Waste Sector:
Solid Waste Disposal (asset)

Krsna Raniga[1], Aaron Davitt[1], Lekha Sridhar[1], Christy Lewis[1], Lee Gans[1], and Gavin McCormick[1]
1) WattTime

1. Introduction

Methane, or CH_4, is a potent greenhouse gas, responsible for 30% of global warming since the industrial revolution (Ayandele et al., 2022). It is short-lived compared to carbon dioxide but is 28 times more potent over a 100 year timespan. Among the sectors most responsible for increasing human-generated atmospheric methane, solid waste disposal sites (SWDS) are the third largest source worldwide, after fossil fuels and enteric fermentation. For example, landfills accounted for 11% of global anthropogenic methane emissions in 2020 (Ayandele et al., 2022). In the United States alone, landfills contribute 14.5% of greenhouse gas generation (EPA, 2022).

There are two primary types of waste disposal practices: landfilling and open dumping. Globally, 37 percent of waste is landfilled, while 33 percent is openly dumped (Kaza et al., 2018). Within landfills, sanitary practices will vary from highly managed, covered, methane-collecting facilities, to simple designated locations within which all collected municipal waste is disposed. Here, the line can blur between dumpsites and landfills, the former of which, in their most common form, are patches of emergent waste. These are most common in low-income nations, where 90 percent of waste is openly dumped or burned and landfilling is usually not yet available (Kaza et al., 2018).

Data coverage is severely limited on waste sites. Datasets often rely on facility-level self-reporting or crowd-sourcing. Within these, the U.S. Environmental Protection Agency (EPA) Landfill Methane Outreach Program (LMOP) dataset is massive in scope and detail, including generated and recovered CH_4 emissions for many locations, but this data is limited to the United States (https://www.epa.gov/lmop). Globally, the Waste Atlas dataset is currently the best available ranked list of the world's largest waste sites (http://www.atlas.d-waste.com/). This site reports known waste capacities and publishes important data such as fractions of organic sub-components, recovered methane (if any), and oxidation factors. However, Waste Atlas does not provide nuanced insight into sanitary practices, and a more comprehensive global landfill database, akin to the EPA LMOP, is lacking.

An alternative approach to identify and provide detailed landfill and open waste information is the application of remote sensing. Previous studies highlight the ability of remote sensing to identify waste disposal sites and, in specific cases, quantify methane emissions from these sources (Iacoboaea and Petrescu 2013; Beaumont et al., 2014; Cusworth et al., 2020). Additionally, the Global Plastic Watch (GPW; https://globalplasticwatch.org/) is an organization that employs machine learning and remote sensing imagery to identify plastic-containing open waste aggregations. The GPW data provides waste site locations and sizes with additional metadata including nearby population numbers, soil composition and coverage, and water proximity, with emphasis on public health impacts. However, the GPW focus on plastic-detection may miss open waste sites with low plastic concentrations.

Climate TRACE sought to combine the best available information from various sources into a model that could be deployed globally to estimate solid waste emissions. Due to variance in the types of information contained in the limited range of available data sources, the burden of calculating emissions estimates was contingent on high-level modeling to compare and highlight the scales of *known* large waste aggregations and their consequent potential emissions. Importantly, the most significant features influencing emissions factors were type of waste site, operational status, and region of the world. In this case, the Intergovernmental Panel on Climate Change (IPCC) guidelines on solid waste emissions estimation were used, with some modifications (discussed in Section 2), to derive an equation that was linear with annual incoming waste, or annual capacity (IPCC, 2006; IPCC, 2019).

2. Materials and Methods

The emission source for waste sites is anaerobic decomposition of organic waste contained in solid waste disposal sites (SWDS). Many waste sites will contain mixed waste, but the primary contributor is municipal solid waste (MSW), or everyday "trash" or "garbage". The IPCC first-order decay (FOD) method was used as the starting point to estimate 2021 CH_4 emissions resulting from this decomposition, using the 2019 refinement to their IPCC guidelines (IPCC, 2006; IPCC, 2019). This method was modified to account for unavailable or inefficiently aggregated historical waste-site activity data. In this case, the total emissions estimate was calculated as the product of annual incoming waste, referred to as annual capacity, and the emissions factor. The capacity was treated as the primary input, while the emissions factor served as an effective scaling coefficient, requiring its own set of inputs. Generating 2021 CH_4 emissions estimates required calculating both components; the methods for each are highlighted in Sections 2.2 and 2.3.

The waste site datasets used by Climate TRACE were selected based on comprehensive data availability for three key parameters: land area, waste-in-place (or total tonnes of waste at the site), and annual capacity (meaning annual incoming waste). These SWDS were classified by two categories: type of site, either open dumpsite or landfill, and operating status, either active or inactive. These classifications were based on the last available information about the location, noting that some sites may have closed since the last update to the dataset, or an open dumpsite may have been reclaimed for sanitary landfilling. A site was treated as active if it was explicitly reported as "active" or "under construction", or if no status was published.

2.1 Datasets employed

2.1.1 Facility-level and crowd source reporting

U.S. EPA: Waste disposal site data was collected from the Landfill Methane Outreach Program (LMOP; https://www.epa.gov/lmop). LMOP data sources include facility self-reports, LMOP partner reports, and publicly available data (EPA, 2022). This is the most comprehensive, detailed, and updated dataset publicly available on U.S. waste disposal sites, with the last reporting year being 2020. In particular, beyond just area, waste-in-place, annual capacity, and operating years, many sites published landfill gas (LFG) generation and collection values. For those sites, the net published LFG values (LFG generated minus LFG collected) were utilized for the reported year's methane emission value and rescaled to provide an updated estimate for 2021, as described in Section 2.2.

Waste Atlas: This database was initially released in 2013, designed as an interactive map with data compiled from crowdsourcing and scientific research (http://www.atlas.d-waste.com/). Since its initial release, and a 2014 report highlighting in-depth statistics about the world's 50 largest dumpsites, additional dumpsites and a sanitary landfill catalog have been added (Mavropoulos et al., 2014). The information available per location is a mix of waste-in-place, annual capacity, waste year (last year with updated data), and operating status.

Other: Eight waste sites not contained in any of the three datasets used (EPA, Waste Atlas, GPW) were also included, with data aggregated from additional sources (see Section 5, Supplementary Materials), as their sizes would rank them among the largest currently known.

2.1.2 Remote-sensing identified

Global Plastic Watch (GPW): This database consists of plastic-containing waste sites identified by a system of neural networks created to analyze spectral, spatial, and temporal components of Sentinel-2 satellite data to identify terrestrial waste aggregations (Kruse et al., 2022; https://globalplasticwatch.org/). Following identification, the footprint of each site was calculated and monitored at monthly intervals. As part of the initial effort, this approach has detected nearly 3,000 waste sites in 26 countries.

The GPW method generated contours to estimate the areas of plastic-containing fractions of detected waste aggregations, updated at a monthly cadence from the Sentinel-2 observation period between 2017 and early-2021. Since the contours only accounted for plastic waste, these areas needed to be rescaled to estimate the total area-cover of all waste types at the site, the approach for which is outlined in Section 2.3.1. As no specific timestamp was available for each site when the data was downloaded to use for this work, all areas were treated as the most updated values as of early-2021.

Figure 1 (Top) The layout of the GPW website, highlighting countries with locations identified in yellow. (Bottom) An example of a large waste site in Indonesia, showing available site attributes when selected (top left corner of the image).

2.2 Emissions Model

The 2021 value for tonnes of CH_4 emissions was a product of two values, the emissions factor and the annual capacity (Eq. 1):

$$CH_4\ Emissions_{\,2021} = (CH_4\ emissions\ factor) \cdot (Capacity_{\,2021})$$

Capacity was the primary data-derived input, incorporating waste-site areas from GPW and known waste-in-place and capacity values from Waste Atlas and the EPA. Differing approaches were required to estimate 2021 capacity, based on the information reported per site per dataset. The Climate TRACE dataset also published total waste-in-place in addition to annual capacity and emissions estimates, so the procedures highlighted additionally served to calculate or update waste-in-place. As a result, waste quantities were estimated even for those EPA sites which reported LFG generated directly.

