Modeling the impact of tolling in large-scale regional networks: a case study for DVRPC

PI: Sean Qian (ORCID: 0000-0001-8716-8989)

Lead Research Assistant: Qiling Zou (ORCID: 0000-0003-2146-558X)

Research Assistant: Pengji Zhang

DISCLAIMER

All data generated from this project can be accessed from https://github.com/psychogeekir/Philly

Final Report

Modeling the impact of tolling in large-scale regional networks: a case study for DVRPC

Qiling Zou[1] and Sean Qian*[1,2]

[1]Department of Civil and Environmental Engineering, Carnegie Mellon University
[2]Heinz College, Carnegie Mellon University

January 4, 2023

1 Introduction

Recent years have witnessed a sharp decline in a national highway trust fund that is primarily used to support infrastructure construction, expansion, and retrofit. This is a challenging and critical issue for almost all states, including Pennsylvania. The primary reason is due to the decline of gas tax collected at gas pumps as a result of the adaption of high fuel efficiency vehicles. The most recent and ongoing COVID-19 crisis adds an additional burden to this lack of public funds for infrastructure since the overall travel demand declines drastically. Most of the states are currently evaluating the implications of tax loss, and proactively developing plans to collect funds in a more equitable and effective manner. For example, the Pennsylvania Department of Transportation (PennDOT) has proposed to add tolls on nine candidate bridges in the PennDOT PAthways Major Bridge P3 Initiative [1], as a way to collect funding to support infrastructure.

However, the transportation network is a complex system where many different vehicle classes (primarily cars and trucks) and travel modes (primarily private vehicles and ride-hailing vehicles) co-exist. Tolling can impact the traffic for each class and each mode differently. In addition, different communities in a region may be affected by tolling differently, leading to social equity issues. It is yet to quantify the influence of tolling on the network performance for each community and incorporate equity into the decision-making of tolling. And the pricing and locations of tolls need to be carefully designed to justify the infrastructure projects in the greatest need for a region. Therefore, it is important yet very challenging to effectively and accurately evaluate the societal consequences of the tolling of various forms, including social equity, congestion delay, emission, fuel use, and potential toll revenue.

To this end, this research project aims to develop a large-scale multi-class network modeling and simulation framework, that holistically models the spatiotemporal behaviors of private cars, ride-hailing cars, and freight trucks, respectively. Based on the framework and the data collected and processed by CMU's Mobility Data Analytics Center (MAC) from multiple sources, the traffic dynamics in the Delaware Valley Regional Planning Commission (DVRPC) region are simulated in which individual cars' and trucks' route choices and travel times are modeled. The result includes the prediction of travel time, travel delay, vehicle-mile-traveled, and emissions for each of those vehicle classes, either at road and intersection level or averaged

*Email: seanqian@cmu.edu

at the system level by time of day. Potential tolling strategies, such as locations and pricing, can be evaluated and deployed, with the consideration of both system mobility and social equity.

The project is divided into three major tasks:

- Task 1: Collecting traffic data for dynamic network modeling

  The following data are collected, processed, and integrated for network modeling and further analysis:

  - Network data (GIS model) for the DVRPC region
  - Traffic volume data for the DVRPC region
  - Traffic speed data for the DVRPC region
  - Vehicle registration data for the DVRPC region
  - Tolling program information

- Task 2: Establishing a dynamic network model for the DVRPC region

  The mesoscopic dynamic network model, MAC-POSTS, developed by the CMU's MAC is used to simulate the dynamic traffic flows over time in the DVRPC region. Based on the data collected in Task 1, the multi-class dynamic origin-destination demand estimation (MCDODE) is first carried out in order to estimate the day-to-day origin-destination demand among all Traffic Analysis Zones (TAZs) that vary by time of day. The DVRPC regional network is coded into the dynamic network model. Baseline travel demand is estimated using the integrated traffic data on typical weekdays without tolling on the proposed candidate locations. With the estimated demand, the network model is then able to replicate the close-to-real-world traffic dynamics. It also has the capacity to model dynamic traffic evolution with the consideration of travel control and traffic demand management. This model adopts state-of-the-art traffic models and is much more computationally efficient than other microscopic models that are extremely labor-intensive to establish.

- Task 3: Modeling and evaluating different tolling plans

  The scenarios of different tolling plans are modeled on the calibrated regional network model. The simulation adopts the historical traffic demand and their pre-scribed route choices from the dynamic network model established in Task 2. The overall traffic impacts induced by different tolling plans are evaluated using both system-level performance metrics (e.g., total traffic delay, average travel time, emissions, energy use, and vehicle miles traveled) and local-level traffic flow changes.

2 Data collection and pre-processing

The Greater Philadelphia Region is traffic data rich compared to other metropolitan areas in the U.S. Various data sets in this region, including traditional traffic sensors (loops, cameras, etc.) and cutting-edge sensors (Bluetooth, GPS probe, ride-hailing vehicle samples, etc.), are available and have been archived for quite long time. The rich data sets allow us to learn travelers' behavior accurately and develop an in-depth understanding of the traffic impact of tolling systems in large-scale networks.

This section briefly discusses the multiple data sources used in this project, including network topological data, traffic count and speed data, and vehicle registration data.

