Predicting Vehicular Emissions by Converging Direct Measurements and Mobility Science

Ayşe Tuğba Öztürk[1], Helen Fitzmaurice[2], Olga Kavvada[3], Philippe Calvez[3], Ronald C. Cohen[4], Marta C. González[1,5,6*]

[1]Department of Civil and Environmental Engineering, University of California, Berkeley, CA 94720, USA
[2]Graduate School of Education, University of California, Berkeley, CA 94720, USA
[3]ENGIE Lab CRIGEN, Computer Science and Artificial Intelligence Lab(CSAI) Paris, France
[4]Department of Chemistry, University of California, Berkeley, CA 94720, USA
[5]Energy Technologies Area, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[6]Department of City and Regional Planning, University of California, Berkeley, CA 94720, USA

*To whom correspondence should be addressed, e-mail: martag@berkeley.edu.

Vehicle emissions pose a significant challenge for cities worldwide, yet a comprehensive analysis of the relationship between mobility metrics and total vehicle emissions at a high resolution remains elusive. In this work, we introduce the Mobile Data Emission System (MODES), a pioneering framework that integrates various sources of individual mobility data on an unprecedented scale. Our model is validated with direct measurements from a network of high-density sensors analyzed before and during the COVID-19 pandemic shelter

in place orders. MODES is used as a laboratory for scaling analysis. Informed by individual trips, we estimate the traffic CO_2 emissions at a metropolitan scale with a combination of 3 accessible metrics: vehicle kilometers traveled (VKT), congestion levels, and vehicle efficiency. Given their ranges of variation, VKT has the greatest role in amplifying vehicular emissions up to 500%, followed by vehicle efficiency that would range 20% to 300% of the average passenger combustion vehicle. In comparison, congestion amplifies vehicle emissions of individual travels up to 50%. We confirm that cities in the Bay Area with high population density are consistently characterized by low per-person VKT. Nevertheless, high population density comes at the expense of increased congestion. Since VKT is the governing factor, overall densifying of the urban landscape reduces transportation emissions despite its impacts in congestion.

Transportation is a sector that is difficult to decarbonize while it represents 29% of the greenhouse gas (GHG) emissions in the United States (*1*). Policy efforts to date have focused on improving vehicle efficiency, alternative fuel technologies, and expanding the range of electric vehicle batteries. Despite the progress in supply-side solutions, per-capita vehicle kilometers travelled (VKT) and transportation CO_2 emissions have been in the rise in many regions worldwide (*2, 3*). Curbing emissions in the transportation sector has proven to be highly challenging compared to the other sectors (*4–6*). Confinements during the COVID-19 pandemic bent the emission curves with estimated reductions of 12.9% in United States and 7% worldwide during 2020, mainly due to social and travel behavior changes (*7*). Highway emissions are estimated to have decreased by 48% in the Bay Area where strict confinement policies were adopted (*8*). Even though the emissions rebounded by 2021, there are still lessons to be learned from the changes in this period.

Current transportation emission inventories of cities rely on fuel sales data aggregated over large regions and time frames. These methods have been shown to underestimate on-road emissions by 28.1% on average in US cities (*9*). Sensor-aided monitoring approaches are gaining importance. Among these methods, highway sensors (*10, 11*) are used to get vehicle count estimates on roads. Networks of high-density sensors together with Bayesian inversion methods are reported to attribute emissions to the source activities (i.e. transportation, home heating, industry) (*12, 13*). While monitoring methods are crucial in constructing and validating data-driven emission inventories, there are two shortcomings that call for individual mobility estimates. First, sensors are expensive to deploy and maintain over large regions to directly inform city inventories. Second, sensor based estimates are not capable of attributing emissions to the origin and destinations associated with the travel activity. The later is particularly important for transportation planning authorities for measuring carbon emissions by source activity at urban scale to achieve decarbonization goals (*14*).

With the increasing data availability, quantification methods combining various data sources on travel demand and road network conditions are gaining importance. These methods can be further divided into three groups based on the mobility data being used: a) local traffic volumes (***?***, *15, 16*), b) aggregated origin-destination (OD) flows (*17*), and c) individual mobility data (*18–20*). To understand the difference between these three approaches, we need to differentiate the outputs of these models as *local emission estimates* and *emission production estimates*. Local emission estimates quantify emission within a region, route or road segment. Methods using local traffic volumes estimate the emissions in the place where they occur. While these models provide accurate estimates locally, they lack information regarding the traveler, source and destination of the trip. This information holds significant importance for policy-making and demographic analysis. *Emission production estimates* quantify CO_2 emitted across the travel network by the origin of the individual travelers and link them to their locations of residence. This type of quantification is possible with individual mobility data or OD flows.

As vehicle emissions depend on the interplay between the urban structure, travel demand, vehicle efficiency, and road conditions, the task at hand is to leverage mobility science to inform climate change mitigation strategies. Mobility science approaches utilize large scale data sources that are passively collected from information and communication technologies. These data sources are used to compare and understand differences among cities (*21, 22*). To that end, they model urban or mobility phenomena as aggregated properties that are functions of a system's variables. For example, studies on multiple cites have uncovered scaling laws relating population to distribution of facilities and socioeconomic activities at macroscopic scale (*21, 23–27*). These studies have reported that the more populated the cities are, the more efficient they are in their per capita energy consumption (*23, 24*), or that professional diversity and the productivity of cities can be modeled as social networks embedded in space (*28*). In the mobility front, universal laws govern the collapse of traffic networks (*29*) or traffic management strategies via smart phone applications (*22*).

This work utilizes mobility science methods in the analysis of vehicle emission estimates. With this in mind, we present MODES, a portable framework to estimate vehicle emissions using various individual data sources. To that end, we integrate, at unprecedented scale, call detail records (CDR's) for the Bay Area (*30*), aggregated Location Based Service data from SafeGraph (*31*) and Uber Movement Speeds data (*32*).

To validate our model, we utilize data from both before and after the COVID-19 shelter-in-place (SIP) orders in the San Francisco Bay Area. SIP orders significantly altered individuals' travel behavior, making it an ideal opportunity to assess our model's performance under different conditions. Therefore we performed estimates for these two periods separately. We estimate the travel behavior and the road congestion for six weeks before and after the SIP orders, implemented on March 16th 2020. We combine the traffic estimates with the StreetSmart fuel consumption model (*33, 34*) and convert these values to tail-pipe CO_2 emissions. We compare our estimates with direct CO_2 sensor measurements and a hybrid model of highway vehicle counts and emission factors.

To our knowledge, this is the first study that utilizes individual mobility extracted from mobile phones in data-driven emissions estimates. This allows us to link CO_2 emissions to the home locations of travelers. We investigate the interplay of population density and road network structure on individual travel demand and their associated emissions. We find that cities with high population density are characterized by low per-person VKT, low per-person CO_2 emissions, and high congestion levels.

Analyzing individual trips at varying percentages of congestion, we uncover a bottom-up law that incorporates VKT, congestion, and vehicle efficiency to estimate emissions from millions of individual trips. The parameters we utilize are not specific to a particular study region, allowing our results to be applicable to trips in any city. For a given fleet of vehicles, VKT emerges as the primary contributor to amplify emissions. Current scaling laws depend on top down estimates of the form $tCO_2 \sim N^{\beta}$, where N is the population size and its emission production ins tons (tCO_2) (*35*). Our approach is an advancement over the existing laws by incorporating parameters directly effecting the vehicular emissions while maintaining the generalizability of the findings.

A critical step is the validation of our estimates. Existing data-driven emission estimation frameworks are validated with other data-driven emission inventories. These inventories often have similar underlying assumptions. In contrast, to validate and test the robustness of MODES, we used completely independent measurements from the Berkeley Air Quality and CO_2 Network ($BEACO_2N$). We performed a comprehensive validation and calculated statistics in regions with different characteristics (ie. highway length, residential road length) and resolutions (ie. from 1 to 25 km^2 spatial resolution, averaging from 1 to 5 measurement days). We validated the results from before and COVID-19 separately ensuring accuracy under different demand levels. We found that at a spatial resolution of 9 km^2 and 5 days of measurement, the median difference between MODES and $BEACO_2N$ is 32 %.

