The urban share and regional urban dynamics of global greenhouse gas emissions

Kevin Gurney

Kevin.Gurney@nau.edu

Northern Arizona University https://orcid.org/0000-0001-9218-7164

Bilal Aslam

Northern Arizona University

Pawlok Dass

Northern Arizona University https://orcid.org/0000-0003-3957-4055

Shanthi Podhili

Northern Arizona University

Anna Kato

northern arizona university https://orcid.org/0000-0003-2758-6817

Article

Keywords:

Posted Date: June 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9868204/v1

Additional Declarations: There is **NO** Competing Interest.

The urban share and regional urban dynamics of global greenhouse gas emissions

Kevin Robert Gurney[1*], Bilal Aslam[1], Pawlok Dass[1], Shanthi Swaroopa Hari Podhili[1], Anna Kato[1]

Affiliations

[1]School of Informatics, Computing, and Cyber Systems, Northern Arizona University, Flagstaff, AZ, USA

*Correspondence to: kevin.gurney@nau.edu

Abstract: The share of global GHG emissions in urban areas, home to the majority of global population, is a central metric in establishing the importance and prioritization of cities for global GHG emissions mitigation efforts. However, the urban share of global GHG emissions remains poorly quantified due to a paucity of accurate and comprehensive urban GHG emissions data and lack of clarity on both geographical and system boundaries. Here, we use a new machine learning-based global fossil fuel CO_2 ($FFCO_2$) emissions dataset (*Vulcan-UrbanML*) to estimate the global urban share of territorial $FFCO_2$ emissions. We quantify the urban share for multiple geographic boundaries, system boundaries, and IPCC regional domains over the 2010 to 2022 time period. We find a 2022 global territorial $FFCO_2$ emissions urban share range of 63.6-69.9%, depending upon the urban boundary considered. The consumption-based global urban share is estimated as 65.3%-79.3%. Per capita urban emissions are significantly lower than their rural counterparts in the Developed Countries, the Middle East, and Eastern Europe and West-Central Asia regions. The global territorial $FFCO_2$ emissions urban share increased 1.4% over the 13 year time period examined here.

One sentence summary: Cities are responsible for between 63.6% and 69.9% of global fossil fuel CO_2 emissions.

Main: Cities and urban areas, home to the majority of global population, occupy a pivotal planetary role as they face the risks and responses to climate change. Chief among these is the potential to mitigate GHG emissions through a wide collection of policies and measures (Lwasa et al., 2022). Much of the justification for this recognition in recent literature and assessment reports, including the upcoming IPCC Special Report on Cities, is driven by the recognition that cities and urban areas constitute a large share of global GHG emissions and are thereby crucial to global progress on combating climate change (Gurney et al., 2022).

In spite of its importance, surprisingly few estimates of the global GHG emissions urban share have been made. More importantly, the few estimates that have been made vary considerably and are sensitive to a series of underlying assumptions and approximations. The first, and perhaps most important limitation, is that there is no globally comprehensive, bottom-up collation of urban GHG emissions though national and regional subsets exist (Ahn et al., 2023; Huo et al., 2022; Kona et al., 2021; Kongboon et al., 2022; Liu et al., 2024; Moran et al., 2022; Nangini et al., 2019). The aggregation of gridded global GHG emissions data to urban areas offers a "top-down" alternative estimate of city GHG emissions, and two studies have taken this approach. The first, aggregated a global 250m consumption-based CO_{2eq} (CO_2+CH_4) emissions dataset into cities and towns estimating a global urban share of 68% in the year 2015 (Moran et al., 2018).

This was later extrapolated to 2020 arriving at a 67-72% range for the global urban consumption-based CO_{2eq} share (Gurney et al., 2022; Lwasa et al., 2022).

A global 0.1° x 0.1° gridded territorial CO_2 emissions dataset was aggregated into "urban centers" (akin to a central core in cities, leaving out peri-urban or suburban portions of a city) and estimated the global urban anthropogenic CO_2 emissions share at 35% (Crippa et al., 2021). Earlier estimates, scaled from mean regional per capita urban emission estimates synthesized by the IPCC, reported an urban share range of 71-76% for territorial CO_2 emissions for the early 2000s time period (Seto et al., 2014).

The variation in these estimates owes to three fundamental choices. Given that there is no single standard database of urban GHG emissions, the first element of variation in the estimation of the global urban GHG emissions share is the emission dataset used, its time period, and how the spatial scaling (i.e., downscaling from nation, aggregation from grid cells) is performed (Shi et al., 2025; Sun et al., 2026; Gurney et al., 2026).

The second choice is the specification of system boundaries. System boundaries refer to which elements of the within- and trans-boundary GHG emission flows are to be included in a given urban GHG emissions account (Chen et al., 2020). Examples of within-boundary choices are which sectors/processes are included (e.g., residential, railroad, industrial, combustion-only), whether to include biological GHG exchange, and which gases or categories of gases are included (e.g., CO_2 versus CO_2eq). An example of a trans-boundary GHG emissions flow the GHG emissions in city A associated with the supply chain of goods consumed within a city B. Trans-boundary emission differences are conventionally classified as territorial emissions (physically entering the atmosphere within a city) and consumption-based emissions (accounting for emissions regardless of location driven by the consumption of goods and services within a city).

The final key ingredient is the choice of urban geographical boundaries. Since the definition of what constitutes a city or urban area is often dependent upon political, physical, and social norms which vary across the globe, common or systematically agreed upon geographic extent has been difficult to establish (Uchiyama & Mori, 2017).

The differences in these three fundamental aspects of generating an estimate of the global urban share of GHG emissions makes the few existing estimates incomparable and can be misleading if used without qualifications as to the boundaries and data used.

Here, we overcome the shortcomings of the existing studies by building a new machine-learning-based, city-scale global estimate of annual fossil fuel CO_2 ($FFCO_2$) emissions trained on a well-established, high-resolution, atmospherically tested U.S. $FFCO_2$ emissions estimate spanning the 2010-2022 time period. We estimate these territorial $FFCO_2$ emissions for four globally available urban geographic boundaries. We also aggregate an existing consumption-based global gridded GHG (CO_2 + CH_4) emissions dataset into the same urban boundary ensemble to provide insight into the implications of differing system boundaries. We estimate the global and regional urban share and further examine the relationship between population, economics, and land area in relation to the global and regional urban $FFCO_2$ emissions share across the 13 years of urban $FFCO_2$ emissions.