2.1 Network description

The network topological data from the DVRPC covers nine counties in the Greater Philadelphia Region with Philadephia city in the central area, as shown in Figure 1. The original network data is trimmed so that there are no isolated nodes and links. The isolated nodes and links represent the parkways in the real world, and the absence of such nodes and links does not affect the network analysis. In addition, some neighboring links with small lengths and the same speed limit are further consolidated, which can substantially reduce the network size. More importantly, network consolidation has great potential in improving the accuracy of network analysis [2].

The final network model used for the following analyses contains 41,586 road segments, 18,294 intersections, and 205 traffic analysis zones (TAZs). Among TAZs, a total of 41,820 origin-destination (OD) pairs are considered.

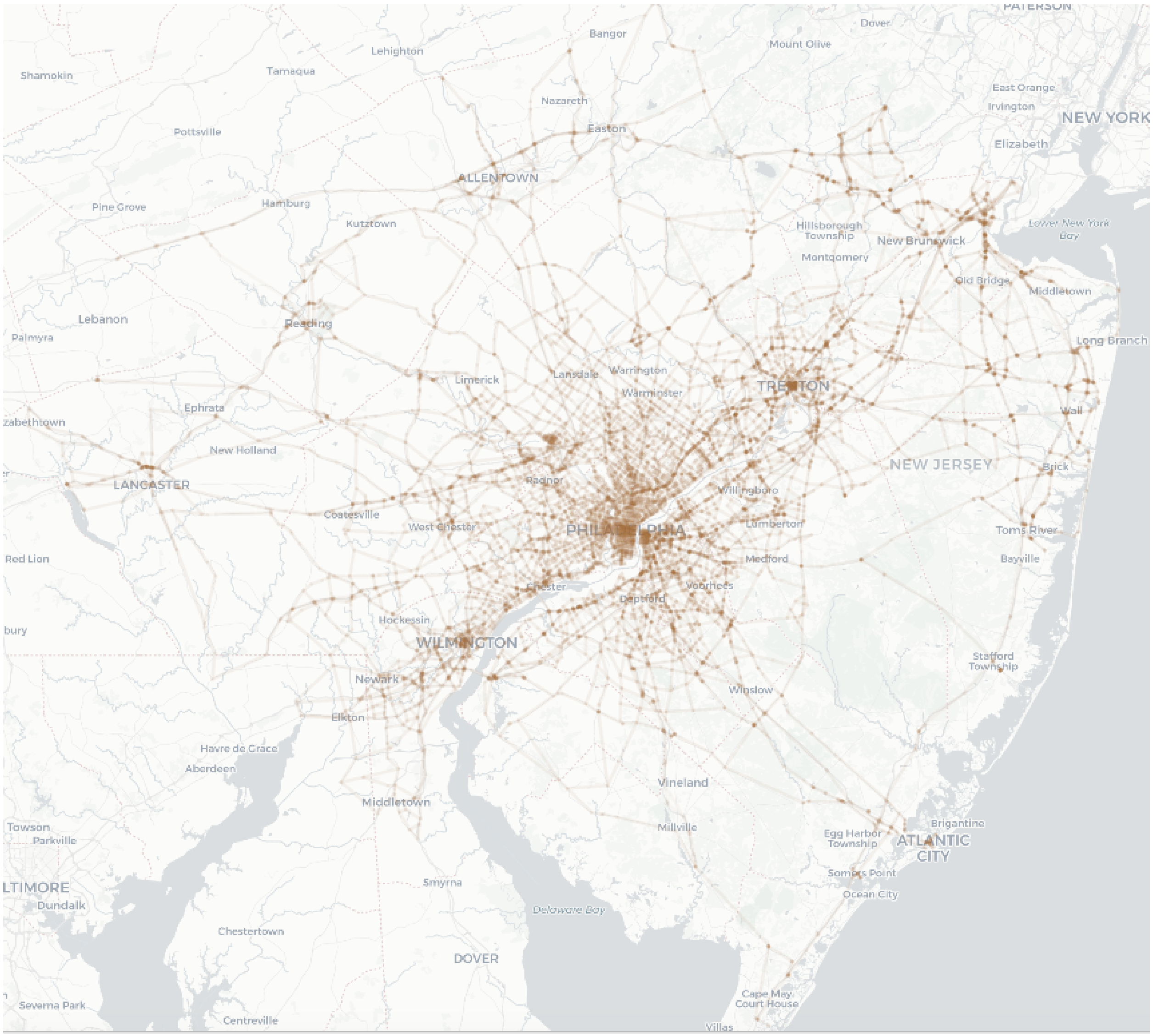

Figure 1: An overview of the DVRPC network

2.2 Traffic counts

Traffic count data represents the vehicle counts passing by a certain location, and it is usually collected by loop detectors, tubes, or manual counting. In this project, the count data is also provided by the DVRPC. The original count data contains 15-minute traffic volume counts for different vehicle types at selected locations in this region each day in 2019. The count data is carefully examined and cleaned and matched to the links in the transportation network. Two vehicle types, i.e., cars and trucks, are counted separately in the data, which represent smaller private or ride-hailing vehicles and larger freight trucks, respectively. In total, there are 1,920 locations with valid car and truck volumes, as shown in Figure 2.

