NREL
Transforming ENERGY

Estimating Travel Energy Consumption Uncertainty Based on Inferred Travel Mode and Sensed Travel Length

Michael Allen and K. Shankari

Bridging Transportation Researchers

August 9-10, 2023

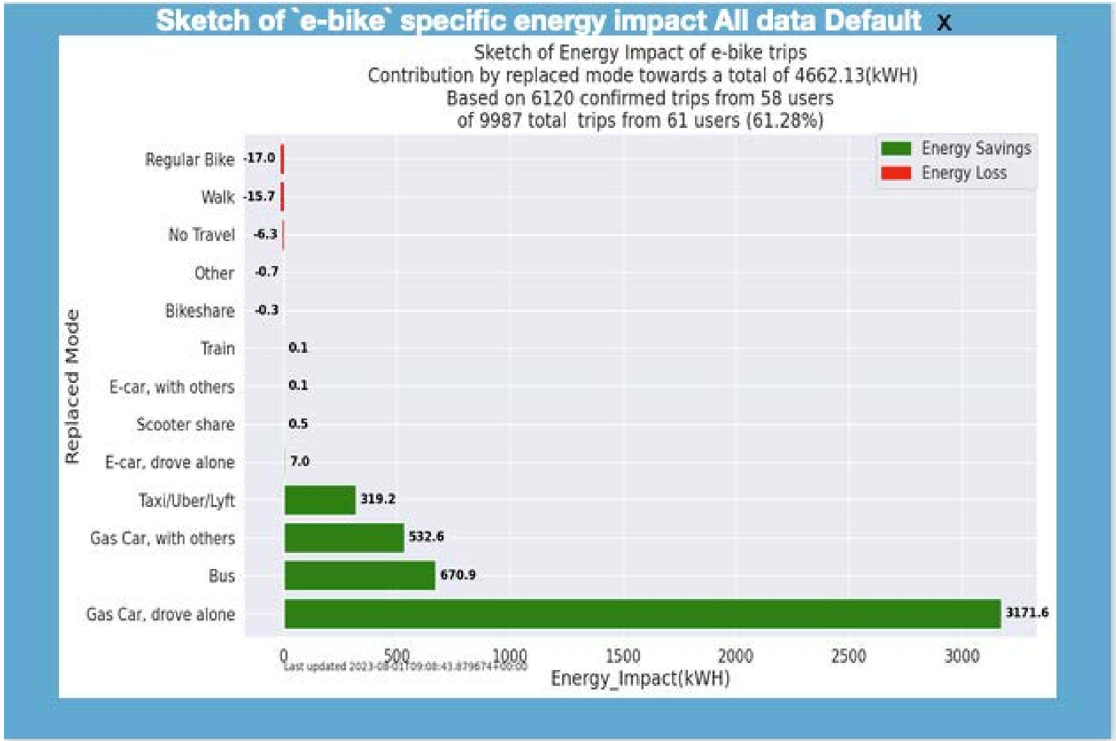
Photo from iStock-627281636

Vehicle Energy Consumption

Vehicle	Energy Impact Score	Greenhouse Gas Emissions (tailpipe)
Compare 	5.0 barrels/yr	10 1 10 Best 151 grams/mile
Compare 	5.4 barrels/yr	9 1 10 Best 162 grams/mile

