CHEER: Calculating Holistic Energetic Emissions on Route — A Fine-Grained, Quasi-Real-Time System for Computing the Energy and Emission Impact of Personal Transportation

1st Abigail Wheelis
National Renewable Energy Laboratory
Golden, CO, USA
abigail.wheelis@nrel.gov
ORCID: 0000-0003-1635-0526

2nd Natalie Schultz
National Renewable Energy Laboratory
Golden, CO, USA
natalie.schultz@nrel.gov
ORCID: 0009-0006-8397-1487

3rd Jack Greenlee
National Renewable Energy Laboratory
Golden, CO, USA
jack.greenlee@nrel.gov
ORCID: 0009-0009-9640-9145

4th Jacques Fleischer
National Renewable Energy Laboratory
Golden, CO, USA
jacques.fleischer@nrel.gov
ORCID: 0000-0002-1102-1910

5th Louis Grassi
National Renewable Energy Laboratory
Golden, CO, USA
louis.grassi@nrel.gov
ORCID: 0009-0003-3869-8096

6th K. Shankari, PhD
National Renewable Energy Laboratory
Golden, CO, USA
k.shankari@nrel.gov
ORCID: 0000-0002-7046-5570

***Abstract*—Precisely calculating the energy and emissions footprint for an individual's travel behavior is a complex task, but critical to understanding the impact of efforts to reduce the footprint of the highest emitting sector in the United States. Each individual takes many trips, spread across any number of modes, and each mode operates at a distinct energy efficiency. Precise energy consumption calculations should incorporate the efficiency of each vehicle and the number of travelers in the vehicle. Emissions rates depend on the fuel used, and with the rise of renewable electricity in particular, even how and when that fuel was produced. Available literature on travel footprints often elides the specific details of calculation, relies on generalized, frequently outdated, averages, and can be limited to a specific region of interest. In this paper, we outline the data necessary to compute a precise footprint; explain the limitations of a more naïve approach that previous calculations, including ours, have relied upon; and propose an improved approach that leverages the limited available data to improve accuracy. We then analyze the differences between methods using three sample travel datasets. Our primary improvements are the support of custom modes, region specific transit and grid mix data, and active maintenance of data dependencies. While this approach is still limited, it is an improvement over the naïve method. We plan to publish our calculations as an open-source software library, benefiting the broader research community and supporting ongoing maintenance and improvements. We call on researchers to address the data limitations by advocating for increased data collection and publication.**

***Index Terms*—energy, emissions, sustainability, footprint calculations, travel behavior, travel diary.**

I. Introduction

The UN Secretary-General, Ban Ki-Moon, has described climate change as *"the defining challenge of our age"* [1]. The International Panel on Climate Change (IPCC) has stressed that measures to lower Greenhouse Gas (GHG) emissions and limit the global increase in temperature to below 2° C are essential to avoid detriments to public health, decreased food production, and rising sea levels [2]. The IPCC models that this will require GHG emissions declining from the current estimate of $\approx$ 59 Gigatonnes to net zero by 2050 or 2070, which will require dramatic reductions across all sectors and regions.

As of 2022, transportation is the largest source of GHG emissions (28%) in the United States of America, greater than electricity generation (25%) [3]. In line with IPCC recommendations, the U.S. government has developed an ambitious blueprint to remove all emissions from the transportation sector by 2050 [4], as demonstrated in Figure 1. The blueprint outlines a comprehensive strategy, ranging from community design and land-use planning to improved vehicular drivetrains. To implement this strategy and track progress towards the goal, we need fine-grained, actionable metrics that are updated regularly to reflect the impact of changes. Federal, state, local, and individual goals will need to work in unison to achieve decarbonization, with reliable metrics to measure and coordinate progress.

However, the land-use and transportation behavior-oriented aspects of the transportation decarbonization strategy lack such fine-grained, actionable metrics. There is detailed data on fuel consumption in different drivetrains, providing actionable metrics for *vehicular energy efficiency improvements* [5]. There is detailed, longitudinal data on *electricity consumption* at the household level, enabling a thriving body of both research and practice on reducing electricity consumption through behavior change [6]–[8]. There are programs to encourage



Fig. 1: Summary figure from the U.S. Transportation Decarbonization Blueprint [4]

transportation mode shifts [9] which rely on *coarse-grained, point-in-time surveys* and a lot of extrapolation to estimate impact. The absence of a fine-grained, longitudinal, individual or household-level ***transportation meter*** can be a barrier to implementing behavior-based transportation energy efficiency strategies at scale.

Smartphone apps are a promising option for collecting user-level travel behavior data and motivating behavior change by communicating directly with participants. Several projects have incorporated trip-level emissions, either during trip planning or as retrospective carbon footprint calculators (Section II). However, the projects are either: (i) closed source implementations which do not reveal details, or (ii) research projects which highlight their tracking capabilities and/or choice architecture, but elide the details of the algorithms used to compute energy and emission impacts. **To the best of our knowledge, there is no rigorous description in the literature on the complexities involved in computing carbon footprints at an individual level.**

Concretely, this paper has three main contributions:

- it outlines the complexities involved in accounting for energy and emission impacts from transportation and highlights the data shortcomings that prevent us from achieving that ideal, both in the U.S. and globally,
- it describes the shortcomings of a naïve approach that we have used before, and that is similar to existing implementations (Section II), and
- it introduces CHEER (Calculating Holistic Energetic Emissions on Route), a more sophisticated approach that accounts for ongoing changes in fleet composition, grid decarbonization efforts, and incentive programs for mode shift, and their variations across geographical regions.

The rest of this paper is structured into five sections. We start by examining the literature for examples of carbon footprint calculations (Section II), and then outline the factors that are relevant for fine-grained metrics, inclusive of new mobility (Section III). We then recap two algorithms for computation (Section IV), outline their limitations, and highlight the difference between the calculation methods across travel datasets from three US locations (Section V). We conclude with a list of future improvements, and a call to action for better data collection (Section VI).

II. Literature Review

We conducted a detailed literature review to unearth prior examples of a transportation energy/emission footprint calculator that supports multi-modal travel. We considered both commercial and research projects, searching on app stores, looking through articles claiming to know the "best footprint tracking apps", and scouring the academic literature for deployed projects. The related work can be broadly categorized into: (a) "sketch" projects that use user-reported summary data, (b) smartphone applications for trip planning and tracking, (c) system-level emissions calculations within the transportation domain, and (d) rigorous calculations in non-transportation domains. **We were not able find prior work in the literature that dives deep into the factors that affect transportation carbon footprints and develops methods to compute them in a changing transportation context**.

A. Broad-based, self-reported motivational footprint calculators that include transportation

An increasing sense of urgency around climate change has led to a growing set of web or smartphone applications promising users a way to estimate their carbon footprints and motivate them towards a more sustainable lifestyle. However, they typically sketch out broad calculations across multiple domains based on self-reported data (Table I). These are examples of projects that could benefit from a more fine-grained calculation of transportation carbon footprint that would support habit-forming, actionable recommendations and suggestions.

B. Smartphone app-based tracking

Smartphone app-based platforms provide a unique opportunity to not only track trips and trajectories, thus providing a fine-grained, continuously updated transportation footprint, but also communicate that footprint in real-time to users, supporting actionable goals and progress towards those goals. However, most of these app platforms, including ours (NREL OpenPATH) [10], [11], focus on low-power tracking, trip processing algorithms, and user interaction. The accuracy and rigor of the carbon footprint calculation are an afterthought, and if described at all, are typically a naïve linear combination of fixed intensities with trip length (Table I). As a result, they

TABLE I: Carbon emission computations in mobility tracking apps, separated by distance reporting method. "Factors updated" based on App Store reporting of updates, dates of study, and/or date of publication.

Mobility App	Calc Method	Similar to Naïve	Distance Report	Factors Updated	Supported Modes	Notes	Currently Functional
Earth Hero [12]	Unknown	NA	Point-in-time survey/quiz	Unknown	Walk, car, bike	"building tools to help people tailor actions" [12]	✓
Commons (née Joro) [13], [14]	Y	N	Point-in-time survey/quiz	Unknown	Car, Airplane, Public transit, Trains, Bike share	Requires users to connect personal financial records	✓
Cool Climate [15]	Y	Y	Point-in-time survey/quiz	2013	Car (gas, diesel, electric), bus, rail (light, heavy, commuter), Amtrak, air (various lengths)	Holistic online calculator; recommends ways to reduce individual emissions; latest documentation is 2013, API developer docs are broken	✓/ X
Klima [16]	Y	N	Point-in-time survey/quiz	2022	Car (3 types), airplane, bus/train, motorbike, boat	Survey-based footprint estimate; built-in support for climate offset purchases	✓
PEACOX [17], [18]	Y	Y	Trip Planner	2013	Walk, bike, Metro (DART), Tram (LUAS), bus, car	Trip planner, considered active modes, transit and car, fixed, Ireland-specific EI for transit	X
Ciclogreen [19]	Unknown	Unknown	Tracking	Unknown	Walk/bike, shared car, public transport	Users starts trip and selects mode, GPS tracks trip; claims to "effortlessly quantify" emissions; must work at a partnered company	✓
UbiGreen [20]	N	NA	Tracking	2009	Walk, bike, public transit	Trips classified into "green" and "non-green" modes; user starts trip and selects mode [20]	X
Quantified Traveler [21]	Y	Y	Tracking	2012	Walk, cycle, bus, light / heavy rail, car	"Used a CO2 calculator for driving" mode; other EIs from [22]	X
MatkaHupi [23]	Unknown	NA	Tracking	2013	Car, bus, train, walk, bike	"targets behavior change through personalized challenges" [23]	X
OPTIMUM [24]	Y	Y	Tracking	2017	Walk, bike, subway, heavy rail, bus, car	6 week pilot study in Austria; heavy focus on persuasion algorithm [24]	X
Capture [25], [26]	Unknown	Unknown	Tracking	2019	Walk, car, bus, train, carpool motorbike, airplane	"predicts your carbon footprint from mobility using GPS information" [25]; apps have been removed from stores, likely defunct. Monthly fee.	X
NREL OpenPATH [10], [11]	Y	Y	Tracking	Multiple iterations from 2014 - 2024	17 modes (Table 2)	Multi-modal, automated trip tracking and carbon footprint estimation app. Approach currently used in our work; details in Section IV-A	✓

do not capture the current fluidity of the transportation landscape, with institutional incentives to encourage electrification, mode shift and land use changes.

C. Carbon footprint calculations within transportation

To better understand the effect of transportation on carbon emissions, Nocera et al. developed a framework to calculate footprint based on traffic flow in Seville, Spain [27]. After filling gaps in existing curating traffic flow data using methods such as simple imputation models and machine learning models, the carbon emissions are calculated on a yearly, per-vehicle basis that depends on the distance traveled, fuel type, size and load of vehicle, road type, average speed, road gradient, European emissions standard, and number of such vehicles on the road. Notably, the work uses models that fill in missing data from actual real-time traffic data.

Furthermore, a methodology by Liu et al. calculated passenger-level carbon emissions based on trip origin and destination, as well as travel modes [28]. Specifically, user GPS data is the first segment of their "bottom-up" approach that combines user data into road segments, and then finds the entire daily carbon emissions for the city of Hangzhou, China. Specifically, the carbon calculation is on a per-trip basis that

multiplies the travel distance of the trip by an emission factor and by a correction factor. The correction factor depends on vehicle speed and other characteristics, such as load factor for public transportation. This method also takes into account the fuel type, such as gasoline, electric, or hybrid.

