Ambulance Pre-Assignment Problem: A Case Study of Jersey City, New Jersey

EunSu Lee
New Jersey City University

Natalia De La Fuente
New Jersey City University

Melanie McDonald
New Jersey City University

Erin O'Neill
New Jersey City University

Pre-assignment is one of the typical operational strategies for an ambulance in service for quick response to a call. By doing so, the ambulance can be assigned in advance for a call to a predicted area in order to avoid traffic congestion depending on the frequency and type of calls. This study proposes a model minimizing total annual travel time from ambulance stations to incidents while considering historical call volume. The propositioned model can accomplish noteworthy results of coverage compared to the classic use of ambulance stations. It will be beneficial especially where the most vulnerable population resides.

Keywords: ambulance service, public health, pre-assignment, spatial analysis, GIS, emergency medical service

INTRODUCTION

In New Jersey counties, an emergency medical service (EMS) has an average of 90th percentile of 13 minutes response time to reach the scene of an emergency. In fact, the 90th percentile of response time in any NJ counties are over 9 minutes, according to the September 2019 report from the Office of Emergency Medical Services (NJ Department of Health, 2019). In the City of Jersey City, the results are not different. The Office of Emergency Medical Services (OEMS), New Jersey Department of Health (NJDOH) reports that the average response time of Jersey City is 9 minutes and 11 seconds in September 2019 (Figure 1). This data was submitted by Advanced Life Support (ALS) and Basic Life Support (BLS) EMS agencies and considers only the emergency/911 calls organized into the counties in which the emergency occurred. The report contains approximately 85% of all BLS and 100% of all ALS 911 pre-hospital emergency medical request records in NJ where EMS has responded to an actual patient (NJ Department of Health / Office of Emergency Medical Services, 2017). To analyze the response time of each city in NJ, NJDOH divided the cities into geographical areas based on population density. Towns are divided by the town size and the population per square miles. Jersey City belongs to Group 1 which is a Town Size of 5+ squared miles (NJ Department of Health / Office of Emergency Medical Services, 2017).

FIGURE 1
90TH PERCENTILE EMERGENCY MEDICAL SERVICE (EMS) RESPONSE TIME OF GROUP 1 IN NEW JERSEY IN SEPTEMBER 2019 (NJDOH / OEMS, 2019)

