Biking&Breathing: A Dangerous Combination?

Assessing Bicycle Commuters’ Exposure to Air Pollution in Downtown and West Oakland, CA

Data Revolution

Over the past 30 years, numerous studies have attempted to assess the air pollution exposure associated with urban cycling [1]. For the lack of a better alternative, most of these studies have relied on an extremely small sample of air pollution data collected over few select routes over the course of a couple of days [1]. Others have attempted to model air pollution and bicycle traffic volume via land-use-based models [2]. However, when compared with real-life data, these models fail to explain over 50% of the variability within air pollution and bicycle traffic volume levels [2]. Research in this field has been hindered by the lack of widely-available large-scale datasets.

However, the data revolution is underway and big data has the potential of revolutionizing research in this field. Over the past decade, Motivate has operated numerous bike share programs in metropolitan areas all over the United States and collected detailed data on every single trip taken with one of their share bicycles. This data can be used to estimate share bicycle traffic volumes for all the areas Motivate operates in. Additionally, in 2014, Google launched a pilot project monitoring street-level air pollution data via sensor-equipped Google Street View vehicles in parts of Oakland, CA. The raw data gathered in this project was aggregated by a research team at the University of Texas at Austin. The first results of the project – detailed air pollution data for 30-foot road segments in Downtown, East, and West Oakland was released to the public in late 2018 [3]. The following examines how this dataset can be combined with share bicycle trip data from Ford GoBike – the Bay Area bike share program operated by Motivate – to assess bicycle commuters’ exposure to air pollution in Downtown and West Oakland.

Methodology

Data on all share bicycle trips taken within the Ford GoBike system in the year 2017 was acquired form Motivate. This dataset included the timestamp, origin bike share station, destination bike share station, and low-level user information for every trip. Share bicycle traffic volume levels for all the bikable road segments in the East Bay were estimated as follows. Trips confined to San Francisco and San Jose were removed from the dataset. Trips where the origin and destination stations were the same were also removed as they do not provide any information on the movement of the bicycle. The rest of the trips were assumed to be commuter trips. Thus, it was assumed users would take the most time-optimal route between their origin and destination stations and prefer routes with bike infrastructure. The Mapbox Bicycle Routing Application Program Interface (API) was used to estimate the optimal route between the two endpoints for every trip. This API was chosen due to its performance and ease of use and because it optimizes for time and prefers routes with bicycle infrastructure. To increase performance, directional information was removed and it was assumed that all trips with the same endpoints would follow the same route.

A network of bikable roads in the East Bay was compiled from street data acquired from Alameda County and the United States Census Bureau. To match routes generated by the Mapbox Bicycle Routing API with the bikable road network and aggregate the number of annual trips for every road segment, the following was done. The study area was divided into a 50×50-foot grid. This cell size was chosen due to the unidirectional nature of the route data – most bidirectional bikable roads in the area have a width of 50-feet or less. All the cells that did not intersect a bikable road segment were discarded. All the routes generated by the Mapbox API fell within the remaining cells – justifying the chosen cell size. The annual number of share bicycle trips was aggregated for each cell, generating a share bicycle traffic volume estimate for each 50-foot road segment in the East Bay.

The study area for further analysis was determined based on the extent of the available air pollution data. The dataset included annual mean BC, NO, and NO2 levels for all drivable road segments in Downtown, West, and East Oakland with a 30-foot resolution. The dataset for East Oakland was omitted due to the lack of bike share stations in the area. Because the dataset did not include measurements for some bike paths within the study area that lack car access, all of the air pollution indicators were interpolated to ensure full coverage of all the bikable roads within the area. Inverse distance weighting (IDW) with a linear coefficient was used as it has been shown to be the most accurate at interpolating high-density datasets [4]. The resulting overlays had a spatial resolution 50×50-feet to match the road network and were clipped to the extent of the original air pollution data points. Because the spatial distribution of the air pollution indicators was not uniform, an Air Pollution Index (API) was calculated for each data point to normalize the data and condense air pollution into a single indicator. The API assigned equal weights to each indicator and was calculated as follows:

$$API = (BC/BCMAX + NO/NOMAX + NO2/NO2MAX)/3$$

The API ranges from zero to one with one indicating a data point with the highest air pollution in the area according to all indicators and zero indicating a data point with the lowest possible air pollution in the area. All interpolated air pollution indicators, including the API were joined with their corresponding 50-foot road segments.