The different waste-in-place and capacity scenarios are outlined here, with their methods explained in Sections 2.2.1 and 2.2.2:

Waste Atlas and EPA:
- Waste-in-place known, capacity known; active vs inactive status
- Waste-in-place known, capacity unknown; active vs inactive status
- Capacity known, waste-in-place unknown; active vs inactive status
- Capacity unknown, waste-in-place unknown: *site omitted, not published*

Global Plastic Watch:
- Sites with waste-site area reported (section 2.2.2)
- Sites with no area reported: *site omitted, not published*

Emissions factor calculations are explained in detail in Section 2.2.3. Preliminarily, values for the coefficients contributing to the total emissions factor were decided based on the following criteria:

- Type of waste-site: sanitary landfill vs dumpsite
- World region of the waste-site per the World Bank (in fact, also necessary to update waste-in-place and capacity): East Asia & Pacific, South Asia, Europe & Central Asia, Middle East & North Africa, Sub-Saharan Africa, North America, Latin America & the Caribbean (Kaza et al., 2018)
- Human Development Index (HDI) of the waste-site country

The final scenario was the case of EPA locations where landfill gas (LFG) generated was reported, which did not use Eq. 1; however, their waste-in-place and capacity values were still updated as outlined above. These reported LFG values were converted to the appropriate units and directly scaled up to estimate 2021 emissions. This case is explained in detail in section 2.2.4.

2.2.1 Estimating Waste-in-Place & Annual Capacity for Waste Atlas and EPA

The scenarios used to estimate the $Capacity_{2021}$ variable in Eq. 1 are displayed in Figures 2 to 4 and explained below. Terms used in the scenarios are:

- "*r*" is the average regional annual waste generation growth rate for the world region in which the site is located, calculated from the World Bank's projection for 2016-2030 waste growth (Kaza et al., 2018);
- W = waste-in-place, or the total tonnes of waste contained at the site;
- C = annual capacity, or tonnes of annual incoming waste;
- y = year last reported;
- y_c = year closed;
- y_i = year opened.

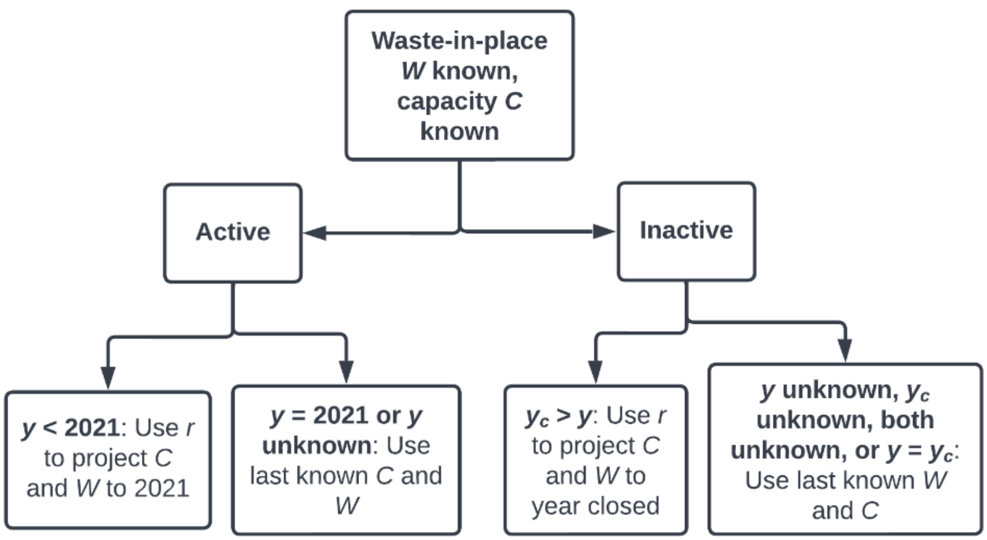

Figure 2 Scenario 1 flowchart to estimate capacity: Waste-in-place known, capacity known.

Scenario 1: Where waste-in-place, annual capacity, and waste-in-place year were all known (Figure 2). If the status was considered active (either explicitly active, or labeled unknown or under construction), then either:

- If the year last reported year *"y"* was before 2021, then the waste-in-place "*W*" was projected forward by adjusting the initial capacity "*C*" yearly up to 2021 and adding to the initial waste-in-place using growth rate "*r*". The final 2021 waste-in-place estimated here represents the summation of previous years waste-in-place for a final 2021 total. The 2021 adjusted capacity used for the summation was kept as final capacity.

- If the year that was last reported *"y"* was 2021 or *"y"* was unknown, then the last reported capacity and waste-in-place were kept as final values.

Where the status was known as inactive, then either:

- If the year closed "y_c" was after the known year last reported "y", then the waste-in-place "W" was projected forward by adjusting the initial capacity "C" yearly up to "y_c" and adding to the initial waste-in-place using growth rate "r". The "y_c"-adjusted capacity used for the summation was kept as final capacity.
- If "y_c" and "y" were unknown, then the last known "W" and "C" were used.

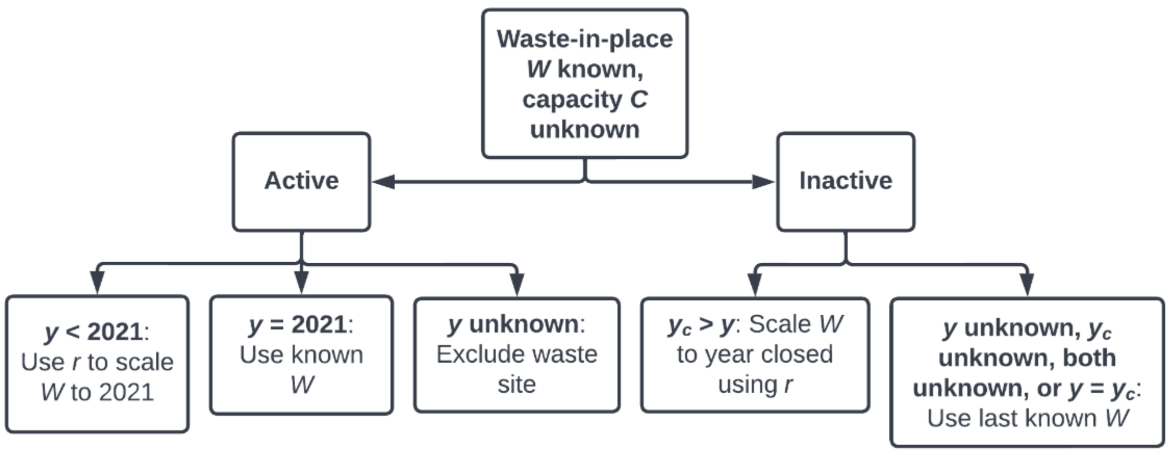

Figure 3 Scenario 2 flowchart to estimate capacity: Waste-in-place known, capacity unknown.

Scenario 2: Where waste-in-place and waste-in-place year were both known but capacity was unknown (Figure 3). If the status was active, then either:

- If the year last reported *"y"* was before 2021, waste-in-place "W" from that year was directly scaled up to 2021 using growth rate "r", instead of adding scaled capacity each year up to 2021.
- If the last reported year *"y"* was 2021, then known "W" was used.
- If "y" was unknown, the site was excluded from emission estimates.

Where the status was known as inactive, then either:

- If year closed was known, waste-in-place "W" was scaled up using "r" from the last reported year *"y"*to the year closed "y_c".
- If "y_c" and "y" were unknown, then the last known "W" was used.

Since capacity was unknown in scenario 2, it was derived from the waste-in-place estimated above. First, the capacity/waste-in-place ratios were calculated for any sites where both were known, and then average ratios by world region were calculated. The estimated waste-in-place values were multiplied by these regional ratios to estimate the unknown capacities.

