Travel speed and routes of the bike-share system: An analysis of Bike Share Toronto

Zehui Yin [redacted]

1. Background

Cycling, as one of the fastest-growing transportation modes worldwide, can help alter the long-established car-oriented urban spatial structure for society and promote physical activity for individual health. Bike-sharing, which is a special case of cycling that originated from a concept in the revolutionary 1960s, shares the benefits of cycling as a whole while creating a larger cycling population and mitigating the first-mile/last-mile problem of other traditional public transit modes (DeMaio, 2009). In the meantime, the Bike-sharing program owns and operates the bicycles transferring the theft or vandalism concerns from riders to the program, which offers riders more confidence in cycling. Currently, most of the Bike-sharing programs operate on a docked system and only collect "as-the-crow-flies" data, which is the straight line distance between origin and destination (OD) pairs, from the global positioning system (GPS) installed on bikes.

The publicly available data from a typical Bike-sharing program do not reveal the true routes, distances, and speeds of the bike trips. Efforts have been made by researchers to better understand these bike trip characteristics. However, there is no existing consensus on the average cycling speed of typical riders. In the existing literature, the cycling speeds are mostly either directly measured by a speed gun or inferred with high-frequency GPS data. A US study conducted by Thompson et al. (1997) found the mean travel speed of cyclists aged 14 and above is around 16 kilometres per hour (km/h). Other two European studies showed the average cyclist travelling speeds are 13.6 km/h (Dozza & Werneke, 2014) and 20.4 km/h (Gustafsson & Archer, 2013), respectively. The large differences in estimated mean travel speeds imply that external factors like the built-environment and climate in different locations might play an important role in determining cycling speed. Thus, case studies are needed to find reliable results for a specific location. Route choices have also received extensive attention from previous research. Studies suggest separate paths and lanes are preferred by cyclists over sharing roads with especially fast travelling motorized traffic (Buehler & Dill, 2016). A more cycling-friendly urban spatial structure with more bikeways could help encourage more drivers to make a switch from cars. Therefore, whether one place has sufficient cycling infrastructures like dedicated bike lanes can be investigated to determine whether more support or improvements are needed to better accommodate cyclists.

The Toronto city government introduced the local bike-sharing program named Bike Share Toronto back in 2011 aiming to complement the existing transit system while providing an alternative high-mobility option for the downtown core area (El-Assi et al., 2017). A previous project (Yin, 2022) found that there are seasonal spatial pattern variations in ridership, while stations located in the connecting area suffer the most from low ridership problems. Trip-level characteristics in the Bike Share Toronto system and whether the cycling infrastructures in the city provide sufficient support to it have not been extensively investigated.

2. Research Questions

In this project, I further examine the Bike Share Toronto system focusing on trips level characteristics to fill the research gap by answering the following two research questions:

RQ1. What are the travel speeds of cyclists in the Bikeshare Toronto System?

RQ2. To what extent do the existing bike lanes complement the Bikeshare Toronto System?

3. Data

My study area is the City of Toronto which is a highly urbanized city and the capital of Ontario, Canada. It has one of the best public transit systems in North America consisting of buses, streetcars, subways, wheel-trans, and commuter trains operated by the Toronto Transit Commission and Metrolinx. Bike Share Toronto operated by Toronto Parking Authority is the local bike-sharing program which currently runs 653 stations with a full capacity of 12411 bikes.

Table 1: Data sources

Name	Description	Format	Source (URL)	Access Time
Bikeshare ridership 2021	Toronto bike-share trip data in 2021	CSV	City of Toronto Open Data	2023-1-11
Bikeshare station data	Toronto bike-share station information in 2023	JSON	Toronto Parking Authority through GBFS	2023-1-11
OpenStreetMap data	OpenStreetMap data for Ontario, Canada	PBF	OpenStreetMap through Geofabrik	2023-1-11
Bikeways	The existing cycling network across Toronto	Shapefile	City of Toronto Open Data	2023-1-15

Table 1 above summarizes the data used in this project. Bike-share ridership data contain all the trip information in 2021 including importation attributes origin and destination station names as well as trip durations. Bike-share station data provide current station names and locations. OpenStreetMap data provide current street network information for the whole province of Ontario. Meanwhile, bikeways data contains spatial information of the existing bike lanes in Toronto.

4. Methods

The data first go through a data cleaning process. Since the Bike-share ridership data and station data are not collected in the same year, some current stations did not operate in 2021, while some stations in the past are not in operation currently. Therefore, the trips that start or end at a currently not-in-operation station are filtered out of analysis due to no spatial data. Meanwhile, trips that start and end at the same station are also excluded from the analysis due to impossible route inferences. Then all the trips in 2021 are classified into four seasons: spring (March, April, May), summer (June, July, August), autumn (September, October, November), and winter (December, January, February), based on the standard season interpretation.