Source: www.fueleconomy.gov

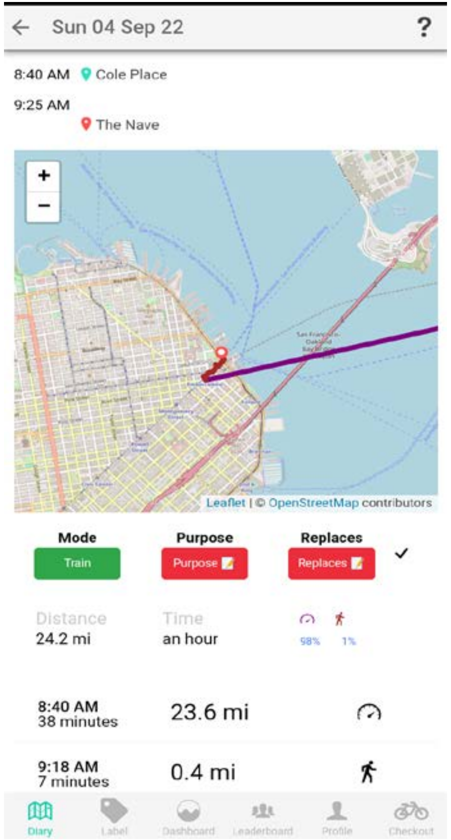
Traveler Energy Consumption



Energy savings when e-bike replaced the modes on the left.
Data collected with OpenPATH in Durham, NC

Smartphone Travel Diaries

- Smartphones can collect travel diaries
  - Carry phone during travel
- Participants get tired of labeling trips
  - Only 38% (94,000) of OpenPATH trips labeled
- Lots of unused information (146,000 trips)
- What if we try to use that information for decisions?
  - Infer travel mode with sensors or past labels



Screenshot from CanBikeCO app

Key Points

- Travel diaries from smartphones can tell you about travel behavior in a region
- Mode inference (machine learning) models can fill in some info that travel monitoring participants leave out
  - Makes data collection easier
- BUT: need to know how well those models perform
- We looked at uncertainty for one metric – energy consumption

Accounting for Uncertainty

- Energy = Energy Intensity * Travel Length = I*L
- For a travel diary or set of trips D:

$$E = \sum_{t \in D} I_t L_t$$

- Need uncertainty in both inputs
- Used ground truth smartphone travel data set (MobilityNet (Shankari et al, 2020))
- Energy Intensity/Mode uncertainty – confusion matrix
  - Can find for any mode classification algorithm
  - Accuracy not enough
- Length uncertainty – relative length error

Accounting for Uncertainty

- Mode: confusion matrix
  - Given predicted mode
  - Mean and variance of energy intensity (I)
- Length: relative length error
  - Given measured length
  - Mean and variance of actual length (L)
- Combine with variance propagation

$$\sigma_E^2 \approx \sigma_I^2 \mu_L^2 + \sigma_L^2 \mu_I^2$$

	Predicted Mode				
Actual Mode		Car	Ebike	Bike	Walk
	Car	7	2	1	0
	Ebike	2	5	3	1
	Bike	1	2	4	1
	Walk	0	1	2	8

Example confusion matrix.
Entries are number of occurrences.
More modes were present in actual

- Convert confusion matrix columns to probability distributions
  - P(actual mode | predicted mode)
- One approach: Divide each entry by the sum of its column
- After that, assign each an energy intensity
  - Based on the Transportation Energy Data Book from Oak Ridge National Laboratory

		Predicted Mode			
		Car	Ebike	Bike	Walk
Actual Mode	Car	0.7	0.2	0.1	0
	Ebike	0.2	0.5	0.3	0.1
	Bike	0.1	0.2	0.4	0.1
	Walk	0	0.1	0.2	0.8

