

Transportation Sector - Road Emissions

Estimation Methodology

Derek Rollend¹, Tomek Kott¹, Kevin Foster¹, Rohita Mocharla¹, Rai Muñoz¹, Neil Fendley,¹ Chace Ashcraft¹ and Marisa Hughes¹

1) The Johns Hopkins University Applied Physics Laboratory (JHU/APL)



1. Introduction

Transportation contributed 27% of anthropogenic greenhouse gas (GHG) emissions in the U.S. for 2020, higher than any other sector, and 12.6% of all global GHG emissions in 2019 (U.S. EPA 2022a; World Resource Institute 2022). The primary source of transportation sector emissions are on-road vehicles, accounting for approximately 74% of global transportation emissions in 2018 (International Energy Agency 2022). Quantifying the distribution of on-road transportation emissions and creating timely emissions inventories are vital to identify trends, track mitigation efforts, and inform policy decisions.

Previous efforts have developed detailed bottom-up on-road emission inventories for the U.S. (Gately, Hutyra and Wing 2019; Gurney *et al.*, 2020), but do not easily extend globally due to the reliance on vehicle traffic and road data that is not always readily available. The Emissions Database for Global Atmospheric Research (EDGAR) provides a global inventory for transportation that uses road density as a proxy to spatially distribute emissions (Crippa *et al.*, 2019). However, some emission estimates for urban centers in EDGAR deviated from other bottom-up inventories by 500%, indicating that road density is not a sufficient proxy for global high-resolution inventories (Gately, Hutyra and Wing 2015). Carbon Monitor is a global emissions inventory that utilizes a variety of activity data to estimate daily GHG emissions, however the reliance on proprietary traffic data in the ground transportation sector limits the ability to extend to locations where this data is not available (Liu *et al.*, 2020). Other methods have used machine learning (ML) to directly predict emissions, but their ability to generalize globally is unclear (Mukherjee *et al.*, 2021; Scheibenreif, Mommert and Borth 2021).

JHU/APL has developed an approach to estimate road transportation emissions, and applied this methodology to 500 prioritized global cities, representing some of the largest cities in the world. Our “hybrid” algorithm leverages the strengths of ML applied to remote sensing data, in addition to incorporating region-specific emissions factors data to create scalable and transparent emissions estimations globally. A detailed description of our method, data sources, and validation results is contained herein.

2. Data and Methods

2.1. Overview

Our hybrid emissions estimation method was primarily composed of two parts: a set of ML models trained to predict road transport activity, and an emissions factors (EFs) pipeline that converts activity predictions to emissions estimates. This approach combines the strengths of satellite imagery and ML with traditional “bottom-up” emissions inventories that directly incorporate vehicle fleet mix, fuel efficiency, and other EF data. Having these two, primarily independent parts affords continuous improvement of each as newer and better data become available. A high-level system architecture is shown below in Figure 1.

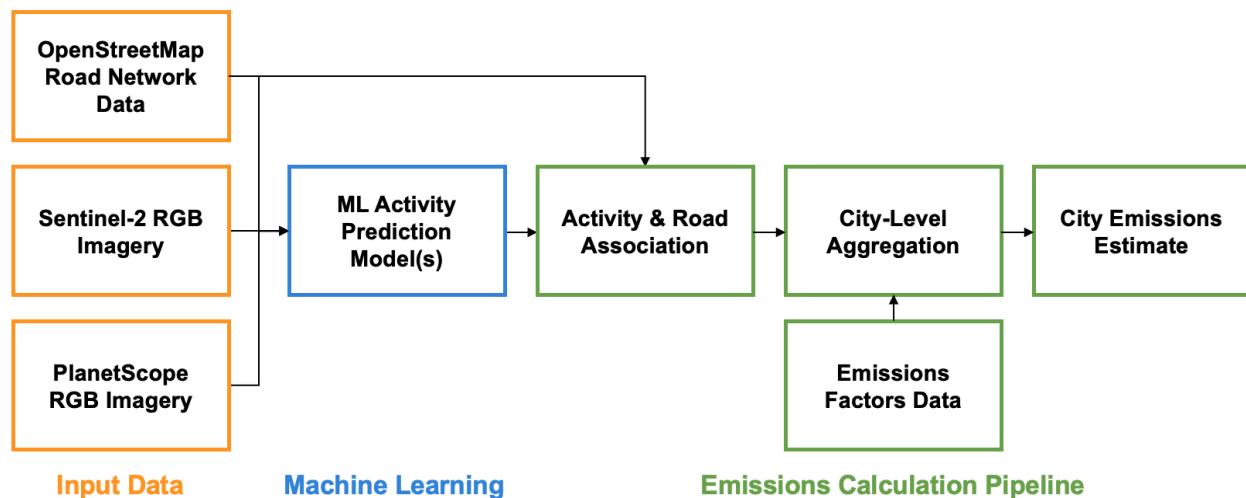


Figure 1 Emissions estimation architecture overview. Description of figure is provided in the text.

Remote sensing and geospatial data were fed to ML models to provide more complete and higher frequency estimates of average annual daily traffic (AADT). These AADT predictions were then assigned to their corresponding road segment based on the known geographic location of the underlying road network. Emissions factors were computed a priori from a database of road and vehicle-related data for a specific region, assigning EF values to each type of road in a city. Estimated AADT was converted to total vehicle kilometers traveled (VKT) using the known length of each road segment, and then multiplied by the appropriate EF for this road type. This process was repeated and summed over all road segments in a city to calculate the final, total emissions estimate for that city for 2021. This process is shown below in Figure 2.

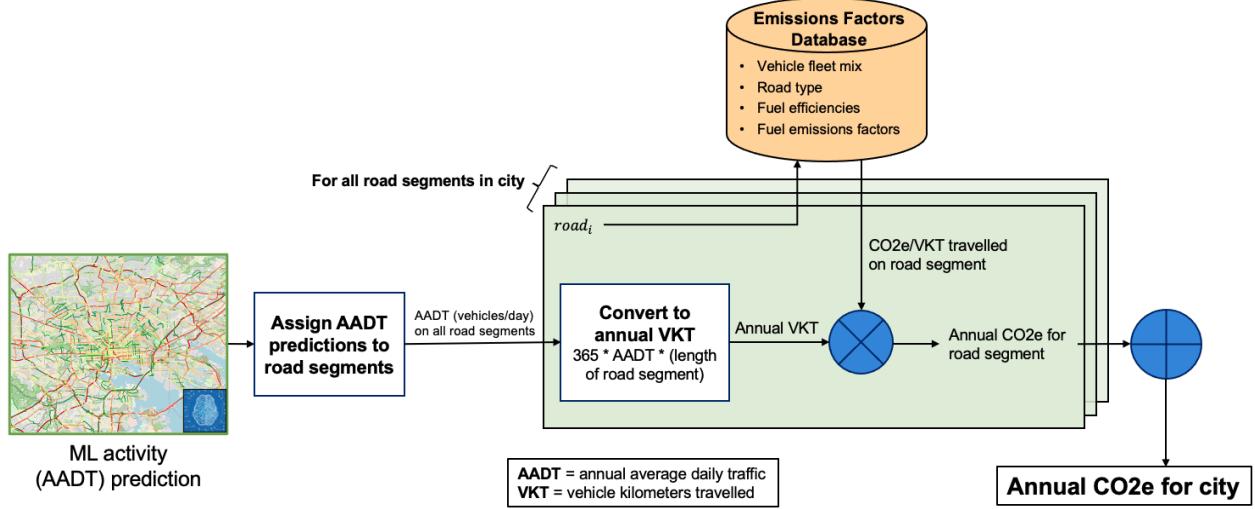


Figure 2 Converting ML-predicted road traffic to emissions.

2.2. Data

2.2.1. City Selection

To prioritize the set of 500 cities, we utilized the European Union Joint Research Center Global Human Settlement Layer Urban Centers Database (GHSL-UCDB) dataset (Florczyk *et al.*, 2019) for a globally consistent representation of city extent. This database contains approximately 13,000 cities worldwide, and utilizes a definition of city/urban center based on population density and built up area. Specifically, an urban center was defined as “the spatially-generalized high-density clusters of contiguous grid cells of 1 km² with a density of at least 1,500 inhabitants per km² of land surface or at least 50% built-up surface share per km² of land surface, and a minimum population of 50,000.” (Florczyk *et al.*, 2019). Due to this definition, city geometries in UCDB often have significantly different shapes and sizes as compared to official administrative bounds, e.g., from OpenStreetMap (OSM; OpenStreetMap contributors 2021) or Global Administrative Areas (GADM) (Global Administrative Areas 2022). Examples of these differences are shown below in Figure 3-Figure 5.

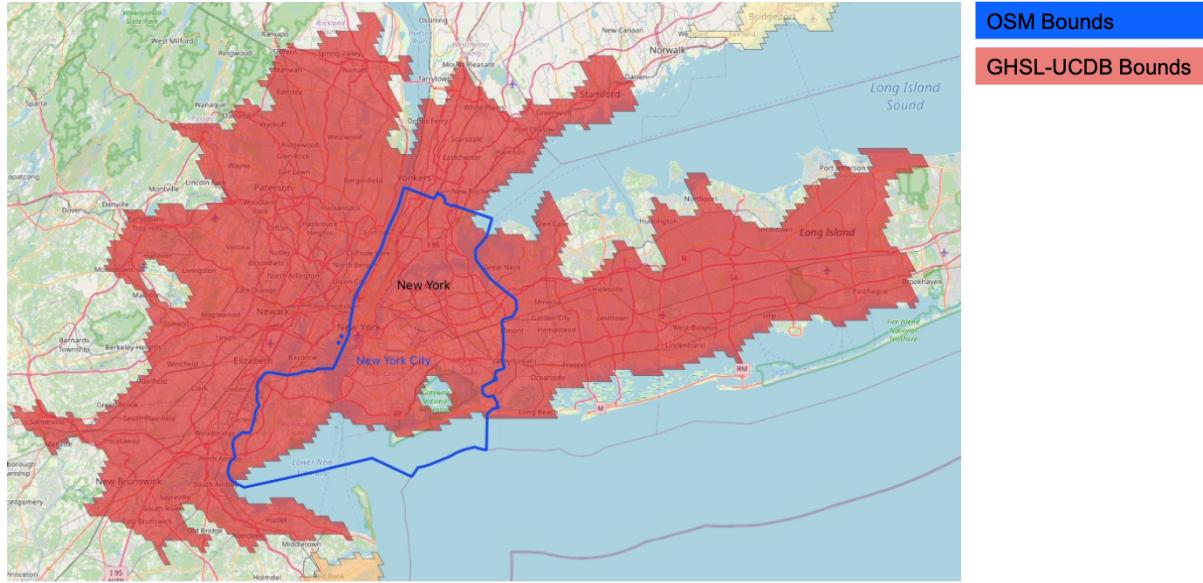


Figure 3 Comparison of OSM administrative bounds (blue boundary) and GHSL-UCDB city bounds (red area) for New York City, USA.

