam was missed in the prior emissions but captured by the observational constraints, it is likely to have caused an underestimation in emissions for a subset of road segments for 1 to 2 hours, but it should not affect errors in emissions from other sectors in the vicinity.

Ultimately, this sector-specific, spatially explicit covariance structure leverages the strength of INFE^{2}R—combining dense observational constraints and fine-resolution transport modeling—to deliver highly resolved and observationally constrained emission estimates. These posterior estimates form the foundation for the downstream analyses presented in this study, including evaluations of sectoral contributions, temporal trends, exposure disparities, and robustness to data limitations.

Supplementary Materials

This PDF file includes:

Supplementary Text
Figs. S1 to S22
Tables S1 and S2
References

REFERENCES

1. J. D. Marshall, E. Nethery, M. Brauer, Within-urban variability in ambient air pollution: Comparison of estimation methods. *Atmos. Environ.* **42**, 1359–1369 (2008).
2. J. S. Apte, C. Manchanda, High-resolution urban air pollution mapping. *Science* **385**, 380–385 (2024).
3. A. J. Turner, J. Kim, H. Fitzmaurice, C. Newman, K. Worthington, K. Chan, P. J. Wooldridge, P. Köehler, C. Frankenberg, R. C. Cohen, Observed impacts of COVID-19 on urban CO_2 emissions. *Geophys. Res. Lett.* **47**, e2020GL090037 (2020).
4. A. J. Turner, A. A. Shusterman, B. C. McDonald, V. Teige, R. A. Harley, R. C. Cohen, Network design for quantifying urban CO_2 emissions: Assessing trade-offs between precision and network density. *Atmos. Chem. Phys.* **16**, 13465–13475 (2016).
5. H. Nesser, D. J. Jacob, J. D. Maasakkers, A. Lorente, Z. Chen, X. Lu, L. Shen, Z. Qu, M. P. Sulprizio, M. Winter, S. Ma, A. A. Bloom, J. R. Worden, R. N. Stavins, C. A. Randles, High-resolution US methane emissions inferred from an inversion of 2019 TROPOMI satellite data: Contributions from individual states, urban areas, and landfills. *Atmos. Chem. Phys.* **24**, 5069–5091 (2024).
6. P. Bergamaschi, A. Segers, D. Brunner, J.-M. Haussaire, S. Henne, M. Ramonet, T. Arnold, T. Biermann, H. Chen, S. Conil, M. Delmotte, G. Forster, A. Frumau, D. Kubistin, X. Lan, M. Leuenberger, M. Lindauer, M. Lopez, G. Manca, J. Müller-Williams, S. O'Doherty, B. Scheeren, M. Steinbacher, P. Trisolino, G. Vítková, C. Yver Kwok, High-resolution inverse modelling of European CH_4 emissions using the novel FLEXPART-COSMO TM5 4DVAR inverse modelling system. *Atmos. Chem. Phys.* **22**, 13243–13268 (2022).
7. M. Göckede, A. M. Michalak, D. Vickers, D. P. Turner, B. E. Law, Atmospheric inverse modeling to constrain regional-scale CO_2 budgets at high spatial and temporal resolution. *J. Geophys. Res. Atmos.* **115**, D15113 (2010).
8. I. Lopez-Coto, X. Ren, O. E. Salmon, A. Karion, P. B. Shepson, R. R. Dickerson, A. Stein, K. Prasad, J. R. Whetstone, Wintertime CO_2, CH_4, and CO emissions estimation for the Washington, DC–Baltimore metropolitan area using an inverse modeling technique. *Environ. Sci. Technol.* **54**, 2606–2614 (2020).
9. V. A. Lang, S. F. Camilleri, S. van der Lee, G. Rowangould, B. Antonczak, T. M. Thompson, M. H. Harris, C. Harkins, D. Q. Tong, M. Janssen, Z. E. Adelman, D. E. Horton, Intercomparison of modeled urban-scale vehicle NO_x and $PM_{2.5}$ emissions—Implications for equity assessments. *Environ. Sci. Technol.* **59**, 4560–4570 (2025).
10. S. F. Camilleri, A. Montgomery, M. A. Visa, J. L. Schnell, Z. E. Adelman, M. Janssen, E. A. Grubert, S. C. Anenberg, D. E. Horton, Air quality, health and equity implications of electrifying heavy-duty vehicles. *Nat. Sustain.* **6**, 1643–1653 (2023).
11. G. H. Kerr, M. Meyer, D. L. Goldberg, J. Miller, S. C. Anenberg, Air pollution impacts from warehousing in the United States uncovered with satellite data. *Nat. Commun.* **15**, 6006 (2024).
12. N. G. Asimow, A. J. Turner, R. C. Cohen, Sustained reductions of bay area CO_2 emissions 2018–2022. *Environ. Sci. Technol.* **58**, 6586–6594 (2024).
13. Y. Zhu, M. Y. Patel, A. R. Winter, N. G. Asimow, R. C. Cohen, Observational inferences of NO_x and CO emission factors for vehicles and homes in the San Francisco Bay Area. *ACS EST Air* **2**, 1478–1487 (2025).
14. V. Isakov, J. S. Touma, J. Burke, D. T. Lobdell, T. Palma, A. Rosenbaum, H. Özkaynak, Combining regional- and local-scale air quality models with exposure models for use in environmental health studies. *J. Air Waste Manage. Assoc.* **59**, 461–472 (2009).
15. D. K. Joe, H. Zhang, S. P. DeNero, H.-H. Lee, S.-H. Chen, B. C. McDonald, R. A. Harley, M. J. Kleeman, Implementation of a high-resolution source-oriented WRF/Chem model at the port of Oakland. *Atmos. Environ.* **82**, 351–363 (2014).