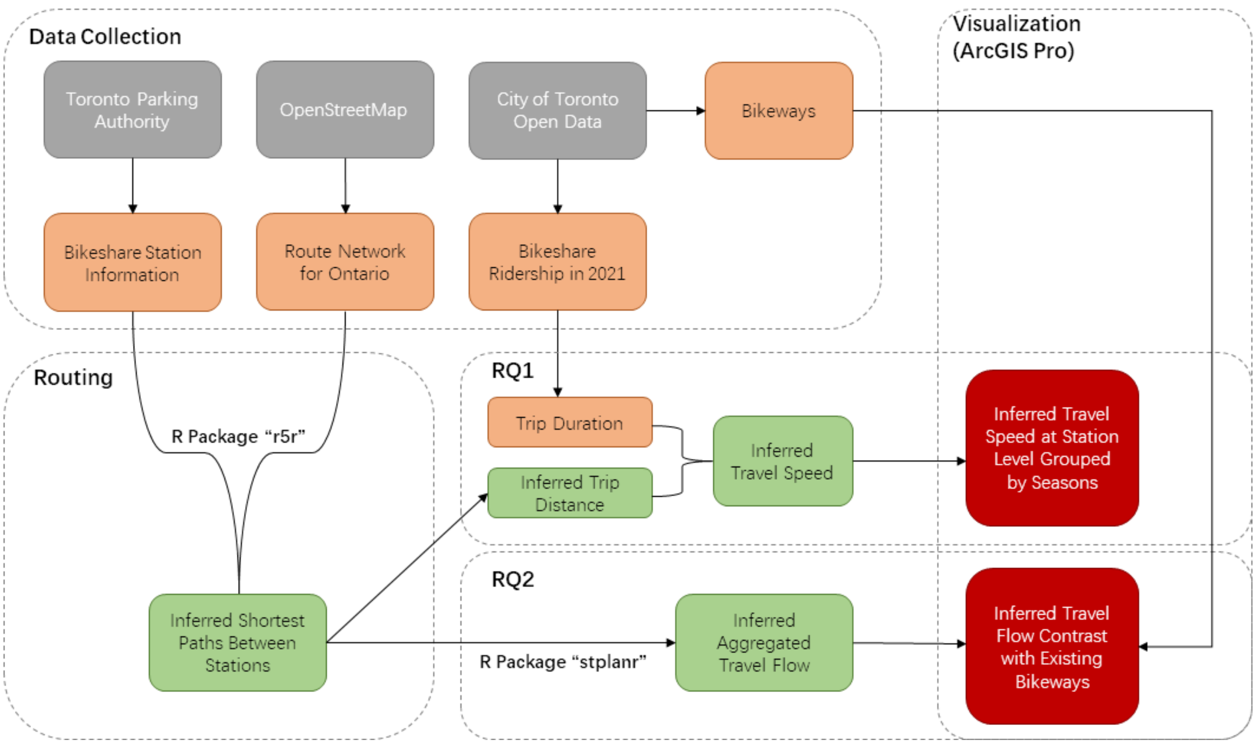

Figure 1: Workflow diagram

4.1 Routing

To address both RQ1 and RQ2, the shortest cycling paths between all the possible combinations of stations are inferred based on Bike-share station data and the OpenStreetMap road network. The routing is computed in R with the package "r5r" with the maximum level of traffic stress setting to 2 which represents the routes that are tolerable for the mainstream adult population (Pereira et al., 2021).

4.2 Travel speed inference

To address RQ1, the average travel speed originating at a specific station in a particular season is estimated by the inferred route distance from the routing process and recorded trip durations directly from the Bike-share ridership dataset. For each station and season, the mean travel speed is calculated with the total estimated travel distance divided by the total trip duration originating from that station and season or by Equation 1:

$$S_{iw} = \frac{\sum_{j=1}^{M-1} n_{ijw} * d_{ij}}{\sum_{z=1}^{N_{iw}} t_{iz}} \quad (1)$$

Where:

S_{iw} = the mean travel speed of trips originating from i^{th} station in w^{th} season (m/s)

M = the number of total stations

N_{iw} = the total number of trips originating from i^{th} station in w^{th} season

n_{ijw} = the number of trips originating from i^{th} station to j^{th} station in w^{th} season

d_{ij} = distance travelled from i^{th} station to j^{th} station

t_{ij} = the time it took by z^{th} trip to travel from i^{th} station

To quantify the seasonal differences between estimated mean travel speed at the station level, a relative variation metric, which accounts for station travel speed differences, is calculated for each station with Equation 2:

$$RV_i = \frac{\sqrt{\frac{\sum_{w=1}^{4}(S_{iw}-\overline{S_i})^2}{4}}}{\overline{S_i}} \quad (2)$$

Where:

RV_i = the relative variation of mean travel speed of trips originating from i^{th} station

$\overline{S_i}$ = the mean travel speed of trips originating from i^{th} station in all seasons

4.3 Travel flow inference

The total travel flow within Bike Share Toronto system for the whole year 2021 is generated from the inferred routes and trip counts from the Bike-share ridership dataset to answer RQ2. All the trips' inferred routes are aggregated based on the overlapping road segments and the total number of trips that went through that particular road segment is recorded in R with package "stplanr" (Lovelace & Ellison, 2018). Then, the existing bikeways are visualized together with the inferred travel flows to investigate whether they complement the Bike Share Toronto system. Also, to quantify the spatial patterns of the travel flow, some spatial descriptive statistics, such as mean centre, median centre, standard distance, and standard deviational ellipse, as well as local and global Moran's I based on eight nearest neighbours are computed and visualized.

5. Results

In this section, I first present the descriptive statistics of the Bike-share ridership, mean travel speeds at stations, inferred cycling paths, and travel flows within the system. Then I further examine and interpret the results with map visualizations.