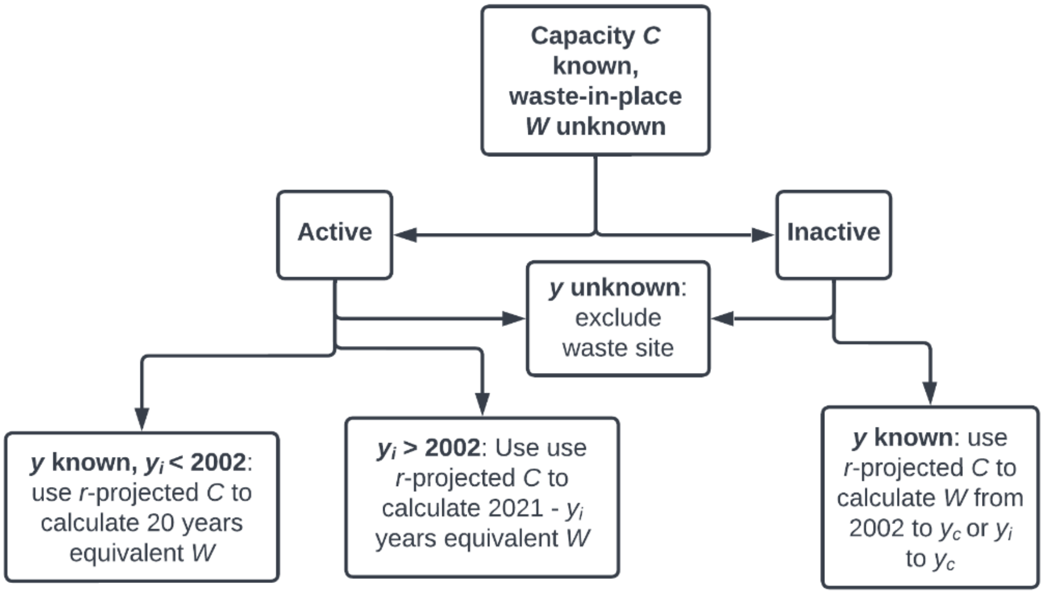

Figure 4 Scenario 3 flowchart to estimate capacity: Capacity known, waste-in-place unknown.

Scenario 3: Where waste-in-place was unknown, but capacity was known (Figure 4). If the status was active, then either:

- If the year last reported "*y*" was known or year opened "y_i" was before 2002, then growth rate "*r*" was used to project capacity "*C*" from 2002 to 2021. This produced an effective waste-in-place calculated by summing up these projected capacity values.
- If the year opened " y_i" was after 2002, growth rate "*r*" was used to project capacity "*C*" from "y_i" to 2021.

If the status was inactive and year last reported *"y"* was known:

- Growth rate "*r*" was used to project capacity "*C*" from year opened "y_i" to 2021, if 2002 < "y_i" < 2021.
- If "y_i" < 2002, then "y_i" = 2002, and growth rate "*r*" was used to project capacity "*C*" from 2002 to 2021.

The time window of 2002 and 2021 was chosen to calculate 20 years' worth of effective waste-in-place, a period over which all methane is effectively emitted from solid waste (Agency for Toxic Substances & Disease Registry, 2001). Lastly, if "*y*" was unknown, regardless of operational status, then these waste sites were excluded from emissions estimates.

2.2.2 Estimating Waste-in-Place & Annual Capacity from Global Plastic Watch

A priori, the belief was that total waste-in-place and site size should be correlated. As area was the most informative site activity value available from the GPW dataset, and volume and waste density values were not available, a relationship was needed between area and waste-in-place for the final emissions calculation. Figures 5 and 6 display the relationship between area and waste-in-place data, using the Waste Atlas and EPA datasets as ground truth to establish the strength of the correlation. These figures indicate that the amount of waste-in-place increases as the dumpsite or sanitary landfill area increases.

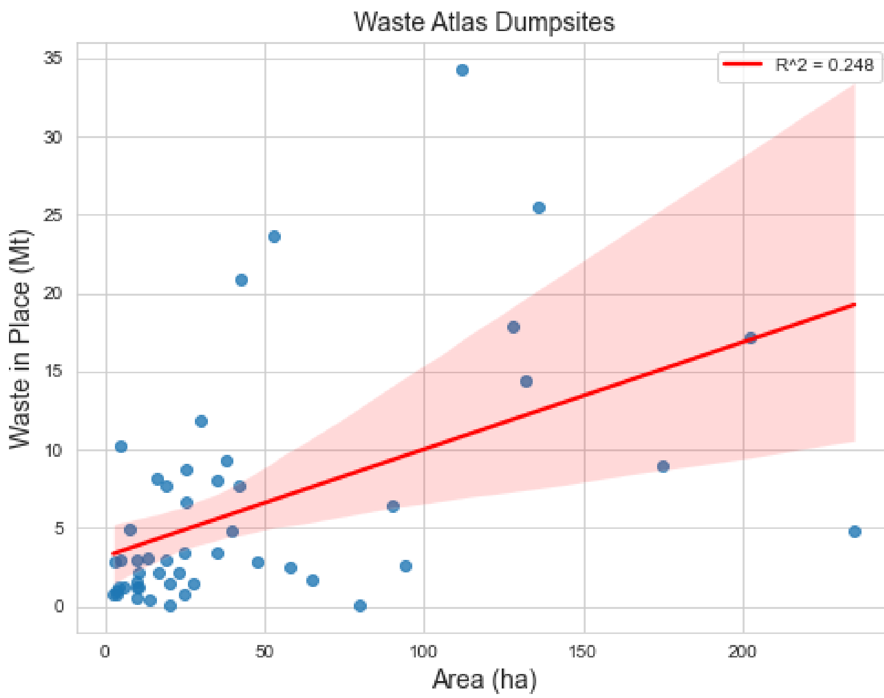

Figure 5 Waste Atlas waste-in-place (Mt) and area (ha) relationship for open dumpsites (R^2 = 0.248). Shaded area represents the 95% confidence interval.

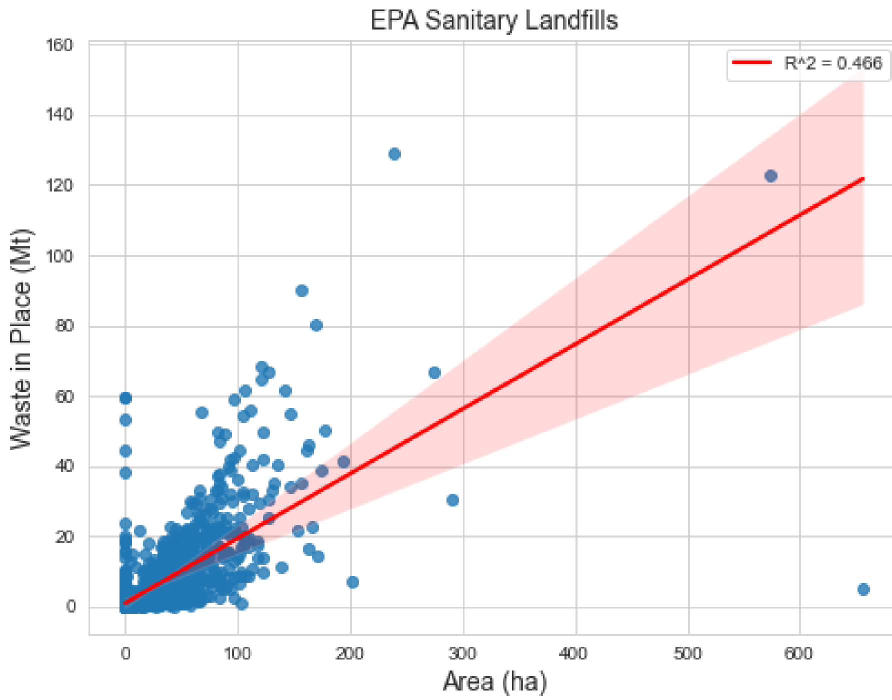

Figure 6 EPA waste-in-place (Mt) and area (ha) relationship for sanitary landfills ($R^2 = 0.466$). Shaded area represents the 95% confidence interval.