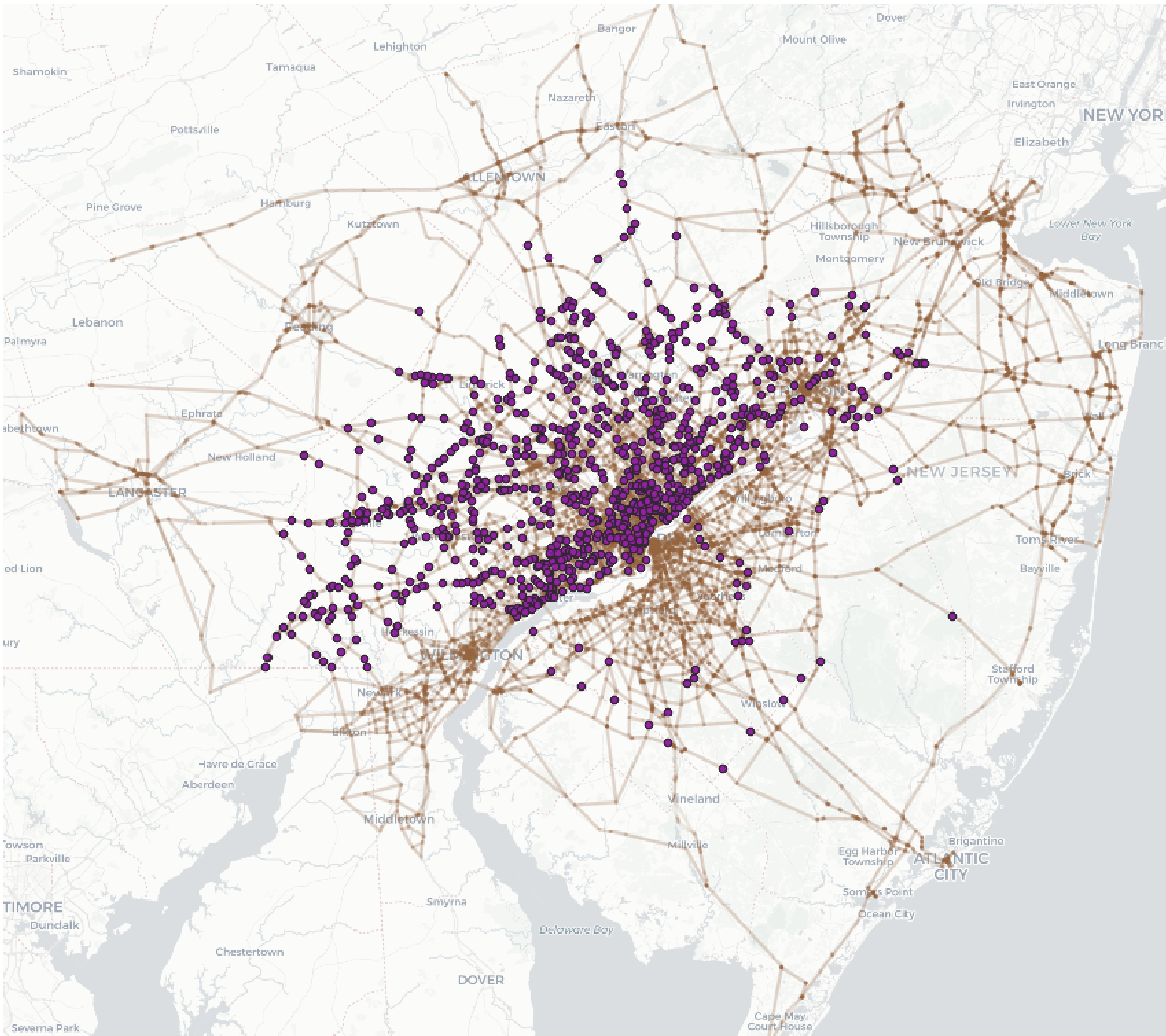

Figure 2: An overview of the traffic count locations.

2.3 Traffic speed

Traffic speed data is obtained from INRIX for the year 2019. Speeds of different vehicle types are measured separately, and hence both passenger car speeds and freight truck speeds are available. All the speed data is measured every 5 minutes of each day, and we average the data for different days in 2019 and aggregate the data to 15-minute intervals. There are a total of 3,420 locations with valid car and truck speed measurements, as shown in Figure 3.

