Optimizing London Fire Station Resources to Better Serve the Community

BIA-650
Section A
Professor Ted Stohr
Fall 2018

Sonali Johari

Table of Contents

Abstract

This paper uses historical data to simulate real-time emergency response for London fire incidents. Historical fire incident data for January through October 2017 and fleet list and fire station location information from the London Fire Brigade were combined to build the model in R. The simulation consisted of 100 days' worth of 15-minute simulation periods (9,600 total simulation periods). The number of fire incidents was assigned to each period using a zero-truncated Poisson distribution, and the actual incidents for each period were selected using a random draw from a subset of the full 85,000-incident historical data set. An Integer Programming optimization model was solved for each simulation period in order to minimize the total effective distance (a combination of actual distance and a random delay factor) traveled between the fire stations and each fire incident for that simulation period. The actual distances between the fire stations and incidents were calculated using the "Great Circle" distance formula.

The output of the model was the simulation results for each period, showing how the fire engines were dispatched to each incident, as well as the minimized total effective distance. The simulation assumed that fire engines dispatched in the previous two periods could not be redeployed in the current period, and therefore there were over 400 periods in the base simulation that did not have sufficient remaining fire engines to be dispatched to all incidents. In addition, two sensitivity analyses were run: the first added one fire engine to each fire station's initial fleet and the second added two fire engines to each fire station's initial fleet. Adding one engine provided enough resources to dispatch to all incidents in all but one simulation period. The additional fire engines also changed the distribution of how the fire engines were dispatched, increasing dispatch from certain fire stations while decreasing dispatch from others.

Introduction

When emergencies happen, efficient emergency response is critical. It is crucial that emergency response teams have sufficient resources that can be deployed in an optimal manner. This project uses information from the London Fire Brigade in order to examine the current and possible future state of emergency response. By simulating real-time emergency scenarios, potential opportunities to improve emergency response and better serve the community can be identified.

Problem Description

A historical database of over 85,000 fire incidents from 2017 (from January to October) were obtained from data provided by the London Fire Brigade as well as information from Kaggle. The historical incidents were combined with locations of fire stations in London and actual fleet list information in order to simulate real-world emergency scenarios, with the goal of showing how fire station resources can be efficiently deployed to each incident while minimizing overall travel distance for the fire engines. A sensitivity analysis was also run to assess the impact of adding additional resources to existing fire stations.

Data Preparation

For this project, public data sets from the London Datastore were used, as well as fire incident data from Kaggle. This information was cleaned and combined in R prior to be used in the simulation and optimization models.

Incident Data

The incident data from Kaggle had many interesting data points, including the type of incident (fire, special service or false alarm), the type of building (residential or commercial) and the actual station from which the fire engines were dispatched. Ultimately, the variables used in the analysis were incident number, date/time of the incident and location of the incident (both the borough and the postal code).

As part of the cleaning process, the data were filtered to exclude any locations without sufficient geographic information (i.e., any location that did not have a full postal code such as EC1Y 8LZ). In addition, any historical incidents that did not have a fire engine dispatched were also removed.

The *ggmap* package in R was used in order to geocode the location of each incident based on the full postal code value using the Google Maps API. These latitude and longitude coordinates were later used to calculate the master distance matrix between each incident and each fire station.

Fire Station Location Data

Using the *rgdal* package in R, a KML file (Google Maps proprietary format) containing the location of each London fire station was parsed into a data frame. This data frame included station name, street address, latitude/longitude coordinates, station call sign, status (open or closed) opened dates and closed dates (if applicable). The station list was filtered to include only currently open stations, as well as removing the Lambeth River Station, which does not have any fire engines (this station only dispatches rescue boats). This left a total of 102 fire stations in London.

Figure 1 below shows the total number of incidents in each London borough, as well as the 102 currently open fire stations.