The main contributions of our work include:

(1) We establish a scaling law that quantifies the relationship between emissions, VKT, congestion, and vehicle efficiency, utilizing easily available metrics in numerous cities worldwide. This approach greatly simplifies the estimation of emissions.

(2) Scaling analysis, highlights the VKT as the main parameter influencing emissions. Building upon this universal finding, we quantify the interplay between emissions VKT, congestion, and vehicle efficiency. This advancement enhances our understanding of the complex dynamics influencing emissions in transportation systems.

(3) MODES represents a significant advancement over local emission estimates by quantifying emissions at the individual traveler level. This capability allows MODES to directly inform policy decisions regarding vehicle electrification, equity in air pollution exposure, and land use planning.

(4) Unlike other models that lack validation or are validated across large metropolitan regions, MODES stands out by being validated through direct measurements of CO2 emissions at various spatial resolutions. This validation process ensures the accuracy and precision necessary to meet the requirements of individual policy-making use cases.

(5) MODES is the first study utilizing state-of-art travel models informed by passively collected mobile phone activity data in emissions estimates.

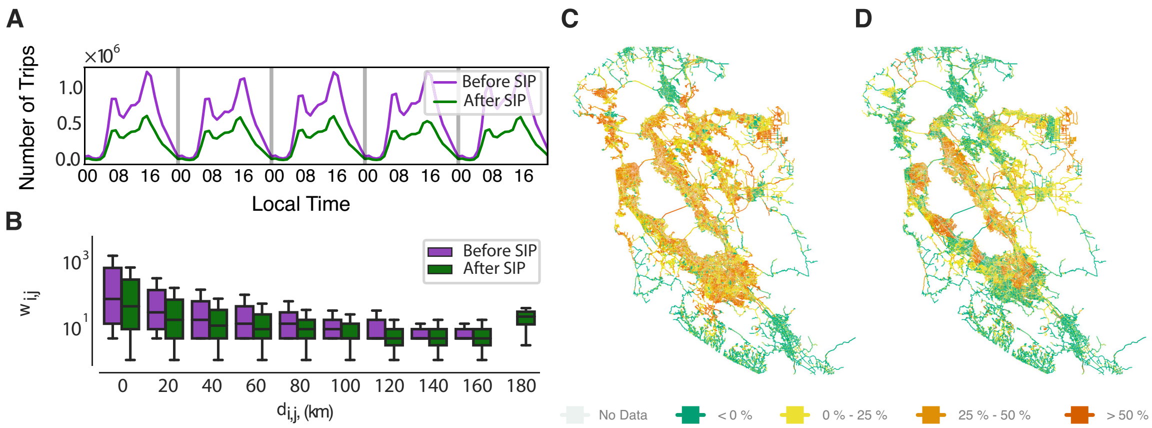

Figure 1: **Travel demand and congestion levels before and during shelter-in-place orders**. Congestion level is defined as percentage change in travel time compared to the free-flow travel time. Congestion levels are calculated using hourly average speeds for each road segment in the network. (A) Weekday travel demand extracted from CDR and LBS data, aggregated hourly. (B) Travel demand between i (origin) and j (destination) grouped by Haversine distance. Congestion level in the road network on Monday 5PM before shelter-in-place orders (C) and during shelter-in-place orders (D).

Results

Congestion estimates and travel behavior

We characterize the typical weekday travel behavior before and during the shelter in place (SIP) orders in the Bay Area road network. All Bay Area counties imposed SIP orders starting on March 16th, 2020. We defined the before SIP period as six weeks until this date and during SIP period as six weeks following this date. The primary purpose of utilizing the post-COVID-19 period is to conduct a comprehensive validation process.

The travel demand used in this work is based on the urban mobility model TimeGeo, a simulation of individual mobility using CDRs (*36*). The individual mobility patterns of the simulated users align well with the 2010–2012 California Household Travel Survey (CHTS) and the 2009 National Household Travel Survey (NHTS) (*30*). Travel surveys, while valid, are costly methods of data collection that only capture a limited portion of the population and may not encompass all individual trips undertaken. Hence TimeGeo presents a comprehensive and valid estimate of the individual travel demand. This model represents the typical weekday travel behavior for the before SIP period. To extend this model to each day of the week for before and during SIP periods, we used SafeGraph data. We multiply the hourly travel demand between each origin and destination pair with a scaling coefficient retrieved from SafeGraph (see Supplementary Notes S2). To use the resulting travel demand for vehicles, the standard is to estimate vehicle trips from the total trips by scaling total trips with the vehicle usage rate of the home census tract (see Supplementary Notes S3). The resulting vehicle demand is 9.6 million vehicle trips out of 13 million passenger trips within the Bay Area for a typical weekday. The hourly and weekly distribution of the weekday trips for before and during SIP is shown in Fig 1A. We find that due to the SIP orders the total number of trips decreased by 52% and 59% during the morning peak. In Fig 1B, we show the number of trips between each origin and destination pair grouped by the distance. The variations in number of trips are due to the differences in travel demand during different times of the day. Interestingly, while the number of trips decreased in total, there was an increase in the share of trips above $150km$. The total vehicle kilometers travelled (VKT) decreased by 44% from 234M km/day to 131M km/day. The comparison of these findings are consistent with the report of the Bureau of Transportation statistics COVID-19 travel behavior changes (*37*).

We perform road network analysis within a bounding box formed by the locations of mobile phone users. We modeled the San Francisco Bay Area road network as a weighted, directed graph where edges represent the road segments and the edge weights are the speeds, travel times and road lengths. Road geometries are retrieved from Open Street Map (OSM) and corresponding hourly speeds are provided by the Uber Movement Speeds API. Travel time on each edge is calculated using the road geometries and hourly speed values. We identified the free-flow speed as the 85th percentile of all speed values observed on a road segment during a week. Then the congestion level in a region is defined as the percentage change in the travel time in comparison with the free-flow travel time. Most of network-wide congestion occurs at 5 PM before the SIP orders and ranges between 30 and 37% depending on the day of the week. During the SIP orders, the 5PM's network-wide congestion dropped to 14 to 19%. The study region and the congestion level per road segment is illustrated in Figures 1 C and D.

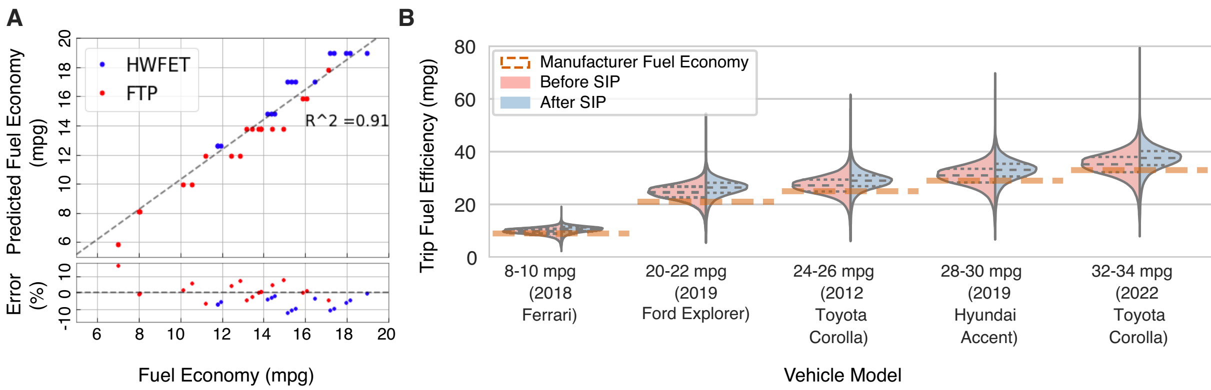

Figure 2: **Vehicle Fuel Consumption Model** (A) Validation of the StreetSmart fuel efficiency estimates with Autonomie vehicle fuel consumption simulations. The fuel economy estimations use EPA's FTP-75 (urban) and HWFET (highway) standard speed profiles. (B) Variations in on-road fuel economy by manufacturers reported fuel economy in the same analysis period. The differences between the actual fuel economy and the manufacturers' value can be as high as 70%, as shown in the distributions, illustrating the impact of the road network traffic. Due to a decrease in congestion, before SIP's order trips have a 6-7% lower median fuel efficiency than the during SIP order