Sanitary landfill waste management practices, especially in the United States, will often tend to strategically layer and condense waste. In contrast, open dumpsites are often emergent phenomena and are most common in developing regions of the world, especially where municipal waste collection and centralization practices are inadequate (Kaza et al., 2018). This makes them better candidates for random and unpredictable patterns of waste distribution within a site. These features may at least partly account for the relatively weaker linear relationship for dumpsites, other than the fact that there were more EPA sanitary landfills available to plot in Figure 6. Nonetheless, the correlation explored above established that there was a relationship between area and waste-in-place.

The area-waste relationship was used to estimate both waste-in-place and capacity for GPW sites. First, to estimate actual land use from the reported plastic-containing areas, the GPW areas were rescaled. For the 100 largest sites where plastic-containing areas were published, contours were manually drawn on Google Earth. The average scale factor of Google Earth-estimated to GPW-reported areas was then calculated, and the remaining site areas were multiplied by that factor.

Second, for Waste Atlas locations that overlapped with GPW, of which there were 25, their last known area to last known waste ratios were calculated to obtain area densities, assuming a constant waste density over time. The rescaled GPW site areas were multiplied by the average of these Waste Atlas area densities to estimate total waste-in-place. Then, using the same regional average capacity/waste-in-place ratios as developed in Scenario 2 in Section 2.2.1 for sites with unknown capacities, the capacity was calculated as a fraction of the estimated waste-in-place.

2.2.3 Calculating Emissions Factors

For the general case where emissions values were not reported, the IPCC first-order decay method was used to calculate emissions factors, which were then applied to Eq. 1 (IPCC, 2006; IPCC, 2019). This included all Waste Atlas data, GPW data, the eight extraneous sites ("Other" in Section 2.1), as well as any EPA sites where LFG generated was not reported. The equation was modified to account for changes in annual capacity every year within a 20 year window (see Section 2.2.1, Scenario 3) based on the latest estimated capacity. Additionally, the equation was reversed in time running from the last known year to 2002 (or where facility age was younger), such that with increasing *i*, capacity decreased, but emissions increased. The methane emissions factor was estimated through the following equation (Eq. 2):

$$\mathrm{CH_4\ emissions\ factor} = L_0 \cdot (1-R) \cdot (1-e^{-k}) \cdot \sum_{y=y_i}^{y_f} e^{-k(y_f-y_i)} \cdot (1+r)^{-(y_f-y_i)}$$

where "y_i" is the starting year of the summation and "y_f" is the final year calculated. "k" is the methane generation rate constant, related to the decay half-life of waste at the site. A default value of $k = 0.05$ (indicating a half-life of 14 years), as suggested by IPCC best practices, was used for all locations (IPCC, 2006; IPCC, 2019). "r" is the same average regional annual waste generation growth rate as in Section 2.2.1. "R" is the fraction of methane captured by landfill gas recovery systems, which was published for a large number of EPA sites. For Waste Atlas sites, where the site was a sanitary landfill, or EPA locations where recovery values were not reported, a best practice default value of $R = 0.2$ was used. For dumpsites and GPW locations, where practices were unknown, $R = 0$, assuming no methane was captured by landfill gas recovery systems (IPCC, 2019).

"L_0" is the methane generation potential, defined as (Eq. 3):

$$L_0 = MCF \cdot DOC \cdot DOC_f \cdot F \cdot \frac{16}{12}$$

where the methane correction factor "*MCF*" accounts for degree of aerobic decomposition due to waste layering, effectively rescaling the emissions factor based on level of waste management at the site. For all EPA locations and for any Waste Atlas sites labeled as sanitary landfills, $MCF = 1$ was used, representing managed disposal sites. To categorize the remaining sites to select their "*MCF*" value, the Human Development Index (HDI) of the waste site's country was used. According to the United Nations (U.N.), HDI as a metric captures people's quality of life and wellbeing, in contrast to the Gross Domestic Product (GDP) alone, which considers only national economic activity (United Nations Development Programme, 2022). Most importantly, HDI's focus on quality of life across the population may be more closely tied to local waste management practices. Generally, as HDI increases, the amount of waste generated increases (Wilson et al., 2012). For Eq. 3, for Waste Atlas dumpsites and GPW locations in countries where the HDI < 0.8, the "*MCF"* was set to 0.4, indicating shallow unmanaged disposal sites. For the remaining locations in countries where HDI > 0.8, the "*MCF"* was set to 0.6, signifying uncategorized waste sites where the country's average practice was not known to be unmanaged open waste disposal. These values were chosen based on 2019-updated IPCC guidelines on methane correction classifications for types of solid waste disposal sites (IPCC, 2019).

Degradable organic carbon "*DOC"*, or the fraction of carbon available for decomposition, is broken down into paper and textiles, food waste, non-food organic waste, and wood and straw. The relative fractions of each, represented by the "*W"* in Eq. 4, were chosen based on regional average compositions from the World Bank report (Kaza et al., 2018). Food and non-food organics were combined into total organics, based on categories available from the World Bank. The calculation was then a weighted average of the carbon contents of different waste types "*W"* (Eq. 4):

$$DOC = 0.4 \cdot W_{paper,textiles} + 0.32 \cdot W_{organics} + 0.3 \cdot W_{wood}$$

Fraction of degradable organic carbon dissimilated "DOC_f" is an additional scale factor, accounting for the fraction of carbon that is actually released from decomposition. An IPCC best practice value of $DOC_f = 0.5$ was used (IPCC, 2006).

The fraction of methane in overall landfill gas "*F"* was taken to be the typical value of 0.5 for sites lacking their own fractions. Many LMOP sites published their own fractions "*F"*. Finally, 16/12 is the conversion factor for carbon to methane (IPCC, 2006).

2.2.4 Emissions for EPA Where LFG is Reported

For those sites where the EPA reported landfill gas (LFG) generated, values were published in units of millions of standard cubic feet per day (mmscfd), and the net CH_4 generated was calculated as follows (Eq. 5):

$$CH_{4,mmscfd,2021} = F \cdot (LFG_{generated} - LFG_{collected}) \cdot \sum_{y=y_i}^{y_f} (1+r)^{-(y_f - y_i)}$$

Once again, "*F"* is the fraction of methane in total LFG and "*r"* is the average annual waste generation growth rate for the United States (r = 1.0121). In most cases, the reported year was 2020 (and in some cases earlier), so the net LFG was scaled using the summation factor in Eq. 5.

Where LFG collected was not reported, LFG generated was scaled by a factor of 0.8 (using the IPCC best practice of R = 0.2), treating the fraction of methane captured by landfill gas recovery systems the same as in all other sanitary landfills across the various datasets used. Lastly, the emissions in mmscfd were converted to metric tonnes/yr. To calculate and report the emissions factor for the Climate TRACE platform, the resultant 2021 emissions were divided by the 2021 capacity values determined as outlined in Section 2.2.1.

2.4 Verification of Approach

In total, emissions estimates were modeled for 2021 for 4,609 solid waste disposal sites across 85 countries. Out of these, the EPA published LFG generation values for 991 U.S. locations; the methods in Section 2.3.1 were used to estimate CH_4 emissions to compare against the scaled published values described in Section 2.2.1. Therefore, the EPA dataset itself served as a source of ground truth validation, since large facility-level emissions datasets were scarce.

