9th Symposium of the European Association for Research in Transportation

September 1-3, 2020, Lyon, France

Hotspot Identification, Ranking and Impact Estimation of Illegal Parking Using Spatial Association and Queueing Model

Jingqin Gao[a], Kaan Ozbay[b]

[a] Department of Civil and Urban Engineering, New York University
[b] Department of Civil and Urban Engineering &Center for Urban Science and Progress, New York University

---

Keywords: Illegal parking, Double parking, Hotspot Identification, Queueing models, Traffic impact

1. Introduction and Background

Urban cities have long experienced traffic congestion as well as high demand for curbside parking. Recently, there is on-street parking competition from a variety of usages, including transit, bicycles, delivery trucks due to e-commerce, and shared mobility services. The increasing demand and limited supply for on-street parking have consequences such as illegal parking, for example, double parking. Every year, over 10 million parking violation tickets are issued in New York City (NYC) alone. Anyone who has driven in a large city knows the frustration of encountering a street blocked by illegal-parked vehicles, especially double-parked vehicles. More importantly, double-parked trucks can worsen the problem by urging mandatory lane merging to bypass blockages and making many streets impassable. Unfortunately, this type of road impediments due to illegal parking routinely happens on the busiest urban streets that can hardly afford to lose any capacity. The ability to estimate the impacts of illegal parking, particularly double parking, is crucial for daily traffic operation and management.

Various studies used microscopic simulation models *(1-4)* to quantify the impact of illegal parking. However, the use of a specific microscopic simulation tool might be limited at the city level and not always be feasible due to time and budget constraints. Thus, there is a need to identify a reliable analytical methodology that has lower complexity and higher computational speed compared with microscopic simulation models so that a large number of sites can be studied efficiently. On the other hand, this analytical methodology needs to be robust and capable to quantify traffic impacts under different frequency and severity of the illegal parking conditions.

The objective of this paper is to explore the potential of using big data and macroscopic queueing models together in advancing the on-street illegal parking analysis, including hotspot identification and quantification of the illegal parking impacts with a focus on double parking. As the first step of the identify, rank and quantify impacts methodology proposed in this paper, NYC parking citation data was geocoded by a custom-developed software to capture local patterns of spatial association. A local spatial autocorrelation analysis based on the Local Indicators of Spatial Association (LISA) *(5)* statistics was applied to identify double parking hotspots. This type of spatial cluster map can identify promising locations for efficient enforcement in resolving conflicts over demand and supply of parking.

The second step is to rank the hotspots with highest citations. Neighborhood tabulation area (NTA) was used as the analysis unit. The highest ranked NTA was then targeted for field data collection. Next, as the third step in the proposed methodology, we applied an M/M/∞ queueing model to estimate average link travel time that incorporates interruptions due to illegal on-street parking. The proposed model was tested and empirically validated using field data collected from three study sites chosen from the highest ranked hotspot – Midtown Manhattan, NYC. Different levels of illegal parking conditions were investigated (high, moderate, low frequency). The proposed queuing model assumes multiple road segments to be considered as individual servers. It describes a system that has an infinite number of servers and its arrival follows a Poisson process with a service time that follows an exponential distribution. Compare with microscopic simulation models, queueing model is less expensive and easier to develop and implement.

This study can provide traffic agencies a potential approach to quantify the impact of on-street illegal parking violations in a large-scale network, therefore better enforcement, pricing strategies, and policies can be gained for effective parking management.

2. Data Description and Study Sites

Parking violation tickets in 2017 in NYC were obtained from NYC OpenData portal *(6)* and were projected onto a map to identify violation "hotspots". As this study's focus is mainly on double parking activities, here we used double-parking as representative instead of all illegal parking citations. In 2017, double parking violation had 636,946 ticketed summons and was ranked seventh among all the parking violations in New York City *(6)*. It should be noted that the total number of violations may be underestimated as most of such activities may not even be recorded or ticketed.

After identifying the hotspots, three case study sites, with different levels of illegal parking frequency, were chosen specifically (Table 1).

