ENVIRONMENTAL RESEARCH
LETTERS

LETTER • OPEN ACCESS

Assessing the accuracy of the Climate Trace global vehicular CO_2 emissions

To cite this article: Kevin R Gurney *et al* 2026 *Environ. Res. Lett.* **21** 094018

IOP Publishing Environ. Res. Lett. **21** (2026) 094018 https://doi.org/10.1088/1748-9326/ae6355

ENVIRONMENTAL RESEARCH
LETTERS

OPEN ACCESS

RECEIVED
29 October 2025

REVISED
27 February 2026

ACCEPTED FOR PUBLICATION
22 April 2026

PUBLISHED
5 May 2026

LETTER

Assessing the accuracy of the Climate Trace global vehicular CO_2 emissions

Kevin R Gurney*, **Bilal Aslam** and **Pawlok Dass**

School of Informatics, Computing, and Cyber Systems, Northern Arizona University, Flagstaff, AZ 86011, United States of America
* Author to whom any correspondence should be addressed.

E-mail: kevin.gurney@nau.edu

Keywords: CO_2 emissions, machine learning, on road

Supplementary material for this article is available online

Abstract

Accurate estimation of greenhouse gas (GHG) emissions at the infrastructure scale remains essential to climate science and policy applications. Vehicle emissions often dominate GHG emissions in urban areas and are rapidly increasing globally. Climate Trace (CT), co-founded by former U.S. Vice President Al Gore, is a new AI-based effort to estimate roadway-scale GHG emissions. However, limited independent peer-reviewed assessment has been made of this dataset. Here, we compare CT on road CO_2 emissions in U.S. urban areas to atmospherically calibrated, multi-constraint estimates of on road CO_2 emissions from the Vulcan Project. Across 260 urban areas in 2021, we find a mean relative difference (MRD) of 70.4%. These large differences are driven by biases in CT's machine learning model, fuel economy values, and fleet distribution values. We conclude that sub-national policy guidance or climate science applications using the on road CO_2 emissions estimates made by CT should be done so with caution.

1. Introduction

The emission of greenhouse gases (GHGs) into the earth's atmosphere is the primary driver of current and projected climate change [1]. Accurate estimation of GHG emissions is not only essential for planning and deploying emissions mitigation activities, but remains fundamental to questions in carbon cycle science and climate change projections [2, 3]. Estimation of these fluxes has a long history and reflects efforts at multiple space and time scales, using different estimation approaches, and aimed at different questions across the climate science and policy landscape [4–6].

As GHG emissions mitigation activity has increasingly focused on sub-national spatial scales, with particular emphasis on urban, provincial, and facility/-corporate activities, the need for a mixture of high-resolution estimation that is also globally comprehensive has grown [7]. In response, a variety of datasets have been produced mostly from the research community using rigorous science-based approaches [8–14]. Most of these are expressed in a gridded format ranging in resolution from approximately 0.01×0.01 to $1 \times 1^{\circ}$ spatial resolution. This is often done with the expressed aim to interface with atmospheric modelling research. Scientific studies have attempted to quantify emissions at the scale of emitting infrastructure representing emissions for buildings, roadways, or powerplants [15–18]. These efforts have mostly been limited to domains representing a single country or smaller (e.g. province, city) due to the limited data and techniques able to resolve information at these very fine space/time scales.

The first attempt to mix asset-scale GHG emissions estimation with global coverage is the results provided by the Climate Trace (CT) coalition (https://climatetrace.org/). Established in 2019 and co-founded by former Vice President Al Gore, CT estimates emissions by '…training AI algorithms to fuse data across multiple wavelengths and from more than 300 satellites; 11 100 air-, land-, and sea-based sensors; and other data streams to identify all of the largest point sources and track them over time' [19]. The granularity was further emphasized in a recent data release, 'This expanded data set also includes inventories for more than 72 000 state and local regions, providing comprehensive GHG data for over 90% of the world's cities' [20]. And further claiming 'unprecedented granularity that pinpoints nearly every major source of GHG emissions around the world and provides independently produced estimates of how much each emits'[21]. These impressive claims are attracting important companies such as Boeing, Tesla, and General Motors to use CT data [22].

A recent study examined the CT emissions estimation for powerplant fossil fuel carbon dioxide ($FFCO_2$) emissions in the U.S. [23]. In 2022, power production accounted for 36% of global $FFCO_2$ emissions and 34.3% of the same emissions in the U.S., emphasizing the importance of this sector when considering emissions management [14, 24, 25]. When compared to Vulcan, a well-calibrated, high-resolution estimate of $FFCO_2$ emissions in the U.S., the study found a mean relative difference (MRD) of 50% across 1726 paired power plants. More surprisingly, the study concluded that only 3.7% of facility emissions estimation used AI techniques. A simple and error-prone approach was used for the remaining 96.7% of the facilities analysed. Of these facilities, only 8.7% agreed to within ±20% of the Vulcan estimates.

The publicity and promotional efforts by the CT coalition, the critical need for this type of data, combined with the large and persistent differences found in the previously completed power plant CO_2 emission assessment, warrants further examination of the CT GHG emissions data.

In this study, we focus on analysis of the on road $FFCO_2$ emissions from recent CT coalition data releases. On road $FFCO_2$ emissions account for 30% of the U.S. total $FFCO_2$ emissions and hence, are the second largest single emitting sector in the U.S. behind power production [25]. Within cities the on road emissions share is typically larger. For example, on road $FFCO_2$ emission accounted for 43% of the total $FFCO_2$ emissions in the Los Angeles megacity (California, USA) domain [16]. In Salt Lake city (Utah, USA), the on road $FFCO_2$ share was estimated at 37.6% and in Indianapolis (Indiana, USA), the share was estimated at 42% [15, 26].

As with the earlier study comparing power production facilities, the comparison provided here is made to the Vulcan Project version 4.0 on road $FFCO_2$ emissions [25, 27]. We provide a succinct description of the CT and Vulcan methods used to estimate on road $FFCO_2$ emissions and the comparison metrics relies upon. The differences are reported and an attempt is made to distinguish what elements of the estimation procedure account for the differences found. We end by recommending some best practices for improving the CT coalition estimates.

2. Methods

2.1. CT

The CT urban on road $FFCO_2$ emissions data for the years 2021 and 2022 were downloaded from the CT website (https://climatetrace.org/). Two versions of the CT emissions estimates were analysed here and we refer to these as 'version 1' and 'version 2'. CT initially published version 1 in 2024. An update to the urban on road $FFCO_2$ emissions estimates was published in 2025, reflecting methodological modifications [28]. However, since the release of version 2, version 1 was removed from the CT website. We retain analysis of both versions here, as the evolution of the results in comparison to Vulcan offers additional insights beyond sole comparison to the more recent version. We provide the earlier methodology document as a supplementary information attachment and note that both versions also refer to a study by Rollend *et al* [29].

CT estimates on road $FFCO_2$ emissions for 613 U.S. urban polygons in version 1 and 260 urban polygons in version 2. Both vintages of polygon boundaries were supplied by CT enabling a direct comparison with Vulcan. In version 1, the polygons represented subdivisions of urban areas while in version 2, the polygons were more representative of entire urban areas (see SI, figure S1). This partly accounts for the differing number of urban polygons between the two CT versions. However, version 1 is not a strict subdivision of the urban areas in version 2; while overlap occurred there are also instances where strict overlap was not present.

CT calculates on road $FFCO_2$ emissions partly using a machine learning (ML) approach. Their two ML models use the red/green/blue bands of Sentinel-2 satellite imagery (Level-2A product at 10 m × 10 m resolution) and are trained on a U.S. dataset of annual average daily traffic (AADT) at the road-segment scale to predict AADT in urban areas [30]. CT uses the OpenStreetMap (OSM) road network as a road basemap, which provides road segment geometry and length [31]. The ML approach effectively averaged outcomes of both a convolutional neural network model and a graph neural network model to predict AADT. The predicted AADT and road length are used for calculating vehicle kilometres travelled (VKT) for each road segment in the target urban areas (i.e. VKT = AADT × length). We convert the VKT from kilometres to miles vehicle miles travelled (VMT) in the comparison analysis performed in this study and refer to both AADT and VMT as a measure of 'vehicle activity'.