Results and Discussion:

The urban share of global and regional $FFCO_2$ emissions for the year 2022 across all four urban boundary definitions and two different system boundaries using the Vulcan-UrbanML estimates are shown in Table 1 (see SI, Table S1 for results across all years). Globally, the urban share of territorial $FFCO_2$ emissions ranges from 63.6% - 69.9%, reflecting the range of values across four urban boundary definitions. This contrasts with the recent estimate of 35% based on urban aggregation of the Emissions Database Global Atmospheric Research (EDGAR) global gridded anthropogenic CO_2 emissions (Crippa et al., 2024). A key distinction is that the urban boundary used in that study, the Global Human Settlement - Urban Centre Database (GHS-UCDB) boundary, is representative of the urban center or "core" area of cities (Melchiorri et al., 2024). Using that boundary with the Vulcan-UrbanML dataset developed here arrives at an urban share of 40.6% (see SI, Table S2).

For the consumption-based emissions, the 2022 global urban share ranges from 65.3% - 79.3%, somewhat larger than the territorial emissions share. This consumption-based range is consistent with a previously reported consumption-based urban GHG emissions share of 67-72% (Gurney et al., 2022; Lwasa et al., 2022).

Regionally, the 2022 territorial urban $FFCO_2$ emissions shares range from 44.4%-56.0% (Eastern Europe & West-Central Asia) to 69.8-72.9% (Asia & Developing Pacific). The consumption-based equivalent emissions range from 55.1%-72.2% (Africa) to 70.9%-85.7% (Developed Countries).

The absolute 2022 global territorial urban $FFCO_2$ emissions range from 5.42-5.96 GtC. In the six IPCC regions, the values range from 0.16-0.22 GtC (Africa) to 2.96-3.09 GtC (Asia and Developing Pacific) (Figure 1 and SI Table S1).

The relationship between land area, population, GDP, and $FFCO_2$ emissions across the six regional aggregations in the year 2022 is presented in Figure 3. The urban emissions share value is commensurate with the population share in roughly half of the six global regions. Where the population share exceeds the urban $FFCO_2$ emissions share, the urban per capita emissions are, by definition, less than the per capita value for the remainder of the region (the rural land area), and vice-versa. This occurs near-consistently for the Developed Countries (excepting the GUB boundary), the Eastern Europe & West-Central Asia, and the Middle East regions where urban per capita values are lower than the regional rural per capita value. In the Asia & Developing Pacific, Latin America & Caribbean, and Africa regions, the urban per capita value exceeds the regional rural per capita value.

All regions, with the exception of the GCTB boundary, suggest that urban economic productivity, represented by urban GDP, occupy a greater share of the regional total than their population share. This is particularly evident for the FUA urban boundary and this may reflect the inclusion of outlying industrially active areas with a disproportionately smaller population that is found in the less expansive GUB urban boundary (Bradley et al., 2025; Kerr et al., 2024). In the case of the GCTB boundary, the inclusion of smaller settlements ("towns") may result in an increase in population with limited increase in commercial or industrial activity.

The amount of land within the urban boundaries as a share of the regional total is consistently largest for the Developed Countries and Asia & Developing Pacific regions, regardless of the boundary considered.

Between 2010 and 2022, the global urban $FFCO_2$ emissions grew at average rate of 0.88%-0.90%/year (±0.08%/year) which was dominated by an annual decline of -1.08% to -1.14%/year (±0.16%/year) in the Developed Countries region and a more than compensating increase of +2.25%-2.50%/year (±0.16%/year) in the Asia & Developing Pacific region. The Developed Countries and Asia & Developing Pacific regions together account for 81.1%-84.2% of the global urban $FFCO_2$ emissions. The urban share grew 1.4% (±0.03%/year) over the 13 year time period examined here.

The juxtaposition of four trend values: the urban $FFCO_2$ emissions share, the urban per capita $FFCO_2$ emissions, the absolute urban $FFCO_2$ emissions, and the absolute regional $FFCO_2$ emissions, provide additional insight into the regional urban dynamics (Figure 4). Firstly, the Eastern Europe & West-Central Asia and the Middle East regions show mostly insignificant trends for the urban share and the urban per capita emissions but significant positive and commensurate trends in absolute urban and region-wide $FFCO_2$ emissions. This suggests increasing urban population with limited efficiency or economies of scale associated with the urban emissions.

By contrast, the Africa region shows positive significant absolute urban emissions coupled with a decline in both the urban per capita and urban share values. The total Africa region $FFCO_2$ emissions are also growing slightly faster than the urban component. This is suggestive of increasing urban economies of scale whereby growth in urban emissions occur in spite of decline per capita values.

The Asia & the Developing Pacific region shows a combination of positive significant growth in the urban per capita $FFCO_2$ emissions and both the absolute urban and region total $FFCO_2$ emissions but insignificant changes for the urban share. This is suggestive of a combination of urban population growth and increasing per capita fossil fuel consumption potentially tied to increasing individual wealth. The lack of significant change in the urban share is consistent with the aligned growth in both the urban and region absolute $FFCO_2$ emissions.

The Developed Countries region shows declines in the urban per capita emissions, the absolute urban emissions, and absolute regional emissions but an increase in the urban share. This suggests both urban and rural emission declines, somewhat great in urban areas. This could be driven by increasing economies of scale in urban areas.

Finally, the Latin America & Caribbean region shows all four metrics declining, though with large uncertainties. This mixture of trends implies multiple potential change dynamics from overall urban decline to gains in urban energy consumption efficiency outpacing regional trends.

The distribution of $FFCO_2$ emissions in cities across the planet is lognormally distributed. The cumulative distribution of 2022 urban territorial $FFCO_2$ emissions indicates that, out of 48,262 cities, the top 2 cities (Guangzhou and Shanghai, China) emit an amount equivalent to the national total of Germany (Figure 5). The top 15 cities and amount equal to the national total of India, the top 42 cities an amount equal to the United States and the top 323 cities equivalent to the national total $FFCO_2$ emissions of China.

Though there is no independent means to verify the Vulcan-UrbanML urban emission estimates, we compare the predicted values to a curated set of individual self-reported city emissions (N=77) from the Carbon Disclosure Project (see Methods) (Nangini et al., 2019). The comparison shows a mean relative difference (calculated as: [CDP-Vulcan

UrbanML]/mean[CDP,Vulcan UrbanML]) for the $FFCO_2$ emissions of -13.4% (±7.8%) and a mean relative difference of -5.1% (±8.9%) on a per unit area basis (acknowledging that the urban polygon boundaries cannot be identically matched). Additional information is provided in the Supplementary Information.