16. T. Oda, R. Bun, V. Kinakh, P. Topylko, M. Halushchak, G. Marland, T. Lauvaux, M. Jonas, S. Maksyutov, Z. Nahorski, M. Lesiv, O. Danylo, J. Horabik-Pyzel, Errors and uncertainties in a gridded carbon dioxide emissions inventory. *Mitig. Adapt. Strat. Glob. Chang.* **24**, 1007–1050 (2019).
17. G. Lu, E. A. Marais, K. Vohra, R. P. Horner, D. Zhang, R. V. Martin, S. Guttikunda, Near-automated estimate of city nitrogen oxides emissions applied to South and Southeast Asia. *J. Geophys. Res. Atmos.* **130**, e2024JD041000 (2025).
18. A. Thomasson, P. Roldin, N. Schutgens, B. George, H. Denier van der Gon, G. Monteil, M. Scholze, Towards better black carbon emission estimates in Europe: Assimilating observations with a Bayesian inversion framework [Preprint] (2025). https://doi.org/10.5194/egusphere-2025-1568.
19. A. Mols, K. F. Boersma, H. Denier van der Gon, M. Krol, An improved Bayesian inversion to estimate daily NO_x emissions of Paris from TROPOMI NO_2 observations between 2018–2023. *Atmos. Chem. Phys.* **26**, 1497–1513 (2026).
20. N. Evangeliou, R. L. Thompson, S. Eckhardt, A. Stohl, Top-down estimates of black carbon emissions at high latitudes using an atmospheric transport model and a Bayesian inversion framework. *Atmos. Chem. Phys.* **18**, 15307–15327 (2018).
21. J. J. Guerrette, D. K. Henze, Four-dimensional variational inversion of black carbon emissions during ARCTAS-CARB with WRFDA-Chem. *Atmos. Chem. Phys.* **17**, 7605–7633 (2017).
22. A. Valencia, S. Arunachalam, V. Isakov, B. Naess, M. Serre, Improving emissions inputs via mobile measurements to estimate fine-scale Black Carbon monthly concentrations through geostatistical space-time data fusion. *Sci. Total Environ.* **793**, 148378 (2021).
23. C. Lee, R. V. Martin, A. van Donkelaar, H. Lee, R. R. Dickerson, J. C. Hains, N. Krotkov, A. Richter, K. Vinnikov, J. J. Schwab, SO_2 emissions and lifetimes: Estimates from inverse modeling using in situ and global, space-based (SCIAMACHY and OMI) observations. *J. Geophys. Res.* **116**, D06304 (2011).
24. D. G. Streets, T. Canty, G. R. Carmichael, B. de Foy, R. R. Dickerson, B. N. Duncan, D. P. Edwards, J. A. Haynes, D. K. Henze, M. R. Houyoux, D. J. Jacob, N. A. Krotkov, L. N. Lamsal, Y. Liu, Z. Lu, R. V. Martin, G. G. Pfister, R. W. Pinder, R. J. Salawitch, K. J. Wecht, Emissions estimation from satellite retrievals: A review of current capability. *Atmos. Environ.* **77**, 1011–1042 (2013).
25. J. K. Hedelius, J. Liu, T. Oda, S. Maksyutov, C. M. Roehl, L. T. Iraci, J. R. Podolske, P. W. Hillyard, J. Liang, K. R. Gurney, D. Wunch, P. O. Wennberg, Southern California megacity CO_2, CH_4, and CO flux estimates using ground- and space-based remote sensing and a Lagrangian model. *Atmos. Chem. Phys.* **18**, 16271–16291 (2018).
26. A. J. Turner, D. J. Jacob, K. J. Wecht, J. D. Maasakkers, E. Lundgren, A. E. Andrews, S. C. Biraud, H. Boesch, K. W. Bowman, N. M. Deutscher, M. K. Dubey, D. W. T. Griffith, F. Hase, A. Kuze, J. Notholt, H. Ohyama, R. Parker, V. H. Payne, R. Sussmann, C. Sweeney, V. A. Velazco, T. Warneke, P. O. Wennberg, D. Wunch, Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data. *Atmos. Chem. Phys.* **15**, 7049–7069 (2015).
27. T. S. Jones, J. E. Franklin, J. Chen, F. Dietrich, K. D. Hajny, J. C. Paetzold, A. Wenzel, C. Gately, E. Gottlieb, H. Parker, M. Dubey, F. Hase, P. B. Shepson, L. H. Mielke, S. C. Wofsy, Assessing urban methane emissions using column-observing portable Fourier transform infrared (FTIR) spectrometers and a novel Bayesian inversion framework. *Atmos. Chem. Phys.* **21**, 13131–13147 (2021).
28. D. H. Cusworth, A. A. Bloom, S. Ma, C. E. Miller, K. Bowman, Y. Yin, J. D. Maasakkers, Y. Zhang, T. R. Scarpelli, Z. Qu, D. J. Jacob, J. R. Worden, A Bayesian framework for deriving sector-based methane emissions from top-down fluxes. *Commun. Earth Environ.* **2**, 242 (2021).
29. A. Mols, K. F. Boersma, H. Denier van der Gon, M. Krol, An improved Bayesian inversion to estimate daily NO_x emissions of Paris from TROPOMI NO_2 observations between 2018–2023 [Preprint] (2025). https://doi.org/10.5194/egusphere-2025-49.
30. J. R. Pitt, I. Lopez-Coto, A. Karion, K. D. Hajny, J. Tomlin, R. Kaeser, T. Jayarathne, B. H. Stirm, C. R. Floerchinger, C. P. Loughner, R. Commane, C. K. Gately, L. R. Hutyra, K. R. Gurney, G. S. Roest, J. Liang, S. Gourdji, K. L. Mueller, J. R. Whetstone, P. B. Shepson, Underestimation of thermogenic methane emissions in New York City. *Environ. Sci. Technol.* **58**, 9147–9157 (2024).
31. Bay Area Air Quality Management District, West Oakland Environmental Indicators Project, "Owning Our Air: The West Oakland Community Action Plan" (West Oakland, 2019); https://www.baaqmd.gov/community-health/community-health-protection-program/west-oakland-community-action-plan.