At a more low-level methodology, Huang et al. utilize a simulation model to calculate realistic driving habits. The model calculates a driver's acceleration, speed, distance from the next vehicle, and the vehicle's carbon emission factor. However, for electric vehicles in particular, the equation for electricity consumption uses factors such as the vehicle weight, road slope, speed, and the resulting chemical power that is used by the cell [29]. Similar to the preceding papers, the intention was to calculate the emissions for an entire city—this case study, in particular, on Suzhou, China.

Lastly, a study by Kholod et al. surveyed the city of Murmansk, Russia in a method that the authors recommend for cities with little-to-no data availability [30]. To find black carbon emissions, the authors conducted video surveys (with cameras that record traffic flows), parking-lot questionnaires, and used preexisting models such as the Computer Programme to Calculate Emissions from Road Transport (COPERT) [31]. Then, Annual Average Daily Traffic (AADT) is used with the Euro vehicle classification standard to find diesel vehicles' effect on yearly black carbon emissions.

While these methods are useful for informing policy and providing high-level views of urban areas, there is a need to combine this broad perspective with a user-catered view, allowing individual users to better understand their own habits.

D. Complex calculations in other domains

An evaluation of the effect of emission factor assumptions on GHG emission calculations for buildings found that emissions associated with electricity consumption vary significantly by location, season, and time, as related to the grid mix and demand peaks [32]. Case studies presented for different zones in Canada indicate wide variance in average emissions factors between zones, and peak values far higher than average. OPOWER, a company delivering Home Energy Reports to customers, has "developed statistical algorithms to automate complex multi-variable regression analysis that combines energy usage, housing, and weather data to estimate the amount of heating and cooling energy used by each household" [7]. They deliver specific metrics comparing energy usage between homeowners and their neighbors, showcasing the potential benefits of increasing home energy efficiency.

Even in the transportation domain, vehicle-oriented analyses have addressed complexity and the changing mobility landscape by performing fine-grained experiments on drivetrain efficiencies to determine the optimal hybridization level [5], accounting for the charging timing [33] and geographic grid mix [34], the ambient temperature [35], or route-level characteristics [36]. These have even been used for deploying "eco-driving" driving behavior changes at scale [37], but have not been expanded to a holistic, multi-modal view of travel behavior.

To summarize, the problem of generating accurate and fine-grained, *actionable* metrics for per-user energy and emissions footprint calculations has not been studied before. We now (in Section III) outline some of those complexities, with a particular focus on factors that are impacted by strategies in the U.S. Transportation Decarbonization Blueprint [4].

III. Discussion

The calculation for an individual carbon footprint is conceptually:

$$\begin{aligned} EN_u &= \Sigma_{t \in T} EN_t \\ EN_t &= EI_t \times L_t \\ &= \Sigma_{s \in t} EI_s \times L_s \end{aligned}$$

$$\begin{aligned} EMI_s &= EI_s \times EF_s \\ EM_u &= \Sigma_{t \in T} \Sigma_{s \in t} EMI_s \times L_s \end{aligned}$$

where:

- EN_u is the energy footprint for a particular user u,
- EN_t is the energy footprint for a particular trip t in the travel history T for the user u,
- EI_t is the energy intensity of the particular trip
- L_t is the length of the trip
- EI_s and L_s are the energy intensity and length, respectively of each section s with mode m in the trip t
- EMI_s is the emissions intensity of a segment
- EF_s is the emission factor, including production and combustion, of the fuel for section s
- EM_u is the emission footprint for a particular user, computed, similarly to the energy footprint EN_u by the emissions for a section s, summed for the emissions of a trip, and again for the emissions of a user u. The main difference from EN_u is that the emission intensity for a section s is used instead of the energy intensity.

The data model used in this section is consistent with prior work focused on evaluating travel diary creation using smartphone apps [38].

A. Computing L_t

The trip level length L_t can be challenging to obtain from classic point-in-time surveys. While travelers can have a rough sense of the distances they travel on an ongoing basis ("I live 5 km from work"), they are unlikely to know the exact length (5.1 km vs. 5.25 km) or minor differences in length between alternate routes (4.5 km versus 5 km versus 5.75 km). Fortunately, distances are well-suited to automatic measurement using sensors, and vehicle odometers or bike computers have long been used to measure aggregate distances traveled by vehicles.

However, users can split their travel across multiple modes, so computing L_t requires installing sensors in all vehicles,

and linking distances across multiple modes to the same user. An easier approach would be to use a smartphone app installed on the user's phone, which would effectively have the sensor associated with the user instead of a vehicle. If the smartphone app can track location in the background, it can reconstruct all user trajectories, so computing ΣL_t for a user is straightforward. Computing the individual L_t and L_s needed for fine-grained computation requires robust trip and section segmentation algorithms. While trip segmentation algorithms using dwell times have been well-studied in the literature [39], section segmentation algorithms, crucial for multi-modal trip analysis, are less robust [38].

B. Computing EI_s and EMI_s

Computing the energy/emissions intensity of travel is significantly more complicated than computing the length. Factors affecting the energy/emissions intensity can be static (e.g. mode) or can vary across time and geography (e.g. load factor, electrification status, or grid mix). The energy/emissions intensity can even vary across the course of a single trip, depending on the presence of hills, or the maximum speed limit along the route.

1) Effect of travel mode: The primary determinant of the energy intensity is the travel mode since the EI_s, and by implication, EN_t, can be as low as zero for active transportation modes. The ability for the calculation algorithm to support a rich set of travel modes, such as distinguishing between regular bicycles and e-bikes, or between commuter trains and subways, can improve the estimate of the associated energy and emission intensity. For the highest accuracy, we would want to know the exact vehicle used in the trip, such as a hybrid bus versus a diesel bus, or an electric two-seater versus a pickup truck.

2) Single versus shared occupancy modes: Once the mode has been determined, the simplest approach is to map it to a single EI_s for the section. This can work reasonably well for single occupancy modes, where the passenger energy intensity is the same as the vehicle energy intensity. However, handling shared occupancy modes is significantly more complex because the passenger energy intensity depends on the load factor in the vehicle.

Formally,

$$EI_{s_p} = \frac{EI_{s_v}}{P}$$

where:

1) EI_{s_p}: is the energy intensity per passenger
2) EI_{s_v}: is the energy intensity of the vehicle
3) P is the number of passengers in the vehicle, also known as the *load factor*.

As we can see, the load factor can dramatically lower the energy intensity of travel - an internal combustion engine (ICE) minivan ($\approx$ 22 MPG) with 7 passengers has a higher passenger mileage ($\approx 22 \times 7 = 154$) than a battery electric vehicle (BEV) with a single driver ($\approx$ 120 MPGe). This is particularly relevant for transit trips since transit vehicles are typically larger and less energy efficient than passenger vehicles, and rely on large load factors for their passenger efficiency [40], [41].

3) Fleet composition and efficiency: A direct mapping between the trip and the vehicle in which it was taken (e.g. "2013 Nissan Leaf") can support more fine-grained and accurate calculations than a generic vehicle type (e.g. "car"). However, if we are not able to map trips to individual vehicles, particularly in shared fleet settings, then knowing the *fleet composition*, the number of vehicles of each type, and the distance traveled in each can provide an intermediate level of accuracy. This level can assist fleet managers to assess the impact of electrification initiatives [42]–[44] without significantly increasing the reporting burden.

4) Decoupling energy and emissions: When vehicles are primarily powered by fossil fuels, energy and emissions change in sync, and the pathway to reduce emissions is to reduce energy. However, electrification has the potential to reduce emissions while keeping energy consumption unchanged, as long as the new electricity comes from carbon-free sources. This implies that sustainability metrics should report on energy and emissions separately. While the emissions metric is directly relevant to climate goals, energy consumed and energy efficiency measures can also be important. This is because a reduction in energy consumed for transportation can make more energy available for other, harder to decarbonize sectors.

5) Variation across time and geography: As governments at various levels introduce initiatives to electrify vehicles, both personal [45] and fleet [43], both large [44] and small [46], energy and emission intensities will change over time. For example, fleet electrification initiatives will change the fuel mix used for travel, which will in turn affect the energy efficiency, since electric drivetrains are more efficient than traditional ICE ones. However, the impact will vary across geographies in response to varying incentives and varying levels of uptake. Similarly, the emission intensity of electrified trips also depends on the sources of electricity to the grid, and the related emissions [33] (Section III-B). This *grid mix* will change over the *long-term* as efforts to decarbonize electricity generation take effect [47]. It will also fluctuate in the *short-term*, even over the span of a few hours, as renewable energy generation ebbs and falls with weather conditions [32], [48].

6) Variation at the trip-level: Even if the mapping between the trip and the vehicle is known and fixed, factors such as the ambient temperature, number of stops and starts, and altitude gain or loss can affect trip-level energy consumption. While the effects of *eco-driving* behavior choices are smaller [49] than switching to a different mode or modifying the load factor, they can be easier to adopt. If location-based trip tracking is used to estimate trip length (Section III-A), the resulting trip trajectory can be used to account for these factors [36] and improve the accuracy of the calculations.

To summarize, the ideal calculation of the energy and emissions footprint of individual transportation is complex,

depends on availability of data, and can vary across time and geography. Our next step (Section IV) is to discuss a practical implementation of this calculation that outlines the challenges in greater detail and provides a practical approach to addressing them.

IV. Data Sources and Algorithms for the Computation

If it is specified at all, the most common approach to estimating energy and emissions footprints is a linear combination of energy and emission factors with trip length (Section II). This is conceptually similar to the ideal computation (Section III), but elides several of the complexities involved in estimating energy and emission factors. In particular, the set of supported modes is limited, and the intensity factors are typically fixed based on the time and geography for which the project was created, leading to inconsistencies. This section highlights how we used this naïve approach in our own work (Section IV-A), the limitations we uncovered (Section IV-B), the new approach that we evolved to address them (Section IV-C), the data sources used (Section IV-D), and the limitations (including lack of data) of the new approach (Section IV-E).

A. Naïve approach summary

The naïve approach, consistent with the related work (Section II), leverages the recorded distance and mode, mapped to fixed energy/emission intensities to estimate CO_2 emissions generated by a given trip (Table II). The naïve approach is the one that we first adopted in our own work, and one that we have evolved organically over time to take on additional layers of sophistication.

2014 : The calculations used only emission factors, and were based on LCA estimates from the literature [34]. The estimate was from 2008, and was based on underlying data sources from 2004-2006. The carbon footprint was displayed in-app, and based on simple modes automatically determined from sensor data (walk,bike,car,bus,train,air)[a]. The platform did not allow users to specify trip-level mode labels at the time.

2019 : A group in Europe made an open-source contribution, adding support for emission intensity factors from TREMOD [50] and HBEFA [51], and the ability for the user to choose the mapping to use in the computation[b].

2020 : As part the analysis of an e-bike program in Colorado funded by the Colorado Energy Office (CEO) [52], the platform began to support a richer set of user specified, trip-specific mode labels, including e-bikes, carpooling and taxis. The analysis incorporated both energy and emission impacts, and refreshed the intensity factors with recent data from the Transportation Energy Data Book (TEDB). It also supported new mobility modes such as e-bikes. To synchronize energy and emission factors, the computation treated the energy intensity as primary [c] [d] and converted to emission intensity based on a static fuel emissions table (Table III)[e]. This computation was used to generate reports [52], [53] and an in-review paper [54]. It was also used to create an automatically updated public dashboard of energy and emissions impact.

