Policy Research Working Paper 10297

Scalable Tracking of CO_2 Emissions

A Global Analysis with Satellite Data

Susmita Dasgupta
Somik Lall
David Wheeler

WORLD BANK GROUP
Development Economics
Development Research Group
February 2023

Policy Research Working Paper 10297

Abstract

This paper extends recent research on satellite-based carbon dioxide measurement to an easily updated template for tracking changes in carbon dioxide concentrations at local and regional scales. Using data from the National Aeronautics and Space Administration's Orbiting Carbon Observatory-2 satellite platform and a large sample of urban areas, a comparison of trend estimation models suggests that the template can use a simple model that estimates trends directly from satellite data pre-filtered to isolate local concentration anomalies. Illustrative applications are developed for a long-period trend model and a short-period model focused on change in the most recent year. In addition, the paper estimates carbon dioxide emissions for thousands of urban areas and identifies cities whose emissions performance is above or below expectation. Although the tracking model is "simple," it requires software and hardware that are beyond the means of many interested stakeholders. For this reason, the World Bank's Development Economics Vice Presidency has established an open web facility that pre-filters data from the National Aeronautics and Space Administration's Orbiting Carbon Observatory-2satellite and publishes monthly mean concentration anomalies for all terrestrial cells of a 25-kilometer global grid. The website will also publish annual carbon dioxide tracking reports for urban areas and provide information that links the 25-kilometer global grid cell IDs to IDs for urban areas and national administrative units (levels 0, 1, and 2).

This paper is a product of the Development Research Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at sdasgupta@worldbank.org.

Scalable Tracking of CO_2 Emissions: A Global Analysis with Satellite Data

Susmita Dasgupta*
Somik Lall
David Wheeler

World Bank

Keywords: CO_2 emissions, OCO-2, Urban pollution, Emissions tracking

JEL Classification: Q53, Q54, Q58, R11, R40, R48, R58

* Authors' names in alphabetical order.

1. Introduction

The World Meteorological Organization forecasts that the current greenhouse gas (GHG) emissions trend will increase global temperature 3-5 degrees C by 2100 (Reuters 2018). This would far overshoot the 2-degree limit pledged by the 2015 Paris Climate Accords (COP 21) and might have a catastrophic impact (Steffen et al. 2018; World Bank 2012). In response, several industrial nations pledged very steep emissions reductions at the recent Leaders' Summit on Climate (April 22-23, 2021).

Unfortunately, these pledges confront a striking information shortfall at the outset: near-total absence of directly-measured local and regional GHG data for problem diagnosis, program design and performance assessment. Recently, the advent of satellite-based GHG measurement has greatly expanded the potential for empirical assessment. High-resolution observations of atmospheric GHG concentrations are now available from several platforms, including NASA's OCO-2 and OCO-3 instruments, the European Space Agency's METOP-A and TROPOMI (Sentinel-5P) platforms, China's TANSAT and the Japan Space Exploration Agency's GOSAT and GOSAT-2. Detailed technical assessments of measures from these platforms have verified that they provide useful and comprehensive information for global carbon emissions analysis (Weir et al. 2021; Nassar et al. 2021; Pan et al. 2021; Wu et al. 2020; Hakkarainen et al.; Labzovskii et al. 2019).

This paper extends recent research on satellite-based measurement to an easily-updated template for tracking changes in atmospheric CO_2 concentrations at local and regional scales. Using observations from NASA's OCO-2 platform, we develop the template from the data filtering techniques and econometric analysis employed by Dasgupta, Lall and Wheeler (2022). Our prior work estimates an econometric model that relates satellite-based CO_2 measures to georeferenced emissions sources. In this paper, we develop and compare two versions of the CO_2 tracking template. The first tracks changes in residuals after fitting the econometric model to satellite-based observations, while the second version simply tracks changes in the observations themselves. Using data from thousands of urban areas, we find an extremely close correspondence between results for the two versions. We opt for simplicity and select the second version for template development. We also introduce a technique for identifying city-level changes that are distinct from regional changes induced by broader atmospheric circulation patterns. We provide several illustrative applications for urban areas, while noting that the same approach could be used for any other areas of interest.

In an additional exercise, we use our regression model results to compute expected emissions from urban areas. The regression residuals identify the directions and relative magnitudes of departures from expected values for individual areas. We convert the residuals to their emissions equivalents using high-resolution gridded information from the EDGAR global database (Crippa et al. 2020). For 1,306 urban areas with populations greater than 500,000, we find a rough balance between cities whose emissions are higher and lower than their expected values. Converting deviations to percentages of expected values, we find that percent deviations are typically greater in absolute value for cities with lower-than-expected emissions.

The remainder of the paper is organized as follows. Section 2 motivates the comparison of tracking models by reviewing the econometric model and supporting data from Dasgupta, Lall and Wheeler (2022). Section 3 develops the candidate tracking models and Section 4 compares their results for a large sample of urban areas. Section 5 applies the selected model to illustrative urban cases, while

Section 6 computes expected emissions values and deviations from those values for a large number of urban areas. Section 7 summarizes and concludes the paper. The Appendix describes the World Bank's development of an online platform to support future work.

2. The CO_2 Emissions Model

2.1 Model Specification

An extensive body of empirical literature has explored the determinants of growth in CO_2 emissions. Attention has focused primarily on the drivers of CO_2 emissions from fossil-fuel combustion and cement production (e.g., Raupach et al. 2007; Jotzo et al. 2012). Recent research has also estimated more precisely the CO_2 emissions from fires associated with agriculture and land-use change (Gasser et al. 2020; Winkler et al. 2021). However, emissions drivers in this sector have received less attention than work on industrial determinants (Sanchez and Stern 2016). For the industrial sectors, most of the available estimates are inferred from survey-based activity measures that may be incomplete, particularly for developing countries.

This study takes a completely different approach, employing direct CO_2 observations from satellites. The dependent variable in our model is the atmospheric CO_2 concentration above a cell in a global grid with 25 km resolution. We employ 25-km grid cells defined by Van der Werf et al. (2017), who provide the data on CO_2 emissions from fires for this analysis. The dominant component of the atmospheric concentration is the global stock of CO_2 molecules that have accumulated since the Industrial Revolution. The second component is seasonal CO_2, reflecting differential absorption and release by vegetation over the annual cycle. The seasonal CO_2 component is latitudinal, differing by hemisphere because the Northern Hemisphere has more plant life than the Southern Hemisphere.

The third component is local CO_2 emissions, reflecting the time lag between local emissions of CO_2 molecules and their full dispersion into the global mix. We follow convention by terming this the "concentration anomaly" since it measures the local deviation from the global background CO_2 concentration. In our global emissions model, we classify the determinants of local concentration anomalies in three categories. The first comprises activities in the most significant CO_2-emitting industry sectors. The Intergovernmental Panel on Climate Change (IPCC) (Gale et al. 2005) has identified four dominant industrial sources of CO_2 emissions: (i) power plants, (ii) steel mills, (iii) cement plants, and (iv) oil refineries. The second category includes CO_2 emissions from agricultural and forest fires. The third category comprises population-related emissions other than those directly associated with CO_2-intensive industrial activity. These include motor vehicle emissions, which are not measured reliably at the spatial resolution required for our analysis.[1] We would expect traffic emissions per capita to increase with income per capita, all else being equal. Population-related factors also include CO_2 emissions from household heating and cooling, which are not captured by data for central power plants.[2] The subway component of the model affects population-related emissions by

[1] Recent research has used Google Traffic to infer vehicular emissions from high-resolution traffic congestion data for some cities (Heger et al. 2018; Dasgupta, Lall, and Wheeler 2021). However, no currently available technology enables direct estimation of global vehicular emissions at 25 km resolution.

[2] Household air conditioning is powered by fossil-fired home generators in many hot low-income areas where utility-scale power is either nonexistent or unreliable. For a detailed assessment, see Lam et al. (2019).

reducing the demand for motor vehicle transport and promoting denser settlements that are more easily served by utility-scale energy sources.

We specify the econometric model as follows:

$$(1)\ CO2_{it} = \beta_0 + \beta_1 I_{it} + \beta_2 DI_{it} + \beta_3 F_{it} + \beta_4 DF_{it} + \left(\beta_5 P_{it} H_{it} + \beta_6 P_{it} C_{it} + \beta_7 P_{it} Y_{it}\right) e^{\beta_8 S_{it}(L_{it}, A_{it})} + \varepsilon_{it}.$$

Expected signs: $\beta_1, \beta_2, \ldots \beta_7 > 0$, $\beta_8 < 0$

For grid cell i in period t, CO_{2it} is the satellite-measured mean CO_2 concentration anomaly. I_{it} stands for CO_2 emissions from industrial sources; DI_{it} represents wind-displaced industrial CO_2 emissions from other cells; F_{it} equals CO_2 emissions from agricultural and forest fires; and DF_{it} is wind-displaced fire CO_2 emissions from other cells. P_{it} stands for population and Y_{it} represents income per capita. H_{it} and C_{it} represent heating and cooling degree days, respectively. S_{it} is the subway impact index, which is a function of system scale (L) and age (A); ε is a random error term.

In this equation, the atmospheric CO_2 anomaly is related to emissions from industrial sources, fires, and non-industrial population sources. Spatially-referenced variables in the model are translated to consistent measures by resampling to centroids for our 25-km grid cells. The core model is additive because emissions from the three sources contribute separately to the accumulation of CO_2 molecules in the atmosphere. The anomaly recorded for a grid cell by a satellite platform includes emissions from sources within the cell and the "spillover" emissions created by wind displacement from sources in neighboring cells. For industry and fires, the model includes both cell-specific emissions (I_{it}, F_{it}) and wind-displaced emissions (DI_{it}, DF_{it}).

In the population-related component of the model, the marginal impact of population (P_{it}) is a function of heating degree days (H_{it}), cooling degree days (C_{it}), and income per capita (Y_{it}). For a subway city, their composite effect is conditioned by an exponential function of the scale (L_{it}) and age (A_{it}) of the subway system. The exponential constrains the multiplier to a range from 1 (no subway: $L_{it} = 0$, $A_{it} = 0$) to 0. The multiplier value should decline with both age and scale, the latter measured by the length of operating subway lines.

2.2 Data

Data from several satellite platforms that provide CO_2 measures have been collected by various instruments over different periods, with different resolutions and observation repeat cycles and widths of area coverage along orbital paths (Pan, Yuan, and Jieqi 2021). The data are also accessible in varying degrees. Combining observations from multiple sources could present difficulties that are as yet little-explored. For this exercise, prudence has dictated the choice of one platform, NASA's OCO (Orbiting Carbon Observatory)-2, because it offers open access (JPL/NASA 2021); a long panel of consistently measured, daily observations (beginning on September 6, 2014); and the highest spatial resolution among the available sources (1.29 × 2.25 km).

The design of OCO-2 supports comparative exercises like our analysis. It follows a sun-synchronous near-polar orbit, crossing the equator in ascending mode around 1330 hours local time. This means that the OCO-2 observations for our study are collected between 1200 and 1500 local time for all

cities in the sample, providing a consistent mid-day activity benchmark for comparing CO_2 concentration anomalies.[3] OCO-2 has an observation repeat time of 16 days. We have downloaded georeferenced measures of XCO_2 (the column-averaged dry air mole fraction of CO_2).

We filter the XCO_2 data for local concentration anomalies, or differences between observed and background CO_2 at each point. We calculate background CO_2 using the methodology of Hakkarainen et al. (2019), which incorporates both temporal and geographic elements. As Hakkarainen notes, the available data are insufficient for estimating daily medians at resolutions higher than 10 degrees of latitude. We compute the daily median XCO_2 for each 10-degree latitude band and linearly interpolate the result to each OCO-2 observation with 1-degree resolution. Following Hakkarainen, we use the median as the representative value because it is not skewed by extreme observations. We subtract this background value to compute the local anomaly for each observation. Then we compute monthly mean values of concentration anomalies for the 25-km grid cells in our database.

We use georeferenced facility-level global databases to obtain capacity measures and technology specifications for power plants (Byers et al. 2021), steel mills (GEM 2021), cement plants (McCaffrey et al. 2021), and oil refineries (Auch 2017). We convert capacity measures to annual CO_2 emissions using standard emissions factors for power production by fuel source (USEIA 2021), steel mills (World Steel Association 2021), cement (IEA 2020), and refineries (Jing et al. 2020). Van der Werf et al. (2017) provide monthly estimates of carbon emissions from agricultural and forest burning at 25 km resolution.

