Building sector: Estimating Global, High-resolution Onsite Building Emissions

Paul Markakis[1,2,4], Trey M. Gowdy[1,4], Zion Sheng[3,4], Brittany Lancellotti[1,4], Jordan M. Malof[3,4], and Kyle Bradbury[1,2,4]

1. Nicholas Institute for Energy, Environment & Sustainability, Duke University; 2. Dept. of Electrical & Computer Engineering, Duke University; 3. Dept. of Electrical Engineering and Computer Science, University of Missouri; 4. Climate TRACE

Abstract

Globally, buildings account for approximately 26% of energy sector greenhouse gas (GHG) emissions by end-use and 8% of all direct emissions (International Energy Agency). Yet, the building sector lacks high spatial and temporal resolution emissions estimates which could help to drive action by informing planning at the local and subnational level through emissions inventories. We present a methodology that super-resolves the gridded EDGAR v8 dataset's building emissions estimates from a spatial resolution of approximately 11km-by-11km down to a 1km-by-1km resolution: a level that enables building data to be included in comprehensive municipal-level emissions inventories. We perform this disaggregation of the higher resolution EDGAR v8 emissions data using building floor area as a spatial proxy and subdivide emissions data into residential and non-residential subsectors, both based on satellite-derived data products from the European Commission's Global Human Settlement Layer. Lastly, these data are provided monthly using heating degree days to allocate the time-varying portion of emissions. These data are provided from 2015 through early 2025. We also estimate the potential reduction of building emissions based on the replacement of fossil-based heating and cooking systems with higher-efficiency technologies.

1. Introduction

1.1 Existing emissions inventories are inadequate

Accurate and up-to-date greenhouse gas (GHG) emissions inventories are a valuable tool for climate change mitigation and emissions reduction goal-setting; for example, such inventories underpin the effectiveness of the Paris Agreement (Umemiya and White 2023). However, multiple studies have found the state of emissions inventories at the national and subnational levels to be wanting (Umemiya and White 2023, Luers et al. 2022). Luers et al. went as far as to state that the "existing patchwork of greenhouse-gas inventories is woefully inadequate. From governments to businesses, information on these emissions is inconsistent, incomplete, and unreliable" (Luers et al. 2022). In many cases, emissions inventories at the national level are outdated, with a gap of 10 years not being uncommon (Pearce 2024). In other cases, the reliability of those inventories has come under scrutiny, and in a study evaluating promptness, accuracy, and consistency, 69 countries were deemed to be struggling in one or more of those three categories, representing roughly half of developing countries (Umemiya and White 2023).

1.2 The buildings sector GHG share is substantial and hard to estimate accurately

For a GHG emissions inventory to be valuable, each sector included needs to be accurately represented so that priorities for GHG emissions reduction targets can be accurately set. Additionally, the resolution of the information has to be sufficiently high for establishing actionable emissions mitigation plans. The buildings sector accounts for 8% of direct global emissions and an additional 18% of indirect emissions from the production of electricity and heat used in buildings (International Energy Agency n.d.); a substantial component, globally. This sector's emissions are perniciously difficult to estimate accurately, however, due to the highly distributed nature of building energy consumption and the general lack of ground truth for model calibration.

Some existing GHG inventory approaches include building sector information, but those approaches have limitations around spatial resolution and the availability of subsector information. Existing inventories include IEA's energy consumption data (which can be translated into emissions), the Open-source Data Inventory for Anthropogenic CO2 (ODIAC) (Oda et al. 2016), the Community Emissions Data System (CEDS) (McDuffie et al. 2020), the Emissions Database for Global Atmospheric Research (EDGAR) (Crippa et al. 2024), the Global Carbon Grid (Tong et al. 2018), and the Global Gridded Daily CO2 Emissions Dataset (GRACED) (Dou et al. 2022). The highest spatial resolution of these datasets are EDGAR and GRACED, which are available at a resolution of about 11 km x 11 km (0.1 degree by 0.1 degree), which for context means about 3 data points for all of the Island of Manhattan (as shown in Figure 1): an insufficient resolution for local decision-making, planning, and monitoring. Similarly, there is limited subsector information, with most inventories grouping all of the building subsectors into one category rather than breaking down the estimates into, for example, residential and non-residential buildings.

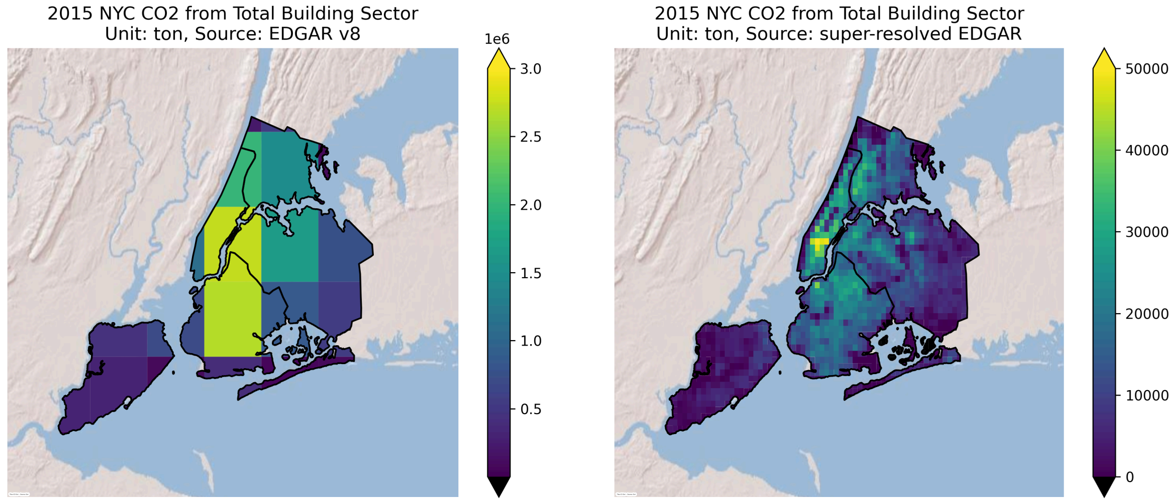

Figure 1. Visualization of EDGAR (left subplot), 0.1 degree by 0.1 degree (~11km-by-11km at the equator) building CO_2 emissions data centered on New York City and the island of Manhattan. In contrast, super-resolved EDGAR (right subplot) is provided at the 30 arc second (~1km-by-1km at the equator) resolution.

To make building sector information more actionable, the ability to identify larger sources of emissions is important. Commercial buildings vary greatly in terms of their energy use intensity (i.e., energy use per unit area), and therefore, emissions. Yet it is not possible to distinguish between residential and non-residential building emissions using existing datasets. Furthermore, hospitals, grocery stores, and restaurants may consume 2 to 8 times as much energy as residential facilities (Energy Star 2023), but current data do not provide sufficient detail for distinguishing between types of non-residential buildings. Therefore, providing higher resolution information separately for residential and non-residential buildings, would make these building sector data more directly actionable.

Direct (a.k.a. Scope 1) emissions from residential and commercial buildings result from onsite fuel use such as natural gas, oil, and kerosene consumption. Residential buildings primarily use these fuels for thermal comfort (space and water heating), cooking, and other equipment and appliances; in commercial buildings, space and water heating generally remain the largest end-use. The magnitude of these end uses is correlated with, for example: building size, climate conditions, and the number of occupants (Gonzalez 2022). However, timely and accessible high-spatial-resolution data on these are limited.

Existing activity-based methods to estimate emissions leverage data typically not accessible at scale, such as information on a building's insulation, appliances in a building, and information on occupant behavior (Kavgic 2010). These approaches may rely on building physics (DOE, 2014) or statistical approaches, but the data requirements for these approaches prevent them from being scaled globally.

Some activity-based approaches use spatial disaggregation techniques to reallocate emissions estimates to finer (i.e., sub-national) scales. Inventories that use these techniques include the Open-source Data Inventory for Anthropogenic CO2 (ODA 2018), the Community Emissions Data System (McDuffie 2020), the Emissions Database for Global Atmospheric Research (EDGAR) (Janssens 2019), the Global Carbon Grid (Tong 2018), and the Global Gridded Daily CO2 Emissions Dataset (GRACED) (Dou 2022). While GRACED data are published near-monthly, most other key data are produced with a year or more of latency. Additionally, the highest resolution of these data is 0.1 decimal degrees or roughly 11 km near the equator. Lastly, few of these models break down emissions into residential and commercial subsectors as well as separate emissions estimates into individual greenhouse gasses. These differences are summarized in Table 1. The lack of spatial resolution in these existing approaches makes it difficult to attribute emissions to building sources with the specificity required to inform prioritization at the local level that would ultimately drive global emissions reduction actions.

Table 1. Summary of building emissions data sources.

Emissions Data Source	Spatial Resolution	Aggregated Residential and Non-Residential	GHGs Covered	Update Frequency
UNFCCC	Country-Level	Disaggregated	CO2, N2O, CH4	Annual
CAIT	Country-Level	Aggregated	CO2, N2O, CH4	Annual
EDGAR	0.1 deg x 0.1 deg	Aggregated	CO2, N2O, CH4	Annual
GRACED	0.1 deg x 0.1 deg	Aggregated	CO2	Monthly

Here, we present a methodology for increasing the spatial and temporal resolution and sectoral detail of existing building emissions data to enhance its actionability. Furthermore, we describe a method for estimating potential building emissions reducing solutions (ERSs) under a high-efficiency retrofit scenario.

2. Materials and Methods

Our approach increases the spatial and temporal resolution and sectoral detail of building emissions data by by: (1) disaggregating each ~11km-by-11km EDGAR v8 grid cell for each GHG (CO_2, CH_4, and N_2O) into ~1km-by-1km grid cells by using an estimate of building energy consumption, (2) further decomposing each grid cell into residential and non-residential emissions using Global Human Settlement Layer data building type categorizations, and (3) allocating the time-varying component of annual GHG emissions to 12 months of the year based on heating degree days derived from Copernicus ERA5 globally gridded temperature data. At the time of publication, EDGAR v8 country-level is available through 2022 and gridded data is available through 2023. We therefore use backed-out emissions factors from the most recently available year for each emissions data type combined with extrapolated building area data to estimate emissions covering periods beginning in January 2024.

In this section, we introduce the datasets used for both building emissions estimation and the high-efficiency energy system retrofits approach, and then describe the steps of each process in detail.

2.1 Datasets employed for building emissions estimation

2.1.1 Energy and emissions datasets

EDGAR v8: EDGAR v8.0 uses spatial data on global human presence and air temperature to allocate country-level residential and non-residential building emissions estimates to a 0.1° by 0.1° resolution global grid. These datasets include the GHSL and data on heating degree days from the ERA5 climate reanalysis (Crippa et al., 2023). In this work, we super-resolve the EDGAR v8 building emissions dataset from the provided resolution of 0.1-degree-by-0.1-degree which near the equator is close to 11km by 11km onto a 30 arc second (~1km-by-1km) grid using the GHSL building area dataset. To accomplish this, for every nonzero EDGAR estimate of CO_2, CH_4, and N_2O, we allocate those emissions proportional to the fraction of building floor area in each 1km-by-1km grid cell region. This ensures the sum of the original EDGAR data is equivalent to the super-resolved dataset at a global and national level, while providing additional resolution to enable municipal-level inventories more effectively. We exclude the EDGAR emissions from biomass for CO_2 as those are accounted for in other sectors within Climate TRACE. However, it is not specified in the EDGAR v8 tool whether the sector-specific gridded data for N_2O and CH_4 contain or exclude biomass.