2021 : Incorporated trip-specific mode labels into the in-app dashboard.[f]

2022 : Added support for e-cars[g] and hybrid cars[h], specified through additional options for user labels.

B. Limitations of the naïve approach

Although our current implementation of the naïve approach is significantly more sophisticated than the literature (Table I), it does not address its fundamental drawbacks. The evolution of our implementation showcases these challenges, which fall into three broad categories.

It supports a small, fixed set of modes:
A simple lookup table that maps modes to energy intensities works well when the set of modes is fixed. But as new mobility modes (e.g., e-bikes, on demand transit, hydrogen cars...) are introduced, capturing the differences between them can improve the accuracy of the computation (Section III-B1). A generic computation method should also support modes that are popular in some countries, such as funiculars in hilly regions and motorcycles in Asia.

It is specific to the region that was the focus of the original data collection:
While creating the mapping for a project, it is tempting to use the local region (e.g., SF Bay Area in 2014 and Denver in 2021). However, particularly for shared mobility options (Table II), there can be significant regional variations in both fleet composition and load factor, which can have a large impact on energy and emissions intensity (Section III-B2).

It is static and does not incorporate changes over time to the EI or EMI factors:
As the strategy outlined in the U.S. Transportation Decarbonization Blueprint [4], and similar efforts documented in other Nationally Determined Contributions (NDCs) take effect, we expect to see differences in land use, mode choices, fleet composition, and electricity generation (Section III-B5) that will affect the energy intensities of the modes used.

[a] https://github.com/e-mission/e-mission-server/blob/27e0a7ac1bfec22eef8e92142e71fcbc74406683/front/server/carbon_calc_details.html

[b] https://github.com/e-mission/e-mission-phone/pull/597

[c] https://github.com/e-mission/em-public-dashboard/pull/9

[d] https://github.com/e-mission/em-public-dashboard/blob/8e0c6b056fa8d4891842c31de783e74bb8153948/viz_scripts/auxiliary_files/energy_intensity.csv

[e] https://github.com/e-mission/em-public-dashboard/pull/26

[f] https://github.com/e-mission/e-mission-docs/issues/688

[g] https://github.com/e-mission/e-mission-phone/commit/2cbcb8198e45ef6d426a3a3003d92fe2b065d262

[h] https://github.com/e-mission/e-mission-phone/pull/1075

TABLE II: Energy Intensity by Mode

Mode	(App, 2014)	(Dashboard, 2020)	CHEER	App & Dashboard Notes	CHEER Notes
ICEV	172.78 g/pkm	941.49 Wh/pkm[a]	898.77 Wh/pkm[a]	Mean of cars and trucks from TEDB	Mean of GREET[b] baseline gas car and gas truck
HEV	NA	NA	625.22 Wh/pkm[a]	Did not represent non-gas vehicles	Mean of GREET grid-independent gas HEV car and truck
PHEV	NA	NA	0.58×289.1 Wh/pkm + 0.42×663.5 Wh/pkm[a]	Did not represent non-gas vehicles	GREET grid-connected PHEV electric and gas modes, weighted by operation assumption, mean of cars and trucks
BEV	NA	NA	286.8 Wh/pkm[a]	Did not represent non-gas vehicles	Mean GREET electric car and electric truck
Ridehail	NA	1,313.71 Wh/pkm	1,198.36 Wh/pkm	Productive travel assuming 1.5 riders + equal distance deadheading [55]	Productive travel with assumed 1.5 riders + deadheading of 2/3 distance [56]
Train	57.17 g/pkm	227.64 Wh/pkm	NTD[c] lookup	Denver light rail specific from TEDB Figure 7.1	Varies by city; e.g., 144 Wh/pkm in NYC, 419 Wh/pkm in Chicago
Bus	165.94 g/pkm	830.4 Wh/pkm	NTD lookup	"Transit bus" from the TEDB	Varies by city; e.g., 618 Wh/pkm in NYC, 1074 Wh/pkm in Chicago
Air	134.86 g/pkm	NA	493.91 Wh/pkm	Non relevant to mode-shift focused calculations	Average passenger plane energy per kg, assumed passenger + luggage 90kg
Walk	0	0	0	human-powered modes	human-powered modes
Bike	0	0	0		
E-bike	NA	13.67 Wh/km	13.67 Wh/km	From [57]	From [57]
Bikeshare	NA	0 Wh/km + 90 Wh/trip	0 Wh/km + 90 Wh/trip	In 2020, Denver only had human-powered bikeshare; rebalancing penalty	Assuming human-powered bikes with energy for rebalancing
Scootershare	NA	16.78 Wh/km + 4.1 Wh/trip	16.78 Wh/km + 4.1 Wh/trip	From [57]	From [57]

[a]If user selects "shared ride," an assumed load factor of 2 is used, and energy intensity is cut in half.
[b]Greenhouse gases, Regulated Emissions, and Energy use in Technologies [58].
[c]National Transportation Database [59].
[d]Transportation Energy Data Book [60].

C. CHEER 2024: Improvements to support custom modes, time, and geographic variations

In this first iteration (CHEER 2024), we focus on improvements that can address the limitations of the naïve approach while using existing data sources:

Fixed modes → custom modes In addition to expanding the default set of modes to include hybrid and battery powered vehicles, deployers can introduce custom modes when adapting the platform to a new region. The custom set of modes can include modes left off of the default list, such as funiculars or motorcycles, and will include energy and emissions intensity for each novel mode. They can also adapt existing default modes to specific contexts with additional energy or emissions data. The data collection layer can then offer custom modes as label options to users, and when passed in, CHEER will use the specified factors in calculations.

Single region → region-specific factors CHEER now derives transit efficiencies and load factors from the NTD (National Transportation Database) based on the location of the trip. For a given trip, the algorithm aggregates the efficiencies for the labeled mode (bus, train) between transit agencies in that area based on ridership. CHEER determines the load factor by the ratio of reported passenger miles to vehicle miles. A similar aggregation is performed for emissions if multiple fuels are used for the mode (bus, train) in that area. CHEER obtains the emissions due to electricity from eGRID (Emissions & Generation Resource Integrated Database) based on the location of the trip. This update accounts for the fact that in places where renewable energy has a higher presence in the grid mix, emissions are lower than places where electricity is generated by burning fossil fuels.

Snapshot → ongoing updates The improved calculations currently use the most recent data from eGRID (2022) [61], the NTD (2022) [62], and GREET (2023) [58]. Each of these data sources is regularly updated, and we will update our calculations to use the new data as it is released. Maintaining grid mix, transit fleet, efficiency, and emissions data will ensure the calculations reflect continuous improvements in renewable energy availability, fleet composition and utilization, and vehicle and fuel technologies. Other changes in mode

TABLE III: Emissions Factors by Fuel Type

Fuel Type	(Dashboard, 2020)	Calibrated Approach	Naïve Notes	Calibrated Notes
Gasoline	243.3 kg/MWh	324.183 kg/MWh	EIA (2020)	based on passenger cars, average across ICEV, HEV, and PHEV (CD mode)
Diesel	249.65 kg/MWh	325.073 kg/MWh	EIA (2020)	based on transit buses
Jet Fuel	NA	304.354 kg/MWh		based on passenger (non-freight) plane types
CNG	NA	271.024 kg/MWh		based on transit buses
Hydrogen	NA	332.852 kg/MWh		based on transit buses
Electricity	528.89 kg/MWh	eGRID lookup [61]	EIA data for Rocky Mountain region	Varies by region; e.g. 125 kg/MWh in upstate NY, 723 kg/MWh in Puerto Rico

representations, such as the updated research that reduced the deadheading proportion from 50% [55] to 40% [56] of all ridehail driving, can be monitored and reflected in further iterations of the calculations.

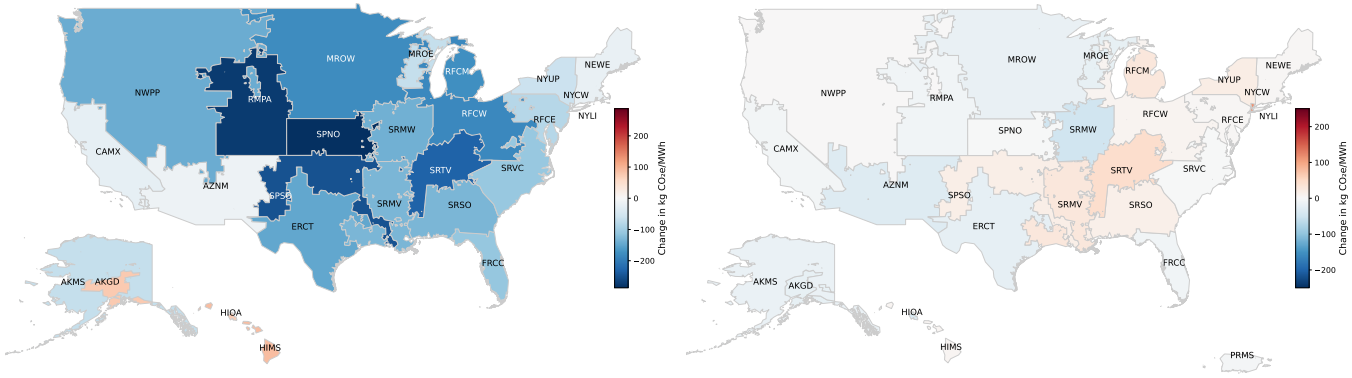
D. CHEER Data Sources (eGRID and NTD)

To achieve the specific emissions depending on a trip's location and time, CHEER retrieves emissions intensity data from eGRID (Emissions & Generation Resource Integrated Database) provided by the Environmental Protection Agency at the subregion level and transit data the National Transit Database (NTD) provided by the Federal Transit Administration at the Urban Area Census (UACE) code level.

Figure 2 shows the United States' subregion change in grid cleanliness, as measured in kilograms of carbon dioxide emissions per megawatt hour of electricity generated, over time, illustrating the subregion and year granularity of eGRID data used in CHEER calculations. Before eGRID annual emissions data can be visualized or used in trip-level calculations, it must be processed to standardize across the varied publishing formats. These methods are included in the CHEER codebase. Figure 2a depicts the change in grid cleanliness from the year in which the App (2014) method was developed to the year the Dashboard (2020) method was developed. Figure 2b demonstrates a change from the year that Dashboard (2020) was developed to the latest year that eGRID provides, which is 2022 at the time of writing. The large magnitude of these shifts demonstrates the need for dynamic factors to calculate carbon emissions.

The NTD is a vast repository, with dozens of features reported by transit agencies across the United States. In order to narrow this data into a format which is easily managed in runtime calculations, CHEER provides a mechanism to download NTD data, ingest it into a standardized readable format, and store it in JSON (JavaScript Object Notation, a format that is both human and machine-readable) files. The process for calculating efficiency figures for each transportation agency and mode is documented below. Figures 3 and 4 exemplify the efficiency metrics, including the underlying occupancy data, by mode, fuel type, and UACE for select states.