Vehicle efficiency variations in traffic

In this work, we employed the StreetSmart (*33*) model to estimate each trip's fuel consumption. This model requires the speed profiles on the road segments and accounts for the vehicle efficiency variations under different speed and acceleration levels. The model requires four variables as shown in Equation 1.

$$FC_{gal} = k_1 T_{idle} + k_2 T_{move} + k_3 \int_x |a| dx + k_4 L \tag{1}$$

The first term accounts for the idling energy consumption; the second term accounts for the energy depending on the time of the movement; the third term accounts for the acceleration and deceleration over a given distance, and the final term accounts for the energy use for the distance traveled. The k's in the model are vehicle efficiency-specific multipliers. The model has over 96% accuracy in tests performed with vehicles GPS coordinates and output fuel consumption values read from on board diagnostics (OBD-II) devices. To use this model for any vehicle in the Bay Area fleet, we calibrated the vehicle efficiency constants and further validated the calibrated model. For calibration of the constants *k*s, we used fuel efficiency values provided by the Environmental Protection Agency's (EPA) 2021 report (*38*) and FTP-75 speed profiles used in EPAs urban fuel economy testing procedure (*39*). (see Supplementary Notes S4 for a list of calibrated coefficients). We validate our calibration by comparing them with the results of the software Autonomie, a fuel consumption simulator developed by the Argonne National Laboratory (*40*). We tested the model for five vehicles in the software under EPA's FTP-75 and HTP-75 drive cycles (see Supplementary Note S4). We recorded the mean absolute error of all tests, obtaining 5% on average, as shown in Fig. 2A.

To use this model to estimate city-wide fuel consumption, we assigned each vehicle trip a route on the network, based on the edges with the shortest travel time. Each trip's travel time, speed, and distance is calculated from the edge weights. The idling time is imputed only if a stop sign, traffic light, or crossing is present at the nodes traversed. The acceleration variable in the model is approximated since we do not have high-resolution speeds along the edges. We use constant speeds along the roads as extracted from the Uber Movement Speeds. We then included the acceleration/deceleration if a stop or a speed change between edges is present (for details see Supplementary Notes S5). For example, on an average Monday, we have 9.6 million daily trips before SIP orders and 4.2 million distinct origin and destination pairs. We assigned 5.3 million distinct routes and recorded associated speed profiles.

The fuel consumption of all the routes are calculated for different manufacturer-reported fuel economies with the calibrated StreetSmart model. Due to traffic conditions, the average of our estimated fuel efficiency of the vehicles varies from the manufacturer-reported fuel economy by 17%, and the differences in the actual fuel economy of the same car can be as high as 70% as it is shown in the distributions of Figure 2B. These variations illustrate the impact of the road network traffic on the actual vehicle efficiency. We observe that trips before the SIP orders have a 6-7% lower median fuel efficiency than the trips during the SIP orders. The difference arises from the less congested states of the roads.

Emissions comparisons with sensor-based estimates

In urban areas, disagreement between on-road emission inventories can be as large as 50-250% (*41*). The uncertainty arises from the model assumptions and differences in underlying data sources such as vehicle efficiency, emission factors, magnitude and spatial distribution of VKT, and travel speeds. In MODES, emissions are calculated only for the personal vehicles for an average Bay Area vehicle with 25 mpg efficiency (*42*). Trucks fuel consumption is not included due to a lack of data on heavy-duty vehicles' temporal and spatial distribution. In order to get emissions, fuel consumption estimates for the personal vehicles are then converted to CO_2 emissions assuming 8,887 [g CO_2] is emitted per gallon of fuel burned (*43*) (See Supplementary Notes S8).

We compared our resulting emission estimates with the direct measurement of the Berkeley Air Quality and CO_2 Network ($BEACO_2N$) and also with the Emissions Factor model (EMFAC2017) of the California Air Resources Board (CARB) (*44*) applied to vehicle flows acquired from CalTrans Performance Measurement System (PeMS) (*45*). After COVID-19 SIP estimates are mainly used for validation purposes, making the results robust to different travel demand levels within the network.

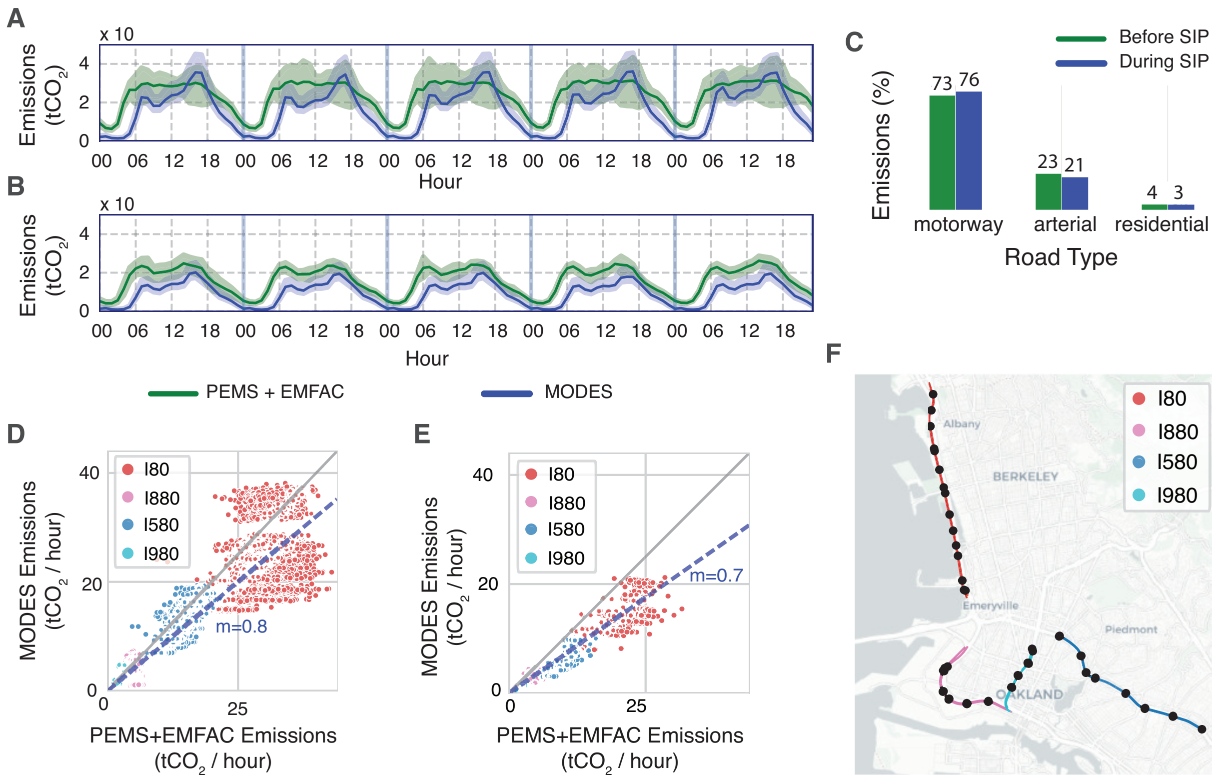

Figure 3: **On-road vehicle emissions validation**. Comparison of PEMS + EMFAC estimates and MODES on I-80 highway for (A) an average weekday before the SIP order and (B) for an average weekday during the SIP period. (C) Percentage of total emission estimates for each road class before and after SIP estimated by MODES. Highways have the largest share of emissions both before and after SIP. (D) Comparison of hourly PEMS + EMFAC and MODES emission estimates on 4 highway segments before SIP. The best-fit line $y = 0.8x$ (blue) and $y = x$ (gray) are shown. (E) Comparison of hourly PEMS + EMFAC and MODES emission estimates on 4 highway segments after SIP. The best-fit line $y = 0.7x$ (blue) and $y = x$ (gray) are shown. (F) Highway segments used in part D.