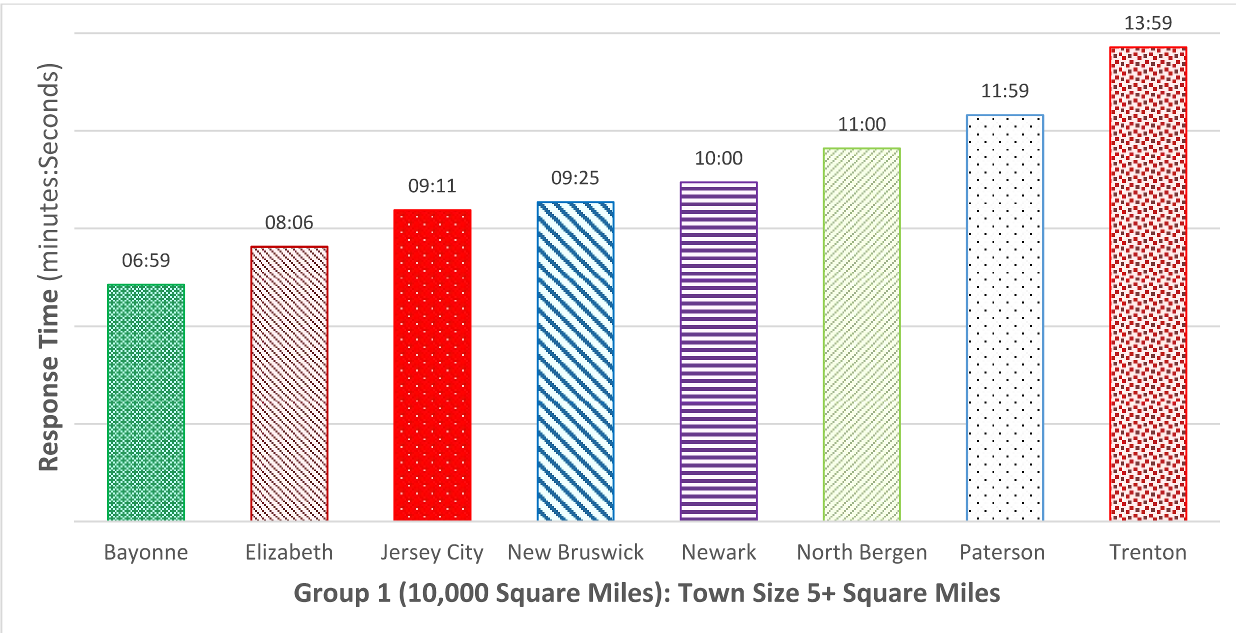


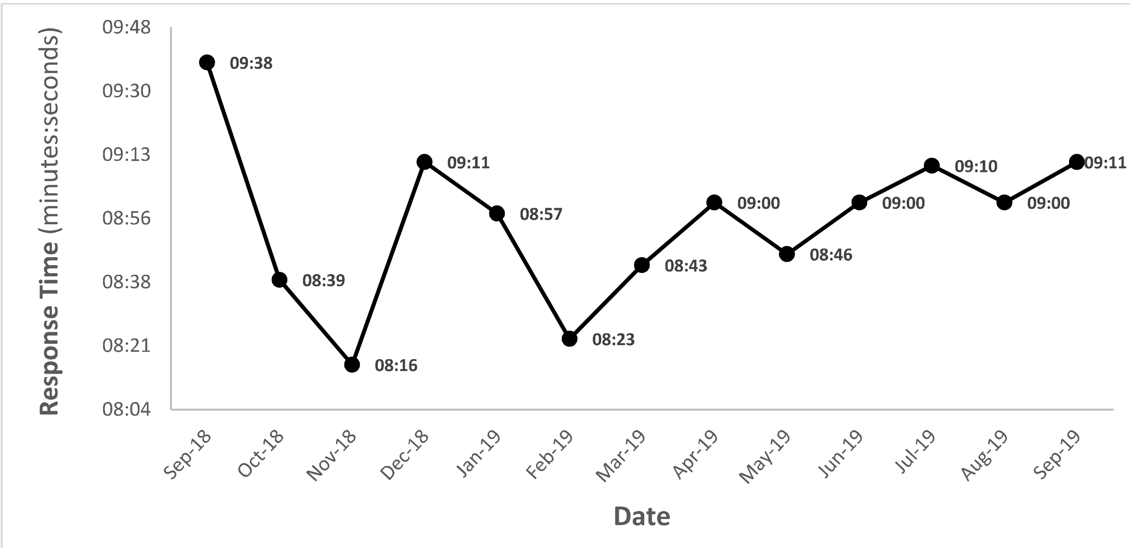
Figure 2 shows the response time of Jersey City from the period of September 2018 through September 2019. In the late of 2018, the response time was up and down, while the response time in 2019 steadily increased. The increase may be caused by population growth, resource allocation, call patterns, etc. Regardless of the factors, the EMS response team would control and maintain the response time in a good status. The systems may consider strategic resource allocation in an innovative way, while sustaining the targeted service level. As part of operational innovative strategy, pre-assignment or dynamic allocation systems can be considered in the area. Pre-assignment approach is widely adopted in the industry but locating the ambulances at right places is the crucial task. Finding the right locations is one of the biggest problems because it is difficult to choose the best place to locate emergency facilities.

This study examines the distribution of EMS response time based on existing EMS system and historical response data in Jersey City. The study proposes a mixed integer linear programming (MILP) by minimizing system-wide total response time (i.e. travel time) between the pre-assigned ambulance service locations and centroids of census tracks as demand locations. The public health planners and the ambulance operators would benefit from the study. The methods proposed in this study can be replicated at other urban areas by changing such parameters as travel speed, ambulance capacity, required response time, etc.

LITERATURE REVIEW

Response time is important when dealing with emergency cases. Swift responses characterized by minimal delays improve the quality of outcome of emergency and incident response activities. On the other hand, delays deter or decrease the quality of services when it comes to emergency response and resources. Even though many references are available in a wide range of applications of EMS, we reviewed literature which are very closely related to this study.

FIGURE 2
90TH PERCENTILE EMERGENCY MEDICAL SERVICE (EMS) RESPONSE TIME JERSEY CITY IN THE PERIOD OF SEPTEMBER 2018 – SEPTEMBER 2019 (NJDOH / OEMS, 2019)



Saisubramanian et al. (2015) focused on ambulance allocation and dynamic reallocation for the complete ambulance fleet. In order to facilitate the detection of patterns and to prevent experimental worst case planning, they used a data-driven approach. Firooze et al. (2018) focused on emergency vehicle location and relocation problems with the purpose of dealing with emergency demands considering service time, or response time and return time.

A mixed integer linear programming (MILP) is widely applied for ambulance location and relocation problem (Firooze et al., 2018; Liu et al., 2017; Pérez et al., 2015; Saisubramanian et al., 2015). The objective of their studies was to minimize the total response time and relocation time that they used for the case and penalty costs according to the results (Firooze et al., 2018). Pérez et al. (2015) explained reducing response time using linear programming (LP) model to locate ambulances. Liu et al. (2017) determined that the response time with the patient that will be sent to the nearest hospital to receive the appropriate service which should not exceed in thirty minutes. Saisubramanian et al. (2015) minimized the maximum response time value instead of reducing estimated response time, so that the chance of discovering requests with a higher value is bounded. In their dynamic relocation model in order to allow for "live" reallocation of ambulances, they delivered a decomposition method based on Lagrangian Relaxation.

Saisubramanian et al. (2015) presented a risk-based optimization criterion for analyzing past incident results. Pérez et al. (2015) used p-medians and set covering approach helped to shorten the distance between the user and the ambulance. Aringhieri et al. (2007) proposed a three-step method that can help to improve ambulance location and reduce response time during emergencies. Swalehe and Gunay (2016) used an ArcGIS network analyst position allocation tool to identify best emergency stations from which to hit ambulance demand areas within 5 minutes of driving time. Saisubramanian et al. (2015) revealed that if a consistent sequence of applications emerges within a few hours, it would be advantageous to adjust the allocation determined from a large number of requests. Through using request sets for the next few hours from multiple days, they determined that it is practical to not move the same ambulance repeatedly over two locations.