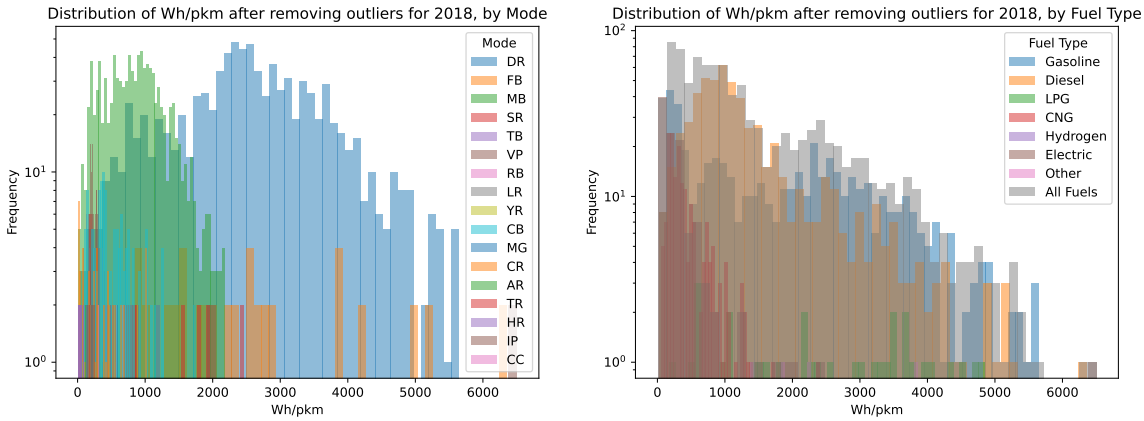
Figure 3 uses NTD data to showcase efficiency on per-mode and per-fuel basis. The list of all modes provided by the NTD and their corresponding acronyms are enumerated in Table IV. Within the per-mode figures, a higher watt hour per passenger kilometer (Wh/pkm) figure signifies less efficient transportation. Demand response (DR) tends to have



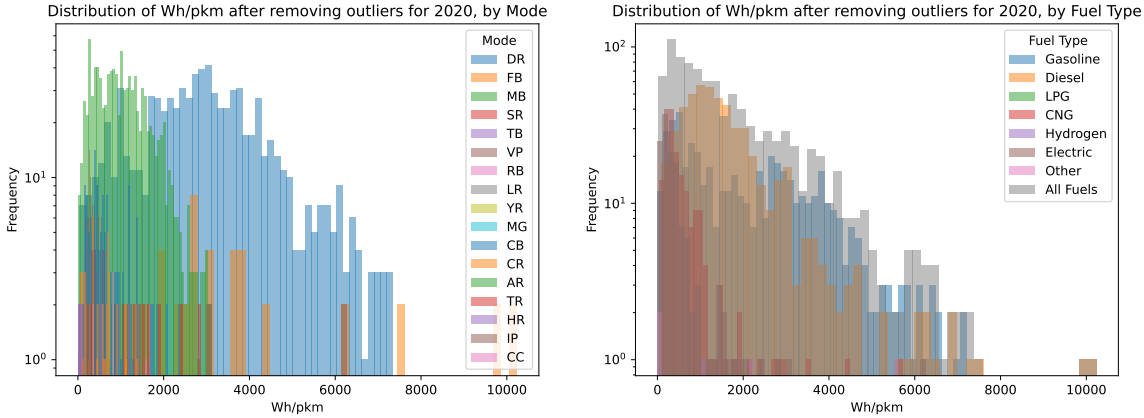
(a) Change in CO_2 Intensity from 2014 to 2020

(b) Change in CO_2 Intensity from 2020 to 2022

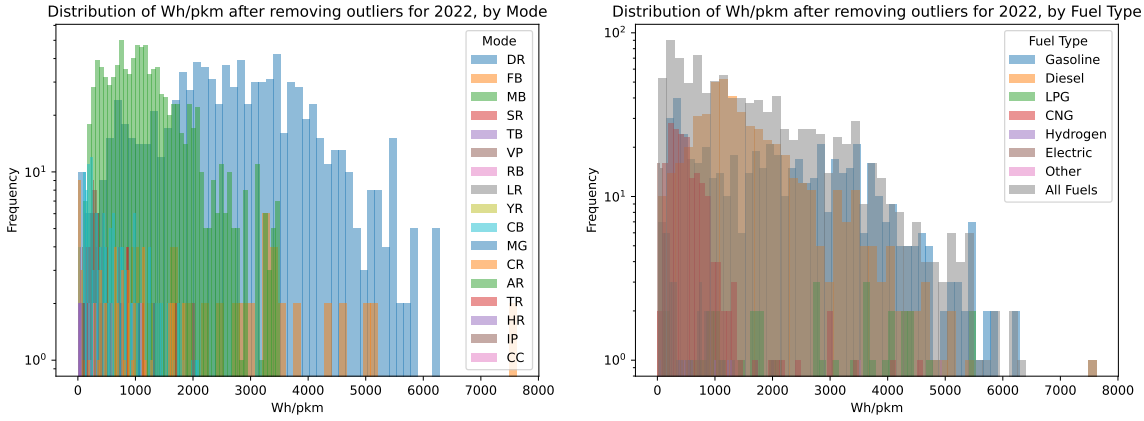
Fig. 2: Delta comparison of grid cleanliness in eGRID subregions over two periods (kg CO_2e/MWh).



(a)



(b)



(c)

Fig. 3: Distribution of Wh/pkm after removing outliers for selected years (2018, 2020, 2022).

Mode Name	Acronym
Aerial Tramway	TR
Alaska Railroad	AR
Bus	MB
Bus Rapid Transit	RB
Cable Car	CC
Commuter Bus	CB
Commuter Rail	CR
Demand Response	DR
Ferryboat	FB
Heavy Rail	HR
Hybrid Rail	YR
Inclined Plane	IP
Jitney	JT
Light Rail	LR
Monorail/Automated Guideway	MG
Público	PB
Streetcar Rail	SR
Trolleybus	TB
Vanpool	VP

TABLE IV: NTD transit modes and acronyms. [63]

a higher energy cost per mile whereas buses (MB) tend to be more efficient. Further, the per-fuel figures show that vehicles operating on compressed natural gas (CNG) are more efficient whereas gasoline vehicles are less energy efficient.

Occupancy, or average number of passengers, energy efficiency, and fleet composition are gathered from the NTD for the resulting `All Fuels (Wh/pkm)` (watt hours per passenger kilometer) metric which is stored in the JSON files and used by CHEER when estimating emissions for a transit trip. The efficiency metric is calculated as follows:

1) Obtain amount of fuel consumed for each fuel, convert from units of fuel (e.g. gallons of gasoline) to kWh.
2) Obtain vehicle distance traveled for each fuel type, convert to kilometers.
3) Obtain total passenger distance traveled for fuel type, covert to kilometers.
4) Derive intensity by dividing energy consumed by vehicle distance in kWh/km.
5) Divide passenger distance by vehicle distance to yield Average Passengers.
6) Divide kWh/km by Average Passengers to get passenger-kilometers.
7) To calculate a weighted efficiency for all fuels:
   - Divide each individual fuel's distance by total distance for an agency to get fuel weight.
   - Multiply fuel efficiency (kWh/km) by weight.
   - Sum weighted fuel efficiencies across all fuels.

The UACE represents areas with multiple agencies so the values are weighted by the number of passenger trips taken under each agency. Once the data has been normalized, the trends over time are standardized into a digestible format. Figure 4 shows the change over time in both NTD-reported occupancy and overall efficiency from 2018 to 2022, using three NREL OpenPATH programs' locations: Colorado, North Carolina, and Massachusetts. Both changes in occupancy, as a proxy for utilization, and changes in the fleet composition, such as replacing diesel buses with electric buses, contribute to these changes.

Figure 4b shows that there are areas with buses that do not provide data about their fuel and energy use to the NTD. For instance, Mesa County, Colorado operates buses in Grand Junction; however, they do not report such information, whereas Denver's bus system does [59]. Therefore, the UACE for Grand Junction (at latitude -39 and longitude -108.5) is greyed out, showing No Data. CHEER handles these data gaps by defaulting to the national average for that transportation mode and year, which is calculated with the NTD data.

E. Limitations of CHEER 2024

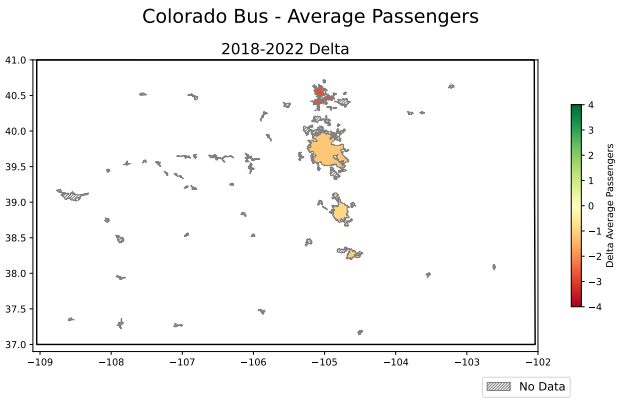
The improved approach is more accurate and generalizable than the naïve approach but it also has limitations. These can be broadly categorized into: (i) limitations that can be addressed by a more sophisticated algorithm and (ii) limitations that need additional data sources.

1) Limitations in algorithm sophistication: Examples of assumptions in the algorithm that can be addressed by improved methods include:

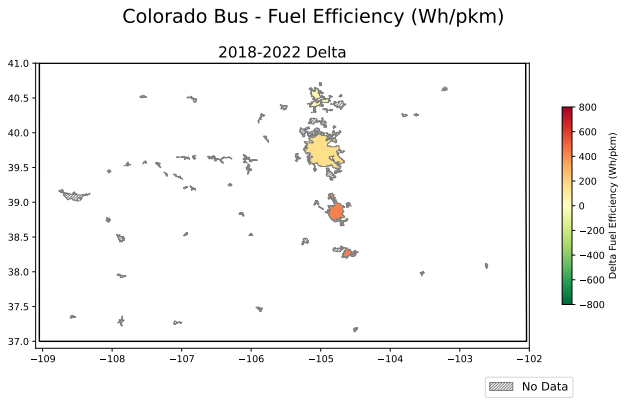
1) using marginal impacts where appropriate, including in transit and replaced mode calculations,
2) accounting for fleet composition in micromobility modes as well,
3) accounting for trip level variability using trajectories and ambient temperature, and
4) improving the trip $\leftrightarrow$ vehicle mapping for personal vehicles through user-specific configuration.