Due to the lack of widely-accepted safe limits of BC, NO, and NO2, the methodology from [2] was adapted to assess bicycle commuters’ exposure to air pollution. All 50-foot bikable road segments were ranked based on their air pollution indicators and annual bicycle traffic volume estimates and divided into corresponding quartiles. Based on these quartiles, all road segments were placed into one of four categories – Sweet Spot, Active and Exposed, Inactive and Clean, and Sour Spot. Road segments in the Sweet Spot category are within the bottom 50% in regards of air pollution (assessed via API) and the top 50% in regards of ABMT (derived from bicycle traffic volume). These road segments are clean and popular with bicyclists and their use is encouraged. Road segments in the Active and Exposed category are within the top 50% in terms of both air pollution and ABMT. These are polluted roads with high bicycle traffic volumes. The pollution of these road segments should be addressed or an alternate route for cyclists should be provided. Road segments in the Inactive and Clean category are within the bottom 50% when it comes to both air pollution and ABMT. These are clean roads with low bicycle traffic and their usage by cyclists and pedestrians should be encouraged while keeping automobile traffic low. Road segments in the Sour Spot category are within the top 50% in terms of air pollution and in the bottom 50% in regards of ABMT. These are polluted roads actively avoided by cyclists. Their use by cyclists should be discouraged and motor traffic from more popular cycling routes should be diverted to these road segments if possible.

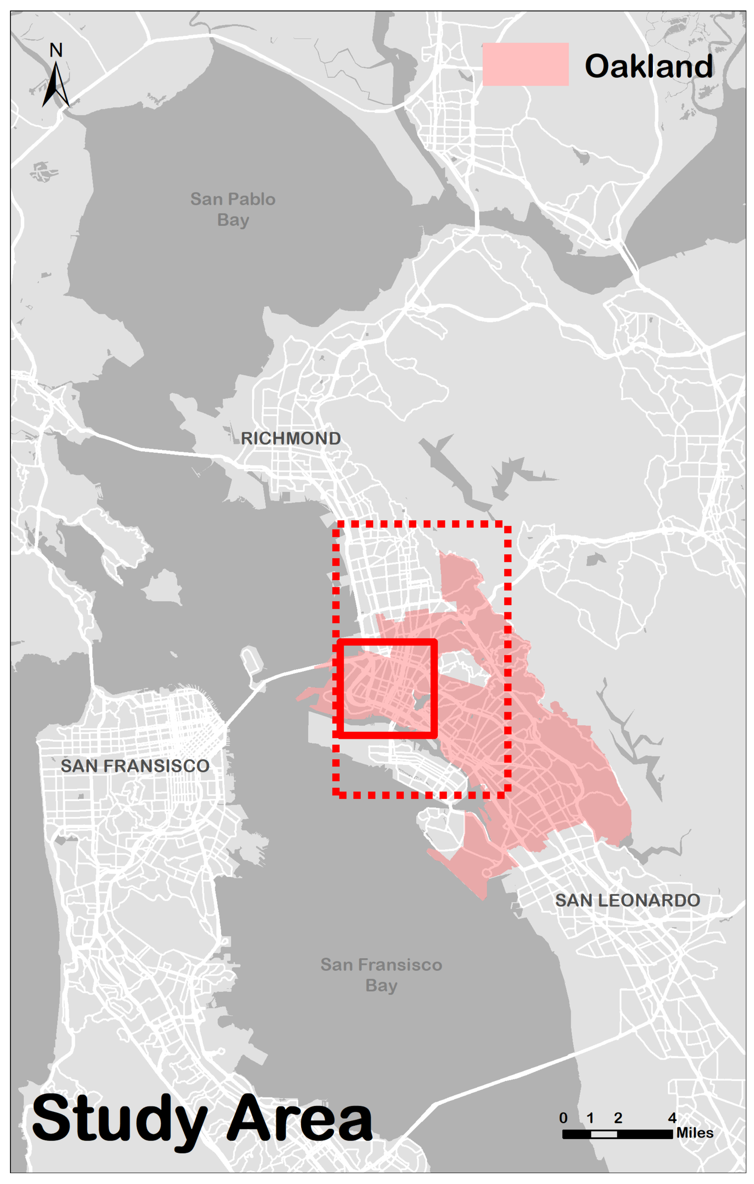
The distribution of total bikable road mileage and ABMT over these categories was examined. Additionally, the distribution of ABMT over the air pollution indicator quartiles was studied. Finally, the air pollution indicators for each unique route between two bike share stations were aggregated and the relationship between route popularity (annual trips along route) and route pollution (average air pollution indicator values along the route) was examined.