Energy Use Intensity data from the World Bank: Energy Use Intensity (EUI) is an indicator of energy use relative to building size. EUI data originates from the CURB Tool (World Bank, 2019) which contains data for about 400 cities around the world. The EUI values by end-use and by city are presented as kWh/m^2/year. The data are provided for eight residential and non-residential building types. We include all end uses (including those typically associated with electricity) at this stage of the model, accounting for the removal of electricity consumption later in the model, as electricity emissions are accounted for elsewhere in the Climate TRACE inventory. We labeled the locations using identification numbers from the Geonames database to assign the EUI value to geographic coordinates closest in proximity. Using residential and non-residential building floor area data from the United States, we developed a weighted average of EUIs across building types to develop a single residential and single non-residential estimate for each city (see supplementary materials for more details on this process).

2.1.2 Remote sensing datasets

Building Area: The Global Human Settlement Layer (GHSL) is provided openly by the European Commission's Joint Research Centre (European Commission 2023) and we use this as our source of building-related data. These data, derived from Landsat and Sentinel 2 satellite imagery, offer global coverage of key building-related factors at a desirable spatial resolution. GHSL provides building-related datasets associated with surface areas, volumes, and heights at varying levels of resolution and across several different years. To generate these data, GHSL used machine learning to infer various building-related attributes from multiple distinct sources of satellite imagery collected across the entire surface of the earth. Additional detail on the GHSL data and methodology is available via GHSL (European Commission 2023).

We rely upon the built-up volume area data product, BUILT-V (Pesaresi 2023c), at 30 arc second resolution (~1km-by-1km at the equator) and convert it to total floor area by assuming a 3m average height for a typical story of a building. This allows us to create an estimate of the total floor area contained within the building volume in each region. As the source data are already subdivided into residential and non-residential building categories based on GHSL modeling, we directly use the GHSL categorization of the data to assign building-related emissions to residential and non-residential buildings. GHSL offers select years of data including 2015 and 2020. For the remaining years, we linearly interpolate and extrapolate at a grid-cell level to estimate building floor area for 2016-2019, and 2021-2025. See the supplemental materials for more details about building data preprocessing.

Copernicus ERA5 2m temperature data: The European Commission's Copernicus program provides gridded climate reanalysis data through its suite of ERA5 data products (Copernicus 2024, Hersbach 2023). We use ERA5 2m temperature data as our primary source of daily temperature for computing heating and cooling degree days to allocate emissions monthly. These data provide a highly accurate temperature estimate at an hourly cadence, globally. The data are available on a lat-lon grid with 0.25deg x 0.25deg spatial resolution.

2.1.3 Validation data

True ground truth of building emissions data would be measurements of the emissions emitted from individual buildings or groups of buildings. While there are measurements of *energy* consumed by individual buildings, there are not, to our knowledge, any known measurements of actual emissions from individual buildings or neighborhoods of buildings, and certainly none that we are aware of at a wider scale. At the national level, there are emissions inventories from the United Nations Framework Convention on Climate Change (UNFCCC) (UNFCCC, 2024) and the latest version of the Climate Action Indicators Tool (CAIT) curated by the World Resources Institute (Washington, DC: World Resources Institute, 2024). However, these two data sources both rely upon member nations to respond to surveys related to energy consumption and emissions so we do not treat them as empirical ground truth. At the municipality level, we rely upon region-specific estimates of building sector emissions data at the municipality level from the Data Portal for Cities (DPFC) tool as aggregated by the Global Covenant of Mayors for Climate & Energy (Global Covenant of Mayors for Climate & Energy, 2019). While also not empirically measured ground truth, since these estimates are performed at the municipality-level, we use the DPFC data as a source of local comparison and evaluation for our estimates at a high spatial resolution. These comparisons are further described in Section 2.4.

2.1.4 Emissions reducing solutions

There are numerous examples from homes to large commercial buildings installing the high efficiency building technologies outlined in our emission reduction solutions. This includes a library in Alaska installing a heat pump water heater (National Renewable Energy Laboratory, 2024). In the United States alone, 4.2 million heat pumps were sold in a 12-month span between 2023-2024 (RMI, 2025). The European Commission calls heat pumps “key to enabling the clean energy transition” (European Commission n.d.). The International Energy Agency notes that in some countries (such as Norway), a majority of buildings have heat pumps installed (IEA) and that “China continues to be the largest market for new sales” (International Energy Agency 2022). Electric cookstoves (i.e. stove) are a common cooking appliance, installed in homes and commercial buildings for well over a century (National MagLab, n.d.). Electrification of cooking, such as from gas to electric stoves, is a well-established emission reduction solution - such as in Ecuador, where more than 750,000 households made the transition from gas to electric cooking (in this case induction stoves) between 2015 and 2021 (Gould et al. 2023).

2.2.1 Datasets employed for high-efficiency building energy system retrofits

Energy and emissions datasets: Using EUI data from the CURB Tool (described above), we estimate the share of total building energy demand attributable to space heating, water heating, and cooking, separately for residential (houses and apartments/flats) and non-residential buildings (hotels, retail spaces, offices, hospitals, and warehouses).

We use the World Health Organization Household Energy Database (WHO, n.d.) to estimate the share of traditional cooking technologies that could be retrofitted. Specifically, we use the dataset to calculate the proportion of the population with primary reliance on fossil-based technologies for cooking.

2.2 Models

Activity-based approaches for estimating building emissions generally follow the 2006 Intergovernmental Panel on Climate Change (IPCC) guidelines, which combine information on energy consumption and emissions factors for a given region (IPCC, 2006). Of course, neither energy consumption nor emissions at the individual or neighborhood level are directly observable at scale. What is observable, however, are buildings themselves, and we leverage this to create a three-part activity based model which combines building floor area (in square meters), or A; building energy use intensity (Joules per meter), or EUI; and emissions factors (tonnes of gas per Joule), or EF. If each of these quantities is accurate in a given region, we can estimate tonnes of emissions, E, as $E = (A)(EUI)(EF)$.

Since this work is focused on super-resolving EDGAR's gridded emissions estimates to a higher spatial and temporal resolution, we assume a regional (country-level) value of EUI (World Bank 2016) and assume that regionally (within a ~11km-by-11km EDGAR grid cell), that emissions factors are approximately constant. Therefore the variation of emissions is due to the differences in building energy consumption; that assumption we can use for disaggregating the emissions.

The GHSL data we use for building floor area also provides a classification of building volume into residential and non-residential classes. We use these classifications from the GHSL data and adopt the same definition of residential buildings as the GHSL, which is the European Commission's INSPIRE definition (INSPIRE 2024) of the residential sector:

Residential: "Areas used dominantly for housing of people. The forms of housing vary significantly between, and through, residential areas. These areas include single family housing, multi-family residential, or mobile homes in cities, towns and rural districts if they are not linked to primary production. It permits high density land use and low density uses. This class also includes residential areas mixed with other non-conflicting uses and other residential areas." (INSPIRE)

While an explicit definition of non-residential is not provided, we define it as follows to be mutually exclusive and collectively exhaustive of buildings and assume that the emissions from **non-residential** buildings are any direct building emissions not included within the residential sector and include emissions from commercial and municipal buildings.

One challenge with the definition of residential is that it will encompass some commercial building activity by including mixed-use and other building types that happen to exist within an area of predominantly residential buildings. This will lead to an overestimation of residential building emissions and an underestimation of non-residential emissions, which we do see in our data.

2.3 Methods

2.3.1 Building emissions estimation

The process we use for generating the emissions estimates involves estimating activity data (energy consumption), disaggregating EDGAR emissions to a higher spatial resolution using the activity data, allocating emissions into residential and non-residential categories, extrapolating data to estimate future year emissions, and generating higher temporal estimates of the data by allocating the variable component of emissions using heating degree days.

Our process is used identically to disaggregate both gridded EDGAR data and country-level EDGAR data. We provide monthly emissions estimates using EDGAR gridded data as the source for the primary GHGs: CO_2, N_2O, and CH_4. We are also able to use EDGAR gridded data to provide monthly emissions estimates for the CO_2 biomass contributions (a.k.a., CO2bio). However, we are only able to isolate biomass emissions from N_2O and CH_4 within the EDGAR country-level data, which we instead use as a source for those biomass emissions estimates. We also provide monthly emissions estimates for various air pollutants – NMVOC, CO, NOx, PM10, PM2.5, and SO2 - using the EDGAR country-level data for those air pollutants. Section S.4.3 provides further detail related to inclusion and exclusion of biomass.

Before we begin detailing the steps of this methodology, we will include some definitions that we use for describing the emissions estimation procedure. In Table 2, we introduce the key indices that we use to reference our variables, and in Table 3 introduce the variables themselves.

Table 2. Definition of key indices used in the methodology.

Variable	Definition	Possible values
r	A region in space, as represented by the index of any grid cell in the output emissions dataset (representing a ~1km x 1km region at the equator)	The index of a grid cell in the output emissions dataset
R	The region covered by an EDGAR grid cell (e.g., representing a ~11km x 11km region at the equator)	The index of a grid cell in the EDGAR emissions gridded dataset
c	Building type category	residential (res), non-residential (nres)
g	Greenhouse gas	Carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O)

Table 3. Definition of key variables used in the model.

Variable	Units	Description
$A(r)$	m^2	Building floor area *Source: Global Human Settlement Layer, Building Volume data (BUILD-V)*
$F(r, c)$	*unitless*	Fraction of building type *Source: Global Human Settlement Layer, Building Volume data (BUILD-V)*
$EUI(r, c)$	J/m^2	Energy use intensity *Source: World Bank CURB tool*
$\hat{EF}(r, c, g)$	tonnes/J	Emissions factor
$\hat{EN}(r, c)$	J	Energy
$E^{EDGAR}(R, g)$	tonnes	EDGAR emissions *Source: EDGAR v8*
$\hat{E}(r, c, g)$	tonnes	Super-resolved emissions

Step 1. Calculate the activity data.

Here, we calculate the product of building floor area data (converted from building volume assuming 3m floor heights), $A(r)$, dividing it up into residential and non-residential floor area based on GHSL BUILT-V data classifications, $F(r, c)$, and multiply that product by the EUI value for the given region and building type $EUI(r, c)$, which are unique for each country, computed as the average of EUI values within that country. Additional assumptions and processing for the EUI data can be found in the supplementary materials. This process yields estimates of energy consumption for every ~1km-by-1km region of the world and by building type category (residential, non-residential), $\hat{EN}(r, c)$.

$$\hat{EN}(r, c) = A(r)F(r, c)EUI(r, c)$$

This process is shown in Figure 2A.

Step 2. Disaggregate emissions.

Using the activity data, the EDGAR v8 data, $E^{EDGAR}(R, g)$, are allocated proportionally based on the fraction of energy (activity) in each grid cell. This estimation process is completed for each gas and each category of buildings:

$$\hat{E}(r,c,g) = \left[\frac{\hat{EN}(r,c)}{\sum_{r\in R}\hat{EN}(r,c)}\right] E^{EDGAR}(R,g)$$

This process is shown in Figure 2B using gridded EDGAR data as the source R.