Country-level emissions inventories were more common, notable sources being the Emissions Database for Global Atmospheric Research (EDGAR) (https://edgar.jrc.ec.europa.eu/), Climate Watch from World Resources Institute (CAIT) (https://datasets.wri.org/dataset/cait-country), and the United Nations Framework Convention on Climate Change (UNFCCC) (https://unfccc.int/topics/mitigation/resources/registry-and-data/ghg-data-from-unfccc). The UNFCCC data was split by Annex-1 and Non-Annex 1 countries. As defined by the U.N., Annex-1 countries are industrialized nations who belonged to the OECD (Organisation for Economic Co-operation and Development) in 1992, in addition to “economies in transition”; Non-Annex 1 countries are mostly developing nations (UNFCCC, 2022). The EDGAR, CAIT, and UNFCCC datasets were used to compare against TRACE site emissions aggregated to country-level totals.

3. Result Highlights

Figure 7 shows hotspot waste sites, filtered for emissions within the 90th percentile of Climate TRACE estimates. Since the EPA dataset comprised 1,462 out of 4,609 total sites used for the TRACE dataset, the map unsurprisingly shows a large density of landfill locations in the U.S. Out of all GPW sites incorporated in the TRACE dataset, 25% of locations are in India and 13% are in Indonesia, and 77% of sites included in the Climate TRACE dataset from GPW are located in East, Southeast, or South Asia, explaining the prevalence of dumpsites visible in Asia in Figure 7.

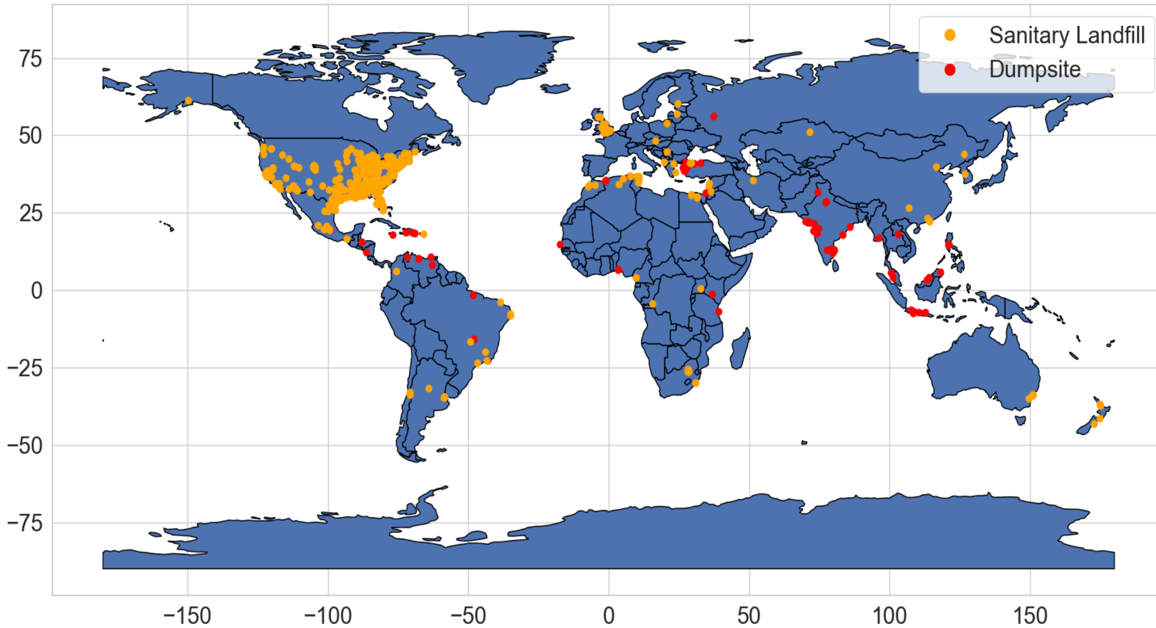

Figure 7 Identified waste sites worldwide whose emission estimates are in the 90th percentile. Orange dots represent Sanitary landfills and red dots represent Dumpsites.

Figure 8 compares CH_4 values derived for 2021 from net LFG emissions reported by the EPA using Eq. 5 in Section 2.2.4 (x-axis), and the full approach of Eq. 1 using Section 2.2.1 and Eq. 2 (y-axis). Comparing the regression line to the 1:1 line shows that TRACE values tend to underestimate reported values. The coefficient of determination is found to be $R^2 = 0.50$, quantifying the goodness of fit, which can be interpreted as saying that 50% of variation in values of y can be explained by values of x.

The variety of EPA's LMOP data sources (listed in Section 2.1.1) mean there will be a spread in how LFG quantities are measured, at what cadence, and at which location within the facility. The benefit of LMOP is its access to detailed, site-specific waste activity and emissions information. In contrast, the novel benefit of the Climate TRACE approach in Eq. 1 is that emissions estimation is consistent across sites once the input capacity is estimated, and each facility is evaluated on an equal footing. Due to these methodological differences between estimation techniques, both some alignment and some discrepancy between values is expected, as shown in Figure 8.

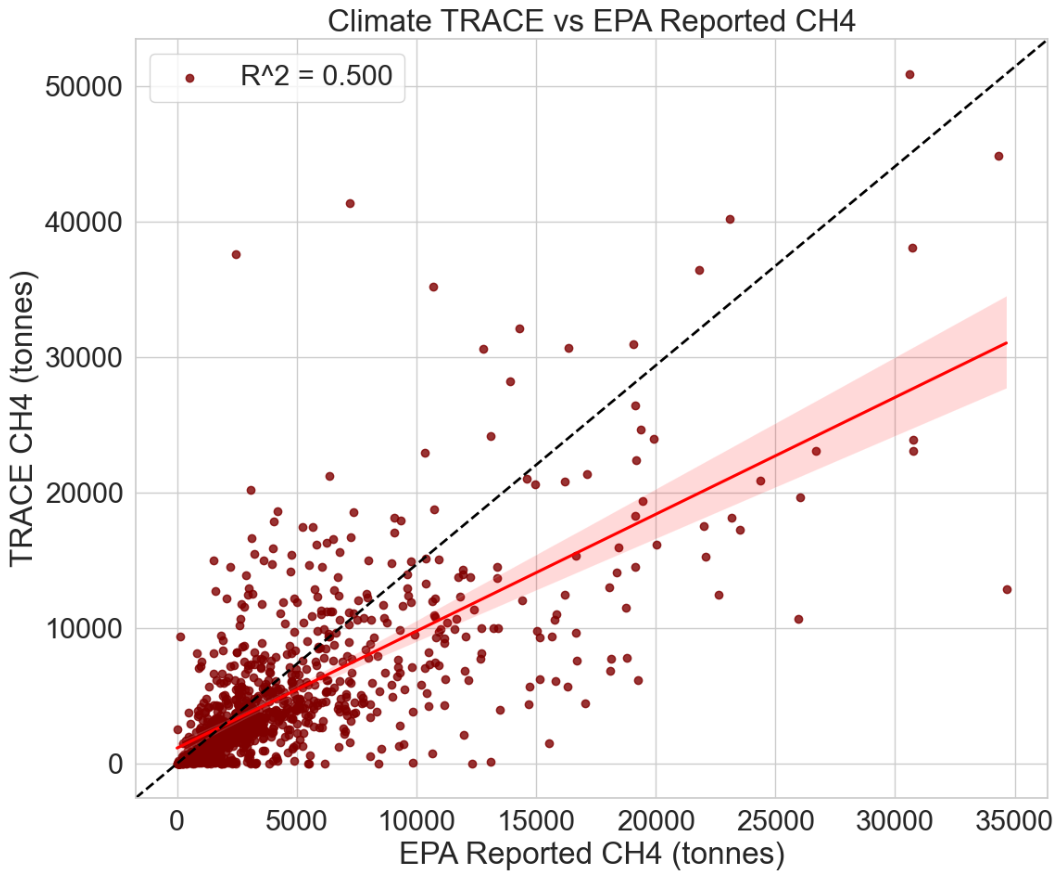

Figure 8 Climate TRACE (Eq. 1) vs EPA LFG reported (Eq. 5) emissions estimates. $R^2 = 0.500$, shaded region is the 95% confidence interval. Black dotted line is the 1:1 line.