Mistry (2019) has provided global estimates of monthly heating and cooling degree days at 25 km resolution.[4] We compute population at 25 km resolution by aggregating data from CIESIN (2021) at 5 km resolution. Monthly estimates are interpolated from data provided for 2010, 2015, and 2020. We use two sources to construct our georeferenced measure of income per capita. From the G-Econ database (Nordhaus et al. 2006), we obtain GDP per capita in 2005 purchasing power parity for a global grid with 100 km resolution. Each grid cell is assigned to its geographically dominant country by G-Econ. For each cell in a country, we compute the ratio of cell GDP per capita to the national mean for all cells. We merge the results with annual UN estimates of GDP per capita in constant $US 2015 (UN 2021), and use the cell ratios to estimate annual GDP per capita for each cell. We resample these cells to 25 km for compatibility with the rest of our database.

We draw our subway data from two sources. The first is a global subway survey by Turner and Gonzalez-Navarro (2018), which includes 137 systems installed prior to 2011.[5] The survey includes digital subway maps at five-year intervals from 1930 to 2010. The second source is our own survey of 55 subway systems installed since 2010. We have constructed digital maps for these systems using

[3] CO_2 measurement during the full daily activity cycle will improve as systems like OCO-3 observe each area at more widely varying times.

[4] Mistry's data terminate in December 2019. We extend the domain for regression analysis by computing monthly means for each 25 km cell using the data for 2014–19.

[5] Gendron-Carrier et al. (2020) provide the following definition: "These data define a 'subway' as an electric powered urban rail system isolated from interactions with automobile traffic and pedestrians. This excludes most streetcars because they interact with vehicle and pedestrian traffic at stoplights and crossings, but underground streetcar segments are counted as subways. The data do not distinguish between surface, underground, or aboveground subway lines as long as the exclusive right of way condition is satisfied. To focus on intraurban subway transportation systems, the data exclude heavy rail commuter lines (which tend not to be electric powered). For the most part, these data describe public transit systems that would ordinarily be described as 'subways', e.g., the Paris metro and the New York city subway, and only such systems."

information from OpenStreetMaps (OSM 2021). We overlay the digital subway maps on our 25 km grid and compute the total length of subway lines in each grid cell and year. Both subway age and line length measures are highly right-skewed, so we apply the inverse hyperbolic sine transformation prior to estimation.[6]

2.3 Accounting for Wind Displacement of CO_2 Emissions

Space-based observations detect higher CO_2 concentrations over emissions sources because atmospheric diffusion is not instantaneous. As the prevailing winds displace emissions from their sources, deviations from background concentrations persist for some time. City-level or plant-level estimates have commonly employed measures of wind direction to model these effects (Nassar et al. 2017; Wu et al. 2020; Ye et al. 2020). We replicate this exercise at global scale, using ERA5 monthly wind direction data for all grid cells in the database (Hersbach et al. 2019). For emissions from each grid cell, we determine the wind-directed path across neighboring cells. We compute monthly wind bearings at 0.25° resolution from 10-m u and v components and then resample to our 25 km grid.[7] Wind paths are calculated in sequence. For each origin cell (A) in the sequence, the destination cell (B) is determined by the wind bearing in cell A. Using each grid cell as a source, we determine the sequential path across nearby cells through 20 iterations.

Theory provides no guidance on local atmospheric persistence as wind displacement proceeds, so we address the issue empirically. In preliminary regression experiments, we perform a grid search across two variables. The first is the duration decay function, modeled as the inverse of the iteration sequence number raised to a power that varies in increments of 0.1 between 0 and 2.0. The second is the number of iterations, which varies from 1 to 20. Our grid search yields best fits for decay and iteration parameters of 1.0 and 10, respectively.

Using these parameters, we incorporate wind displacement effects as follows. For each year and month, we use our industrial and fire emissions data to compute total CO_2 emissions separately for industry and fires in each grid cell. We route these emissions across nearby cells through 10 sequential iterations, identifying the destination cells by iteration. Once this process is complete for all cells in the grid, we proceed cell-by-cell. For each cell, we add across observations for displaced CO_2 from every other cell, with separate totals by iteration step. We weight these totals for industry and fire CO_2 by the inverse iteration step number (which incorporates the decay function). Next, we add across the weighted totals to obtain the overall decay-weighted totals for wind-displaced CO_2 in each cell. These are the variables *DI* and *DF* in the econometric model (equation [1]).

[6] We use the IHST rather than the logarithmic transformation because most of the observations in the data set are zeros (see Burbidge, Magee and Robb (1988) and Layton (2001)).

[7] The bearing calculation formula can be viewed at https://www.movable-type.co.uk/scripts/latlong.html.

2.4 Model Results

Our econometric results are presented in Table 1. The table includes results for alternative estimators that incorporate different assumptions about the structure of the stochastic error term (ε_{it}) in the model. These techniques produce the same point estimates for model parameters, but their differing estimates of standard errors (and the accompanying t-statistics) may lead to very different inferences about the statistical significance of model variables. We replicate the point estimates in columns (1) – (3) to aid interpretation of the t-statistics. We include results for standard nonlinear (NL) regression, NL with robust standard errors (SE) and NL with SE adjusted for 3,074 clusters defined by level-1 administrative units (states, provinces, etc.) for the 190 countries in the regression database.

As the table shows, all results have the expected signs. The mean anomaly for satellite-observed CO_2 concentrations in a 25-km grid cell is positively related to direct emissions from industry and fires; wind-displaced emissions from industry and fires in neighboring areas; and population-related emissions. The marginal impact of population is positively related to heating needs (heating degree days), cooling needs (cooling degree days) and income per capita, and declines with the interaction of subway length and scale. All variables meet classical significance tests in all three regressions, with the exception of cooling degree days in the cluster-adjusted regression.

3. Alternative Templates for CO_2 Emissions Tracking

As noted in the introduction, recent research indicates that satellite-based observations can support an objective, spatially-referenced system for tracking CO_2 trends in local areas and regions. Pre-filtering by the method of Hakkarainen et al. (2019) or similar methods enables measurement of CO_2 concentration anomalies -- the locally determined components of atmospheric concentrations. Temporal considerations are also important because satellite-based measurements, like observations from ground-based monitors, include random components that hinder short-period trend identification. At the present state of the art, trend estimates for periods shorter than a year seem problematic.

Table 1: Determinants of CO_2 concentration anomalies

Dependent Variable: XCO_2 Anomaly (parts per billion)

	NL	NL (Robust)	NL(Cluster)[a]
Industry CO_2 Emissions ['000 Tons]	0.277*** (26.19)	0.277*** (18.85)	0.277*** (9.88)
Industry CO_2 Wind-Displaced Emissions ['000 Tons (Weighted)]	0.275*** (44.59)	0.275*** (31.60)	0.275*** (13.82)
Fires CO_2 Emissions ['000 Tons]	0.359*** (24.79)	0.359* (2.35)	0.359* (2.03)
Fires CO_2 Wind-Displaced Emissions ['000 Tons (Weighted)]	0.573*** (61.85)	0.573*** (8.64)	0.573*** (4.46)
Population ['000] x Heating Degree Days	5.362*** (69.53)	5.362*** (17.25)	5.362*** (6.18)
Population ['000] x Cooling Degree Days	0.398*** (5.48)	0.398*** (5.29)	0.398 (1.32)
Population ['000] x Income Per Capita [$US '000]	13.54*** (24.45)	13.54*** (15.94)	13.54*** (3.75)
IHST[b] [Subway Scale] + IHST[b] [Subway Age] [Scale: Track Length in km]	-0.193*** (-37.48)	-0.193*** (-22.29)	-0.193*** (-10.51)
Constant	-193.4*** (-192.01)	-193.4*** (-153.66)	-193.4*** (-10.32)
Observations	1,961,754	1,961,754	1,961,754

[a] GADM (2021) Level 1 Administrative Divisions (States, Provinces)
[b] IHST: Inverse hyperbolic sine transformation

t statistics in parentheses
* p<0.05 ** p<0.01 *** p<0.001

At temporal resolutions of one year or greater, it may be useful to augment the Hakkarainen technique with a second filter based on an econometric model like (1) above. The model relates local CO_2 concentration anomalies to local emissions from industry, fires and non-industrial population sources. Model parameters provide estimates of characteristic marginal relationships between local emissions and satellite-recorded concentration anomalies, while model residuals measure the deviations of local anomalies from their expected values. In principle, model estimation could be viewed as a useful post-Hakkarainen filter, with trends in model residuals used for tracking changes in local CO_2 emissions intensities. However, the present case is complicated by the presence of both static and dynamic variables in the model. The industry components are basically fixed effects because they are derived from fixed plant capacities, not variable outputs. Among the non-industrial components, population, income per capita and the subway variables are interpolated from observations over intervals of a year or longer. The only variables with one-month periodicity are fires, heating degree days and cooling degree days. Under these conditions, it is not clear whether the econometric post-Hakkarainen filter adds significant value to an emissions tracking analysis.

We test the filtering utility of econometric model estimation with data for 6,142 Functional Urban Areas (FUAs) with populations greater than 100,000, as defined by Schiavina et al. (2019). Our exercise controls for potential biases introduced by limited sampling within FUAs. Although the OCO-2 satellite platform provides the best available database, its coverage for our 25-km grid cells is limited by its 16-day repeat cycle, relatively narrow observation track, and the frequent occurrence of cloud cover over some areas. Within an FUA, typical concentration anomalies may differ substantially across grid cells. To cite one possible consequence, a naive trend analysis could generate spuriously-positive results in cases where early-period observations are more numerous in lower-anomaly cells and later observations are more concentrated in higher-anomaly cells. To test for this potential source of bias, our exercise includes regressions with dummy variable controls for grid cells.

For each FUA (j), we estimate the following tracking models:

CO2 Models

(2) $CO2_{itj} = \gamma_{0j} + \gamma_{1j}t + \varepsilon_{itj}$

(3) $CO2_{itj} = \gamma_{0j} + \sum_{i=1}^{N_j} \rho_i D_i + \gamma_{1j}t + \varepsilon_{itj}$

Residuals Models

(4) $[CO2_{it} - \widehat{CO2}_{it}]_j = \beta_{0j} + \beta_{1j}t + \varepsilon_{itj}$

(5) $[CO2_{it} - \widehat{CO2}_{it}]_j = \beta_{0j} + \sum_{i=1}^{N_j} \theta_i D_i + \beta_{1j}t + \varepsilon_{itj}$

where, for grid cell i in month t:

$CO2_{it}$	=	Mean CO_2 anomaly (after pre-filtering by the method of Hakkarainen et al. (2019))
$\widehat{CO2}_{it}$	=	Prediction from model (1) above
D_i	=	Dummy variable for grid cell i[8]
t	=	Time from initial period in months
ε_{it}	=	Random error term

For an FUA, changes in the CO_2 concentration anomaly over the sample period are judged from the sign, size and statistical significance of $\widehat{\gamma_{1j}}$ and $\widehat{\beta_{1j}}$.

4. Results for Tracking Models

For model evaluation, we select the 507 urban areas with populations greater than 100,000 that have sufficient observations to yield 60 degrees of freedom after accounting for the number of dummy variable controls in equations (3) and (5).[9] All models are estimated for the period September 2014 to October 2021. We are particularly interested in testing the efficacy of (2), the simplest possible tracking model, which is estimated directly from the pre-filtered data with no grid cell dummy variables or econometric-model-based controls for local emissions sources. Our tests are performed for the change parameters $\widehat{\gamma_{1j}}$ and $\widehat{\beta_{1j}}$.

Table 2 displays correlation coefficients between change parameters, ordered by structural "distance" from model (2). We focus on column (2), which tabulates correlation coefficients for model (2). They are all very high, declining slightly from 0.99 to 0.97 for (5), the most structurally-distant model, which includes filtering with econometric residuals and dummy variable controls for grid cells. Figure 1 displays the accompanying point scatter for model (2) vs model (5), while Table 2 presents the associated regression results.

[8] No subscript j is needed because grid cells are uniquely assigned to FUAs.

[9] The familiar classical 95% significance criterion is t=2.00 with 60 degrees of freedom for estimation.