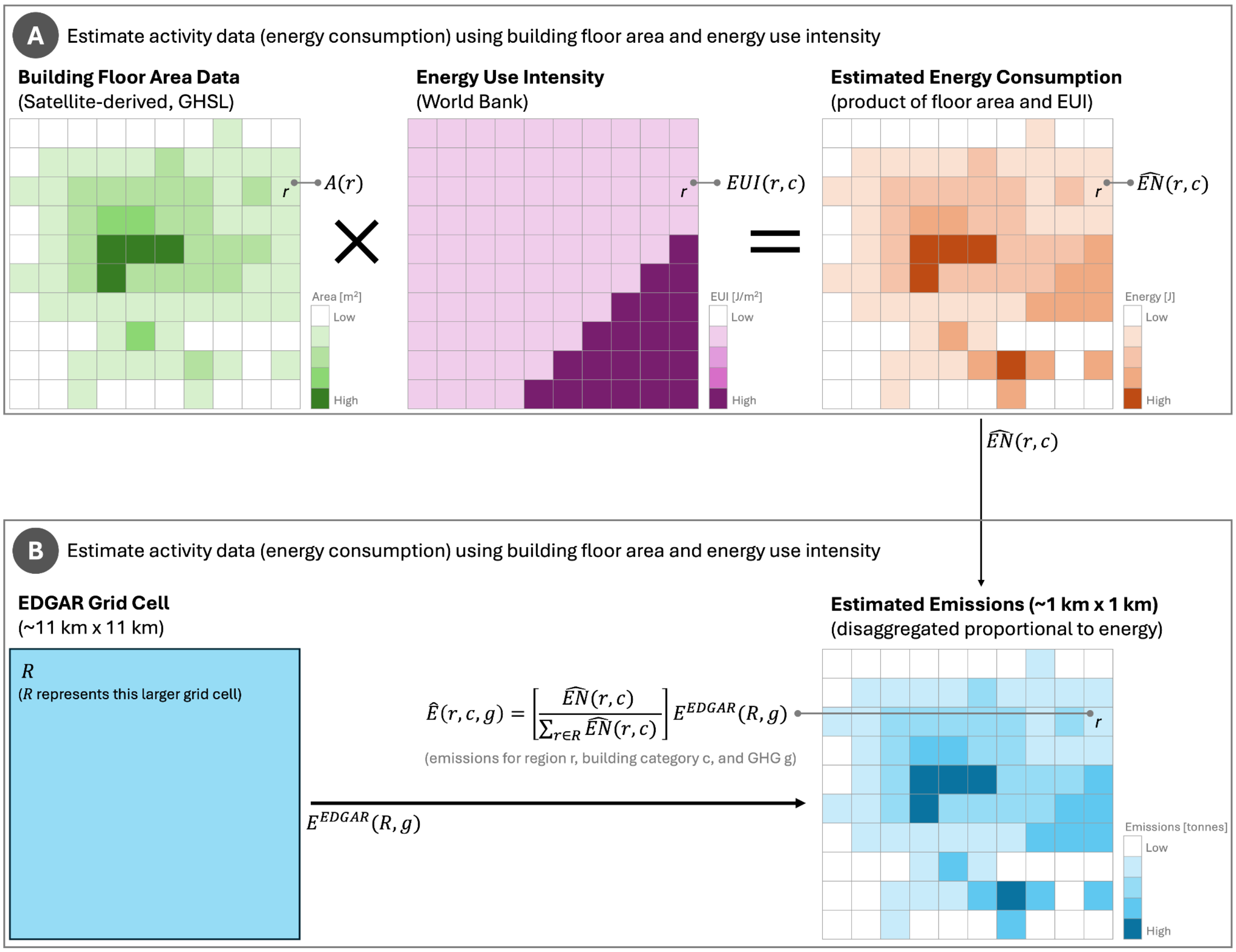

Figure 2. EDGAR disaggregation methodology.

Step 3. Back out emissions factors.

Through steps 1 and 2, we do not need emission factors to calculate the disaggregated emissions. However, EDGAR generally has a latency of more than a year, therefore, we back out emissions factors for each year so that we can use them for extrapolating emissions estimates to future years. Our estimates of emissions factors are the emissions estimates divided by the energy estimates for each building category and GHG:

$$\hat{EF}(r,c,g) = \frac{\hat{E}(r,c,g)}{\hat{EN}(r,c)}$$

Step 4. Extrapolate to future years.

Because GHSL has not yet published building area data more recent than the year 2020, we linearly extrapolate the building floor area data based on the 2015 and 2020 estimates derived from the GHSL. Using the extrapolated building floor area, we can calculate emissions as the product of the extrapolated building floor area, EUI (which we assume is unchanged from the most recently-available year), and the emissions factors in Step 3 which we estimated from the most recently-available building emissions data. For example, to estimate 2024 emissions, we first linearly extrapolate the GHSL data to the year 2024 and then use the EUI and Emissions Factor data from 2023 (the most recently-available building emissions year) to estimate 2024 emissions.

Step 5. Estimate monthly emissions.

As space heating is a major driver of energy consumption in buildings, we make the assumption that heating needs are the primary driver of variable energy consumption in buildings. Globally, variable loads (loads that change seasonally, space heating and cooling) represent around 37.8% of consumption (International Energy Agency 2021). We allocate 37.8% of load based on heating degree days for a given month, $HDD(m)$. Emissions for a given month, which we represent as $\hat{E}(r, c, g, m)$ are then allocated according to the relationship:

$$\hat{E}(r, c, g, m) = \alpha\left[\frac{HDD(m)}{\sum_{m=1}^{12} HDD(m)}\right]\hat{E}(r, c, g) + (1 - \alpha)\left[\frac{\hat{E}(r,c,g)}{4}\right]$$

Here, $\alpha = 0.378$, which is the fraction of variable load. This concept is demonstrated in Figure 3. The grey, baseload emissions are constant throughout the year and are simply $(1 - \alpha)\frac{\hat{E}(r,c,g)}{4}$ each month. The variable load emissions (the fraction represented by α) are distributed proportionally to the quantity of HDDs that occurred that month relative to the total for the year: $\left[\frac{HDD(m)}{\sum_{m=1}^{12} HDD(m)}\right]\hat{E}(r, c, g)$.

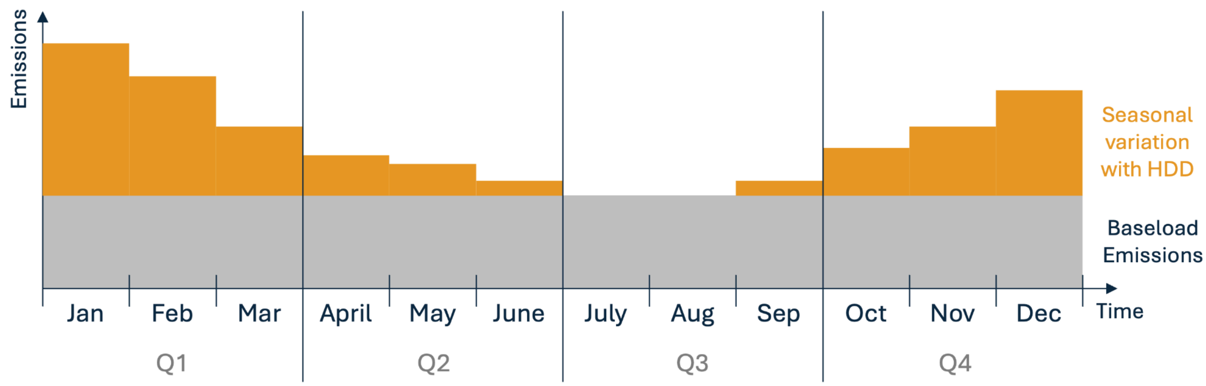

Figure 3. Monthly emissions allocation.

Regions for which there are few HDDs per year may result in unrealistic estimates. For example, if a region only has 1 HDD per year, 37.8% of annual emissions would be shifted in full to the month containing that HDD. To prevent this outcome, we linearly scale up α from 0 to 0.378 based on the total number of HDDs. The value of 0.378 results if there are 1000 HDDs or more for a given region.

2.3.2 ERS - High-efficiency building energy system retrofits

Here, emission reducing solutions (ERSs) for buildings are provided. We estimate the potential reduction in building emissions achievable through high-efficiency energy system retrofits. Because space heating, water heating, and cooking are among the largest sources of direct building emissions, the scenario we model focuses on reducing emissions intensity for these three end uses at the regional level. The retrofit strategies involve replacing fossil-based systems with high-efficiency technologies:

(1) **Space heating**: replacing fossil-fuel (gas or propane) heating systems with high-efficiency heat pumps, improving efficiency from ~75–100% to ~250% (Ciolkosz & Chahal, n.d.).

(2) **Water heating**: replacing fossil-based or electric resistance systems with high-efficiency heat pumps, improving efficiency from ~50–95% to ~175% (Ciolkosz & Chahal, n.d.).

(3) **Cooking**: replacing fossil-fuel cookstoves with electric resistance systems, improving efficiency from ~45–60% to ~75–90% (Ayub et al., 2021).

This retrofit approach is implemented as a percentage reduction in emission factors, representing a decrease in emissions-generating activity. We calculate the resulting fractional decrease in emission factors across space heating, water heating, and cooking for each region as follows, with variable definitions and data sources shown below in Table 4:

$$\alpha(r) = \sum_i s_i(r) \left[m_i(r) \left(\frac{\epsilon_{\text{old},i}}{\epsilon_{\text{new},i}} \right) + (1 - m_i(r)) \right]$$

For each end use, only the portion of technologies eligible for retrofitting (i.e., buildings that do not already have high-efficiency systems) contributes to a reduction in the emission factor. The remaining share retains the existing regional emission factor. Our retrofit scenario models the potential reduction in direct building emissions from switching to higher-efficiency cooking and heating technologies. It does not factor in the potential change in emissions associated with additional electricity demand, in lieu of on-site combustible fuels, that would result from these technology retrofits. On-site combustible fuels (e.g. natural gas, oil, biomass) often have higher emissions factors than grid-supplied electricity, so factoring in fuel switching would likely increase emissions reductions estimates further, but this is dependent on local electricity generation sources. This analysis is planned as a future release.

Table 4. Definitions of variables used in the high-efficiency building energy system retrofits approach.

Symbol	Definition	Notes	Data source(s)
i	Subscript denoting end uses	Examples include space heating, water heating, and cooking	
r	Variable representing regions	Regions will be country-specific	
$s_i(r)$	Share of activity for end use i in region r	These sum to 1 for each region and range from 0 to 1	World Bank CURB Tool 2019
$m_i(r)$	Market share of traditional technologies for end use i (the fraction of end use that could be retrofitted)	These values will range from 0 to 1	Heating: IEA 2024 Water Heating: Rewiring America, RMI, Cooking: World Health Organization Household Energy Database 2022, Eurostat 2022, EIA RECS 2020, Statistics Canada 2006
$\epsilon_{old,i}$	The energy efficiency of the traditional technologies for end use i	The units are output energy per unit of input energy. These values will be greater than 0, but	Space Heating: Penn State Extension n.d.(a), Water Heating: PennStateExtension n.d. (b), Cookstoves: Ayub et. al 2021

Symbol	Definition	Notes	Data source(s)
		could be above 1	
$\epsilon_{new,i}$	The energy efficiency of the high-efficiency technologies for end use i	The units are output energy per unit of input energy. These values will be greater than 0, but could be above 1	Space Heating: Penn State Extension n.d.(a), Water Heating: PennStateExtension n.d. (b), Cookstoves: Ayub et. al 2021
$\alpha(r)$	Fraction of emissions remaining for region r after retrofit	These values will range from 0 to 1	

Our approach makes several assumptions to conduct this approach globally, several of which are intended to make up for a lack of higher resolution regional data:

1. The proportion of activity for each end use that results in direct emissions (space heating, water heating, and cooking appliances) is assumed to be constant across the country. For example, for all of Norway, 14% of building energy demand is attributable to cooking appliances.
2. All building systems that can be retrofitted are retrofitted, so there is a 100% conversion to the more energy-efficient technology.
3. The market penetration of traditional versus high-efficiency technologies is constant, globally.
4. The energy efficiency of traditional technologies can be represented as an average across several common technology types and is constant globally.
5. Several of the technology approaches also involve a switch to electricity as the fuel source. We do not factor in the electrification aspect of that switch; we limit these estimates to the potential reduction in emissions due to the increased efficiency of end-use technologies themselves, regardless of fuel source.
6. The emissions factors associated with the GHGs in a region are impacted equally (same reduction factor for all three GHGs).
7. We do not consider retrofits to biomass-based systems, we only consider retrofitting fossil fuel- and less efficient electric-based technologies for space heating, water heating, and cooking to high-efficiency heat pumps (space heating, water heating) and electric resistance stoves (cooking).