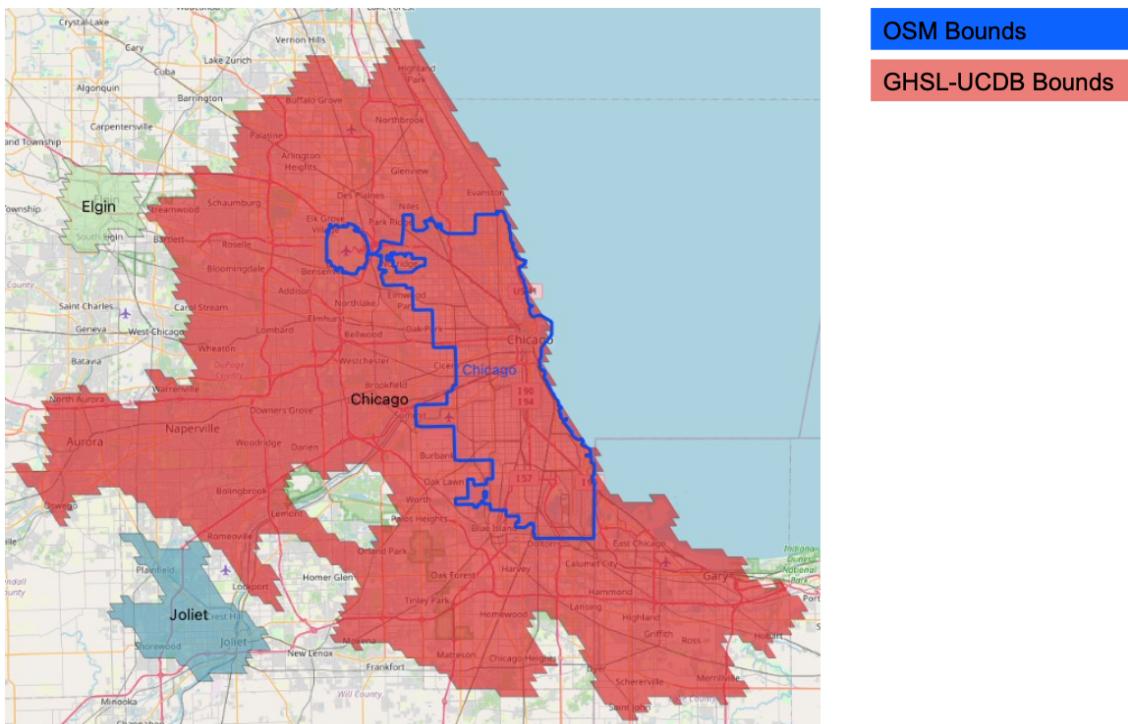


Figure 4 Comparison of OSM administrative bounds (blue boundary) and GHSL-UCDB city bounds (red area) for Chicago, USA.

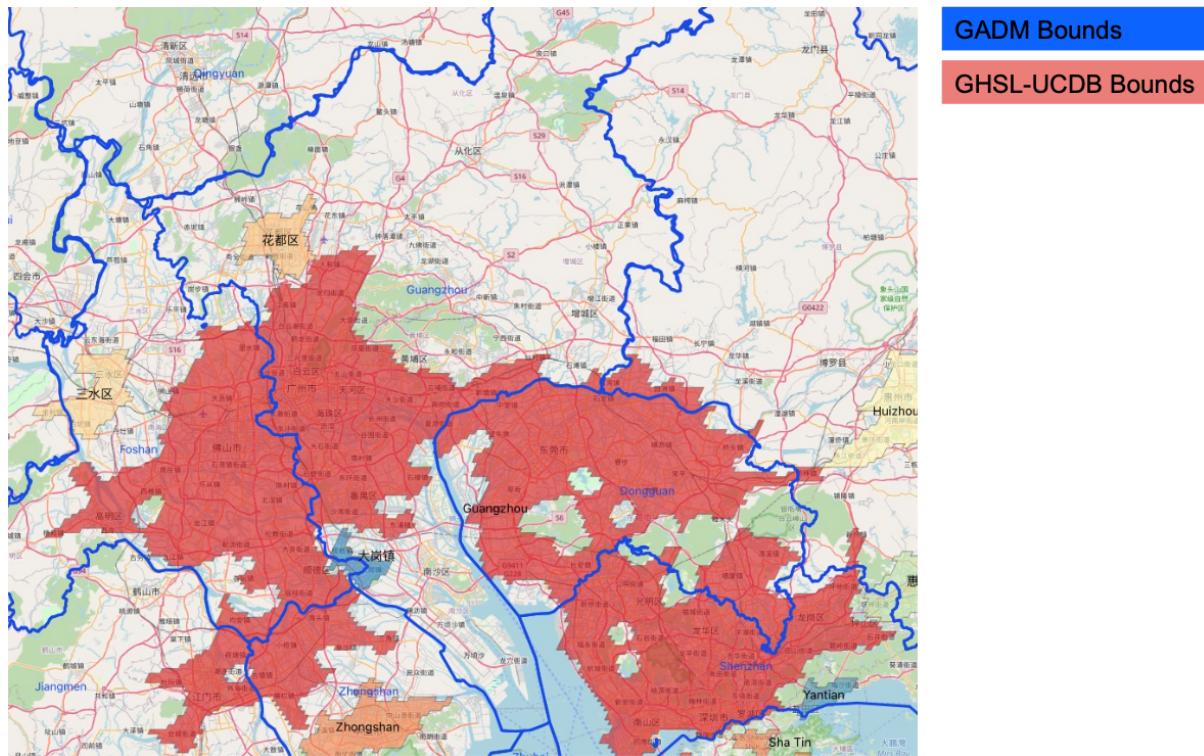


Figure 5 Comparison of OSM administrative bounds (blue boundary) and GHSL-UCDB city bounds (red area) for Guangzhou, China. The OSM boundary is in the middle of the figure.

UCDB spatially combines urban center bounds with a variety of associated metadata related to geography, socio-economic, environment, disaster risk, and sustainable development goals. This metadata includes EDGAR V5.0 (Crippa *et al.*, 2019) emissions estimates within urban center bounds for 1975, 1990, 2000, and 2015. We used the 2015 transport sector total CO₂ emissions from non-short-cycle organic fuels (fossil fuels, “CO2_excl_short-cycle_org_C” in EDGAR) to sort and select the largest 500 cities for this work. The distribution of the selected cities across continents is shown below in Table 1.

Table 1 Region representation in the top 500 cities.

Region	Proportion of Top 500 Cities
Asia	42.6%
Europe	18.8%
North America	17.8%
Latin America and the Caribbean	11.2%
Africa	8.2%
Oceania	1.4%

2.2.2. Visual Satellite Imagery

Remote sensing imagery from the Sentinel-2A/B satellites and PlanetScope constellation were used as input data in our ML modeling approach to predict road transportation activity. Below, a description of each satellite and imagery processing steps are provided.

The European Space Agency's (ESA) Sentinel-2 mission comprises two satellites- Sentinel-2A, launched in 2015, and Sentinel-2B, launched in 2017 (Main-Knorn *et al.*, 2017). Each Sentinel-2 satellite has a 10-day revisit time with a 5-day combined revisit. Both satellites are equipped with a multispectral (MSI) instrument which provides 13 spectral band measurements, blue to shortwave infrared (SWIR) wavelengths (~442 nm to ~2202 nm) reflected radiance. We used the Sentinel-2 Level-2A product at 10 m x 10 m resolution, using bands 4 (red), 3 (green), and 2 (blue) (Drusch *et al.*, 2012).

Planet Lab's PlanetScope satellite constellation consists of approximately 130 individual satellites, called "Doves", with the first launch of this constellation in 2014 (Planet Labs 2022). Each PlanetScope satellite images the earth's surface in the blue, green, red and near-infrared (NIR) wavelengths (~450 nm to ~880 nm). We acquired PlanetScope ~3 m monthly and quarterly mosaics (Planet Labs 2022) over the various geographic extents of the cities we used for training, validation, and global inference.

2.2.3. Road Network Data

Using the open-source tools Osmium (Topf 2022) and OSMnx (Boeing 2017) and custom filters, OpenStreetMap data for each city was cropped and filtered to the set of roads that carry normal vehicular traffic. A multi-directed graph of the road network was created from that data, and the total edge length (meaning the road length for each direction of traffic is counted separately) was computed to yield the reported road network length (in kilometers). Example road networks are visualized in Figure 6.

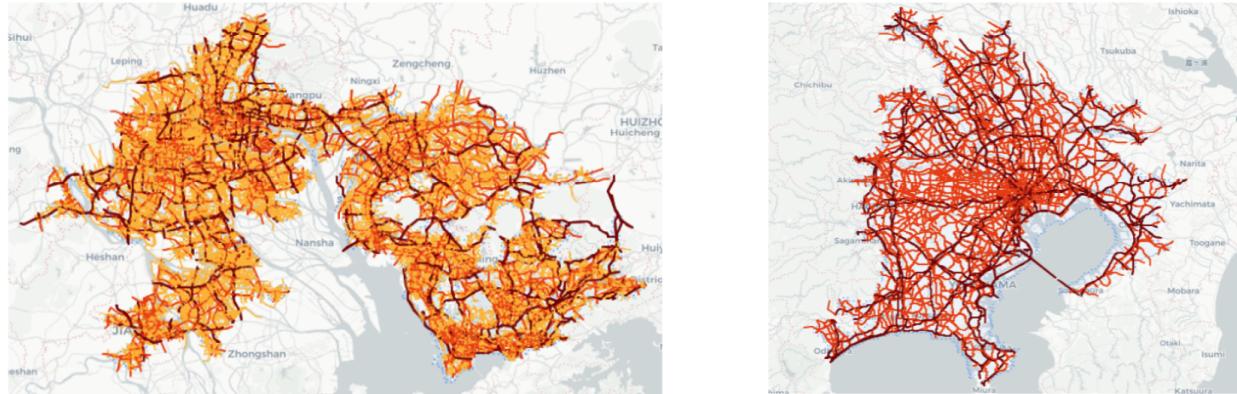


Figure 6 Example road network data for Guangzhou, China (left) and Tokyo, Japan (right). Highways are in dark red, arterial roads in orange, and local roads in yellow. Local roads are intentionally not displayed for Tokyo due to their large count.