2) Limitations on quality and quantity of data: However, limitations in the freshness and availability of data are harder to address. For example:

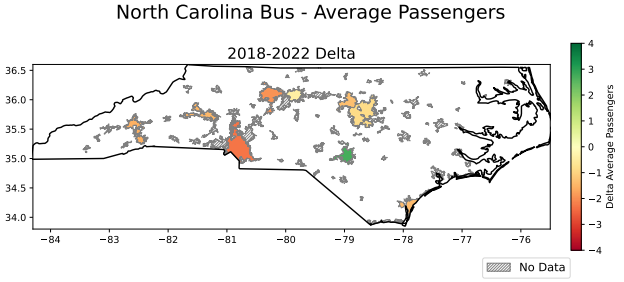
1) both the eGRID and NTD databases currently publish data from 2022. This multi-year lag in reporting means that the energy and emission factors computed using them will need to be extrapolated and would ideally be marked as provisional until the final numbers are released,
2) the NTD only provides aggregated data at the *vehicle type* granularity for distance and at *service type* for ridership. This elides the differences in load factor between highly popular and less popular routes, and complicates a full understanding of energy efficient system design,
3) while the NTD data is stale and coarse-grained, similar data in other regions of the world is non-existent. We could not find open data, even at aggregate, country level, on public transit ridership, although we checked sources such as the World Bank or the United Nations. The only source that we could find was from the UITP [64], but: (i) their data is available to members only, (ii) their most recent bus fleet data (2023) is only from Europe, and (iii) their last Global Bus Survey is from 2019, and includes only data sampled from "selected operators". Fortunately, grid mix data is available from the International Energy Agency (IEA) [65] although it is only available at the country level and is not as fine-grained as the US sources,



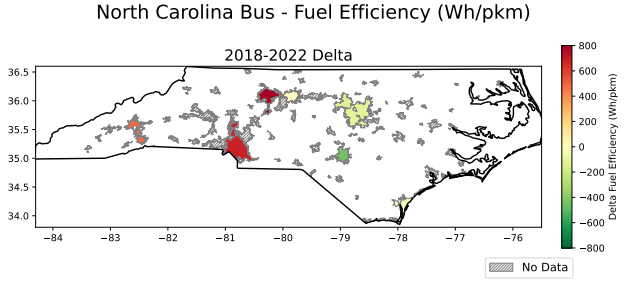
(a)



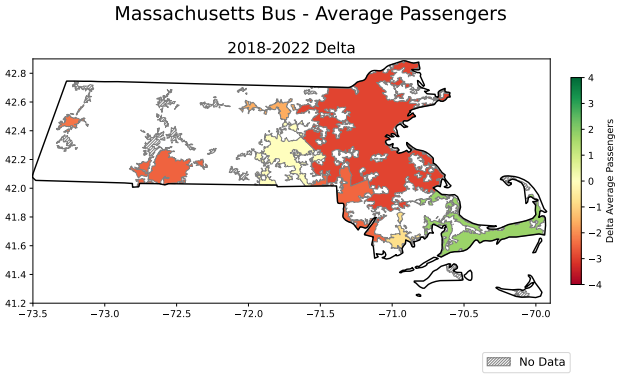
(b)



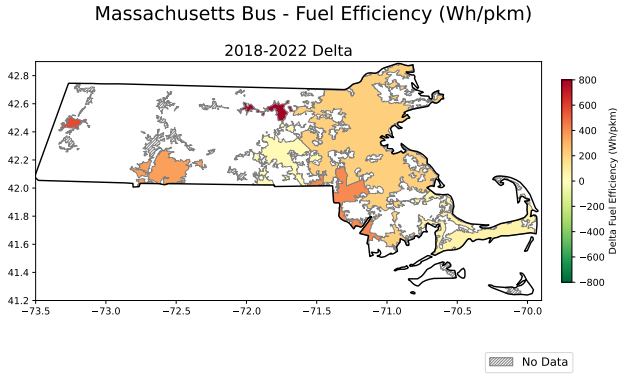
(c)



(d)



(e)



(f)

Fig. 4: Change in occupancy and fuel efficiency by state for 2018-2022 using NTD data [59].

4) we currently have no way of capturing the time at which users charge their electric or hybrid vehicles, so there is no way to account for the difference in grid emissions during peak versus off peak hours. Even if we captured the charging time, eGRID does not offer time-specific data. More detailed data about charging habits and grid emissions would be required to properly account for grid changes within a day,
5) data from eGRID is aggregated into "eGRID subregions" which may not accurately characterize every location. The power profiler determines each region's mix by assigning primary and secondary power sources for each zip code. A given power source may be the main power *delivery* system for a zip code, but the community may be utilizing an alternative, clean, electricity source [66], [67], and
6) as noted in Table II, estimates of deadheading penalties for ridehailing services vary. General estimates derived from limited data sources [55], [56] generalize data which varies by location. If ridehail providers shared data, even if highly aggregated, for service areas, these estimates could be spatially resolved and more reliable.

V. Results

To better understand the impacts of the differing calculation methods, we leverage three different programs' datasets from our app-based travel diary collection platform, NREL OpenPATH. Carbon emission calculations are performed on trip information from three datasets: CanBikeCO [53], [54], Bull eBike (Durham, NC) [68], and MassCEC (Massachusetts Clean Energy Center) [69]. The data was retrieved from the Transportation Secure Data Center [70].

To begin, we analyze data from a program titled CanBikeCO, created with the intention of providing e-bikes to essential workers within Colorado during the COVID-19 pandemic [53], [54]. This program provides trip data that can be analyzed using both the naïve and CHEER emission calculation methods. The change in emissions intensity over time, calculated using the CHEER method with CanBikeCO data (kg CO_2 per km), is shown in Table V.

TABLE V: Average CHEER CO_2 emissions in kg per km by mode and year for CanBikeCO.

	2021	2022	% Change
Bus	0.5518	0.4583	-16.9461
E-bike	0.0072	0.0070	-2.9286
Free Shuttle	0.5408	0.4532	-16.1851
Gas Car, drove alone	0.2914	0.2914	0.0000
Gas Car, with others	0.1457	0.1457	0.0000
Scooter share	0.0094	0.0085	-9.0866
Taxi/Uber/Lyft	0.3885	0.3885	0.0000
Train	0.0984	0.0958	-2.6568

Between 2021 and 2022, nearly all modes decreased in emissions intensity. Rates decreasing from year-to-year are not necessarily solely because of cleaner energy, but can also be due to changes in utilization or efficiency. Referring to Figure 2b, this notion is corroborated by eGRID data, where there was a slight decrease achieved in emissions due to electricity generation within the Rocky Mountain Power Area (which includes Colorado) from 2020 to 2022.

Table VI demonstrates the percentage difference in emissions intensity (kg of CO_2/km) between naïve and CHEER.

TABLE VI: Percentage differences in emissions intensity (kg/km) by mode for each dataset between naïve and CHEER, for the average trip taken of that mode.

Mode	CO	NC	MA
Bus	146.51%	67.98%	64.78%
E-bike	-1.52%	-46.24%	-53.60%
E-car, drove alone	–	29.53%	-14.42%
E-car, with others	–	27.96%	-14.42%
Free Shuttle	92.45%	87.40%	53.93%
Gas Car, drove alone	27.20%	27.20%	27.20%
Gas Car, with others	27.20%	27.20%	27.20%
Scooter share	-20.29%	-46.31%	–
Taxi/Uber/Lyft	21.54%	21.54%	21.54%
Train	-29.14%	-72.76%	-29.69%

CanBikeCO's cumulative bus emissions (signified by CO in the column name) are an example of a stark difference between the two methods. To understand why buses were found to have 146.51% higher emissions intensities with CHEER, we perform calculations to yield the bus energy intensities (kg/km CO_2) across methods using the CanBikeCO program. The calculations proceed as follows:

- We calculated the bus energy intensity with CHEER to be 0.32487 kg/km for the Denver-Aurora, CO area.
  - First, the transit intensities are fetched from the previously downloaded and parsed NTD data by looking up the UACE for a particular set of coordinates.
  - Intensities corresponding to the trip mode are retrieved from the set of intensities.
  - The mode intensities are weighted according to their frequency within the bus fleet for that UACE.
  - During 2021, the UACE code 23527 (Denver-Aurora, CO) has a weighted Wh/km of 992.54:

$$\begin{aligned}\text{Overall Wh/km} &= (\text{Diesel Wh/km} \times \text{Diesel Weight}) \\ &\quad + (\text{Electric Wh/km} \times \text{Electric Weight}) \\ &= (999.237 \times 0.989) \\ &\quad + (390.449 \times 0.011) \\ &= 988.245 + 4.295 \\ &= 992.54 \text{ Wh/km}.\end{aligned}$$

  - A lookup dictionary is used to convert megawatt hours to kilograms of CO_2 depending on fuel (325.073 kg/MWh for diesel, taken from Table III). However, if the fuel-type is electric, the kg/MWh is fetched from eGRID, which is 528.994 for the Denver-Aurora UACE.

– Using the overall weighted energy intensity (992.54 Wh/km), we calculate the carbon emissions in kg per kilometer:

$$\begin{aligned}\text{kg/km } CO_2 &= \frac{\text{Overall Wh/km}}{1000} \times \text{Weighted kg/MWh}\\ \text{Weighted kg/MWh} &= (0.989 \times 325.073_{\text{Diesel}})\\ &\quad + (0.011 \times 528.994_{\text{Electric}})\\ &= 321.497_{\text{Diesel}} + 5.819_{\text{Electric}}\\ &= 327.316\,\text{kg/MWh}\\ &= 0.99254\,\text{kWh/km} \times 327.316\,\text{kg/MWh}\\ &= 0.99254\,\text{kWh/km} \times 0.327316\,\text{kg/kWh}\\ &= \underline{0.32487\,\text{kg/km}}.\end{aligned}$$

It is important to note that CanBikeCO participants took trips outside of Denver, where the grid generation differs and trips outside of any UACE entirely, where specific NTD data is not available. Within the dataset, 83.55% of bus trips defaulted to the national average (0.3104 kg/km) as their starting coordinates did not fall within the bounds of a UACE code. 10.46% of bus trips occurred in the Denver-Aurora region (UACE 23527), 2.68% were in the Fort Collins area (UACE 30628), and 1.31% were in the Los Angeles–Long Beach–Anaheim metropolitan area (UACE 51445). Remaining trips took place in other UACEs but are too negligible in frequency to note. Because of the diversity of trips within datasets, this change in average emission across areas reinforces the need for spatially attuned calculations.

Now, we perform the calculation for the Dashboard (2020) naïve emission method.

- We calculated the bus energy intensity with the naïve method to be 0.2073 kg/km.
  - Buses were assumed to use diesel fuel, with an energy intensity factor of 4560 and a CO_2 factor of 161.3, taken from Table 2.14 of the TEDB [60].
  - The CO_2 emission formula for diesel buses per mile is given by:

$$\begin{aligned}CO_2\ \text{(lbs/mile)} &= 4560\,\text{BTU}_{\text{PMT}} \times 10^{-6}\,\text{MMBTU}_{\text{BTU}}\\ &\quad \times 161.3\,\text{lb}_{\text{CO2/MMBTU}}\\ &= 0.735528\,\text{lbs/mile}.\end{aligned}$$

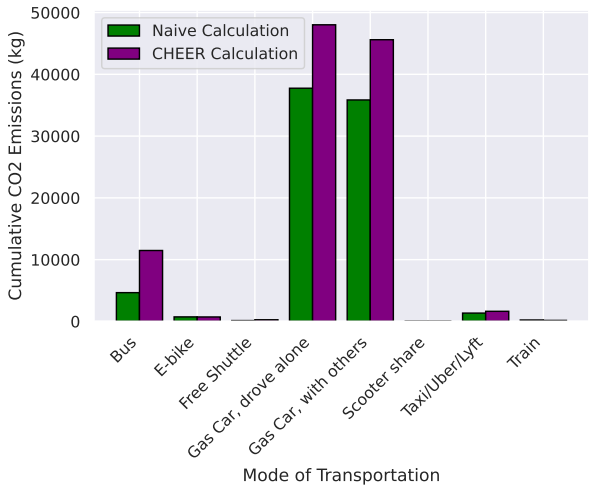
  - Convert from lbs/mile to kg/km:

$$\begin{aligned}CO_2\ \text{(kg/km)} &= CO_2\ \text{(lbs/mile)} \times \frac{0.453592\,\text{kg}_{\text{lb}}}{1.60934\,\text{km}_{\text{mile}}}\\ &= 0.735528 \times \frac{0.453592\,\text{kg}_{\text{lb}}}{1.60934\,\text{km}_{\text{mile}}}\\ &= 0.735528 \times 0.2818_{\text{Conversion Factor}}\\ &= \underline{0.2073\,\text{kg/km}}.\end{aligned}$$