With segment-scale estimates of VKT, CT employs class-specific U.S. average fuel economy and fuel shares from the climate action for urban sustainability (CURB) database and an urban national average vehicle fleet distribution from the U.S. Federal Highway Administration (FHWA) to distribute the total VKT into a series of vehicle classes specific to fuel type [32–34]. The total distance driven on each segment divided by the fuel economy provides an estimate of the amount of fuel consumed. Application of a fuel carbon content (gasoline: 8.78 kgCO2/gal; diesel: 10.21 $kgCO_2$/gal; CNG: 0.054 $kgCO_2$/gal) provides an estimate of the total carbon emissions by vehicle class and road segment within the target urban areas. These are aggregated to provide the total on road $FFCO_2$ emissions in the chosen urban areas. Further specifics on the source or calculation method supporting the CURB fuel economy values are not available.

Aside from the differing urban geometry, the CT version 2 urban on road $FFCO_2$ estimation reflects two additional changes from the version 1 methodology. The version 2 methodology includes an adjustment to the fuel economy values in an attempt to reflect the additional fuel economy burden associated with vehicle cabin heating/cooling and related processes.

More importantly, the version 2 methodology scales the predicted AADT by a single 'activity scale factor', 0.3255, in order to generate consistency with country-scale on road CO_2 emissions from the global emissions dataset, EDGAR v8.0 [13, 28].

2.2. Vulcan

Vulcan version 4.0 on road $FFCO_2$ emissions are derived from emissions reported at the county scale from the 2017 national emissions inventory (NEI) [35–38]. The NEI estimate, in turn, is derived from simulations performed by the U.S. environmental protection agency (EPA) using the Motor Vehicle Emissions Simulator (MOVES v2014b) model across all counties except the state of California [39]. The NEI quantifies on road emissions by 13 vehicle types (which include a mix of vehicle classes and fuel types) and five road types, as well as differentiating between urban and rural roads. These simulations rely on a county database (CDB) compiled and gap-filled by the EPA from submissions by State, Local, and Tribal agencies. The CDB contains important inputs to the model emissions estimation, such as vehicular registrations, fleet characteristics, age, average speed, and fuel properties [40]. California on road $FFCO_2$ emissions are generated using a California-specific emissions model, EMFAC (Emission Factors), which are submitted to the EPA [41].

To estimate annual emissions for years 2021 and 2022, the State Energy Data System transportation fuel consumption was used to scale annual on road $FFCO_2$ emissions relative to 2017 [42]. $FFCO_2$ emissions at the county scale are downscaled to non-local and local road segments defined using a road basemap which integrates the highway performance monitoring system and OSM road networks [30, 31]. The sub-county emissions distribution to non-local roads uses a road segment-scale estimate of VMT_{AADT} derived from 2015 AADT data and road segment lengths:

$$\mathrm{VMT}\left(l\right) = \mathrm{AADT}\left(l\right) \times L\left(l\right) \tag{1}$$

where AADT is the 2015 traffic volume (vehicles/day) on road segment l, and L is the l^{th} road segment length (miles/segment) [30]. The share of the 2015 VMT on each road segment to the county total is used to apportion the county-scale total on road $FFCO_2$ emissions for each of the non-local road classes. This assumes that traffic volume is the primary distribution metric for non-local roads at the sub-county scale and ignores effects such as sub-county spatially varying congestion (to the extent that it does no correlate with volume), spatially varying vehicle fleet distribution, and spatially varying driving behaviour (e.g. speeding, aggressive stop/start).

Because local roads have very limited traffic volume (AADT) data, $FFCO_2$ emissions are distributed from the county to local roads based on road segment length instead of traffic volume. This latter downscaling assumes an even distribution of vehicles on local roads. This assumption is ameliorated by the fact that the total on road $FFCO_2$ on local roads is 9.7% and 11.6% % of the total $FFCO_2$ emissions in U.S. cities for the polygons in CT version 1 and 2, respectively.

$FFCO_2$ emissions associated with parking locations (e.g. truck stops, rest areas) was explicitly included and used directly from the NEI 2017 reporting.

Vulcan on road uncertainty is based on [43] and the 95% CI is estimated as ±14.2% for all road types [44]. This value is, in turn, referenced to Gately *et al* [45] and Mendoza *et al* [46]. This uncertainty is applied at the county/annual space/time scale and are propagated through all subsequent steps such as the allocation to sub-county road segments. No increase in the uncertainty is made during this allocation step.

Vulcan has been compared to independent estimates at multiple scales using multiple independent methods. For example, the total U.S.-wide Vulcan total $FFCO_2$ emissions were within 1.4% of an atmospheric radioisotope CO_2 inversion result [47]. In a comparison of the Vulcan U.S national total on road $FFCO_2$ to the U.S. EPA estimate, the relative difference, after accounting for identifiable differences due to assumed fuel heating values, was 2.2% [48]. Comparison of the Vulcan transportation sector $FFCO_2$ emissions to those published by the U.S. EPA at the state scale show a MRD of 4.4% [25].

Additional evaluation of the Vulcan on road emissions is performed at the road segment scale via comparison to emissions based on the FHWA Continuous Counting Stations (CCS). The FHWA measures hourly traffic volume by vehicular class at stationary CCS locations across the U.S. Details of the method to estimate $FFCO_2$ emissions at these observing sites and comparison outcomes is provided in the supplementary information text S2. The core result is that the MRD between the widely-distributed 810 CCS estimated $FFCO_2$ emissions and Vulcan at the sampled road segment locations is +0.13% (1 sigma range: +8.93%−9.48%) slightly larger in equivalent terms than the aggregate scale 95% CI of ±14.2% previously described.

Further details regarding Vulcan's methods of computing on road $FFCO_2$ emissions can be found in [25, 27, 44].

To compare Vulcan to the CT urban on road $FFCO_2$ emissions, the 2021 and 2022 Vulcan v4.0 on road $FFCO_2$ emissions estimates at the road segment scale were aggregated into the CT urban polygons (both versions) by bisecting roads where they were intersected by the CT urban polygon boundary. The bisected roads at the polygon boundaries were apportioned by share of road segment length within the polygon.

Since the 2017 NEI reporting does not explicitly produce VMT, a comparison of the Vulcan VMT to the CT VMT requires inferring VMT from a combination of the VMT_{AADT} produced according to equation (1) (based on AADT and road segment lengths) and VMT derived from the $FFCO_2$ values available from the 2017 NEI reporting. The VMT_{AADT} estimated in equation (1) is not directly used in the Vulcan processing since it is calculated only for the relative sub-county (i.e. road segment) spatial distribution and may not be consistent with the sub-county magnitude of the Vulcan $FFCO_2$ emissions.

For non-local roads, the road segment-scale VMT calculated in equation (1) is summed for every U.S. state and scaled to match the state-scale VMT reported by the FHWA [49]. The adjusted road segment-scale non-local VMT is aggregated within the CT version 1 and CT version 2 urban boundaries.

To estimate the local component of the total VMT and scale the 2015 VMT to the two comparison years (2021 and 2022), we calculate the local and non-local vehicle class-specific VMT in 2015 and 2021/2022 from the urban area Vulcan $FFCO_2$ emissions using the FHWA national fuel economy specific to vehicle class. We refer to these VMT values as VMT_{vulc} and calculate this as follows:

$$\mathrm{VMT_{vulc}}\left(v,f,a,y\right)=\frac{\mathrm{Emit_{vulc}}\left(v,f,a,y\right)\times \mathrm{FE_{NA}}\left(v\right)}{\mathrm{EF}\left(f\right)} \tag{2}$$

where $Emit_{vulc}$ represents the Vulcan emissions specific to vehicle class, v, fuel type, f, urban area, a, and year, y. FE_{NA} refers to the national average fuel economy specific to vehicle class and EF refers to the emission factor (tonnes carbon/gallon) specific to gasoline and diesel fuel (8.45 and 10.10 kgCO2/gal, respectively) [50]. The EF values used in Vulcan are 3.7% and 1.1% lower than the EFs used by CT, a difference negligible to the differences shown later in this study. The use of the national average fuel economy requires harmonization of the vehicle class categorization in the Vulcan output (from the 20 017 NEI reporting) and the vehicle class categorization used in the FHWA fuel economy statistics. Details are shown in supplementary information table S1.