Downloaded from https://www.science.org on October 29, 2026

32. S. D. Hamilton, R. A. Harley, High-resolution modeling and apportionment of diesel-related contributions to black carbon concentrations. *Environ. Sci. Technol.* **55**, 12250–12260 (2021).
33. L. MacIver, J. London, N. Sampson, M. Gordon, R. Grow, V. Eady, West Oakland's experience in building community power to confront environmental injustice through California's Assembly Bill 617. *Am. J. Public Health* **112**, 262–270 (2022).
34. D. E. Goin, S. Sudat, C. Riddell, R. Morello-Frosch, J. S. Apte, M. M. Glymour, D. Karasek, J. A. Casey, Hyperlocalized measures of air pollution and preeclampsia in Oakland, California. *Environ. Sci. Technol.* **55**, 14710–14719 (2021).
35. S. E. Alexeeff, A. Roy, J. Shan, X. Liu, K. Messier, J. S. Apte, C. Portier, S. Sidney, S. K. Van Den Eeden, High-resolution mapping of traffic related air pollution with Google street view cars and incidence of cardiovascular events within neighborhoods in Oakland, CA. *Environ. Health* **17**, 38 (2018).
36. C. Zhao, A. E. Andrews, L. Bianco, J. Eluszkiewicz, A. Hirsch, C. MacDonald, T. Nehrkorn, M. L. Fischer, Atmospheric inverse estimates of methane emissions from Central California. *J. Geophys. Res. Atmos.* **114**, D16302 (2009).
37. D. K. Henze, J. H. Seinfeld, D. T. Shindell, Inverse modeling and mapping US air quality influences of inorganic $PM_{2.5}$ precursor emissions using the adjoint of GEOS-Chem. *Atmos. Chem. Phys.* **9**, 5877–5903 (2009).
38. C. V. Preble, T. E. Cados, R. A. Harley, T. W. Kirchstetter, In-use performance and durability of particle filters on heavy-duty diesel trucks. *Environ. Sci. Technol.* **52**, 11913–11921 (2018).
39. T. R. Dallmann, S. J. DeMartini, T. W. Kirchstetter, S. C. Herndon, T. B. Onasch, E. C. Wood, R. A. Harley, On-road measurement of gas and particle phase pollutant emission factors for individual heavy-duty diesel trucks. *Environ. Sci. Technol.* **46**, 8511–8518 (2012).
40. T. R. Dallmann, T. W. Kirchstetter, S. J. DeMartini, R. A. Harley, Quantifying on-road emissions from gasoline-powered motor vehicles: Accounting for the presence of medium- and heavy-duty diesel trucks. *Environ. Sci. Technol.* **47**, 13873–13881 (2013).
41. J. G. Powers, J. B. Klemp, W. C. Skamarock, C. A. Davis, J. Dudhia, D. O. Gill, J. L. Coen, D. J. Gochis, R. Ahmadov, S. E. Peckham, G. A. Grell, J. Michalakes, S. Trahan, S. G. Benjamin, C. R. Alexander, G. J. Dimego, W. Wang, C. S. Schwartz, G. S. Romine, Z. Liu, C. Snyder, F. Chen, M. J. Barlage, W. Yu, M. G. Duda, The weather research and forecasting model: Overview, system efforts, and future directions. *Bull. Am. Meteorol. Soc.* **98**, 1717–1737 (2017).
42. J. C. Lin, C. Gerbig, S. C. Wofsy, A. E. Andrews, B. C. Daube, K. J. Davis, C. A. Grainger, A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model. *J. Geophys. Res. Atmos.* **108**, 4493 (2003).
43. C. Manchanda, R. A. Harley, J. D. Marshall, A. J. Turner, J. S. Apte, Integrating mobile and fixed-site black carbon measurements to bridge spatiotemporal gaps in urban air quality. *Environ. Sci. Technol.* **58**, 12563–12574 (2024).