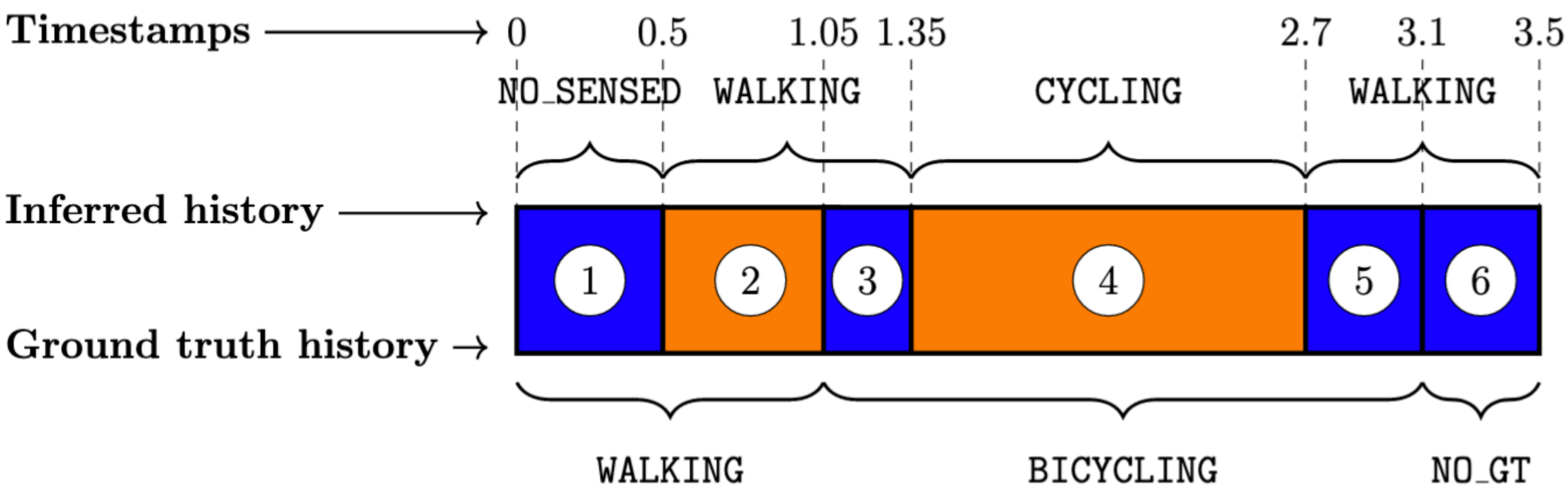
Now a table of mode probabilities in columns

	predict car	predict ebike	predict bike	predict walk
1.512 (car)	0.7	0.2	0.1	0
0.022 (ebike)	0.2	0.5	0.3	0.1
0 (bike and walk)	0.1	0.3	0.6	0.9

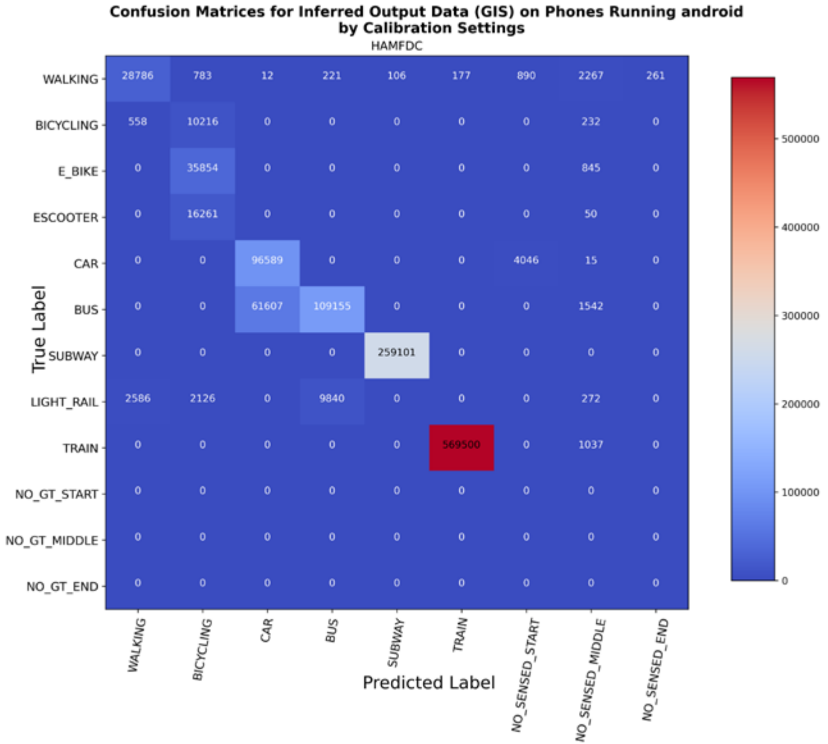
Table of energy intensity (kWH/PMT) probabilities