With the VMT_{vulc} quantities calculated, the 2015 local component of the total VMT is estimated via the product of 2015 VMT_{AADT} and the ratio of local to non-local VMT_{vulc}. The resulting 2015 total VMT (local + nonlocal) is scaled to the two comparison years (2021 and 2022) via the 2021/2015 and 2022/2015 ratios of the total VMT_{vulc} magnitudes. We refer to these quantities as VMT_{hybrid} to denote the mixture of observed AADT data and the Vulcan $FFCO_2$ output.

This combination (i.e. VMT_{hybrid}) of VMT derived from observed AADT data and VMT inferred from the Vulcan output is to limit the use of the national average fuel economy to the relative measures of local to non-local VMT and VMT growth/decline over time.

In order to calculate the fleet distribution for the Vulcan output, we use the share of VMT_{vulc} by vehicle class to the total VMT_{vulc}. The vehicle class-specific Vulcan fuel economy is calculated as,

$$\mathrm{FE_{vulc}}\left(v\right)=\frac{\mathrm{VMT_{hybrid}}\times EF\left(f\right)}{\mathrm{Emit_{vulc}}\left(v,f,a,y\right)}. \tag{3}$$

On road $FFCO_2$ estimates for Vulcan are limited to the U.S. domain and hence the urban areas compared here are limited to U.S. cities. We further limit analysis to urban areas for which the total on road $FFCO_2$ estimates are greater than 100 tC as it was evident that emission values this small lead to instability in both data products and this only removed 20 and 0 urban areas from version 1 and version 2, respectively. This results in a total of 599 comparable cities in version 1 and 260 comparable cities in version 2. We focus primarily on comparison results for the year 2021 but include the equivalent results for the year 2022 in the supplementary information (figures S4–S8), noting that the comparison conclusions are little impacted by the choice of analysis year.

3. Results

Multiple statistical measures were used to compare CT and Vulcan urban on road $FFCO_2$ emissions (table 1) [51]. The total relative difference (TRD) in 2021 between CT version 1 and Vulcan was 55.7% with the CT total larger than Vulcan by 96.9 MtC. In 2022 the differences were of a similar magnitudes at 57.6% and 101.0 MtC, respectively (table 1). This represents the difference in the sum of the 599 individual paired urban areas. Using the Vulcan urban on road

$FFCO_2$ emission estimate for these urban areas, the total emissions in 2021 account for 30.9% of the U.S. on road $FFCO_2$ total. The same calculation but using the CT total comes to 54.7% of the on road U.S. total in 2021.

When examined at the individual urban scale, the mean difference (MD) between CT version 1 and Vulcan is 161.8 ktC and 168.9 ktC for 2021 and 2022, respectively (table 1, figure 1(a)). A similar magnitude is found for the mean absolute difference (MAD) reflecting the fact that CT version 1 shows consistently larger on road $FFCO_2$ emissions at the scale of individual urban areas than found in Vulcan. In relative terms, the urban on road $FFCO_2$ MRD is 63.1% and 65.2% for 2021 and 2022, respectively (table 1, figure 1(a)). Close ordination of the estimates at the individual urban scale is revealed by the large Pearson correlation value of 0.99 for both years. This suggests a systematic difference between the two datasets (see figure 1(b)). Out of the 599 version 1 urban polygons, 583 are outside of the Vulcan estimated 95% CI.

The version 2 results from CT exhibit somewhat larger total on road $FFCO_2$ emission differences compared to the Vulcan estimates (2021/2022: −119.3/−116.2 MtC) with TRD values for both years of opposite sign (2021/2022: −69.0/−67.3%) to that calculated for version 1 (table 1). At the individual urban scale the differences are somewhat larger and of opposing sign to the results using version 1 (figure 1(a)). The MD between CT version 2 and Vulcan for 2021 and 2022 is −459.0 ktC and −447.1 ktC, values much larger than those found when comparing to version 1. This owes mostly to the fact that the urban areas in version 2 are fewer in number but larger in areal extent than version 1. When measured on a relative basis, however, the differences are of a similar magnitude (though opposite in sign) to the version 1 comparison (2021/2022: −70.4%/−68.6%). Close ordination of the estimates at the individual urban area scale, as with version 1, is revealed by the large Pearson correlation values of 0.99. This again suggests a systematic difference between the two datasets, in spite of the opposing sign difference (figure 1(b)). For version 2, 258 out of 260 urban polygons fall outside the 95% CI bounds of Vulcan.

Comparison can be made to CT version 2 with removal of the 0.3255 activity scale factor (see Methods). Here, the MRD drops by nearly a factor of 2 and of opposing sign to 37.2/39.2 for 2021 and 2022, respectively (table 1). The resulting urban on road $FFCO_2$ emissions are a closer match to the Vulcan results than found in the comparison to CT version 1 (figure 1).

The frequency distribution of the individual urban area relative difference (RD) further emphasizes the systematic offset between the two datasets, the large shift in CT estimation magnitudes from version 1 to version 2, and the impact of the scaling factor used in CT version 2 (figure 2). Close correspondence between the mean and median RD confirms the symmetry of the RD values around the central tendency.

In absolute units, the difference between the Vulcan and CT urban on road $FFCO_2$ emissions is significantly related to the mean magnitude (figure 3(a)). The direction of the relationship is consistent with the MD sign difference between the CT version 1 and version 2 (table 1). When expressed in percentage relative terms, the relationship is far less directional (figure 3(b)), suggesting once again that the urban on road $FFCO_2$ emission differences are most likely methodologically systematic and scale systematically with magnitude.

4. Discussion

In order to better understand the differences between the CT and Vulcan urban on road $FFCO_2$ emissions, we examine key elements of the CT calculation. We first examine estimated urban VMT from both CT and Vulcan, a core component of the activity-based emission estimates. Figure 4 shows the paired comparison of total VMT across the urban areas in CT version 1, CT version 2, and an unscaled CT version 2. Table 2 provides the numerical difference statistics.

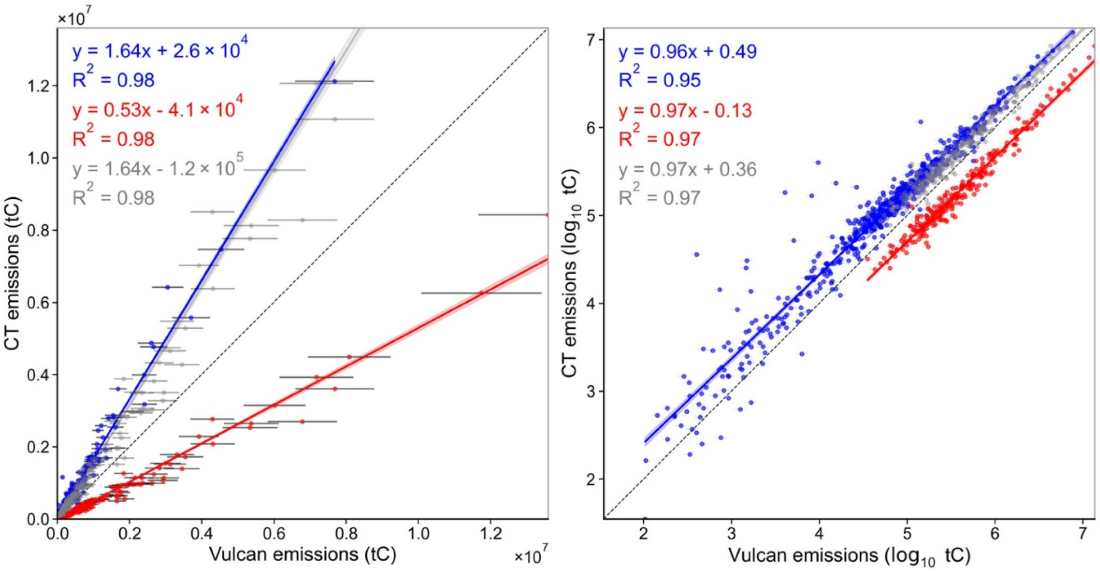
These results confirm that the ML-based prediction of on road urban vehicular activity (VMT) in version 1 and the unscaled version 2 were much closer to the Vulcan estimate than that found in the version 2 estimate (2021 version 1 MRD: 35.7%; version 2 unscaled MRD: 10.8%; version 2 MRD: −92.6%). However, version 1 increases the difference from Vulcan when proceeding from the estimate of vehicular activity to the estimate of $FFCO_2$ emissions, judging from the change in the VMT slope coefficients in figure 4 (slope = 1.29) to the $FFCO_2$ emissions slope coefficient in figure 1 (slope = 1.64). Version 2, by contrast, moves closer to the Vulcan estimates when proceeding from the VMT to $FFCO_2$ emission estimates (VMT slope of 0.39 to $FFCO_2$ slope of 0.53). The same dynamic is reflected in the contrasting MRD paired values in table 2 versus table 1.