The EMFAC2017 model provides emissions factors for each vehicle class and speed level. However, getting the aggregate emissions for road segments requires further knowledge of vehicle counts and speeds. We followed the method presented in Fitzmaurice et al. (*42*) and combined the emissions factors with the vehicle count, truck percentage, and speed data obtained from PeMS. The PeMS data is only available for highways. Therefore, we first validated our results only on highways which account for 73% and 76% of all vehicle emissions before and during the SIP. Selected highway segments; I580 I80, I880 and I980 are presented in Figure 3. These segments were selected due to the differences in length and average truck percentages. We observe that our estimates are within 35% of the PeMS-EMFAC2017 model for daytime emissions between 7 am and 9 pm. Before and during SIP emission comparisons for a typical weekday for a 5 km stretch of the I80 are presented in Figures 3 A-B. Emissions comparisons for the rest of the highways are provided in the Supplementary Figure S9.

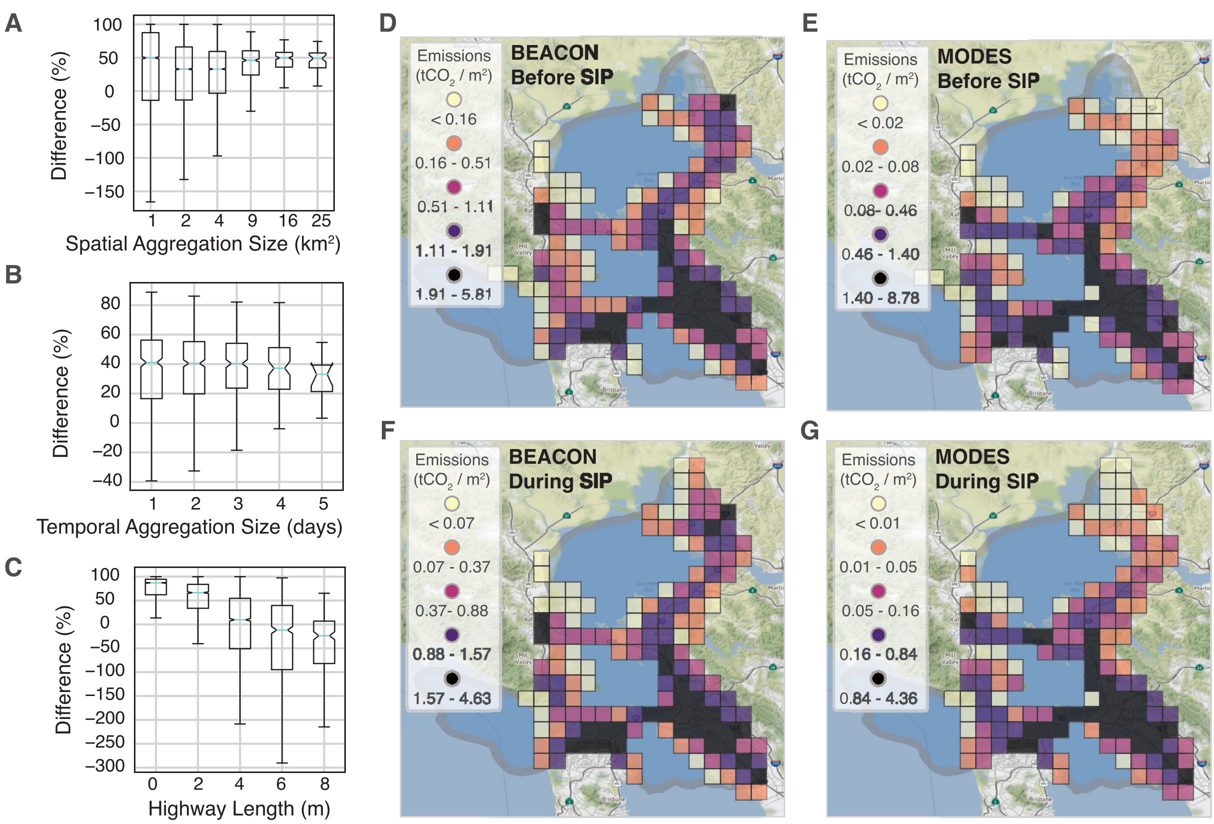

Figure 4: **Comparison of MODES CO_2 emissions estimates with $BeaCO_2N$ measurements.** (A) Comparison under different spatial aggregation units. (B) Comparison under different temporal aggregation units. The difference between $BeaCO_2N$ minus MODES is calculated for 9 km^2 grids. (C) Comparison in cells with different highway lengths. As the highway span increases, our estimates overpredict. (D-G) Local emission estimates within the $BEACO_2N$ domain (Gray line indicates the region that contains the largest 40% of the total network influence). Emissions are spatially aggregated into 9 km^2 cells and temporally aggregated within the 6 weeks before and after the SIP analysis period. Quantile breaks are adopted to demonstrate the spatial distribution of emission hot spots.

To evaluate the performance of MODES in non-highway emissions, we compared the results with the $BEACO_2N$ estimates. Comparison of data-driven models with the direct emissions estimates are expected to have differences as there are uncertainties involved in both processes. The CO_2 concentration measurements are recorded by a $BEACO_2N$'s dense network of sensors and attributed to a source using a Bayesian inversion method (*8*). We performed analysis on transportation emissions in the 40% CO_2 influence region (*42*), shown in Figure 4 (D-F). Percentage difference is calculated as $\frac{(CO_2^{BeaCO_2N} - CO_2^{MODES})}{CO_2^{BeaCO_2N}}$.

Hourly airflow models at high spatial resolution can vary significantly due to uncertainty in wind speed and direction (*46*). In order to find a good validation resolution, we tested the difference between $BEACO_2N$ and MODES estimates at varying aggregation levels and highway lengths. Results are shown in Figure 4. We first aggregated the transportation emissions zones from 1 to 25 km^2 areas to account for the possible errors in the airflow direction model. Spatial aggregation significantly decreased the range and median of the difference between the two models up to 9 km^2. We further aggregated the results temporally for 9 km^2 cells. We take the average of the weekdays available for each measurement hour based on the sensor measurements. Temporal aggregation further reduced the range and median of the difference. Finally, we found that MODES underpredict regions with fewer highways.

The spatial variation of the emissions for the average 3 PM travel behavior before and after the SIP is illustrated in Figure 4 D-G. The results shown are for 9 km^2 cells and aggregated temporally as described above. The cells are colored based on the quantile breaks of each model and measurement period. When we compare Figure 4 D and E, MODES and $BEACO_2N$ have a similar spatial distribution of the emissions. MODES generally underpredict. The underprediction of MODES is expected since we do not have heavy-duty vehicles (HDV). Our high prediction in highways could be due to the high flow values assigned to the highways in the vehicle assignment method. When we compare Figs. 4 D and F, and Figs. E and G, we observe that the model captured the decrease in emissions during the SIP order.