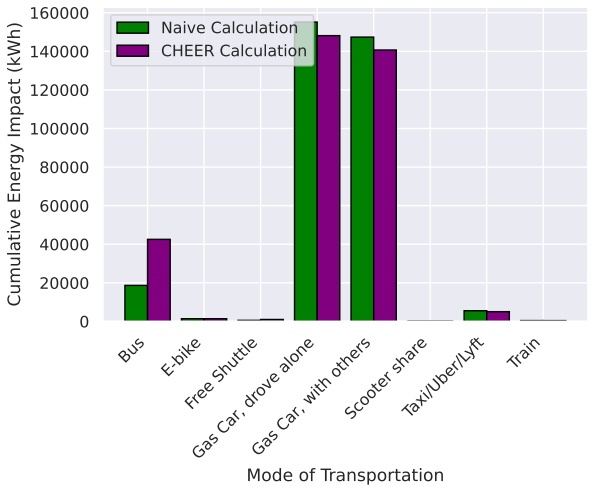
The naïve Dashboard (2020) method underestimates the actual energy intensity compared to CHEER. Throughout the programs' durations, CHEER was able to account for the decrease in ridership during the COVID-19 pandemic, explaining why the CO_2 emissions were considerably higher than the naïve calculation. Specifically, for the naïve method, the average number of passengers was assumed to be 7.7 as per Table 2.14 in the TEDB [60]. On the other hand, CHEER utilizes yearly data from the NTD; for 2021, it reported a national average of 5.76 passengers on buses. Thus, the watt hours per passenger kilometer is found to increase when utilizing CHEER. Referring to Figure 4, the buses within the surveyed states experienced lower efficiency due in part to lower average passengers utilizing the buses. Since efficiency has a direct impact on energy intensity, this explains the change from naïve to CHEER.

Another aspect to consider, aside from the energy intensity, is the cumulative amount of CO_2 emitted and energy consumed over the entire program. Nuances and variations in calculation methods cause significant jumps in cumulative CO_2 and kWh amounts. Figure 5 shows the cumulative amount of CO_2 generated over the entire data collection period, for each transportation mode, with both naïve and CHEER methodologies. This figure does not show the Air or Other modes, as they outweigh the remaining modes and/or are not commonly used. Further, the magnitude of the cumulative amounts differ as the datasets ran for various amounts of time, but the changes across methods remain proportional. As shown in Figure 5, CHEER tends to calculate higher cumulative CO_2 kg emissions in nearly every mode and dataset, while the opposite is true within cumulative kWh impact. In the energy impact plots, the Dashboard (2020) naïve method is higher, with notable exceptions for Bus, Free Shuttle, and Train. Since CHEER finds that cars tend to output more CO_2 and use slightly less kWh than previously thought, the carbon intensity changes. However, any slight decrease in the energy in kWh used is negated by the carbon intensity, showing the need for cleaner grids.

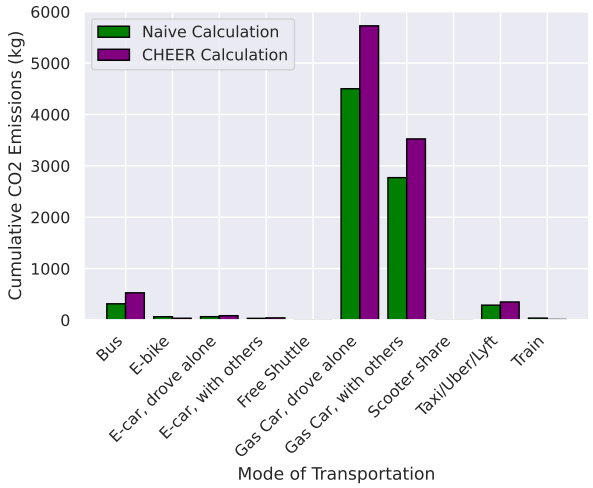
Figure 6 shows the average CO_2 emissions intensities in kg per km for every year that the program was running. Notably, across the years for each dataset, the naïve calculation stays the same because the factors do not change. On the other hand, since CHEER considers the change in the energy grid, the Bus and Free Shuttle modes (among others) are dynamic from year to year. Since distance can be a confounding factor, this comparison isolates the mode's intrinsic efficiency from the trip length. Considering this, the figure demonstrates CHEER calculating a lesser CO_2 per km from cars than buses. If even more people rode the bus, the emissions attributed to each passenger would continue to decrease. As outlined in Section IV-E1, using marginal intensities for shared modes is one way CHEER's methods could be improved. Accounting for utilization, more individuals drive cars alone, and so the cumulative emissions from cars are substantially higher, as demonstrated in Figure 5. Accounting for factors such as energy grid and utilization marks a significant change in the algorithm's accuracy, different from previous iterations.



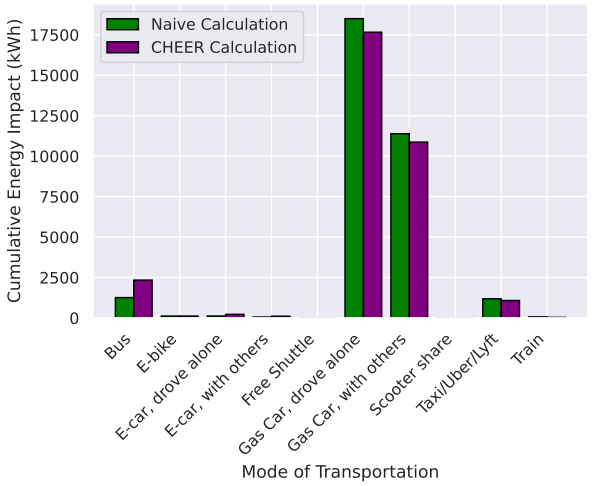
(a) Cumulative CO_2 Emissions (kg) for CanBikeCO



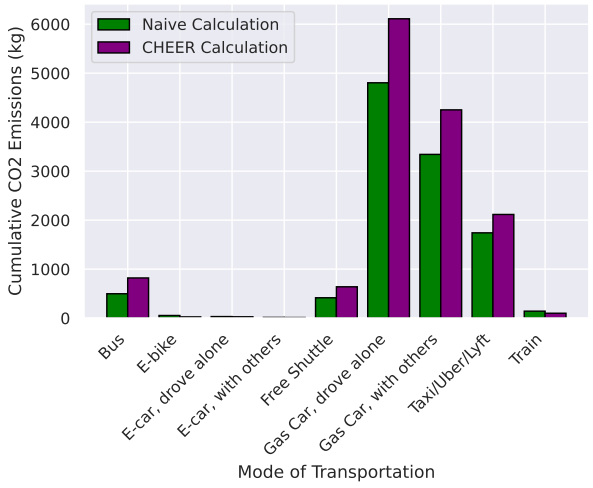
(b) Cumulative Energy Impact (kWh) for CanBikeCO



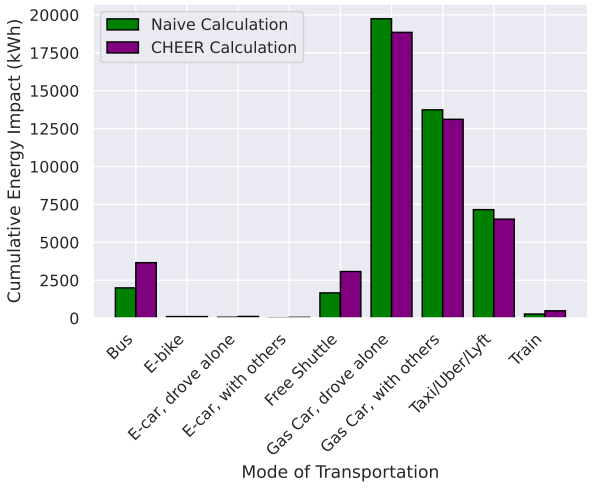
(c) Cumulative CO_2 Emissions (kg) for Bull (Durham, NC) eBike



(d) Cumulative Energy Impact (kWh) for Bull eBike (Durham, NC)

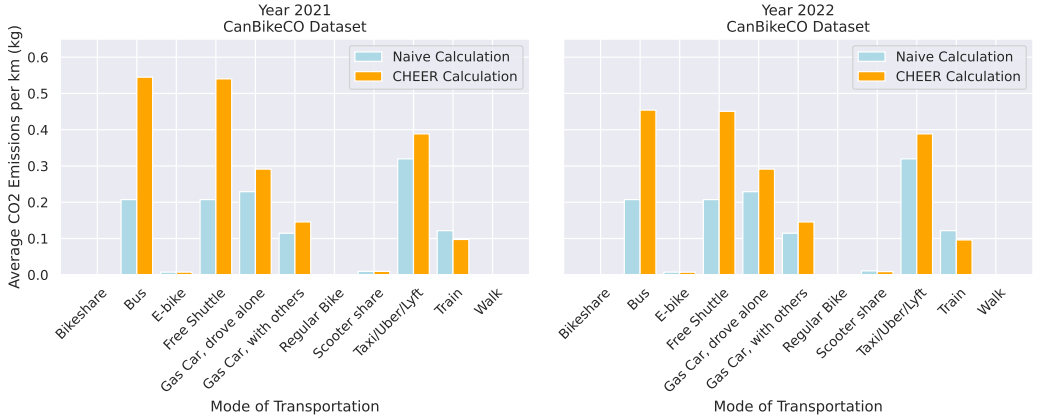


(e) Cumulative CO_2 Emissions (kg) for MassCEC

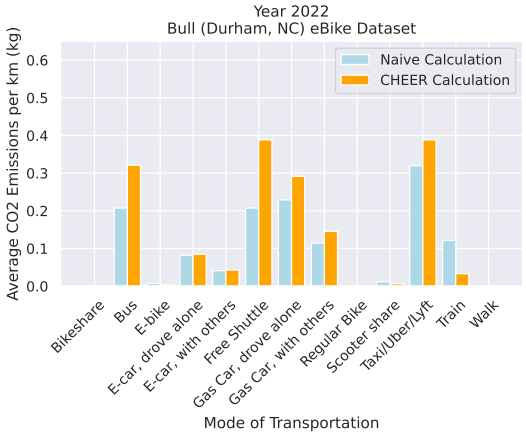


(f) Cumulative Energy Impact (kWh) for MassCEC

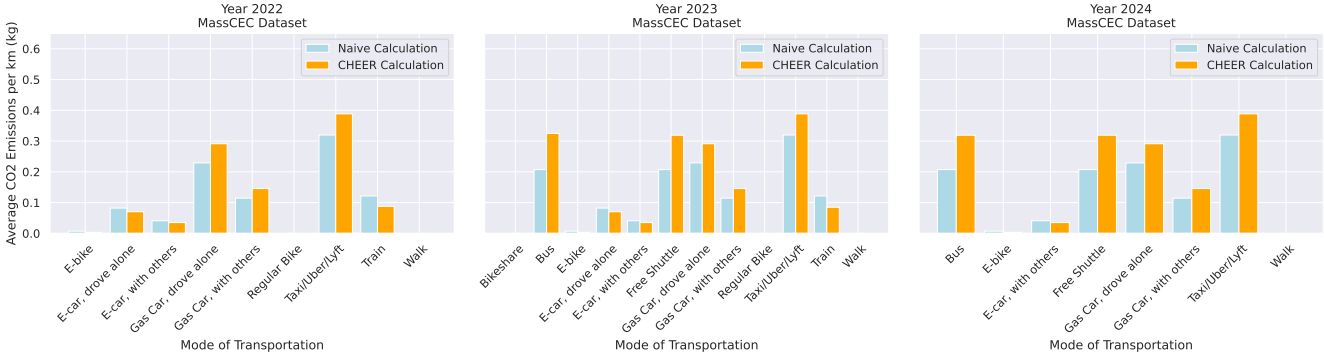
Fig. 5: Comparison of Cumulative Emissions (kg and kWh) across datasets for naïve and CHEER calculations.