The close correspondence between the CT version 1 and unscaled version 2 VMT suggests that the choice to scale the ML-predicted VMT results to more closely match the EDGARv8 country scale on road CO_2 emissions, is the main agent of change between version 1 and version 2 ML-predicted VMT. Most other elements of the ML procedure—input data, training data, and ML algorithm—remain constant across the versions.

However, the polygon choice itself could have an influence on the ML model outcomes. We perform a simple test in which we compute the 2021 Sentinel-2 spectral radiance values using Bands 2/3/4 for a 100 m

Table 1. Vulcan and Climate Trace urban on road $FFCO_2$ emissions comparison statistics. Differences measured as: [Climate Trace—Vulcan].

Version	Year	*N* pairs	Total emit CT (MtC)	Total emit Vulcan (MtC)	TD (MtC)	TRD (%)	MD (ktC)	MAD (ktC)	MRD (%)	MRD StDev (%)	Pearson correlation
Version 1	2021	599	222.7	125.7 (107.7–143.6)	96.9	55.7	161.8	162.6	63.1	32.7	0.99
	2022	599	225.8	124.8 (107.1–142.6)	101.0	57.6	168.9	169.5	65.2	31.9	0.99
Version 2	2021	260	113.4	232.7 (199.7–265.8)	−119.3	−69.0	−459.0	459.1	−70.4	17.6	0.99
	2022	260	114.7	230.9 (198.1–263.7)	−116.2	−67.3	−447.1	447.2	−68.6	17.5	0.99
Version 2 (unscaled)	2021	260	348.3	232.7 (199.7–265.8)	115.6	39.8	444.6	449.5	37.2	19.3	0.99
	2022	260	352.3	230.9 (198.1–263.7)	121.4	41.6	466.9	470.4	39.2	19.0	0.99



(a) (b)

Figure 1. Comparison of 2021 Vulcan and Climate Trace urban on road FFCO$_2$ emissions at the scale of individual urban areas. Included are CT version 1 (blue symbols), CT version 2 (red symbols), unscaled Climate Trace version 2 (grey symbols. Also included are the Vulcan FFCO$_2$ 95% CI values (horizontal bars), linear regression equations and fits (solid lines), the R^2 values, and the 1:1 line (black dashed line).

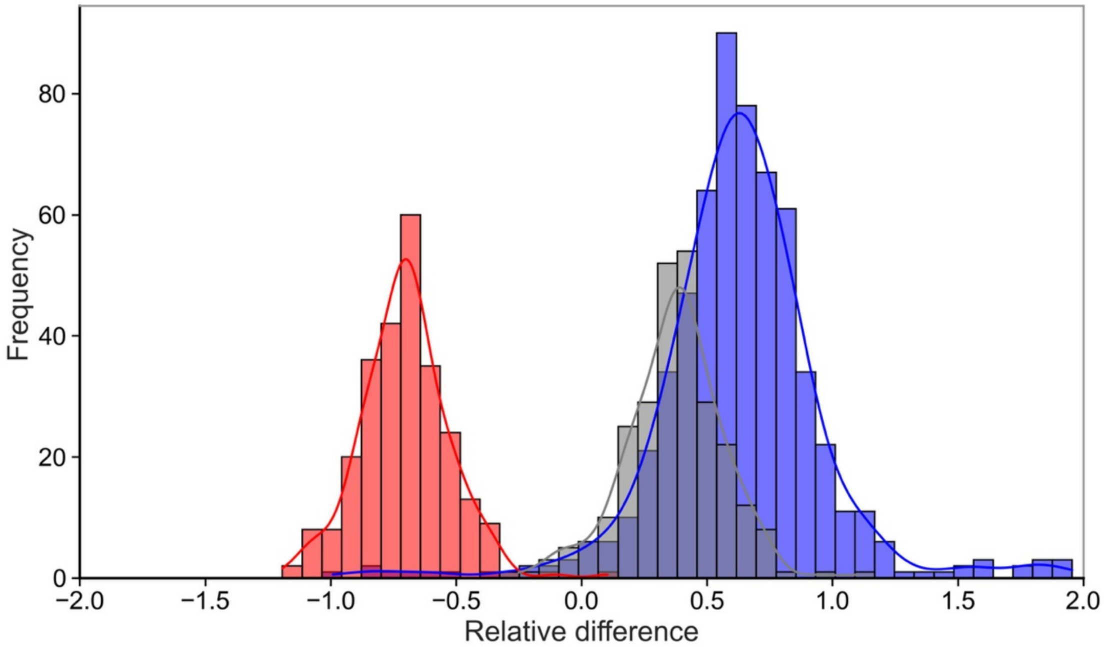


Figure 2. Frequency of the 2021 Climate Trace and Vulcan urban on road FFCO$_2$ emissions relative difference (RD). Blue columns denote CT version 1, red columns denote CT version 2, and grey columns represent the unscaled CT version 2. Relative difference calculated as (CT-Vulcan)/avg (CT, Vulcan).

OSM road buffer in the same manner as performed by CT [52]. The Sentinel-2 median reflectance distribution indicates that the version 1 urban polygons have higher median reflectance (more impervious, built-up surfaces) compared to version 2 (see SI, figure S9). The outer areas display lower and more variable reflectance, confirming that vegetation, bare soil, and non-urban cover dominate that land surface. The clear spectral gradient demonstrates that expanding the CT spatial extent beyond urban cores introduces lower-reflectance, spectrally distinct areas, which likely leads to an AADT underprediction in version 2.

Calculating the CT urban on road FFCO$_2$ emissions from the ML-predicted VMT estimates requires the application of an emission factor, as described in

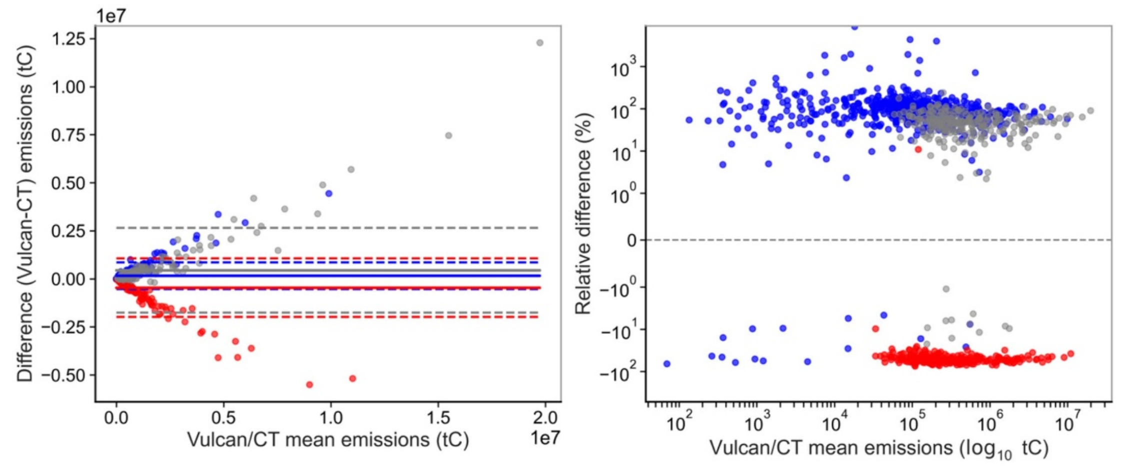


Figure 3. The relationship between the mean Vulcan/Climate Trace urban on road FFCO$_2$ emissions and the 2021 Vulcan versus Climate Trace urban on road FFCO$_2$ emissions RD. Panel b is the logged version of panel a. Blue symbols/lines denote Climate Trace version 1, red symbols/lines denote Climate Trace version 2, and grey symbols/lines denote unscaled Climate Trace version 2. Solid horizontal lines in panel a represent the MRD, dashed lines represent the upper and lower 2-sigma variation of the differences.