5.1 Descriptive Statistics

Bike-share ridership 2021 dataset contains 3575182 trips. However, only 2734053 trips have origins and destinations that are all operational currently. Due to data availability challenges, the station's spatial information is only available for operating stations. The analysis is conducted on this subset of all the trips in 2021. Most of the trips in 2021 are undertaken during summer and autumn, while the median trip's durations are relatively longer during spring and summer. The relative variation for all stations is around 0.15, which means that on average stations' seasonal travelling speeds have a standard deviation of 15% of their whole year mean travelling speeds.

A total number of 577 currently operational stations have been used during autumn and winter, while only 542 of them have been utilized during spring and summer in 2021. Considering 577 total stations, there are 332352 (577×576) possible combinations of origin and destination (OD) pairs excluding the scenario that both the origin and destination are the same station. Among all the possible OD combinations, 152805 of them are truly undertaken by the riders in 2021. Meanwhile, based on inference, 10072 distinct road segments have been used by the riders in 2021.

Table 2: Descriptive statistics and inference results

	N	*Mean*	*Median*	*IQR*
	Duration (s)			
Total Bike-share Ridership	2734053	971.909	760	724
Spring Trips	592722	1111.725	878.5	787
Summer Trips	1065757	1021.28	811	738
Autumn Trips	855635	867.9523	673	646
Winter Trips	219939	760.2994	609	588
	Estimated Mean Travel Speed (km/h)			
All Stations	577	10.4144	10.64293	2.312048
Spring Stations	542	9.481036	9.697407	2.398202
Summer Stations	542	10.24636	10.5607	2.52391
Autumn Stations	577	11.26381	11.62967	2.421905
Winter Stations	577	11.92385	12.03691	2.068494
Relative Seasonal Variation	542	0.1459059	0.1184173	0.1030178
	Inferred Distance (m)			
Cycling Routes (all possible)	332352	8207.216	6815	6262
Cycling Routes (actually used)	152805	4330.338	4028	3095
	The number of trips went through that inferred road segment			
Travel Flows (road segments)	10072	8004.472	1879	6905.75

5.2 Estimated average travel speed at station level

Based on Table 2, surprisingly, the median estimated mean travel speeds at the station level keep increasing from less than 10 km/h in spring to slightly above 12 km/h in winter. In the meantime, winter's estimated station level mean travel speed dispersion which is measured by the interquartile range (IQR) is considerably smaller compared to all the other three seasons.