Methods

Overview: The urban share analysis in this study is based on a machine learning (ML) approach to estimate urban $FFCO_2$ emissions globally. This method includes different geospatial input datasets, harmonized to the same grid/pixel resolution. These datasets are aggregated to four urban polygon definitions and trained on United States Vulcan version 4.0 territorial $FFCO_2$ emissions similarly aggregated to the same urban polygon definitions across the 2010-2022 time period. The predicted urban $FFCO_2$ emissions are analyzed at the IPCC regional and global scales (see SI, Figure S4).

Input and training data: Open-source, globally consistent geospatial data, aggregated to common urban polygons are used as inputs to the Vulcan-UrbanML model (SI, Table S5). The input data includes a mixture of human energy consumption activity proxies (nighttime lights), economic metrics (GDP), transportation infrastructure (road density), demographic distributions (population), urban infrastructure (building volume and build-up area), and pollutant emissions (NOx). Combining these yields a multidimensional representation of $FFCO_2$ emissions. The input datasets are harmonized to a common 1km x 1km resolution and reprojected to the WGS84 geographic coordinate system (EPSG:4326). The aim is to achieve annual input data spanning the 2010-2022 time period. However, the input data represent a mixture of annual data, 5-year interval data (population, building volume, build-up area), in addition to timespans that do not extend to the full length of the 2010-2022 time period (nighttime lights, GDP, NOx). The 5-year interval data is linearly interpolated to annual increments. Shorter time series are extrapolated, fixing the out-years to the closest annual value. Road density is available for a single year, 2018, and hence, has no interannual variation. The urban boundaries are similarly fixed with no growth or decline over the time period examined here (see information below).

The training dataset used in the ML-based prediction model is the Vulcan version 4.0 annual $FFCO_2$ emissions from 2010 to 2022 (Gurney et al., 2025). This training data is only available for the United States. It has the distinct advantage of having been compared to both atmospheric observations and to in situ monitoring. It shows close consistency with observations in numerous analyses (Basu et al., 2020; Dass et al., 2025; Gurney et al., 2025; Lauvaux et al., 2020).

Urban boundaries: Three different boundary definitions are used to predict the urban $FFCO_2$ emissions. The first is the Functional Urban Areas (FUA) dataset (Moreno-Monroy et al., 2021). The FUA is defined using a combination of population density and vehicle commuting flows by combining dense urban cores with their surrounding functional labor-market catchments. Globally, there are 9,030 urban entities in the FUA dataset covering an area of 2.49×10^6 km^2. The second boundary is the Global Urban Boundaries (GUB) dataset (Li et al., 2020). GUB is derived from satellite-observed built-up extent delineating contiguous urbanized land purely on the basis of physical infrastructure and settlement morphology (Li et al., 2020). Both FUA and GUB reflect conditions circa 2015. Globally, there are 65,432 urban entities in the GUB dataset covering an area of 8.1×10^5 km^2. Hence, though there are far fewer FUA urban entities, they occupy a much larger total land area than the GUB boundary. Third, is the Global City and Town Boundaries (GCTB) dataset which relies on a combination of remotely-sensed impervious surface information and 1km^2 population (Bai et al., 2025). It represents annually resolved information and includes 89,034 urban entities ("cities" and "towns") encompassing 8.04×10^5 km^2 of land surface.

Given the different metrics used to delineate these boundaries there are instances where the FUA and GUB urban entities fully or partially overlap and instances in which they are discrete. This is particularly the case for the GUB boundary where there are many GUB urban entities that are not associated with a FUA urban entity. These are typically smaller urban areas that do not achieve the threshold cutoff for the FUA definition but are considered urban entities under the GUB definition. To represent a final more encompassing urban boundary definition, we combined the FUA and GUB boundaries into a fourth urban boundary definition ("FUGUB"). Where a GUB boundary is partly or wholly contained within a FUA boundary the portion contained is fully eliminated from the GUB dataset. GUB delineations that are partly or wholly outside of a FUA urban boundary are retained. This leads to 48,267 FUGUB urban entities covering with an urban area of 2.8×10^6 km^2.

Two additional boundaries were explored, but not used in this study, the GADM, and the UCDB boundaries. See SI, Text S1 for additional details on their consideration.

Model development and evaluation: After upscaling the input and training datasets to the urban boundaries, the resulting dataset is a polygon-year feature matrix, where each observation corresponds to a specific urban area and year. This structure enables supervised learning of relationships between input data (urban characteristics) and the target urban $FFCO_2$ emissions. Input variables are transformed (e.g., logarithmic scaling) to reduce skewness and standardized to improve numerical conditioning.

Vulcan v4 $FFCO_2$ emissions were upscaled from the point, line, and polygon intrinsic spatial scale to the same urban boundaries as the input dataset. Because the Vulcan dataset is only available in the U.S. domain; hence, the model was trained and tested in the U.S. and used to estimate $FFCO_2$ emissions globally.

Since U.S. urban areas may not offer adequate representation of cities in other global geographies, the U.S.-specific input variables aggregated into the urban polygons are matched to the input variables calculated and aggregated to the IPCC region-specific polygons, and only those U.S. urban polygons that have similar input variable characteristics for the region are selected. This is accomplished via the following series of steps: 1) normalize each input variable to standardized units; 2) screen the U.S. urban polygons, retaining only those with standardized input variable values that are within ±1 standard deviation of matched IPCC regional variable distributions; and 3) rank the U.S. urban polygons according to their overall similarity to the IPCC region input variables using Euclidean distance in the standardized feature space (with smaller distances indicating closer matches). The polygons that matched most closely were retained as region-specific training analogs for model development. This is done to minimize the differences due to regionally specific infrastructure, energy consuming behavior, and economic conditions between U.S. cities and cities in the rest of the world. Hence, only the selected polygons for each region and boundary are used to train and test the model.