44. J. J. Caubel, T. E. Cados, C. V. Preble, T. W. Kirchstetter, A distributed network of 100 black carbon sensors for 100 days of air quality monitoring in West Oakland, California. *Environ. Sci. Technol.* **53**, 7564–7573 (2019).
45. S. E. Chambliss, C. V. Preble, J. J. Caubel, T. Cados, K. P. Messier, R. A. Alvarez, B. Lafranchi, M. Lunden, J. D. Marshall, A. A. Szpiro, T. W. Kirchstetter, J. S. Apte, Comparison of mobile and fixed-site black carbon measurements for high-resolution urban pollution mapping. *Environ. Sci. Technol.* **54**, 7848–7857 (2020).
46. M. Kopacz, D. J. Jacob, D. K. Henze, C. L. Heald, D. G. Streets, Q. Zhang, Comparison of adjoint and analytical Bayesian inversion methods for constraining Asian sources of carbon monoxide using satellite (MOPITT) measurements of CO columns. *J. Geophys. Res. Atmos.* **114**, D04305 (2009).
47. D. H. Bennett, T. E. McKone, J. S. Evans, W. W. Nazaroff, M. D. Margni, O. Jolliet, K. R. Smith, Defining intake fraction. *Environ. Sci. Technol.* **36**, 206A–211A (2002).
48. K. R. Smith, Place makes the poison: Wesolowski Award Lecture—1999. *J. Expo. Sci. Environ. Epidemiol.* **12**, 167–171 (2002).
49. L. H. Koolik, R. D. Bullard, E. Min, R. Morello-Frosch, R. Patterson, M. Salgado, N. Wedekind, J. D. Marshall, J. S. Apte, Eliminating air pollution disparities requires more than emission reduction. *Proc. Natl. Acad. Sci. U.S.A.* **122**, e2505888122 (2025).
50. R. Sawyer, R. A. Harley, S. H. Cadle, J. M. Norbeck, R. Slott, H. A. Bravo, Mobile sources critical review: 1998 NARSTO assessment. *Atmos. Environ.* **34**, 2161–2181 (2000).
51. D. Zavala-Araiza, R. A. Alvarez, D. R. Lyon, D. T. Allen, A. J. Marchese, D. J. Zimmerle, S. P. Hamburg, Super-emitters in natural gas infrastructure are caused by abnormal process conditions. *Nat. Commun.* **8**, 14012 (2017).
52. Y. Wang, J. S. Apte, J. D. Hill, C. E. Ivey, R. F. Patterson, A. L. Robinson, C. W. Tessum, J. D. Marshall, Location-specific strategies for eliminating US national racial-ethnic $PM_{2.5}$ exposure inequality. *Proc. Natl. Acad. Sci. U.S.A* **119**, e2205548119 (2022).
53. N. P. Nguyen, J. D. Marshall, Impact, efficiency, inequality, and injustice of urban air pollution: Variability by emission location. *Environ. Res. Lett.* **13**, 024002 (2018).
54. L. H. Koolik, Á. Alvarado, A. Budahn, L. Plummer, J. D. Marshall, J. S. Apte, $PM_{2.5}$ exposure disparities persist despite strict vehicle emissions controls in California. *Sci. Adv.* **10**, eadn8544 (2024).
55. V. A. Southerland, S. C. Anenberg, M. Harris, J. Apte, P. Hystad, A. van Donkelaar, R. V. Martin, M. Beyers, A. Roy, Assessing the distribution of air pollution health risks within cities: A neighborhood-scale analysis leveraging high-resolution data sets in the Bay Area, California. *Environ. Health Perspect.* **129**, 37006 (2021).
56. P. Gu, H. Z. Li, Q. Ye, E. S. Robinson, J. S. Apte, A. L. Robinson, A. A. Presto, Intracity variability of particulate matter exposure is driven by carbonaceous sources and correlated with land use variables. *Environ. Sci. Technol.* **52**, 11545–11554 (2018).
57. T. Nehrkorn, J. Eluszkiewicz, S. C. Wofsy, J. C. Lin, C. Gerbig, M. Longo, S. Freitas, Coupled weather research and forecasting–stochastic time-inverted lagrangian transport (WRF–STILT) model. *Meteorol. Atmos. Phys.* **107**, 51–64 (2010).