Table 2: Correlation coefficients for model parameters (N=507)

Model	Parameter	(2) γ_1	(3) γ_1	(4) B_1	(5) B_1
(2)	γ_1	1.00			
(3)	γ_1	0.98	1.00		
(4)	B_1	0.98	0.97	1.00	
(5)	B_1	0.97	0.98	0.98	1.00

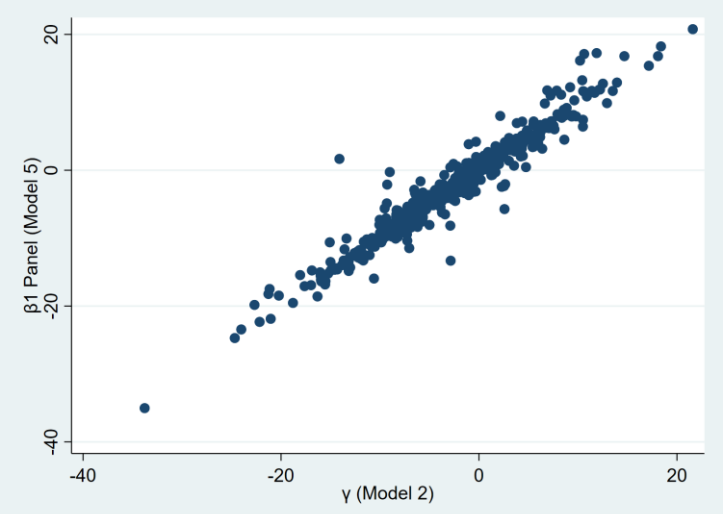
Table 3: Regression results

	B_1 (Model 5)
γ_1 (Model 2)	0.974***
	(88.10)
Const.	0.0906
	(1.05)
R^2	0.94
N	507

t statistics in parentheses

* p<0.05 ** p<0.0 *** p<0.001

Figure 1: B_1 (model 5) vs γ_1 (model 2)



The results in Tables 2 and 3 and Figure 1 strongly suggest that the simplest model (2) is sufficient for tracking trends in local concentration anomalies. This is good news for interested global stakeholders, who can track areas of interest in two simple steps: (1) match the areas to grid cells in the 25 km Hakkarainen-filtered database of concentration anomalies; (2) estimate model (2) for each area.

5. Illustrative Applications

5.1 Long- and Short-Period Tracking Models

We illustrate the methodology with two tracking models for urban areas. The first is (6), reproducing (2) above, which provides trend estimates for an extended period. While these multi-year trends provide useful information, they may lack the immediacy needed to catalyze local action. The second model (7) contributes by estimating the size and significance of changes in the most recent year. Technically, (7) replaces the trend term in (2) with a dummy variable (D_F) for observations in the final year.

(6) $CO2_{itj} = \gamma_{0j} + \gamma_{1j}t + \varepsilon_{itj}$

(7) $CO2_{itj} = \gamma_{0j} + \delta_{1j}D_F + \varepsilon_{itj}$

5.2 Notes on Interpretation

In the following section, we report urban trend results for September 2014 – December 2021 and dummy-variable results for two five-year periods: January 2015 – December 2019 and January 2017 – December 2021. Before presenting the results, we believe that some interpretive notes are warranted. First, we offer a caveat about viewing long- and short-period differences across urban areas as indicators of differential performance, because "performance" implies intentionality on the part of public or private actors. However, even highly-significant changes in local concentration anomalies may reflect non-intentional factors such as changed agricultural practices, blight-induced forest degradation, or additional emissions from traffic congestion during periods when mass transit systems are installed. Measurement anomalies may also intrude, particularly for model (7) because it focuses on relatively short-run changes.[10] To summarize, it may be more useful to view tracking results as guides to detailed local assessments rather than as performance indicators.

A note about measurement units is also warranted. In Dasgupta, Lall and Wheeler (2022), we convert concentration anomalies to emissions estimates using the overall ratio of total global CO_2 emissions (drawn from standard sources) to the total for all grid squares of concentration anomalies predicted from the parameters of our econometric model. This enables us to perform a distribution of predicted CO_2 emissions across all terrestrial cells of the 25 km global grid. However, we have little confidence in our current ability to infer changes in emissions volumes from directly-observed changes in local anomalies that are not linked to identifiable ground sources. For this reason, our trend estimation exercise operates solely with concentration anomalies. This has no practical consequence, since the results provide readily-comparable change estimates.

5.3 Tracking Data

We estimate models (6) and (7) for 1,799 functional urban areas (FUAs) whose data provide at least 30 degrees of freedom for estimation.[11] Table 4 enumerates the outlying observations that are removed prior to estimation. Observations with standardized z-values greater than 5.0 have been identified as outliers.[12] As the table shows, 1,761 of 1,799 FUAs have no outliers. Single outliers have been removed for 35 FUAs, 2 outliers for 2 FUAs and 3 for 1 FUA.

[10] See particularly Weir et al. (2021) for a useful discussion of potential effects from year-to-year variability caused by differences in atmospheric circulation. We address this issue in Section 5.5.

[11] The sample is much larger than the sample used in Section 4 because models (6) and (7) do not absorb degrees of freedom with dummy variable controls for grid cells.

[12] The z-value of an observation is its distance from the mean, measured in standard deviations.

Table 4: Outlier observations removed

Outliers (z>5)	FUAs
0	1,761
1	35
2	2
3	1
Total	1,799

Model estimation yields t-statistics for results classification by degree of significance. We group urban areas into four categories by significance level: $p>.05$ (no significant change at 95% confidence); $p\leq.05$ (95% confidence); $p\leq.01$ (99% confidence) and $p\leq.001$ (99.9% confidence).

Figure 2 illustrates this classification for selected urban areas with negative and positive trends at different levels of significance for the period 2014 – 2021.[13] Each graph includes the regression line (equivalent to model (6)) through the observations. The graphs for the non-significant group have been chosen for FUAs with t-statistics near 1.0. The illustrations highlight the need for statistical analysis, given the role of random variation in satellite-based measurements. Of course, the same is true for ground-based measurements (e.g. Chakraborty et al. 2008).

5.4 Tracking Results: FUA Trends, 2014 - 2021

Table 5 summarizes our results for trend equation (6). Our sample comprises 1,799 urban areas with 30 degrees of freedom or more during the period 2014 - 2021.[14] Of these, 272 have significant decreases in local concentration anomalies and 108 have significant increases. Overall, 380 of 1,799 cities (21.1%) have significant changes during the 8-year period. Table 6 distributes the changes across regions, showing Asia with a disproportionate share of increases and the other regions with disproportionate shares of decreases. Figure 3 maps the same information, showing that within Asia, substantial portions of the decreases and increases are accounted for by India and China, respectively.

[13] Scaling on the y-axis varies across cases because observations vary over different ranges.
[14] FUAs in the database have populations of 50,000 or more.

Figure 2: Trends in monthly concentration anomalies, 2014 – 2021 (ppm)

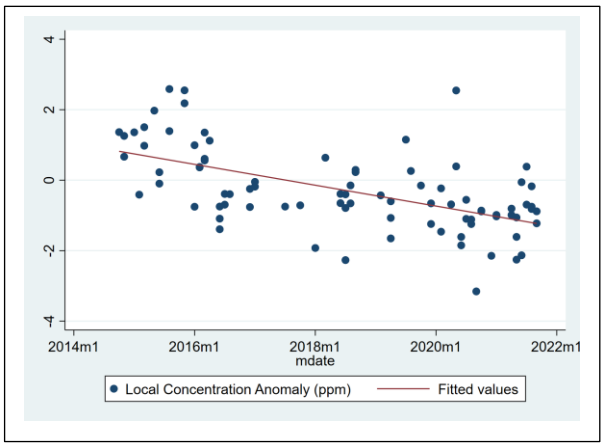
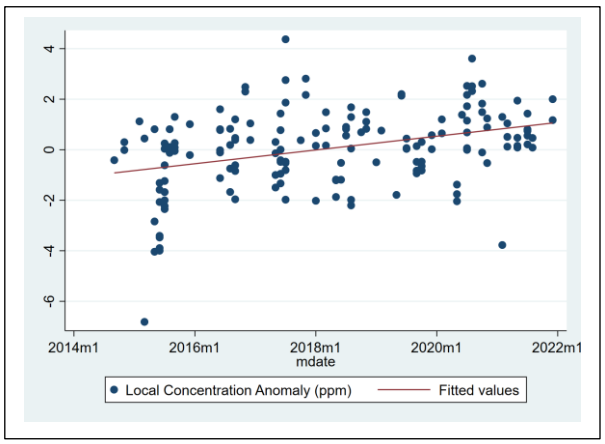
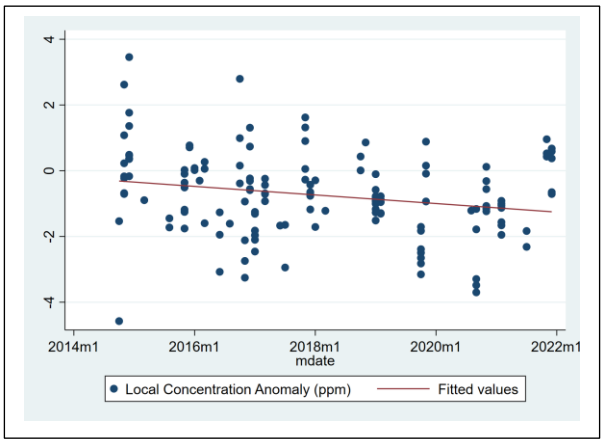
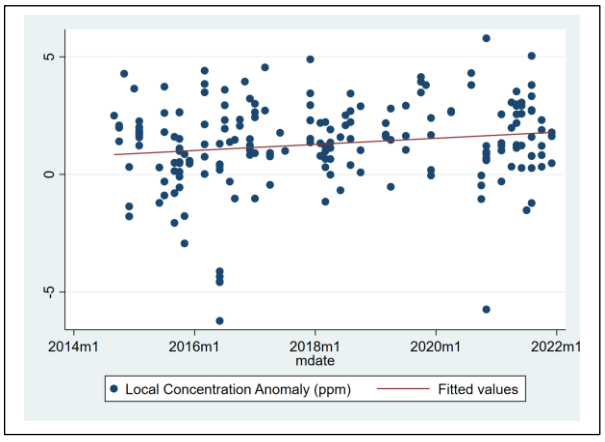
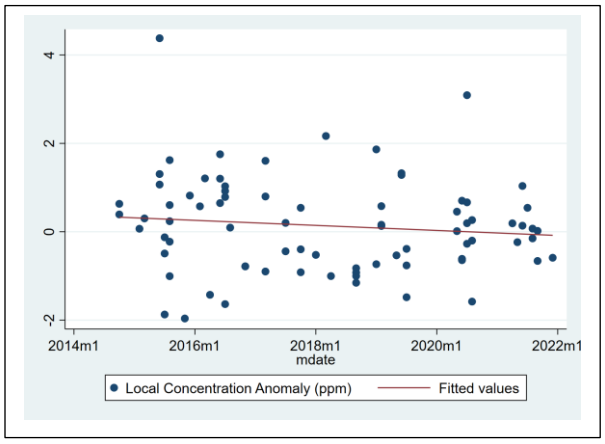
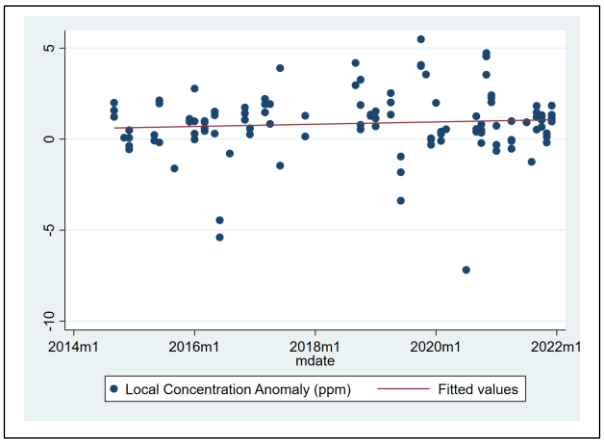
	Negative Trends	Positive Trends
p≤.001	Vila Velha, Brazil	Naples, Italy
		