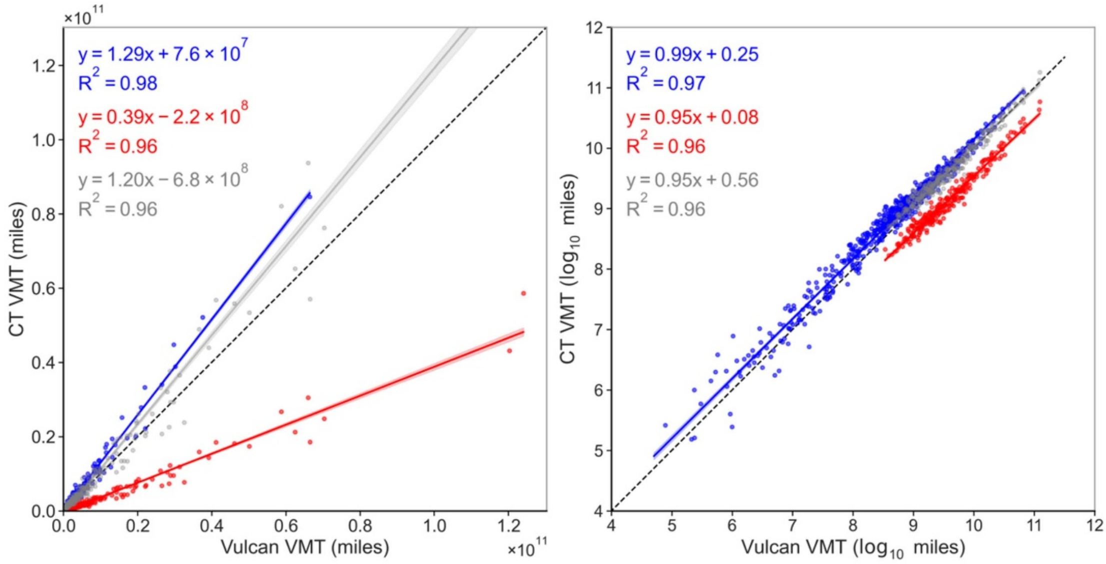


Figure 4. Comparison of 2021 Vulcan and Climate Trace urban on road VMT at the scale of individual urban areas. Included are Climate Trace version 1 (blue symbols, Climate Trace version 2 (red symbols), and the unscaled Climate Trace version 2 (grey symbols). Figure includes the linear regression equations and fit (solid blue line), the R^2 values, and the 1:1 line (black dashed line).

the Methods section. The emission factor is primarily composed of vehicle class-specific national mean fuel economy values combined with an estimate of the carbon content of the fuels consumed. The CT documentation implies that the vehicle class distribution, in both versions 1 and 2, was sourced to the FHWA U.S. national vehicle fleet statistics which provides total VMT by U.S. state in urban versus rural settings across six vehicle classes, in combination with the fuel-specific share within each class, sourced to the CURB tool. Though documentation is not clear (there was no response to multiple inquiries to the CT team), we conclude that the fleet distribution used represents an urban-specific national mean weighted by the state/road class-specific VMT. These shares are as follows (using the FHWA vehicle categories): motorcycles (0.62%), passenger cars (69.5%), light trucks (21.7%), buses (0.49%), single-unit trucks (3.6%), and combination trucks (4.1%).

The Vulcan fleet distribution is estimated at the county scale relying on data submitted by states reflecting vehicle registration records [40]. This results in a varying fleet distribution at the scale of the urban areas analysed here. A comparison of the Vulcan and CT estimates (across the 2 CT versions) demonstrates the underlying variance and mean comparison (figure 5). We focus on the three dominant classes: passenger cars, light trucks, and large

Table 2. Vulcan and Climate Trace urban on road VMT emissions comparison statistics. Differences measured as: [Climate Trace—Vulcan].

Version	Year	*N* pairs	Total VMT CT (Gmiles)	Total VMT Vulcan (Gmiles)	TD (Gmiles)	TRD (%)	MD (Mmiles)	MAD (Mmiles)	MRD (%)	MRD StDev (%)	Pearson correlation
Version 1	2021	599	1556	1172.6	383.4	28.1	640	673.4	35.7	29.8	0.99
	2022	599	1578	1158.3	419.6	30.7	701.7	723.4	38.5	28.8	0.99
Version 2	2021	260	778.8	2143.7	−1365	−93.4	−5249	5249	−92.6	17.4	0.98
	2022	260	785.6	2114.9	−1329.3	−91.7	−5112.6	5112.6	−90.7	17.3	0.98
Version 2 (unscaled)	2021	260	2392.7	2143.7	249.1	11.0	957.9	1631.3	10.8	21.9	0.98
	2022	260	2413.6	2114.9	298.7	13.2	1148.7	1691.6	13.1	21.6	0.98

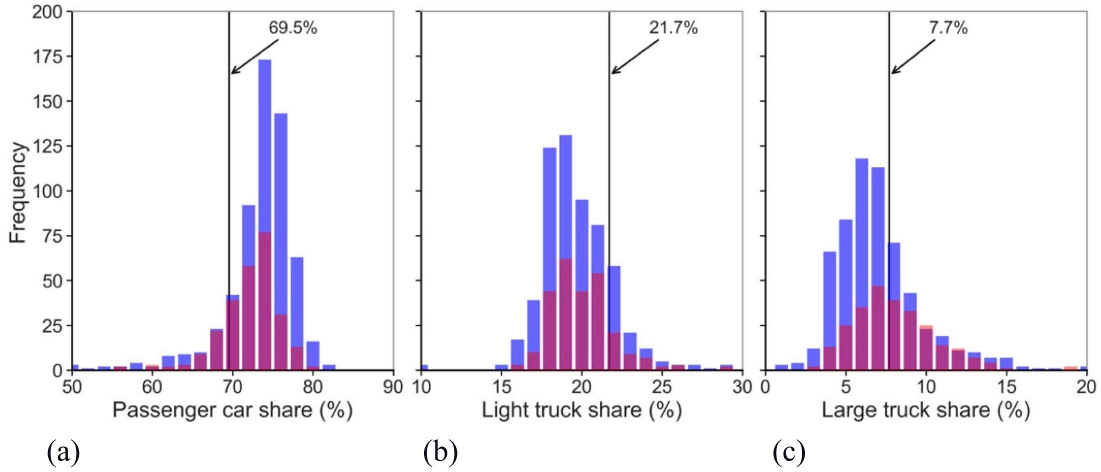


Figure 5. Comparison of the 2021 Vulcan and Climate Trace urban on road VMT vehicle class share at the scale of individual urban areas across the three dominant vehicle types. (a) passenger cars; (b) light trucks; (c) large trucks (single-unit + combination). Blue columns represent the Vulcan vehicle class shares for Climate Trace version 1 cities and red columns represent Vulcan vehicle class shares for Climate Trace version 2 cities. The vertical black lines represent the Climate Trace vehicle class share (a single value as opposed to a distribution) for the U.S.

trucks (single-unit + combination trucks) which together account for 99% of the total national VMT.

The share of VMT by vehicle class indicates that, relative to the Vulcan distribution, CT allocates a smaller proportion of total VMT to passenger vehicles and a larger amount to the light truck and large truck vehicle classes, on average. This is evident for both version 1 and 2, consistent with the stated methodology whereby both versions used the same fleet distribution values.