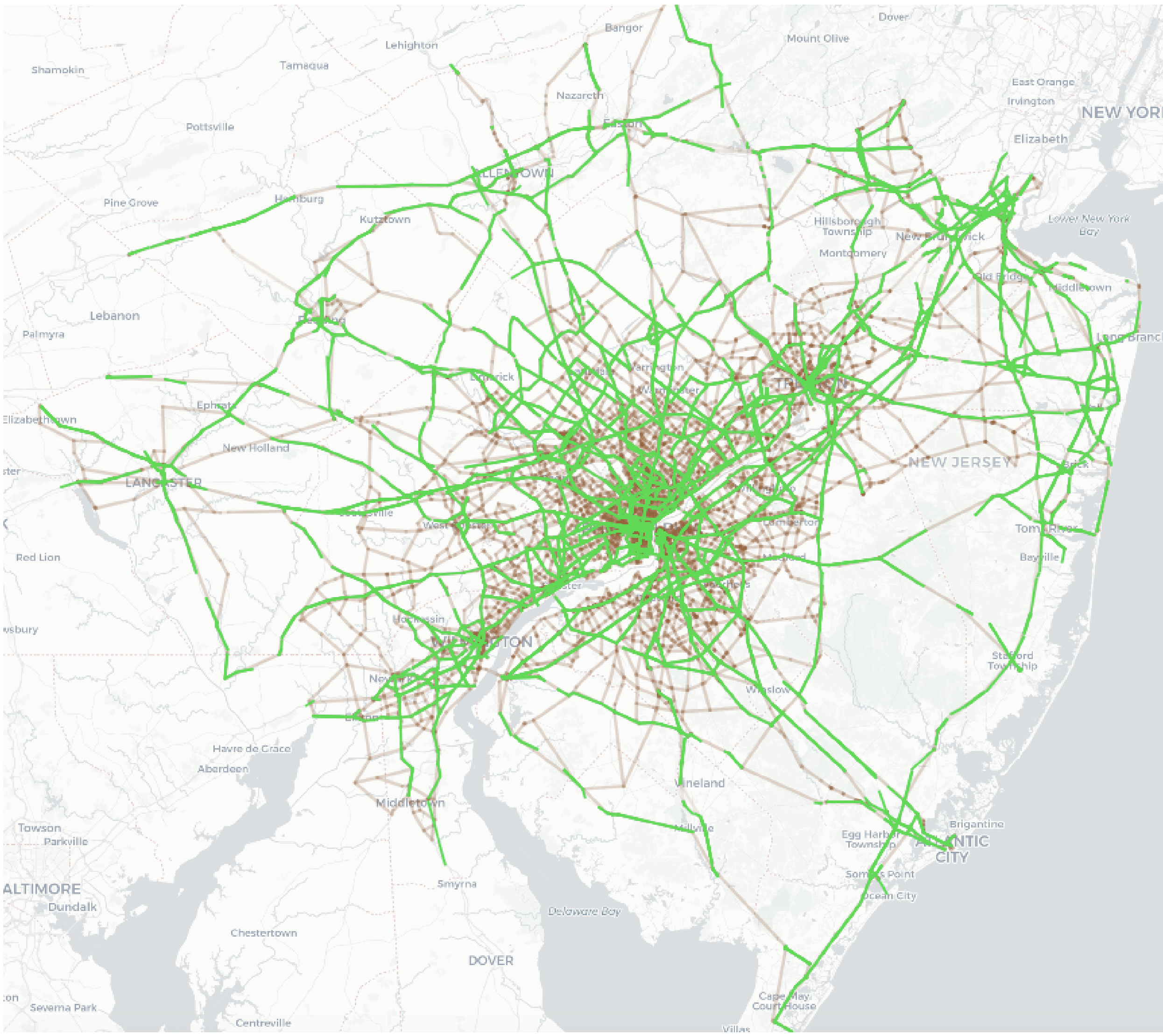

Figure 3: An overview of the speed data.

2.4 Vehicle registration information

In addition to the traditional traffic count and speed data, vehicle registration data is also used to calibrate the dynamic network model. Since vehicle registration is mandatory for all vehicles, nearly all vehicles in the region are covered by this data. The data is usually available from the Department of Motor Vehicles. With this data set, we can estimate the total number of vehicles in each by vehicle classification, and the detailed information regarding each of those vehicles, for example, engine type, vehicle size, year, etc. Integrating this data enables MAC-POSTS to estimate more fine-grained vehicle demands for further different research needs (e.g., more accurate emission estimation based on the vehicle's fuel types).

The vehicle registration data is obtained from PennDOT and contains eight columns:

- Vin: Vehicle Identification Number(VIN), which is a unique ID for a vehicle;
- Zipcode: zip code of the zone where the car owner or lessee live;
- Body type: body type of the vehicle;
- Titestdt: title established data;
- Odomval: the odometer reading at the time of reporting;
- Expdate: the expiration date of the registration;
- Storig: state of origin;
- Vstop: v stop. This field may be empty.

The most important two features are Zipcode and Vin. From Zipcode we can estimate how many vehicles there are in each zone; from Vin we can find details on the vehicles—model, make, engine type, etc. We use these two features to calculate the number of vehicles of different types for each origin and add this as an additional term in the objective function in the MCDODE task.

3 Modeling Current Traffic Conditions

This section describes traffic dynamics modeling for the DVRPC region in the morning peak hours.

3.1 Mesoscopic traffic simulation

In this project, the traffic dynamics in the region are simulated in high spatiotemporal resolutions. The regional model simulates millions of vehicles that depart from their respective origins, arrive at their destinations, and follow different routes. The CMU Mobility Data Analytics Center develops a dynamic network tool (MAC-POSTS) which is capable of simulating network-wide traffic dynamics for any general networks consisting of freeways, arterials, and local streets [3]. MAC-POSTS adopts the state-of-art mesoscopic traffic flow model and can scale up to regional-level transportation networks. MAC-POSTS can be calibrated to replicate real-world traffic conditions and predict the impact of different traffic scenarios, such as tolling, work zones, events, and incidents. The recurrent and stabilized traffic conditions are considered in MAC-POSTS, while the non-recurrent events such as crashes are not considered.