To provide additional comparisons, the TRACE site-level emissions are aggregated and summed by country to compare against UNFCCC, CAIT, and EDGAR totals, seen in Figure 9. The regression lines compared to the 1:1 line show that the TRACE country level sums tend to underestimate emissions relative to all three other inventories. UNFCCC is closest to the 1:1 line, with the largest R^2 value, meaning that UNFCCC and Climate TRACE are the most closely correlated and have the best linear fit. CAIT and EDGAR, on the other hand, have lower R^2 values, indicating greater scatter about their lines of best fit. Climate TRACE underestimation at country-level is expected, since it is known that the collection of EPA, Waste Atlas, and GPW waste sites do not capture all waste sites in any country. In contrast, since UNFCCC, CAIT, and EDGAR are inherently country-level inventories, they likely account for a greater volume of total waste activity per country. Though one informs the other, and in the ideal scenario will produce equal final emissions values, site-level and country-level emissions are ultimately different domains of estimation and analysis.

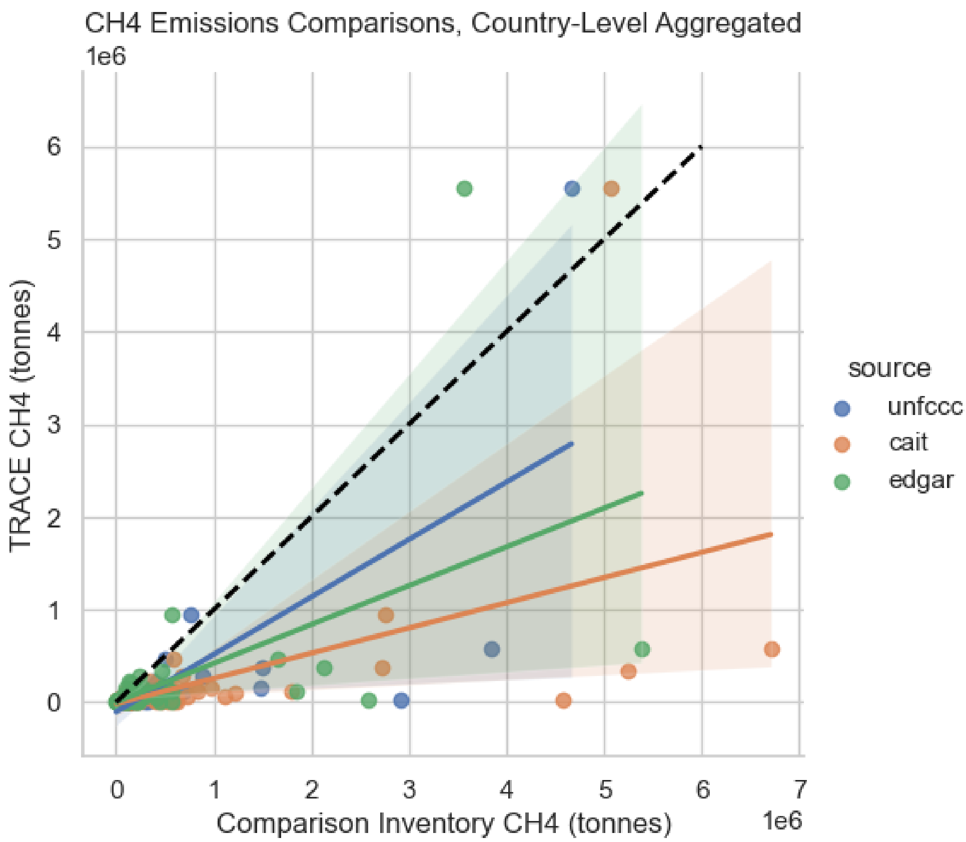

Figure 9 Climate TRACE country-level aggregated emissions vs country-level totals from UNFCCC (R^2 = 0.513), CAIT (R^2 = 0.284), and EDGAR (R^2 = 0.296). Black dotted line is the 1:1 line. Shaded regions are 95% confidence intervals.

The UNFCCC and EDGAR comparisons in Figures 10 and 11 demonstrate the tendency of country-level TRACE sums to comparatively underestimate total country emissions - there is a "pull" to the right of the 1:1 line. Interestingly, from Table 2, the top 10 countries for the UNFCCC comparison can have up to a 50% difference between TRACE and UNFCCC values, whereas EDGAR and CAIT are capped at ~21% and ~23%, respectively. However, as seen in Figure 9, the UNFCCC also has the strongest linear correlation at R^2 = 0.513, and thus still has the lowest overall spread across all countries compared to the other two inventories. The U.S. has the highest absolute emissions out of all the countries listed, so Figures 11 and 12 omit this country for cleaner visualization and to avoid intensive plot-rescaling; the U.S. values are instead reported in Tables 2 and 3. The relatively low percent difference for the U.S. can be attributed to the large number of sites cataloged by the EPA dataset. The numbers of waste sites identified and modeled for the remaining countries were sparser, as there were no additional single-country sources, resulting in higher absolute differences between Climate TRACE emissions and the three inventories compared.

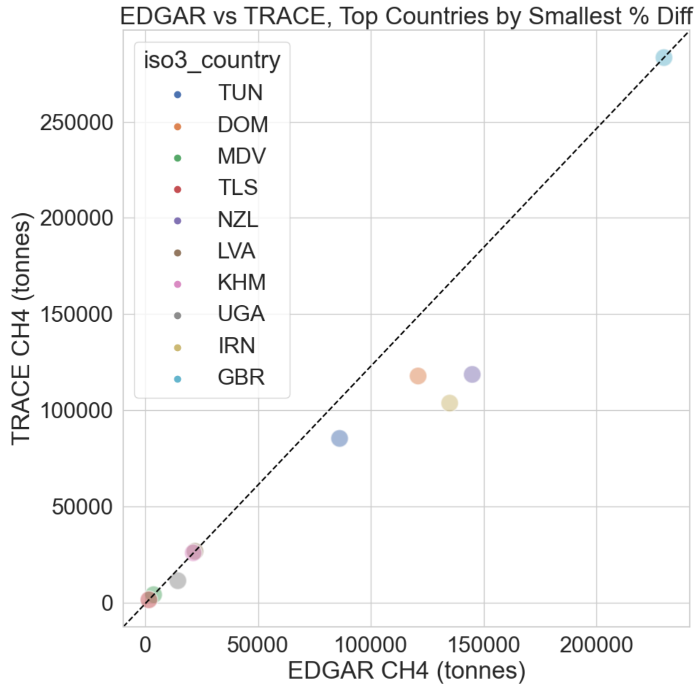

Figure 10 EDGAR vs TRACE, top 10 countries by smallest percent difference in emissions. Black dotted line is 1:1 line. Values are provided in Table 1.

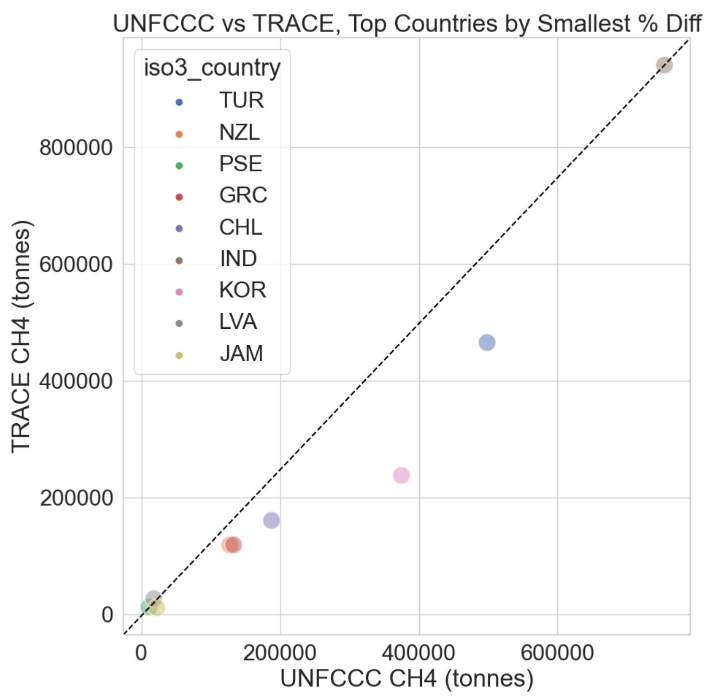

Figure 11 UNFCCC vs TRACE, top 9 countries (top 10 with U.S. excluded) by smallest percent difference in emissions. Black dotted line is 1:1 line. Values are provided in Table 2.