Table 1 NYC Study Sites

Name	Location	Illegal Parking Frequency	# Illegal Parking Events	Video Screenshots
Site A	6th Avenue from 49th Street to 50th Street	High	74 events/hour	
Site B	5th Avenue from 41st Street to 42nd Street	Moderate	29 events/hour	
Site C	5th Avenue from 49th Street to 50th Street	Low	9 events/hour	

Videos were recorded for one day in May 2017 during AM (8-9AM) peak hour for all three sites. Traffic and illegal parking information, including traffic flow, travel time, observed queue, driver behaviors and illegal parking violation type, arrival and departure time and vehicle type, were obtained from the videos. The information was collected and summarized for every 5 minutes. Table 2 shows an example of the traffic information collected for Site A. Observation from the illegal parking violations reveal the following behaviors: 1) About 50-70% of the illegal parking activities at the three study sites were contributed by double-parked vehicles; 2) the average duration of the parking violation is relatively short (about 20 seconds to 1 minute); 3) not all illegal parking events caused a great impact on traffic. Street and traffic characteristics as well as network topology (e.g., number of lanes or presence of bus/bike lane) had an effect as well; and 4) double-parked trucks have relatively larger impacts compared with other type of violations and type of vehicles.

Table 2 Traffic Information for Site A

Time	Volume	HV%	# Illegal parking	Status	Avg Travel Time(s)
8:00-8:05 AM	119	13.4%	5	IP	24.29
				w/o IP	7.65
8:05-8:10 AM	129	14.7%	8	IP	33.60
				w/o IP	8.38
8:10-8:15 AM	92	18.5%	10	IP	23.00
				w/o IP	8.27
8:15-8:20 AM	98	13.3%	5	IP	21.86
				w/o IP	7.11
8:20-8:25 AM	122	10.6%	7	IP	26.55
				w/o IP	10.07
8:25-8:30 AM	108	11.1%	4	IP	31.20
				w/o IP	9.66
8:30-8:35 AM	111	18.0%	4	IP	31.82
				w/o IP	13.90
8:35-8:40 AM	106	11.3%	4	IP	29.17
				w/o IP	10.05
8:40-8:45 AM	95	22.1%	10	IP	24.10
				w/o IP	10.61
8:45-8:50 AM	114	16.7%	6	IP	29.5
				w/o IP	10.10
8:50-8:55 AM	114	15.8%	6	IP	29.3
				w/o IP	12.15
8:55-9:00 AM	80	13.8%	5	IP	26.75
				w/o IP	25.81

*HV% = Heavy vehicle percentage, IP = illegal parking, w/o IP = without illegal parking

3. Proposed Framework

As mentioned in the introduction section, this study proposes an identify, rank and quantify impacts methodology. The proposed framework is illustrated in Figure 1.