Firooze et al. (2018) found that by increasing the number of ambulances from three to four, the objective function is expected to improve. The study focused on emergency vehicle location and relocation problem with the purpose of dealing with emergency demands considering service time or response time and return time. Liu et al. (2017) used the shortest travel time from ambulance stations to the demand points and from demand to hospital on the basis of Euclidean distance, while Lee et al. (2013) analyzed response time with road network travel time considering various travel speed based on road classification. Liu et al. (2017) used a running average speed of an ambulance of 55 km per hour, without considering the different types of speed of different routes, but this speed is consistent with the traffic conditions for most of the Shanghai urban streets. They focused on reducing the time of response to cover all possible demands.

Pérez et al. (2015) explained reducing response time using linear programming model to locate ambulances. The study considered different location of users, costs, and demand of transport in the area, the number of services required, the geographical location, and the capacity of each location are the basic things to consider when seeking to optimize the quality of services and operations in emergency response and disaster management. Thus, to solve the problem that affects many response teams in the world, Pérez et al. (2015) used *p*-medians and set covering approach helped to shorten the distance between the user and the ambulance (Lee, 2014). Aringhieri et al. (2007) proposed a three-step method that can help to improve ambulance location and reduce response time during emergencies. The first step in this method analyzed the real time data on the system to understand its behavior. The second step involves using mixed integer linear programming (MILP), while the third method involves finding new locations. Aringhieri et al. (2007) concluded that the integer method could help to improve emergency responses by reducing response, but the method was subject to further discussion.

Liu et al. (2017) utilized a mixed integer programming with the CPLEX software. They concluded that at least 60% of the calls should be covered in less than eight minutes with a liability of 90%. Noble's (1973) results showed that more than one ambulance in the pre-assignment lists influences location decisions and consequently the system performance. The capacity of ambulances and their maximal workload also impact the decisions, but mostly pre-assignment ones. Indeed, when the system capacity is tight, pre-assignment lists adjust accordingly thus providing a better estimate of system performance and workload balance.

Saisubramanian et al. (2015) evaluated the effectiveness of their findings on ambulance data sets from two large cities in Asia. The approach was able to reduce the response time by at least a few minutes over current practice. The study revealed that if a consistent sequence of applications emerges within a few hours, it would be advantageous to adjust the allocation determined from a large number of requests. Saisubramanian et al. (2015) determined that it is practical to not move the same ambulance repeatedly over two locations. Swalehe and Gunay (2016) directed an analysis in Odunpazari District of Eskisehir Province, Turkey. The fleet of 17 ambulances in the Odunpazari District was deployed differently each six hours to balance ambulance demand and as a consequence, 77.6% of the emergency demand areas were located at the closest ambulance station within 5 minutes of driving time.

The difference between the proposed study and the existing literature is that the existing literature mostly focuses on analyzing the functioning of the entire emergency response systems in developed areas, while the current study examines emergency response time and outcomes in Jersey City, taking into account income level and fewer facilities. For this reason, the approaches used in the existing literature might not work for this study. Therefore, this research will help fill the gap left by the existing literature when it comes to addressing emergency services in low income areas.

MODEL DEVELOPMENT

Ambulance locations are needed in order to improve emergency medical services (EMS) and manage efficient public health systems for a sustainable society. Thus, this study investigates optimal pre-assigned ambulance locations in urban areas using historical 911 call locations. For the optimal locations minimizing total response time (i.e. travel time), the study proposes a mixed integer linear program

(MILP). Thus, the decision variables of the study are two sets: the first set of decision variables are the potential ambulance locations responding to 911 calls (y_i), and the other set of decision variables are the number of calls from each demand area, which each potential location covers (x_{ij}). In this model, we assume that the coordinates of ambulance location i can be the same with the coordinates of a centroid of demand area j before assigning any ambulances.

x_{ij} : The number of 911 calls x sourced from a centroid of demand area j are being served by ambulance location i

$$y_i : \begin{cases} 1, \text{If an ambulance is preassigned at location } i \\ 0, \text{otherwise} \end{cases}$$

The objective function is to minimize annual total travel time responding to call volumes per year (1). This is a system optimum solution instead of finding a solution for each zone. Using sum of total annual trips from ambulance location (i) to scenes (j) multiplied by travel time between i and j. The travel time (t_{ij}) is the quickest travel time, which is travel distance (d_{ij}) divided by the estimated travel speed *en route* (s_{ij}). Over the region, the travel speed varies by the road types and the local traffic between ambulance location i and scenes j.

$$Min\ Z = \sum_{i=1}^{I} \sum_{j=1}^{J} x_{ij}\, t_{ij} \tag{1}$$

where:

- i : Preassigned ambulance location, $i = \{1, 2, \ldots, I\}$
- j: Location (scenes) of 911 call demand, $j = \{1, 2, \ldots, J\}$
- t_{ij}: Travel time in minutes from preassigned ambulance location (i) to a centroid of demand area (j)
- d_{ij}: Travel distance from ambulance location (i) to a centroid of demand area (j)
- s_{ij}: Travel speed in minutes