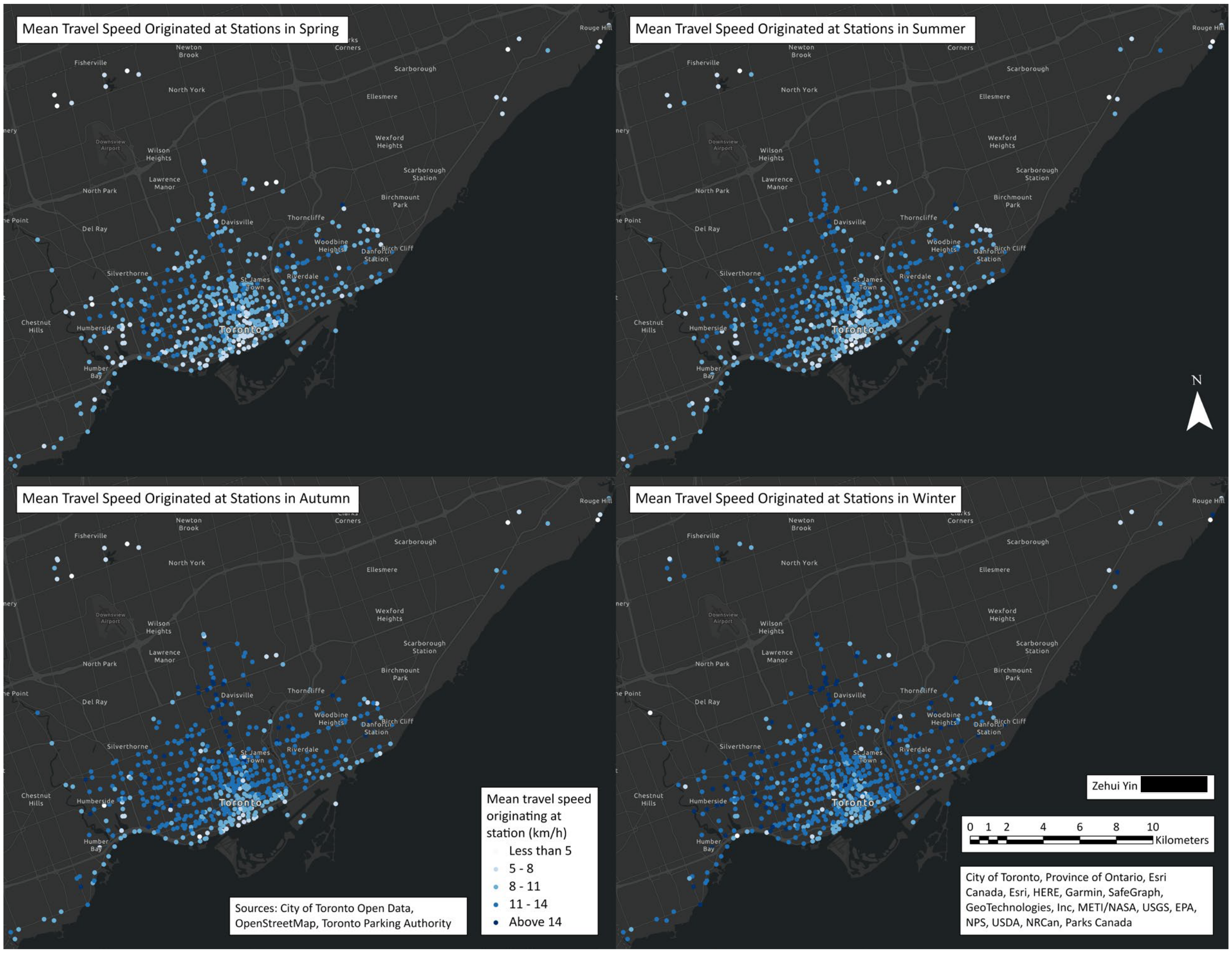

Figure 2. Estimated seasonal mean travel speed at station level

The trend observed in descriptive statistics is confirmed by the maps shown in Figure 2 below. The overall speed is relatively higher in autumn and winter compared to the counterpart in spring and summer. Also, the stations located in the downtown area and at the edges of the whole bike-sharing system have significantly lower average speeds compared to other stations. This spatial disparity in average travelling speed is larger in spring and summer than other two seasons.

Figure 3 below visualizes the magnitude of seasonal variations in estimated average travelling speed. The stations located at the downtown core area and the edges of the whole system have the largest variation in travelling speed, while the remaining station situated in between have relatively stable travelling speeds throughout the year.

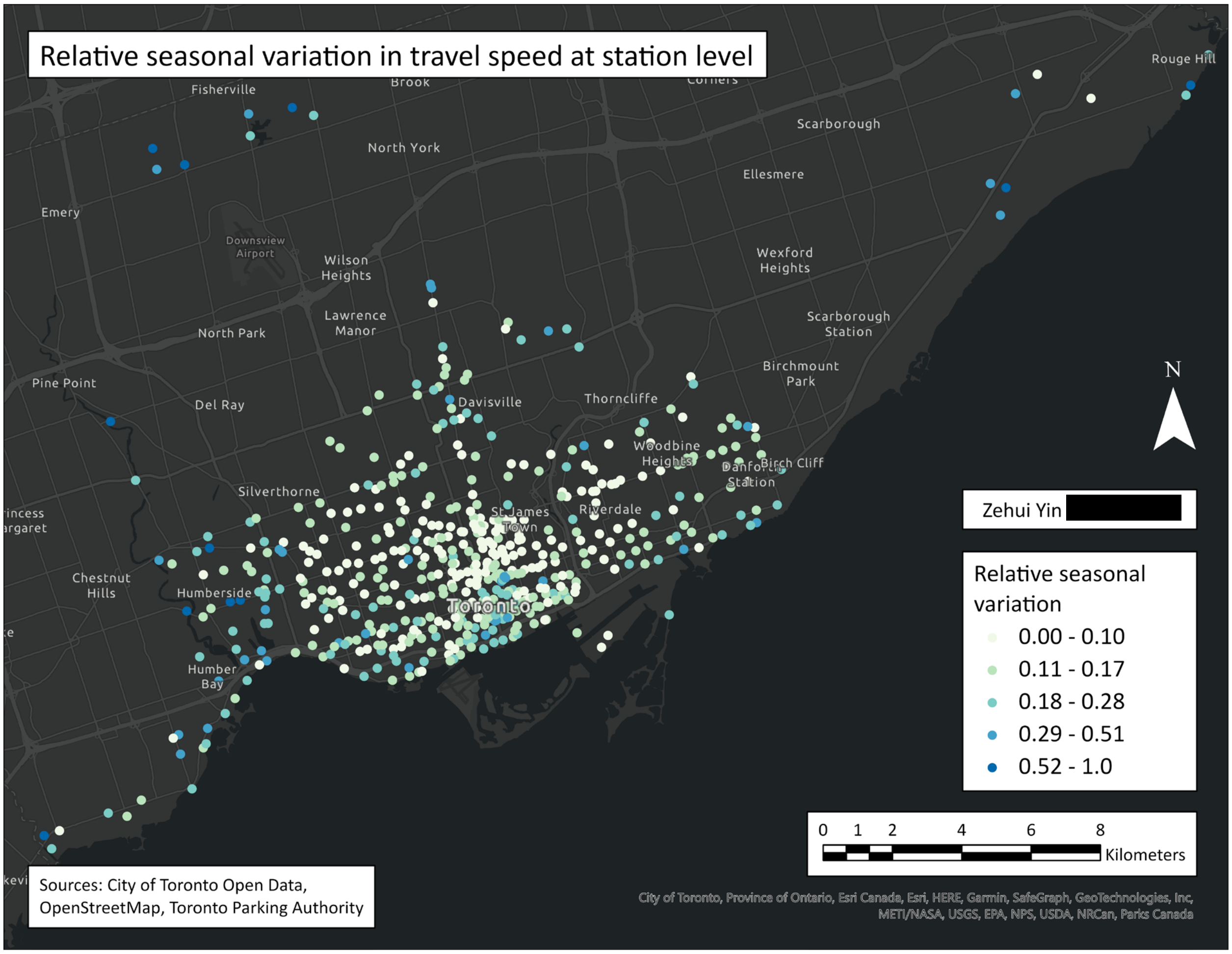

Figure 3. Relative variation in estimated seasonal mean travel speed at station level

5.3 Estimated routes used by the riders

The routes in the downtown core area enjoy the highest usage. The usage decline quickly as the roads locate farther from the centre of Toronto. Some estimated routes only had 1 trip in the whole year, while the highest route was used by 236145 riders. Interestingly, the route along Lake Ontario, which is not located in the downtown core area, had been used extensively.