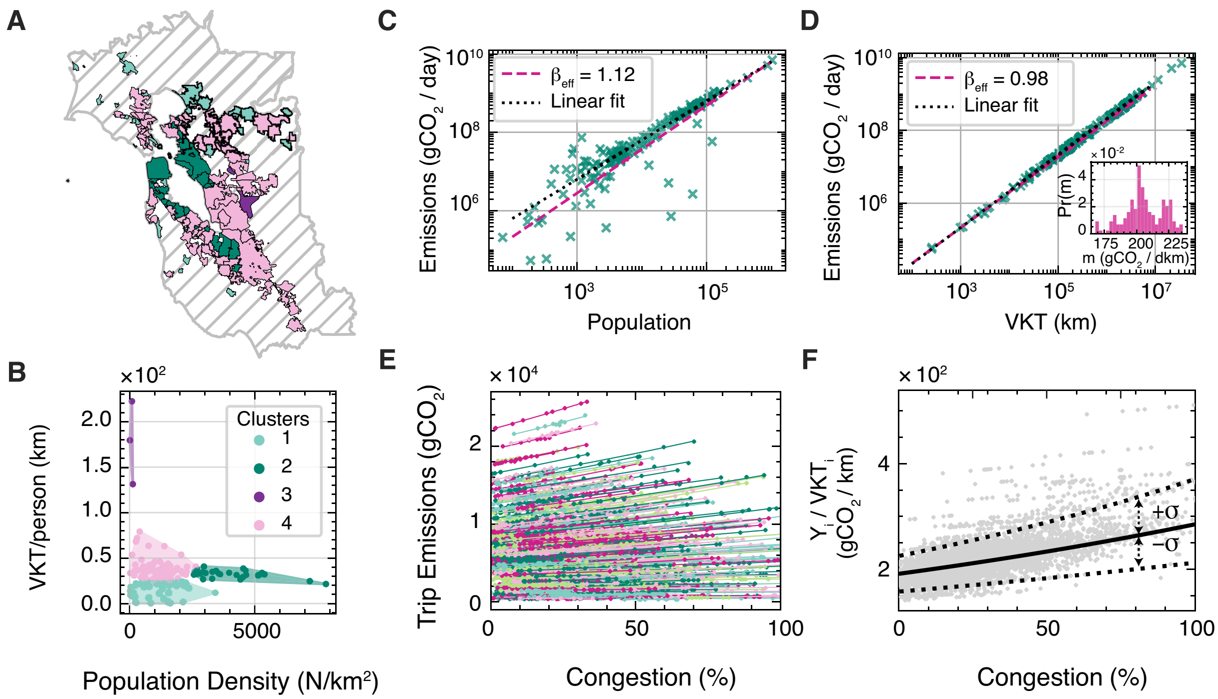

Figure 5: **Scaling of on-road vehicle emissions** (A-B) Cities in the analysis region are clustered into four groups (different colors) based on population density and per-person VKT within the travel network. In cluster 2, we can observe that low VKT / person uniquely characterizes high population density. (C) The power-law relationship between the local population and the network-level vehicle emissions produced in the Bay Area. A super-linear scaling with $\beta_{eff} = 1.12 \pm 0.13$ is found. (D) The linear relationship between the total VKT by the city residents and total emissions produced by them. The inline plot shows the distribution of the slope (gCO_2/km) per city, and the median is at 209 gCO_2/km. (E) Emissions of trips at the different congestion levels, keeping everything else (stop time, acceleration, distance) constant. Each line is Eq. 3 (F) Emissions per trip by the law presented in Eq. 4 with a vehicle efficiency of 25 mpg (10.6 km/l). The region within the dotted lines is the one standard deviation confidence interval.

Scaling of vehicular emissions

This section investigates the relationship between vehicular emissions and urban metrics. We analyze the role of urban form and travel patterns. Scaling laws in urban systems have been studied widely in various domains. Response of the quantity Y to a change in the independent variable, X, is represented in the form of a power law:

$$Y = \alpha X^{\beta} \tag{2}$$

Data from cities present scalings of the form of Eq. 2 where X is population and Y various socio-economic indicators (*23, 24, 28*). In the urban domain, scaling of emissions over a region with its population size has attracted significant attention (*47, 48*). Yet no consensus has been reached on the relationship. The value of the exponent β differs for the same data set with different urban borders (*35, 49, 50*). The presence of noise and lack of enough orders of magnitude on axes can pose a problem in establishing a significant scaling relationship (*51*).

Street length, congestion, and VKT are other scaling relationships explored (*48, 52*). Yet these existing studies lack reliable emission estimates or the resolution of the data is too low to establish a relationship with trips over the road networks.

We investigate scaling relationships between the daily emissions (gCO_2/day) production within the network in relation to vehicle kilometers traveled (VKT) and population. The daily emissions are calculated based on the home locations and the trips extracted from the TimeGeo model (*30*). CO_2 emission production by travelers across the Bay Area are aggregated into the home cities as defined by the US Census Bureau. The CO_2 emission production calculations were made based on an average vehicle efficiency of 25 mpg in the Bay Area.

To calculate the local scaling exponent, we generated scaling tomography plots suggested by Barthelemy et al. for noisy data (*51*) (see Supplementary Figure 5A). This method allows us to identify threshold effects and multiple scaling exponents that are not detectable by classical least squares fitting. A super-linear scaling with $\beta_{eff} = 1.12 \pm 0.13$ is found between the city population and the CO_2 emissions produced by them as shown in Fig 5 C.

As shown in Fig. 5D, we find a linear relationship between the total vehicle kilometers traveled (VKT) and the total CO_2 emissions produced by the city residents. We observe a clear linear relationship and no threshold effect in the tomography plots (see Supplementary Figure 5B). The slope of the best-fit line is the emission efficiency (gCO_2/km) of a trip and it varies among different cities. We find 209 gCO_2/km as the mean emission efficiency of the cities in our domain and the value can range between 175 to 225 $gCO_2/km.$ The difference between the cities arises from the different congestion levels experienced and the stop and go traffic.

We also analyze the population density as a function of $VKT/person.$Cities cluster into four groups within the Bay Area travel network in the analysis region, as shown in Fig. 5 A-B. K-means clustering is used to generate the clustering. We observe that high population density is associated with low VKT per person, but in low population density zones, VKT per person varies. Also, high population density results in high congestion levels experienced by their residents (see Supplementary Figure 6).

To quantify the impact of congestion and VKT, we further check emissions at the trip level for the same route at varying congestion levels. Figure 5E illustrates $1,500$ trips where each dot represents the trip's emissions for a given congestion level. The best fit of gCO_2 emissions (Y_i) of trip i is

$$Y_i = \alpha_i \exp\left(\beta_c X\right) \tag{3}$$

with $\beta_c = 0.004 \pm 0.001$ for all trips. At the same congestion level, emissions for a trip can vary significantly due to the VKT. The route difference presents itself in the intercept α_i, of each trip's emissions best-fit line. The α_i is, in turn, a function of the VKT and the fuel economy (FE). For trips with a 25 mpg vehicle, $\alpha_i = \gamma VKT_i$, and $\gamma = 190.6 \pm 33.3[gCO_2/km]$

Figure 5F shows Eq. 3 divided by VKT_i, where 89.7% of the trips lie within the confidence interval. The trips outside this region are short in distance and have much longer idling times per VKT. Supplementary Figure 8 illustrates the distribution of VKT, idle time per VKT, and acceleration per VKT for the trips outside and inside the confidence interval. We show that the trips outside the confidence interval have a mean idling time per VKT of 60 seconds, whereas this value is 18 seconds for the rest of the trips.

This relationship can be generalized to any fuel economy when we write α_i as a function of the manufacturer's reported fuel economy of the vehicle. Supplementary Figure 9 illustrates that the best fit for this relation is $\gamma = 2006.53 \pm 314.32[gCO_2/l]/FE$ where FE stands for the fuel economy in [km/l]. We get the general expression per trip emissions:

$$Y_i = \frac{2006.53[gCO_2/l]}{FE[km/l]} VKT_i \exp^{\beta_c X} \tag{4}$$

The median vehicle travel distance in the Bay Area is 15 km, and 95% of the trips are below 70 kms. This means VKT can vary emissions by 10% to 500%. Vehicle efficiency approximately ranges between 10 mpg to 119 mpge (mpg equivalent for electric vehicles), that would change the estimates of Eq. 4 from 20% to 300% of the mean efficiency for a passenger combustion vehicle. Varying the congestion level from free-flow to 100% can increase the emissions by 1.5 times or 50%. However, 100% congestion is rarely experienced, 95% of the trips in the Bay Area experience congestion levels below 62% which corresponds up to 28% enhancement of vehicular emissions.

Equation 4 is derived at the individual trip level, independent of specific city characteristics. While the above analysis is conducted for the Bay Area travel network, the equation can be universally applied to any city. The availability of VKT, congestion, and efficiency parameters enables this equation to provide a robust and straightforward estimation of transportation emissions with high spatial and temporal granularity.