The ambulance response time from i to j should be less than upper bound response time (UB_RT) to reflect local policy and requirement (2). The lower bound of t_{ij} should be very close to zero (e.g. $t_{ij} \geq 0.00001$ second). For the purpose of model development, the lower bound of t_{ij} can be zero since the ambulance location is placed at the centroid of demand zone. Each ambulance cannot serve more than maximum calls (MAXCALLS) per year due to the limited capacity of an ambulance (3). In addition, there should be less than a certain number of pre-assigned ambulances (MAXAMB) in the region to receive a call due to a budget and resource constraint (4). The demand of 911 calls can be served by one or more ambulances (5). However, all 911 calls should be served by ambulances within the required response time. The number of calls assigned to an ambulance service should be non-negative and integer (6). The decision variables for preassigned location should be binary variable (7).

$$0 \leq t_{ij} \leq \text{UB\_RT} \tag{2}$$

$$\sum_{i=1}^{I} x_{ij} - \text{MAXCALLS} \cdot y_i \leq 0, for\ \forall j = \{1,2, \ldots, J\} \tag{3}$$

$$\sum_{i=1}^{I} y_i - MAXAMB \leq 0 \tag{4}$$

$$\sum_{i=1}^{I} x_{ij} - CALL_j = 0, for\ \forall\, j = \{1,2, \ldots, J\} \tag{5}$$

$$x_{ij} \geq 0 \text{ and integer} \tag{6}$$

$$y_i = \{0, 1\} \tag{7}$$

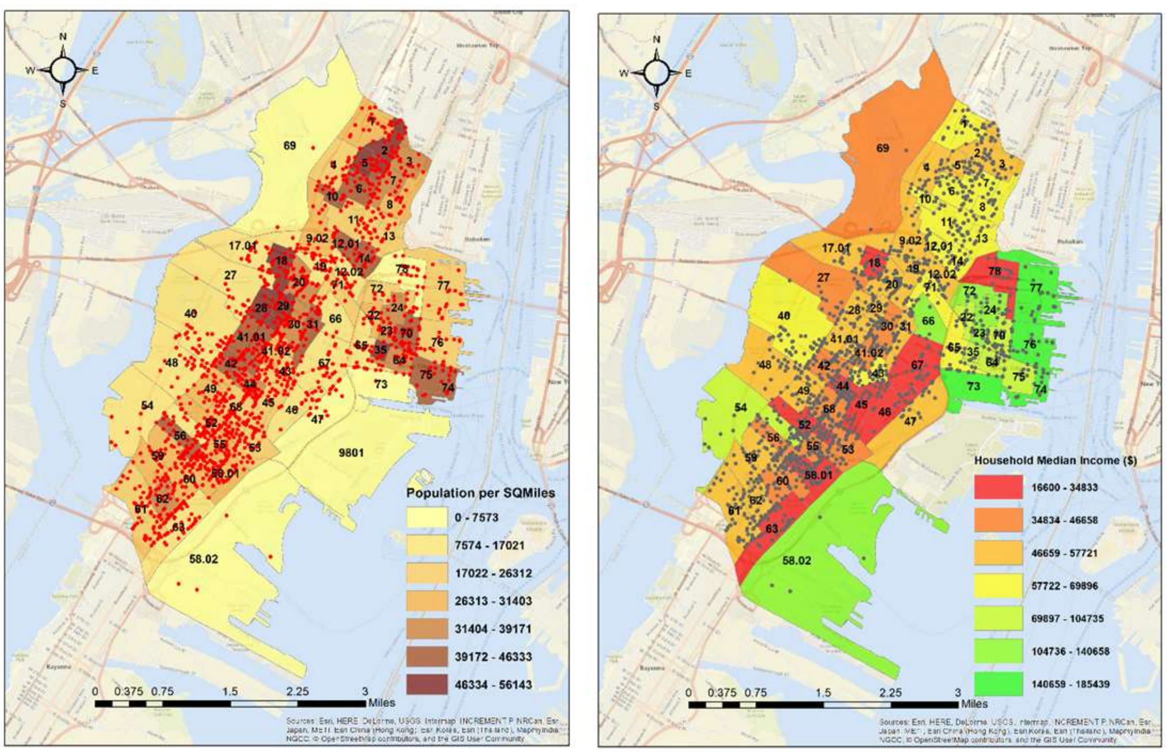
CASE STUDY

Study Region

The City of Jersey City (JC) is located in the North Eastern part of the State of New Jersey. The city is located in Hudson County. Jersey City has a population of about 16,736 per square mile in 2010. On the other hand, the land covered per square mile is approximately 14.79 in 2010. The eastern waterfront is located in a strategic position where it faces the Hudson River (Washko & Heightman, 2012). This is also a place where the waterfront meets the Upper New York Bay. The city has eleven neighborhoods: Bergen-Lafayette, Downtown, Greenville, Hackensack Riverfront, Journal Square, Liberty Park, Lincoln Park, McGinley Square, The Heights, The Waterfront, and West Side. Jersey City counts with 67 Census Tracts ($J = 67$).