Three ML models are used, and the best-performing model is selected for global urban $FFCO_2$ emissions estimation. Ridge regression is used as a linear baseline (Hoerl & Kennard, 1970). Histogram-based Gradient Boosting Regressor (HGBR) (Friedman, 2001) and Light Gradient Boosting Machine (LightGBM) (Ke et al., 2017) are used as ensemble ML techniques considered very good for transfer learning. Model performance was assessed using k-fold cross-validation (k = 5–10), with hyperparameter tuning conducted via grid and random search strategies (Bergstra & Bengio, 2012). Feature importance was evaluated using permutation importance and SHAP (SHapley Additive exPlanations) values (Lundberg & Lee, 2017).

Coefficient of determination (R^2), mean absolute error (MAE), root mean squared error (RMSE), mean absolute percentage error (MAPE), and bias are the model performance assessment parameters. Among the tested models, LightGBM consistently achieved superior predictive performance and was selected for estimating global $FFCO_2$ emissions. Detailed diagnostics are shown in Table S6.

A complete description of the Vulcan-UrbanML model and results are in Aslam and Gurney (2026).

Calculation of results: The trained LightGBM model was applied globally to all urban polygons across the four urban boundaries for each year from 2010 to 2022. This yielded spatially estimates of urban territorial $FFCO_2$ emissions worldwide (Figure 2). Predicted emissions were aggregated to regional and global scales using the IPCC AR6 regional classification framework (IPCC, 2021). The 6 regions and the member countries are shown in SI, Table S7. Feature importance indicated that the most influential variables were Nightime lights, population, road density and the two built metrics (volume and area), in that order (SI, Figure S4). GDP and NOx contributed information, but the relative feature importance of these two variables was less than 5%.

To identify the urban $FFCO_2$ emissions share, the EDGAR v8.0 national totals are aggregated to each of the six IPCC regions and these are used as the share denominator (Crippa et al., 2024). The urban share is thus calculated as:

$$\text{Urban FFCO2 emissions share (\%)} = \frac{\sum \text{Predicted urban FFCO2 emissions}}{\text{EDGAR total FFCO2 emissions}} \times 100 \quad (1)$$

It is also important to note that, since Vulcan does not include fugitive emissions, fuel exploitation, and agricultural emissions, these sectors are removed from the EDGAR national totals to match the sector composition of the predicted urban $FFCO_2$ emissions.

Consumption-based CO_{2eq} emissions: The consumption-based GHG emissions used in this study is based on the work of Moran et al., (2018). This 250m resolution consumption-based CO_2eq emissions dataset is upscaled into the

four urban boundaries considered in this study. The gases included in this dataset are somewhat different than the predicted values in the Vulcan-UrbanML model results. The consumption-based account includes CO2 and CH4 and therefore, in calculating the consumption-based urban share, the EDGAR regional totals include both anthropogenic CO_2 and CH_4 are used.

Independent check: We obtained city self-reported inventories from Nangini et al. (2017), which have data from 373 global cities. We matched the cities using fuzzy string-matching, which subsequently reduced the number of comparable cities to 117 for 2016 and 2017. Because the polygon boundaries in this data differ from the geographic urban boundaries used in the Vulcan-UrbanML boundaries, we choose the boundary in the Vulcan-UrbanML results that provide the closest match in total urban area. However, a two-standard-deviation (2SD) filter is subsequently applied to exclude cities with unusually large boundary-area differences. The final dataset included 77 cities worldwide, of which 61 were outside the U.S.. Regionally, the fewest cities were available in the Middle East, with just 1.

We paired these urban emissions to the Vulcan-UrbanML estimates and calculated both an absolute emissions and emissions per area relative difference. This cannot be considered a form of validation as the self-reported inventories can contain large biases and the system boundaries are not consistently reported (Gurney et al., 2021).

Acknowledgements: Thanks to Jocelyn Turnbull and Angel Hsu for insightful discussion. This work was made possible through the generous support of the Builders Bridge/Grantham/Sant Foundations combined grant.

Author Contributions: KRG conceptualized the research and wrote the manuscript. BA performed analysis and calculations. PD, SP, and AK performed elements coding, data collection and contributed to the manuscript development.

Competing Interests: The authors declare no competing interests.

Data availability: The complete Vulcan-UrbanML output is publicly available at (repository TBD).

Table 1. 2022 Regional and global urban share of territorial and consumption-based $FFCO_2$ emissions. Three urban boundaries are shown (Global Urban Boundaries - GUB; Global Cities and Town Boundaries - GCTB; Functional Urban Area - FUA; hybrid FUA/GUB - FUGUB).

	Territorial (%)				Consumption (%)			
Region	**GUB**	**GCTB**	**FUA**	**FUGUB**	**GUB**	**GCTB**	**FUA**	**FUGUB**
Africa	62.3	50.4	55.7	69.2	55.1	64.0	65.2	72.2
Asia & Developing Pacific	69.8	72.6	72.9	73.0	60.8	67.3	62.4	74.1
Developed Countries	69.3	54.6	59.2	69.1	74.2	79.0	70.9	85.7
Eastrn Eurpe & West-Cntrl Asia	50.6	53.3	44.4	56.0	72.9	89.8	63.7	80.7
Latin America & Caribbean	57.2	64.6	68.4	69.9	68.5	83.1	69.1	73.0
Middle East	43.8	59.0	59.7	63.6	69.4	88.0	78.1	83.3
Global	**65.6**	**64.0**	**63.6**	**69.9**	**65.3**	**75.8**	**65.9**	**79.3**