Note: Only rank 1 strategies are provided for assets on the Climate TRACE website and additional strategies will be made available in future releases.

2.4 Evaluating modeled emissions estimates

2.4.1 Building emissions estimates

A major contribution of this work is the ability to provide sufficiently finely-resolved data to be useful not only at the national level, which several existing datasets accomplish, but particularly at the subnational and municipal level, which currently is a gap in the literature. In this section, we describe the process used to conduct evaluation of our super-resolved emissions estimates and original EDGAR data compared with independent municipal-level emissions estimates.

As there is no global sensor-measured empirical ground truth for direct onsite building emissions estimates, we instead compare our modeled emissions estimates to existing inventories at municipality level, where we compare the estimates to 19,998 individual municipality-level estimates from Data Portal for Cities to evaluate the efficacy of the disaggregation process as compared to the original EDGAR data. At the national level, our estimates are identical to those of EDGAR v8.0 since this process super-resolves the EDGAR data.

For our municipality-level analysis, we aggregate our modeled emissions by municipality and compare them against emissions data from DPFC. We first collect the raw emissions data for each municipality in DPFC, and align those emissions with a geographic boundary associated with each municipality. Because DPFC does not include any building emissions data contributed from biofuels, we used a disaggregated version of EDGAR v8.0 gridded data for CO_2 and a disaggregated version of EDGAR country-level fossil fuel only data for N_2O, and CH_4 for municipality-level comparison. Additional detail related to our method for the inclusion or exclusion of emissions from biofuels is included in Section S.4.3.

One challenge is that both the EDGAR v8.0 data and our super-resolved data are gridded data, however, we need to allocate those raster emissions data to vector polygon municipal boundaries. Existing software tools by default use a centroid-based method to align gridded raster data with geographic boundary polygons, wherein if the center-point (a.k.a., centroid) of the grid cell falls within the boundary the grid cell is assigned to that polygon and otherwise the grid cell is excluded. While a centroid-based method is sufficiently accurate to allocate gridded emissions data to country-sized boundaries, municipalities are too small relative to the size of a grid cell to show a sufficient level of accuracy.

While we show results using centroid-based allocation of the raster data to municipalities, we also compare results using an area-based proportional allocation method to allocate emissions from gridded data to municipality boundary polygons; the proportional allocation in this case is based on the percentage difference in boundary area between a grid cell and the component of the municipality's boundary that the grid cell overlaps. For example, if a grid cell has a 10%

overlap with a portion of a municipality's boundary, the relative emissions contribution from that portion of the municipality boundary is inferred to be 10% of that grid cell's emissions. We repeat this process for all overlapping grid cells across the entirety of the municipality, and assign the sum total of all such component-wise emissions as the total emissions attributed to that municipality from the gridded data.

At the country level we compare our super-resolved EDGAR emissions against both UNFCCC and CAIT for "total" buildings (i.e., aggregated residential and non-residential buildings). Figure 4 maps the presence of building-related emissions data at the country level for UNFCCC and CAIT, respectively, in the year 2020, which is the most recent year available for these two data sources.

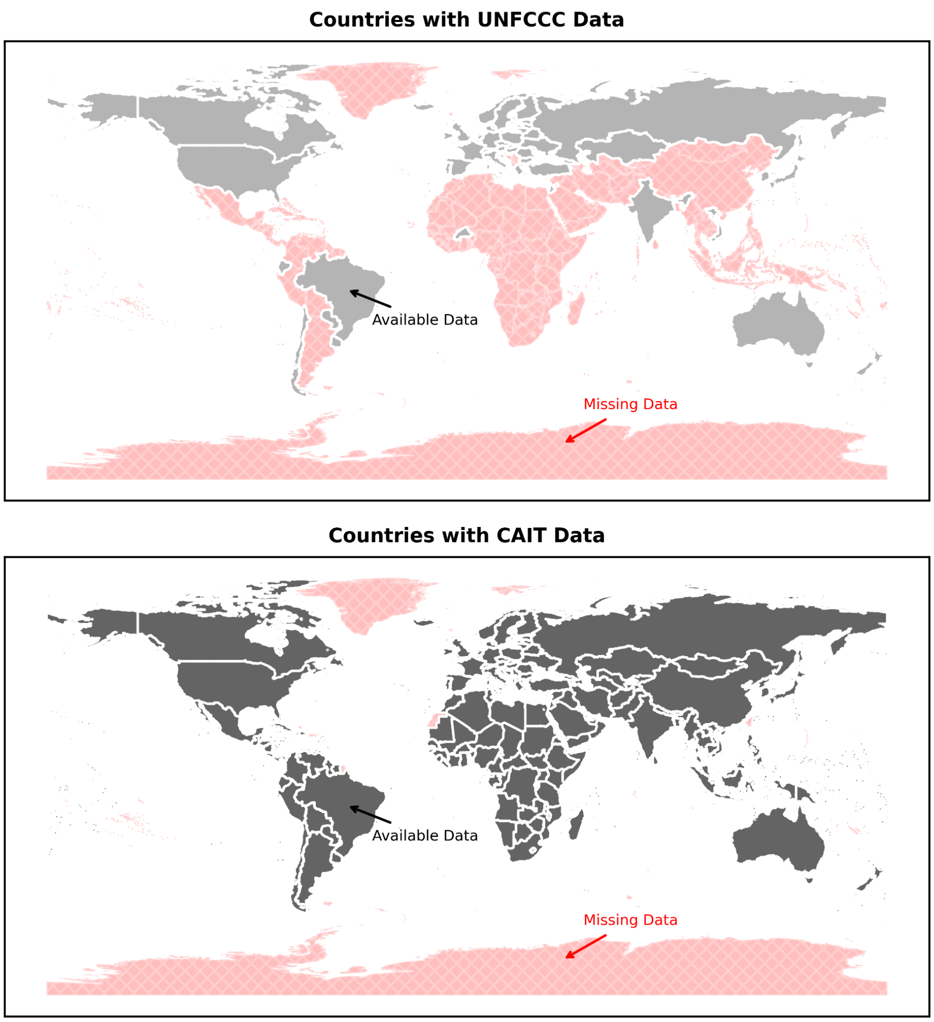

Figure 4. The two maps above show the countries containing 2020 onsite direct building emissions data from UNFCCC (top) and CAIT (bottom). The countries included within each dataset are in gray, whereas the countries missing from each dataset are in red.

2.4.2 High-efficiency building energy system retrofits

We verify the emissions reductions from the retrofits by confirming that all country-level values fall within the range bounded by the lowest and highest plausible reductions given the relative efficiency of the conventional systems and the high-efficiency systems.

3. Results

3.1 Emissions Comparison Results

At the country level, we are able to provide per-country comparisons across all three primary GHGs (CO_2, N_2O, and CH_4) as well as for two different years (2015 and 2020). It is important to note that while the original EDGAR v8.0 and our super-resolved EDGAR data are generally identical at the country-level, there are some smaller countries (e.g., Monaco) for which the original EDGAR data appears missing when aggregated to the country-boundary; remedying this limitation is part of our ongoing work. These comparisons for 2015 and 2020 are shown in Figures 5 and 6, respectively.

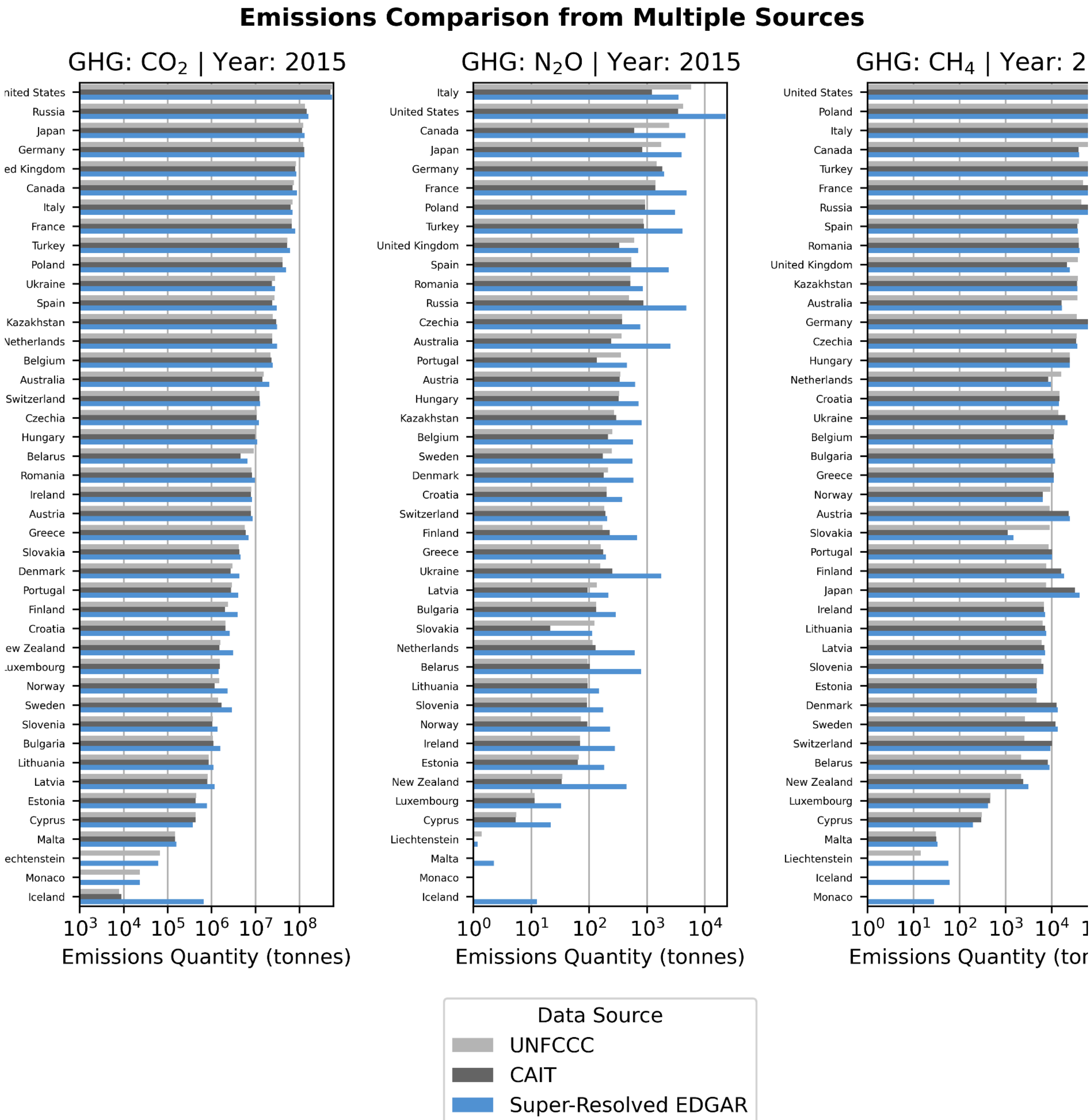

Figure 5. We compare the quantity of direct onsite building emissions (in tonnes) across each of the countries included in the UNFCCC inventory for 2015. We include emissions inventory data from UNFCCC and CAIT for each such country, as well as emissions estimates from our model's super-resolved version of EDGAR. We show this emissions comparison for CO_2, N_2O, and CH_4.