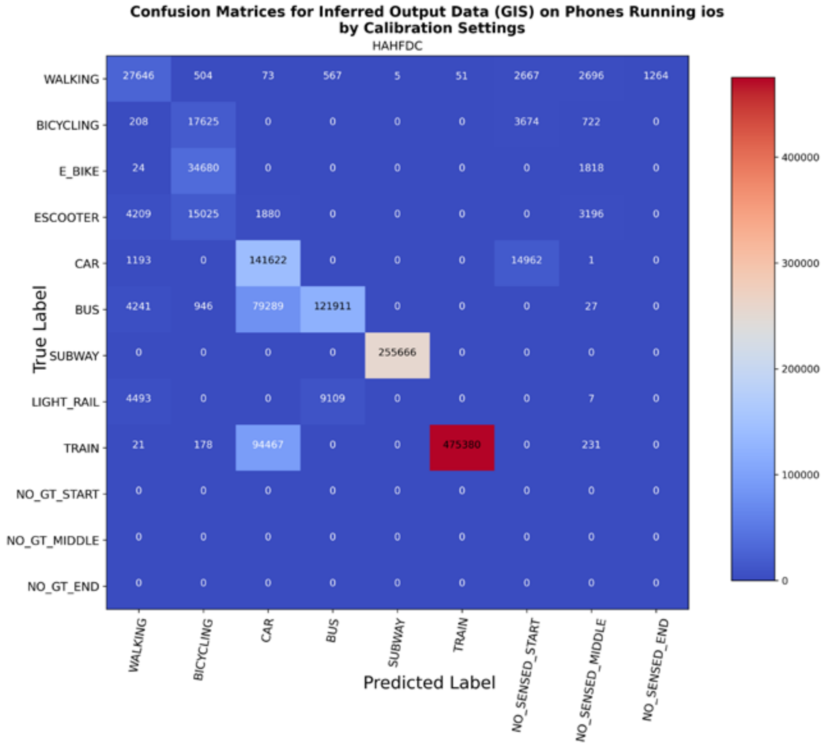
PMT = passenger miles traveled

Align Segments

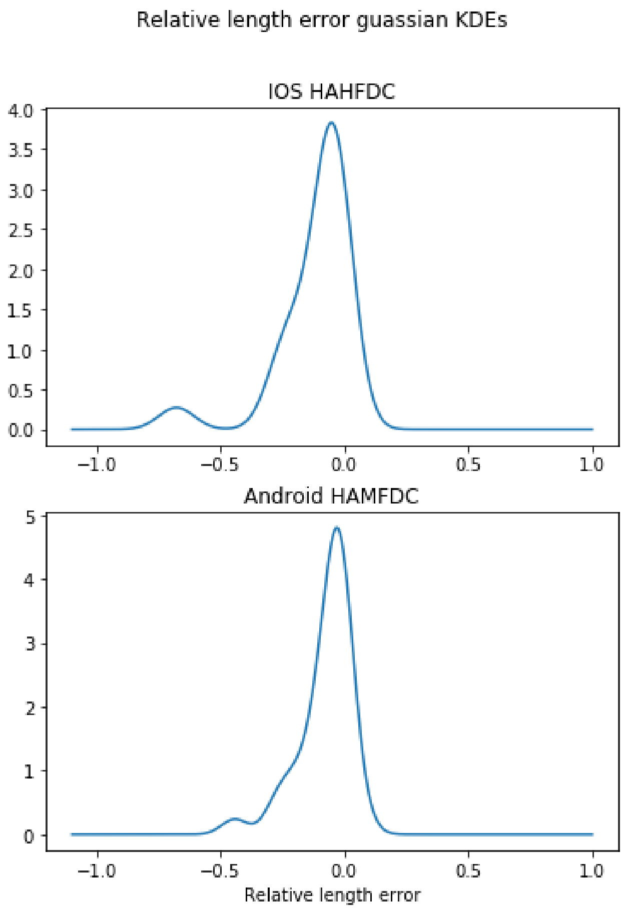


G. Kosmacher and K. Shankari, 2022





- Confusion matrix for phones running Android or iOS for a sensor-based mode inference model. Entries are in meters (found with methods based on G. Kosmacher and K. Shankari).
- HAHFDC and HAMFDC: phone configurations for recording data.
    - HAHFDC: High Accuracy, High Frequency, Duty Cycling
    - HAMFDC: High Accuracy, Medium Frequency, Duty Cycling



$$L_{actual} = L_{measured} * \frac{1}{1+R}$$

- Sampled from relative length error (R) distributions to find mean and variance of actual length given a 1 unit measured length

	$\bar{x}_R$	s_R^2	μ_L MCS	σ_L^2 MCS
android HAMFDC	-0.0776	0.1051	1.1060	0.03214
ios HAHFDC	-0.1397	0.2018	1.2224	0.3305

Acronyms:

- KDE: Kernel Density Estimate
  - Finds a distribution to represent the variation in data