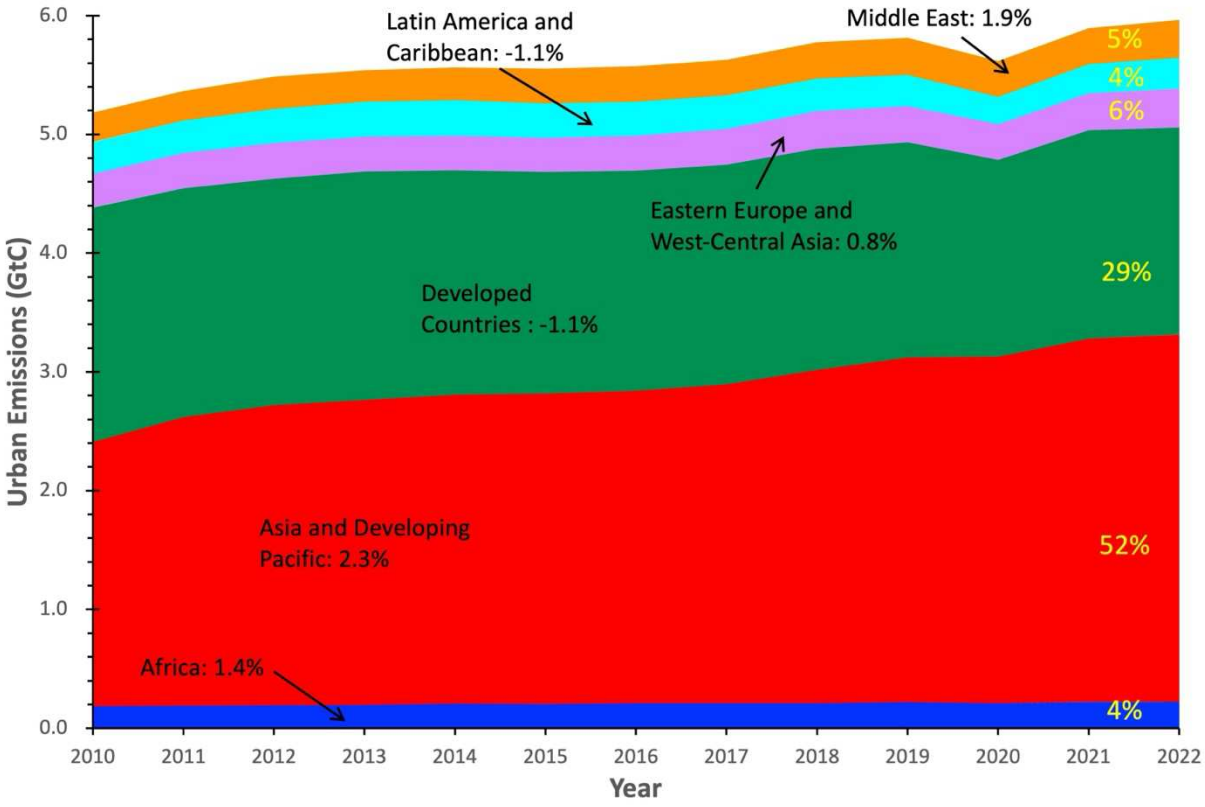

Figure 1. *Vulcan-UrbanML urban $FFCO_2$ emissions (FUGUB boundary) from 2010-2022 annotated with mean urban $FFCO_2$ emission growth rate and regional share of total urban $FFCO_2$ emissions (yellow type). (see SI, Table S1 for tabular values).*

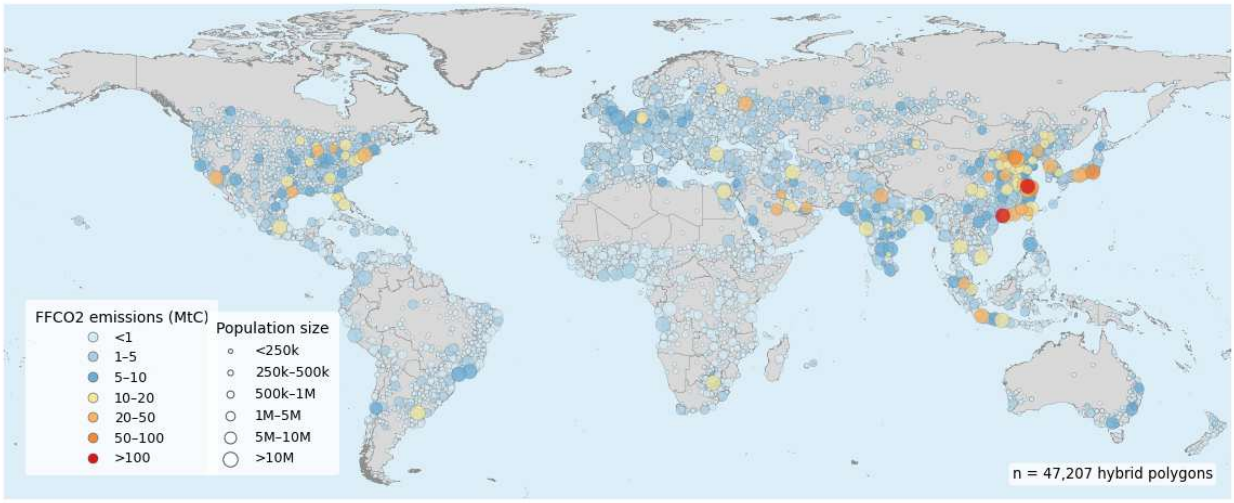

Figure 2. *Urban $FFCO_2$ emissions (MtC) by city and population from the Vulcan-UrbanML model and the FUGUB boundary (N=48,267) for the year 2022.*