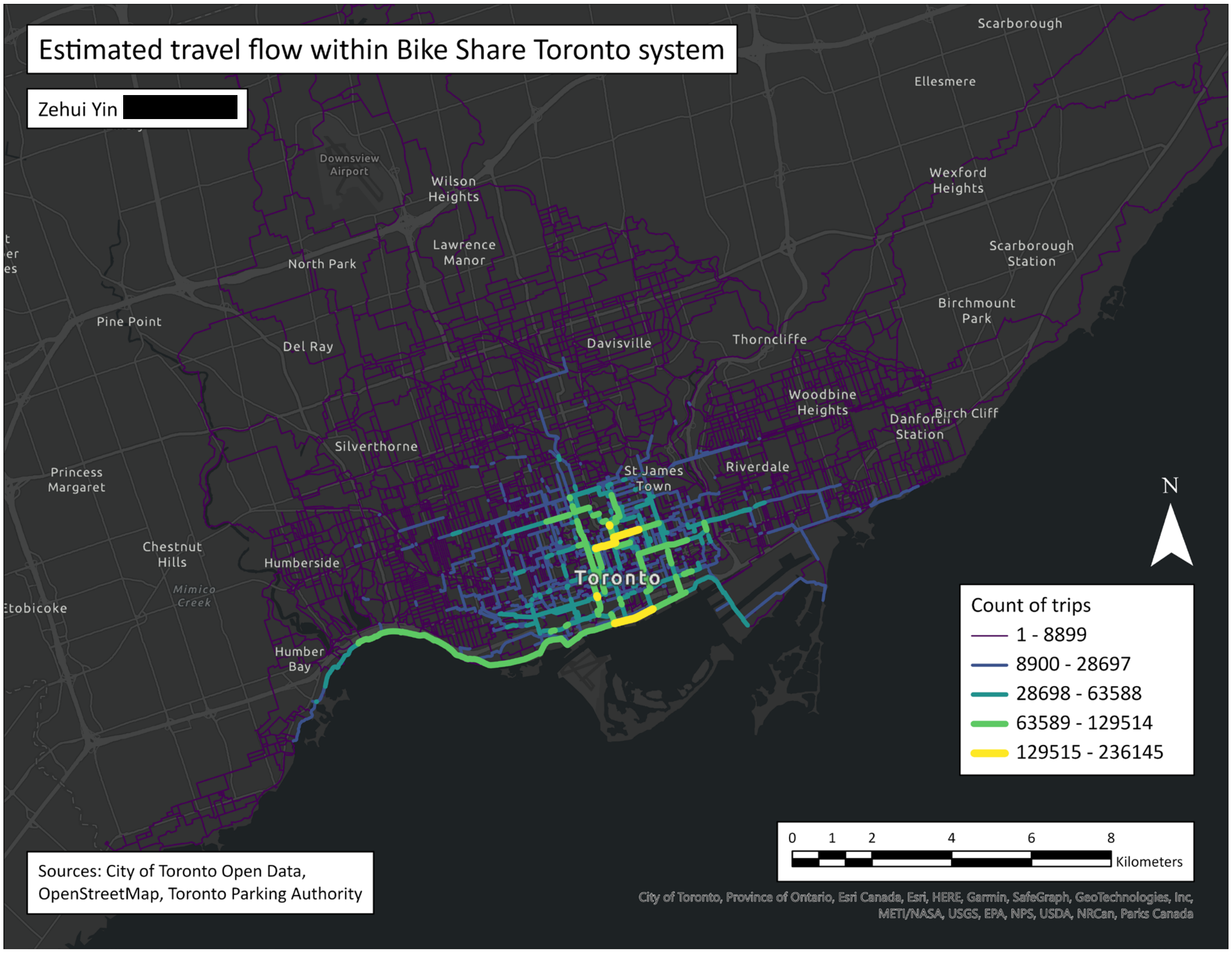

Figure 4. Estimated travel flow within the Bike Share Toronto system

The spatial descriptive statistics and local Moran's I results visualized in Figure 5 confirm that the centre of the estimated travel flow distribution is at the downtown core area of Toronto. Meanwhile, the variation of the distribution is not very large, as one standard deviation circle and ellipse can only cover a small area of the city centre. The travel flow distribution also extends more in the Southwest and Northeast directions based on the ellipse shape. The local Moran' I indicates that the high-value clusters are mostly located within the one standard deviation around the centre, while roads outside the circle and ellipse are mostly low-value clusters. The global Moran's I for the travel flow is 0.457 with an almost 0 p-value, indicating that there is a statistically significant clustering pattern in the estimated travel flow in the bike-sharing system.