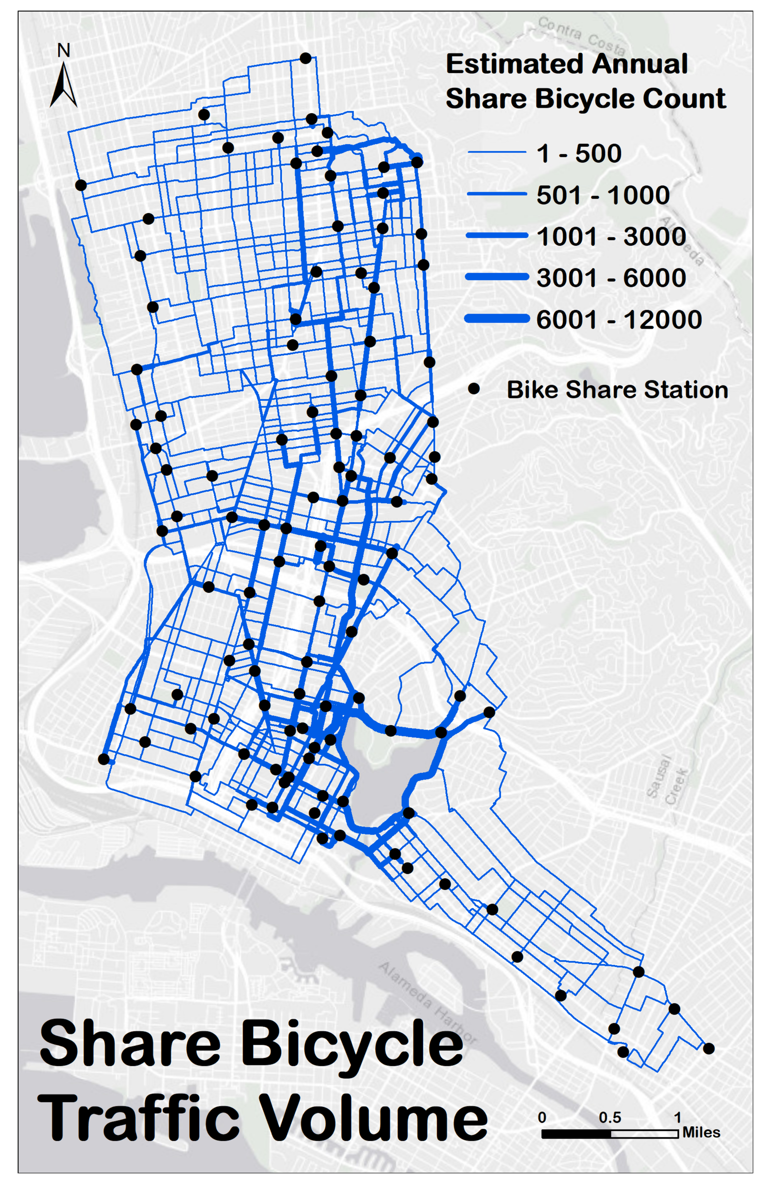
Results

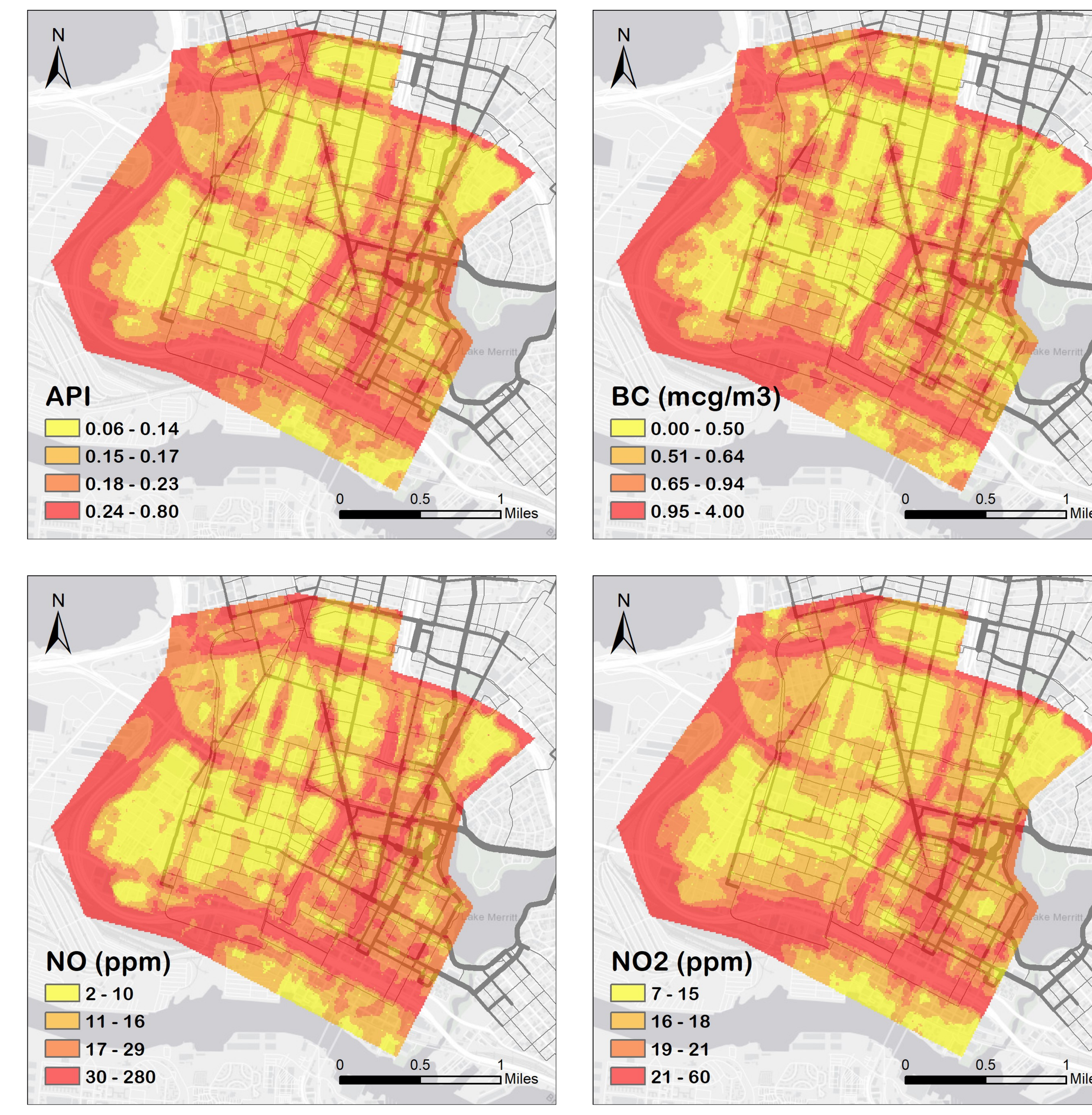
Results indicate that the majority of share bicycle traffic in Downtown and West Oakland take place on roads with high levels of air pollution. About 60% of all ABMT are on roads that are within the top two quartiles in regards to air pollution levels as assessed by the API. When looking at individual air pollution indicators, we see that about 54% of all ABMT are on roads within the top two quartiles in terms of BC levels, roughly 68% of all bicycle traffic takes place on roads within the top two quartiles in terms of NO levels, and about 59% of all ABMT are on roads within the top two quartiles in terms of NO2 levels.