Moreover, the MODES framework can be replicated to conduct in-depth analyses and reveal location-specific emission trends. The mobile phone activity data used in MODES is readily available in modern cities, and average road speeds are obtained through the Uber Movement Speeds API. While we utilize average speed data, which is more widely accessible compared to

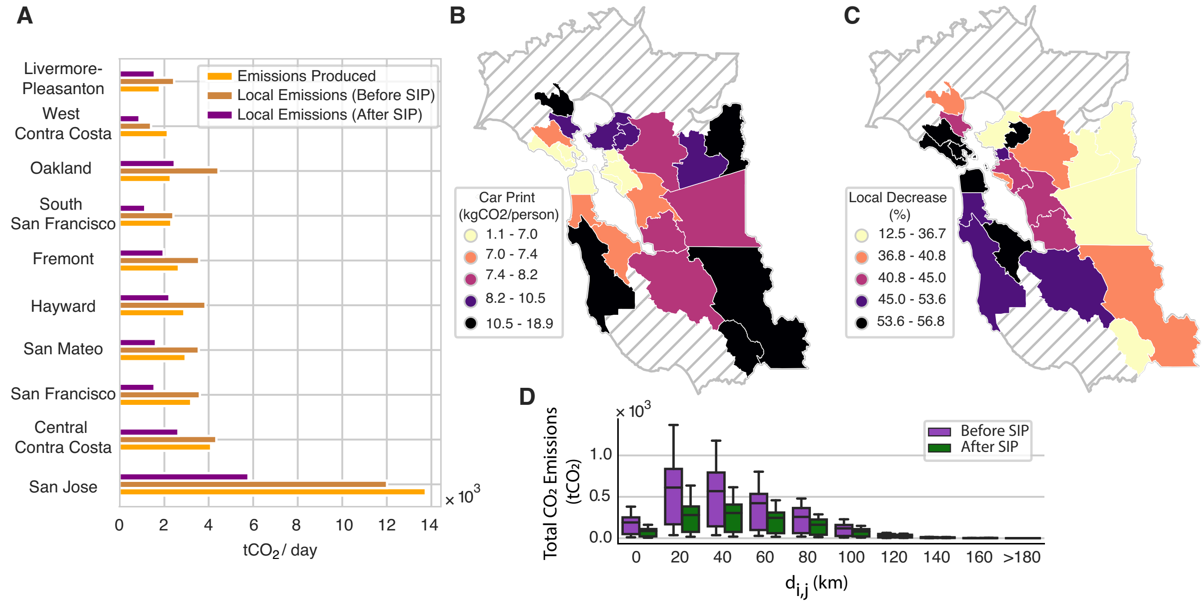

Figure 6: **Spatial analysis of the CO_2 emissions in the Bay Area.** (A) Comparison of total emissions produced by the county subdivision residents before SIP and the local emissions within the county subdivisions. The 10 regions with the most daily total emissions produced before shelter in place are shown (dark-yellow bar). (B) Per person emissions produced and its spatial variation by county subdivision. (C) Change in local emissions before and after SIP decisions and its spatial variation. (D) Total CO_2 emissions by trip distance in the periods of analysis. The 20-80 km range is where the majority of emissions occur. An important reduction of emissions is observed in this distance range. The SIP decisions did not affect the emissions of the long-distance trips.

instantaneous speed values, our validation process ensures the reliability and robustness of the obtained results.

Spatial analysis of on-road emissions and air quality disparities

Estimated emissions can vary noticeably for the same region depending on the calculation method. Transportation emissions can be either attributed to the regions where they are emitted (Scope-I) or to the region that the transportation activity started (Scope-II). The former is well quantified by regional monitoring of emissions and it is useful in understanding exposure. From an emission mitigation perspective, the latter is crucial to make policies directly targeting high emission-producing driver groups, their routes, and activities.

We compared the Bay Area county subdivisions in term of the daily CO_2 emissions (tCO_2/day) emitted locally in the region and the daily CO_2 emissions produced (tCO_2/day) by the residents anywhere within the network. Vehicular emissions increase the formation of smog and gasses harmful to respiratory health. Therefore, we introduce an equity perspective to the analysis by demonstrating the disparities between local emission exposure and emissions produced.

San Jose has the majority of emission production. Its emission production is about 2000 tones higher than the local emissions in the region. On the other hand, Oakland, San Francisco, and San Mateo have more local emissions exposure than the emissions produced by their residents. Particularly in Oakland, local emissions are about twice its emission production.

In Figure 6B, we show the spatial distribution of the emission production per capita. San Francisco, Oakland, Berkeley, Alameda, San Rafael, and Novato have the lowest emission production. Lower emission production of the mentioned regions can be attributed to the proximity of San Francisco, where most employment and leisure activities are present. As the distance to the activity center increases, the emission production per capita increases.

SIP orders decreased local emissions significantly. However, the spatial distribution of this change is not uniform as shown in In Figure 6C. The rural areas did not benefit from the emission reduction of remote work as much as employment centers like San Francisco and Sunnyvale. Interestingly East Bay county subdivisions Oakland, Hayward, and Fremont, experienced less reduction than Marin County subdivisions of San Rafael and Novato.

When we analyze the emissions over trip distances, we find that a 44% decrease in the VKT resulted in a 47% decrease in the transportation emissions in the Bay Area. Comparing the trips with different travel distances in Figure 6D, we observe that trips between $20km$ to $80km$ are responsible for most of the emissions across the travel network. Electrification of the trips at this range can decrease tail-pipe emissions significantly. An average electric vehicle battery with $135 km$ range can provide enough energy for a round trip for this distance bin.

Discussion

This work presents MODES, a data-driven framework for estimating vehicle emissions from mobile phone data validated by direct CO_2 measurements of the Berkeley Air quality and CO2 Network $(BEACO_2N)$. While both estimates have various sources of uncertainties, they have comparable results for 9 km^2 spatial aggregation and 5 days of temporal aggregation. MODES estimates are lower than the $(BEACO_2N)$ estimates, mostly because of our use of average passanger vehicles with 25 mpg efficiency and lack of trucks (on average 7 mpg efficiency in the Bay Area) in our model.

Our mobility science informed emissions model is further evaluated against a hybrid model developed with highway vehicle counts, highway speeds and truck percentages acquired from PeMS highway sensors and emissions factors acquired from the CARB's EMFAC2017 model. Different than the validation with $(BEACO_2N)$ data, we used PEMS heavy duty vehicle percentage estimates in MODES. We assigned a percentage of the mobile phone travelers as trucks and assumed a 7 mpg efficiency for them. Comparisons of the hourly emission estimates on different highway segments showed that results of the two model's are within 35% of each other before and during the SIP orders.

The use of before and after COVID-19 SIP data is crucial for our validation. We show that the framework successfully captures the changes in mobility and associated emissions in the San Francisco Bay Area six weeks before and during the shelter-in-place order due to COVID-19. In our analysis, we show that the work-from-home behavior decreased the total number of weekday trips. However, the share of long-distance trips increased. Remote working improved the road conditions by lowering daily average network-wide congestion from 28% to 15%. The impact of the lowered congestion on vehicle efficiency is captured by the proposed framework.

During the SIP, the median fuel efficiency of the trips increased by 6-7% per vehicle type due to decreased congestion. Furthermore, we showed that on-road vehicle efficiency can vary by 70% from the manufacturer fuel economy depending on the vehicle speeds and stops.

We quantify the relationship between vehicular emissions, congestion, and VKT with Equation 4. The importance of the equation lies in the accessibility of the used metrics. TomTom Traffic Index (*53*) provides live and historical congestion values. VKT can be approximated by vehicle sensor counts, fuel sales, or surveys and vehicle registrations provides insights on the manufacturing fuel economy of the vehicles used in a city. Using Equation 4, we found VKT as the main factor affecting emission production for a fleet of vehicles, compared to the effects of the VKT.

Using individual mobility data allows us to relate emissions produced across the network to the home locations of travelers. This provides valuable insights into sustainable urban forms that local emission measurements can not provide. We observe that cities with high population density have lower VKT and CO_2 emission production per person. Nevertheless, the congestion levels are higher for high population density regions. Considering the relationship obtained in Eq. 3, high population density is desirable from a climate change mitigation perspective.

Another advantage of the use of individual mobility data in the emissions estimates is the ability to quantify the disparities between the CO_2 exposure vs. its production by different location sources. The ability to capture origin and destination pairs with highest emissions have the potential to enrich policy interventions. From an emission reduction perspective, the residents of San Jose have the highest total emissions. Therefore, this region should be prioritized in mobility decarbonization efforts. From an equity perspective, MODES allows to identify the trips responsible for the emissions in a region and guides necessary actions for air quality improvements.