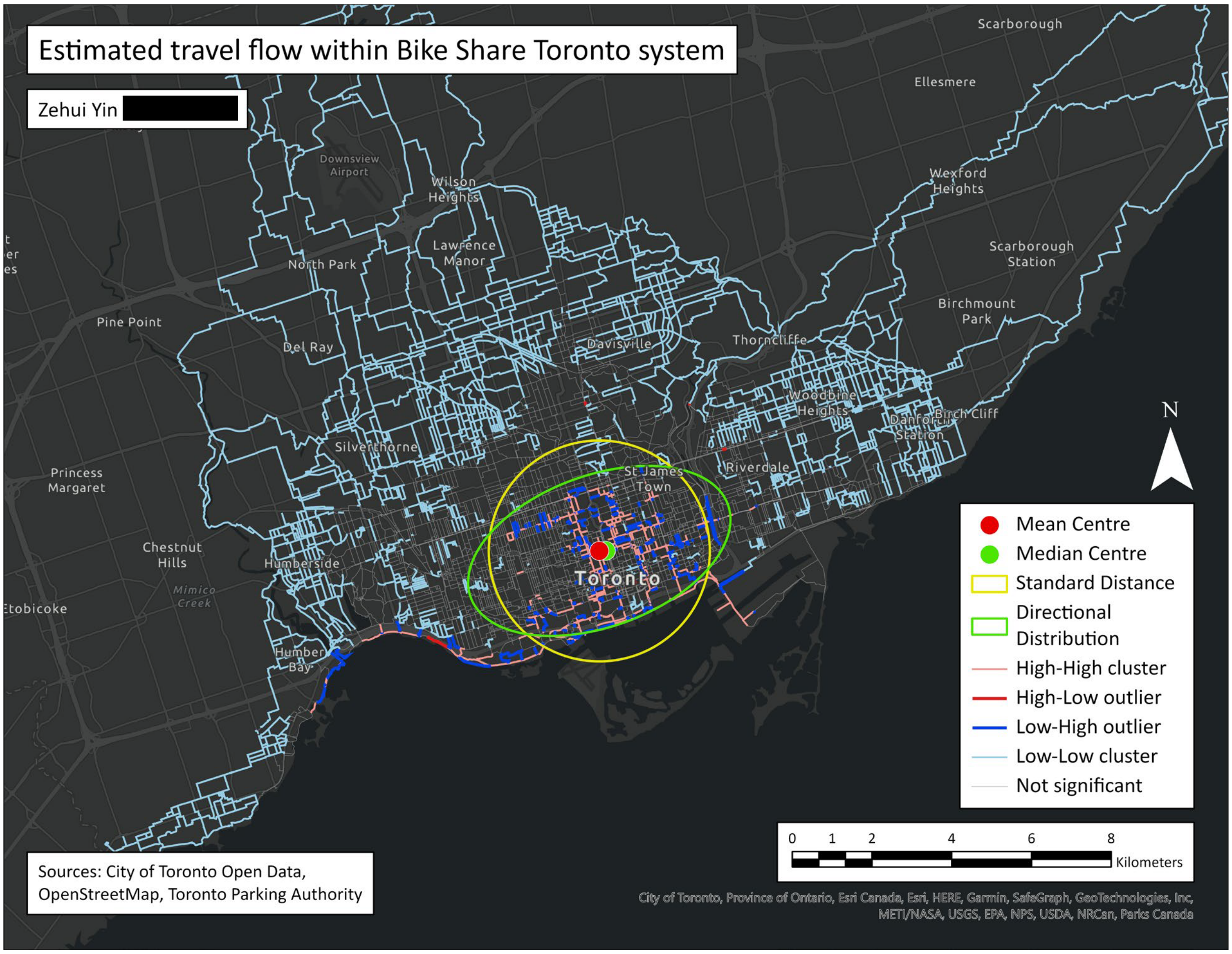

Figure 5: Spatial descriptive statistics and local Moran's I

Figure 6 visualized the estimated travel flow within the Bike Share Toronto system along with the existing bike lanes in the city. The bike lanes do not align well with the estimated travel flows. There are long-length bikeways built at the edges of the bike-sharing system where little usage is estimated. For the routes estimated to have the highest demand in the downtown area, only a small portion of them are lucky enough to have an existing bike lane. In the meantime, the special route along Lake Ontario which enjoys high demand doesn't have bike lanes to accommodate cyclists.