The application of the fuel economy to the class-specific VMT provides a comparison of the total fuel consumption by vehicle class across the urban areas considered here, and thereby illuminates the impact of the chosen fuel economy values. As noted in Methods, CT version 1/version 2 base the fuel economy values, specific to vehicle class, on the CURB database. CT version 2 further adjusts the fuel economy by surface temperature. Vulcan, using the MOVES-based NEI output, derives county-scale, class-specific fuel economy based on a variety of factors that modify new model fuel economy values. These include vehicle characteristics (e.g. age, weight, engine type), driving conditions (city versus highway, speed), maintenance, and driving behaviour. The emergent fuel economy by vehicle class can be calculated from the Vulcan output, as outlined in the Methods. Figure 6 shows a comparison, across the two CT versions, of the vehicle class specific fuel economy values for Vulcan and CT.

For passenger cars and light trucks, CT mean fuel economy value is lower than the fuel economy apparent in the Vulcan results (resulting in larger $FFCO_2$ emissions, all else being equal). For example, the median passenger car fuel economy value in the version 1 urban areas from Vulcan is 27.2 miles/gallon while the value for CT is 20.9 miles per gallon, a mean difference of 26.5%. The equivalent values for light trucks are 19.8 versus 14.3 miles/gallon (RD: 32.7%), respectively. In the case of the large truck category, CT assigns a higher fuel economy value (9.0 miles/gallon) compared to Vulcan (7.6 miles/gallon) or a relative difference of 16.2%. The version 2 results are similar in this regard to those in version 1 through the application of a temperature adjustment to fuel economy values is evident (values for the three classes are: 20.9 ± 1.6, 14.3 ± 1.1, 9.0 ± 0.7 miles/gallon for passenger cars, light trucks, and large trucks, respectively).

The CURB fuel economy values, used by CT, are noted in the CURB documentation as having been updated in the year 2016 [32]. This suggests that the vintage of fuel economy reporting may provide an explanation for the large fuel economy differences found here. However, examination of the FHWA data between 2016 and 2022 suggests this is likely only a small contributing factor. The FHWA national fuel economy reporting shows an increase of 3.7%, 3.7% and 12.9% for passenger cars, light trucks and large trucks, respectively [53, 54]. This is considerably smaller than the fuel economy differences found here.

The combination of the vehicle class share and the fuel economy dictates, in conjunction with the on road activity (VMT), the differences in the final on road $FFCO_2$ estimation of the two approaches. Since CT uses constant fleet distribution and fuel economy values for each of the vehicle classes, these two constants can be combined into a single emission rate (mass of $FFCO_2$ per vehicle mile travelled). The same can be applied to the Vulcan on road $FFCO_2$ emissions across the urban areas in both CT versions. Figure 7 shows a comparison between these emission rates demonstrating the large variance of emission rates in the Vulcan output compared to the much

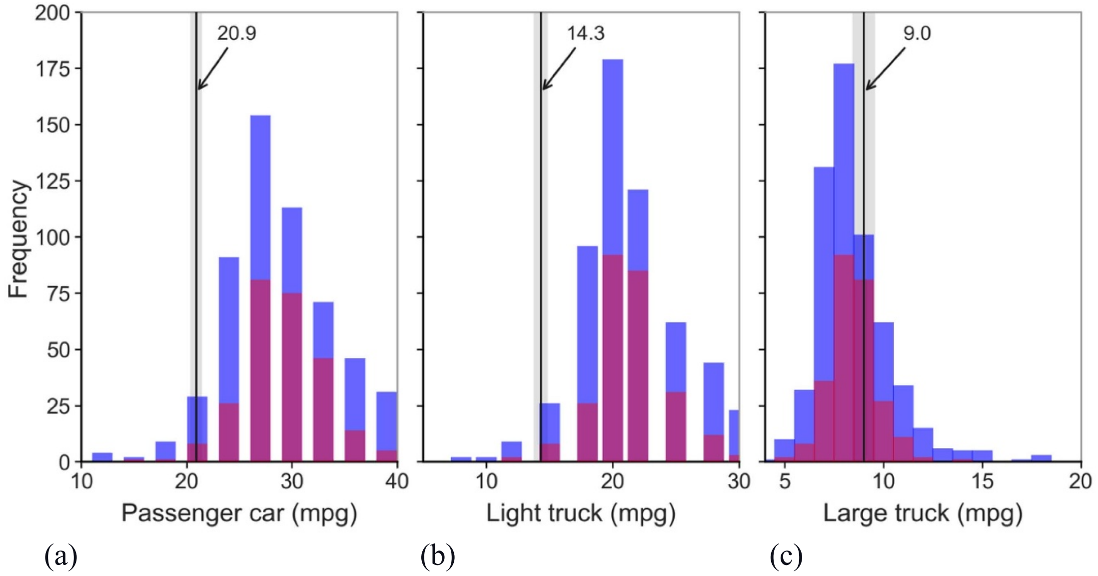


Figure 6. Comparison of 2021 Vulcan and Climate Trace urban fuel economy (MPG) at the scale of individual urban areas across the three dominant vehicle types. (a) passenger cars; (b) light trucks; (c) large trucks (single-unit + combination). Blue columns represent the Vulcan fuel economy for Climate Trace version 1 cities and red columns represent Vulcan fuel economy for Climate Trace version 2 cities. The vertical black lines represent the Climate Trace fuel economy (a single value as opposed to a distribution) for the U.S. grey shaded area represents the variation in Climate Trace fuel economy due to temperature adjustments.

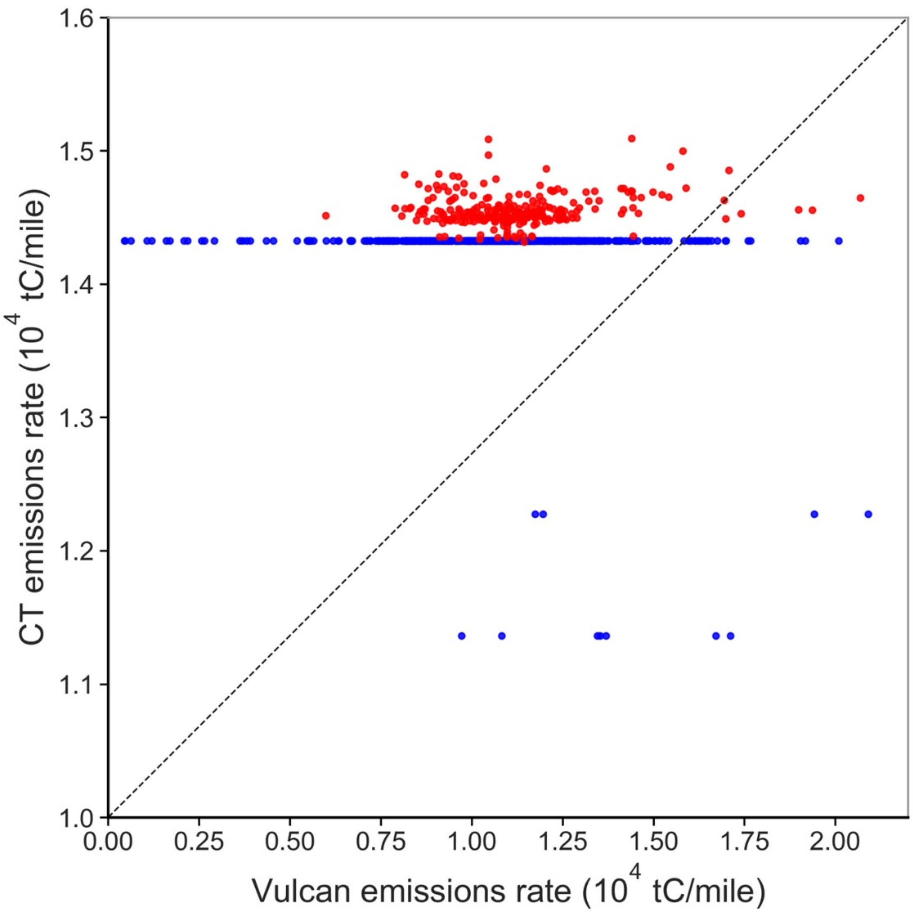


Figure 7. Comparison of 2021 Vulcan and Climate Trace urban on road $FFCO_2$ emission rate values at the scale of individual urban areas. Included are Climate Trace version 1 (blue symbols) and Climate Trace version 2 (red symbols).

narrower range of emission rate values for CT. In the case of version 2, the variation in the on road $FFCO_2$ emission rate is presumably driven by the temperature adjustment applied. In version 1, a few values inexplicably differ from the constant value of 1.43×10^{-4} tC/mile (no temperature adjustment was performed for version 1). The median on road $FFCO_2$ emission rate for Vulcan across the urban areas is 1.08×10^{-4} and 1.10×10^{-4} tC/mile for version 1 and version 2, respectively. Hence, in both versions the CT emission rate is larger than Vulcan with a few individual instances of Vulcan using a larger emission rate (8.8% of urban areas in version 1 and 2.7% in version 2).