58. X. Feng, L. J. Mickley, M. L. Bell, T. Liu, J. A. Fisher, M. V. Martin, Improved estimates of smoke exposure during Australia fire seasons: Importance of quantifying plume injection heights. *Atmos. Chem. Phys.* **24**, 2985–3007 (2024).
59. N. Dadheech, A. J. Turner, Simulating out-of-sample atmospheric transport to enable flux inversions. *Atmos. Chem. Phys.* **26**, 427–441 (2026).
60. J. C. Lin, B. Fasoli, L. Mitchell, R. Bares, F. Hopkins, T. M. Thompson, R. A. Alvarez, Towards hyperlocal source identification of pollutants in cities by combining mobile measurements with atmospheric modeling. *Atmos. Environ.* **311**, 119995 (2023).
61. B. Fasoli, J. C. Lin, D. R. Bowling, L. Mitchell, D. Mendoza, Simulating atmospheric tracer concentrations for spatially distributed receptors: Updates to the Stochastic Time-Inverted Lagrangian Transport model's R interface (STILT-R version 2). *Geosci. Model Dev.* **11**, 2813–2824 (2018).
62. J. J. Caubel, T. E. Cados, T. W. Kirchstetter, A new black carbon sensor for dense air quality monitoring networks. *Sensors* **18**, 738 (2018).
63. J. S. Apte, K. P. Messier, S. Gani, M. Brauer, T. W. Kirchstetter, M. M. Lunden, J. D. Marshall, C. J. Portier, R. C. H. Vermeulen, S. P. Hamburg, High-resolution air pollution mapping with Google Street View cars: Exploiting big data. *Environ. Sci. Technol.* **51**, 6999–7008 (2017).
64. C. D. Rodgers, *Inverse Methods for Atmospheric Sounding* (World Scientific, 2000), vol. 2.
65. V. Yadav, A. M. Michalak, Improving computational efficiency in large linear inverse problems: An example from carbon dioxide flux estimation. *Geosci. Model Dev.* **6**, 583–590 (2013).
66. K. P. Messier, S. E. Chambliss, S. Gani, R. Alvarez, M. Brauer, J. J. Choi, S. P. Hamburg, J. Kerckhoffs, B. Lafranchi, M. M. Lunden, J. D. Marshall, C. J. Portier, A. Roy, A. A. Szpiro, R. C. H. Vermeulen, J. S. Apte, Mapping air pollution with Google Street View cars: Efficient approaches with mobile monitoring and land use regression. *Environ. Sci. Technol.* **52**, 12563–12572 (2018).
67. South Coast Air Quality Management District, Rule 2305 – Warehouse Actions and Investments to Reduce Emissions (WAIRE) program (2021); https://aqmd.gov/home/rules-compliance/compliance/waire-program.
68. S. Manson, J. Schroeder, D. Van Riper, T. Kugler, S. Ruggles, IPUMS National Historical Geographic Information System: Version 16.0. IPUMS [Preprint] (2021); http://doi.org/10.18128/D050.V16.0.
69. S. E. Chambliss, C. P. R. Pinon, K. P. Messier, B. LaFranchi, C. R. Upperman, M. M. Lunden, A. L. Robinson, J. D. Marshall, J. S. Apte, Local- and regional-scale racial and ethnic disparities in air pollution determined by long-term mobile monitoring. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2109249118 (2021).
70. N. Dadheech, T.-L. He, A. J. Turner, High-resolution greenhouse gas flux inversions using a machine learning surrogate model for atmospheric transport. *Atmos. Chem. Phys.* **25**, 5159–5174 (2025).
71. N. Balasus, D. J. Jacob, A. A. Bloom, J. D. East, L. A. Estrada, S. E. Hancock, M. He, T. A. Mooring, A. J. Turner, J. R. Worden, 2019–2024 trends in African livestock and wetland emissions as contributors to the global methane rise. *Atmos. Chem. Phys.* **26**, 4601–4617 (2026).