- MCS: signifies that an estimate came from Monte Carlo Simulation

Variance Propagation

- At first, we summed the variances across trips:
  - $\mathrm{E} =$ variance of energy consumption for all trips
  - $var(E) = \sum_{t \in D} \sigma_t^2$
- However, most program (set of participants) estimates were above 7 standard deviations (sd) from the truth
  - With this, we would be placing more certainty on our estimates than we should

Variance Propagation

A better approach:

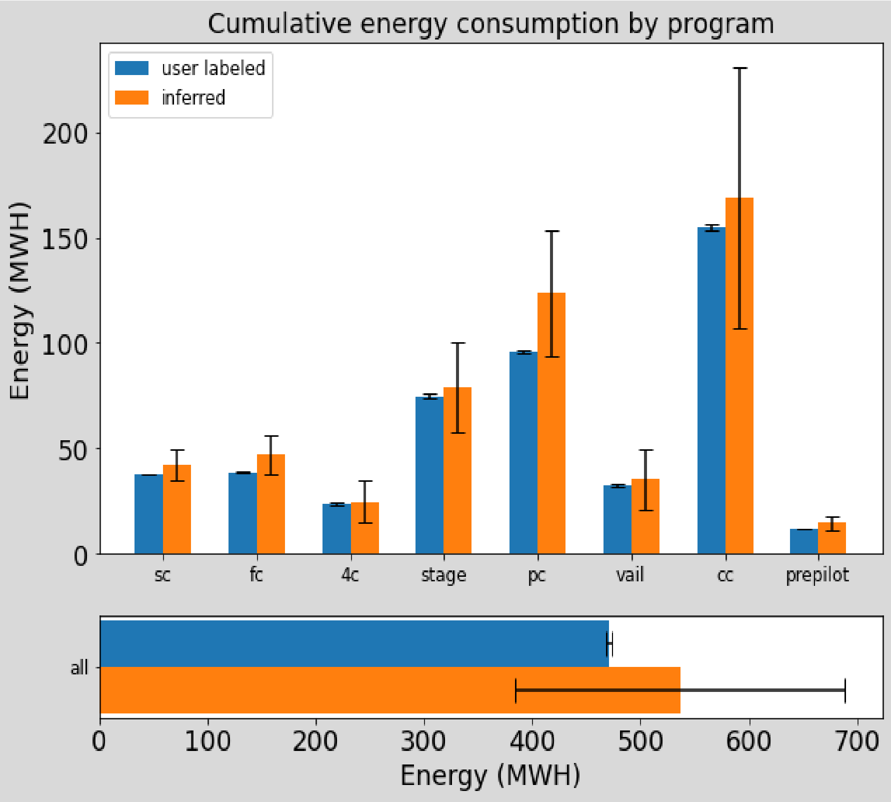
- Group trips together by predicted mode
- Number the modes (m) as {1, 2, …, k}
- I_m, L_m: Intensity for mode m, total travel length where mode m was predicted
- $var(E) = var(\sum_{m=1}^{k} I_m L_m) = \sum_{m=1}^{k} var(I_m L_m)$
  - $var(I_m L_m) \approx \sigma_{I_m}^2 \mu_{L_m}^2 + \sigma_{L_m}^2 \mu_{I_m}^2$
- Considers that predictions are not isolated events
  - Some mistakes might be repeated often