Plots of both road network length and density are provided in Figure 7-Figure 9 (note that the term “edge” is equivalent to “road”, and comes from a description of the underlying road graph network with nodes and edges). In Figure 7 and Figure 8, only the top 20 cities for each metric were displayed for visual clarity. Future analysis may examine the dependence of total estimated emissions on road network length and road network density.

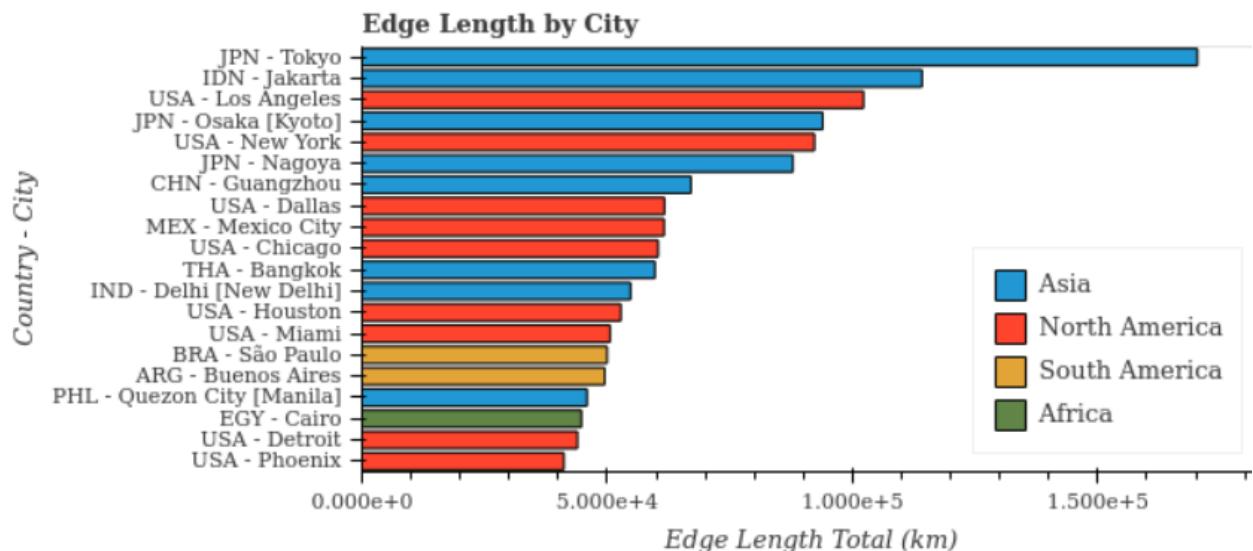


Figure 7 Total road length (Edge Length) for the top 20 cities by region- Asia (blue), North America (red), South America (orange) and Africa (green).

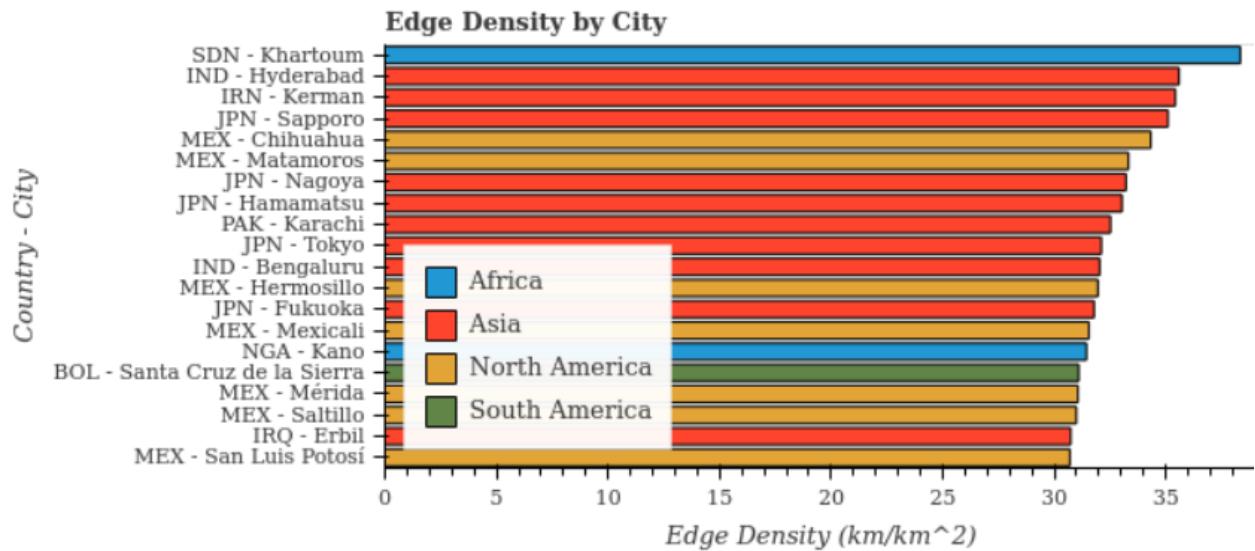


Figure 8 Road density (Edge Density, km/km²) for the top 20 cities by region- Africa (blue), Asia (red), North America (orange) and South America (green).

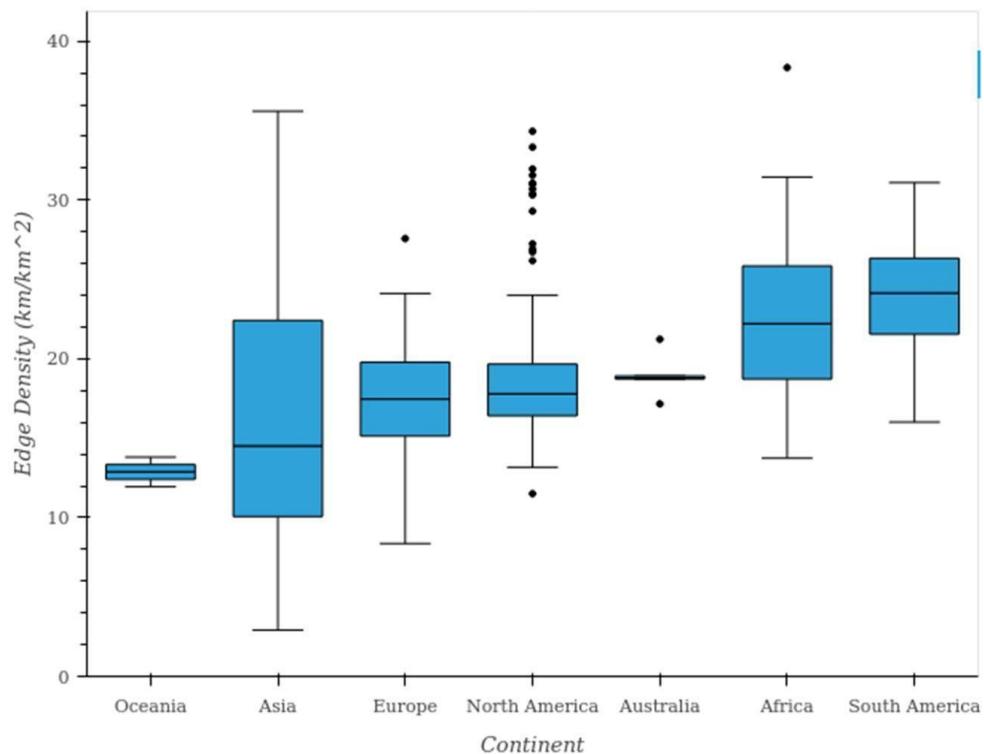


Figure 9 Distribution of road density (Edge Density, km/km²) by continent.

2.2.4. Emissions Factors

Calculating emissions factors for a city was a complex process, due to the fact that a city can actually contain potentially hundreds of thousands of individual road segments, which might be considered “sub-assets”. Transportation emissions factors are dependent on many variables, including (but not limited to) road category, vehicle type, fuel type, and fuel efficiency. Data collection for each of these variables across 500 cities was a significant undertaking. Thus, the initial version of estimated emissions factors focused on collecting data at the country level for the 86 countries in which the top 500 cities were located. A detailed description of the city-level emissions factor calculation is provided in section 2.2.9. Sources for each type of data required for the emissions factor calculation are shown below in Table 2, with descriptions of each data type provided in subsequent sections.

Table 2 Primary emissions factors data sources used.

Data Type	Source(s)
Road Segment Type	OpenStreetMap (OpenStreetMap contributors 2022)
Vehicle Fleet Mix	Various, see Vehicle Fleet Mix
Fuel Type	CURB (World Bank Group 2019)
Fuel Efficiencies	CURB
GHG Emissions Factors	U.S. EPA GHG Emissions Factors Hub (U.S. EPA 2022b)

2.2.5. Road Segment Type

Road segment types (categories) were derived from OpenStreetMap data for all 500 cities. The current supported road types are highway, arterial, and local, which were chosen to align with other similar emissions inventories and traffic-related databases. The mapping between these road types and their respective OSM tags is provided in Table 3. Road type categorization is important in the emissions factor calculation for a given road segment as other emissions factors variables, including vehicle fleet mix and fuel efficiency, can vary significantly across different types of roads.

Table 3 Road segment type association with OpenStreetMap tags.

Road Class	OpenStreetMap Tags
Highway	motorway, motorway_link, trunk, trunk_link
Arterial	primary, primary_link, secondary, secondary_link
Local	tertiary, tertiary_link, residential, living_street, unclassified

2.2.6. Vehicle Fleet Mix

Vehicle fleet mix refers to the distribution of total vehicles in a given country across various vehicle types. The supported vehicle types were: passenger cars, light duty trucks, single unit trucks, combination trucks, motorcycles, and buses. Country-specific vehicle fleet data was used for 36 countries, while a U.S. urban area average derived from U.S. Federal Highway Administration (FHWA) data (U.S. FHWA 2020) was used for the remaining 50 countries until further country-specific data sources are identified.

A full listing of the 36 countries with country-specific data and their respective sources was provided in the supplementary material (Table 10). For many of these countries, the vehicle fleet mix was derived from vehicle registration data. The same type of vehicle registration data may be available at the city-level, which was left as a future data exploration and collection task. Due to differences in the vehicle type taxonomy across countries, vehicle types were manually re-categorized to the supported vehicle types as well as possible. Vehicle fleet mix values are currently the same across all supported road types but will be updated as sources of road type-specific data are identified.