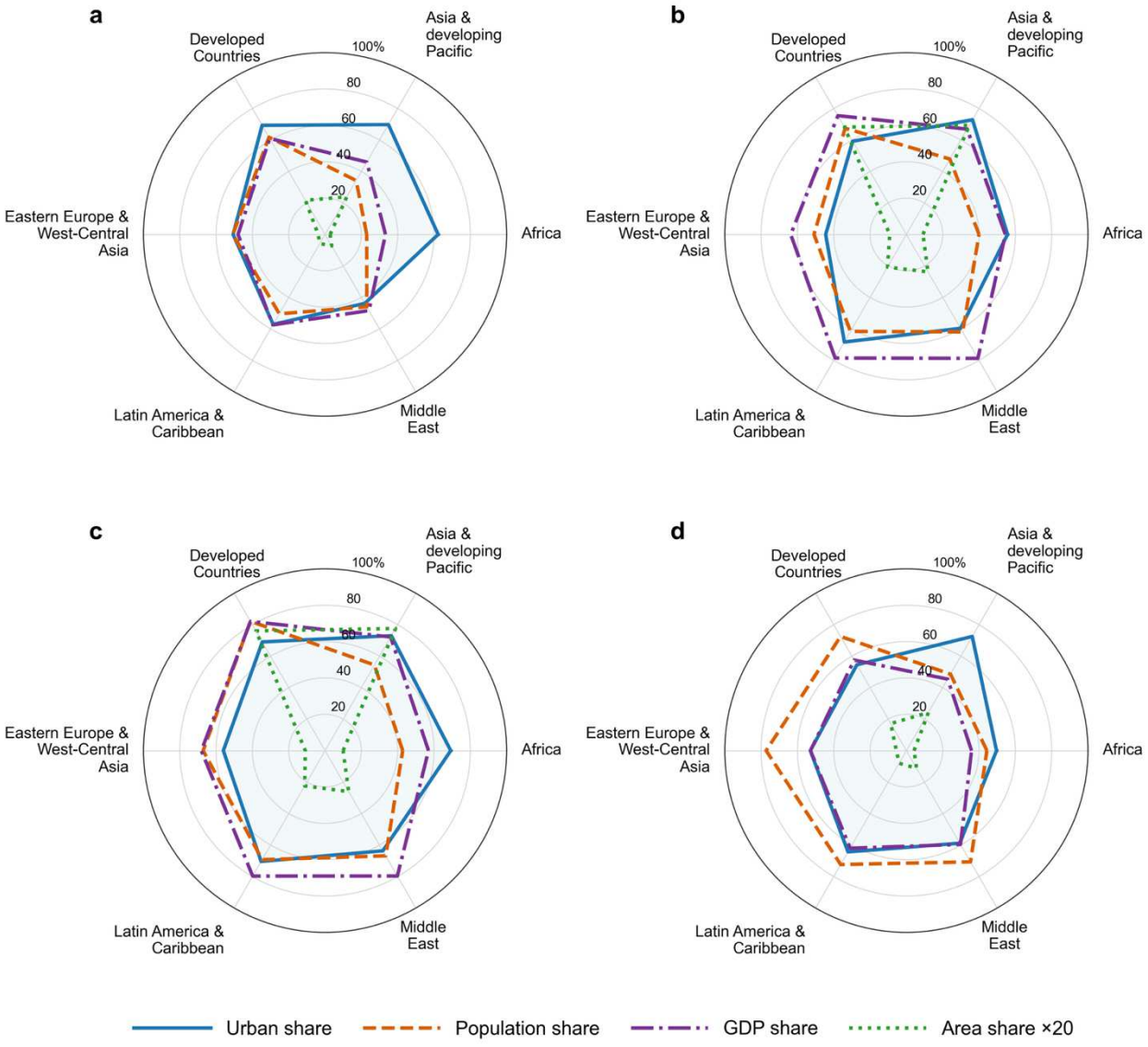

Figure 3. *2022 Regional representation of the $FFCO_2$ urban share, urban population share, GDP share, and the urban land area share for the six IPCC regions and four urban boundaries: a) GUB; b) FUA; c) FUGUB; d) GCTB. See SI, table S3 for tabular information.*

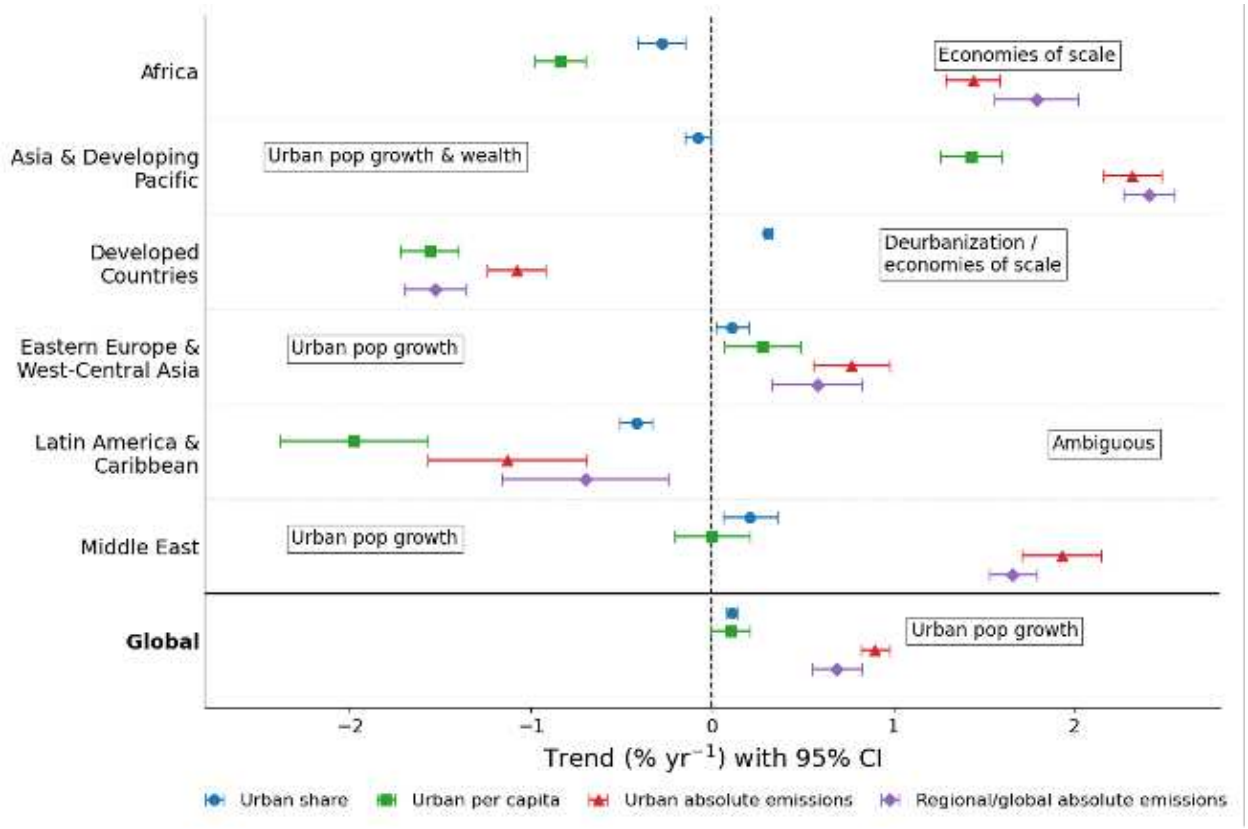

Figure 4. *Trends (and 95% CI) associated with the regional urban $FFCO_2$ emissions share, urban per capita $FFCO_2$ emissions, absolute urban $FFCO_2$ emissions, and absolute regional $FFCO_2$ emissions over the 2010-2022 time period using the FUGUB boundaries. See SI, table S4 for tabular information.*