.01<p≤.05	Dhaka, Bangladesh	Osaka, Japan
		
Not Significant (t≈1.0)	Nice, France	Jinzhou, China
		

Table 5: Time trends for FUAs, 2014 – 2021*

Dependent variable: Local CO_2 Anomaly (ppb)

	Trend		
Significance Level	Decrease	Increase	Total
Not Significant	804	615	1,419
.05	151	72	223
.01	87	27	114
.001	34	9	43
Total Significant	272	108	380
Total	1,076	723	1,799

* FUAs with degrees of freedom ≥ 30

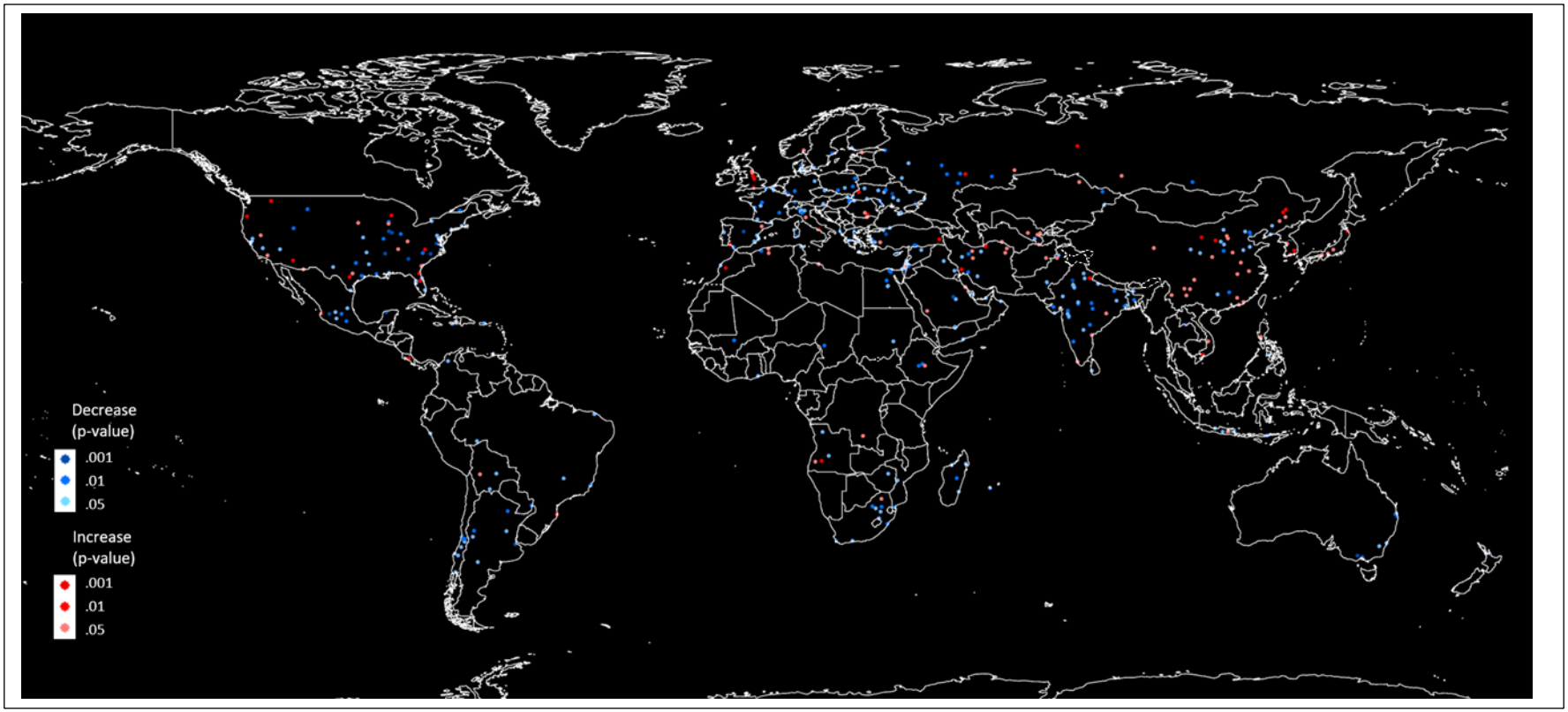
Table 6: Regional distribution of FUA anomaly trends, 2014 – 2021 *

[Count / Column Pct]

Region	Decrease	Increase	Total
Africa	32	9	41
	(11.8)	(8.3)	(10.8)
Americas	72	21	93
	(26.5)	(19.4)	(24.5)
Asia	86	58	144
	(31.6)	(53.7)	(37.9)
Europe	75	20	95
	(27.6)	(18.5)	(25.0)
Oceania	7	0	7
	(2.6)	(0.0)	(1.8)
Total	272	108	380

* Column percents in parentheses

Figure 3: Functional Urban Areas: Trends in local concentration anomalies, 2014 - 2021



5.5 Accounting for Atmospheric Circulation

Although the Hakkarainen filter removes annual and seasonal changes from CO_2 concentrations, it does not account for region-scale changes in atmospheric circulation that can alter measurements by several parts per million, even if local emissions remain constant (Weir et al. 2021). By implication, some of our FUA trend results may reflect regional atmospheric circulation effects rather than changes in local emissions. We test for this effect by creating two groups within 100 km of each FUA centroid: grid squares inside the FUA and those lying outside. If regional circulation plays an important role, then trend results should be very similar for the two groups. Formally, we estimate separate regressions for 100-km-radius grid squares inside (I) and outside (O) of FUA j:

(8) Inside Regression: (2I) $CO2_{ijtI} = \gamma_{0I} + \gamma_{1jI}t + \varepsilon_{itjI}$

(9) Outside Regression: (2O) $CO2_{ijtO} = \gamma_{0O} + \gamma_{1jO}t + \varepsilon_{itjO}$

The relevant null hypothesis holds that atmospheric circulation explains a significant result for γ_{1jI}. The null hypothesis is rejected if:

sign(γ_{1jI}) ≠ sign(γ_{1jO}) or

p(γ_{1jI}) ≤ p(.05); p(γ_{1jO}) > p(.05)

To summarize, regional atmospheric circulation is rejected as the source of a significant change in an FUA concentration anomaly if the outside change parameter has the opposite sign from the inside parameter and/or the inside parameter is significant while the outside parameter is not.

Table 7 provides evidence for 335 FUAs from our estimation sample that have populations greater than 100,000 and inside change parameters with p⩽.05. Among the corresponding outside change parameters, 92 (27.5%) have sign reversal (sign(γ_{1jI}) ≠ sign(γ_{1jO})). Among the 243 outside parameters that do not have sign reversal, 136 (56.0%) are not significant at 95% confidence. In summary, the null hypothesis is rejected for 228 (92 + 136) (68.1%) of the 335 urban areas in the sample.

Table 7: FUA concentration anomaly trends: inside/outside test results ($p(\gamma_{1jI}) \leq p(.05)$; FUAs with populations > 100,000)

Conditions for Rejection of H_0: $sign(\gamma_{1jI}) \neq sign(\gamma_{1jO})$ or $p(\gamma_{1jI}) \leq p(.05)$; $p(\gamma_{1jO}) > p(.05)$

Sign Reversal	Count	Percent
Yes	92	27.5
No	243	72.5
Total	335	
If No Sign Reversal: Significance (95%)		
Yes [$p(\gamma_{1jO}) \leq p(.05)$]	107	44.0
No [$p(\gamma_{1jO}) > p(.05)$]	136	56.0
Total	243	
Atmospheric Circulation Hypothesis		
Accepted	107	31.9
Rejected	228	68.1
Total	335	

Table 8 incorporates this factor, presenting percent changes in CO_2 concentration anomalies for 83 urban areas with populations greater than 100,000, trend results significant at $p \leqslant .01$, and absence of general circulation effects.[15] The table reveals substantial geographic diversity, with 44 countries represented. The United States has 13 entries, followed by India (9), the Russian Federation (5), China (4) and South Africa (4). Table 9 shows that cities with decreasing concentration anomalies outnumber cities with increasing anomalies in all regions. At the same time, African and European cities have disproportionate decreases, American cities have disproportionate increases, and Asian cities are about equally represented. Overall, 71% of the cities have decreasing trends and 29% have increasing trends.

5.6 Short-Period FUA Changes

Model (7) estimates the size and significance of final-year deviations from typical concentrations in previous years. For this illustration, we employ rolling five-year time series for urban areas. The change indicator is δ_{1j}, the parameter for the final-year dummy variable in model (7). Figure 4 provides illustrations for six urban areas during the period 2017-2021.[16] Warsaw and Orlando represent the highest confidence class ($p \leqslant .001$), with Warsaw concentration anomalies in 2021 significantly below the five-year line and those of Orlando significantly above. The results for Hengyang/Mumbai ($p \leqslant .05$) meet standard classical significance standards but are somewhat less robust, because estimated final-year deviations are smaller and/or observational variance is larger than in the most robust cases. Casablanca and Bukhara have decreases and increases, respectively, but neither is significant by classical standards.

Tables 10 and 11 summarize our results for five-year periods ending in 2019 and 2021, respectively.[17] The patterns of statistical significance and negative/positive distributions are similar for the two periods; significant changes are identified for 225 urban areas in 2015-2019 and 235 areas in 2017-2021.

[15] We calculate percent changes after subtracting each city's minimum concentration anomaly from all of its observations. This permits percent calculations by transforming negative anomalies to positive values. For each city, we use its regression result for model (6) to predict concentrations in the first and last years. These predictions are used to calculate percent changes.

[16] Again a cautionary note: y-axis scaling differs substantially across cases.

[17] The results are for urban areas with at least 30 degrees of freedom for estimation.

Table 8: Trends in CO_2 concentration anomalies, 2014 – 2021
(population ≥ 100,000; p ≤ .01; atmospheric circulation effects absent)

City	Country	% Change	City	Country	% Change
Shivamogga	India	-38.2	Matala	Angola	61.0
Trier	Germany	-32.7	Marrakesh	Morocco	53.8
Lublin	Poland	-31.5	Haldwani	India	49.5
Cherkasy	Ukraine	-31.5	Dimitrovgrad	Russian Federation	49.0
Maputo	Mozambique	-30.8	Hachinohe	Japan	46.7
Balurghat	India	-30.6	Winston-Salem	United States	44.8
Vientiane	Lao PDR	-30.5	Eugene	United States	42.0
Bindi	Pakistan	-30.1	Ho Chi Minh City	Vietnam	42.0
Valence	France	-29.9	Huelva	Spain	41.7
Izmir	Türkiye	-29.4	San José	Costa Rica	39.0
Durban	South Africa	-28.7	Sora-myeon	Korea, Rep.	32.5
Tiraspol	Moldova	-28.5	Orlando	United States	32.1
Le Tampon	Reunion	-28.5	Van	Türkiye	28.9
Adama	Ethiopia	-28.4	Bucharest	Romania	26.2
Fayetteville	United States	-28.2	Ahwaz	Iran, Islamic Rep.	26.2
Odense	Denmark	-27.8	Puerto Vallarta	Mexico	22.8
Jabalpur	India	-27.5	Lviv	Ukraine	21.6
Río Piedras [San Juan]	Puerto Rico	-26.5	Portsmouth	United Kingdom	21.4
Ulyanovsk	Russian Federation	-25.6	Florianópolis	Brazil	20.6
Cuernavaca	Mexico	-25.5	Daegu	Korea, Rep.	18.9
Jerez	Spain	-24.8	Hangyulu	China	14.8
Ciudad del Este	Paraguay	-24.5	Appleton	United States	14.5
Irkutsk	Russian Federation	-24.5	Jacksonville	United States	14.4
Ogden	United States	-24.4			
Chenzhou	China	-24.3			
Yicheng	China	-23.7			
Relizane	Algeria	-23.6			
Mariupol	Ukraine	-23.3			
Warangal	India	-23.0			
Balaghat	India	-22.6			
Sterlitamak	Russian Federation	-22.3			
Janesville	United States	-22.2			
Eskisehir	Türkiye	-22.1			
Champaign	United States	-21.3			
Cape Town	South Africa	-21.2			
Middelburg	South Africa	-21.2			
Maumere	Indonesia	-21.1			
Behbahan	Iran, Islamic Rep.	-21.0			
Rustenburg	South Africa	-20.8			
Katowice	Poland	-20.5			
Saint Petersburg	Russian Federation	-20.3			
Iloilo City	Philippines	-19.1			
Athens	Greece	-19.0			
Melbourne	Australia	-18.9			
Guadalajara	Mexico	-18.8			
Nicosia	Cyprus	-18.8			
Rajgarh	India	-17.5			
Sherbrooke	Canada	-16.8			
La Rochelle	France	-16.7			
St. Louis	United States	-16.7			
Lafayette	United States	-16.4			
Buenos Aires	Argentina	-16.1			
Rajkot	India	-15.8			
Muzaffarpur	India	-15.0			