Jersey City has the largest population in Hudson County and the second largest in the state of New Jersey (World Population Review, 2019). The city is part of the New York metropolitan area, and it is known for its distribution, busiest transportation and manufacturing center situated with one of the busiest urban ports. Population density is high in the middle of the region and downtown near the waterfront (Figure 3). The map also demonstrates how most of the calls are originated from the highly populated census tracts.

FIGURE 3
DISTRIBUTION OF 911 CALLS WITH POPULATION DENSITY OF 2010 (A) AND HOUSEHOLD INCOME OF 2017 (B) BY CENSUS TRACT IN JERSEY CITY



(A) POPULATION DENSITY **(B) HOUSEHOLD INCOME**

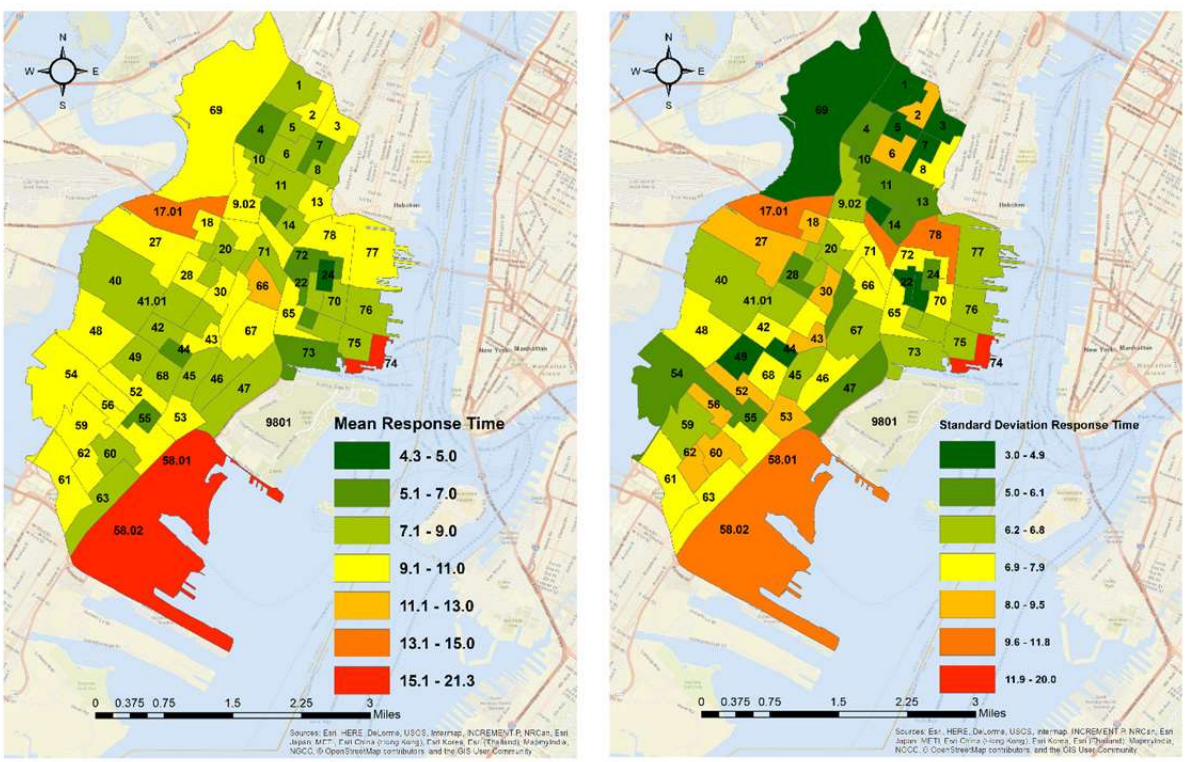
Previous studies suggest that the response time of ambulance services in low income areas take longer, leading to disparities in prehospital delivery of care over time (Hsia et al., 2018). This information is crucial to understand if there are any links between poverty levels and ambulance services (See Figure 4). This study also collected median income and population from each census tract in the period of 2017

(Data USA, 2019). In Jersey City, households have a median annual income of $66,264, which is more than the median annual income of $60,336 across the entire United States (Data USA, 2019). As well, in 2017, the location of Census Tract 74 shows the highest Median Household Income in Jersey City with a value of $185,439, followed by Census Tract 76 and Census Tract 77, with respective values of $181,923 and $152,325 (Data USA, 2019). On the other hand, the place with the lowest value of $16,600 is from Census Tract 44.

According to the Office of Emergency Medical Services, New Jersey Department of Health, following the 90th percentile methodology, the response time in September 2019 was 9 minutes and 11 seconds in Jersey City (NJ Department of Health / Office of Emergency Medical Services, 2019). Visual analytics helps us to understand the distribution of the average response time and variability of the ambulance response time of the city of Jersey City. Figure 4-A also displays the average response time to the 911 calls of 2017. Average response time is slow near the waterfronts including Hudson River on the East and Hackensack River on the West. The areas of slow response time will have extra ambulances in motion to quickly respond to any future calls.