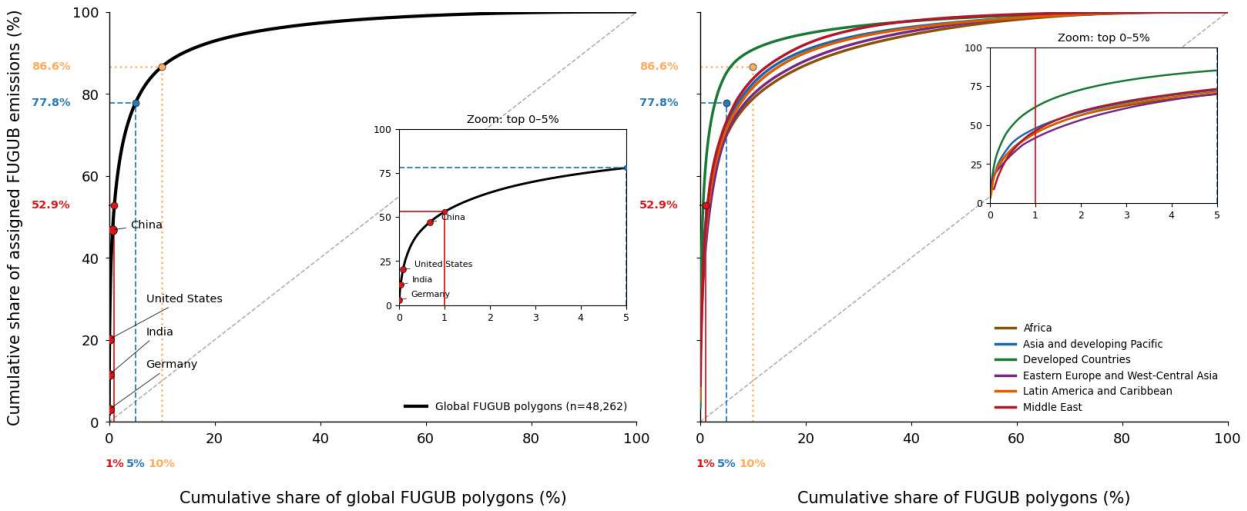

Figure 5. *The relationship between the accumulation of urban $FFCO_2$ emissions (y-axis) and the accumulation of urban entities using the FUGUB boundaries (N=48,267). a) global accumulation with additional information regarding the number of accumulated cities equivalent to the given national total; b) regional accumulation.*

References and Notes

Ahn, D., Goldberg, D. L., Coombes, T., Kleiman, G., & Anenberg, S. C. (2023). CO_2 emissions from C40 cities: Citywide emission inventories and comparisons with global gridded emission datasets. *Environmental Research Letters*. https://doi.org/10.1088/1748-9326/acbb91

Aslam, B., K.R. Gurney (2026). Global City CO2 emissions based on Machine Learning, *in preparation.*

Bai, M., Zhang, X., Ai, W., Jing, X., & Liu, L. (2025). A high-resolution global annual city and town boundaries dataset (2000–2022) derived from GLC_FCS30D product. *Scientific Data*, *13*(1), 50. https://doi.org/10.1038/s41597-025-06368-9

Basu, S., Lehman, S. J., Miller, J. B., Andrews, A. E., Sweeney, C., Gurney, K. R., Xu, X., Southon, J., & Tans, P. P. (2020). Estimating US fossil fuel CO_2 emissions from measurements of^{14} C in atmospheric CO_2. *Proceedings of the National Academy of Sciences*, *117*(24), 13300–13307. https://doi.org/10.1073/pnas.1919032117

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, *13*(2).

Bradley, A. C., McNulty, M. L., Downey, L. C., & de Gouw, J. A. (2025). How historical siting decisions promote modern pollution disparities in Denver, Colorado. *Environmental Sociology*, *0*(0), 1–15. https://doi.org/10.1080/23251042.2025.2534255

Chen, S., Long, H., Chen, B., Feng, K., & Hubacek, K. (2020). Urban carbon footprints across scale: Important considerations for choosing system boundaries. *Applied Energy*, *259*, 114201. https://doi.org/10.1016/j.apenergy.2019.114201

Crippa, M., Guizzardi, D., Pagani, F., Schiavina, M., Melchiorri, M., Pisoni, E., Graziosi, F., Muntean, M., Maes, J., Dijkstra, L., Van Damme, M., Clarisse, L., & Coheur, P. (2024).

Insights into the spatial distribution of global, national, and subnational greenhouse gas emissions in the Emissions Database for Global Atmospheric Research (EDGAR v8.0). *Earth System Science Data*, *16*(6), 2811–2830. https://doi.org/10.5194/essd-16-2811-2024

Crippa, M., Guizzardi, D., Pisoni, E., Solazzo, E., Guion, A., Muntean, M., Florczyk, A., Schiavina, M., Melchiorri, M., & Hutfilter, A. F. (2021). Global anthropogenic emissions in urban areas: Patterns, trends, and challenges. *Environmental Research Letters*, *16*(7), 074033. https://doi.org/10.1088/1748-9326/ac00e2

Dass, P., Gurney, K., Song, Y., Kato, A., Aslam, B., & Gawuc, L. (2025). Advancing High-resolution U.S. onroad fossil fuel carbon dioxide emissions estimates. *Journal of Cleaner Production*, *under review*.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, *29*(5), 1189–1232.

Gurney, K.R. C. Kai, H. Sun, P. Das, A. Kato, B. Aslam (2026) On the urban representation of global gridded greenhouse gas inventories, *in preparation*.

Gurney, K. R., Dass, P., Kato, A., Gawuc, L., Aslam, B., & Sun, H. (2025). Vulcan version 4.0 high-resolution annual carbon dioxide emissions in the US for 2010–2022. *Scientific Data*, *12*(1), 1946.

Gurney, K. R., Kılkış, Ş., Seto, K. C., Lwasa, S., Moran, D., Riahi, K., Keller, M., Rayner, P., & Luqman, M. (2022). Greenhouse gas emissions from global cities under SSP/RCP scenarios, 1990 to 2100. *Global Environmental Change*, *73*, 102478. https://doi.org/10.1016/j.gloenvcha.2022.102478

Gurney, K. R., Liang, J., Roest, G., Song, Y., Mueller, K., & Lauvaux, T. (2021). Under-reporting of greenhouse gas emissions in U.S. cities. *Nature Communications*, *12*(1), 553. https://doi.org/10.1038/s41467-020-20871-0

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*.