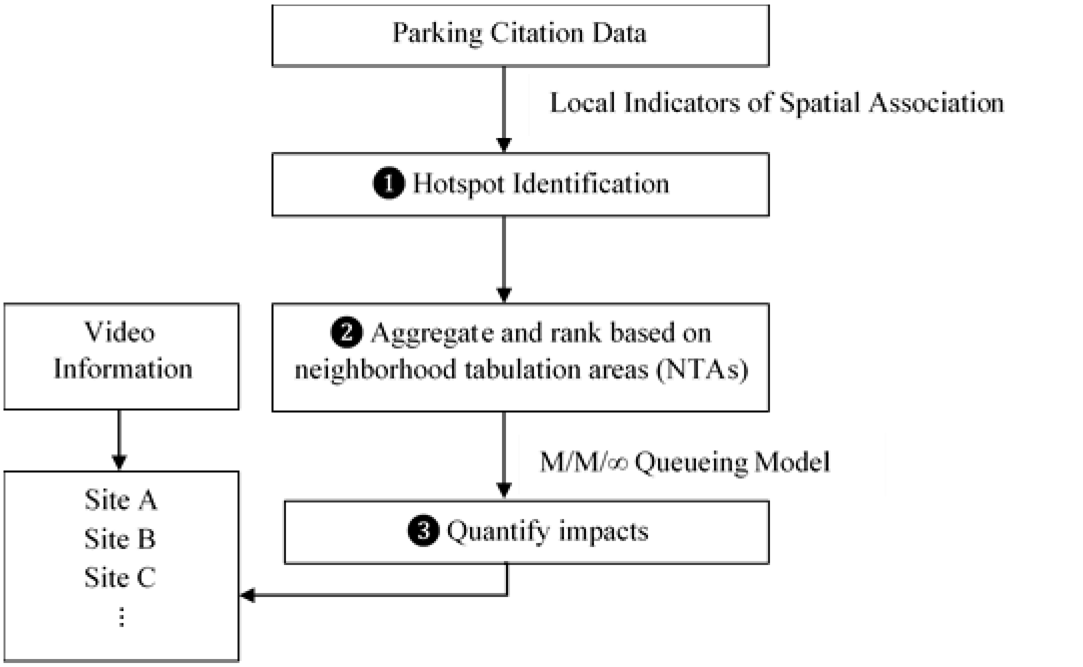


Figure 1 Identify, Rank and Quantify Impacts Methodology

4. Spatial Cluster Parking Maps

The original address information from the citation data is in plain text format. To convert them into geographic coordinates, a batch geocoding program was developed based on U.S. Census Bureau Geocoding Services Web Application Programming Interface (API) *(7)* and Google Geocoding API *(8)*. Next, the map of Manhattan is uniformly split into 6,204 equally sized grid cells (300 feet × 300 feet). Using equally sized grid cells as the basic geographical units eased the bias from different analysis unit sizes and were proved to be effective in previous studies *(9-11)*.

Global Moran's I statistic proposed by Moran *(12)* is the most commonly used indicator of global spatial autocorrelation. It is a "cross-product statistic between a variable and its spatial lag, with the variable expressed in deviations from its mean" *(13)*. Global Moran's I statistic can be computed as follows:

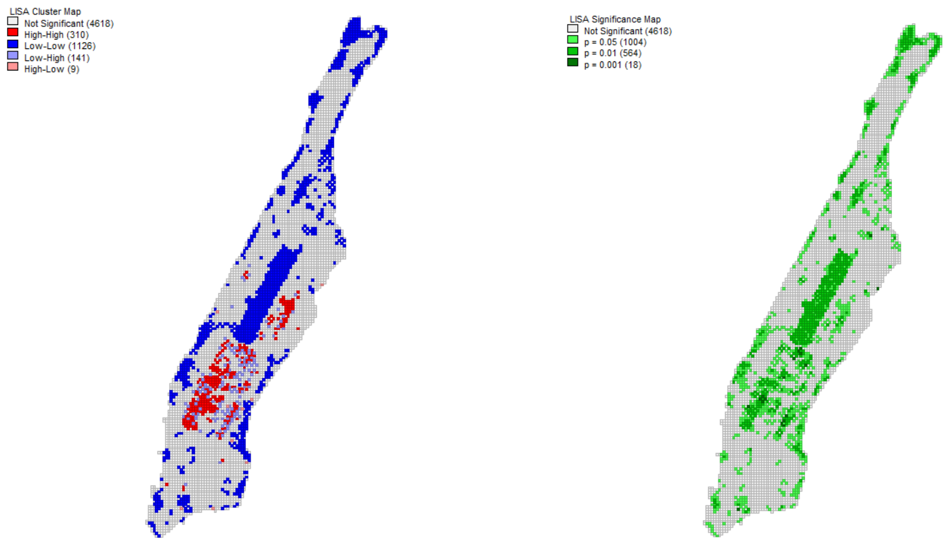
$$I = \frac{\sum_i \sum_j w_{ij} z_i z_j / S_0}{\sum_i z_i^2 / N}$$

Where z_i, z_j are the observation in location i, j, w_{ij} is the weights matrix, S_0 is the sum of all the weights, and N is the number of observations.