For municipality-level model validation as compared to the DPFC data, we focus on measuring the quality of our modeled emissions estimates using the Weighted Absolute Percentage Error (WAPE) and Mean Absolute Percentage Error (MAPE) metrics. While the WAPE metric weights error in regions that have a greater magnitude of emissions more heavily than in regions with fewer emissions (e.g., percentage error estimating municipality-level emissions in New York City would be weighted more heavily than error estimating municipality-level emissions in a small low-emitting town), MAPE considers error across all comparisons as equally important (and will be subject to small municipalities that are outliers in the data skewing the results). These two metrics each consider error differently, and collectively provide a more comprehensive understanding of estimation model error than by using either metric on its own.

Because DPFC does not contain data related to biofuels, we are unable to evaluate super-resolved EDGAR v8.0 gridded data for CH_4 and N_2O – since both of these datasets contain both fossil fuels and biofuels emissions data – so we limit our municipality-level validation exclusively to a comparison of CO_2 emissions. Our comparison is between DPFC, EDGAR v8.0 gridded emissions data at 0.1-degree-by-0.1-degree spatial resolution, and our super-resolved EDGAR gridded emissions data at 30-arcsecond-by-30-arcsecond spatial resolution. We compare four approaches to estimating municipal-level emissions.

- (1) EDGAR v8.0 + Centroid Allocation: Using original EDGAR v8.0 gridded data with centroid allocation to the municipal level.
- (2) EDGAR v8.0 + Proportional Allocation: Using original EDGAR v8.0 gridded data with proportional allocation to the municipal level.
- (3) Super-Resolved EDGAR + Centroid Allocation: Using our super-resolved emissions estimates with centroid allocation to the municipal level.
- (4) Super-Resolved EDGAR + Proportional Allocation: Using our super-resolved emissions estimates with proportional allocation to the municipal level.

When comparing these four approaches across all DPFC validation municipalities for CO_2 emissions, we observe that the combination of our super-resolved EDGAR data and our proportional allocation method leads to both the lowest WAPE and lowest MAPE. We also observe that for both the lower-spatial-resolution EDGAR v8.0 data and higher-spatial-resolution super-resolved EDGAR data, our proportional allocation method leads to lower WAPE and MAPE for municipality-level emissions estimation as compared to centroid allocation. This summary is shown in Figure 7 below.

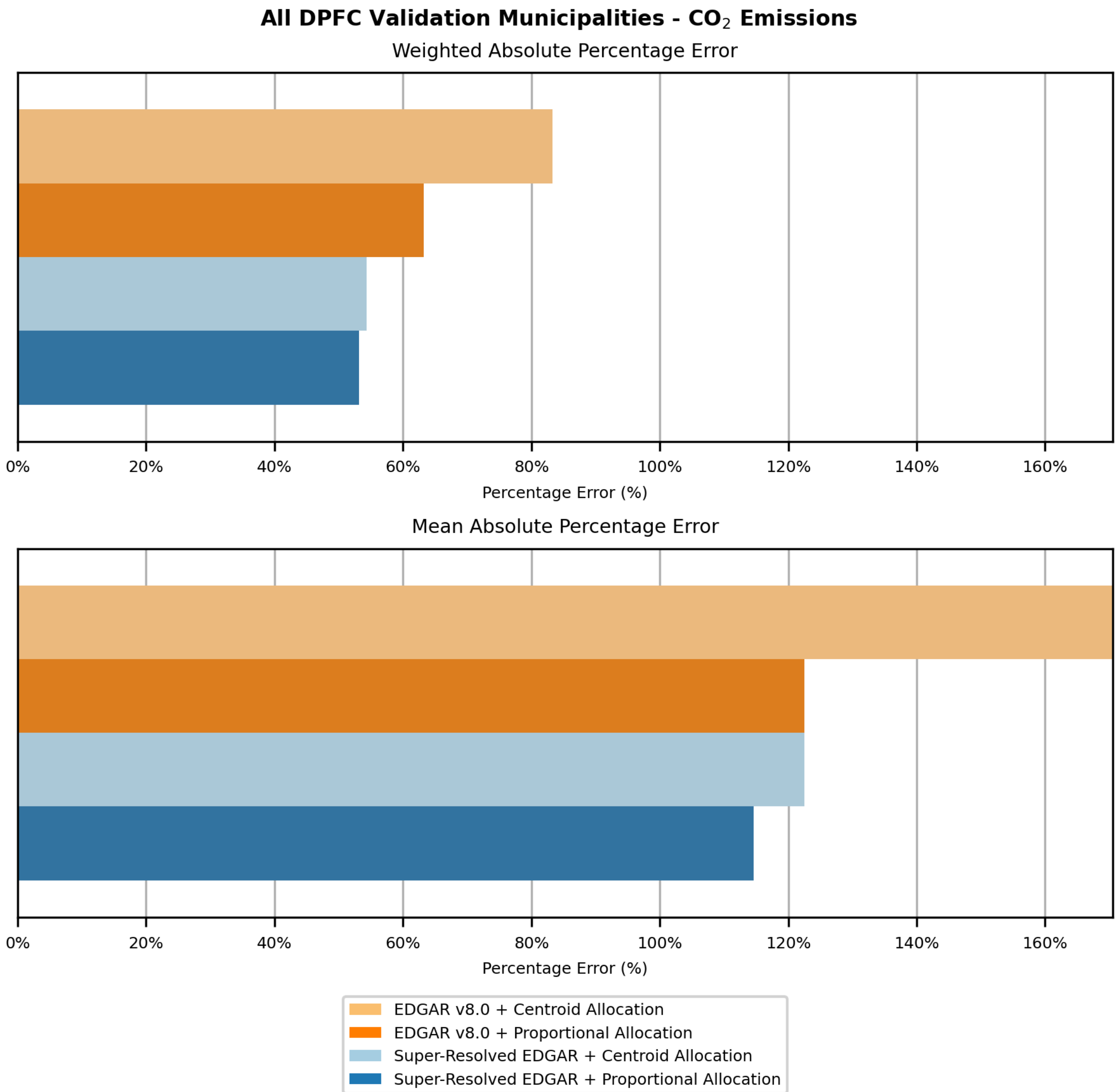

Figure 6. We summarize the Weighted Absolute Percentage Error (WAPE), top, and Mean Absolute Percentage Error (MAPE), bottom, across all municipalities in DPFC for CO_2 *emissions. Our comparison contains both gridded emissions datasets under consideration (EDGAR v8.0 and our Super-Resolved EDGAR) and both gridded raster data allocation methods (Centroid and Proportional).*

3.2 Visualization of emissions

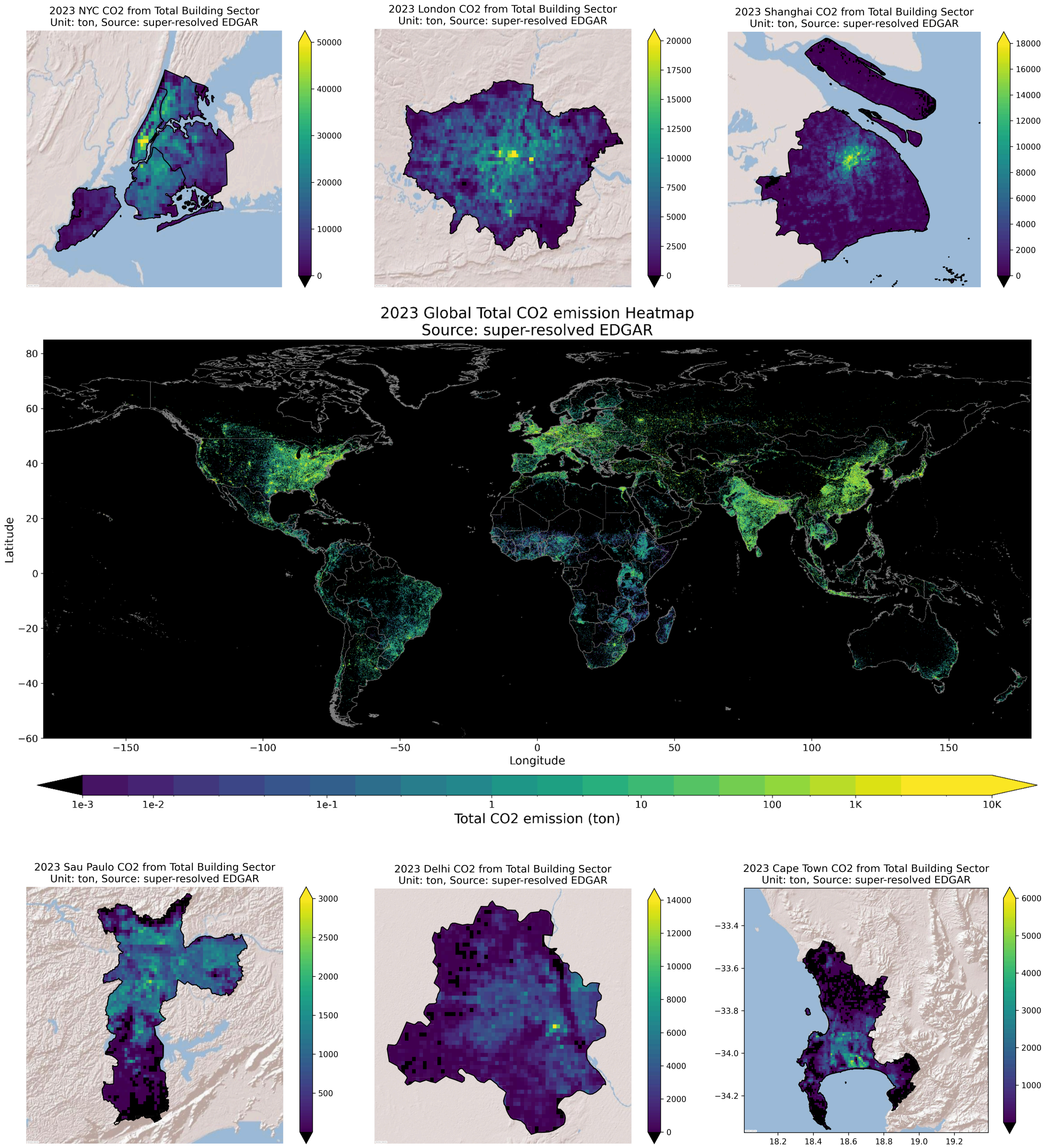

Figure 7. Global and city-level CO_2 emission heatmap. The value of each grid cell is the estimated CO_2 emission from the total building area in that 1km by 1km region in 2023 given by super-resolved EDGAR (unit: metric tonnes). As a shorthand, we use "Total CO_2" in the plots to denote the whole-year CO_2 emission from the total buildings (the sum of residential and non-residential buildings). From the top left to bottom right, selected cities are NYC, London, Shanghai, Sao Paulo, Delhi, and Cape Town. Places with zero building area (zero building-generated CO2 emission) are colored in black.