Additionally, the examination of the relationship between the average air pollution levels along a unique route between two share bicycle stations and the estimated number of annual trips along said route concluded a positive correlation between to two in terms of all air pollution indicators, including the API. Meaning that more popular bike share routes tend to have increased levels of air pollution. This again enforces the finding that most share bicycle traffic in the area takes place on roads with high levels of air pollution.

However, the findings suggest there is room for improvement. 99% of current share bicycle traffic utilizes only a third of the available bikable road mileage. About 35% of the bikable road mileage falls within the Inactive and Clean category and sees only 1% of ABMT. To reduce bicycle commuters’ exposure to air pollution and thus increase public health, future planning efforts should focus on utilizing these road segments for bicycle traffic while keeping motorized traffic volumes low.







Glossary

ABMT – Annual Bicycle Miles Traveled

API – Air Pollution Index

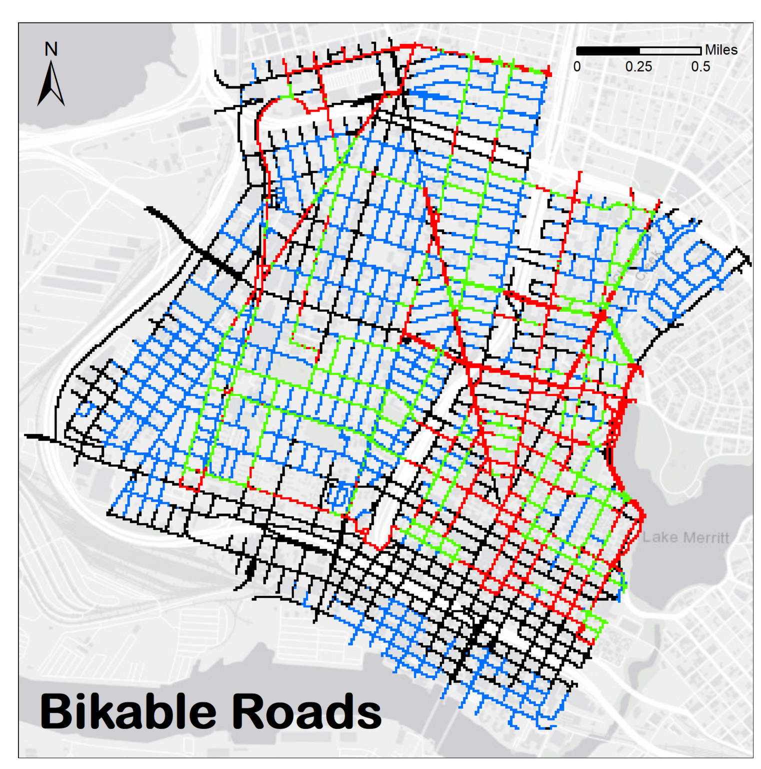
BC – Black Carbon

NO – Nitric Oxide

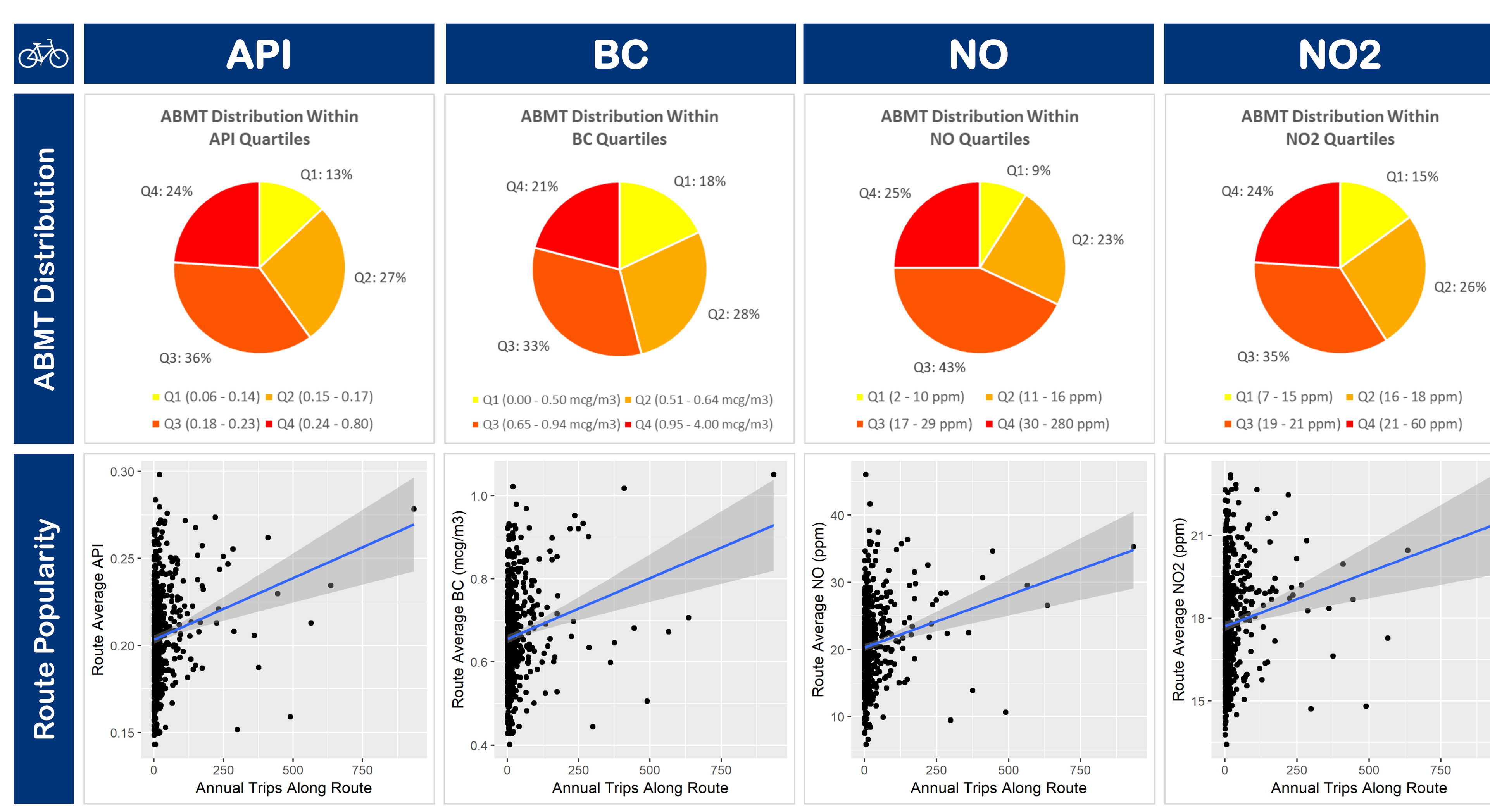
NO2 – Nitrogen Dioxide

mcg – micrograms

ppb – parts per billion



Color	Type	AMBT	API	% Total Bikable Road Mileage	% Total AMBT
Green	Sweet Spot *High Bicycle Traffic Volume and Low Air Pollution*	Top 50%	Bottom 50%	15%	39%
Red	Active and Exposed *High Bicycle Traffic Volume and High Air Pollution*	Top 50%	Top 50%	17%	60%
Blue	Inactive and Clean *Low Bicycle Traffic Volume and Low Air Pollution*	Bottom 50%	Bottom 50%	35%	1%
Black	Sour Spot *Low Bicycle Traffic Volume and High Air Pollution*	Bottom 50%	Top 50%	33%	0%