LISA – a Local Moran's I test is derived from the assumption that global Moran's I is a summation of individual cross-products and computes a measure of spatial association for every observation *(5)*. Local Moran's I statistics for each location i *(5)* can be computed as follows:

$$I_i = \frac{z_i}{\left(\frac{\sum_i z_i^2}{N}\right)} \sum_j w_{ij} z_j$$

Significance information (LISA significance map Figure 2(b)) combined with the location of each observation in the Moran Scatterplot allows for a classification of the significant locations (LISA cluster map (Figure 2 (a)) *(13)*. These significant locations include high-high and low-low spatial clusters as well as high-low and low-high spatial outliers. The high-high and low-low locations suggest a clustering of similar high or low local Moran's I values. High–high clusters, like Midtown Manhattan and the Upper East Side, were regarded as hotspots for double parking.



(a) LISA Cluster Map (b) LISA Significance Map

Figure 2 LISA Cluster and Significance Map

5. Hotpots Ranking

The high-high relationship cells identified by LISA were aggregated into NTAs and ranked by the number of total citations. Table 4 presents the top 10 NTAs in Manhattan, NYC. Since Midtown-Midtown South neighborhood has the highest rank, we choose three sites from this neighborhood for the next step to quantify illegal parking impacts.

Table 3 Hotspot Ranked by Highest Citations

Rank	NTA Code	NTA Name
1	MN17	Midtown-Midtown South
2	MN40	Upper East Side-Carnegie Hill
3	MN12	Upper West Side
4	MN13	Hudson Yards-Chelsea-Flatiron-Union Square
5	MN32	Yorkville
6	MN20	Murray Hill-Kips Bay
7	MN24	SoHo-TriBeCa-Civic Center-Little Italy
8	MN15	Clinton
9	MN31	Lenox Hill-Roosevelt Island
10	MN21	Gramercy

6. M/M/∞ Queueing Model

We consider an M/M/∞ queueing model with an arrival rate μ governed by Poisson process with an infinite number of servers and experiences random disruptions of exponentially distributed durations. The traffic disruption due to illegal parking is assumed to arrive according to a Poisson process with rate *f*, and the repair time is exponentially distributed with rate r. The stochastic process *{X(t), U(t)}* describes the state of the system at time t, where *X(t)* is the number of vehicles in the system at *t*, and *U(t)* is the status of the system. When illegal parking occurs, the system subjects to either partial failures where service rate μ decreases to μ' or complete system breakdown where μ decrease to μ'=0 (*U(t)* = "Failure (F)"). The system resumes to normal state when the traffic disruption due to illegal parking is cleared (service rate back to μ) (*U(t)* = "Normal (N)"). The system is in state (i, F) if there are i vehicles in the system that are interrupted by illegal parking. Similarly, if the system has i vehicles without any illegal parking disruptions, it is assumed to be in state (i, N). Transition-rate diagram for M/M/∞ Model with two server states are shown in Figure 4.

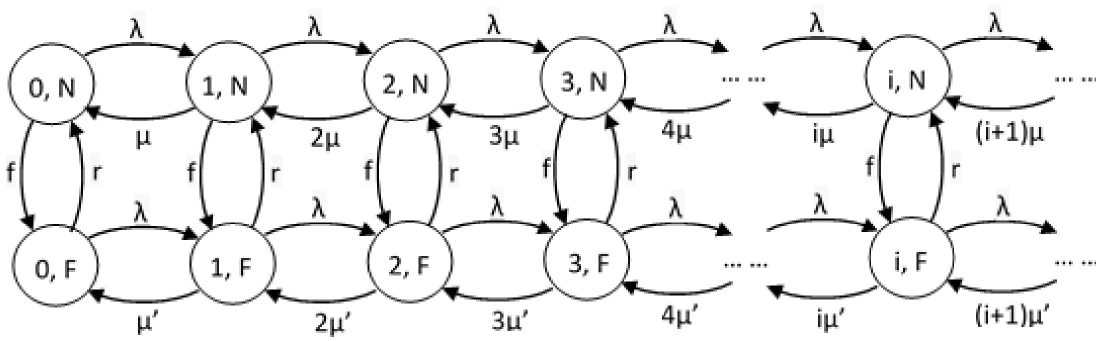


Figure 3 Transition-rate Diagram for M/M/∞ Model with Two Server States

From the transition-rate diagram, it is easy to derive the steady-state balance equations for the state probabilities $P_{i,F}$ and $P_{i,N}$ yielding:

$$(\lambda + i\mu' + r)P_{(i,F)} = (i+1)u' P_{i+1,F} + \lambda P_{i-1,F} + fP_{i,N} \qquad (i = 1, 2, \ldots)$$

$$(\lambda + i\mu + f)P_{(i,N)} = (i+1)uP_{i+1,N} + \lambda P_{i-1,N} + rP_{i,F} \qquad (i = 1, 2, \ldots)$$

One important result from Baykal-Gursoy and Xiao *(14)* showed that the expected number of vehicles on the link can be represented as below when the M/M/∞ queueing system is experiencing service interruptions:

$$\mathrm{E(X)} = \frac{\lambda}{\mu} + \frac{\lambda f(\mu-\mu')}{\mu^2(r+f)}\left(1+\frac{(f+\mu)(\mu-\mu')}{(r\mu+f\mu'+\mu\mu')}\right)$$

Based on Little's Theorem, the average travel time on the link can be calculated as follows:

$$\mathrm{W} = \frac{E(X)}{\lambda} = \frac{1}{\mu} + \frac{f(\mu-\mu')}{\mu^2(r+f)}\left(1+\frac{(f+\mu)(\mu-\mu')}{(r\mu+f\mu'+\mu\mu')}\right)$$

With the intention of accommodating commercial vehicle violations, a correction factor c_j is applied to depict the impact of illegal-parked trucks. Moreover, illegal parking impact can vary among different street segments due to traffic volume, number of lanes, gross leasable area (GLA) of commercial properties and so on. To capture these effects, the above equations are modified. Let's denote $\boldsymbol{A}$ be the set of roadway links in the study area, $V=\{P,T\}$ to indicate the vehicle type, where P stands for passenger cars and T stands for commercial trucks. To reflect road segment characteristics, the above equations are reformatted to compute the average number of vehicles on the roadway link i:

$$N_i = c_i \frac{D_i L_i}{v_i}\left[1+\frac{\sum_j c_j f_{ij}\left(1-\frac{v'_i}{v_i}\right)}{\frac{1}{d_i}+\sum_j c_j f_{ij}}\left(1+\frac{\left(\sum_j c_j f_{ij}+\frac{v_i}{L_i}\right)(v_i - v'_i)}{\frac{v_i}{d_i}+\sum_j c_j f_{ij}\, v'_i+\frac{v_i v'_i}{L_i}}\right)\right],\quad (\mathrm{i} \in \mathrm{A}, \mathrm{j} \in \mathrm{V})$$

Subsequently, the average link travel time is:

$$t_i = c_i \frac{L_i}{v_i}\left[1+\frac{\sum_j c_j f_{ij}\left(1-\frac{v'_i}{v_i}\right)}{\frac{1}{d_i}+\sum_j c_j f_{ij}}\left(1+\frac{\left(\sum_j c_j f_{ij}+\frac{v_i}{L_i}\right)(v_i - v'_i)}{\frac{v_i}{d_i}+\sum_j c_j f_{ij}\, v'_i+\frac{v_i v'_i}{L_i}}\right)\right],\quad (i \in A, j \in V)$$