The approach presented here, which super-resolves EDGAR using high resolution building floor area data, provides global greenhouse gas emissions estimates at a higher spatial resolution (~1km by 1km) than any previous inventory for building emissions. The data shown in Figure 7 include both a global map of total building emissions and city-level data for six major cities across the globe.

The global plots show, as expected, the density of building emissions in some of the most populated regions of the world including the eastern United States, Europe, China, and India. What is new and different about this dataset is the ability to show differences within cities themselves. Emissions from Manhattan stick out clearly in New York, while the coastal regions of Cape Town clearly show a higher concentration of emissions. This level of spatial granularity enables prioritization in local and regional planning and decision-making when paired with other emissions sectors.

3.3 ERS results of the high-efficiency retrofit analysis

Informed by national-level residential and non-residential building sector fuel mixes (United Nations Energy Statistics Database n.d.), we identified substantial variation across countries in fuels used directly within buildings and across sectors in the same country. We developed strategies to reduce onsite combustion of fossil fuels towards electrification of three key building end uses (space heating, water heating, cooking). Our retrofit analysis indicates that country-level potential emissions reductions range from 22-68% for residential buildings and from 12-59% for non-residential buildings (Figure 8). In our framework, the differences between residential and non-residential reduction potentials arise solely from the distribution of energy shares (s_i), since all other variables are held constant between building types. Slightly higher reduction potential for residential buildings is likely a result of differences in how energy is allocated across end uses. For both building types, countries where space heating accounts for less than 1% of fossil-based total energy demand (e.g., Laos, Mali, Niger, Guinea-Bissau) exhibit the lowest reduction potential, since emissions reductions from space heating are effectively zero. By contrast, countries such as Ukraine and Romania exhibit among the highest reduction potentials, with 81% and 90% of their population, respectively, relying on fossil-based cooking technologies (Figure 8). This represents a substantial opportunity for emissions reductions through high-efficiency cooking technology retrofits. High-income countries, such as the U.S., Canada, U.K., Japan, and Australia, exhibit moderately high potential emissions reductions, averaging 49% for residential buildings and 53% for non-residential buildings, despite differences in climate across these countries (Figure 8).

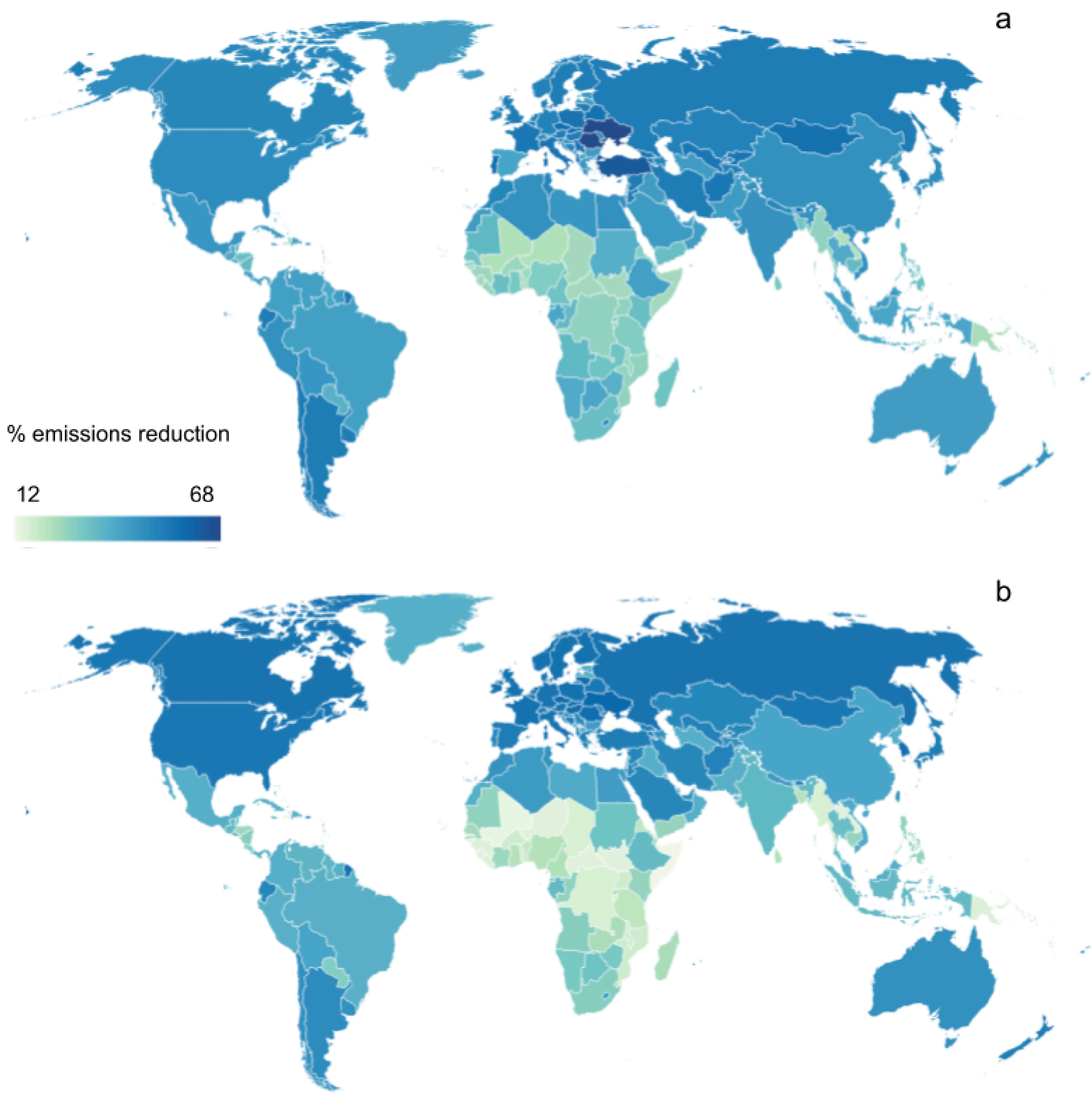

Figure 8. Global maps showing country-level potential emissions reductions from (a) residential buildings and (b) non-residential buildings resulting from high-efficiency retrofits for space heating, water heating and cooking technologies. Buildings employing specific strategies that have been aggregated to the country-level in this figure don't necessarily match what is employed on the Climate TRACE website.

4. Discussion

4.1 Contributions

Our work introduces a model that provides high-spatial-resolution global emissions for CO_2, N_2O, and CH_4. These emissions from our model are further disaggregated into residential and non-residential building subsectors and each year's emissions are reflected monthly. Using activity data to spatially super-resolve the data from EDGAR v8 into 30-arcsec grid cells (roughly 1km x 1km grid cells) not only provides a two-order-of-magnitude increase in the information content for each 0.1 deg x 0.1 deg (roughly 11km x 11km) grid cell from the original EDGAR data, but it proportionally allocates the EDGAR emissions based on the underlying activity data most representative of where emissions are concentrated within each region. This work is also the first attempt to provide gridded data disaggregated into residential and non-residential subsectors. Finally, using temperature data, this work provides gridded data monthly enabling the EDGAR emissions to be allocated temporally based on changing space heating needs.

While the results at the national level are the same as EDGAR v8 at the national level, there were municipal-level improvements in CO_2 estimates for the majority of countries for which municipal data are available. The differences—and in many cases, reduced error—from our super-resolved version of EDGAR are observable in the measurement of WAPE, MAPE, and RMSE at the municipality level compared to DPFC data.

4.2 Limitations

Comparison rather than validation. Although there are clear benefits to our model's temporal and spatial resolution improvements as compared to existing models, a significant challenge is the limited data available for validation to quantify model performance. In fact, there is no true ground truth data for building emissions, so we rely on comparisons to other inventories at the national and municipal level. UNFCCC and CAIT emissions data are both only available with annual values at a country level; these are insufficient to fully measure performance at subnational levels. While DPFC data are available at the municipality-level and thus can be used for sub-country-level analysis, they are only included in the source data for a subset of municipalities represented by eight countries in the world and only exist for a single year and annual set of emissions.

Monthly emissions estimates. We currently do not have data available for validating or comparing monthly emissions estimates.

Biomass exclusion is unclear in source and comparison data. Across inventories used for comparison and for CH_4 and N_2O for the gridded EDGAR data, it is unclear whether biomass is aggregated with fossil fuel emissions in these inventories. This inconsistency complicates accurate emissions accounting and comparison, especially for N_2O and CH_4.

Residential and non-residential definitions don't perfectly align with other inventories. By adopting the definitions of residential regions including the GHSL data, which categorizes all regions whose primary purpose is the housing of people as residential (this includes all mixed-use regions), more regions and buildings will be categorized as residential. This means the estimates for residential emissions will be a bit higher and non-residential a bit lower, than if alternative definitions of residential and non-residential buildings are used.

Each of these limitations has led to areas of active research to overcome these challenges in future iterations of this work.

4.3 High-efficiency retrofit analysis

Our ERS for high-efficiency building retrofit analysis provides insight into the potential emissions reductions from global adoption of higher-efficiency technologies targeting the most energy-intensive building activities. While these retrofits are infeasible to occur all at once, due to the long lifespans of building heating and cooling systems, our analysis reveals the reductions that could eventually be achieved as older systems are being replaced and/or with efforts to accelerate retrofits (e.g. increased financial incentives or policies). On average, we find that a country could reduce emissions from residential buildings by 42% and from non-residential buildings by 38%. These reductions are substantial and could contribute significantly to climate mitigation efforts.

While our analysis highlights substantial emissions reductions that could be achieved through building energy system retrofits from appliance-level energy efficiency improvements alone, there are several challenges that may limit adoption. For example, high upfront costs for retrofits may slow adoption despite long-term energy savings. Subsidies or financing mechanisms could accelerate uptake. Additionally, there is uncertainty around the actual efficiency gains achievable in diverse building types or countries, which may affect the realized reduction potential.

We estimate the change in direct building emissions that would result from heating and cooking technology retrofits. Replacing fossil-based space or water heating systems with heat pumps or switching to electric cooking technologies increases electricity demand, but our estimated emissions reductions do not account for this increased electricity demand. The emissions associated with operating a heat pump depend on the emissions intensity of the electricity used to power it. Even with the current electricity generation mix, the switch to heat pumps for space heating markedly reduces GHG emissions in all major heating markets (IEA, 2022). Furthermore, under a moderately decarbonized grid scenario, the switch to heat pumps is projected to yield substantial reductions in space heating emissions in the U.S., even within the first year after installation (RMI, 2023). Since they do not directly emit air pollutants (Reeg et al., 2025), heat pump retrofits can also reduce indoor and outdoor air pollutant emissions (e.g., nitrogen oxides, carbon monoxide, and particulate matter), representing a significant public health benefit. For example, heat pump retrofits have been projected to significantly reduce emissions in urban areas, like Rome, Italy (Carella & D'Orazio, 2021). There are also indoor air quality benefits for switching to electric stoves from fossil fuel options, resulting in improved public health outcomes related to reduced nitrogen dioxide exposure (Kashtan et al. 2024) and reduced methane and nitrous oxide leaks (Lebel et al. 2022).