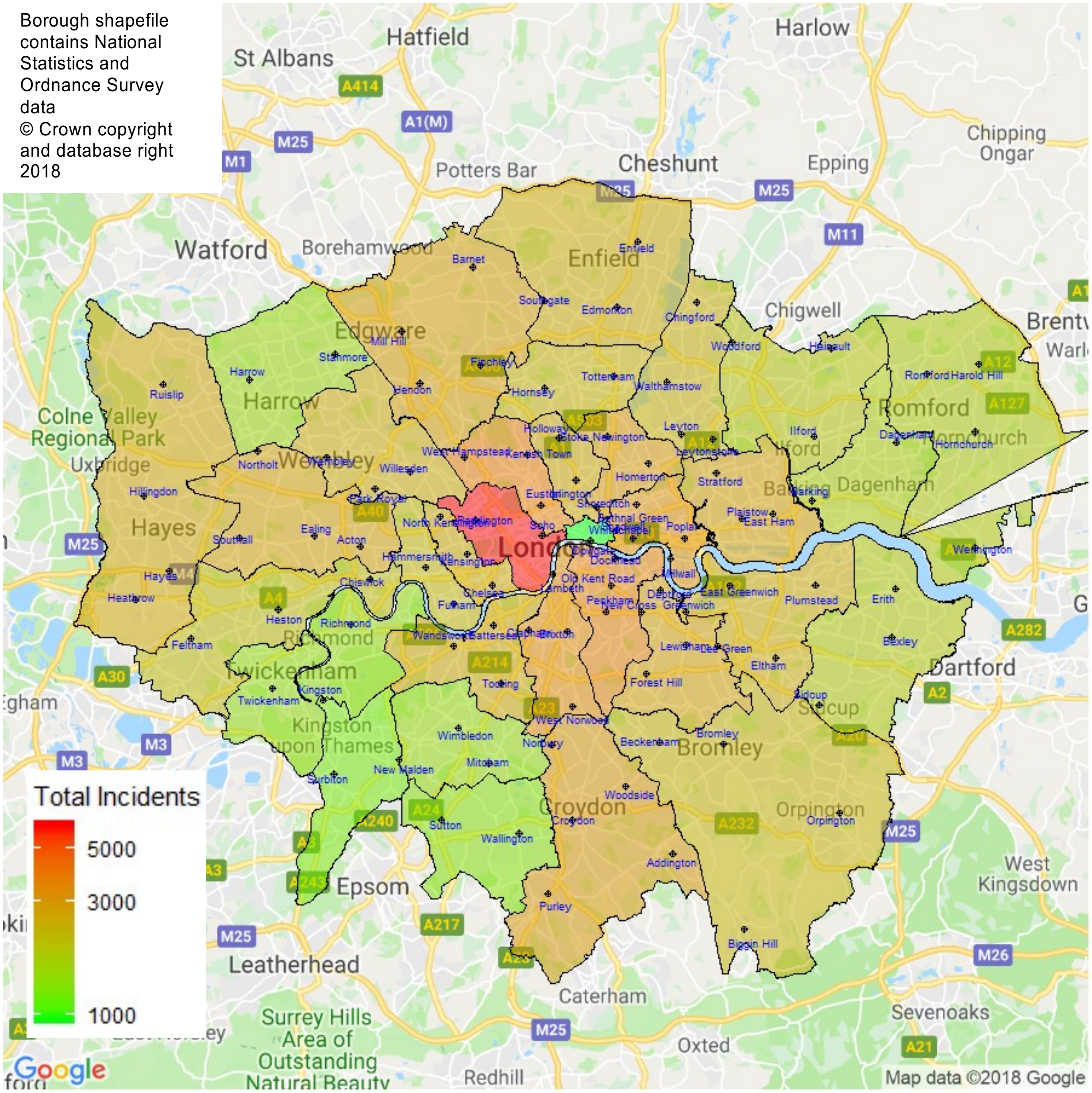

Figure 1: Historical Incidents and Fire Stations by Borough for January through October 2017

Distance Calculations

The original plan was to use the Google Maps API to calculate how far away each incident was from each station; however, each query to Google Maps took three seconds, which would take far too long to calculate the distances. In order to make the incident data set more manageable (and reduce the number of distance calculations required), only incidents in the "Northern Command" districts of Westminster, the City of London, Barnet, Camden, Enfield, Haringey and Islington were kept. The fire stations were also filtered to include only the following eight stations: Dowgate, Euston, Holloway, Islington, Kentish Town, Paddington, Soho and West Hampstead.

Even after creating a subset of incidents and fire stations, there were still about 12,000 incidents in the data set. Multiplied by the eight fire stations, this meant the master distance matrix required close to 100,000 entries, which would still require a large time investment in order to utilize the Google Maps API. Therefore, the Great Circle distance was used to calculate the distances.

The Great Circle formula calculates the distance between two points on the surface of a sphere. This method of calculating distances is quick and effective, which allowed the information needed to be computed in a reasonable amount of time.

The Great Circle distance formula is as follows:

$$d = 2r\sin^{-1}\left(\sqrt{hav(\phi_2 - \phi_1) + \cos(\phi_1)\cos(\phi_2)\, hav(\lambda_2 - \lambda_1)}\right)$$

where ϕ_1 and ϕ_2 are the latitude values in radians of location 1 and location 2, respectively, λ_1 and λ_2 are the longitude values, and r is the radius of the Earth (6,371 kilometers).
The haversine formula is:

$$hav(\theta) = \frac{1 - \cos(\theta)}{2}$$

This was implemented as a function in R, which was then called to calculate the distance between each incident and each fire station.

Model Methodology

Simulation

In real-world situations, emergency dispatch has to decide how best to deploy the limited resources of the London Fire Brigade when fielding many calls in a short period of time. In order to replicate these conditions, a simulation model was created that optimized the deployment of available fire engines from each fire station to the incidents.

For this model, a simulation set of 9,600 periods (100 days, split into 15-minute periods), was built in R. The simulation set used a fixed random seed, which ensured that the results were consistent every time the simulation was run.