2.2.7. Fuel Type and Efficiencies

Due to the fact that different fuel types have different emissions factors, it is important to know the relative mix of fuel types for each type of vehicle traveling on a given road segment. The types of supported fuels are: gasoline, diesel, compressed natural gas (CNG), liquefied petroleum gas (LPG), plug-in hybrid, battery electric vehicle (BEV), and other fuels (e.g., biogas, ethanol). The primary source of this data is the Climate Action for Urban Sustainability (CURB) tool (World Bank Group 2019), which provides a global database of fuel type mix by country. Future updates may include updated country or city-specific fuel type data.

CURB was also the primary source of fuel efficiency data for all 86 countries. CURB fuel efficiency values are reported in units of kilometers per liter and were extracted for all supported fuel and vehicle types described above. Fuel efficiencies were the same across all supported road types (highway, arterial, and local) in this release, but may be continuously updated as better

country or city-specific datasets are located.

2.2.8. Vehicle Greenhouse Gas (GHG) Emissions Factors

GHG emissions factors refer to how much of a given gas is emitted per unit of fuel burned and varies by fuel type. Our data focuses on carbon dioxide (CO₂), nitrous oxide (N₂O), and methane (CH₄) emissions factors, using data from the U.S. Environmental Protection Agency (U.S. EPA 2022). For nitrous oxide and methane, the emissions factors for each gas were given in units of grams of each gas per mile driven. This was different from the data for carbon dioxide, which was given as grams per liter. To normalize all greenhouse gas emissions factors to “grams per liter”, we used fuel efficiency data (given in “liters per km”) to generate data for nitrous oxide and methane as grams per liter.

2.2.9. Road Segment Type Emissions and Emissions Factors

Total emissions were first computed for each road segment within a city, and then summed to estimate total city emissions (CE) for each greenhouse gas:

$$CE = \sum_i SE_i \quad (1)$$

where SE_i is the “segment emissions” for a road segment *i*. Each SE_i was calculated as:

$$SE_i = 365 \cdot AADT_i \cdot l_i \cdot \sum_{v,f} \eta_{v,f,s_i} \cdot m_{v,f,s_i} \cdot g_{v,f,s_i} \quad (2)$$

where:

- AADT_{*i*} is the average annual daily traffic (unitless vehicles),
- *l_i* is the length of the road segment *i*, in units of km (see section 2.2.3),
- η_{v,f,s_i} is the fuel efficiency, in units of liters per km (see section 2.2.7), for a vehicle type *v*, fuel type *f* and a road segment category *s_i* of the road segment *i*,
- m_{v,f,s_i} the vehicle mix, as a fraction, typically present on the road segment based on the vehicle type *v*, fuel type *f* and a road segment category *s_i* of the road segment *i*.

Specifically, we require that $\sum_{v,f} m_{v,f,s_i} = 1$ for each road segment category *s_i*.

- Finally, g_{v,f,s_i} is the greenhouse gas emissions factor, in grams of gas per liter, for the vehicle type *v*, fuel type *f* and a road segment category *s_i* of the road segment *i*.

All three of η_{v,f,s_i} , m_{v,f,s_i} , and g_{v,f,s_i} are look-up tables based on data gathered from several sources. A list of supported road segment types is provided in section 2.2.5, supported vehicle types in section 2.2.6, supported fuel types in section 2.2.7, and greenhouse gas factors in 2.2.8. We calculated separate CEs for every greenhouse gas, and used IPCC GWP20 and GWP100 values (Forster *et al.*, 2021) to convert methane and nitrous oxide emissions into carbon dioxide equivalent emissions.

To calculate a city emissions factor (CEF), we calculated the total city emissions and divided by the total activity, defined as the sum over all road segments of the AADT times length for each segment:

$$CEF = CE / (365 \cdot \sum_i AADT_i \cdot l_i), \quad (3)$$

As with the CE calculation, we calculate a separate CEF for each of the three major greenhouse gasses. Thus, three emissions factor values were provided for each city, representing the average amount of each greenhouse gas emitted per kilometer traveled by a single vehicle on any road segment within that city. The units of each provided city emissions factor (CEF) are metric tons (tonnes) of greenhouse gas per vehicle kilometer traveled (VKT).

2.2.10. U.S. Emissions Inventories

We utilized three other GHG inventories in the U.S. for validation of our emissions estimates (section 3.3.1): Google Environmental Insights Explorer (EIE; Google 2022a), Database of Road Transportation Emissions (DARTE; Gately, Hutyra and Wing 2019), and Vulcan v3.0 (Gurney *et al.*, 2020). Google EIE leverages trip data from Google Maps (Google 2022b) in combination with emissions factors data to provide emissions estimates for multiple modes of transportation in 42,000+ cities worldwide. We used the publicly available 2018 EIE data in the U.S. for our comparison. DARTE uses reported vehicular traffic data combined with Census TIGER (Marx 1986) road network information to estimate regional on-road emissions and disaggregate them among mapped road networks. We compared our estimates to both DARTE 2015 and 2017 data. Vulcan is a national-scale, multi-sectoral, hourly inventory from 2010-2015 with a resolution of 1 km². Vulcan transportation emissions are based on EPA county-level on-road emissions estimates, further downscaled using data from the Federal Highway Administration. We selected Vulcan data from 2015 for comparison.

2.2.11. International Emissions Inventories

Comparison against two international emissions inventories has been performed for the global 500 cities emission estimates: EDGAR (Crippa *et al.*, 2019) and Carbon Monitor (Liu *et al.*,

2020). EDGAR provides a global inventory for transportation that uses road density as a proxy to spatially distribute emissions. We used EDGAR data for 2015 for comparison, as it was included in the metadata for all 500 cities we selected from the GHSL-UCDB dataset (Florczyk *et al.*, 2019). Carbon Monitor (Liu *et al.*, 2020) is a global, multi-sector emissions inventory that utilizes a variety of activity data to estimate daily GHG emissions since 2019. We used the Carbon Monitor city level data for comparison and found a set of 50 cities that overlapped with our set of 500 global cities.

2.3. Machine Learning Models

As part of our hybrid modeling approach, machine learning models were trained to estimate road activity from satellite imagery, road network data, and other geospatial datasets. These estimates were required in the absence of global, openly available, and high-quality traffic activity data.

2.3.1. Ground Truth Road Activity Data

To train our ML models, we utilized the U.S. Highway Performance Monitoring System Average Annual Daily Traffic (AADT) data from 2017 (U.S. FHWA 2017). This AADT data was recorded using road-side devices and was provided by each U.S. state independently. We utilized the total AADT measure in our work, measured in vehicles per day. AADT data was not available for every road segment and was typically only recorded on major highways and arterial (collector) roads. For CNN-based models, we used AADT data for 3,708 cities to train our models and 115 cities for validation. There was minimal geographic overlap between this set of cities with the 14 cities used for emissions validation in section 3.3.1. For GNN models, we aligned the AADT data with OpenStreetMap road segments in 80 U.S. cities, and used 59 cities for training, 7 for validation, and 14 for test (the same 14 cities used for validation in section 3.3.1).

2.3.2. Convolutional Neural Network (CNN)

Our first machine learning approach used semantic segmentation CNNs to predict AADT. Specifically, we input visual satellite imagery in combination with rasterized road network data to predict AADT on a per-pixel basis (Figure 10). This approach was informed by previous work in directly regressing road transport CO₂ emissions (Mukherjee *et al.*, 2021). Models were trained using two sources of imagery: Sentinel-2 Level-2A product at 10 m x 10 m resolution, using bands 4 (red), 3 (green), and 2 (blue) (Drusch *et al.*, 2012), and Planet Labs PlanetScope ~3 m monthly and quarterly mosaics (Planet Labs 2022).

Road network data was retrieved and rasterized for the corresponding extent of each visual image tile. Each road type (highway, secondary, local) was rasterized independently, and the resulting raster channels are concatenated together to form a three channel image. This image

was then combined with the visual image to form a six channel input image that was input to the CNN. Thus, the CNN model is tasked with predicting the number of vehicles traveling on a road segment as a function of visual satellite image features and road location and type information. We primarily used MAnet-based architectures (Fan *et al.*, 2020) for our segmentation models, based on the findings of similar previous work (Mukherjee *et al.*, 2021).

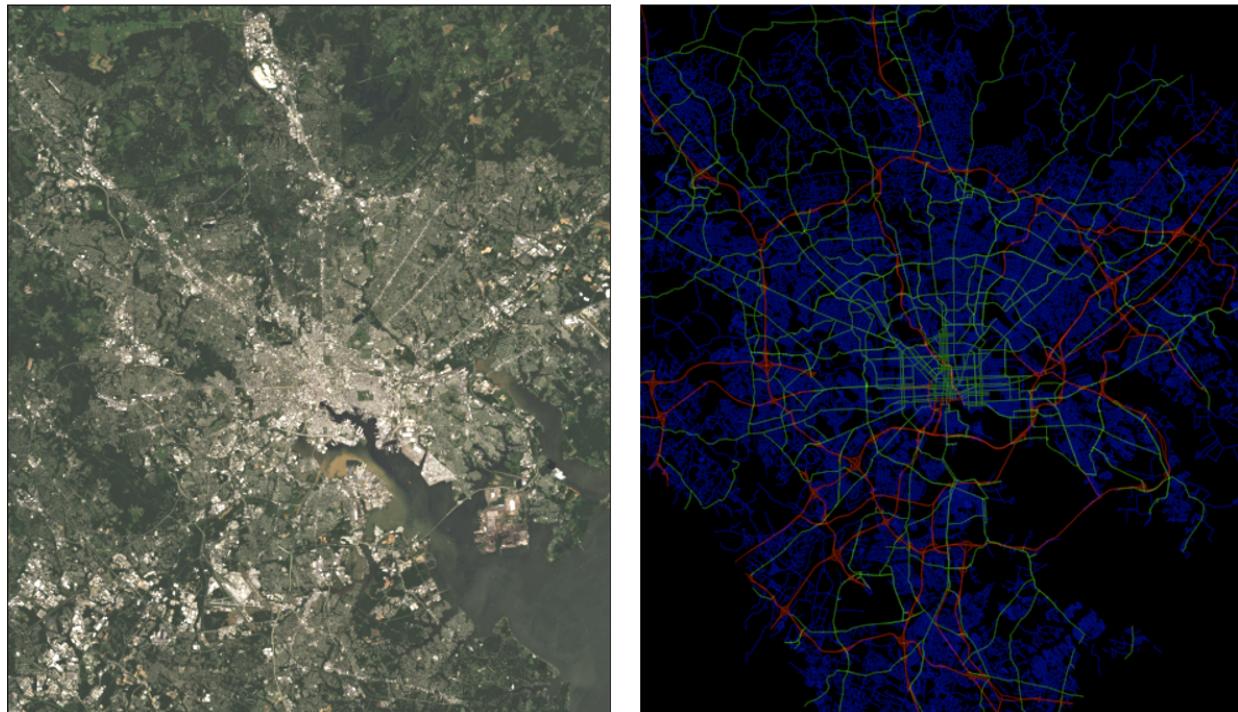


Figure 10 Example Sentinel-2 visual (RGB) satellite image (left) and corresponding rasterized OSM road network data (right) for Baltimore, Maryland USA. The OSM raster colors represent different road types: highways (red), secondary roads (green), and local roads (blue).