Our high-efficiency retrofit scenario is grounded in real-world adoption trends, such as the marked increase in global heat pump sales from 2020-2022 (IEA, 2025). Although heat pump sales have decreased since 2023, likely due to household budget constraints resulting from higher borrowing costs and inflation (IEA, 2025), this trend demonstrates both the practicality of our retrofit scenario and the economic and policy challenges that may shape its implementation.

5. Conclusion

This work introduced a methodology that builds on EDGAR v8 by using high-resolution building data derived from satellite imagery to create the highest resolution estimates of building emissions presently available. This work also provides subsector disaggregation for residential and non-residential buildings and monthly emissions estimates. The data are compared at the national level to other global emissions inventories and, for the first time, to municipal-level estimates.

The data at the national level is equivalent to EDGAR v8 since the super-resolution process disaggregates the gridded EDGAR v8 data. The super-resolved EDGAR data, however, is closer to municipal emissions estimates from Data Portal for Cities emissions values for 5 of 8 countries included in the data across the major metrics of comparison for CO_2 data, with one additional country being approximately the same as EDGAR v8 for a comparison with DPFC. While there are greater disparities for CH_4 and N_2O, this is likely because EDGAR v8 contains aggregated fossil fuels and biofuels emissions estimates for CH_4 and N_2O, whereas DPFC only contains data for fossil fuels; because biofuels have disproportionately larger emissions factors for CH_4 and N_2O as compared to fossil fuel emissions factors, the discrepancy between these data sources is the likely cause of our performance for CH_4 and N_2O. Ongoing research is focused on overcoming this challenge.

These high-resolution data are the first to enable municipal-level building emissions inventories at a global scale. These data can also enable the development of a building emissions inventory for any arbitrarily defined region. We hope that these data can be an additional tool to guide planning and prioritization for climate change mitigation strategies at several organizational levels.

We estimated potential reductions in building emissions from high-efficiency energy system retrofits. We did this by quantifying how much space heating, water heating, and cooking activity could be lowered through replacing less efficient technologies with more efficient alternatives. We found that these retrofits could result in marked emissions reductions (on average 38% and 42% for non-residential and residential buildings, respectively). Though higher-efficiency technologies are being adopted, our modeled emissions reductions would realistically take several years, as these technologies have long lifespans. Their adoption also depends on current policies, economics, technology awareness, and likely other building-specific considerations.

6. Acknowledgments

The authors would like to WattTime and the Climate TRACE coalition for their organizational support, and Climate TRACE funders – Al Gore, Generation Investment Management, Google.org, and the Patrick J. McGovern Foundation.

We also thank the following students who have assisted with this project - without their help this work would not have been possible: Pablo Salazar Armella, Lawson Garner, Ya-Yun Huang, Julia Kourelakos, Pragya Raghuvanshi, Nicholas Sommer, Nicholas Sorokin, Weilin Wang, Yuanjing Zhu.

S. Supplementary materials

S.1. Building floor area

We also identified that the GHSL volume data, when converted to total floor area of buildings (by assuming an average floor height of 3m per story) and combined with energy use intensity to compute energy consumption data, tended to overestimate energy consumption at the national level, so a scaling factor of 0.12 was applied to the building floor area estimates based on a regression of our derived energy consumption estimates and national level energy consumption from IEA data. Since our goal is to estimate emissions, not energy consumption directly, what is most important is that the estimates of energy consumption are proportional to actual energy consumption and globally consistent.

S.2. Energy Use Intensity

The data sources used by the World Bank CURB Tool for these values are the 'DOE Model' and EDGE,' model year(s) unknown. The CURB Tool was released in 2016. Using building stock information from the United States Energy Information Administration's Commercial Building Energy Consumption Survey (CBECS) (U.S. Energy Information Administration 2023a) and Residential Energy Consumption Survey (RECS) (U.S. Energy Information Administration 2023b), we developed a proportional mix of values for residential vs. non-residential building categories to create a weighted average of building EUI totals. We then developed weighted averages across socioeconomic tiers for the residential sector (homes - houses and apartments/flats) to create a sole residential sector value for each location.

Table S.1. Building Floor Area Weighted Average Data Preparation. This table describes the weighted averages derived from the U.S. Energy Information Administration (EIA) data to allocate floor area to the building types found in the World Bank CURB Tool for energy use intensity values. The EIA values originate from the Residential Energy Consumption Survey (RECS) and Commercial Building Energy Consumption Survey (CBECS).

Building Category	**Percentages**	**Specific Building Type**	**Sq. Footage (US) - To inform percentage (Millions SF)**	**Percentages (based on US data)**	**Notes/Source**
Residential	71.12%	Home	237,400.00	71.12%	RECS 2015 HC10.1 - includes all Single Family Homes, Apartment Units, and Mobile Homes.
Non-residential	28.88%	Hospital	4,018.00	1.20%	CBECS 2018 B12 - categories based recategorization of CBECS data categories to align with CURB Tool building types
		Hotel	6,976.00	2.09%	
		Office	18,200.00	5.45%	
		Retail	19,412.00	5.82%	
		Warehouse	17,483.00	5.24%	
		Other	30,333.00	9.09%	

There are a few limitations with this process. The EUI values represent only a single point in time (the exact year for each location is unknown) and the values are associated primarily with major population centers only (potentially reducing accuracy for rural and suburban areas). Building-use percentages are based on a US average, and then applied globally. Identifying more localized detailed building use percentage estimates that can be applied globally will improve accuracy. Further refinement of EUI values by recency and spatial resolution (more locations overall, including rural and suburban areas) provide opportunity for improvement to this model. These data are then converted from kilowatt hours (kWh)/square meter to megajoules (MJ)/square meter.

Prior to that conversion, we clip outlying data points. We clip all values to be between 50 and 400 kWh/m^2/year to eliminate potential outliers, before then converting them MJ/m^2/year.

S.4. Known issues

S.4.1. Energy Use Intensity

To address challenges with the CURB Tool data, Agartala (India) was removed since there are no residential sector values in the original data, duplicate values for Juba (labeled "Sudan" for "South Sudan") from CURB tool dataset were removed, and Panama City's (Panama) data was removed as there was a likely error with the hotel EUI datapoint that impacted the non-residential data totals.

S.4.2. Final Estimates

In the Democratic Republic of the Congo, we are aware of possible missing data from the EDGAR data that impacts our estimates and may lead to underestimation in this location in the EDGAR super-resolved version of the model.

S.4.3. Inclusion/Exclusion of Biomass

Biofuels (e.g., wood) are a fuel type used across the building energy sector and – particularly for N_2O and CH_4 – has a disproportionately large emissions factor as compared to other fuel types. However, biofuels are not consistently included (or excluded) by default in various inventories and emissions estimation models. IPCC standards (IPCC, 2019) dictate that for accounting purposes, CO_2 biofuel-related emissions from harvested wood biomass used directly as energy feedstocks, to include harvested wood biomass burnt directly as fuel wood on a residential, commercial, or industrial scale, as part of the Agriculture, Forestry, and Other Land Use (AFOLU) sector (IPCC, 2019) whereas N_2O and CH_4 biofuel-related emissions are accounted for within the Energy sector alongside fossil fuel-related emissions. As some sources (e.g., UNFCCC and CAIT) follow this IPCC convention in a non-separable way, we cannot choose to entirely exclude biofuels by default – nor can we add it to sources such as DPFC that contain no biofuels-related data. We thus use the following variations of our modeling approach to overcome differences across comparison data related to the inclusion or exclusion of biomass.

Where possible, we use the EDGAR v8.0 gridded data (with 0.1degree-by-0.1degree spatial resolution) as our source of EDGAR v8.0 data against which we apply our super-resolution method. However, to isolate the emissions contributions from onsite biofuels for N_2O and CH_4 we instead apply our super-resolution method to the EDGAR v8.0 country-level data; we thereafter compute the difference between the super-resolved gridded data reflecting all onsite combusted fuels and super-resolved country-level data reflecting onsite biofuels to reflect the emissions contributions from onsite fossil fuels. These data sources are shown in Table S.2.

Table S.2. Sources of EDGAR v8.0 data, by GHG and fuel type.

GHG	***All Onsite Combusted Fuels***	***Onsite Fossil Fuels***	***Onsite Biofuels***
CO_2	EDGAR v8.0 Gridded Data (0.1degree-by-0.1degree)	EDGAR v8.0 Gridded Data (0.1degree-by-0.1degree)	EDGAR v8.0 Gridded Data (0.1degree-by-0.1degree)
N_2O	EDGAR v8.0 Gridded Data (0.1degree-by-0.1degree)	**Computed*	EDGAR v8.0 Country Data
CH_4	EDGAR v8.0 Gridded Data (0.1degree-by-0.1degree)	**Computed*	EDGAR v8.0 Country Data

**Computed as the difference between super-resolved gridded data for all onsite combusted fuels and super-resolved country-level data for onsite country-level biofuels.*

S.4.4. Verifying modeled emissions estimates

EDGAR's Energy for Buildings data includes not only residential and commercial buildings. As EDGAR's Energy for Buildings data uses data from the IPCC's 1996 (1A4) and 2006 codes (1A4+1A5) on 'Other Sectors,' this includes agricultural off-road vehicles (1.A.4.c.ii), fishing (1.A.4.c.iii), and marine vessels (1.A.5.b) that do not exactly align with our data estimates. From our review, these are small values that have a minor impact on estimates, but are key to note as we compare our model to EDGAR for validation.

S.4.5. Municipal-Level Validation Boundary Summary

For municipal-level validation, we discovered that there is not a consistent set of geospatial boundaries from a common source at a corresponding municipality level. GADM is an open-source project that "wants to map the administrative areas of all countries, at all levels of sub-division" (GADM, 2024) and our primary source for country-level boundaries represented by GADM Level 0. While GADM tends to have near-complete global coverage within their data at the Level 0 administrative subdivision representing each nation, their data was not globally comprehensive at the municipality level. For example, the GADM Level 2 boundaries for the United States were not consistent with the size of each municipality contained within the United States DPFC data, and GADM does not contain municipality-level boundaries for the United States at a level corresponding to a more detailed subdivision.

To conduct our model validation at the municipal level, we identified distinct geospatial boundary files for each municipal region under consideration using authoritative data sources for each region. We found that GADM boundaries were sufficient for Chile, Canada, Costa Rica, and Denmark. For Mexico, Japan, and India, we identified corresponding municipality-level boundaries from the Humanitarian Data Exchange (The Humanitarian Data Exchange, 2024). In the United States, we were required to identify a boundary file independently for each of the 50 states in the country from authoritative state government websites. Table S.3 below summarizes the source for each such set of boundaries.

Table S.3. Boundary source data used for municipal comparison by country.