Acknowledgments: We are grateful for the contributions of M. Gordon, B. Beveridge, A. Franco, K. Tuxen-Bettman, T. Kirchstetter, C. Preble, N. Asimow, L. H. Koolik, Y. Wang, and S. Hamilton. **Funding:** C.M., R.C.C., R.A.A., T.M.T., M.H.H., and J.S.A. were supported by S. Ostby and S. Cook, Valhalla Foundation. **Author contributions:** Writing—original draft: C.M. and J.S.A. Conceptualization: C.M., R.A.A., J.D.M., R.A.H., and J.S.A. Investigation: C.M. and J.S.A. Writing—review and editing: All authors. Methodology: C.M., R.C.C., T.M.T., J.D.M., R.A.H., and J.S.A. Resources: R.C.C. and J.S.A. Funding acquisition: R.C.C., J.S.A., and R.A.A. Data curation: C.M. and J.S.A. Validation: C.M., R.C.C., R.A.A., and J.S.A. Supervision: R.C.C. and J.S.A. Formal analysis: C.M., A.J.T., and J.S.A. Software: C.M., R.C.C., and J.S.A. Project administration: R.C.C. and J.S.A. Visualization: C.M. and J.S.A. **Competing interests:** C.M., R.A.H., A.J.T., J.D.M., and J.S.A. are inventors on a patent application (US Provisional Patent Application No. 63/864,040, filed 14 August 2025, and PCT International Application No. PCT/US2026/041149, filed 13 August 2026) submitted by UC Berkeley that covers the hyperlocal sensing and inverse modeling methods described in this study. M.H.H. serves as a member of the Community Health and Environmental Research Initiatives Advisory Council for Health Effects Institute. All other authors declare that they have no competing interests. **Data, code, and materials availability:** All data and code needed to evaluate and reproduce the results in the paper are present in the paper and/or the Supplementary Materials and are available in the Dryad repository at https://doi.org/10.5061/dryad.0gb5mkmfx. This study did not generate new materials.

Submitted 26 August 2025
Accepted 28 September 2026
Published 28 October 2026
10.1126/sciadv.aeb7647

Downloaded from https://www.science.org on October 29, 2026

ScienceAdvances

Hyperlocal sensing and inverse modeling reveal community impacts of urban air pollutant emissions

Chirag Manchanda, Ronald C. Cohen, Ramon A. Alvarez, Tammy M. Thompson, Maria H. Harris, Alexander J. Turner, Julian D. Marshall, Robert A. Harley, and Joshua S. Apte

Sci. Adv. **12** (44), eaeb7647. DOI: 10.1126/sciadv.aeb7647

View the article online
https://www.science.org/doi/10.1126/sciadv.aeb7647
Permissions
https://www.science.org/help/reprints-and-permissions

Downloaded from https://www.science.org on October 29, 2026

Use of this article is subject to the Terms of service

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.

Copyright © 2026 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).