There are various avenues in which this work can be extended. Lack of truck data in our model causes underestimation of the local emissions. For policy evaluation, vehicle efficiencies, flow values, and road network variables (stops, congestion etc.) can be altered to perform scenario analysis. For example, it is possible to illustrate local and network-level emission reduction benefits of electric vehicle adoption per driver source. Also, adding trip purposes on the current results can reveal patterns of emission intensity from work and leisure trips. Our framework does not incorporate the measurement of instantaneous acceleration along the roads due to the unavailability of such data. Consequently, there is a possibility of underestimating some trip emissions. However, our results still provide valuable insights for comparing trips among different individuals. While the inclusion of acceleration data would enhance the framework, our main contribution lies in the development of an alternative function that relies on three macroscopic trip variables: VKT, congestion, and fuel efficiency. This feature enables our model to estimate emissions without relying on detailed speed profiles. Despite the absence of acceleration data, our validation results demonstrate good comparability with direct measurements from $(BEACO_2N)$ and EMFAC estimates.

Methods

Individual Mobility Model

Mobile phone activity data and call detail records (CDR's) have been widely adopted in mobility modeling studies. In this work we utilized two datasets developed by mobile phone activity analysis: (1) TimeGeo urban mobility model (2) SafeGraph mobility data (*31*).

TimeGeo model in the San Francisco Bay Area built and validated against NHTS and CHTS (*30*) for a typical weekday. We adopted the trips in this model as the base scenario for the travel behavior before the shelter in place decisions. Each trip is associated with a user ID, home census tract for the user, trip purpose, timestamp and origin-destination coordinates.

To extend this model to the each day of the week for before and after shelter in place decisions we followed a simple scaling process using the SafeGraph data. We first extended the average weekday behavior before SIP to 5 days of the week. Then a daily-hourly flow change as a percentage between each OD is calculated.

SafeGraph, a data company that aggregates anonymized location data from numerous applications in order to provide insights about physical places, via the SafeGraph Community. To enhance privacy, SafeGraph excludes census block group information if fewer than five devices visited an establishment in a month from a given census block group (*31*). Due to the aggregation steps towards meeting privacy requirements hourly origin destination (OD) flows are not provided in any of the SafeGraph data products. We retrieved data from Neighborhood Patterns and Social Distancing Metrics between Jan 1-May 31, 2020. We inferred the hourly OD flows for each day of the week before and after shelter in place decisions combining these two datasets. Neighborhood Patterns data is one of the main products of the company which provides monthly analysis of census block group and point of interest visits. We used the *stops_by_each_hour* variable reporting the number of stops starting at each hour throughout the month for a CBG. Social Distancing Metrics product has been released following the COVID-19 confinements. For each origin CBG, destination CBG counts aggregated over a day is provided. We created the hourly OD matrix M as using the daily hourly origin destination matrix D and the total hourly stops matrix S for each destination as following:

$$M_{i,j,t,d} = D_{i,j,d} S_{j,t,d} \tag{5}$$

Where *i*'s are origin CBGs, *j*'s are destination CBGs, *t* is the hour of the day and *d* is the day of the month. The inferred daily hourly OD flows between CBGs are separated as before and after shelter in place flows. The OD flows then averaged for each weekday and hour pair to infer a representative behavior.

Road Network and Speeds

We obtain the San Francisco Bay Area road network as a directed graph where each edge is the road and each node is an intersection. Road and intersection geometry is acquired from Open Street Map (OSM). OSM provides road type, maximum speed and number of lanes associated with each road. For the intersections we retrieved traffic signals, stops, crossings, and junctions to model the stop and go traffic. To retain this information no simplification is performed on the final graph.

The hourly speed values are retrieved from the Uber Movement Speeds API which provides hourly mean speed values of the street segments in several metropolitan areas starting from 2018. In this project, the data between 06/2019 and 4/2020 is retrieved. Its important to note that, Uber Movement API stopped reporting data after 4/2020. Due to Uber's data collection and reporting protocol, we don't have hourly speed values on all of the streets. Only the road segments that attracted enough traffic in each hour are included to protect customer privacy. Therefore to get the missing speed values, we applied a K-nearest neighbors (KNN) data imputation method to have a network with complete information. (see Supplementary Notes S1)

Fuel Consumption and Vehicle Emissions Model

Fuel consumption of the trips is estimated with the regression based model StreetSmart. This model is developed using the OBD-II and GPS data collected from 600 miles of driving. Authors included 4 variables in the final model after testing a range of parameters. The final model with a mean accuracy of 96% in the tested scenarios is presented in Equation 1.

The model predicts the fuel consumption in US gallons. T_{idle} is the stopping time, T_{move} is the moving time in seconds, a is acceleration in $\frac{m}{s^2}$ integrated over the distance and L is the distance driven in km. The k's in the model are vehicle efficiency specific constants and they are calibrated for different vehicle efficiency bins. The model presented is detailed enough to capture the efficiency changes due to the speed variations and stop-and-go traffic. We further validated this model with Autonomie simulations and EPA speed profiles. The results are presented in Figure.

We converted the gallons of fuel consumption to CO_2 emissions in grams using the CO_2 intensity of the fuel. We take the fuel density as $0.75\frac{g}{ml}$ and carbon intensity of the fuel by weight as $0.86\frac{gC}{gFuel}$. Carbon is converted into carbon dioxide in the combustion process with a weight ratio of 12 to 44.

CO_2 Sensor Measurements

We use hourly CO_2 observations from the Berkeley Air quality and CO_2 Network $(BEACO_2N)$. The CO_2 observations are converted to hourly emissions with Stochastic-Time Inverted Lagrangian Transport (STILT) model, coupled with a Bayesian inversion. (*8*)

The inversion process requires meteorology data and prior emission estimates from different sources. Prior emissions sources included in the $BEACO_2N - STILT$ inversion are home heating data distributed spatially according to population density, fuel sales data distributed spatially according to vehicle counts (*10*) and biogenic inventory derived using Solar Induced Fluorescence (SIF) satellite data (*54*). The resulting posterior emissions are stored in $1km^2$ grid cells. Transportation emissions are estimated by subtracting the non-transportation priori sources from posterior emissions. $BEACO_2N - STILT$ inversion is estimated to be precise to at least 30% for a line source (*55*).

The influence region of the sensors are the regions that the emissions are likely to be originated from. In this work we followed Fitzmaurce et. al. (*42*) and included the 40% influence region in the analysis.

References

1. EPA. Inventory of us greenhouse gas emissions and sinks: 1990-2017. Technical Report, U.S. Environmental Protection Agency (2019).

2. 2018 progress report: California's sustainable communities and climate protection act (2018).

3. Agency, E. E. Trends and projections in europe 2021: Tracking progress towards europe's climate and energy targets. Technical Report, European Environment Agency (2020).

4. Shusterman, A. A. *et al.* The berkeley atmospheric co_2 observation network: initial evaluation. *Atmospheric Chemistry and Physics* **16**, 13449–13463 (2016).

5. Peters, G. *et al.* Carbon dioxide emissions continue to grow despite emerging climate policies. *Nature Climate Change. DOI* **10** (2019).

6. Banister, D. Cities, mobility and climate change. *Journal of Transport Geography* **19**, 1538–1546 (2011). Special section on Alternative Travel futures.

7. Liu, Z. *et al.* Near-real-time monitoring of global co 2 emissions reveals the effects of the covid-19 pandemic. *Nature Communications* **11**, 1–12 (2020).

8. Turner, A. J. *et al.* Observed impacts of covid-19 on urban co2 emissions. *Geophysical Research Letters* e2020GL090037.

9. Gurney, K. R. *et al.* Under-reporting of greenhouse gas emissions in u.s. cities. *Nature Communications* 553.

10. McDonald, B. C., McBride, Z. C., Martin, E. W. & Harley, R. A. High-resolution mapping of motor vehicle carbon dioxide emissions. *Journal of Geophysical Research: Atmospheres* 5283–5298.