Assumptions

- Mode error and length error independent
- Ignore energy model specification error
  - We are applying an average intensity to the entire trip
- No error in user’s mode labels
- Error in MobilityNet applies to other geographies
- Assume an underlying mode prevalence (chances of each mode)
- Equal chances of drove alone and shared ride (sensing model does not distinguish these)

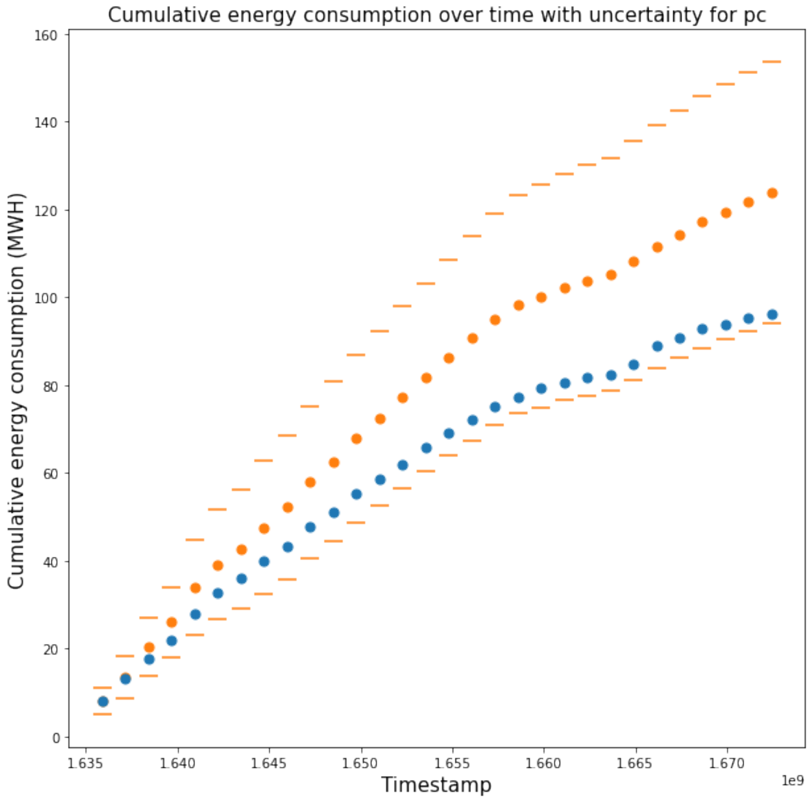
Results

- Benchmark on CanBikeCO
  - Low-income participants given electric bicycles
  - 1.5 years of data
  - Overall energy consumption error: 8% (bottom bars)
    - 13% if basing estimates on sections
  - Each program within 1 estimated standard deviation
  - At the user level: 82% within 2 estimated standard deviations