Applications

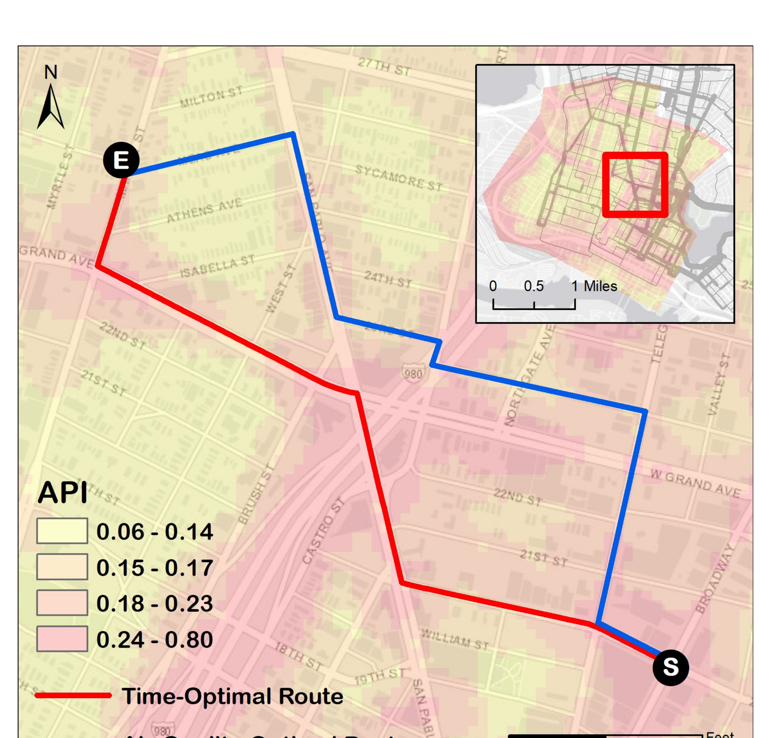
To help increase bicycle traffic on the 35% of all available bikable roads within the Inactive and Clean category, the potential for an air-quality-optimal routing system was explored. Most modern bicycle routing systems take into account a number of variables like distance, terrain, traffic, and the availability of dedicated cycling infrastructure to generate a time-optimal route between two desired endpoints. However, as the findings of this study suggest, these routes could have much higher levels of air pollution than their somewhat longer or slower alternatives. With the increasing emergence of sensor-equipped lamp and traffic sign posts, buildings, bus stop pavilions, and monitoring vehicles, detailed air pollution data for urban areas is becoming increasingly available. And often this data is available in real-time. This opens up the possibility of developing new bicycle routing systems that take into account either current or historic air pollution data and solve for the most optimal route both in regards of time and air quality.

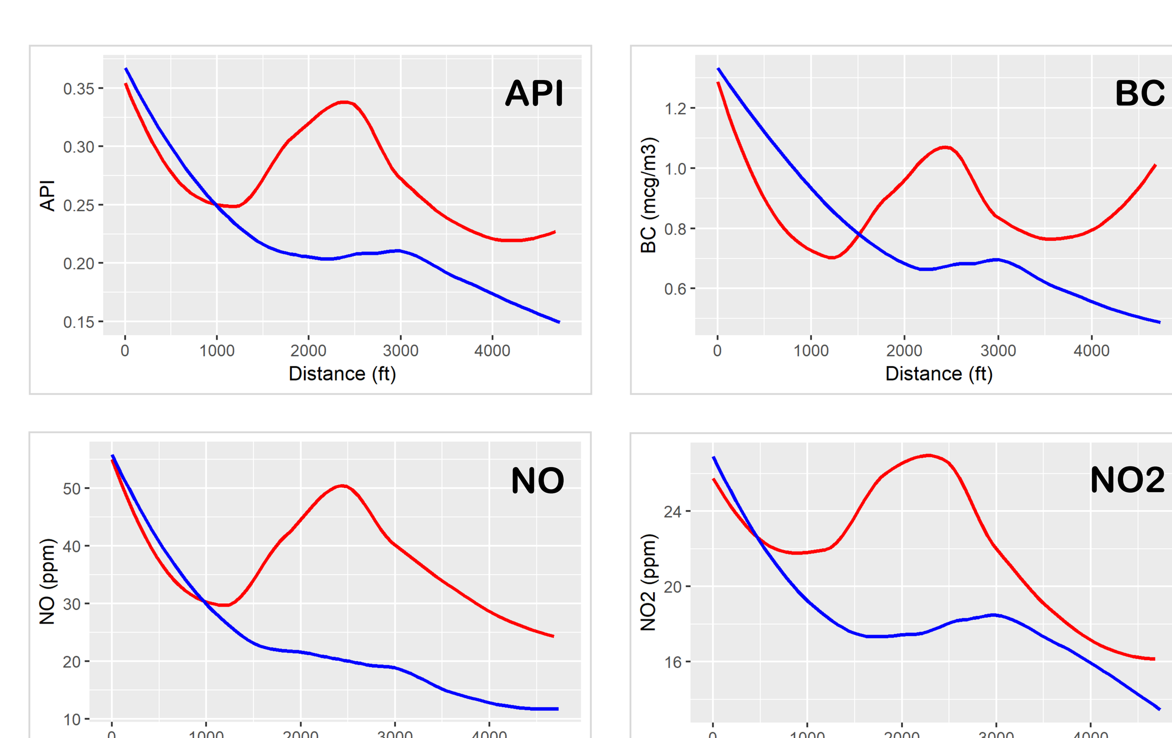
The potential for such a routing system was explored using the derived data and the cost path calculator. A sample route was chosen that was in the highest quartile both in terms of air pollution and annual trip count. An alternate route was calculated using a cost path calculator with the objective of minimizing the average API along the route. The new route was only 2% longer than the original and resulted in a 20% decrease of average API along the route. A detailed analysis of the various air pollution indicators reveal that the new route had higher pollution levels that the original for short sections (especially in terms of BC), but overall there was a notable decrease of air pollution over distance in terms of all four air pollution indicators. This suggests that air-quality-optimal bicycle routing systems are feasible and yield positive results. Further resources should be allocated towards researching and implementing air-quality-optimal routing systems for cyclists.