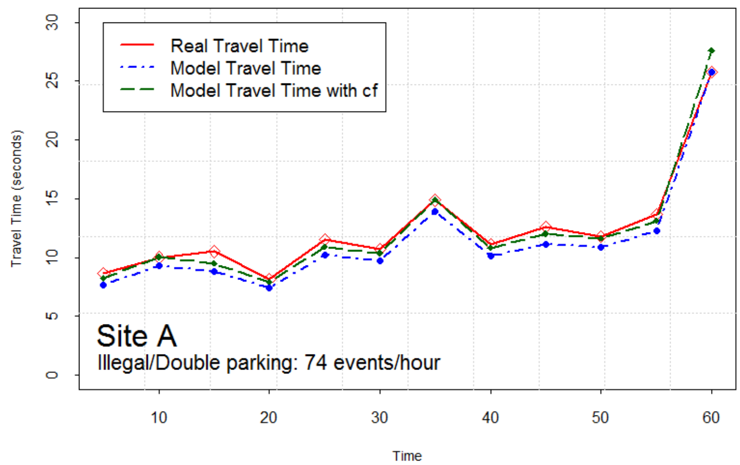
Where L_i=length of link *i* (mile); D_i=hourly traffic demand on link *i* (veh/h); f_{ij}=frequency of illegal parking for vehicle type *j* on link *i* (events/hour), d_i=average duration time of illegal parking on link *i* (hour)*;* v=average speed without illegal parking on link *i* (mph); v'_i=average speed with illegal parking on link *i* (mph), c_i = correction factor for link *i* and c_j = correction factor for vehicle type *j*. In this study, we use $c_j=2$ and c_i=1.07 based on the observed data from a previous study *(1)*.

Root mean square error (RMSE) is used as the performance measure. Figure 4 and Table 4 provide a summary of the model results. Overall, the result implies that the M/M/∞ queueing model produced a good fit with the field data in terms of average link travel times during the majority of time for the three study sites. For Site B during 8:30-8:40, a noticeable gap was identified between the field data and the queueing model output. After investigating the recorded video data, we found there were five vehicles that arrived within one minute and illegally parked on the left-most lane. Furthermore, one of the vehicles parked illegally for almost 10 minutes. This resulted in the loss of one lane capacity of Site B during 8:30-8:40. M/M/∞ queueing model was found to underestimate the impact under this extreme condition.

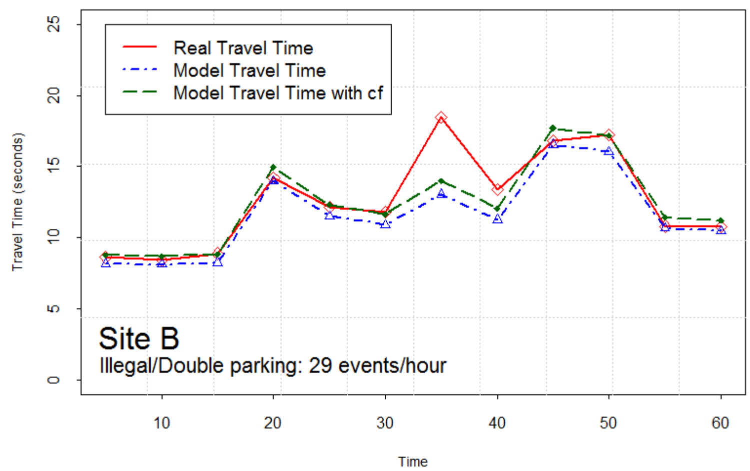
The correction factors that account for vehicle types and spatial locations were proved useful, especially for Site A (high illegal parking activities) and Site B (moderate illegal parking activities). Nevertheless, it did not improve the RMSE for Site C (low illegal parking activities).

Table 4 Queueing Model results

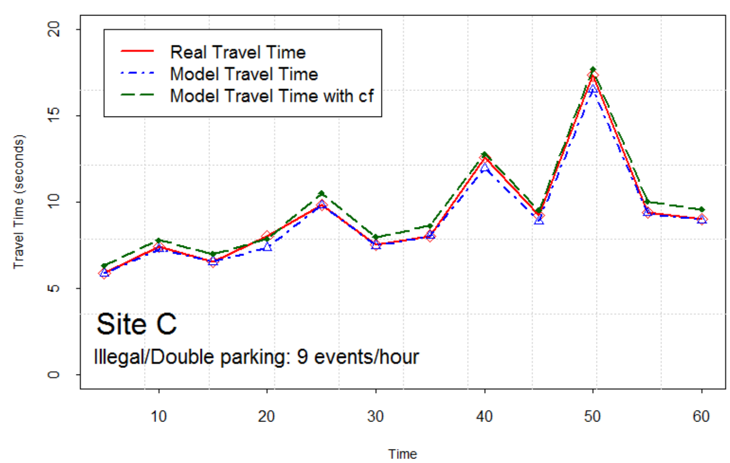
RMSE	M/M/∞ Model	M/M/∞ Model with correction factor
Site A	1.0992	0.6977
Site B	0.3665	0.4534
Site C	1.7628	1.4151