(a) CanBikeCO

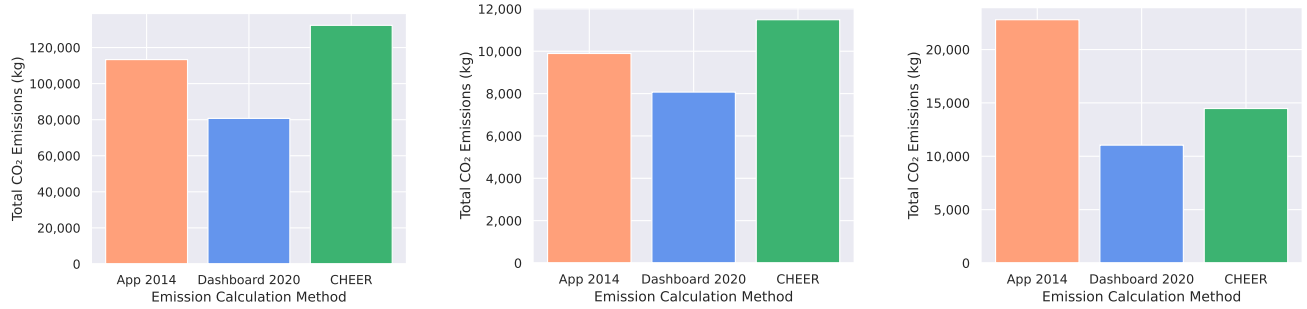


(b) Bull (Durham, NC) eBike



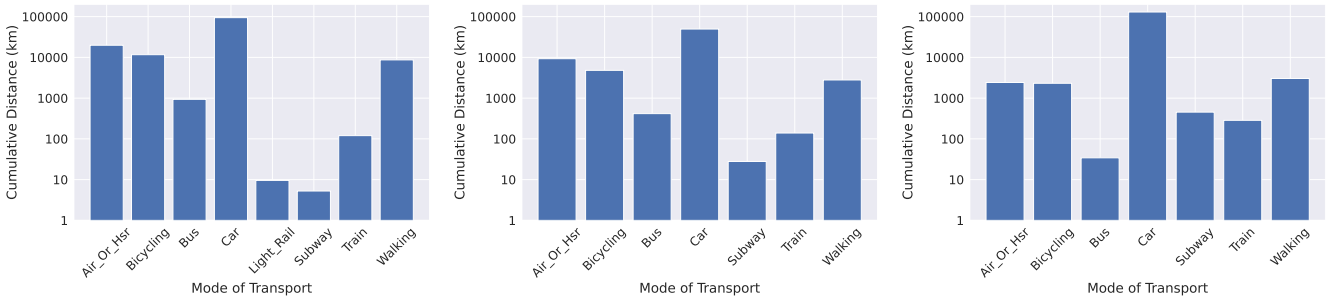
(c) MassCEC

Fig. 6: Comparison of yearly CO_2 emissions per km across different datasets.



(a) CanBikeCO (b) Bull Durham, NC (eBike) (c) MassCEC

Fig. 7: Comparison of three emissions calculation methods for each program.



(a) CanBikeCO (b) Bull eBike in Durham, NC (c) MassCEC

Fig. 8: Cumulative distance km (App, 2014) in logarithmic scale, for each mode, across each program.

To compare all iterations of our algorithms, Figure 7 compares three methods of calculation: App (2014), Dashboard (2020), and CHEER. This figure has higher cumulative values than the sum of the individual calculations in Figure 5 because it also factors the trips classified as Air or Other mode. Within this comparison, Dashboard (2020) consistently produces lower emission results then CHEER when calculated cumulatively on the same datasets. However, App (2014) is variable; for CanBikeCO and Bull eBike, it is in between the other two methods, but for MassCEC, it shows a considerably higher cumulative amount of CO_2 kg emitted. This incongruence could be due to the nature of App (2014), which uses a sensed mode taken from the phone sensors; if the user had specified and confirmed that the actual mode used during a trip was different from what was sensed, then the Dashboard (2020) and CHEER methods will differ as they use the confirmed mode for calculations.

The App (2014) method predates naïve and CHEER. To understand this method further, Figure 8 shows the cumulative amount of distance traveled in km for each mode, using a logarithmic scale. The utilization of modes is seen to vary across the locations; for example, MassCEC shows more subway and train utilization in Massachusetts, whereas there is more Air, High Speed Rail (HSR), and Bus usage in Colorado. Insights into mode utilization can help guide policy making to focus on the most utilized modes. However, App (2014) is limited in that it does not consider energy grids or localized factors such as occupancy. Our iterated and improved algorithms account for these factors. More accurate metrics can further inform policy decisions such that more sustainable modes are prioritized, in ways that distance alone does not capture.

VI. Conclusion

Most prior work from the literature related to computing energy or emissions footprints uses a variation of a naïve algorithm that combines static energy/emission intensity factors with trip lengths (Section II). We have used a version of this algorithm in our personalized carbon footprint app, and have iterated on it over 10 years (Section IV-A). These iterations have enabled us to identify fundamental limitations in the naïve approach that affect accuracy and granularity, and make it hard to generate actionable metrics (Section IV-B). We have identified a set of improvements (CHEER 2024) that address these limitations (Section IV-C), including updateable data sources for the computation (Section IV-D).

While these updates have increased the accuracy of our energy and emissions calculations, there are definitely areas that could be improved upon. We plan to address limitations in *algorithm sophistication* (Section IV-E1) in future iterations of this project (CHEER 2025 and beyond). As the success of this project is contingent on data, we encourage the research community to come together to address the limitations in data quality and freshness (Section IV-E2). With the data currently available, we have visualized the tangible changes across

these iterations thus far (Section V). In particular, we would encourage researchers outside the United States to advocate for the collection and publication of transportation statistics in their local regions. Data-driven techniques for proposing decarbonization strategies, and tracking their progress, can only work in the presence of reliable data.

These computations are currently integrated into our smartphone-based travel diary data collection platform [11], but we plan to make them available as an open-source library in the near future. We hope that the open-source, easy-to-use library, combined with high quality, locally relevant data sources can be used to inform data driven analyses and solutions for the global problem of climate change.

VII. Acknowledgements

This work was authored by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. This work was supported by the U.S. Department of Energy, Vehicle Technologies Office as part of the Energy Efficient Mobility Systems program's Core Tools activity area. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

VIII. Author contributions

A.W., J.G. and N.S. investigated data sources, J.G. and L.G. implemented the system, including data ingestion, processing and computation. A.W. and N.S. took the lead in writing the manuscript. J.F. expanded the literature review and methodology sections and performed data visualization. K.S. conceived the original idea, contributed data sources, supervised the project and assisted in writing. All authors provided critical feedback and helped shape the system design, implementation and manuscript.

References

[1] United Nations, "Climate change poses 'defining challenge' of our time, ban says," https://news.un.org/en/story/2008/10/276912, 2008, accessed: 2024-07-26.

[2] IPCC, *Summary for Policymakers*. Cambridge, UK: Cambridge University Press, 2022, p. In Press, accessed: 2023-07-26.

[3] Office of Transportation and Air Quality, "Fast Facts: U.S. Transportation Sector Greenhouse Gas Emissions," https://nepis.epa.gov/Exe/ZyPDF.cgi?Dockey=P101AKR0.pdf, U.S. Environmental Protection Agency, Fast Facts EPA-420-F-24-022, 2024, accessed 2024-07-29.

[4] Matteo Muratori, Tatjana Kunz, Aaron Hula, Michael Freedberg, "Fact Sheet: The U.S. National Blueprint for Transportation Decarbonization," https://www.energy.gov/sites/default/files/2023-01/EERE_TranspoDecarb_factsheet-508_0.pdf, U.S. DOE, DOT, EPA, HUD, Fact Sheet DOE/EE-2675, 2023, accessed 2024-07-29.

[5] S. Lukic and A. Emadi, "Effects of drivetrain hybridization on fuel economy and dynamic performance of parallel hybrid electric vehicles," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 2, pp. 385–389, 2004.

[6] Z. Guo, K. Zhou, C. Zhang, X. Lu, W. Chen, and S. Yang, "Residential electricity consumption behavior: Influencing factors, related theories and intervention strategies," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 399–412, 2018.

[7] A. Laskey and O. Kavazovic, "Opower," *XRDS: Crossroads, The ACM Magazine for Students*, vol. 17, no. 4, pp. 47–51, 2011.

[8] J. Lin, K. R. Marshall, S. Kabaca, M. Frades, and D. Ware, "Energy affordability in practice: Oracle Utilities Opower's business Intelligence to meet low and moderate income need at Eversource," *The Electricity Journal*, vol. 33, no. 2, p. 106687, 2020.

[9] M. McQueen, J. MacArthur, and C. Cherry, "The E-Bike Potential: Estimating regional e-bike impacts on greenhouse gas emissions," *Transportation Research Part D: Transport and Environment*, vol. 87, p. 102482, 2020.

[10] K. Shankari, M. A. Bouzaghrane, S. M. Maurer, P. Waddell, D. E. Culler, and R. H. Katz, "e-mission: An open-source, smartphone platform for collecting human travel data," *Transportation Research Record*, vol. 2672, no. 42, pp. 1–12, 2018.

[11] National Renewable Energy Laboratory, "NREL OpenPATH: Open Platform for Agile Trip Heuristics," https://www.nrel.gov/transportation/openpath.html, Accessed: 2024-07-31.

[12] EarthHero, https://www.earthhero.org/, Accessed: 2024-07-31.

[13] Commons, https://www.thecommons.earth/about-us, Accessed: 2024-07-31.

[14] S. Pal, "Joro is now commons: Why climate change needs a rebrand," *Commons Blog*, 2023.

[15] Cool Climate Network, https://coolclimate.org/index, Accessed: 2024-07-31.

[16] Klima, https://klima.com/faq/, Accessed: 2024-07-31.

[17] S. Prost, J. Schrammel, K. Röderer, and M. Tscheligi, "Contextualise! personalise! persuade!: a mobile HCI framework for behaviour change support systems," *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services*, pp. 510–515, 08 2013.

[18] W. Brazil, B. Caulfield, and E. Bothos, "Transport emissions information: Lessons from the peacox project," *Proceedings of the ITRN2015*, 08 2015.

[19] Ciclogreen, https://www.ciclogreen.com, Accessed: 2024-07-31.

[20] J. Froehlich, T. Dillahunt, P. Klasnja, J. Mankoff, S. Consolvo, B. Harrison, and J. A. Landay, "Ubigreen: investigating a mobile tool for tracking and supporting green transportation habits," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '09. New York, NY, USA: Association for Computing Machinery, 2009, p. 1043–1052.