Table 9: Trends by region *

Region	Trend: Decrease	Trend: Increase	Total
Africa	8	2	10
	(13.6)	(8.3)	(12.1)
Americas	14	8	22
	(23.7)	(33.3)	(26.5)
Asia	20	9	29
	(33.9)	(37.5)	(34.9)
Europe	16	5	21
	(27.1)	(20.8)	(25.3)
Oceania	1	0	1
	(1.7)	(0.0)	(1.2)
Total	59	24	83
	(71.1%)	(28.9%)	

* Column percents in parentheses

Table 10: FUAs, 2019 deviations from 5-year means, 2015-2019*

Dependent variable: Local CO_2 Anomaly (ppm)

Significance Level	2019 Deviation: Negative	2019 Deviation: Positive	Total
Not Significant	446	401	847
.05	71	56	127
.01	42	29	71
.001	16	11	27
Total Significant	129	96	225
Total	575	497	1,072

* FUAs with degrees of freedom ≥ 30

Table 11: FUAs, 2021 deviations from 5-year means, 2017-2021*

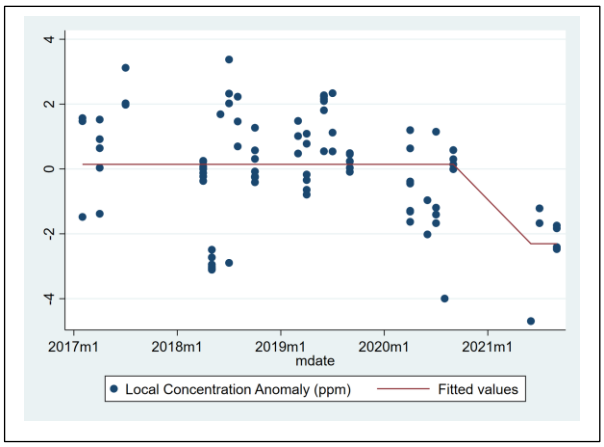
Dependent variable: Local CO_2 Anomaly (ppm)

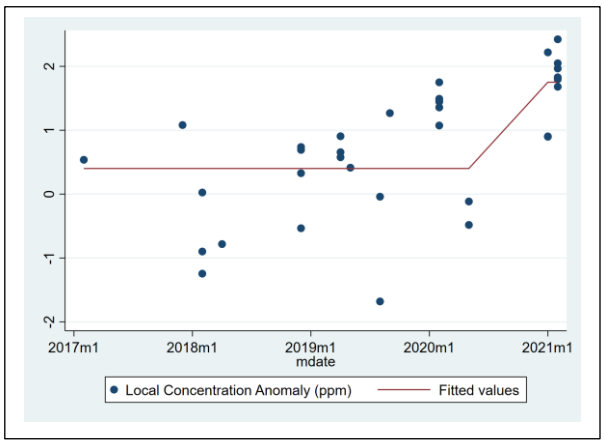
Significance Level	2021 Deviation: Negative	2021 Deviation: Positive	Total
Not Significant	418	431	849
.05	76	70	146
.01	37	25	62
.001	15	12	27
Total Significant	128	107	235
Total	546	538	1,084

* FUAs with degrees of freedom ≥ 30

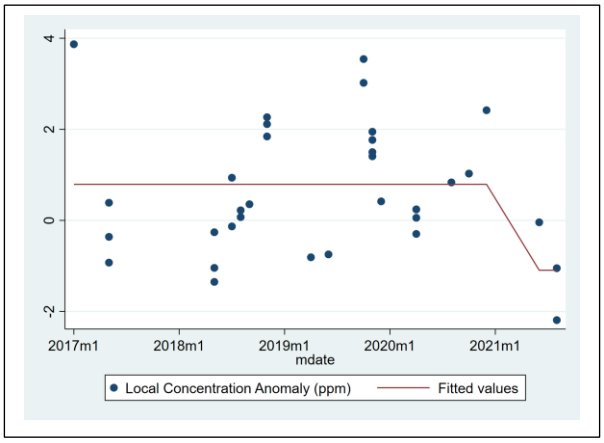
Figure 4: Monthly local concentration anomalies, 2017 - 2021 (ppm)

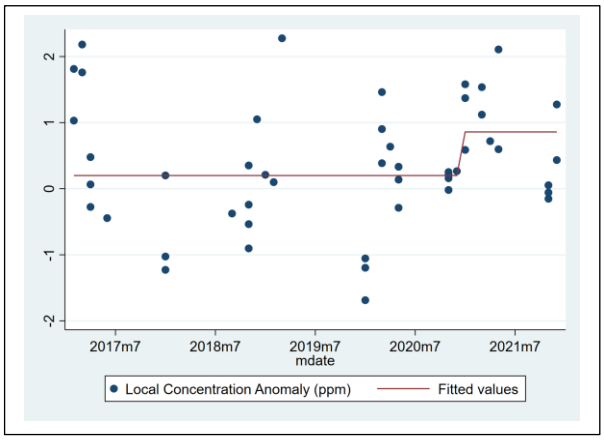
	Negative Deviation	Positive Deviation
p≤.001	Warsaw, Poland	Orlando, United States



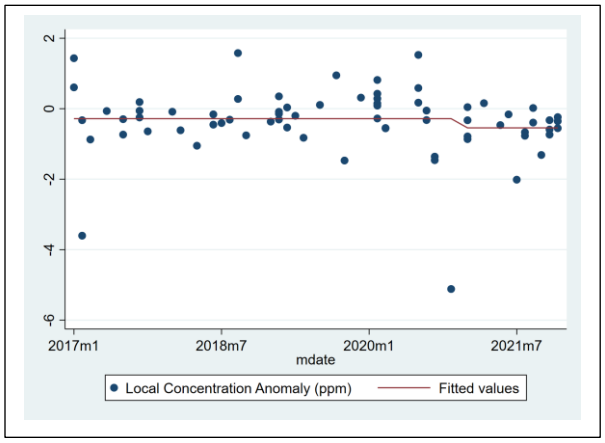


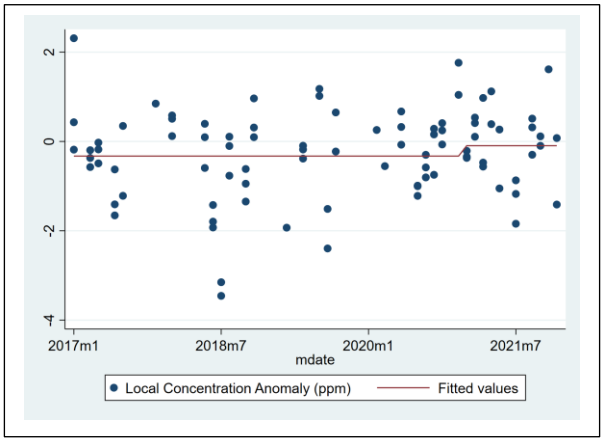
.01<p≤.05	Hengyang, China	Mumbai, India





Not Significant (t≈1.0)	Casablanca, Morocco	Bukhara, Uzbekistan





5.7 Short-Period Changes: Accounting for Atmospheric Circulation

We introduce atmospheric circulation effects using the inside/outside methodology that we have applied to annual trend estimation. Table 12 provides regional summaries for qualifying urban areas with populations greater than 100,000 that have significant short-period changes and absence of atmospheric circulation effects. Regional representation remains roughly stable in the two periods, with the possible exception of Europe.

Tables 13 and 14 display qualifying urban areas for 2015-2019 and 2017-2021, with both decreases and increases displayed in descending order by absolute value. Percent changes are calculated from regression-predicted values for the first and final years. Among the 31 represented countries, 16 appear in only one table and 15 in both. United States urban areas are most numerous in both tables (13 and 12 respectively), followed by China (8 and 7) and France (5 and 3). Figures 5 and 6 map the same cities and identify their change status.

Of the 100 cities represented in Tables 13 and 14, 6 appear in both periods. Table 15 shows that Melbourne, Madrid, La Rochelle and Fayetteville have consistent decreases, while Greensboro and Warsaw switch from increases in 2015-19 to decreases in 2017-21.

Table 12: Regional representation of urban areas without atmospheric circulation effects

	2015-19		2017-21	
Region	Count	Percent	Count	Percent
Africa	2	3.5	2	3.9
Americas	18	31.6	18	35.3
Asia	15	26.3	15	29.4
Europe	21	36.8	15	29.4
Oceania	1	1.8	1	2.0
Total	57	100	51	100

Table 13: 2019 Deviation from 5-Year mean concentration anomaly, 2015 – 2019 (population ≥ 100,000; p ≤ .01; atmospheric circulation effects absent)

City	Country	% Change	City	Country	% Change
Montevideo	Uruguay	-53.2	Las Cruces	United States	63.5
Yekaterinburg	Russian Federation	-43.1	Quebec	Canada	62.4
Lipetsk	Russian Federation	-35.5	Pune	India	48.0
Simferopol	Ukraine	-32.1	Xi'an	China	43.6
Lawrence	United States	-24.9	San Antonio	United States	42.9
Minsk	Belarus	-23.8	Bayannur	China	36.8
Bydgoszcz	Poland	-22.3	Sora-myeon	Korea, Rep.	36.2
Debrecen	Hungary	-21.5	Lorient	France	34.5
Boston	United States	-20.7	Kunming	China	31.7
Acapulco	Mexico	-20.6	Saarbruecken	Germany	31.0
Grand Rapids	United States	-20.0	Rockford	United States	30.5
Fayetteville	United States	-19.2	Bellingham	United States	30.5
Coimbra	Portugal	-17.7	Bojnurd	Iran, Islamic Rep.	28.0
Hohhot	China	-17.5	Sioux City	United States	27.8
Fula'erji	China	-17.4	Greensboro	United States	21.8
Aarhus	Denmark	-17.0	Fenyang	China	21.4
Pindi Gheb	Pakistan	-16.6	Khemis Miliana	Algeria	20.5
Nantes	France	-16.6	Suzhou	China	19.9
La Rochelle	France	-15.3	Quetta	Pakistan	18.5
Barcelona	Spain	-14.9	Pratapgarh	India	18.1
Kryvyi Rih	Ukraine	-14.9	Waterloo	United States	17.7
Madrid	Spain	-14.7	Warsaw	Poland	17.7
Orléans	France	-14.5	Seville	Spain	13.9
Sherbrooke	Canada	-13.8	Birmingham	United States	11.8
Buenos Aires	Argentina	-13.8	Port Elizabeth	South Africa	11.7
Berlin	Germany	-13.1			
Eskisehir	Türkiye	-12.7			
Bordeaux	France	-11.9			
Da'an	China	-11.4			

Table 14: 2021 Deviation from 5-year mean concentration anomaly, 2019 – 2021 (population ≥ 100,000; p ≤ .01; atmospheric circulation effects absent)

City	Country	% Change	City	Country	% Change
Warsaw	Poland	-37.9	Baoshan	China	62.2
Brescia	Italy	-34.9	Kemerovo	Russian Federation	42.5
Le Tampon	Reunion	-30.7	Chiplun	India	36.7
Ogden	United States	-30.7	Besançon	France	35.3
Sevastopol	Ukraine	-30.4	Konch	India	32.5
Jakarta	Indonesia	-29.1	Canton	United States	29.9
Trier	Germany	-28.7	Daegu	Korea, Rep.	29.9
Dajie	China	-28.2	Poyang	China	29.4
Dayton	United States	-25.7	Murcia	Spain	28.6
Sheyang	China	-24.1	Kenitra	Morocco	26.5
Bursa	Türkiye	-23.3	Orlando	United States	24.2
Fayetteville	United States	-22.0	Adiyaman	Türkiye	23.0
Greensboro	United States	-20.5	Ardakan	Iran, Islamic Rep.	22.1
Washington D.C.	United States	-18.8	Hamburg	Germany	21.5
Daytona Beach	United States	-18.5	Anápolis	Brazil	20.6
San Luis Potosí	Mexico	-18.5	Copenhagen	Denmark	18.2
Ciudad Acuña	Mexico	-18.4	Toulon	France	18.2
Havana	Cuba	-17.8	Novorossiysk	Russia Federation	17.3
Karlsruhe	Germany	-17.1	Hangyulu	China	15.9
McAllen	United States	-15.6	Salmas	Iran, Islamic Rep.	15.7
La Rochelle	France	-14.9	Belo Horizonte	Brazil	14.9
Madrid	Spain	-14.7	Jacksonville	United States	14.8
Melbourne	Australia	-14.7	Liaoyuan	China	13.1
Zagreb	Croatia	-14.1	New York	United States	13.0
Seattle	United States	-12.6			