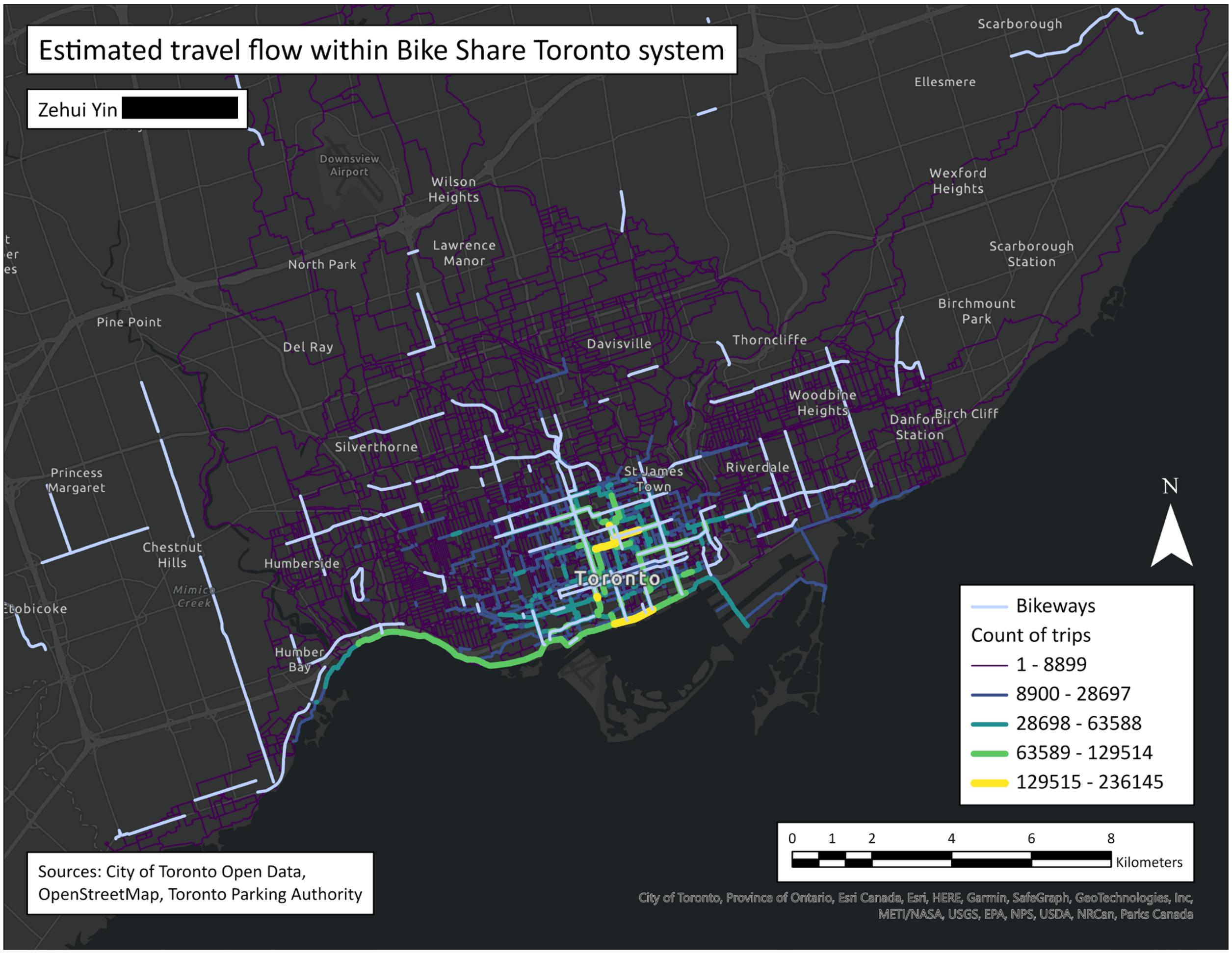

Figure 6. Estimated travel flow within the Bike Share Toronto system and existing bikeways in Toronto

6. Discussion

The counter-intuitive trend observed in the estimated seasonal average travel speed indicates that riders were travelling faster in the winter when the road conditions are bad compared to spring. However, this result is only an estimation based on the inferred travel routes from OD pairs, which means an assumption is made that riders are taking the shortest path within the system. Therefore, such a trend could be interpreted alternatively or more realistically as riders in spring and summer are less likely to take the shortest paths compared to autumn and winter.

Riders when undertaking recreational or leisure purposes trips are less sensitive or considerable about whether they are taking the quickest route, as most of the time they do not have a precise destination in their minds. In the contrast, commuters who are using bike-sharing services always have a clear destination and hope to minimize their commute times. Thus, the shortest cycling path assumption is more likely to hold when commuters or non-recreational riders are dominating or taking up most of the users of the bike-sharing services.

Therefore, there are higher proportions of recreational riders in the spring and summer, while commuters or non-recreational riders dominate the autumn and winter usage of the Bike Share Toronto system. A station that is dominated by non-recreational riders for the whole year would have a small variation in estimated seasonal travelling speed. Based on the relative variation of seasonal average travelling speed, the station located at the downtown core area and edges of the system enjoys more recreational riders in the spring and summer.

To address RQ1, the inferred median travelling speed for the winter 12 km/h would be the most reliable estimation for the true travelling speed of the riders in the system, as the shortest path assumption is most likely to hold and not prone to outlier station influences. Considering the poor road conditions in the winter, it is reasonable to argue that the average travelling speed in the whole year is at least above 12 km/h.

When it comes to RQ2, the existing bike lanes do not complement the Bikeshare Toronto System to a large extent. Only some of the heavily used routes have bike lane coverage, while a great proportion of the existing bike lanes locate on seldom-used roads. Also, the existing bikeways have relatively better coverage for northwest to southeast roads. More bikeways, especially for southwest to northeast directions, should be built in the downtown core area and along Lake Ontario to better support Bike Share Toronto riders.