[21] "Jariyasunant, Jerald, and Abou-Zeid, Maya and Carrel, Andre and Ekambaram, Venkatesan and Gaker, David and Sengupta, Raja and Walker, Joan L.", "Quantified traveler: Travel feedback meets the cloud to change behavior," *University of California Transportation Center*, 09 2013.

[22] M. Chester, "Life-cycle environmental inventory of passenger transportation in the united states," *Institute of Transportation Studies at UC Berkeley*, 2008.

[23] A. Jylhä, P. Nurmi, M. Sirén, S. Hemminki, and G. Jacucci, "Matkahupi: a persuasive mobile application for sustainable mobility," *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, 2013.

[24] E. Anagnostopoulou, J. Urbančič, E. Bothos, B. Magoutas, L. Bradesko, J. Schrammel, and G. Mentzas, "From mobility patterns to behavioural change: leveraging travel behaviour and personality profiles to nudge for sustainable transportation," *Journal of Intelligent Information Systems*, 10 2018.

[25] Climate Technology Solutions Pte. Ltd., https://web.archive.org/web/20220326083812/https://www.thecapture.club/, Accessed: 2024-07-31.

[26] P. Sawers, "Capture automatically tracks your CO2 emissions and offsets for a monthly fee," *VentureBeat*, 2020.

[27] S. Nocera, C. Ruiz-Alarcón-Quintero, and F. Cavallaro, "Assessing carbon emissions from road transport through traffic flow estimators," *Transportation Research Part C: Emerging Technologies*, vol. 95, pp. 125–148, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0968090X18306430

[28] J. Liu, J. Li, Y. Chen, S. Lian, J. Zeng, M. Geng, S. Zheng, Y. Dong, Y. He, P. Huang, Z. Zhao, X. Yan, Q. Hu, L. Wang, D. Yang, Z. Zhu, Y. Sun, W. Shang, D. Wang, L. Zhang, S. Hu, and X. M. Chen, "Multi-scale urban passenger transportation

co2 emission calculation platform for smart mobility management," *Applied Energy*, vol. 331, p. 120407, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261922016646

[29] K. Huang, P. Zhou, Z. Liu, T. Tang, H. Zhang, and W. Jiang, "The calculation and distribution of cav carbon emissions on urban transportation systems: A comparative analysis of renewable and non-renewable energy sources," *Renewable Energy*, vol. 230, p. 120884, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0960148124009522

[30] N. Kholod, M. Evans, E. Gusev, S. Yu, V. Malyshev, S. Tretyakova, and A. Barinov, "A methodology for calculating transport emissions in cities with limited traffic data: Case study of diesel particulates and black carbon emissions in murmansk," *Science of The Total Environment*, vol. 547, pp. 305–313, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0048969715313061

[31] L. Ntziachristos, D. Gkatzoflias, C. Kouridis, and Z. Samaras, "Copert: A european road transport emission inventory model," in *Information Technologies in Environmental Engineering*, I. N. Athanasiadis, A. E. Rizzoli, P. A. Mitkas, and J. M. Gómez, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 491–504.

[32] W. O. Max St-Jacques, Scott Bucking and I. Macdonald, "Spatio-temporal electrical grid emission factors effects on calculated ghg emissions of buildings in mixed-grid environments," *Science and Technology for the Built Environment*, vol. 30, no. 1, pp. 37–50, 2024.

[33] M.-A. M. Tamayao, J. J. Michalek, C. Hendrickson, and I. M. L. Azevedo, "Regional variability and uncertainty of electric vehicle life cycle co2 emissions across the united states," *Environmental Science & Technology*, vol. 49, no. 14, pp. 8844–8855, 2015, pMID: 26125323.

[34] M. Chester and A. Horvath, "Environmental life-cycle assessment of passenger transportation: A detailed methodology for energy, greenhouse gas, and criteria pollutant inventories of automobiles, buses, light rail, heavy rail and air," *UC Berkeley Center for Future Urban Transport*, 2007.

[35] T. Yuksel and J. J. Michalek, "Effects of regional temperature on electric vehicle efficiency, range, and emissions in the united states," *Environmental Science & Technology*, vol. 49, no. 6, pp. 3974–3980, 2015, pMID: 25671586.

[36] L. Zhu, J. Holden, J. Gonder, and E. Wood, "Green routing fuel saving opportunity assessment: A case study on california large- scale real-world travel data," *U.S. Department of Energy Office of Scientific and Technical Information*, 2017.

[37] National Renewable Energy Laboratory, "Google taps nrel expertise to incorporate energy optimization into google maps route guidance," *NREL News and Feature Stories*, 04 2021.

[38] G. Kosmacher and K. Shankari, "Evaluating the interplay between trajectory segmentation and mode inference error," *Transportation Research Record*, vol. 0, no. 0, p. 03611981231208154, 2023.

[39] J. Wolf, W. Bachman, M. S. Oliveira, J. Auld, A. K. Mohammadian, and P. Vovsha, "Applying gps data to understand travel behavior, volume ii: Guidelines," https://nap.nationalacademies.org/read/23436/chapter/7, National Cooperative Highway Research Program, Tech. Rep. 775, 2014.

[40] Kittelson & Associates, Parsons Brinkeroff, KFH Group, Texas A&M Transportation Institute, and Arup, *Transit capacity and quality of service manual*. Transportation Research Board, 2013, no. 165.

[41] R. Murphy, "Identifying micro-transit opportunity areas based on covid-19 ridership analysis and insights," https://www.metrans.org/news/uscs-reaghan-murphy-identifies-micro-transit-opportunity-areas-for-lbt-, University of Southern California, 09 2021.

[42] U.S. Department of Transportation, "Charging ahead: How government agencies can support electrification of private fleet vehicles," https://www.transportation.gov/sites/dot.gov/files/2023-12/OSTX-Momentum-Toolkit-EV-Fleets.pdf, U.S. Department of Transportation, 12 2023.

[43] U.S. Council on Environmental Quality, "Federal Sustainability Plan," https://www.sustainability.gov/federalsustainabilityplan/fleet.html, Council on Environmental Quality, 12 2021.

[44] U.S. Federal Transit Administration, "Low or no emission grant program - 5339(c)," https://www.transit.dot.gov/lowno, Federal Transit Administration, 07 2024.

[45] U.S. Internal Revenue Service, "Identifying Micro-Transit Opportunity Areas Based on COVID-19 Ridership Analysis and Insights," https://www.irs.gov/clean-vehicle-tax-credits, Internal Revenue Service, 05 2024.

[46] Colorado Energy Office, "Community access to electric bicycles rebate program," https://energyoffice.colorado.gov/ebike-rebates, Colorado Energy Office, 2 2024.

[47] U.S. Council on Environmental Quality, "Federal sustainability plan: Carbon pollution-free electricity," https://www.sustainability.gov/federalsustainabilityplan/carbon.html, Council on Environmental Quality, 12 2021.

[48] M. Laughton, "Variable renewables and the grid: an overview," *Renewable Electricity and the Grid*, pp. 1–29, 2012.

[49] J. N. Barkenbus, "Eco-driving: An overlooked climate change initiative," *Energy Policy*, vol. 38, no. 2, pp. 762–769, 2010.

[50] Institut für Energie- und Umweltforschung Heidelberg gGmbH, "TREMOD: Transport Emission Model," https://www.ifeu.de/en/methods-tools/models/tremod/, Institut für Energie- und Umweltforschung Heidelberg gGmbH, 2020.

[51] INFRAS, "The Handbook of Emission Factors for Road Transport," https://www.hbefa.net/, INFRAS, 01 2022.

[52] C. Akcicek, Z. Aemmer, K. Shankari, and A. Duvall, "Freewheeling: What Six Locations, 61,000 Trips, and 242,000 Miles in Colorado Reveal About How E-Bikes Improve Mobility Options," no. NREL/TP-5400-86388, 6 2023.

[53] K. Shankari, L. Boyce, E. Hintz, and A. Duvall, "The CanBikeCO Mini Pilot: Preliminary Results and Lessons Learned," https://www.osti.gov/biblio/1778194, Tech. Rep. NREL/TP-5400-79657, 4 2021.

[54] C. Akcicek, Z. Aemmer, A. Wheelis, K. Shankari, and A. Duvall, "The canbikeco full pilot: Long term results and analysis from an e-bike program in colorado, usa," *International Journal of Sustainable Transportation (under review)*, 2024.

[55] G. S. Nair, C. R. Bhat, I. Batur, R. M. Pendyala, and W. H. Lam, "A model of deadheading trips and pick-up locations for ride-hailing service vehicles," *Transportation Research Part A: Policy and Practice*, vol. 135, pp. 289–308, 2020.

[56] A. Millard-Ball, L. Liu, W. Hansen, D. Cooper, and J. Castiglione, "Where ridehail drivers go between trips," *Transportation*, vol. 50, pp. 1959–1981, Oct 2023.

[57] B. Sun, V. Garikapati, A. Wilson, and A. Duvall, "Estimating energy bounds for adoption of shared micromobility," *Transportation Research Part D: Transport and Environment*, vol. 100, 9 2021.

[58] M. Wang, "R&D GREET® Model," https://greet.anl.gov/, U.S. Department of Energy, Argonne National Lab, 2023.

[59] U.S. Department of Transportation, "NTD Annual Data View: Fuel and Energy by Agency," https://www.transit.dot.gov/ntd/data-product/2023-fuel-and-energy, 2023, accessed: 2024-11-13.

[60] S. C. Davis and R. G. Boundy, "Transportation energy data book: Edition 39." [Online]. Available: https://www.osti.gov/biblio/1767864

[61] U.S. Environmental Protection Agency, "eGRID2022 Unit, Generator, Plant, State, Balancing Authority Area, eGRID Subregion, NERC Region, U.S., and Grid Gross Loss (%) Data Files," https://www.epa.gov/system/files/documents/2024-01/egrid2022_data.xlsx, EPA, 01 2024.

[62] U.S. Department of Transportation, "The National Transit Database (NTD)," https://www.transit.dot.gov/ntd, 2022.

[63] Federal Transit Administration, *2023 NTD Policy Manual: National Transit Database Modes of Service*, 2023, accessed: 2024-11-15. [Online]. Available: https://ftis.org/iNTD-Urban/modes.pdf

[64] The International Association of Public Transport, "UITP Data," https://www.uitp.org/data/, The International Association of Public Transport, 2023.

[65] International Energy Agency, "Emissionsfactors," https://www.iea.org/data-and-statistics/data-product/emissions-factors-2, 2023, licence: Terms of Use for Non-CC Material.

[66] J. Tafoya, T. Nguyen, and M. Heese, https://raw.githubusercontent.com/USEPA/power-profiler/659afc6b6051cc8033cc78f536f10fc157ad1da0/app/data/zip.csv, 2024.

[67] Silicon Valley Clean Energy, https://svcleanenergy.org, Accessed: 2024-07-31.

[68] T. Team, "2022 Bull E-Bike Pilot Program Study." [Online]. Available: https://www.osti.gov/biblio/2323491

[69] K. Shankari, "MassCEC E-bikes for All Program - OpenPATH," https://mm-masscec-openpath.nrel.gov/join/, 2023, accessed: 2024-11-15.

[70] National Renewable Energy Laboratory, "Transportation Secure Data Center," 2021–2022, accessed 21 November 2024. [Online]. Available: https://www.nrel.gov/tsdc