Figure 5: Functional Urban Areas: Local concentration anomalies, 2019 deviations from 5-year means

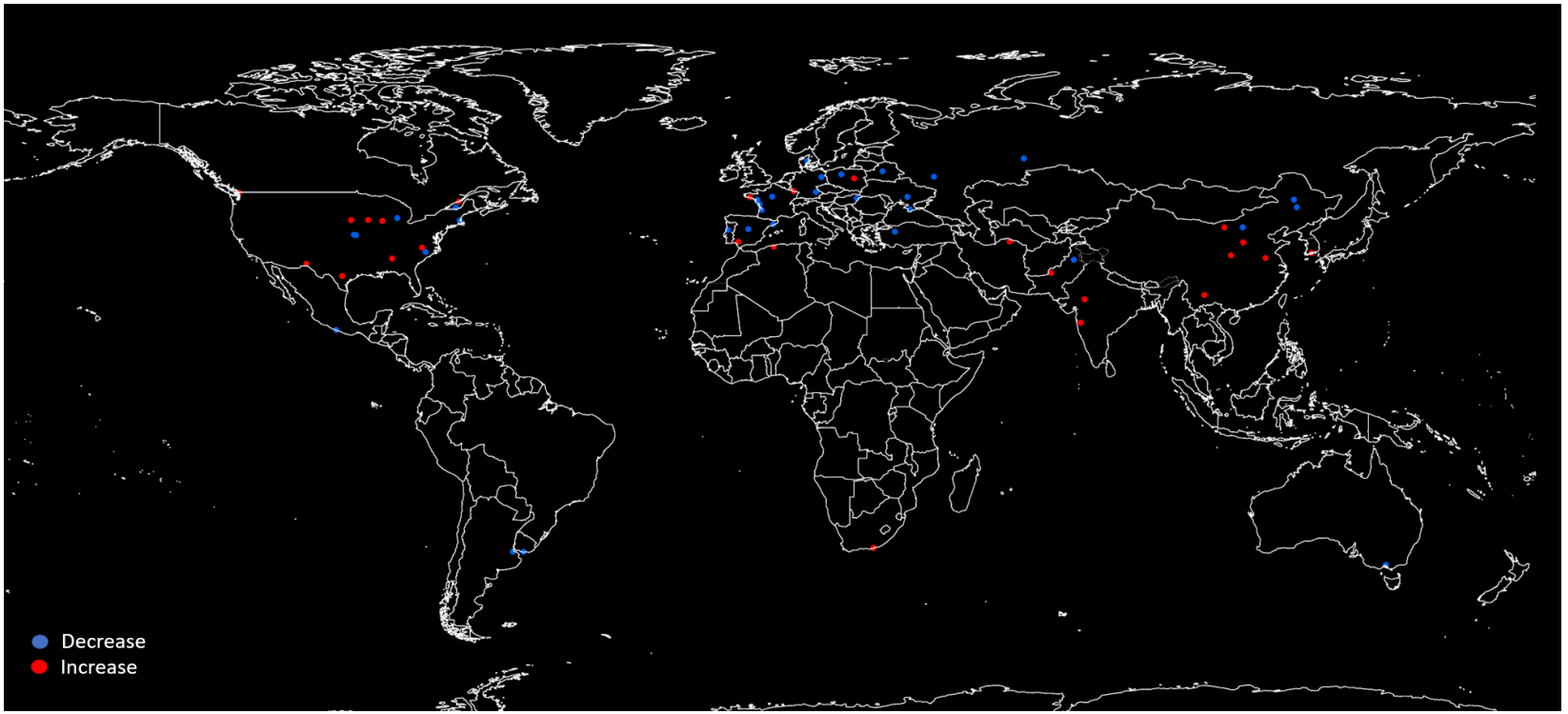


Figure 6: Functional Urban Areas: Local concentration anomalies, 2021 deviations from 5-year means

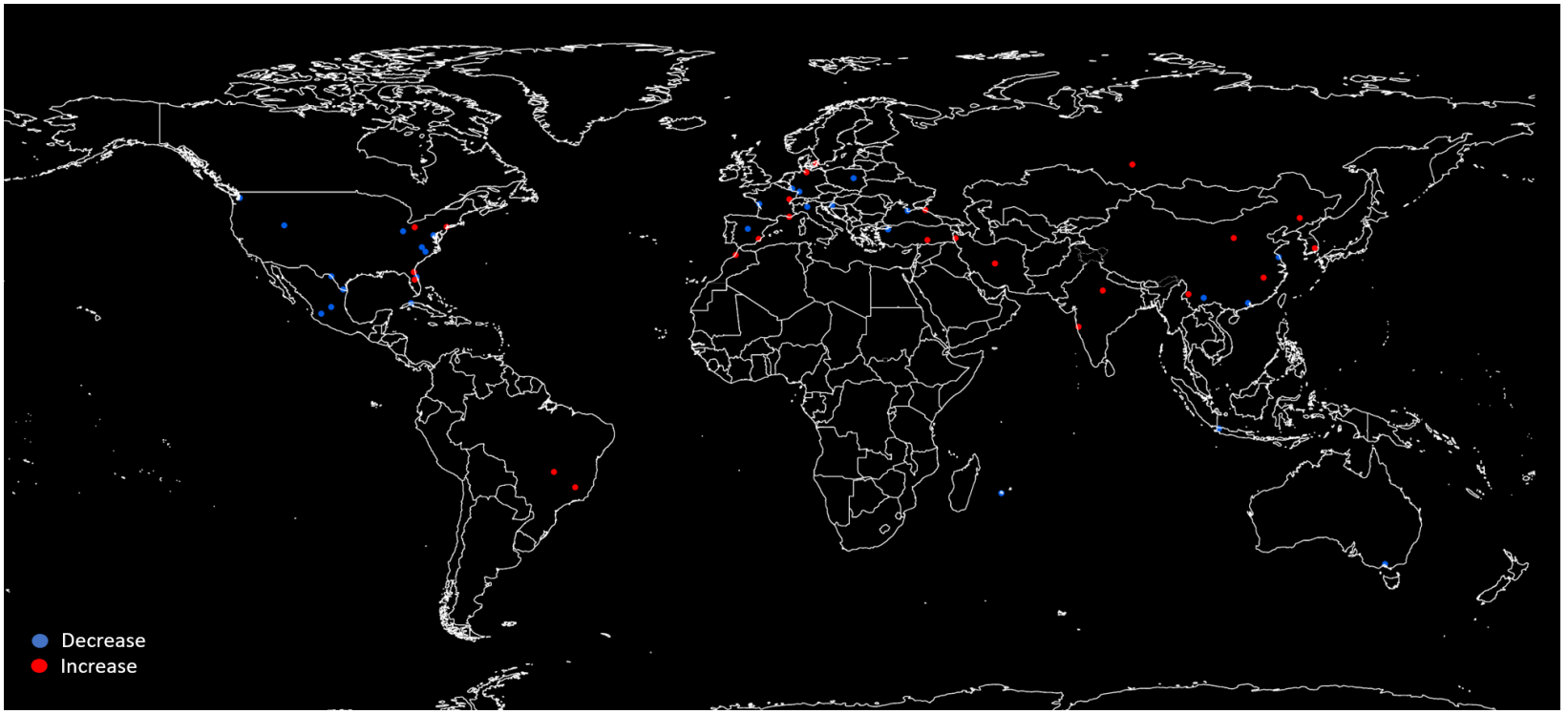


Table 15: Deviations from five-year means

		Percent Deviation	
City	Country	2015-19	2017-21
Greensboro	United States	21.78	-20.46
Warsaw	Poland	17.66	-37.90
Melbourne	Australia	-9.38	-14.70
Madrid	Spain	-14.72	-14.70
La Rochelle	France	-15.25	-14.89
Fayetteville	United States	-19.15	-22.01

6. Estimating Urban Emissions

Our econometric analysis also enables estimation of emissions levels for selected regions or local areas. In our approach, emissions levels have two components. The first, predicted by our econometric model, is the expected value of emissions from an area, given its industry structure, fire occurrences, population, income per capita, climate and mass transit infrastructure. The second component is the deviation from a city's expected value measured by its regression residual. Negative deviations identify areas whose emissions are lower than expected, while positive deviations identify areas with higher-than-expected emissions.

For both components, atmospheric concentration measures can be converted to estimated emissions with exogenous information that provides measurement benchmarks. For the computation of expected values, we employ standard global emissions estimates. For the translation of regression residuals to emissions deviations, we use the EDGAR gridded database of CO_2 emissions estimated from sectoral activity data and standard emissions parameters (Crippa et al. 2020).

6.1 Mapping Global CO_2 Emissions

Our econometric exercise measures marginal effects on atmospheric concentration anomalies from CO_2 emissions by industry, fires, and non-industrial population-related sources. Concentration anomalies are measured in parts per billion; we convert to emissions volumes using a global adjustment. With 2019 as the benchmark year, Ritchie and Roser (2020) estimate global CO_2 emissions of 35.4 gigatons (Gt) from fossil fuels and industrial sources. Bailis et al. (2015) estimate global CO_2 emissions of 1.1 Gt from traditional woodfuels that are still used for heating in many locations. Adding the two yields a total estimate (GCO2) of 36.5 Gt.

Using the estimated model coefficients in Table 1, we predict industry- and population-related concentration anomalies for each of the 266,884 grid cells in our global database. The industry component (ICO2) for each cell is the sum of predicted effects for industry emissions in the cell and wind-displaced industry emissions from neighboring cells. The population component (PCO2) is the sum of predicted effects for heating degree days, cooling degree days, and income per capita, adjusted by the subway impact multiplier. We sum across cells to produce the global aggregate concentration anomaly:

(2) $TCO2 = \sum_{i}^{266,884}(ICO2_i + PCO2_i)$.

We obtain the global conversion factor, as follows:

(3) $gc = \frac{GCO2}{TCO2} = 563.6$.

The conversion factor, *gc*, is dimensioned as tons of CO_2 emissions per 1 part per billion of concentration anomaly.

Fires CO_2 emissions for each cell are drawn from Van der Werf et al. (2017). These are gross emissions rather than estimated net emissions from land-use change that incorporate both carbon-emitting conversions (e.g., from forest to cropland) and carbon-absorbing conversions (e.g., from pasture to forest).[18]

Using the econometric results in Table 1, Figure 7 maps the predicted global distribution of CO_2 emissions using grid cell sums of estimated industry-related emissions ($gc * ICO2_i$), population-related emissions ($gc * PCO2_i$), and fires emissions from Van der Werf et al. (2017). The figure displays mean annual emissions for 2014–21, highlighting the roles of the United States, Europe, China, and the forested regions of South America, Central Africa, and Southeast Asia.

The map illustrates the potential of satellite-based observations for CO_2 emissions monitoring at regional, national, and local scales. Until now, observation-based estimates of CO_2 emissions by source from every location on the globe have simply not been available.