Figure 4-B shows the standard deviation of the response time in the regions. Even if the average response time is high, if the standard deviation is high, the service is not predictable. Thus, high average response time with narrow standard deviation is desirable. In addition to the waterfronts, the standard deviation in the middle of the region is also wide. Thus, the ambulance locations in the middle and near the waterfront area will improve the service overall.

FIGURE 4
AVERAGE RESPONSE TIME (A) AND STANDARD DEVIATION (B) OF JERSEY CITY BASED ON THE 911 CALLS OF 2017



(A) AVERAGE **(B) STANDARD DEVIATION**

Data Description

The Census Tracts were downloaded from Jersey City Open Data. The data set includes x- and y-coordinates, identification codes of a census tract and 911 call locations. The 911 calls correspond to the 2017 period and only include emergency calls (Jersey City Open Data, 2018). The number of 911 calls for ambulances on the census tracts were overlaid and visualized with population density and median household income in Figure 3.

The Jersey City Medical Center operates a fleet of 50 ambulances, about 25 in service each day (Quigley, 2019). However, the ambulances' actual routes are not known and their true locations of being dispatched are not known neither. Thus, the study assumes that the ambulance locations (*i*) are the centroids of census tracks. This study assumes that the travel speed responding to an emergency call is 25 mph over the entire city regardless of road types since the speed limit in the city boundary is 25 mph in general.

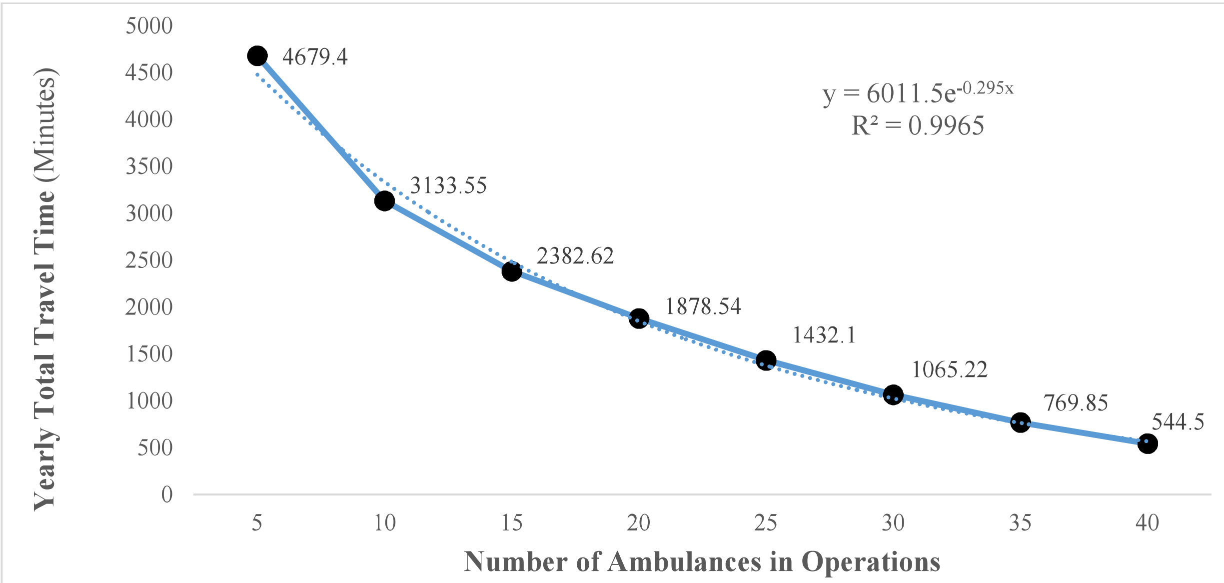
According to the State of New Jersey, an ambulance can "Exceed the maximum speed limits so long as he does not recklessly endanger life or property " (State of New Jersey , 2019). According to the official site of the State of New Jersey: "New Jersey law sets top speed limits for any given road, street , highway , or freeway . The speed limit, unless otherwise posted, is 25 mph in school zones, business, or residential districts; 35 mph in certain low density business and residential districts; 50 mph on all other roadways; 55 mph on certain state highways (as posted) and all interstates; and 65 mph on certain state highways (as posted)" (State of New Jersey, 2019).

RESULTS AND DISCUSSION

While solving the problem using LINGO 18.0 of LINDO® Inc., the study drew a receiver operating characteristic (ROC) curve in regard with the number of locations (i.e. the number of pre-assigned ambulance locations) to respond to any 911 calls in nine minutes (UB_RT = 9). The system measures the yearly total travel time in minutes (*Z*). The number of ambulances in the system is 50, which is the exact number of ambulances operated in 2017. However, 25 ambulances are in operations during each shift (MAXAMB = 25). The quickest travel time from preassigned ambulance stations to 911 call demand points is calculated on the basis of Euclidean distance.

Figure 5 shows a receiver operating characteristic (ROC) curve. The study shows that when five ambulances are in operations, the total travel time of the EMS service is 4679.4 minutes for a year. As the number of ambulances in operations increase, the yearly total travel time decreases exponentially.