Limitations

It is important to note the various limitations and drawbacks of the techniques used in this study. First of all, the study focused on solely share bicycle traffic from the Ford GoBike docked bike share system. Numerous dockless bike share systems also operate within the Bay Area, but data from these systems is not usually publicly available and thus was not included in the study. However, it is important to note that data from dockless bike share systems is incredibly useful for future studies of this nature as these bicycles are GPS-tracked and allow the exact route for each trip to be determined. Bicycle trips taken on personal bicycles were also naturally not accounted for.

Furthermore, it is likely this study overestimated share bicycle traffic volume levels on most road segments. It is highly unlikely that all bicycle commuters travelling between the same two endpoints follow the same time-optimal path proposed by the Mapbox API. A significant portion of cyclists would most likely follow alternative routes, dispersing the ABMT amongst greater bikable road mileage. Also, all trips were automatically assumed to be commuter trips regardless of duration. Some of these trips might have been longer and of a more scenic nature, dispersing the ABMT even more. Also, the omission of directional information increases error in areas with one-way traffic.





Uku-Kaspar Uustalu

ENV-0170 | 6 February 2019 | uku-kaspar.uustalu@tufts.edu

Data Sources

Air Pollution: Aclima, EDF, Google, University of Texas at Austin

Bicycle Trips: Mapbox, Motivate (Ford GoBike)

Shapefiles: Alameda County, California Geoportal, DataSF, Marin County, OakData, San Mateo County, U.S. Census Bureau (TIGER)

Basemap: ESRI, Garmin, HERE, OpenStreetMap

Coordinate System: NAD 1983 StatePlane California III
Projection: Lambert Conformal Conic

References

[1] De Hartog, Jeroen Johan, Hanna Boogaard, Hans Nijland, and Gerard Hoek. "Do the Health Benefits of Cycling Outweigh the Risks?" *Environmental Health Perspectives* 118, no. 8 (2010): 1109-1116.

[2] Hankey, Steve, Greg Lindsey, and Julian D. Marshall. "Population-Level Exposure to Particulate Air Pollution During Active Travel: Planning for Low-Exposure, Health-Promoting Cities." *Environmental Health Perspectives* 125, no. 4 (2016): 527-534.

[3] Messier, Kyle P., Sarah E. Chambliss, Shahzad Gani, Ramon Alvarez, Michael Brauer, Jonathan J. Choi, Steven P. Hamburg et al. "Mapping Air Pollution with Google Street View Cars: Efficient Approaches with Mobile Monitoring and Land Use Regression." *Environmental Science & Technology* 52, no. 21 (2018): 12563-12572.

[4] Wong, David W., Lester Yuan, and Susan A. Perlin. "Comparison of spatial interpolation methods for the estimation of air quality data." *Journal of Exposure Science and Environmental Epidemiology* 14, no. 5 (2004): 404.

Tufts UNIVERSITY