The AM peak hours (5:00 AM - 10:00 AM) are considered for simulation. MAC-POSTS simulates the movements of all vehicles in the studied network with high spatial (50 meters) and temporal (5 seconds) resolution. As with the information provided, we assume 65% of travelers are adaptive to the traffic information, while 35% of travelers will stick to the pre-determined routes when they travel.

3.2 Network calibration

3.2.1 Multi-class dynamic OD demand estimation

Before applied to practical applications, the dynamic network model needs to be calibrated in order to approximately reproduce the actual traffic conditions. To this end, multiple data sources collected in Section 2 are used and a data-driven calibration framework is adopted to calibrate the model. The adopted framework estimates the time-dependent traffic demands and travelers' behaviors, which are the two critical inputs to MAC-POSTS, and the traffic conditions (e.g. traffic volumes, traffic speed, delays) outputted by MAC-POSTS can reflect reality to some extent. To be precise, the performance of MAC-POSTS is measured by how well it can replicate the observed traffic data. This is referred to as the MCDODE problem. The objective function consists of five terms

$$\begin{aligned}
\min_{\{\mathbf{q}_{\text{car}},\mathbf{q}_{\text{truck}}\}} \quad \mathcal{L} =& \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 + \mathcal{L}_4 + \mathcal{L}_5 \\
=& w_1(\|\mathbf{y}'_{\text{car}} - \mathbf{y}_{\text{car}}\|_2^2) \\
&+ w_2(\|\mathbf{y}'_{\text{truck}} - \mathbf{y}_{\text{truck}}\|_2^2) \\
&+ w_3(\|\mathbf{z}'_{\text{car}} - \mathbf{z}_{\text{car}}\|_2^2) \\
&+ w_4(\|\mathbf{z}'_{\text{truck}} - \mathbf{z}_{\text{truck}}\|_2^2) \\
&+ w_5(\|\mathbf{g}' - \mathbf{g}\|_2^2)
\end{aligned} \tag{1}$$

where $\mathbf{q}_{\text{car}}$ and $\mathbf{q}_{\text{truck}}$ are the car and truck demands, respectively; $\mathbf{y}'_{\text{car}}$ and $\mathbf{y}_{\text{car}}$ are the observed and estimated car flows, respectively; $\mathbf{y}'_{\text{truck}}$ and $\mathbf{y}_{\text{truck}}$ are the observed and estimated truck flows, respectively; $\mathbf{z}'_{\text{car}}$ and $\mathbf{z}_{\text{car}}$ are the observed and estimated car travel times, respectively; $\mathbf{z}'_{\text{truck}}$ and $\mathbf{z}_{\text{truck}}$ are the observed and estimated truck travel times, respectively; $\mathbf{g}'$ and $\mathbf{g}$ are the observed and estimated vehicle demands at origins, respectively; w_1, w_2, w_3, w_4, and w_5 are the weights to balance the five terms in the optimization. Note that the origin vehicle demand $\mathbf{g}'$ appeared in $\mathcal{L}_5$ is calculated based on the vehicle registration data in Subsection 2.4.

More details of the calibration framework and the computational-graph-based solution method are omitted, and readers are referred to previous studies [3].

3.2.2 Calibration results

During the AM peak hours, around 2 million vehicles travel on the studied regional network. Simulated traffic conditions are calibrated to match the observed data.

Figure 4 presents the comparison between simulated traffic volumes and observed traffic volumes, in which the y-axis is the simulated counts and the x-axis is the observed counts. The coefficient of determination R^2, as a measure of goodness of fit, is 0.37, 0.42, and 0.97, for the car flow, truck flow, and origin vehicle demand, respectively.

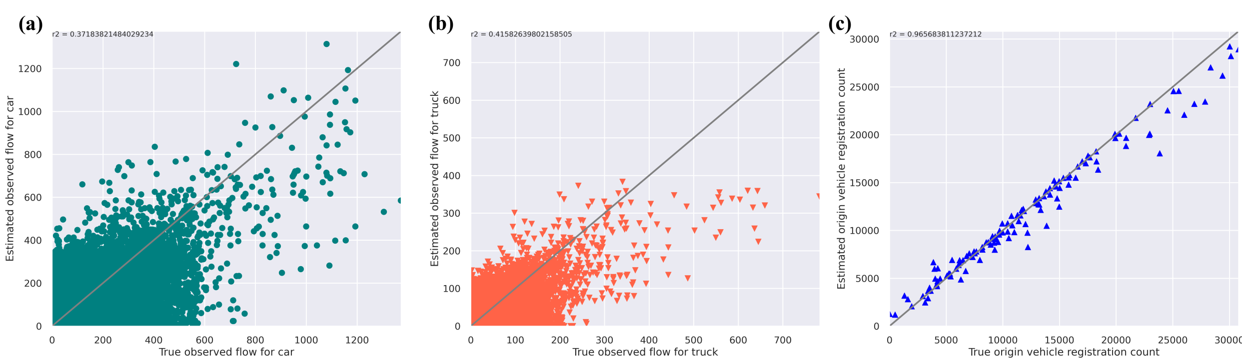

Figure 4: Overall count calibration in AM peak: (a) car flow; (b) truck flow; (c) origin vehicle demand.