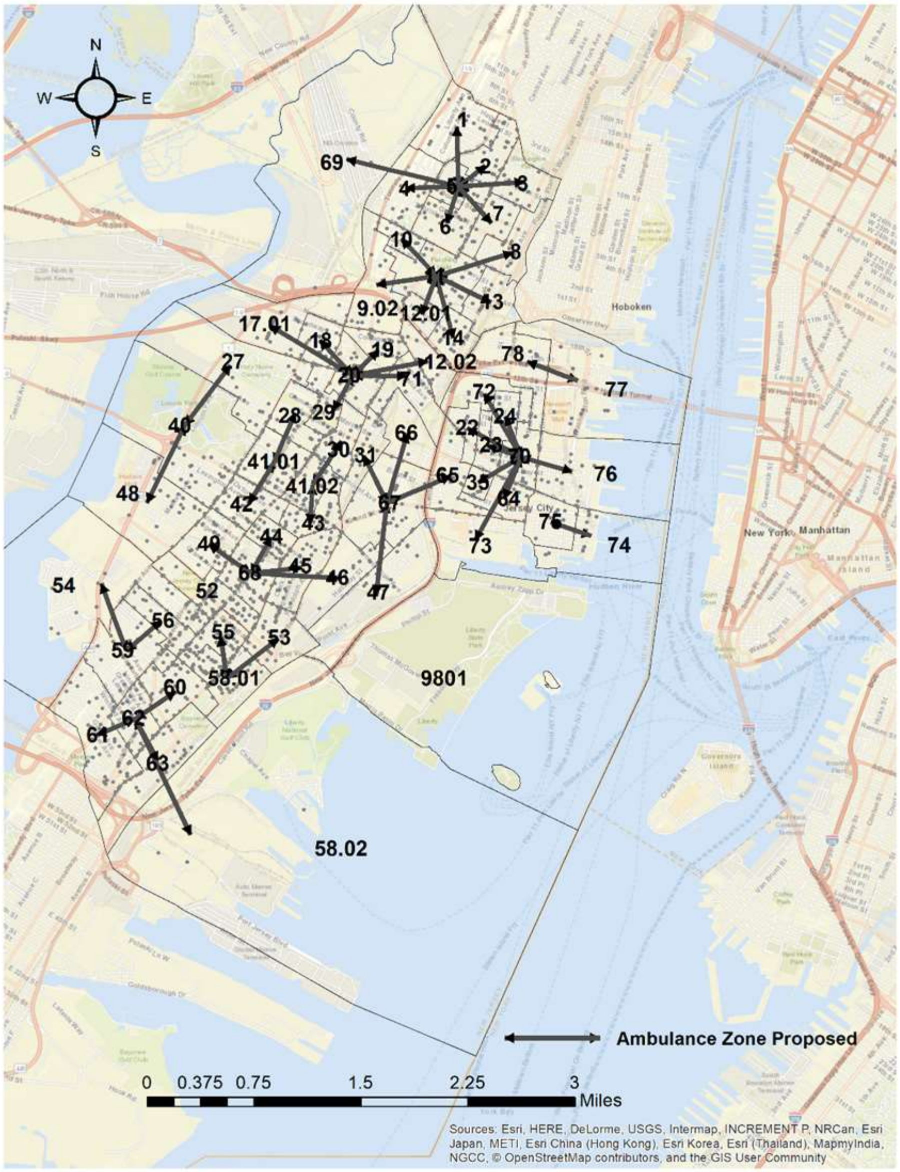
FIGURE 5
RECEIVER OPERATING CHARACTERISTIC CURVE BASED ON THE NUMBER OF AMBULANCES IN OPERATIONS



The R-squared value of the exponential model is 0.9965, which means 99.6% of the data is being explained by the model. The results show that the number of ambulances will provide faster response service to the residents; however, depending on the budget of the city and ambulance, the health planner should decide the service level.

Furthermore, Figure 6 shows the spider lines between a pre-assigned ambulance location and one or multiple census tracts that the ambulance will serve. For example, the figure proposed 15 pre-assigned ambulance locations while meeting the response time of nine minutes in the area. For example, one ambulance is pre-assigned at the center of Census Track 59 and serve any 911 calls from Census Tracks of 59, 54, and 56. Any 911 calls sources from the Census Track 74 (i.e. Downtown of Jersey City) will be served by an ambulance roaming in Census Track 76.

FIGURE 6
PROPOSED 15 AMBULANCE ZONES IN RESPONSE TO THE AVERAGE RESPONSE TIME OF 5 MINUTES



CONCLUSION

This study proposed a mixed integer liner programming model (MILP) minimizing total annual travel time from pre-assigned ambulance stations to incidents of 911 calls while considering historical call volume of 2019 in Jersey City, NJ. After studying and analyzing the ambulance locations for the Emergency Medical Services (EMS) in Jersey City, the study has concluded that the annual travel time is decreased exponentially by increasing the number of pre-assigned ambulances in the areas with 11 minutes as a response time requirement. In addition, it will be extremely beneficial especially where the most vulnerable population resides and the outskirts of the area. By adopting strategic operations, the system is able to serve the community in a sustainable way.

It is noteworthy that the results are based upon only one year's real demands data of 911 calls. The travel speed over the road network in the areas assumed that the ambulance respond is on a straight line

25 mph, not taking in consideration that there are different types of routes. Further study can be done with travel time over a true shape of road network in various time slots of commuting, day, night, weekday, and weekends.