Huo, D., Huang, X., Dou, X., Ciais, P., Li, Y., Deng, Z., Wang, Y., Cui, D., Benkhelifa, F., Sun, T., Zhu, B., Roest, G., Gurney, K. R., Ke, P., Guo, R., Lu, C., Lin, X., Lovell, A., Appleby, K., … Liu, Z. (2022). Carbon Monitor Cities near-real-time daily estimates of CO2 emissions from 1500 cities worldwide. *Scientific Data*, *9*(1), 533. https://doi.org/10.1038/s41597-022-01657-z

IPCC. (2021). *Climate Change 2021: The Physical Science Basis*. Cambridge University Press.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. *NeurIPS*.

Kerr, G. H., Meyer, M., Goldberg, D. L., Miller, J., & Anenberg, S. C. (2024). Air pollution impacts from warehousing in the United States uncovered with satellite data. *Nature Communications*, *15*(1), 6006. https://doi.org/10.1038/s41467-024-50000-0

Kona, A., Monforti-Ferrario, F., Bertoldi, P., Baldi, M. G., Kakoulaki, G., Vetters, N., Thiel, C., Melica, G., Lo Vullo, E., Sgobbi, A., Ahlgren, C., & Posnic, B. (2021). Global Covenant of Mayors, a dataset of greenhouse gas emissions for 6200 cities in Europe and the Southern Mediterranean countries. *Earth System Science Data*, *13*(7), 3551–3564. https://doi.org/10.5194/essd-13-3551-2021

Kongboon, R., Gheewala, S. H., & Sampattagul, S. (2022). Greenhouse gas emissions inventory data acquisition and analytics for low carbon cities. *Journal of Cleaner Production*, *343*, 130711. https://doi.org/10.1016/j.jclepro.2022.130711

Lauvaux, T., Gurney, K. R., Miles, N. L., Davis, K. J., Richardson, S. J., Deng, A., Nathan, B. J., Oda, T., Wang, J. A., Hutyra, L., & Turnbull, J. (2020). Policy-Relevant Assessment of Urban CO_2 Emissions. *Environmental Science & Technology*, *54*(16), 10237–10245. https://doi.org/10.1021/acs.est.0c00343

Li, X., Gong, P., Zhou, Y., Wang, J., Bai, Y., Chen, B., & Zhu, Z. (2020). Mapping global urban boundaries from the global artificial impervious area (GAIA) data. *Environmental Research Letters*, *15*(9), 094044.

Liu, B., Shan, Y., Kuik, R., Ji, X., Chapungu, L., Yang, X., & Hubacek, K. (2024). The landscape of city-level GHG emission accounts in Africa. *Journal of Industrial Ecology*, *28*(6), 1377–1391. https://doi.org/10.1111/jiec.13562

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *NeurIPS*.

Lwasa, S., Seto, K. C., Bai, X., Blanco, H., Gurney, K. R., Kilkiş, S., Lucon, O., Murakami, J., Pan, J., Sharifi, A., & Yamagata, Y. (2022). Urban systems and other settlements. In P. R. Shukla, J. Skea, R. Slade, A. A. Khourdajie, R. van Diemen, D. McCollum, M. Pathak, S. Some, P. Vyas, R. Fradera, M. Belkacemi, A. Hasija, G. Lisboa, S. Luz, & J. Malley (Eds.), *Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press. https://doi.org/10.1017/9781009157926.010

Melchiorri, M., Freire, S., Schiavina, M., Florczyk, A., Corbane, C., Maffenini, L., Pesaresi, M., Politis, P., Szabo, F., Ehrlich, D., Tommasi, P., Airaghi, D., Zanchetta, L., & Kemper, T. (2024). The Multi-temporal and Multi-dimensional Global Urban Centre Database to Delineate and Analyse World Cities. *Scientific Data*, *11*(1), 82. https://doi.org/10.1038/s41597-023-02691-1

Moran, D., Kanemoto, K., Jiborn, M., Wood, R., Többen, J., & Seto, K. C. (2018). Carbon footprints of 13 000 cities. *Environmental Research Letters*, *13*(6), 064041. https://doi.org/10.1088/1748-9326/aac72a

Moran, D., Pichler, P.-P., Zheng, H., Muri, H., Klenner, J., Kramel, D., Többen, J., Weisz, H., Wiedmann, T., Wyckmans, A., Strømman, A. H., & Gurney, K. R. (2022). Estimating CO_2 emissions for 108 000 European cities. *Earth System Science Data*, *14*(2), 845–864. https://doi.org/10.5194/essd-14-845-2022

Moreno-Monroy, A. I., Schiavina, M., & Veneri, P. (2021). Metropolitan areas in the world. Delineation and population trends. *Journal of Urban Economics*, *125*, 103242. https://doi.org/10.1016/j.jue.2020.103242

Nangini, C., Peregon, A., Ciais, P., Weddige, U., Vogel, F., Wang, J., Bréon, F.-M. M., Bachra, S., Wang, Y., Gurney, K., Yamagata, Y., Appleby, K., Telahoun, S., Canadell, J. G., Grübler, A., Dhakal, S., & Creutzig, F. (2019). A global dataset of CO2 emissions and ancillary data related to emissions for 343 cities. *Scientific Data*, *6*, 1–29. https://doi.org/10.1038/sdata.2018.280

Seto, K. C., Dhakal, S., Blanco, H., Delgado, G. C., Dewar, D., Huang, L., Inaba, A., Kansal, A., Lwasa, S., McMahon, J., Mueller, D., Murakami, J., Nagendra, H., & Ramaswami, A. (2014). Human settlements, infrastructure and spatial planning. In *Climate Change 2014:*

Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Intergovernmental Panel on Climate Change. http://pure.iiasa.ac.at/id/eprint/11114/

Shi, Q., Ciais, P., Denier Van Der Gon, H. A., Schoenmakers, E. J., Engelen, R., Dou, X., Ke, P., Zhu, B., Liu, Z., Chevallier, F., & Guevara, M. (2025). Comparison of High-Resolution Gridded Emission Maps of Anthropogenic Carbon Dioxide in Europe: GRACED & CAMS-REG. *Environmental Science & Technology*, *59*(10), 4926–4937. https://doi.org/10.1021/acs.est.4c07289

Uchiyama, Y., & Mori, K. (2017). Methods for specifying spatial boundaries of cities in the world: The impacts of delineation methods on city sustainability indices. *Science of The Total Environment*, *592*, 345–356. https://doi.org/10.1016/j.scitotenv.2017.03.014

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- urban.share.supp.Info.submit.pdf