The larger emission rate explains in succinct terms, why the version 1 comparison, shows an increased difference when first comparing vehicle activity (VMT, figure 4) and then comparing on road $FFCO_2$ emissions (figure 1). Similarly, while the version 2 comparison of on road activity (VMT) has a much larger difference, this is ameliorated somewhat by the larger CT emission rate. The contribution of these two factors to the emissions differences can be directly compared. In version 1, the MRD of the CT VMT and the Vulcan VMT is 35.7% while the MRD for the emission rate comparison is 30.0%, suggesting near-equivalent contribution to the overall on road $FFCO_2$ emissions difference between CT and Vulcan. In version 2, the same two MRD values are −92.6% and 26.8% for the VMT and emission rate, respectively. This suggests that the ML-predicted VMT of CT is the larger source of the emission difference in version 2.

These results demonstrate the challenge of estimating asset-scale $FFCO_2$ emissions when relying on national mean parameters.

These dynamics are similar to a recent study that compared the CT and Vulcan powerplant $FFCO_2$ emissions for the U.S. [23]. That study found facility-scale MRD values ranging from −0.2% to −51.3% depending upon the year and CT method analysed. Together, power production and on road $FFCO_2$ emissions in urban areas account for 53.2% of the total fossil fuel $FFCO_2$ emitted. Hence, estimation of these emissions at the scale of individual cities is of critical importance. Though the Vulcan emissions cannot be considered ground truth, the extensive analysis of Vulcan, intercomparison with other emission data products, and the consistency with atmospheric measurements, suggest Vulcan as a potent comparison estimate capturing city-to-city variance in the on road emissions and their underlying drivers.

In fairness to the CT effort, estimating GHG emissions at infrastructure resolution for the whole planet is a very ambitious goal. If the aim is to use globally available inputs for the estimation procedures (both AI/ML-based or not), one is faced with a relatively short list of input variables and parameters. Reliance on large spatial averages for portions of the model parameter space, such as was used here for the vehicle fleet distribution and the fuel economy, is understandable. Nevertheless, the trade-off between accuracy and a globally uniform methodology might better lean towards accuracy, given the claims and expectations suggested by the CT outreach materials. We recommend the following, in light of the analysis performed in this study:

(1) Application of the CT on road $FFCO_2$ emissions output at urban scales in policymaking and in carbon cycle science studies must be done so with caution and an understanding that the results will likely not reflect the conditions in specific cities. At best they should be aggregated to larger scale (i.e. national/continental) estimates or include detailed caveats to the use and accuracy of the data released.

(2) For countries with a large share of global on road $FFCO_2$ emissions and that archive higher resolution information, such as the U.S., this information should be ingested into the CT estimation procedure rather than relying on the globally uniform approach currently deployed. In this way, the results would reflect a 'mosaic' of the state-of-the-art techniques.

(3) The CT ML-prediction of on road AADT data (subsequently used to estimate on road $FFCO_2$ emissions) is trained on U.S.-only data. However, the same model is used to predict on road vehicle activity outside the U.S. domain. Given the infrastructural and transportation planning differences across the world, separate national or regional models should be explored for better global representative accuracy.

(4) The test of Sentinel 2 spectral reflectance suggests separating urban and peri-urban areas when training the ML models used by CT is important for delivering accurate on road vehicle activity predictions.

(5) The scaling of the version 2 ML-predicted vehicle activity (VMT) is the main driver of difference to the Vulcan on road $FFCO_2$ emissions. While the impact of this choice cannot be explored for regions outside of the U.S., the results within the U.S. suggest that country-mean on road emissions from EDGAR are very different from the Vulcan urban on road $FFCO_2$ emissions and hence, the use of the scaling factor may be a poor choice.

Publication of the CT on road methodology and data output in the peer-reviewed literature is an essential step prior to releasing scientifically-based data. The CT website page that lists peer-reviewed literature authored or co-authored by CT coalition members includes no on road-related literature. Submitting the CT on road $FFCO_2$ emissions methodology to peer-reviewed scrutiny will ensure that scientific rigor, appropriate caveats, uncertainties, and comprehensive documentation are achieved.

Acknowledgments

This work was made possible through support from the National Oceanic and Atmospheric Administration grant NA19OAR4310167.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Supplementary data 1 available at https://doi.org/10.1088/1748-9326/ae6355/data1.

Ethical statement

This research did not involve humans or animals.

Author contributions

Kevin R Gurney 0000-0001-9218-7164
Conceptualization (lead), Formal analysis (lead), Funding acquisition (lead), Investigation (lead), Methodology (lead), Project administration (lead), Supervision (lead), Visualization (equal), Writing – original draft (lead), Writing – review & editing (equal)

Bilal Aslam 0000-0001-7308-5285
Data curation (equal), Formal analysis (equal), Visualization (lead), Writing – review & editing (equal)

Pawlok Dass 0000-0003-3957-4055
Data curation (equal), Methodology (equal), Software (equal)

References

[1] Intergovernmental Panel On Climate Change (Ipcc) 2023 *Climate Change 2021—The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press) (https://doi.org/10.1017/9781009157896)

[2] Lindroth A and Tranvik L 2021 Accounting for all territorial emissions and sinks is important for development of climate mitigation policies *Carbon Balance Manage.* **16** s13021–021–00173–8

[3] National Academies of Sciences 2022 *Engineering, and Medicine. Greenhouse Gas Emissions Information for Decision Making: A Framework Going Forward* (National Academies Press) (https://doi.org/10.17226/26641)

[4] Yin L, Sharifi A, Liqiao H and Jinyu C 2022 Urban carbon accounting: an overview *Urban Clim.* **44** 101195

[5] Yona L, Cashore B and Bradford M A 2022 Factors influencing the development and implementation of national greenhouse gas inventory methodologies *Policy Des. Pract.* **5** 197–225

[6] Guevara M *et al* 2023 Towards near-real-time air pollutant and greenhouse gas emissions: lessons learned from multiple estimates during the COVID-19 pandemic *Atmos. Chem. Phys.* **23** 8081–101

[7] Gurney K R *et al* 2015 Climate change: track urban emissions on a human scale *Nature* **525** 179–81

[8] Asefi-Najafabady S, Rayner P J, Gurney K R, McRobert A, Song Y, Coltin K, Huang J, Elvidge C and Baugh K 2014 A multiyear, global gridded fossil fuel CO_2 emission data product: evaluation and analysis of results *J. Geophys. Res. Atmos.* **119** 10–213

[9] 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 Syst. Sci. Data* **10** 87–107

[10] Liu Z *et al* 2020 Carbon monitor, a near-real-time daily dataset of global CO2 emission from fossil fuel and cement production *Sci. Data* **7** 392

[11] 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 Syst. Sci. Data* **12** 3413–42

[12] Gilfillan D and Marland G 2021 CDIAC-FF: global and national CO_2 emissions from fossil fuel combustion and cement manufacture: 1751–2017 *Earth Syst. Sci. Data* **13** 1667–80