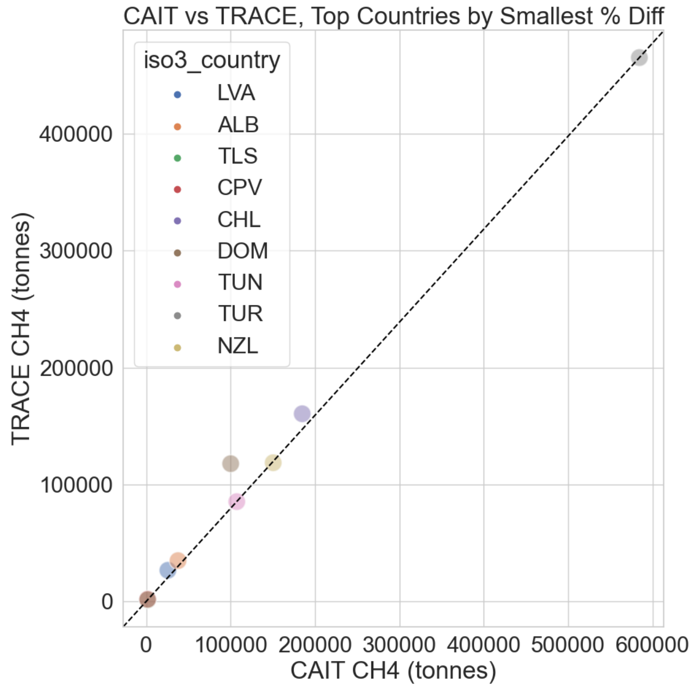

Figure 12 CAIT vs TRACE, top 9 (top 10 with U.S. excluded) countries by smallest percent difference in emissions. Black dotted line is 1:1 line. Values are provided in Table 3.

Tables 1 to 3 reveal some additional patterns. For example, small countries such as Timor-Leste, Maldives, and Cape Verde, though only having one site in the TRACE dataset each, still come close to the country totals estimated by CAIT and EDGAR. On the other hand, though a large country like Iran only has one site entry for TRACE, the emissions value comes within < 24% of EDGAR's country total estimate for Iran. This may suggest an overestimation from TRACE or underestimation from EDGAR. The only country to appear in all three inventory comparison rankings with a consistently low percentage difference is New Zealand, suggesting a high degree of cross-inventory precision in estimating its emissions activity. For example, the largest difference is with CAIT at 21.14%, but as low as 6.98% for the UNFCCC, implying that, unless there is significant waste activity publicly unknown in New Zealand, a range of 100,000 - 160,000 tonnes of CH_4 may be an accurate window for all solid waste emissions from New Zealand.

Altogether, it is evident that all inventories compared, including TRACE, use different approaches for estimating emissions, which result in varying total estimates with different countries appearing in the top 10 ranks in Figures 10-12 and Tables 1-3. For countries that do appear in more than one table, the degree of discrepancy in emissions between TRACE and each inventory are not necessarily similar. For example, Latvia appears in all three tables, but has a 4.03% difference in Table 3 versus a 48.37% difference in Table 2. The only consistency, as per Figure 9, is that TRACE, on average, underestimates country totals compared to the inventories.

Besides the lack of completeness in the TRACE site-inventory, the emissions factors used by UNFCCC, CAIT, and EDGAR may each be smaller compared to TRACE values, or their inputs for waste totals and capacities may be larger. Only more data can illustrate where the precise differences lie, as well as their degrees of accuracy.

Table 1 EDGAR vs TRACE ranked top 10 countries by smallest percent difference.

Country	# of sites ID'd by TRACE	TRACE emissions (tonnes)	EDGAR emissions (tonnes)	Absolute difference	% Difference
Tunisia (TUN)	16	85334.34	86200	865.66	1.00
Dominican Republic (DOM)	62	117799.49	121000	3200.51	2.65
Maldives (MDC)	1	4061.31	3750	311.31	8.30
Timor-Leste (TSL)	1	1400.44	1670	269.56	16.14
New Zealand (NZL)	38	118604.05	145000	26395.95	18.20
Latvia (LVA)	12	26631.40	22100	4531.40	20.50
Cambodia (KHM)	36	25828.83	21300	4528.83	21.26
Uganda (UGA)	1	11326.96	14500	3173.04	21.88
Iran (IRN)	1	103696.80	135000	31303.20	23.19
United Kingdom (GBR)	24	283383.03	230000	53383.03	23.21

Table 2 UNFCCC vs TRACE ranked top 10 countries by smallest percent difference.

Country	# of sites ID'd by TRACE	TRACE emissions (tonnes)	UNFCCC emissions (tonnes)	Absolute difference	% Difference
Turkey (TUR)	75	465039.91	498325.29	33285.38	6.68
New Zealand (NZL)	38	118604.05	127498.94	8894.90	6.98
Palestinian Territories (PSE)	4	12221.41	11074.6	1146.81	10.36
Greece (GRC)	46	118809.01	132983.29	14174.28	10.66
Chile (CHL)	10	160465.15	187700	27234.85	14.51
United States (USA)	1462	5554208.16	4664012.82	890195.34	19.09
India (IND)	691	940047.43	754000	186047.43	24.67
Korea (KOR)	4	237646.64	374980	137333.36	36.62
Latvia (LVA)	12	26631.40	17949.60	8681.80	48.37
Jamaica (JAM)	1	11051.13	22050	10998.87	49.88

Table 3 CAIT vs TRACE ranked top 10 countries by smallest percent difference.

Country	# of sites ID'd by TRACE	TRACE emissions (tonnes)	CAIT emissions (tonnes)	Absolute difference	% Difference
Latvia (LVA)	12	26631.40	25600	1031.40	4.03
Albania (ALB)	38	34877.57	37600	2722.43	7.24
United States (USA)	1462	5554208.16	5060800	493408.16	9.75
Timor-Leste (TLS)	1	1400.44	1600	199.56	12.47
Cape Verde (CPV)	1	1802.14	1600	202.14	12.63
Chile (CHL)	10	160465.15	184800	24334.85	13.17
Dominican Republic (DOM)	62	117799.49	100000	17799.49	17.80
Tunisia (TUN)	16	85334.34	107200	21865.66	20.40
Turkey (TUR)	75	465039.91	584400	119360.09	20.42
New Zealand	38	118604.05	150400	31795.95	21.14

4. Discussion and Conclusion

The approach utilized by Climate TRACE, as a modification of the IPCC's existing paradigm, is a critical step towards global scale, single site-level solid waste emissions estimation. Existing site-level emissions inventories tend to either be executed for specific countries or only probe values such as waste-in-place and nearby population, without estimating emissions. The TRACE approach implements the IPCC framework as a pipeline such that, even from the bare minimum information of land-area size, preliminary emissions may be calculated. An alternative method would have used nearby population numbers and regional economic information (instead of area) to model total waste generated. From that, one could estimate the percentage of that waste disposed of at a particular site. This route was not pursued as the number of detected sites available for estimating emissions outnumbered those where sufficient population data was available. Regardless, a crucial feature of the TRACE method for generating emissions factors is that it is agnostic to the means by which waste-in-place and annual capacity are calculated.