There are some limitations in this project, which mostly come from data challenges. The mismatch between 2021 ridership data and 2023 station data causes almost one-third of the total trips in 2021 cannot to link their origin or destination to a spatial location, which makes route estimation impossible for them. Thus, the analysis is done on only around two-thirds of all trips in 2021 which could introduce biases in the travel speed and travel flow estimations. Meanwhile, due to computation constraints, the travel speed estimation is simplified to calculate mean travel speeds at the station level instead of computing median travel speeds for each trip record. Using mean values to estimate the centre of a distribution is relatively more prone to outliers. That could be part of the reason why the estimated mean travel speeds are relatively low for the spring and summer. Future studies are needed with better-quality station data and higher computing capacity to better understand the travel behaviour within the Bike Share Toronto system.

7. Conclusion

This project provides insights into the travel behaviours of the riders using Bike Share Toronto. I found that the estimated average travelling speed within the Bike Share Toronto system is at least 12 km/h. The stations located in the downtown core area and the edges of the system have relatively higher proportions of recreational riders in spring and summer compared to other stations. Other stations situated in between have small variations in estimated mean travelling speeds which indicate that commuters and non-recreational riders dominate the usage of these stations.

The existing bike lanes in the city do not complement the Bike Share Toronto system to a high extent. Even the routes estimated to have the highest usage do not have full coverage from bikeways let alone other routes. There are high portions of the existing bikeways located at the edges of the bike-sharing system where little demand is estimated. While building bike lanes, more importance should be attached to demands from the Bike Share Toronto system. Since these demands are relatively easier to predict and estimate than other cyclists, the planning can be done with less uncertainty and public funding can be used more effectively. More southwest-to-northeast bikeways should be built in the downtown area and along Lake Ontario to provide better infrastructure for Bike Share Toronto.

The government could provide better services and support for bike-sharing riders with the insights from this study. Commuters might be more price sensitive than recreational riders. Station-based price discounts can be offered to the stations located between the downtown core area and the edges where little seasonal variation in estimated travelling speed is observed. Also, building bike lanes in the downtown area and along Lake Ontario could attract more people to use the bike-sharing services and create a larger cycling population. The estimated travelling speed within the bike-sharing system is somewhere moderately above 12 km/h. Speed heavily influences the competitiveness of a transportation mode. We can explore options such as dedicated cycling traffic lights to reduce the waiting time for cyclists at street intersections in order to increase cycling speed and mode competitiveness.

References

Buehler, R., & Dill, J. (2016). Bikeway networks: A review of effects on cycling. *Transport Reviews*, *36*(1), 9-27. https://doi.org/10.1080/01441647.2015.1069908

DeMaio, P. (2009). Bike-sharing: History, impacts, models of provision, and future. *Journal of public transportation*, *12*(4), 3. http://doi.org/10.5038/2375-0901.12.4.3

Dozza, M., & Werneke, J. (2014). Introducing naturalistic cycling data: What factors influence bicyclists' safety in the real world? *Transportation research part F: traffic psychology and behaviour*, *24*, 83-91. https://doi.org/10.1016/j.trf.2014.04.001

El-Assi, W., Salah Mahmoud, M., & Nurul Habib, K. (2017). Effects of built environment and weather on bike sharing demand: a station level analysis of commercial bike sharing in Toronto. *Transportation*, *44*(3), 589-613. https://doi.org/10.1007/s11116-015-9669-z

Gustafsson, L., & Archer, J. (2013). A naturalistic study of commuter cyclists in the greater Stockholm area. *Accident Analysis & Prevention*, *58*, 286-298. https://doi.org/10.1016/j.aap.2012.06.004

Lovelace, R., & Ellison, R. (2018). stplanr: A package for transport planning. *The R Journal*, *10*(2), 7-23. https://doi.org/10.32614/RJ-2018-053

OpenStreetMap contributors. (2023). *Ontario* [digital resource: PBF]. Geofabrik. http://download.geofabrik.de/north-america/canada/ontario.html

Pereira, R. H., Saraiva, M., Herszenhut, D., Braga, C. K. V., & Conway, M. W. (2021). r5r: rapid realistic routing on multimodal transport networks with r 5 in r. *Findings*, 21262.

https://doi.org/10.32866/001c.21262

Thompson, D. C., Rebolledo, V., Thompson, R. S., Kaufman, A., & Rivara, F. P. (1997). Bike speed measurements in a recreational population: validity of self reported speed. *Injury Prevention*, *3*(1), 43-45. http://dx.doi.org/10.1136/ip.3.1.43

Toronto Parking Authority. (2021). *Bike Share Toronto Ridership Data* [Data set]. City of Toronto Open Data. https://open.toronto.ca/dataset/bike-share-toronto-ridership-data/

Toronto Parking Authority. (2023). *Bike Share Toronto* [GBFS]. City of Toronto Open Data. https://open.toronto.ca/dataset/bike-share-toronto/

Toronto Transportation Services. (2023). *Bikeways* [digital resource: vector]. City of Toronto Open Data. https://open.toronto.ca/dataset/bikeways/

Yin, Z. (2022). *Spatial patterns of seasonal ridership, built environment and social demographics: A taxonomy of Bike Share Toronto stations*. Zehui Yin's Personal Website. https://zehuiyin.github.io/Spatial-patterns-of-seasonal-ridership,-built-enviroument-and-social-demographics-A-taxonomy-of-Bike-Share-Toronto-stations/