ACKNOWLEDGEMENT

The authors would like to thank the reviewers for their scholarly and constructive evaluation on an anonymous referee. In addition, we would like to acknowledge the NJCU President's faculty fellows program and Dr. Saul Petersen from NJ Compact for their support to community-engaged learning (CEL). This work was supported by the SBR 2019-2020 from the NJCU Provost's Office.

REFERENCES

Aringhieri, R., Carello, G., & Morale, D. (2007). *Ambulance location through optimization and simulation: The case of Milano urban area.* Retrieved October 31, 2019, from https://air.unimi.it/retrieve/handle/2434/40782/6875/aor.pdf

Data USA. (2019). Jersey City, NJ. Retrieved November 22, 2019, from https://datausa.io/profile/geo/jersey-city-nj

Firooze, S., Rafiee, M., & Zenouzzadeh, S.M. (2018). An optimization model for emergency vehicle location and relocation with consideration of unavailability time. *Scientia Iranica*, 25(6), 3685-3699. doi: 10.24200/sci.2017.20022.

Hsia, R.Y., Huang, D., Mann, N.C., Colwell, C., Mercer, M.P., Dai, M., & Niedzwiecki, M.J. (2018). A US national study of the association between income and ambulance response time in cardiac arrest. *JAMA Network Open*, 1(7), e185202. doi:10.1001/jamanetworkopen.2018.5202

Jersey City Open Data. (2018). *Census tracts 2018 - Block and tract.* Retrieved November 22, 2019, from https://data.jerseycitynj.gov/map/911_calls/?location=15,40.6994,-74.08154&basemap=jawg.light

Lee, E. (2014). Designing Service Coverage and Measuring Accessibility and Serviceability of Rural and Small Urban Ambulance Systems. *Systems*, 2(1), 34-53. https://doi.org/10.3390/systems2010034.

Lee, E., Vachal, K., & Kayabas, P. (2013). *Phase I: Positioning Emergency Medical Services for Trauma Response for Rural Traffic Crashes: Pilot Case of Williston Basin in North Dakota, DP-270.* North Dakota State University, Fargo: Upper Great Plains Transportation Institute.

Liu, M., Yang, D., & Hao, F. (2017). Optimization for the locations of ambulances under two stage life rescue in the Emergency Medical Service: A case study in Shanghai, China. *Mathematical Problems in Engineering*, *2017* (Article ID 1830480), 14 pages. https://doi.org/10.1155/2017/1830480

NJ Department of Health (NJDOH) / Office of Emergency Medical Services (OEMS). (2017). *EMS Monthly report by County for October 2017.* Retrieved March 8, 2020, from http://www.state.nj.us/health/ems/documents/10.2017.EMSData.pdf

NJ Department of Health (NJDOH) / Office of Emergency Medical Services (OEMS). (2019). *EMS Monthly Report for September 2019.* Retrieved October 31, 2019, from https://www.nj.gov/health/ems/documents/September 2019 EMS Monthly Report Final.pdf

Noble, J. (1973). *Emergency medical services: behavioral, transportation and planning perspectives.* New York: Behavioral Publications.

Pérez, E. S., Flores, I., & Vindel, A. (2015). Ambulance location using linear programming: The case of the National Autonomous University of Mexico (UNAM). In *Proceedings of the European Modeling and Simulation Symposium.* Retrieved November 23, 2019, from https://pdfs.semanticscholar.org/2b4d/54a9da611d81e15a57073f6a7cb5960bc8f8.pdf

Quigley, J. (2019, April 30). *Time to honor EMS workers.* Retrieved November 23, 2019, from https://www.nj.com/hudson/2019/04/time-to-honor-ems-workers-quigley.html

Saisubramanian, S., Varakantham, P., & Lau, H.C. (2015). *Risk based optimization for improving emergency medical systems*. Retrieved March 8, 2020, from http://www.mysmu.edu/faculty/pradeepv/Papers/Ambulance.pdf

State of New Jersey. (2019). *Frequently asked questions [FAQs].* Retrieved November 21, 2019, from https://nj.gov/faqs/drive/

State of New Jersey - Department of Health. (2019). *NJ Emergency Medical Services (EMS) Agencies.* Retrieved October 9, 2019, from https://www.state.nj.us/health/ems/documents/reg-enforcement/nj_ems_agencies.pdf

Swalehe, M., & Gunay Aktas, S. (2016). *Dynamic ambulance deployment to reduce ambulance response times using geographic information systems: A case study of Odunpazari District of Eskisehir Province, Turkey.* Retrieved March 8, 2020, from https://www.sciencedirect.com/science/article/pii/S1878029616302353

Washko, J.D., & Heightman, A.J. (2012). *EMS royal flush: Jersey City Medical Center EMS reinvents itself & achieves five major national EMS recognitions.* Retrieved November 23, 2019, from https://www.firstwatch.net/wp-content/themes/firstwatch/files/040-049_JerseyCity_v2.pdf

World Population Review. (2019). *Jersey City, NJ.* Retrieved October 10, 2019, from http://worldpopulationreview.com/us-cities/jersey-city-population/