11. Kwon, J., Varaiya, P. & Skabardonis, A. Estimation of truck traffic volume from single loop detectors with lane-to-lane speed correlation. *Transportation Research Record* 106–117.

12. Gurney, K. R. *et al.* High resolution fossil fuel combustion co2 emission fluxes for the united states. *Environmental Science & Technology* 5535–5541.

13. Caubel, J. J., Cados, T. E., Preble, C. V. & Kirchstetter, T. W. A distributed network of 100 black carbon sensors for 100 days of air quality monitoring in west oakland, california. *Environmental Science & Technology* 7564–7573.

14. Ramaswami, A. *et al.* Carbon analytics for net-zero emissions sustainable cities. *Nature Sustainability* 460–463.

15. Gurney, K. *et al.* Quantification of fossil fuel co2 emissions on the building/street scale for a large u.s. city. *Environmental Science amp; Technology* **46**, 12194–12202 (2012).

16. Gately, C. K., Hutyra, L. R. & Wing, I. S. Cities, traffic, and co¡sub¿2¡/sub¿: A multi-decadal assessment of trends, drivers, and scaling relationships. *Proceedings of the National Academy of Sciences* 4999–5004.

17. Gately, C. K., Hutyra, L. R., Peterson, S. & Sue Wing, I. Urban emissions hotspots: Quantifying vehicle congestion and air pollution using mobile phone gps data. *Environmental Pollution* **229**, 496–504 (2017).

18. Blaudin de Thé, C., Carantino, B. & Lafourcade, M. The Carbon 'Carprint' of Suburbanization: New Evidence from French Cities (2020). Working paper or preprint, `https://halshs.archives-ouvertes.fr/halshs-02572893`.

19. Miller, E. J. & Ibrahim, A. Urban form and vehicular travel: Some empirical findings. *Transportation Research Record* **1617**, 18–27 (1998).

20. Böhm, M., Nanni, M. & Pappalardo, L. Gross polluters and vehicle emissions reduction. *Nature Sustainability* **5**.

21. Xu, Y., Olmos, L. E., Abbar, S. & González, M. C. Deconstructing laws of accessibility and facility distribution in cities. *Science advances* **6**, eabb4112 (2020).

22. Çolak, S., Lima, A. & González, M. C. Understanding congested travel in urban areas. *Nature Communications* **7**, 10793 (2016).

23. Bettencourt, L. M. A., Lobo, J., Helbing, D., Kühnert, C. & West, G. B. Growth, innovation, scaling, and the pace of life in cities. *Proc Natl Acad Sci U S A* **104**, 7301–7306 (2007).

24. Bettencourt, L. M. A. The origins of scaling in cities. *Science* **340**, 1438–1441 (2013).

25. Newman, M. E. J. Power laws, pareto distributions and zipf's law. *Contemporary Physics* **46**, 323–351 (2005).

26. Depersin, J. & Barthelemy, M. From global scaling to the dynamics of individual cities. *Proceedings of the National Academy of Sciences* **115**, 2317–2322 (2018).

27. Barthelemy, M. The statistical physics of cities. *Nature Reviews Physics* **1**, 406–415 (2019).

28. Bettencourt, L. M. A., Samaniego, H. & Youn, H. Professional diversity and the productivity of cities. *Sci Rep* **4**, 5393 (2014).

29. Olmos, L. E., Çolak, S., Shafiei, S., Saberi, M. & González, M. C. Macroscopic dynamics and the collapse of urban traffic. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 12654–12661 (2018).

30. Xu, Y., Çolak, S., Kara, E. C., Moura, S. J. & González, M. C. Planning for electric vehicle needs by coupling charging profiles with urban mobility. *Nature Energy* **3**, 484–493 (2018).

31. Safegraph. `https://www.safegraph.com`. Accessed 2021-03-01.

32. Uber movement speeds. URL `https://movement.uber.com/?lang=en-US`.

33. Oehlerking, A. L. *StreetSmart : modeling vehicle fuel consumption with mobile phone sensor data through a participatory sensing framework*. Ph.D. thesis, Massachusetts Institute of Technology. Dept. of Mechanical Engineering. (2011).

34. Kalila, A., Awwad, Z., Clemente, R. D. & González, M. C. Big data fusion to estimate urban fuel consumption: A case study of riyadh. *Transportation Research Record* **2672**, 49–59 (2018).

35. Louf, R. & Barthelemy, M. Scaling: Lost in the smog. *Environment and Planning B: Planning and Design* **41**, 767–769 (2014).

36. Jiang, S. *et al.* The timegeo modeling framework for urban mobility without travel surveys. *Proceedings of the National Academy of Sciences* **113**, E5370–E5378 (2016).

37. Bureau of transportation statistics. trips by distance band. `https://www.bts.gov/browse-statistical-products-and-data/covid-related/trips-distance-groupings-national-or-state`. Accessed 2022-01-10.

38. Fuel economy guide: Model year 2021. Tech. Rep., U.S. Department of Energy.

39. Federal test procedure. Tech. Rep., EPA (2021).

40. Yao, J. & Moawad, A. Vehicle energy consumption estimation using large scale simulations and machine learning methods. *Transportation Research Part C: Emerging Technologies* 276–296.

41. Gately, C. K. & Hutyra, L. R. Large uncertainties in urban-scale carbon emissions. *Journal of Geophysical Research: Atmospheres* 11,242–11,260.

42. Fitzmaurice, H. L. *et al.* Assessing vehicle fuel efficiency using a dense network of co_2 observations. *Atmospheric Chemistry and Physics* **22**, 3891–3900 (2022).

43. Agency, U. E. P. Greenhouse gas emissions from a typical passenger vehicle. Technical report, EPA (U.S. Environmental Protection Agency) (2018).

44. California Air Resources Board. *EMFAC2017 Volume I - user guide*.

45. Caltrans performance measurement system. https://pems.dot.ca.gov/.

46. Lauvaux, T. *et al.* High-resolution atmospheric inversion of urban co2 emissions during the dormant season of the indianapolis flux experiment (influx). *Journal of Geophysical Research: Atmospheres* **121**, 5213–5236 (2016).

47. Fragkias, M., Lobo, J., Strumsky, D. & Seto, K. C. Does size matter? scaling of co2 emissions and u.s. urban areas. *PLOS ONE* **8** (2013).

48. Mohajeri, N., Gudmundsson, A. & French, J. R. Co2 emissions in relation to street-network configuration and city size. *Transportation Research Part D: Transport and Environment* **35**, 116–129 (2015).

49. Cottineau, C., Hatna, E., Arcaute, E. & Batty, M. Diverse cities or the systematic paradox of urban scaling laws. *Computers, Environment and Urban Systems* **63**, 80–94 (2017).

50. Arcaute, E. *et al.* Constructing cities, deconstructing scaling laws. *Journal of The Royal Society Interface* **12**, 20140745 (2015).

51. Barthelemy, M. Tomography of scaling. *Journal of the Royal Society Interface* **16**, 20190602 (2019).

52. Louf, R. & Barthelemy, M. How congestion shapes cities: from mobility patterns to scaling. *Scientific reports* **4**, 1–9 (2014).

53. Tomtom traffic index. `https://www.tomtom.com/traffic-index/`. Accessed: 2022-11-07.

54. Turner, A. J. *et al.* A double peak in the seasonality of california's photosynthesis as observed from space. *Biogeosciences* **17**, 405–422 (2020).

55. Turner, A. J. *et al.* Network design for quantifying urban co_2 emissions: assessing trade-offs between precision and network density. *Atmospheric Chemistry and Physics* 13465–13475.

Acknowledgements

This work was supported by Engie SA, the ITS-SB1 Berkeley Statewide Transportation Research Program and the California Air Ressources Board.

Author Contributions

A.T.O. and M.C.G. conceived the research and designed the analyses. A.T.O. developed the code and performed the statistical analyses, created the plots, contributed to the interpretation of the results and wrote the paper. H.F provided the data, contributed to the work methodology and interpretation of the results. A.T.O. and M.C.G. wrote the paper. O.K, P.C, R.C.C, and M.C.G provided general advice and supervised the research.