[13] Crippa M *et al* 2023 Insights on the spatial distribution of global, national and sub-national GHG emissions in EDGARv8.0 *Earth Syst. Sci. Data* **15** 2667–94

[14] Friedlingstein P *et al* 2025 Global carbon budget 2024 *Earth Syst. Sci. Data* **17** 965–1039

[15] Patarasuk R, Gurney K R, O'Keeffe D, Song Y, Huang J, Rao P, Buchert M, Lin J C, Mendoza D and Ehleringer J R 2016 Urban high-resolution fossil fuel CO_2 emissions quantification and exploration of emission drivers for potential policy applications *Urban Ecosyst.* **19** 1013–39

[16] Gurney K R, Patarasuk R, Liang J, Song Y, O'Keeffe D, Rao P, Whetstone J R, Duren R M, Eldering A and Miller C 2019 The Hestia fossil fuel CO2 emissions data product for the Los Angeles megacity (Hestia-LA) *Earth Syst. Sci. Data* **11** 1309–35

[17] Keller E D, Hilton T W, Benson A, Karalliyadda S, Xie S, Gurney K R and Turnbull J C 2023 Mahuika-Auckland: a spatially and temporally resolved fossil fuel CO_2 emissions data product for Auckland, New Zealand *Geosci. Data J.* **10** 347–67

[18] Stagakis S, Feigenwinter C, Vogt R and Kalberer M 2023 A high-resolution monitoring approach of urban CO_2 fluxes. Part 1-bottom-up model development *Sci. Total Environ.* **858** 160216

[19] Gore A 2022 Measure emissions to manage emissions *Science* **378** 455

[20] Climate Trace 2025 About the coalition—climate TRACE (available at: https://staging.c10e.org/about)

[21] Climate Trace 2023 Climate TRACE unveils open emissions database of more than 352 million assets (Climate TRACE) (available at: https://climatetrace.org)

[22] Ma M 2023 Al gore-backed group has a tool to decarbonize supply chains (Bloomberg.com)

[23] Gurney K R, Aslam B, Dass P, Gawuc L, Hocking T, Barber J J and Kato A 2024 Assessment of the Climate Trace global powerplant CO_2 emissions *Environ. Res. Lett.* **19** 114062

[24] International Energy Agency 2025 Electricity 2025 (available at: https://iea.blob.core.windows.net/assets/0f028d5f-26b1-47ca-ad2a-5ca3103d070a/Electricity2025.pdf)

[25] Gurney K R, Dass P, Kato A, Gawuc L, Aslam B and Sun H 2025 Vulcan version 4.0 high-resolution annual carbon dioxide emissions in the U.S. for the 2010–2022 time period *Sci. Data* **12** 1946

[26] Gurney K R, Liang J, Patarasuk R, O'Keeffe D, Huang J, Hutchins M, Lauvaux T, Turnbull J C and Shepson P B 2017 Reconciling the differences between a bottom-up and inverse-estimated $FFCO_2$ emissions estimate in a large US urban area *Elementa* **5** 44

[27] Dass P, Gurney K R, Song Y, Kato A, Aslam B and Gawuc L 2026 Advancing high-resolution U.S. on road fossil fuel carbon dioxide emissions estimates *J. Clean. Prod.* **555** 148150

[28] Kott T *et al* 2024 Transportation sector: global road emissions (available at: https://github.com/climatetracecoalition/methodology-documents/tree/main/2025/Transportation)

[29] Rollend D, Foster K, Kott T M, Mocharla R, Muñoz R, Fendley N, Ashcraft C, Willard F, Reilly E P and Hughes M 2023 Machine learning for activity-based road transportation emissions estimation *Environ. Data Sci.* **2** e38

[30] US FHWA 2020 HPMS public release of geospatial data in shapefile format (Policy | Federal Highway Administration (available at: www.fhwa.dot.gov/policyinformation/hpms/shapefiles_2017.cfm)

[31] Haklay M and Weber P 2008 Openstreetmap: user-generated street maps *IEEE Pervasive Comput.* **7** 12–18

[32] World Bank Group 2019 CURB: climate action for urban sustainability

[33] US FHWA 2023 *Table VM-2—Highway Statistics 2021* (Policy | Federal Highway Administration)

[34] US FHWA 2024 *Table VM-2—Highway Statistics 2022* (Policy | Federal Highway Administration)

[35] US EPA 2016 2014 national emissions inventory (NEI) data (available at: www.epa.gov/air-emissions-inventories/2014-national-emissions-inventory-nei-data)

[36] US EPA 2017 2017 national emissions inventory (NEI) data (available at: www.epa.gov/air-emissions-inventories/2017-national-emissions-inventory-nei-data)

[37] US EPA 2018 2014 national emissions inventory, Version 2 Technical Support Document

[38] US EPA 2021 2017 national emissions inventory: January 2021 updated release, technical support document

[39] US EPA, O 2016 Latest Version of MOtor Vehicle Emission Simulator (MOVES) (available at: www.epa.gov/moves/latest-version-motor-vehicle-emission-simulator-moves)

[40] US EPA 2018 Instructions for submitting on road inputs for the 2017 NEI

[41] CARB 2014 EMFAC2014 volume I—user's guide v1.0.7 April 30, 2014 (available at: https://ww2.arb.ca.gov/sites/default/files/2023-01/emfac2014-vol1-users-guide-052015.pdf)

[42] US EIA. U.S 2025 Energy Information Administration—EIA—independent statistics and analysis State Energy Data System (SEDS)

[43] Gately C K and Hutyra L R 2017 Large uncertainties in urban-scale carbon emissions *J. Geophys. Res. Atmos.* **122** 11–242

[44] Gurney K R, Liang J, Patarasuk R, Song Y, Huang J and Roest G 2020 The vulcan version 3.0 high-resolution fossil fuel CO_2 emissions for the United States *J. Geophys. Res. Atmos.* **125** e2020JD032974

[45] Gately C K, Hutyra L R, Wing I Sue and Brondfield M N 2013 A bottom up approach to on-road CO2 emissions estimates: improved spatial accuracy and applications for regional planning *Environ. Sci. Technol.* **47** 2423–30

[46] Mendoza D, Gurney K Robert, Geethakumar S, Chandrasekaran V, Zhou Y and Razlivanov I 2013 Implications of uncertainty on regional CO2 mitigation policies for the U.S. onroad sector based on a high-resolution emissions estimate *Energy Policy* **55** 386–95

[47] Basu S, Lehman S J, Miller J B, Andrews A E, Sweeney C, Gurney K R, Xu X, Southon J and Tans P P 2020 Estimating US fossil fuel CO_2 emissions from measurements of^{14} C in atmospheric CO_2 *Proc. Natl Acad. Sci.* **117** 13300–7

[48] Kato A, Gurney K R, Roest G S and Dass P 2023 Exploring differences in $FFCO_2$ emissions in the United States: comparison of the Vulcan data product and the EPA national GHG inventory *Environ. Res. Lett.* **18** 124043

[49] US FHWA 2023 *Table VM-4—Highway Statistics 2021* (Policy | Federal Highway Administration) (available at: www.fhwa.dot.gov/policyinformation/statistics/2021/vm4.cfm)

[50] US FHWA 2025 *Table VM-1—Highway Statistics 2021* (Policy | Federal Highway Administration) (available at: www.fhwa.dot.gov/policyinformation/statistics/2021/vm1.cfm)

[51] Gurney K R, Dass P, Kato A, Mitra B and Nematchoua M K 2024 Scope 2 estimates of carbon dioxide emissions from electricity consumption at the US census block group scale *Sci. Data* **11** 1344

[52] Drusch M *et al* 2012 Sentinel-2: ESA's optical high-resolution mission for GMES operational services *Remote Sens. Environ.* **120** 25–36

[53] US FHWA 2018 *Table VM-1—Highway Statistics 2016* (Policy | Federal Highway Administration) (available at: www.fhwa.dot.gov/policyinformation/statistics/2016/vm1.cfm)

[54] US FHWA 2025 *Table VM-1—Highway Statistics 2022* (Policy | Federal Highway Administration) (available at: www.fhwa.dot.gov/policyinformation/statistics/2016/vm1.cfm)