Figure 5 depicts the comparison between simulated travel times and observed travel times, in which the y-axis is the simulated travel time and the x-axis is the observed travel time. The R^2 is 0.87 and 0.89, for the car travel times and truck travel times, respectively.

Note that calibration for such a large regional dynamic network work is known to be a challenging task and the result can be affected by different factors such as noise in the data and the network simplification. The count estimation accuracy is expected to be improved provided more MCDODE iterations. Overall, our model shows relatively good performance in capturing the trend of the observed data and this indicates that the proposed regional model can reflect the actual traffic dynamics in the whole DVRPCS region to some extent.

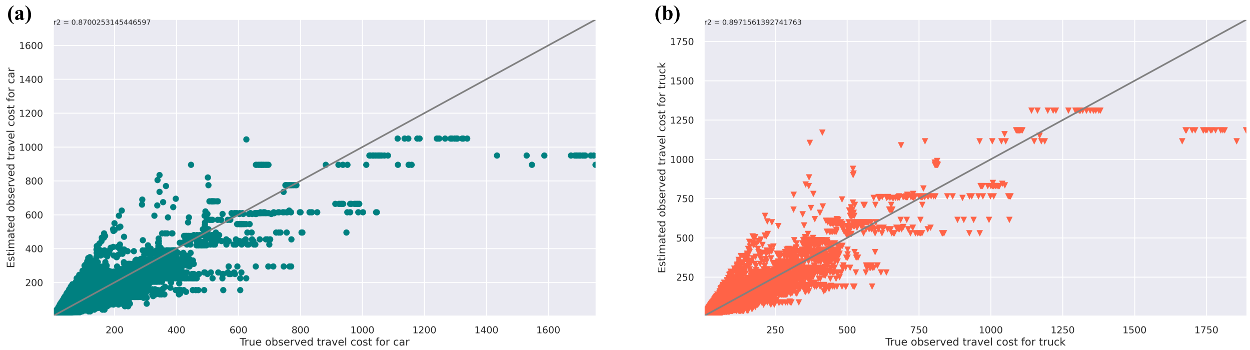

Figure 5: Overall travel time calibration in AM peak: (a) car travel time; (b) truck time.

4 Traffic Impact Analysis for PennDOT Bridge Tolling Program

PennDOT is currently looking at ways to implement various tolling strategies to improve the efficiency, reliability, and safety of Vine Street expressway corridor while having the ability to collect public funds for infrastructure construction projects. The tolling along this corridor will involve considerable traffic pattern changes, detours, and other traffic diversions for both trucks and private and TNC vehicles. It will also impact bus operations across the bridges to/from Center City. This section presents the traffic impact analysis for the bridge tolling program.

4.1 Scenario settings

According to PennDOT, Girard Point Bridge on the I-95 interstate in Philadelphia county (see Figure 6) is proposed to be tolled for both directions of travel and the cost is between $1.00 and $2.00 [4].