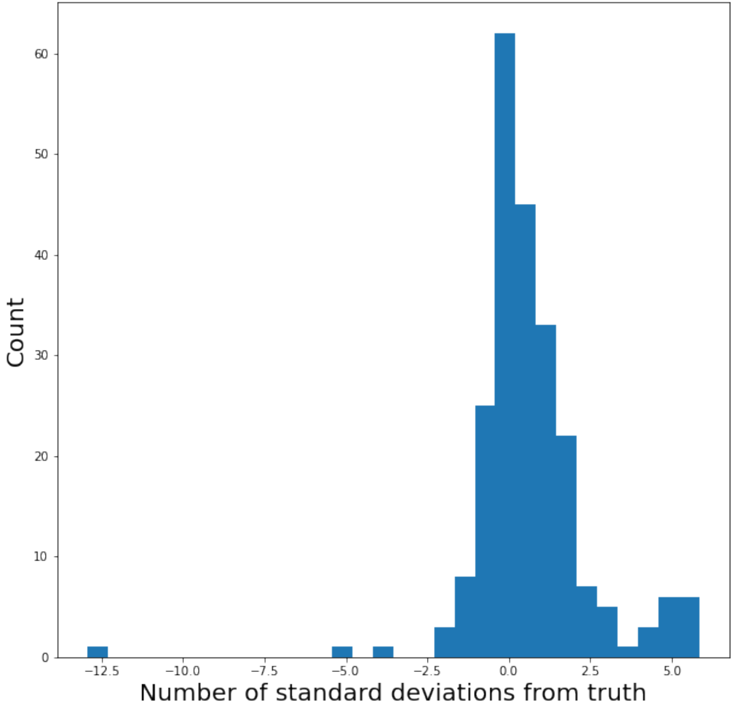


Results

- Cumulative energy consumption over time for Pueblo County (PC)
- PC was the program with the largest error
- Blue: energy based on user labels
- Orange dots: energy based on sensed labels
- Orange dashes: 1 standard deviation from the estimated energy consumption
- Data points are in 30-day increments



Results

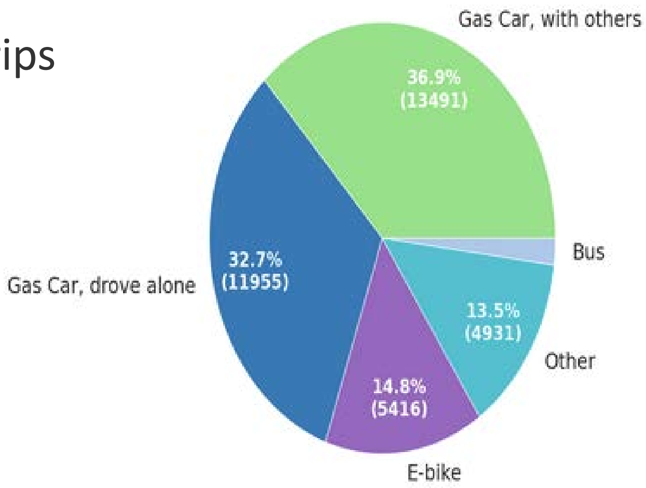


Histogram of user level number of standard deviations from the truth for cumulative energy consumption

Relevance and Future Work

- Inferred labels → Less labeling burden
    - Makes long term studies more viable
- Bias less toward people with free time to label their trips
- Local travel understanding
- Behavior models - better representation in travel demand forecasting

- Estimate mode share, energy/carbon impact, trip purpose share
- Apply methods to other mode inference models
    - e.g., learn from users' past labels
- Make better models (e.g., include route map matching to better distinguish bus and car)
    - → narrow the error bars



Trip miles by mode collected with OpenPATH in Durham, NC

Data Availability

TSDC

Spatial Data ›

An application is required to access latitude and longitude spatial data from transportation studies and surveys.

Transportation Secure Data Center: www.nrel.gov/tsdc

References

1. K. Shankari, J. Fuerst, M. F. Argerich, E. Avramidis, and J. Zhang. *MobilityNet: Towards a Public Dataset for Multi-modal Mobility Research*. ICLR 2020 Workshop on Tackling Climate Change with Machine Learning, page 6, 2020.

2. G. Kosmacher and K. Shankari. *Evaluating the Interplay between Trajectory Segmentation and Mode Inference Errors*. Transportation Research Record (to appear)

Q&A

www.nrel.gov

NREL/PR-5400-86955

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 in part by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTS) under the Science Undergraduate Laboratory Internships Program (SULI). This material is also based upon work supported by the DOE Office of Energy Efficiency and Renewable 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.

NREL Transforming ENERGY