Country	Level for municipalities	Boundary Source
Mexico	Level 2	Humanitarian Data Exchange
Japan	Level 2	Humanitarian Data Exchange
Chile	Level 3	GADM
Canada	Level 3	GADM
Costa Rica	Level 2	GADM
India	Level 5	Humanitarian Data Exchange
Denmark	Level 2	GADM
United States	N/A	State Government Boundaries

Table S.4. Strategy ERS Table. Note: Only rank 1 strategies are provided for assets on the Climate TRACE website and additional strategies will be made available in future releases.

native_strategy_id	Definitions
strategy_name	Technology retrofit
strategy_description	Retrofit traditional building systems with high-efficiency technologies for space heating, water heating, and food preparation.
mechanism	Retrofit
asset_type_new	NA
max_activity_affected_absolute	NA
max_activity_affected_ratio	1
co2_emissions_factor_new_absolute	NA
co2_emissions_factor_new_to_old_ratio	Varies by country from 0.32-0.78 for residential buildings and from 0.41-0.88 for non-residential buildings.
ch4_emissions_factor_new_absolute	NA
ch4_emissions_factor_new_to_old_ratio	Varies by country from 0.32-0.78 for residential buildings and from 0.41-0.88 for non-residential buildings.
n2o_emissions_factor_new_absolute	NA
n2o_emissions_factor_new_to_old_ratio	Varies by country from 0.32-0.78 for residential buildings and from 0.41-0.88 for non-residential buildings.
confidence	Medium
exponential_decay_emissions_factor	NA

native_strategy_id	Definitions
exponential_decay_activity	NA
induced_sector_1	NA
induced_sector_1_activity_conversion_rate	NA
induced_sector_2	NA
induced_sector_2_activity_conversion_rate	NA
induced_sector_3	NA
induced_sector_3_activity_conversion_rate	NA
benchmark_asset_id	NA

References

1. Ayub, H. R., Ambusso, W. J., Manene, F. M., & Nyaanga, D. M. (2021). A review of cooking systems and energy efficiencies. *American Journal of Energy Engineering*, *9*(1), 1-7.
2. Carella, A., & D'Orazio, A. (2021). The heat pumps for better urban air quality. Sustainable Cities and Society, 75, 103314.
3. Ciolkosz, D., & Chahal, A. (n.d.). Online annual heating energy cost analyzer. Penn State Extension. Accessed October 3, 2025 via https://extension.psu.edu/online-annual-heating-energy-cost-analyzer
4. Crippa, M., Guizzardi, D., Pagani, F., Schiavina, M., Melchiorri, M., Pisoni, E., Graziosi, F., Muntean, M., Maes, J., Dijkstra, L. and Van Damme, M., 2023. 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), pp.2811-2830.
5. Copernicus Climate Change Service, Climate Data Store, (2024): ERA5 post-processed daily-statistics on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), DOI: 10.24381/cds.4991cf48 (Accessed on 10-10-2025)
6. Dou, X., Wang, Y., Ciais, P., Chevallier, F., Davis, S.J., Crippa, M., Janssens-Maenhout, G., Guizzardi, D., Solazzo, E., Yan, F. and Huo, D., 2022. Near-real-time global gridded daily CO2 emissions. The Innovation, 3(1).
7. Energy Star, 2023. U.S. Energy Use Intensity by Property Type. https://portfoliomanager.energystar.gov/pdf/reference/US%20National%20Median%20Table.pdf
8. Esch, T., Heldens, W., Hirner, A., Keil, M., Marconcini, M., Roth, A., Zeidler, J., Dech, S. and Strano, E., 2017. Breaking new ground in mapping human settlements from space–The Global Urban Footprint. ISPRS Journal of Photogrammetry and Remote Sensing, 134, pp.30-42.

9. European Commission (n.d.). Energy, Climate Change Environment: Heat pumps. https://energy.ec.europa.eu/topics/energy-efficiency/heat-pumps_en (Accessed 3 October 2025).
10. European Commission Joint Research Centre T1 Unit (2024) INSPIRE Registry. Available at: https://inspire.ec.europa.eu/codelist/HILUCSValue/5_ResidentialUse:1 (Accessed: 30 October 2024).
11. GADM maps and Data (2024) GADM. Available at: https://gadm.org/ (Accessed: 30 October 2024).
12. Geonames (2023): https://www.geonames.org/
13. Global Covenant of Mayors for Climate & Energy, "Data Portal for Cities." Accessed: Sep. 28, 2023. [Online]. Available: https://dataportalforcities.org/
14. C.F. Gould, M.L. Bejarano, B. De La Cuesta, D.W. Jack, S.B. Schlesinger, A. Valarezo, & M. Burke, Climate and health benefits of a transition from gas to electric cooking, Proc. Natl. Acad. Sci. U.S.A. 120 (34) e2301061120, https://doi.org/10.1073/pnas.2301061120 (2023).
15. Hersbach, H., Comyn-Platt, E., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J-N., Cagnazo, C., Cucchi, M. (2023): ERA5 post-processed daily-statistics on pressure levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), DOI: 10.24381/cds.4991cf48 (Accessed on 10-10-2025)
16. Intergovernmental Panel on Climate Change (IPCC). 2006. 2006 IPCC Guidelines for National Greenhouse Gas Inventories: Volume 2, Energy: Chapter 2, Stationary Combustion. Prepared by Gómez, D.R., Watterson, J.D., Americano, B.B., Ha, C., Marland, G., Matsika, E., Namayanga, L.N., Osman-Elasha, B., Kalenga Saka, J.D., and Treanton, K. IGES, Japan.
17. International Energy Agency (no date) Buildings - Energy System, IEA. Available at: https://www.iea.org/energy-system/buildings (Accessed: 30 October 2024).
18. International Energy Agency (2021) Net Zero by 2050: Net Zero by 2050 Scenario - Data product - IEA; as modified by Duke University. https://www.iea.org/data-and-statistics/data-product/net-zero-by-2050-scenario#overview
19. International Energy Agency (2022), The Future of Heat Pumps, IEA, Paris https://www.iea.org/reports/the-future-of-heat-pumps, Licence: CC BY 4.0
20. International Energy Agency. (2025). Heat pumps - Web page. IEA. Accessed October 3, 2025 via https://www.iea.org/energy-system/buildings/heat-pumps
21. Kashtan, Y., Nicholson, M., Finnegan, C. J., Ouyang, Z., Garg, A., Lebel, E. D., ... & Jackson, R. B. (2024). Nitrogen dioxide exposure, health outcomes, and associated demographic disparities due to gas and propane combustion by US stoves. Science Advances, 10(18), eadm8680.
22. Lebel, E. D., Finnegan, C. J., Ouyang, Z., & Jackson, R. B. (2022). Methane and NO x emissions from natural gas stoves, cooktops, and ovens in residential homes. Environmental science & technology, 56(4), 2529-2539.

23. Liu, E., Meng, C., Kolodner, M., Sung, E.J., Chen, S., Burke, M., Lobell, D. and Ermon, S., 2023. Building coverage estimation with low-resolution remote sensing imagery. arXiv preprint arXiv:2301.01449.
24. Luers, A., Yona, L., Field, C.B., Jackson, R.B., Mach, K.J., Cashore, B.W., Elliott, C., Gifford, L., Honigsberg, C., Klaassen, L. and Matthews, H.D., 2022. Make greenhouse-gas accounting reliable—build interoperable systems. nature, 607(7920), pp.653-656.
25. Marconcini, M., Metz-Marconcini, A., Üreyen, S., Palacios-Lopez, D., Hanke, W., Bachofer, F., Zeidler, J., Esch, T., Gorelick, N., Kakarla, A. and Paganini, M., 2020. Outlining where humans live, the World Settlement Footprint 2015. Scientific Data, 7(1), p.242.
26. McDuffie, E.E., Smith, S.J., O'Rourke, P., Tibrewal, K., Venkataraman, C., Marais, E.A., Zheng, B., Crippa, M., Brauer, M. and Martin, R.V., 2020. A global anthropogenic emission inventory of atmospheric pollutants from sector-and fuel-specific sources (1970–2017): an application of the Community Emissions Data System (CEDS). Earth System Science Data, 12(4), pp.3413-3442.
27. National MagLab. (n.d.). Electric Range – 1892. https://nationalmaglab.org/magnet-academy/history-of-electricity-magnetism/museum/electric-range-1892 (Accessed 03 October 2025).
28. National Renewable Energy Laboratory, "Decarbonizing Building Thermal Systems: A How-to Guide for Heat Pump Systems and Beyond." 2024. ttps://betterbuildingssolutioncenter.energy.gov/sites/default/files/attachments/87812.pdf
29. Oda, T., Maksyutov, S. and Andres, R.J., 2018. The Open-source Data Inventory for Anthropogenic CO 2, version 2016 (ODIAC2016): a global monthly fossil fuel CO 2 gridded emissions data product for tracer transport simulations and surface flux inversions. Earth System Science Data, 10(1), pp.87-107.
30. Pearce, F., 2024. Nations Are Undercounting Emissions, Putting UN Goals at Risk. Yale Environment 360.
31. Pesaresi, M., Huadong, G., Blaes, X., Ehrlich, D., Ferri, S., Gueguen, L., Halkia, M., Kauffmann, M., Kemper, T., Lu, L. and Marin-Herrera, M.A., 2013. A global human settlement layer from optical HR/VHR RS data: Concept and first results. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 6(5), pp.2102-2131.
32. Reeg, L., Shea, R., & Tang, J. (2025). Lower bills, cleaner air: Heat pump benefits for homes relying on delivered fuels. Rocky Mountain Institute. https://rmi.org/lower-bills-cleaner-air-heat-pump-benefits-for-homes-relying-on-delivered-fuels/. Accessed 9 October 2025.
33. RMI, "Tracking the Heat Pump & Water Heater Market in the United States." 2025. https://rmi.org/insight/tracking-the-heat-pump-water-heater-market-in-the-united-states/
34. The Humanitarian Data Exchange (2024) Welcome - Humanitarian Data Exchange. Available at: https://data.humdata.org/ (Accessed: 30 October 2024).
35. Shah, A.. Now is the time to go all-in on heat pumps. Rocky Mountain Institute. (2023). https://rmi.org/now-is-the-time-to-go-all-in-on-heat-pumps/. Accessed 9 October 2025.

36. Tong, D., Zhang, Q., Davis, S.J., Liu, F., Zheng, B., Geng, G., Xue, T., Li, M., Hong, C., Lu, Z. and Streets, D.G., 2018. Targeted emission reductions from global super-polluting power plant units. Nature Sustainability, 1(1), pp.59-68.
37. U.S. Energy Information Administration (2015). Residential Energy Consumption Survey (RECS). Accessed 2023. https://www.eia.gov/consumption/residential/.
38. U.S. Energy Information Administration (2018). Commercial Building Energy Consumption Survey (CBECS). Accessed 2023. https://www.eia.gov/consumption/commercial/.
39. Umemiya, C. and White, M.K., 2024. National GHG inventory capacity in developing countries–a global assessment of progress. Climate Policy, 24(2), pp.164-176.
40. United Nations Energy Statistics Database. (n.d.) https://unstats.un.org/unsd/energystats/data/ (Accessed 01 July 2025).
41. UNFCCC, "GHG data from UNFCCC | UNFCCC." Accessed: Jun. 26, 2024. [Online]. Available: https://unfccc.int/topics/mitigation/resources/registry-and-data/ghg-data-from-unfccc
42. Washington, DC: World Resources Institute, "Climate Watch." 2024. Accessed: Jun. 26, 2024. [Online]. Available: https://www.climatewatchdata.org
43. World Bank (2019). CURB: Climate Action for Urban Sustainability tool. Accessed 2023. https://datacatalog.worldbank.org/search/dataset/0042029.
44. World Health Organization (WHO). (n.d.). Cooking fuels: population with primary reliance on fuels and technologies for cooking, by fuel type, proportion (%). Global Health Observatory. Accessed October 3, 2025 via https://www.who.int/data/gho/data/indicators/indicator-details/GHO/proportion-of-population-with-primary-reliance-on-fuels-and-technologies-for-cooking-by-fuel-type