When the model was used within the overall emissions calculation pipeline, pixel-based AADT predictions were post-processed to associate predicted values with their corresponding road segment. All AADT predictions for a given road segment were averaged to produce a single AADT value for every road within the current geographic extent. A road ID to predicted AADT mapping was then saved and fed to the emissions factors pipeline for emissions calculation and aggregation, as described in section 2.2.9.

2.3.3. Graph Neural Network (GNN)

Another type of ML model trained to predict AADT was graph neural networks (GNNs) (Bronstein *et al.*, 2017). Road networks inherently take the form of a graph structure, and GNNs

can capture road activity across a range of scales more easily than the image-based convolutional neural network (CNN) segmentation models. CNN-based solutions constrain the spatial area that can be covered during inference, making it difficult to capture potential dependencies on features in neighboring or distant locations. GNNs can easily leverage various features assigned to nodes and efficiently reason over the full road network graph to provide more robust estimates of on-road activity.

For this work, a GNN was trained using OpenStreetMap road network data, including a number of road features: the number of lanes, road length, road type, link or not (where a link is, e.g., an exit ramp), and the directional angle between roads. The Graph Attention network (GAT; Veličković *et al.*, 2018) architecture was used as it allows for both edge and node input features, and was set up to predict log-AADT values. We note that the GNN does not use visual imagery as input, and is able to learn a relationship between road types and configurations to the density of traffic on those roads.

2.3.4. Model Ensembling

To create a more robust and predictive AADT estimation model, ensembling was performed using the CNN and GNN models. Input data were collected for 2021, and resulting model AADT predictions (measured in vehicles per day) per road segment were averaged before being input to the emissions factors pipeline to estimate emissions for 2021. This capability can be easily extended in the future to include different model architectures, experiment with different combinations of models, and perform further analysis of inter-model variance. An example ensemble AADT output can be seen below in Figure 11.

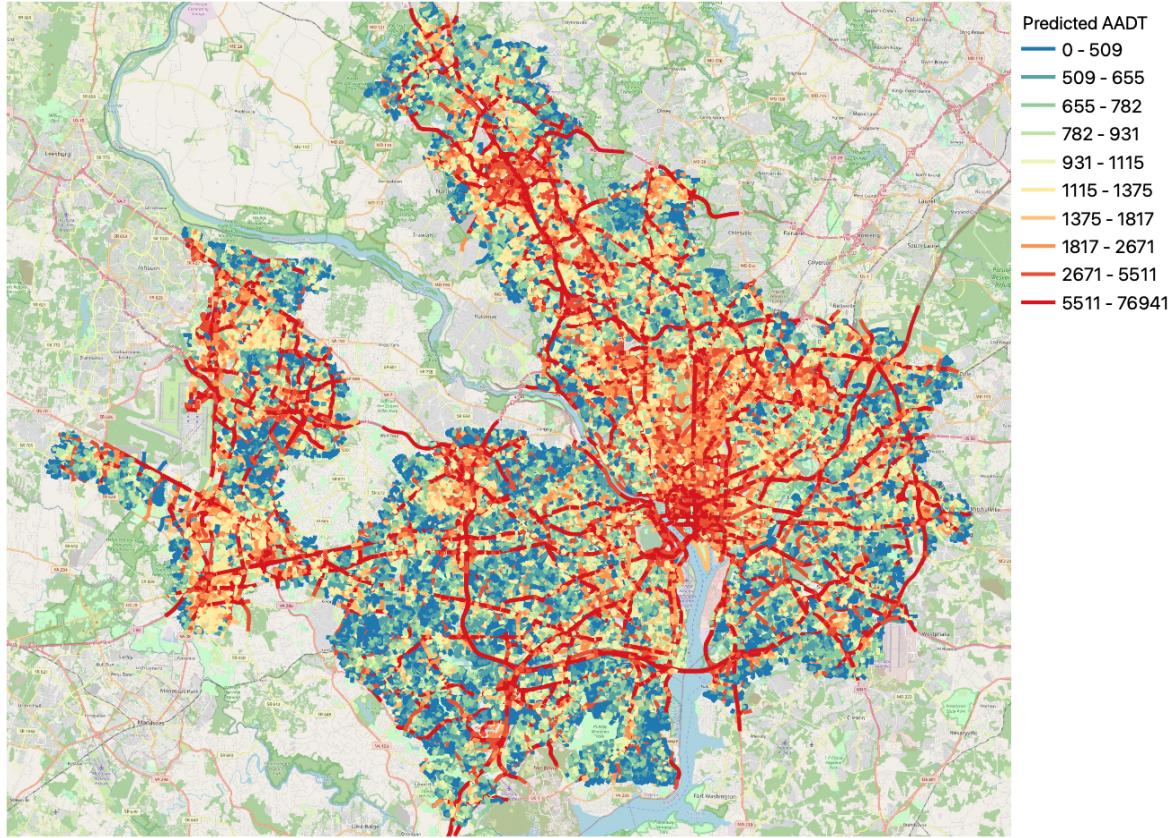


Figure 11 Ensemble-predicted AADT for 2021, measured in vehicles per day, for Washington, D.C.

3. Results

3.1. Metrics

An initial set of metrics are used to quantify differences between various emissions estimates. We note that the use of “error” is kept for consistency with other literature and ease of understanding. However, it should be noted that reconciling the discrepancies and uncertainties across emissions inventories is an open field of research, so “error” may be somewhat pessimistic. The following metrics are used in our validation:

$$\text{Mean Absolute Error (MAE)} = \frac{1}{N} \sum_{i=1}^N \left| \hat{x}_i - x_i \right|$$

$$\text{Mean Error} = \frac{1}{N} \sum_{i=1}^N (\hat{x}_i - x_i)$$

$$\text{Mean Percent Error (MPE)} = 100 * \frac{1}{N} \sum_{i=1}^N \frac{\hat{x}_i - x_i}{x_i}$$

$$\text{Absolute Percent Error (APE)} = 100 * \frac{|\hat{x}_i - x_i|}{x_i}$$

$$\text{Mean Absolute Percent Error (MAPE)} = 100 * \frac{1}{N} \sum_{i=1}^N \frac{|\hat{x}_i - x_i|}{x_i}$$

Where \hat{x}_i is our predicted/estimated value, and x_i is either the emissions value we are comparing against or a ground truth AADT value. Unless otherwise noted, all emissions values are in tonnes CO₂ (tCO₂).

3.2. Machine Learning Model Evaluation

Each ML model is evaluated on a hold-out test set of U.S. cities, for which the ground truth AADT data is known. We then compute MAPE on a per road basis for a fairer comparison between the raster-based CNN models and the graph-based GNN model, as can be seen in Table 4.

Table 4 Machine learning model evaluation metrics.

Model	Per Road MAPE
CNN: S2 + OSM Roads	116%
CNN: Planet + OSM Roads	160%
GNN: OSM Road Features	138%

3.3. Emissions Validation

Our emissions estimates are compared against several other emissions inventories, both within the U.S. and globally. These comparisons are performed for initial validation of our data, and we note that thorough comparison and in-depth discrepancy investigation has not been completed.

3.3.1. U.S. Test Set Validation

A set of 14 cities in the U.S. were selected for validation and comparison with Google Environmental Insights Explorer (EIE; Google 2022a), Database of Road Transportation Emissions (DARTE; Gately, Hutyra and Wing 2019), and Vulcan v3.0 (Gurney *et al.*, 2020; see section 2.2.10 for more details). Both our CNN and GNN-derived emissions estimates were strongly correlated with other inventory values for every city, with mean Pearson R values of 0.97 (CNN) and 0.98 (GNN), and demonstrated the ability to turn ML-predicted activity predictions into reasonable emissions estimates.

Due to the fact that our ground truth AADT data and satellite imagery for this set of cities is from 2017, data from the other emissions inventories were selected from years as close to 2017 as possible. We use the geographic bounds available in the EIE data to retrieve satellite imagery and road network data within each city's bounds. After predicting AADT with our models and associating AADT with each road segment, road geometries are cropped to the city bounds to create an appropriate estimate of VKT and emissions for each road. Corresponding emissions estimates from the DARTE and Vulcan raster products are also selected using each city's EIE bounds.

Several variants of each third party inventory are examined. Google EIE data categorizes trips into three categories: in-boundary, inbound, and outbound. Trips are categorized according to their start and end locations, with in-boundary containing trips that both start and end within city bounds, inbound starting outside and ending inside city bounds, and outbound starting inside and ending outside city bounds. We compare against just in-boundary emissions (EIE_v1_2018), and in-boundary plus 50% inbound and 50% outbound emissions (EIE_v2_2018). For DARTE, we compare against emissions estimates for both 2015 (DARTE_2015) and 2017 (DARTE_2017). Vulcan contains three emissions estimates: the lower 95% confidence interval (VULCAN_lo_2015), mean estimate (VULCAN_mn_2015), and the upper 95% confidence interval (VULCAN_hi_2015).