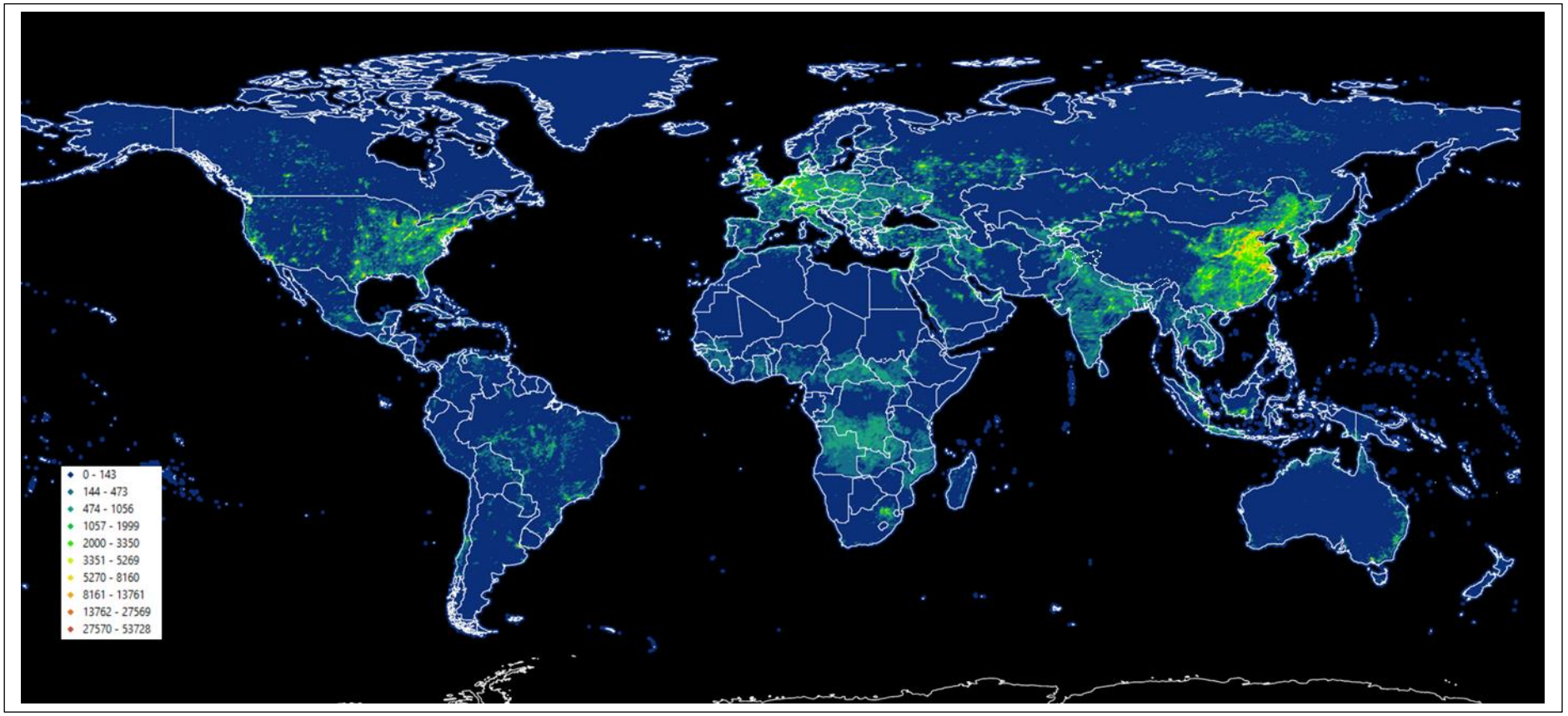
6.2 Emissions from Functional Urban Areas

Using the data described in Section 6.1, we estimate emissions from urban areas by aggregating estimates for grid cells within each Functional Urban Area. These can be viewed as expected emissions values for FUAs, given their characteristics. In reality, the impacts of different policies, individual economic decisions and more fine-grained structural factors will cause FUAs to deviate from their expected values. These deviations should be reflected in the residual concentration anomalies from estimation of model (1).

We translate these residuals to emissions using the EDGAR global database of gridded CO_2 emissions estimated from local activity measures and standard emissions parameters (Crippa et al. 2020). The current EDGAR database terminates in 2018, so we perform the conversion using data for 2015 – 2018, the period of overlap with our OCO-2 database.

[18] For detailed assessments of net carbon emissions from land-use change, see Gasser et al. (2020) and Winkler et al. (2021).

Figure 7: Mean annual CO_2 emissions, 2014-2021
('000 tons/year)
Global grid, 25 x 25 km



After aggregating EDGAR emissions for the grid cells within each FUA, we estimate a regression model that relates annual EDGAR emissions to annual local concentration anomalies for 5,991 FUAs. The result, report in Table 16, is highly significant and indicates that 1 ppb of the local concentration anomaly for an FUA is associated with 394 tons of CO_2 emissions. The R^2 in this case is extremely small, because the "noise" (random variation) in the local anomaly measure is much larger than the "signal" associated with local emissions.

Table 16: Regression results: EDGAR emissions vs. local CO_2 anomalies ('000 tons CO_2)

Dependent variable: EDGAR FUA Emissions ('000 tons)

Local Concentration Anomaly (ppb)	0.394*** (8.88)
Const	3023.2*** (37.42)
R^2:	0.003
N	23,964[a]

[a] 5,991 FUAs x 4 years

t statistics in parentheses

* p<0.05 ** p<0.01 *** p<0.001

Table 17 applies the results in Table 16 to the regression residuals from model (1), converting the residuals to deviations from expected emissions. After ranking urban areas by estimated CO_2 emissions, Table 17 presents the top-40 positive and negative cases. For each urban area, we express the emissions deviation as a percent of emissions. Among the cities with emissions above expected levels (positive deviations), 3 of the cities with deviations greater than 10% are in China (Fuzhou, Jieyang, Xinxiang) and 3 are in the United States (Los Angeles, Seattle, Phoenix), along with Tehran, Islamic Republic of Iran. Two cities have emissions that are 10% or more below expected levels: Buenos Aires, Argentina and Indianapolis, United States.

Table 17: FUAs : Estimated and residual-adjusted CO_2 emissions, 2015-2021

FUA	Country	Pop (Mill.)	CO_2 ('000 tons)	Residual Adjustment ('000 tons)	Pct. Diff.	FUA	Country	Pop (Mill.)	CO_2 ('000 tons)	Residual Adjustment ('000 tons)	Pct. Diff.
Tokyo	Japan	36.5	196,161.8	413.8	0.2	New York	United States	19.5	193,858.0	-733.8	-0.4
Guangzhou	China	45.6	120,382.1	3,302.1	2.7	Chicago	United States	8.8	124,360.9	-3,374.6	-2.7
Houston	United States	6.4	95,739.5	2,662.9	2.8	Dortmund	Germany	5.8	119,878.6	-1,199.2	-1.0
Osaka [Kyoto]	Japan	17.6	93,117.5	5,634.8	6.1	Beijing	China	21.3	107,335.0	-516.3	-0.5
Dallas	United States	7.1	90,937.4	1,251.4	1.4	Moscow	Russian Federation	16.4	91,197.3	-4,426.3	-4.9
Shanghai	China	26.9	90,440.7	1,030.6	1.1	Shenyang	China	6.2	88,373.1	-2,071.9	-2.3
Nagoya	Japan	9.6	88,495.9	1,297.1	1.5	Minneapolis	United States	3.3	73,577.9	-3,712.0	-5.0
Seoul	Korea, Rep.	24.3	87,532.2	5,561.7	6.4	Tianjin	China	8.3	65,981.3	-1,639.8	-2.5
Los Angeles	United States	15.7	86,961.0	14,602.3	16.8	Toronto	Canada	7.3	65,881.2	-42.1	-0.1
Philadelphia	United States	6.1	75,752.3	2,045.5	2.7	Changchun	China	3.9	65,726.8	-905.5	-1.4
Suzhou	China	12.0	70,628.8	4,966.9	7.0	Detroit	United States	4.1	64,425.6	-2,189.8	-3.4
Zibo	China	3.7	61,212.4	2,131.1	3.5	Paris	France	11.2	63,338.8	-106.7	-0.2
Tangshan	China	2.4	55,119.2	914.2	1.7	London	United Kingdom	12.6	61,295.3	-487.9	-0.8
Frankfurt	Germany	3.0	53,977.1	1,123.3	2.1	Atlanta	United States	5.6	56,993.9	-308.2	-0.5
Hangzhou	China	9.6	50,879.0	287.2	0.6	Buenos Aires	Argentina	15.0	56,823.6	-11,623.6	-20.5
Cologne	Germany	3.2	48,279.0	967.8	2.0	Pittsburgh	United States	2.0	55,276.0	-1,434.9	-2.6
Nanjing	China	7.0	46,221.0	443.5	1.0	Washington D.C.	United States	5.6	52,849.9	-6.5	0.0
Seattle	United States	3.9	44,020.1	11,492.8	26.1	Chongqing	China	6.0	48,684.5	-128.2	-0.3
Jieyang	China	12.7	43,030.7	6,067.3	14.1	Istanbul	Türkiye	14.8	48,503.2	-2,787.3	-5.7
St. Louis	United States	2.5	41,553.7	845.4	2.0	Manchester	United Kingdom	3.3	46,283.0	-995.3	-2.2
Xinxiang	China	2.0	41,235.6	6,808.5	16.5	Wuhan	China	8.5	44,688.5	-133.2	-0.3
Linyi	China	2.7	40,904.7	1,467.1	3.6	Chengdu	China	11.8	43,597.9	-233.4	-0.5
Baotou	China	2.2	39,042.1	9.6	0.0	Katowice	Poland	2.9	42,871.1	-1,416.6	-3.3
Phoenix	United States	4.6	38,373.2	10,464.9	27.3	Cincinnati	United States	1.9	42,699.9	-1,411.9	-3.3
Zurich	Switzerland	2.1	37,236.9	2,959.0	7.9	Milan	Italy	5.1	42,150.6	-547.7	-1.3
Jiaxing	China	2.5	36,774.3	909.3	2.5	Jakarta	Indonesia	39.8	42,132.0	-2,002.2	-4.8
Stuttgart	Germany	2.2	36,360.4	578.8	1.6	Leeds	United Kingdom	2.9	42,104.2	-126.2	-0.3
Taizhou	China	4.3	36,250.8	2,840.4	7.8	Tianjiaan	China	1.6	39,360.5	-295.1	-0.7
Kansas City	United States	2.0	36,002.6	816.3	2.3	Birmingham	United Kingdom	3.0	39,274.5	-626.4	-1.6
Portland	United States	2.5	35,873.0	1,558.2	4.3	Montreal	Canada	4.1	38,871.5	-1,135.7	-2.9
Charlotte	United States	2.0	35,542.7	712.4	2.0	Melbourne	Australia	4.5	37,527.0	-309.4	-0.8
Antwerp	Belgium	1.8	33,139.5	303.5	0.9	Wuhu	China	1.4	36,490.7	-288.7	-0.8
Cixi	China	3.2	32,973.4	705.7	2.1	Columbus	United States	2.2	35,512.3	-1,817.0	-5.1
Mannheim	Germany	1.7	32,856.3	652.7	2.0	Datong	China	1.7	33,693.9	-83.5	-0.2
Tehran	Iran, Islamic Rep.	13.4	32,587.6	4,074.3	12.5	Indianapolis	United States	1.9	32,981.6	-3,622.0	-11.0
Shijiazhuang	China	4.2	32,380.4	1,582.2	4.9	Louisville	United States	1.3	32,821.3	-1,701.7	-5.2
Zhangjiagang	China	1.5	32,326.6	152.1	0.5	Jilin	China	1.7	32,262.6	-1,166.0	-3.6
Luoyang	China	2.4	32,214.0	1,089.7	3.4	San Antonio	United States	2.2	29,828.8	-836.8	-2.8
Fuzhou	China	4.8	32,204.2	3,420.4	10.6	Warsaw	Poland	2.9	28,431.0	-1,221.1	-4.3
Anshan	China	1.9	32,134.5	66.0	0.2	Harbin	China	4.6	28,378.2	-1,209.2	-4.3

Table 18 summarizes our adjustment results for 1,306 Functional Urban Areas with populations greater than 500,000. It tabulates percent deviations by range for urban areas with higher- and lower-than-expected emissions. The table reveals the frequency of relatively large deviations in both categories, with the representation of higher-percent ranges significantly greater for urban areas with lower-than-expected emissions.