Many variables are required to model the emissions factors, for which even the coefficients in Eq. 2 are simplifications. Importantly, the TRACE method approximates historical activity for annual waste accumulation at a single site, as such site-by-site data was not available for this work and such data collection in the future would be impractical given the global scope of the Climate TRACE survey. The components of methane generation potential "L_0" (Eq. 3) each require site-specific data on their waste composition and degradation, as well as historical variation in each of these values. Moreover, in Eq. 1, the decay constant "k" took a best practice value for all sites, but per IPCC guidelines, "k" is unique to each waste facility and requires that site's historical waste decay data to derive (IPCC, 2019). As another example, a variable not included at all in the TRACE method was oxidation factor "*OX*", or the fraction of CH_4 that is oxidized due to coverage (such as soil) and layering over the waste (IPCC, 2006). Most of these values can only be probed at ground-level for sites with documented management practices. Assuming there were some way to aggregate better data on practices at managed locations on a global scale, it is still unclear how to systematically gather time-dependent data on open and emergent dumpsites, many of which are undocumented or unstudied, and which comprise a large fraction of Climate TRACE locations.

The most promising and holistic kernel for global scale waste-site activity and emission estimation is remote-sensing data, which is independent of crowd-sourced or self-reported waste-site activity data. For the TRACE emissions estimates, Global Plastic Watch served as the remote-sensing source to provide previously uncatalogued solid-waste aggregations. The best next step is to use remote-sensing technology capable of fully measuring waste-in-place, or at best directly measuring methane plumes, where continuous monitoring over time would illuminate temporal variability due to weather and climate patterns.

A final consideration is that waste sites, particularly open dumpsites in emerging and developing nations, can be volatile and unpredictable locations. Many people live near or even within dumpsites - some picking waste for scraps to sell for subsistence, others merely living near them through no will of their own - all impacted by the continuous release of CH_4 (amid other gasses) and uncollected liquid leachate. These large volumes of gas released from unmanaged waste aggregations can be highly flammable, commonly resulting in massive, toxic fires (Gupta, 2022). Lastly, an additional complication is that some of these waste sites play host to violent gang activity, serving as direct threats to locals while potentially stymying efforts at introducing management and gas collection practices (D'Aubuisson, 2022). Altogether, it is imperative to recognize that the hazards of solid waste disposal sites are substantial in both the short and long-term, and that any attention that the scope of Climate TRACE's solid waste emissions estimates may bring to improving waste-management practices to reduce global greenhouse gas emissions can produce equally substantial public health and safety benefits.

5. Supplementary Materials

As mentioned in Section 2.1.1, the following eight locations were not included in the EPA, Waste Atlas, or GPW datasets, but online searches yielded that they were among the largest known global waste sites. The names, countries, and sources for each are listed below:

1. Malagrotta Landfill (Italy): https://www.worldatlas.com/articles/largest-landfills-waste-sites-and-trash-dumps-in-the-world.html, https://archidiap.com/malagrotta-landfill-initiative-old/malagrotta-landfill-initiative/

2. Laogang Landfill (China): https://www.worldatlas.com/articles/largest-landfills-waste-sites-and-trash-dumps-in-the-world.html
3. Sao Joao Landfill (Brazil): http://www.nyc.gov/html/unccp/gprb/downloads/pdf/SaoPaulo_landfills.pdf
4. Jiangcungou Landfill (China): https://www.bbc.com/news/world-asia-50429119
5. Gonio Landfill (Georgia): https://frontline.thehindu.com/dispatches/life-on-georgia-largest-toxic-landfill/article37209818.ece
6. Meethotamulla (Sri Lanka): https://roar.media/english/life/reports/a-brief-history-of-the-meethotamulla-garbage-dump
7. Thilafushi (Maldives): https://gizmodo.com/thilafushi-maldives-trash-island-plastic-pollution-1848938393/slides/4
8. Yadrovo Landfill (Russia): https://www.nytimes.com/2018/04/05/world/europe/russia-landfills-gases.html

6. References

1. Agency for Toxic Substances & Disease Registry (2001) *Landfill Gas Primer - An Overview for Environmental Health Professionals*. Available at: atsdr.cd.gov/HAC/landfill/html/ch2.html (Accessed 01 September 2022).
2. Ayandele, E., Huffman, K., Jungclaus, M., Tseng, E., Duren, R., Cusworth, D. and Fisher, B., 2022. *Key Strategies for Mitigating Methane Emissions from Municipal Solid Waste*, RMI, https://rmi.org/insight/mitigating-methane-emissions-from-municipal-solid-waste/ (Accessed 01 September 2022).
3. Beaumont, B., Radoux, J. and Defourny, P., 2014. Assessment of airborne and spaceborne thermal infrared remote sensing for detecting and characterizing landfills., *Waste Management and The Environment VII*, 180, pp. 237-248.
4. Cusworth, Daniel H., Duren, R., Thorpe, A.K., Tseng, E., Thompson, D., Guha, A., Newman, S., Foster, K.T. and Miller, C.E., 2020. Using remote sensing to detect, validate, and quantify methane emissions from California solid waste operations, *Environmental Research Letters 15, no. 5*: 054012.
5. D'Aubuisson, J.J.M. (2022) 'How the MS13 Became Lords of the Trash Dump in Honduras', *Insight Crime*, 19 Jan [online]. Available at: https://insightcrime.org/news/honduras-how-ms13-became-lords-trash-dump/ (Accessed 01 September 2022).

6. Gupta, A. (2022) ‘Delhi landfill burning’, *The Times of India*, 25 May [online]. Available at: https://timesofindia.indiatimes.com/blogs/voices/delhi-landfill-burning/ (Accessed 01 September 2022).
7. Iacoboaea, C. and Petrescu, F., 2013. Landfill monitoring using remote sensing: a case study of Glina, Romania. *Waste Management & Research*, *31*(10), pp.1075-1080.
8. IPCC (Intergovernmental Panel on Climate Change), 2006. Chapter 3. *2006 IPCC Guidelines for National Greenhouse Gas Inventories, Volume 5, Waste*. Published: IGES, Japan.
9. IPCC (Intergovernmental Panel on Climate Change), 2019. Chapter 5. *2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories*. Published: IPCC, Switzerland.
10. Kaza, S., Yao, L., Bhada-Tata, P. and Van Woerden, F., 2018. *What a waste 2.0: a global snapshot of solid waste management to 2050*. World Bank Publications.
11. Kruse, C., Boyda, E., Chen, S., Karra, K., Bou-Nahra, T., Hammer, D., Mathis, J., Maddalene, T., Jambeck, J. and Laurier, F., 2022. Satellite Monitoring of Terrestrial Plastic Waste. *arXiv preprint arXiv:2204.01485*.
12. Mavropoulos, A., Mavropoulou, N., Anthouli, A. and Tsakona, M., 2014. *Waste Atlas 2014 Report - The World's 50 Biggest Dumpsites.* 10.13140/RG.2.2.33122.45763.
13. United Nations Climate Change *Parties & Observers*. Available at: https://unfccc.int/parties-observers (Accessed 01 September 2022).
14. United Nations Development Programme (2022) *Human Development Index (HDI)*. Available at: https://hdr.undp.org/data-center/human-development-index#/indicies/HDI (Accessed 01 September 2022).
15. United States Environmental Protection Agency (2022) *LMOP Landfill and Project Database*. Available at: https://www.epa.gov/lmop/lmop-landfill-and-project-database (Accessed 01 September 2022).
16. EPA (United States Environmental Protection Agency) (2022) *Basic Information about Landfill Gas* (2022). Available at: https://www.epa.gov/lmop/basic-information-about-landfill-gas (Accessed 01 September 2022).
17. Wilson, D.C., Rodic, L., Scheinberg, A., Velis, C.A. and Alabaster, G., 2012. Comparative analysis of solid waste management in 20 cities. *Waste management & research*, *30*(3), pp. 237-254.