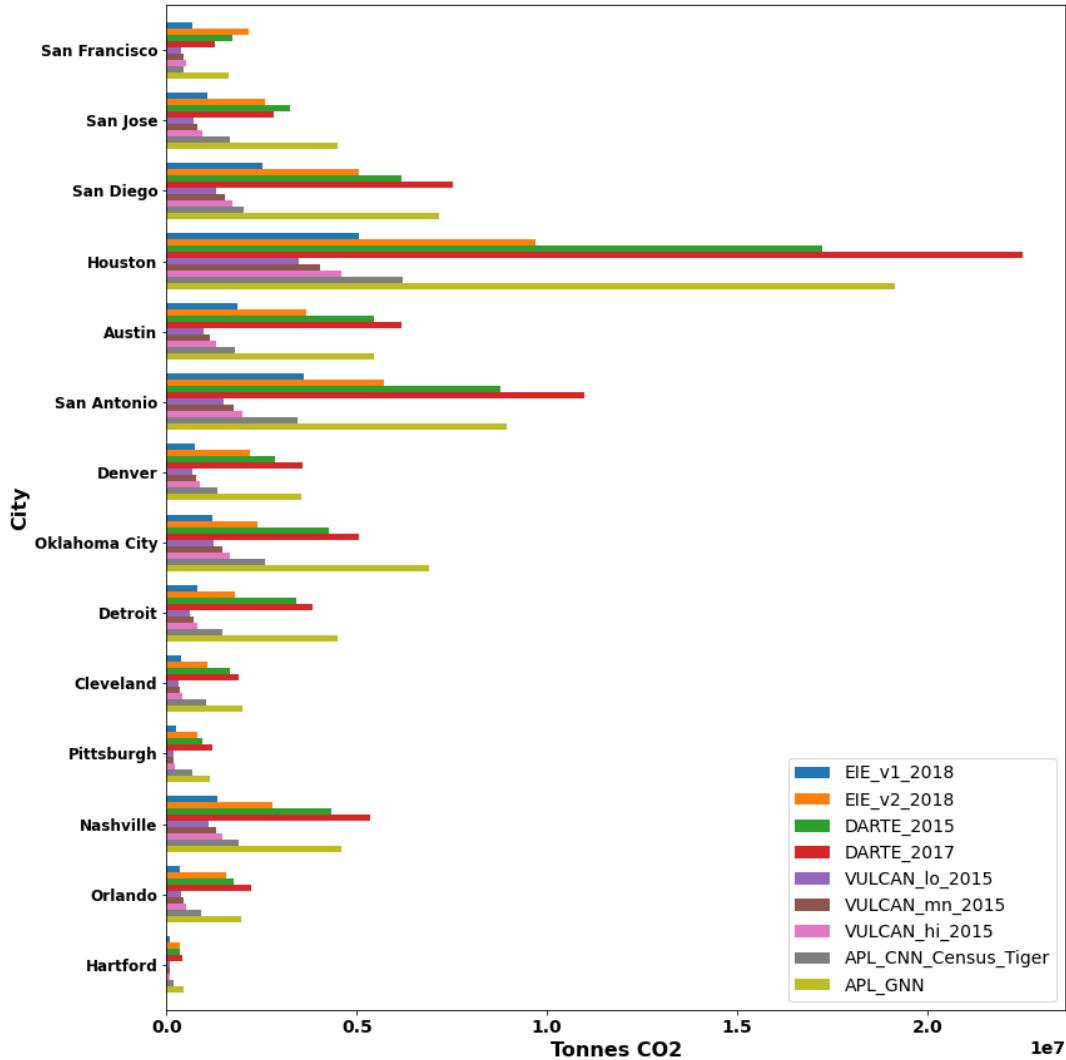


Figure 12 Comparison of our emissions estimates (APL_CNN_Census_Tiger and APL_GNN) with all three other U.S. inventories (EIE, DARTE and Vulcan). See text for description of each entry.

Emissions values for each dataset are shown in Figure 12. Two versions of our estimates are shown: CNN-based (APL_CNN_Census_Tiger, see section 2.3.2) and GNN-based (APL_GNN, see section 2.3.3). The APL_CNN_Census_Tiger is an earlier version of the CNN trained using U.S. Census TIGER road network data (Marx 1986), instead of OpenStreetMap data. However, we expect similar performance for both models. Of note is the wide range of predicted values for some cities, within which our estimates generally fall. The GNN tends to predict higher AADT values than the CNN model, resulting in the larger emissions values shown here. The distribution of emissions values across all datasets is shown in Figure 13 and described in Table 5.

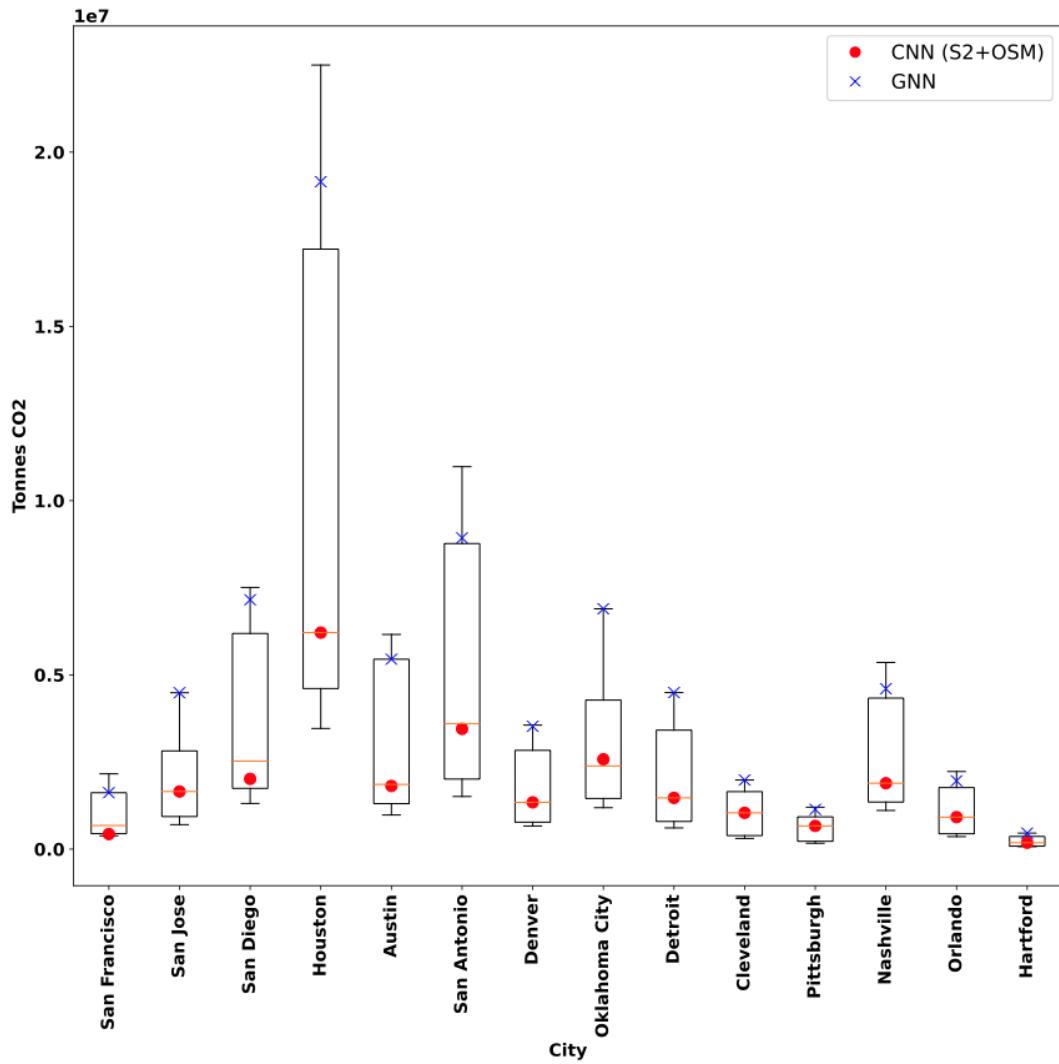


Figure 13 Distribution of emissions estimates for all inventories compared to APL_CNN_Census_Tiger (red dot) and APL_GNN (blue x) emission estimates for each U.S. city.

Table 5 Mean and standard deviations of emissions estimates for 14 U.S. cities across all inventories.

City	Mean (tCO ₂)	Std (tCO ₂)	City	Mean (tCO ₂)	Std (tCO ₂)
San Francisco	1,023,106	678,291	Oklahoma City	2,970,274	2,002,786
San Jose	2,032,675	1,319,900	Detroit	1,990,571	1,513,631
San Diego	3,893,304	2,565,230	Cleveland	1,008,559	689,489
Houston	10,216,277	7,397,669	Pittsburgh	619,213	421,934
Austin	3,091,428	2,107,453	Nashville	2,690,249	1,652,154
San Antonio	5,192,045	3,562,064	Orlando	1,123,018	753,909
Denver	1,837,115	1,213,108	Hartford	231,771	158,412

A summary of metrics for all compared datasets is provided in Table 6. Including portions of the inbound and outbound EIE emissions seems to be a fairer comparison with our emissions estimates, as evidenced by the reduced MAPE values for EIE_v2_2018 vs. EIE_v1_2018 for both types of models. Generally, the GNN tends to overestimate when compared with the CNN model, and produces emissions estimates more in line with DARTE.

Table 6 Comparison metrics for CNN and GNN-based emissions estimates to each emissions dataset.

Emissions Dataset	CNN			GNN		
	MAE (tCO ₂)	Mean Error (tCO ₂)	MAPE	MAE (tCO ₂)	Mean Error (tCO ₂)	MAPE
EIE_v1_2018	544,225	407,997	77.3%	3,706,827	3,706,827	321.5%
EIE_v2_2018	1,180,437	-1,153,065	36.1%	2,223,303	2,145,764	71.3%
DARTE_2015	2,606,389	-2,606,389	53.6%	708,514	692,440	19.8%
DARTE_2017	3,505,472	-3,505,472	59.5%	875,254	-206,642	17.2%
VULCAN_lo_2015	912,134	912,134	124.8%	4,210,964	4,210,964	473.2%
VULCAN_mn_2015	761,213	759,672	93.3%	4,058,502	4,058,502	391.8%
VULCAN_hi_2015	617,740	607,210	71%	3,906,040	3,906,040	330.7%

3.3.2. Global 500 Validation

Comparison against EDGAR (Crippa *et al.*, 2019) and Carbon Monitor (Liu *et al.*, 2020) has been performed for the global 500 cities emission estimates (see section 2.2.11 for more details).

EDGAR 2015 data is available for all 500 cities, and Carbon Monitor city-level data is openly available for 52 cities globally, 50 of which are also contained in our set of 500 cities. We found strong correlation with both inventories, with Pearson R values of 0.74 (EDGAR) and 0.87 (Carbon Monitor), showing the high global accuracy of our method.

The resulting comparison for all 500 cities against EDGAR can be seen in Figure 15. While the Pearson R value of 0.74 indicates strong correlation, the wide variance of the differences is noteworthy and warrants further investigation. The sharp “wall” on the left portion of the plot is caused by the fact that our 500 cities were selected based on thresholded EDGAR 2015 estimates.