Table 18: FUAs: Distribution of Percent Deviations by Performance Category

	Emissions Relative to Expected Value		
% Deviation	Lower	Higher	Total
0-1	64	52	116
	9.28	8.44	8.88
1-5	120	150	270
	17.39	24.35	20.67
5-10	91	104	195
	13.19	16.88	14.93
10-20	123	123	246
	17.83	19.97	18.84
20-50	151	98	249
	21.88	15.91	19.07
50+	141	89	230
	20.43	14.45	17.61
Total	690	616	1,306

7. Summary and Conclusions

This paper has extended recent research on satellite-based CO_2 measurement to an easily-updated template for tracking changes in CO_2 concentration anomalies at local and regional scales. Using observations from NASA's OCO-2 platform, we develop the template from the data filtering techniques and econometric analysis employed by Dasgupta, Lall and Wheeler (2022). For a large sample of urban areas, we compare alternative trend estimation models and conclude that the template can be constructed from a simple model that estimates trends directly from OCO-2 data that are pre-filtered to isolate local concentration anomalies. We also investigate the impact of regional atmospheric circulation on local changes and develop a methodology for identifying urban areas whose changes are independent of regional circulation effects. We present illustrative applications for a long-period trend model and a short-period model that focuses on changes in the most recent year. In both cases we find striking patterns of variation, across regions, within regions, and over time.

We also use our estimation results to compute expected emissions for a large number of urban areas. Then we convert the regression residuals to emissions deviations using information from the EDGAR global database of estimated CO_2 emissions. This enables us to quantify the extent to which urban emissions are greater or less than their expected values. Among 1,306 urban areas with populations greater than 500,000, we find a rough balance between cities with positive and negative emissions deviations. In the distribution of percent deviations, we find typically-higher deviations among cities whose emissions are below their expected values.

We conclude with an acknowledgment that even "simple" trend tracking and emissions estimation require software and hardware that are beyond the means of many interested stakeholders. For this reason, the World Bank's Development Economics Vice Presidency (DEC) has established an open web facility, the XCO2 database,[19] that pre-filters the OCO-2 data and publishes monthly mean concentration anomalies for all terrestrial cells of a 25 km global grid (G25). The website will also publish annual change estimates for urban areas by statistical significance class and provide information that links G25 grid cell IDs to IDs for urban areas and national administrative units (levels 0, 1 and 2). We believe that this web facility will contribute to the global effort to reduce CO_2 emissions as rapidly as possible. We provide more details in the Appendix.

[19] Available online at https://datacatalog.worldbank.org/search/dataset/0062760.

References

Auch, T. 2017. Tracking Oil Refineries and Their Emissions: a complete inventory of global oil refineries. https://www.fractracker.org/2017/12/global-oil-refineries-emissions/

Bailis, R., R. Drigo, A. Ghilardi, A. et al. 2015. The carbon footprint of traditional woodfuels. Nature Climate Change 5: 266–272.

Burbidge, J., L. Magee and A. Robb. 1988. Alternative Transformations to Handle Extreme Values of the Dependent Variable.Journal of the American Statistical Association, 83: 123-127.

Byers, L., J. Friedrich et al. 2021. A Global Database of Power Plants. World Resources Institute.

Chakraborty, N., I. Mukherjee, A.K. Santra et al. 2008. Measurement of CO_2, CO, SO2, and NO emissions from coal-based thermal power plants in India. Atmospheric Environment. Volume 42 (6): 1073-1082.

CIESIN (Center for International Earth Science Information Network, Columbia University). 2021. Documentation for the Gridded Population of the World, Version 4 (GPWv4), Revision 11.

Crippa, M., E. Solazzo, G. Huang, D. Guizzardi, et al. 2020. High resolution temporal profiles in the Emissions Database for Global Atmospheric Research. Sci Data 7: 121.

Dasgupta, S., S. Lall and D. Wheeler. 2021. Traffic, Air Pollution and Distributional Impacts in Dar es Salaam: A Spatial analysis with New Satellite Data, Science of the Total Environment (forthcoming).

Dasgupta, S., S. Lall and D. Wheeler. 2022. Subways and CO_2 Emissions: A Global Analysis with Satellite Data (under review for publication).

GADM. 2021. Maps for the administrative areas of all countries. https://gadm.org/data.html

Gale, J. et al. 2005. Chapter 2: Sources of CO_2, in Metz, B., D. Ogunlade et al., Carbon Dioxide Capture and Storage. Cambridge University Press, UK.

Gasser, T., L. Crepin, Y. Quilcaille, R. Houghton, P. Ciais and M. Obersteiner. 2020. Historical CO_2 emissions from land use and land cover change and their uncertainty. Biogeosciences, 17: 4075–4101.

GEM (Global Energy Monitor). 2021. Global Steel Plant Tracker.

Gendron-Carrier, N., M. Gonzalez-Navarro, S, Polloni and M. Turner. 2020. Subways and Urban Air Pollution. National Bureau of Economic Research, Working Paper 24183.

Hakkarainen, J., I. Ialongo, S. Maksyutov, and D. Crisp. 2019. Analysis of Four Years of Global XCO2 Anomalies as Seen by Orbiting Carbon Observatory-2. Remote Sensing 11, no. 7: 850.

Heger M., D. Wheeler, G. Zensa and C. Meisner. 2018. Motor Vehicle Density and Ambient Air Pollution in Greater Cairo – How Did Fuel Subsidy Removal and Metro Line Extension Effect Congestion and Pollution? Washington DC: World Bank.

Hersbach, H., B. Bell, P. Berrisford, G. Biavati, A. Horányi et al. 2019. ERA5 monthly averaged data on single levels from 1979 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS).
https://cds.climate.copernicus.eu/cdsapp#!/dataset/10.24381/cds.f17050d7?tab=form

IEA (International Energy Agency). 2020. IEA Energy Technology Perspectives. p. 216.
https://iea.blob.core.windows.net/assets/7f8aed40-89af-4348-be19-c8a67df0b9ea/Energy_Technology_Perspectives_2020_PDF.pdf

Jing, L., H. El-Houjeiri, J. Monfort et al. 2020. Carbon intensity of global crude oil refining and mitigation potential. Nat. Clim. Chang. 10: 526–532.

Jotzo, F., P. Burke, P. Wood, A. Macintosh and D. Stern. 2012. Decomposing the 2010 global carbon dioxide emissions rebound. Nat. Clim. Chang. 2(4): 213–214.

JPL/NASA. 2021. OCO-2 Data Set. Jet Propulsion Laboratory, California Institute of Technology. https://co2.jpl.nasa.gov/?mission=oco-2

Labzovskii, L., S. Jeong and N. Parazoo. 2019. Working towards confident spaceborne monitoring of carbon emissions from cities using Orbiting Carbon Observatory-2. Remote Sens. Environ. 233: 111359.

Lam, N., E. Wallach, C. Hsu, A. Jacobson, P. Alstone, P. Purohit and Z. Klimont. 2019. The Dirty Footprint of the Broken Grid: The Impacts of Fossil Fuel Back-up Generators in Developing Countries. International Finance Corporation, Washington DC. September.

Layton, D. 2001. Alternative Approaches for Modeling Concave Willingness to Pay Functions in Conjoint Valuation. American Journal of Agricultural Economics, 83(5): 1314-20.

McCaffrey, R. et al. 2021. Global Cement Directory. Pro Global Media Limited, Octagon House, Surrey, United Kingdom.

Mistry, M. 2019. Historical global gridded degree days: A high spatial resolution database of CDD and HDD. Geoscience Data Journal, 6(2): 214-221.

Nassar, R., T. Hill, C. McLinden, D. Wunch, D. Jones and D. Crisp. 2017. Quantifying CO_2 emissions from individual power plants from space. Geophysical Research Letters, 44: 10,045–10,053.

Nassar, R., J. Mastrogiacomo, W. Bateman-Hemphill et al. 2021. Advances in quantifying power plant CO_2 emissions with OCO-2. Remote Sensing of Environment. 264: 112579.

Nordhaus, W., A. Azam et al. 2006. The G-Econ Database on Gridded Output: Methods and Data. Yale University.

OSM (OpenStreetMap). 2021. OpenStreetMap contributors. (2015) Planet dump [Data file from Feb. 16, 2022]. Retrieved from https://planet.openstreetmap.org.

Pan, G., X. Yuan and M. Jieqi. 2021. The potential of CO_2 satellite monitoring for climate governance: A review. Journal of Environmental Management, 277: 111423.

Raupach, M., G. Marland, P. Ciais, C. Le Quéré et al. 2007. Global and regional drivers of accelerating CO_2 emissions. Proc. Natl. Acad. Sci. 104(24): 10288–10293.

Reuters. 2018. Global temperatures on track for 3-5 degree rise by 2100: U.N. https://www.reuters.com/article/us-climate-change-un/global-temperatures-on-track-for-3-5-degree-rise-by-2100-u-n-idUKKCN1NY186?edition-redirect=uk

Sanchez, L. and D. Stern. 2016. Drivers of industrial and non-industrial greenhouse gas emissions. Ecological Economics, 124:17–24.

Schiavina, M., A. Moreno-Monroy, L. Maffenini and P. Veneri, Paolo. 2019. GHS-FUA R2019A - GHS functional urban areas, derived from GHS-UCDB R2019A (2015). European Commission, Joint Research Centre (JRC) [Dataset] doi:10.2905/347F0337-F2DA-4592-87B3-E25975EC2C95 PID: http://data.europa.eu/89h/347f0337-f2da-4592-87b3-e25975ec2c95.

Steffen, W., J. Rockström and K. Richardson. 2018. Trajectories of the Earth System in the Anthropocene. PNAS 115(33): 8252-8259.

Turner, M. and M. Gonzalez-Navarro. 2018. Subways and urban growth: Evidence from earth. Journal of Urban Economics 108: 85-106.

UN (United Nations). 2021. GDP and its breakdown at constant 2015 prices in US Dollars - all countries for all years - sorted alphabetically. UN Department of Economic and Social Affairs, Statistics Division, National Accounts. https://unstats.un.org/unsd/snaama/downloads

USEIA (United States Energy Information Administration). 2021. How much carbon dioxide is produced per kilowatt hour of U.S. electricity generation? https://www.eia.gov/tools/faqs/faq.php?id=74&t=11

Van der Werf, G. , J. Randerson et al. 2017. Global fire emissions during 1997-2016. Earth Syst. Sci. Data, 9: 697–720.

Weir, B., D. Crisp, C. O'Dell et al. 2021. Regional impacts of COVID-19 on carbon dioxide detected worldwide from space. Science Advances, 7(45).

Winkler, K., R. Fuchs, M. Rounsevell et al. 2021. Global land use changes are four times greater than previously estimated. Nature Communications, 12: 2501.

World Bank. 2012. Turn Down the Heat ~ Why a 4°C warmer world must be avoided. World Bank: Washington, D.C.

World Steel Association. 2021. Sustainability Indicators. https://www.worldsteel.org/steel-by-topic/sustainability/sustainability-indicators.html

Wu, D., J. Lin, T. Oda and E. Kort. 2020. Space-based quantification of per capita CO_2 emissions from cities. Environmental Research Letters, 15: 035004.

Ye, X., T. Lauvaux, E. Kort, T. Oda, S. Feng, J. Lin, E. Yang and D. Wu. 2020. Constraining fossil fuel CO_2 emissions from urban area using OCO-2 observations of total column CO_2. JGR Atmospheres, 125(8).

Appendix: Mobilizing OCO-2 Data for the Stakeholder Community

In this paper, we have shown that a relatively simple tracking model can provide useful information about changes in local CO_2 concentration anomalies for areas of interest. However, we recognize that most stakeholders do not have the requisite hardware and software for mobilizing the OCO-2 data directly. Accordingly, the World Bank's Development Economics Vice Presidency (DEC) has established an open web facility with the following features. We believe that it will contribute to the global effort to reduce CO_2 emissions.

- The XCO2 database (https://datacatalog.worldbank.org/search/dataset/0062760), an OCO-2 panel database for a 25 km grid (G25), in Stata format, beginning in September 2014 and updated regularly. The database includes G25 grid cell ID numbers, cell centroid coordinates, monthly means of measured CO_2 concentrations, and monthly means of Hakkarainen pre-filtered CO_2 concentration anomalies. The JPL/NASA database publishes OCO-2 data with a lag of approximately two months.

- For functional urban areas (FUAs) with sufficient data, annually-updated change parameter estimates for models (6) and (7), with statistical significance categories [$p>.05$, $\leqslant.05$, $\leqslant.01$, $\leqslant.001$].

- Stata files that match G25 grid cell ID numbers for:

    - Functional urban areas (ID numbers from Schiavina et al. (2019));
    - National level 0, 1 and 2 administrative areas (World Bank GADM (2021)).