(a) Site A with high illegal parking activities



(b) Site B with moderate illegal parking activities



(c) Site C with low illegal parking activities

*cf = correction factor

Figure 4 M/M/∞ queueing model results

As an extenion of a previous research *(1)*, this study introduced a three step methodology that can be summarized as "identify, rank and quantify" for illegal parking in urban areas. First, a screening tool which uses Local Moran's I statistics to identify and rank double-parking hotspots is developed. Then, a M/M/∞ queueing model is used to compute average travel times under illegal parking conditions. The proposed model considers the effect of illegal parked commercial trucks with respect to their spatial locations. The previous study only tested the feasibility of using queueing model to quantify double parking effect and was limited by the amount and low-resolution data collected. This study further extends the queueing model approach and validate it by using three study sites in one of the busiest areas in the highly complex New York City network with a different frequency of illegal parking activities (low, moderate, and high level of activities). The proposed model was proved to be effective under most conditions, especially in the presence of double parking events. The current study used a uniformed correction factor generated from the previous study *(1)* and only focused on single block estimation. Future research will concentrate on collecting more data to generlize the correction factor function and consider network effects.

The findings of this study can provide transportation agencies with useful insights on fast estimation of illegal parking impacts on urban streets and the management and alleviation of on-street parking problems.

Reference

1. Gao, J. and K. Ozbay, Modeling Double Parking Impacts on Urban Street. In: Proceedings of the Transportation Research Board 95th Annual Meeting, Washington, D.C., 2016.
2. Holguín-Veras, J., K. Ozbay, A. Kornhauser, A. Shorris, and S. Ukkusuri. *Integrative Freight Demand Management In the New York City Metropolitan Area, Final Report.* DTOS59-07-H-0002. 2010.
3. Kladeftiras, M. and C. Antoniou, Simulation-based assessment of double-parking impacts on traffic and environmental conditions. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2390*, 2013. pp. 121-130.
4. Galatioto, F. and M. Bell, Simulation of Illegal Double Parking: Quantifying the Traffic and Pollutant Impacts. In: Proceedings of the IV International SIIV Congress, 2007. pp. 12.
5. Anselin, L., Local indicators of spatial association—LISA. *Geographical analysis 27(2)*, 1995, pp. 93-115.
6. NYC OpenData, https://nycopendata.socrata.com/, Accessed Jan. 1st, 2017.
7. U.S. Census Bureau. *Geocoding Services Web Application Programming Interface (API).* 2017.
8. Google, Geocoding API, developers.google.com/maps/documentation/geocoding, Accessed July 18, 2016.
9. Gao, J., K. Xie, and K. Ozbay, Exploring the Spatial Dependence and Selection Bias of Double Parking Citations Data. *Transportation Research Record: Journal of the Transportation Research Board (Accepted)*, 2018.
10. Xie, K., K. Ozbay, A. Kurkcu, and H. Yang, Analysis of traffic crashes involving pedestrians using big data: investigation of contributing factors and identification of hotspots. *Risk analysis 37(8)*, 2017, pp. 1459-1476.
11. Xie, K., K. Ozbay, Y. Zhu, and H. Yang, Evacuation Zone Modeling under Climate Change: A Data-Driven Method. *Journal of Infrastructure Systems 0(0)*.
12. Moran, P.A., The interpretation of statistical maps. *Journal of the Royal Statistical Society. Series B (Methodological) 10(2)*, 1948, pp. 243-251.
13. Anselin, L., GeoDa 0.9 user's guide. 2003.
14. Baykal-Gursoy, M. and W. Xiao, Stochastic decomposition in M/M/∞ queues with Markov modulated service rates. *Queueing Systems 48(1-2)*, 2004, pp. 75-88.