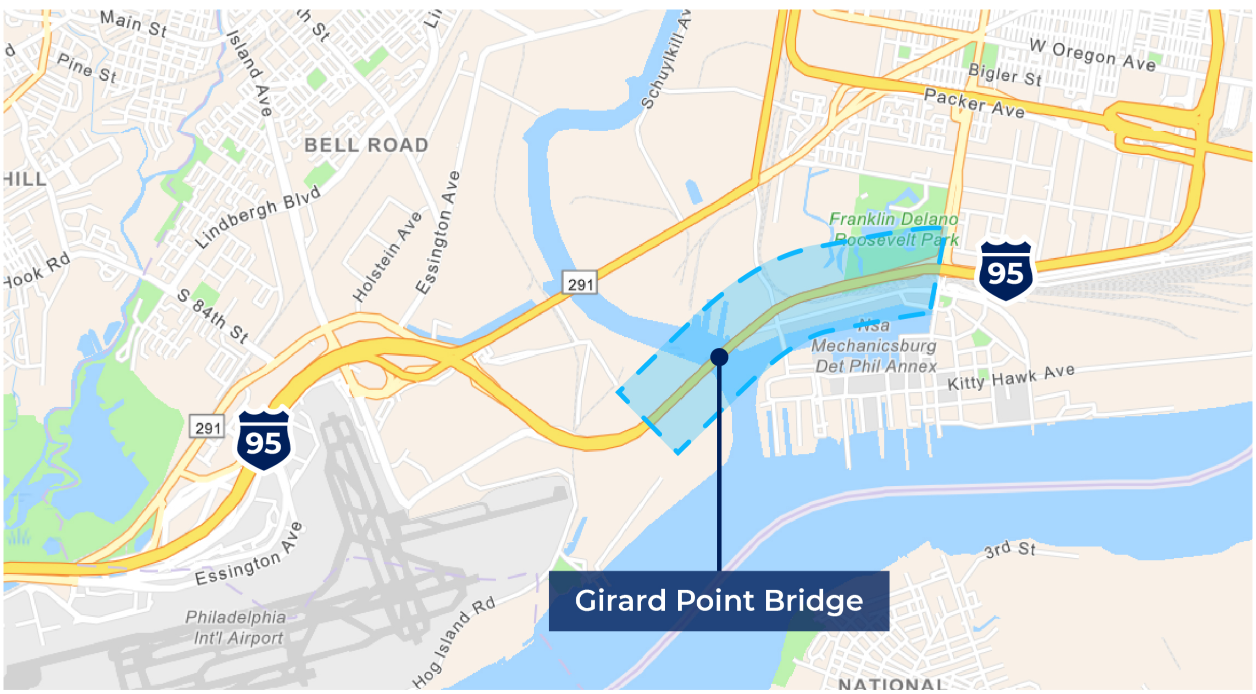

Figure 6: Girard Point Bridge [5].

We set up three scenarios to investigate the impact of the tolling on regional traffic. The baseline scenario represents the current no-toll situation. Scenarios 1 and 2 assume a toll of $1.00 and a toll of $2.00, respectively, for both directions for each vehicle using this bridge. In our simulation, the 65% of travelers will adapt to this tolling and may change their routes while the remaining 35% travelers will still stick to the pre-determined routes.

4.2 Traffic impacts

4.2.1 Generic metrics

The aggregated traffic metrics within the DVRPC region, including travel time, delays, Vehicle Miles Traveled (VMT), and emissions are presented in Table 1. The metric changes from the baseline scenario to different tolling scenarios are shown in Table 2. In each table, the average (total) travel time indicates the average (total) time spent within the DVRPC region, and the average delay represents the average waiting time at each intersection in the DVRPC region.

As can be seen, with the introduction of tolling, travel time, fuel consumption, and emissions are increased. Since this is a large regional network, the relative changes in these metrics at the system level caused by only one bridge tolling may not be very significant, but in the absolute values, it still indicates some substantial impacts on fuel consumption and emissions. It is expected that these metrics changes for a more local area near the bridge can be more significant. And comparing Scenario 1 with Scenario 2, it can be found that the increase in the toll results in an increase in travel time, fuel consumption, and emissions, because facing a higher toll, travelers are more likely to take detours.

The total toll revenues in different scenarios can be estimated by multiplying the toll with the number of passing vehicles and are presented in Table 3.

Table 1: General metrics in the AM peak hours.

		Total vehicles #	**Total travel time hour**	**Average travel time hour**	**Average delay second**	**VMT mile**	**Fuel gallon**	**CO2 ton**	**CO ton**	**HC ton**	**NOX ton**
Baseline	Car	1,604,373.5	2,052,337.1	1.3	11.4	52,975,339.9	1,908,235.5	16,958.5	63.8	44.6	59.5
	Truck	399,947.5	571,819.9	1.4	11.4	17,331,507.4	941,201.6	8,364.5	82.3	30.2	117.6
Scenario 1	Car	1,604,387.5	2,113,781.6	1.3	11.4	53,163,655.5	1,919,841.5	17,061.6	64.1	44.9	59.8
	Truck	400,143.5	585,932.5	1.5	11.4	17,419,338.4	947,672.6	8,422.0	82.8	30.4	118.4
Scenario 2	Car	1,604,330.5	2,157,446.0	1.3	12.0	53,890,143.4	1,947,461.6	17,307.1	64.9	45.8	60.6
	Truck	400,043.0	599,026.1	1.5	12.0	17,636,433.9	960,328.7	8,534.4	83.5	31.0	120.1

Table 2: Change of general metrics in the AM peak hours.

		Total vehicles #	**Total travel time hour**	**Average travel time hour**	**Average delay second**	**VMT mile**	**Fuel gallon**	**CO2 ton**	**CO ton**	**HC ton**	**NOX ton**
Baseline	Car	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Truck	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Scenario 1	Car	14.0	61,444.5	0.0	0.0	188,315.6	11,606.0	103.1	0.3	0.3	0.3
	Truck	196.0	14,112.6	0.1	0.0	87,831.0	6,471.0	57.5	0.5	0.2	0.8
Scenario 2	Car	-43.0	105,108.9	0.0	0.6	914,803.5	39,226.1	348.6	1.1	1.2	1.1
	Truck	95.5	27,206.2	0.1	0.6	304,926.5	19,127.1	169.9	1.2	0.8	2.5

Table 3: Total toll revenues in the AM peak hours ($).

	Scenario 1		Scenario 2	
	Cars	Trucks	Cars	Trucks
Westbound	12,774.5	4,005.5	23,231.0	6,910.0
Eastbound	10,381.0	3,333.5	12,465.0	4,095.0
Total revenue	30,494.5		46,701.0	

4.2.2 Vehicle detours

This subsection examines the primary vehicle detour routes due to the tolling on Girard Point Bridge, in order to help public agencies understand how travelers' behavior changes. The accurate prediction of vehicle detours could help the general public understand the potential impacts of the tolling on the neighborhood community in terms of congestion, safety, and air quality.