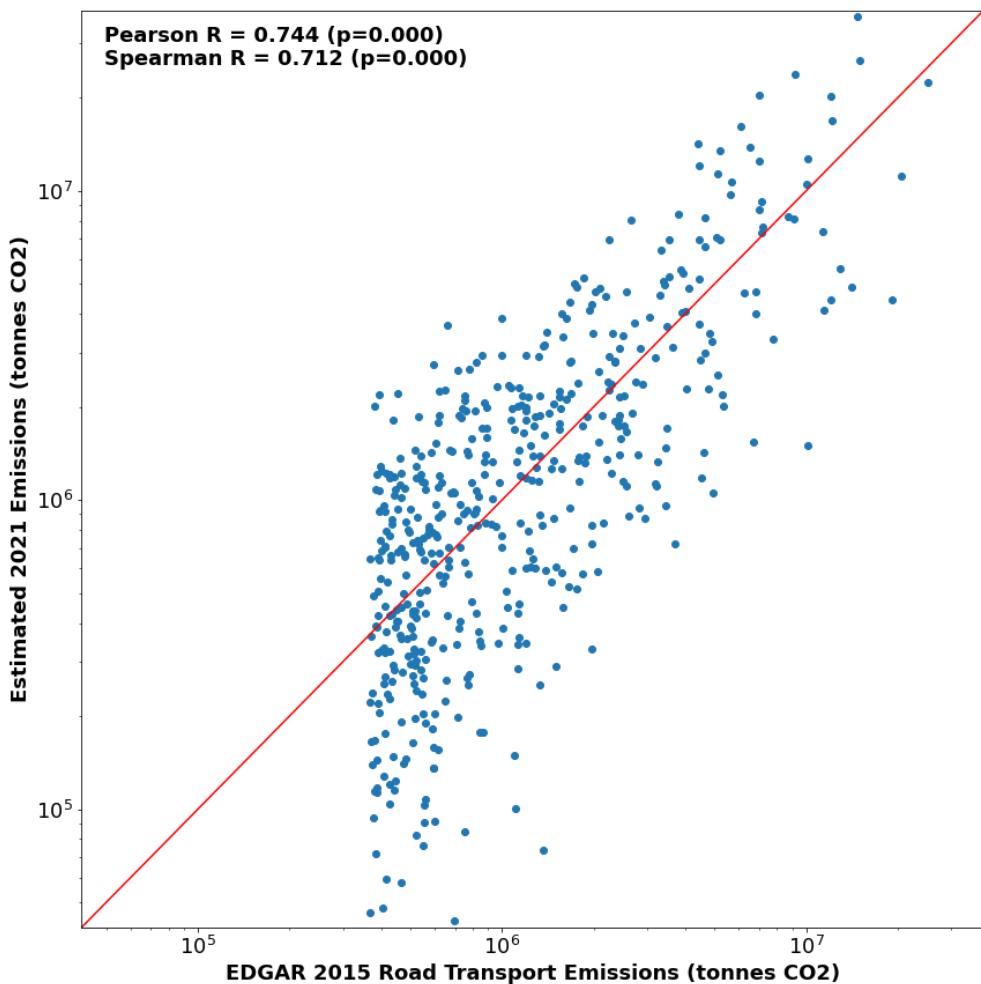


Figure 15 Our emissions estimates compared to EDGAR 2015 data for 500 global cities. Note, x- and y-scales are in log scale.

Further analysis shows the countries for which we see the largest and smallest differences between our emissions estimates and EDGAR values, listed below in Tables 7 and 8, respectively. We suspect the large differences in African and Latin American countries are due to lack of localized emissions factors data in those regions, as well as contrasting visual appearance and vehicle activity levels when compared to our U.S.-based training data.

Table 7 20 countries with the largest absolute differences compared to EDGAR 2015 data.

Country	Absolute Pct. Error	Region
Angola	455%	Africa
Ethiopia	454%	Africa
Tanzania	430%	Africa
Ghana	364%	Africa
Paraguay	316%	Latin America and the Caribbean
Sudan	288%	Africa
Ukraine	208%	Europe
Morocco	202%	Africa
Tunisia	179%	Africa
Bolivia	171%	Latin America and the Caribbean
Jordan	141%	Asia
Chile	131%	Latin America and the Caribbean
Nigeria	123%	Africa
United States	120%	Northern America
Finland	99%	Europe
Argentina	97%	Latin America and the Caribbean
Japan	93%	Asia
South Africa	91%	Africa
Iraq	89%	Asia
Zimbabwe	88%	Africa

Table 8 20 countries with the smallest absolute differences compared to EDGAR 2015 data.

Country	Absolute Pct. Error	Region
Turkmenistan	2%	Asia
Peru	3%	Latin America and the Caribbean
Sri Lanka	4%	Asia
Austria	9%	Europe
Kenya	13%	Africa
Denmark	22%	Europe
Russia	22%	Europe
Algeria	23%	Africa
Ireland	24%	Europe
Panama	24%	Latin America and the Caribbean
Bahrain	27%	Asia
Thailand	27%	Asia
Vietnam	28%	Asia
Israel	28%	Asia
Hungary	28%	Europe
South Korea	29%	Asia
United Arab Emirates	29%	Asia
Palestine	30%	Asia
Czech Republic	30%	Europe
Spain	32%	Europe

The results of the comparison with the 50 overlapping Carbon Monitor cities for 2021 is shown in Figure 16. There is generally good alignment between the two sets of emissions, with some larger differences in France (Nice, Lyon, Marseille), South America (Bogota, São Paulo), Russia (Saint Petersburg, Moscow), and India (Mumbai, Delhi). We also note the larger MAPE for 2020 in Table 9 as compared to 2019 and 2021, likely due to COVID-19 lockdown effects.

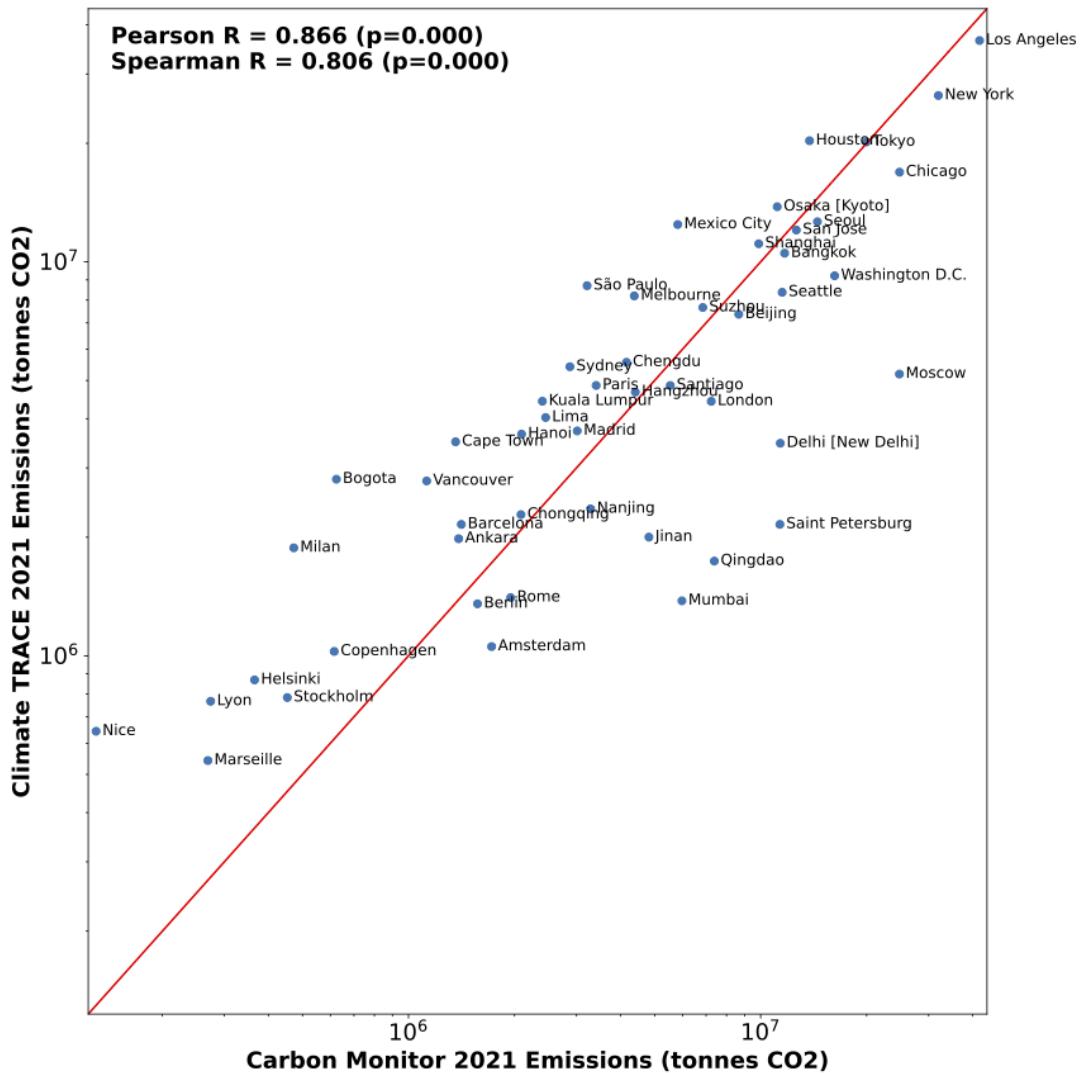


Figure 16 Our emissions estimates compared against Carbon Monitor 2021 data for 50 global cities. Note, x- and y-scales are in log scale.

Table 9 Metrics summary for global 500 comparisons.

Emissions Dataset	# of Cities	MAE (tCO ₂)	MAPE	Mean Error (tCO ₂)	MPE	Pearson ρ
EDGAR 2015	500	1,158,740	68.8%	248,624	23.6%	0.74
Carbon Monitor 2019	50	2,857,690	72.4%	-844,634	44.4%	0.87
Carbon Monitor 2020	50	2,634,598	83.2%	-317,283	55.7%	0.86
Carbon Monitor 2021	50	2,795,294	73.5%	-781,053	42.4%	0.87

4. Discussion

Our emissions estimates are within reasonable distance from and strongly correlated with other emissions inventories, as has been shown by the validation results above. Potential sources of discrepancies between the various emissions inventories warrant further investigation, including timespan differences and varying geographic extents for each city. A key strength of our approach is the ease with which our estimates can be updated as better data become available, as well as the ability to run our pipeline over more cities and/or larger regions. Further international validation for both AADT and resulting emissions estimates is necessary, including a deeper understanding of all potential sources of differences (e.g., emissions factors uncertainty, varying total road network length included, etc.). We aim to incorporate more international AADT datasets, where available, into our model training and validation process. The addition of more real-time data (e.g., traffic and mobility) will help address the temporal ambiguity of our estimates, however finding publicly available data of this type remains a significant challenge.