As shown in Figure 7, there are two primary detour routes travelers can take to avoid the toll on Girard Point Bridge: George C. Platt Memorial Bridge and Passyunk Avenue Bridge.

Figure 7: Vehicle detours.

The traffic flows and the percentage changes in three scenarios are presented in Tables 4 and 5. It can be found that the introduction of a toll on Girard Point Bridge indeed induces travelers to take detours and thus reduce its own traffic flows. More specifically, a toll of \$1.00 in Scenario 1 can decrease the traffic flow in the baseline by almost 50% and a toll of \$2.00 in Scenario 2 can decrease the traffic flow in the baseline by 60%. In comparison, the traffic flows on both George C. Platt Memorial Bridge and Passyunk Avenue Bridge are increased significantly. Due to its closer distance to Girard Point Bridge, George C. Platt Memorial Bridge experiences more traffic growth than Passyunk Avenue Bridge. It can also be found that a higher toll leads to more traffic reduction on Girard Point Bridge and more traffic increase on the other two bridges.

Table 4: Traffic flows of three bridges in the AM peak hours.

		Baseline		Scenario 1		Scenario 2	
		Cars	Trucks	Cars	Trucks	Cars	Trucks
Girard Point Bridge	Westbound	25,113.5	7,444.5	12,774.5	4,005.5	11,615.5	3,455.0
	Eastbound	15,568.0	4,333.0	10,381.0	3,333.5	6,232.5	2,047.5
George C. Platt Memorial Bridge	Westbound	10,112.5	2,634.0	17,836.5	4,663.0	18,173.0	4,716.0
	Eastbound	10,053.5	2,791.0	12,493.5	3,129.5	15,947.0	4,149.5
Passyunk Avenue Bridge	Westbound	9,644.5	911.5	10,331.5	1,218.5	11,583.0	1,238.5
	Eastbound	8,678.5	2,787.5	9,961.5	3,369.5	11,085.5	3,683.0

Table 5: Percentage change of traffic flows of three bridges in the AM peak hours.

		Baseline		Scenario 1		Scenario 2	
		Cars	Trucks	Cars	Trucks	Cars	Trucks
Girard Point Bridge	Westbound	0.0%	0.0%	-49.1%	-46.2%	-53.7%	-53.6%
	Eastbound	0.0%	0.0%	-33.3%	-23.1%	-60.0%	-52.7%
George C. Platt Memorial Bridge	Westbound	0.0%	0.0%	76.4%	77.0%	79.7%	79.0%
	Eastbound	0.0%	0.0%	24.3%	12.1%	58.6%	48.7%
Passyunk Avenue Bridge	Westbound	0.0%	0.0%	7.1%	33.7%	20.1%	35.9%
	Eastbound	0.0%	0.0%	14.8%	20.9%	27.7%	32.1%

5 Conclusion

In this research project, we use multi-source data to create a large-scale multi-class network modeling and simulation framework, that holistically models the spatiotemporal behaviors of private cars, ride-hailing cars, and freight trucks, respectively. The dynamic network model for the DVRPC region is calibrated based on the collected count, speed, and vehicle registration data and is able to real-world traffic dynamics. The developed network model is then used to model and evaluate the traffic impacts of the PennDOT toll program. Three scenarios with different tolls on Girard Point Bridge are created. Results, including both the system-level metrics (e.g., travel time, travel delay, vehicle-mile-traveled, and emissions) and the local-level traffic detours for each of those vehicle classes, are presented and discussed. It is found that the introduction of a toll on the bridge can significantly reduce the traffic on the bridge leading towards downtown area, but meanwhile it increases travel time and emissions for the whole region and also result in a significant traffic flow increase on neighboring bridges due to traffic detours. It indeed generates considerable revenue for the public agencies to fund infrastructure in general. Further decisions on the bridge tolling need to be made by trading off social costs among different communities and the resultant funding for infrastructure.

References

[1] PennDOT P3 Office Bulletin. https://us17.campaign-archive.com/?u=2d9f31707d560d09bf89419da&id=a5517aa34f. Accessed: 2022-12-21.

[2] Sean Qian. Dynamic Network Analysis & Real-time Traffic Management for Philadelphia Metropolitan Area. Final Report WO-004, Carnegie Mellon University, Pittsburg, PA, September 2016.

[3] Wei Ma, Xidong Pi, and Zhen Qian. Estimating multi-class dynamic origin-destination demand through a forward-backward algorithm on computational graphs. *Transportation Research Part C: Emerging Technologies*, 102747, 2020. arXiv: 1903.04681.

[4] PennDOT suggests 9 bridges in the state add tolling, including one on I-79. https://www.wpxi.com/news/top-stories/penndot-suggests-9-bridges-state-add-tolling-including-one-our-area/CS7LROZXHVGMVBHWYZWBYXOV34/, February 2021. Accessed: 2022-12-19.

[5] Girard Point Bridge Improvement Project. https://www.penndot.pa.gov:443/RegionalOffices/district-6/ConstructionsProjectsAndRoadwork/Pages/I-95-Girard-Point.aspx. Accessed: 2022-12-19.