5. Conclusion

We have presented a hybrid road transportation emissions estimation method that is detailed, scalable, and easy to update. The ability to calculate emissions per road segment can be further refined to reach an unprecedented level of detail and global coverage. Where available, the integration of real-time traffic data would increase the temporal resolution and accuracy of our models. We also plan to carry out further analysis of our emissions estimates with other inventories to identify the main causes of discrepancies. As well, we aim to explore open-sourcing our emissions factors schema such that governments and other entities can contribute more up-to-date and accurate emissions factors data to further improve our estimates.

This type of actionable emissions monitoring data will be critical to ensuring we meet global emissions reduction targets and may inspire new ways of mitigating the effects of climate change.

6. Acknowledgements

Special thanks to Dr. Kevin Gurney for reviewing our 2021 methodology and for his collaboration and continuous feedback this year. Additional thanks to Gabriela Volpato, Aaron Davitt, and Lekha Sridhar from WattTime for their help in developing and reviewing our methodology.

7. Supplementary Materials

Table 10 Vehicle Fleet Mix Data Sources by Country.

Country	Source
Argentina	ADEFA Anuario 2020 - Parque Automotor, page 2
Australia	Australian Bureau of Statistics
Austria	ACEA [7]
Azerbaijan	UNECE [8]
Belgium	ACEA
Brazil	DENATRAN; Sindicato Nacional da Indústria de Componentes para Veículos Automotores (from Statista)
Canada	Statistics Canada (https://doi.org/10.25318/2310006701-eng)
Chile	Instituto Nacional de Estadísticas
China	Guangzhou Statistical Yearbook (Guangzhou Bureau of Statistics, 2017)
Czech Republic	ACEA
Denmark	ACEA
Finland	ACEA
France	ACEA
Germany	ACEA
Greece	ACEA
Guatemala	Boletín Estadístico SAT Guatemala, tab G2
Hungary	ACEA
India	India: Road Transport Yearbook
Indonesia	Indonesian National Police; Statistics Indonesia (via Statista)

Ireland	ACEA
Israel	UNEC
Italy	ACEA
Japan	MLIT
Kuwait	CSB (Kuwait) - Statista
Malaysia	Malaysia Department of Environment 2016 Report
Myanmar	MMSIS; MOTC (Department of Road Transport Administration)
Netherlands	ACEA
Poland	ACEA
Portugal	ACEA
Russia	ACEA
Singapore	LTA (lta.gov.sg)
Spain	ACEA
Sweden	ACEA
Switzerland	ACEA
Turkey	ACEA
United Kingdom	ACEA

8. References

1. Boeing, G., 2017. OSMnx: New Methods for Acquiring, Constructing, Analyzing, and Visualizing Complex Street Networks. *Computers, Environment and Urban Systems*, 65, pp.126-139. doi:10.1016/j.compenvurbsys.2017.05.004
2. Bronstein, M. M., Bruna, J., LeCun, Y., Szlam, A., and Vandergheynst, P., 2017. Geometric Deep Learning: Going beyond Euclidean data. *IEEE Signal Processing Magazine* (34), no. 4, pp. 18-42 [online]. Available at: <https://doi.org/10.1109/MSP.2017.2693418>. (Accessed: 6 September 2022).
3. Crippa, M., Solazzo, E., Huang, G., Guizzardi, D., Koffi, E., Muntean, M., Schieberle, C., Friedrich, R., Janssens-Maenhout, G., 2019. High resolution temporal profiles in the Emissions Database for Global Atmospheric Research (EDGAR), *Nature Scientific Data*, [online]. Available at: doi:10.1038/s41597-020-0462-2. (Accessed: 30 August 2022)
4. Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., et al., 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment* (120), pp.25-36 [online]. Available at: <https://doi.org/10.1016/j.rse.2011.11.026>. (Accessed: 30 August 2022)
5. Fan, T., Wang, G., Li, Y., and Wang, H., 2020. MA-Net: A Multi-Scale Attention Network for Liver and Tumor Segmentation. *IEEE Access* (8), pp. 179656-179665 [online]. Available at: <https://doi.org/10.1109/ACCESS.2020.3025372>.
6. Florczyk, A.J., Corbane, C., Schiavina, M., Pesaresi, M., Maffenini, L., Melchiorri, M., Politis, P., Sabo, F., Freire, S., Ehrlich, D., Kemper, T., Tommasi, P., Airaghi, D. and Zanchetta, L., 2019. GHS Urban Centre Database 2015, multitemporal and multidimensional attributes, R2019A. European Commission, Joint Research Centre (JRC) [online]. Available at: <http://data.europa.eu/89h/53473144-b88c-44bc-b4a3-a4583ed1f547e> (Accessed: 30 August 2022)
7. Gately, C., Hutyra, L.R., and Wing, I.S., 2015. Cities, traffic, and CO₂: A multidecadal assessment of trends, drivers, and scaling relationships. *Proceedings of the National Academy of Sciences*, 112 (16) pp.4999-5004 [online]. Available at: <https://doi.org/10.1073/pnas.1421723112> (Accessed: 19 September 2022).
8. Gately, C., Hutyra, L.R., and Wing, I.S., 2019. DARTE Annual On-road CO₂ Emissions on a 1-km Grid, Conterminous USA, V2, 1980-2017. ORNL DAAC, Oak Ridge, Tennessee, USA [online]. Available at: <https://doi.org/10.3334/ORNLDaac/1735> (Accessed: 30 August 2022)
9. Global Administrative Areas, 2022. *GADM database of Global Administrative Areas, version 4.0* [online]. Available at: www.gadm.org (Accessed: 30 August 2022)
10. Google, 2022. *Environmental Insights Explorer* [online]. Available at: <https://insights.sustainability.google> (Accessed: 30 August 2022)

11. Google, 2022. *Google Maps* [online]. Available at: <https://maps.google> (Accessed: 6 September 2022)
12. Gurney, K. R., Liang, J., Patarasuk, R., Song, Y., Huang, J., and Roest, G., 2020. The Vulcan version 3.0 high-resolution fossil fuel CO₂ emissions for the United States. *Journal of Geophysical Research: Atmospheres*, 125 [online]. Available at: <https://doi.org/10.1029/2020JD032974> (Accessed: 30 August 2022)
13. International Energy Agency (IEA), 2022. Transport sector CO₂ emissions by mode in the Sustainable Development Scenario, 2000-2030 [online]. Available at: <https://www.iea.org/data-and-statistics/charts/transport-sector-co2-emissions-by-mode-in-the-sustainable-development-scenario-2000-2030> (Accessed: 16 September 2022).
14. Liu, Z., Ciais, P., Deng, 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 [online]. Available at: <https://doi.org/10.1038/s41597-020-00708-7> (Accessed: 30 August 2022)
15. Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U. and Gascon, F., 2017, October. Sen2Cor for sentinel-2. In *Image and Signal Processing for Remote Sensing XXIII* (Vol. 10427, pp. 37-48). SPIE.
16. Marx, R. W., 1986. The TIGER system: Automating the Geographic Structure of the United States Census. *Government Publications Review* (13), pp.181-201.
17. Mukherjee, R., Rollend, D., Christie, G., Hadzic, A., Matson, S., Saksena, A., and Hughes, M., 2021. Towards Indirect Top-Down Road Transport Emissions Estimation. *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1092-1101 [online]. Available at: <https://doi.org/10.1109/CVPRW53098.2021.00120> (Accessed: 6 September 2022)
18. OpenStreetMap contributors, 2021. *Planet OSM* [online]. Available at: <https://planet.openstreetmap.org> (Accessed: 30 August 2022)
19. Planet Labs, 2022. *PlanetScope Imagery* [online]. Available at: <https://www.planet.com/products/planet-imagery> (Accessed: 30 August 2022)
20. Scheibenreif, L.M., Mommert, M., and Borth, D., 2021. Estimation of Air Pollution with Remote Sensing Data: Revealing Greenhouse Gas Emissions from Space. *ICML 2021 Workshop on Tackling Climate Change with Machine Learning* [online]. Available at: <https://www.climatechange.ai/papers/icml2021/23> (Accessed: 19 September 2022).
21. The European Automobile Manufacturers' Association (ACEA), 2022. *Report – Vehicles in Use, Europe* [online]. Available at: <https://www.acea.auto/publication/report-vehicles-in-use-europe-2022/> (Accessed: 30 August 2022)
22. Topf, J., 2022. *Osmium: Command line tool for working with OpenStreetMap data based on the Osmium library* [online]. Available at: osmcode.org/osmium-tool (Accessed: 30 August 2022)

23. United Nations Economic Commission for Europe (UNECE) Statistical Database, 2022. *Road vehicle fleet at 31 December by vehicle category and fuel type* [online]. Available at: https://w3.unece.org/PXWeb2015/pxweb/en/STAT/STAT_40-TRTRANS_03-TRRoadFleet/03_en_TRRoadFuelFlt_r.px (Accessed: 30 August 2022)
24. U.S. Environmental Protection Agency (EPA), 2022. *Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2020* [online]. Available at: <https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks> (Accessed: 19 September 2022)
25. U.S. Environmental Protection Agency (EPA), 2022. *GHG Emissions Factors Hub* [online]. Available at: https://www.epa.gov/system/files/documents/2022-04/ghg_emission_factors_hub.pdf (Accessed: 30 August 2022)
26. U.S. Federal Highway Administration (FHWA), 2020. *Distribution of Annual Vehicle Distance Traveled - Percentage by Vehicle Type Rural/Urban (Table VM-4)* [online]. Available at: <https://www.fhwa.dot.gov/policyinformation/statistics/2020/vm4.cfm> (Accessed: 30 August 2022)
27. U.S. Federal Highway Administration, 2017. *Highway Performance Monitoring System (HPMS) Data* [online]. Available at: <https://www.fhwa.dot.gov/policyinformation/hpms> (Accessed: 30 August 2022)
28. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., and Bengio, Y., 2018. Graph Attention Networks. *arXiv* [online]. Available at: <http://arxiv.org/abs/1710.10903> (Accessed: 6 September 2022)
29. World Bank Group, 2019. *CURB: Climate Action for Urban Sustainability, Version 2.1* [online]. Available at: <https://datacatalog.worldbank.org/search/dataset/0042029> (Accessed: 30 August 2022)
30. World Resource Institute, 2022. *Climate Watch Historical GHG Emissions* [online]. Available at: <https://www.climatewatchdata.org/ghg-emissions> (Accessed: 19 September 2022).
31. Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J.-L., Frame, D., Lunt, D.J., Mauritzen, T., Palmer, M.D., Watanabe, M., Wild, M., and Zhang, H.,f 2021: *The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 923–1054, doi:10.1017/9781009157896.009. (Accessed: 19 September 2022).