The first step in creating the simulation model was to examine the historical incident data to determine the average frequency of incidents for each 15-minute period. This average frequency was 2.77 for the full set of incidents in the historical data. Figure 2 below shows the number of incidents in the historical data for the first three plus days of January 2017. The frequencies range from 1 incident to 11 incidents, with most of the periods having between 1 and 5 incidents.

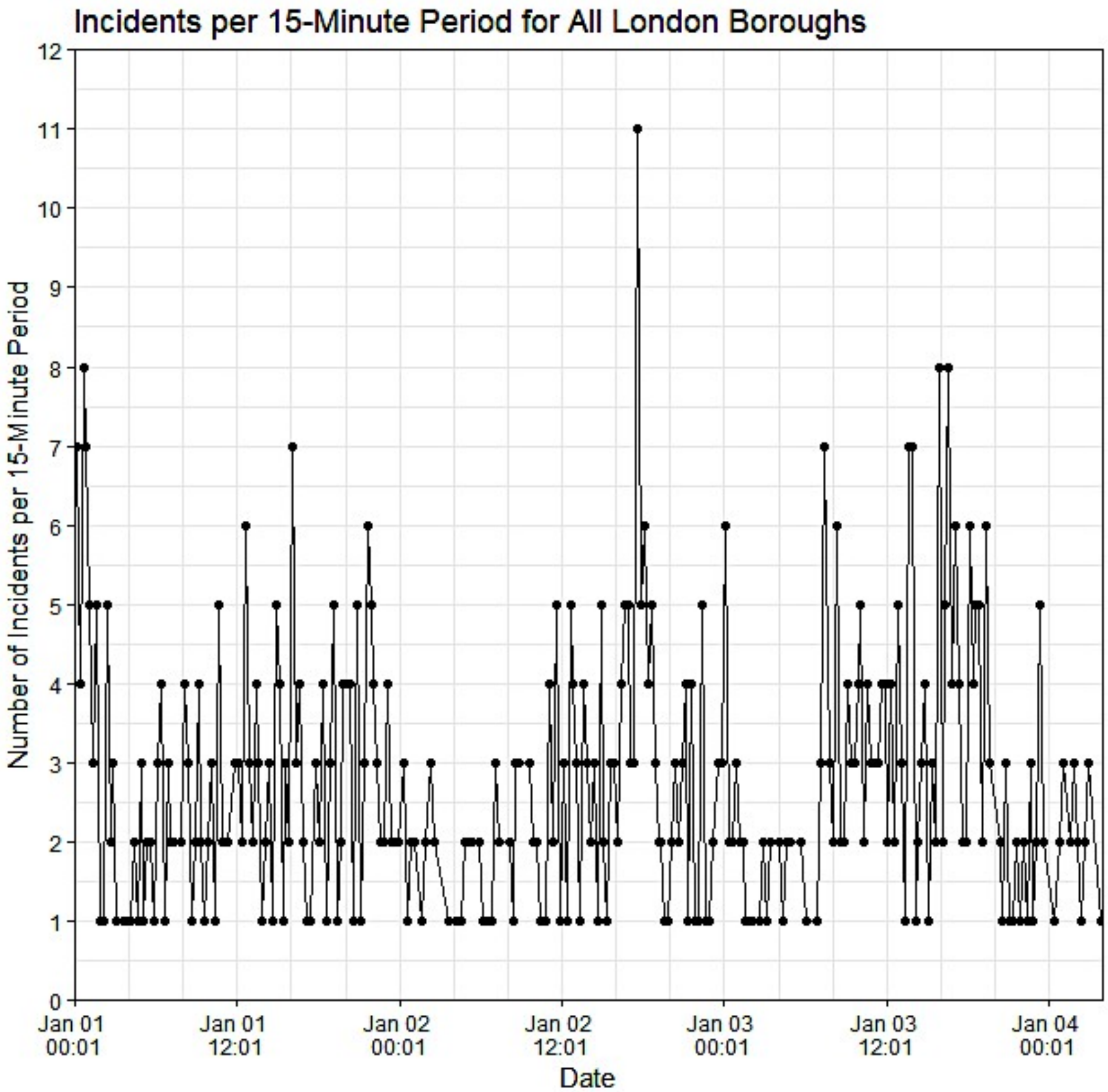

FIGURE 2: INCIDENTS PER 15-MINUTE PERIOD FOR JANUARY 1 TO JANUARY 4

The number of incidents in each simulation period was generated using the zero-truncated Poisson distribution. The standard Poisson distribution calculates discrete probabilities of having k events in a period using the following formula:

$$P(k\ events\ in\ a\ given\ interval) = \frac{\lambda^{A} e^{(B}}{k!}$$

where λ is the average frequency of the event occurring. For this simulation, λ was set to 2.7.

The zero-truncated version of the Poisson distribution does not allow k to ever be 0. This is consistent with the historical data (there was always at least one event in each 15-minute interval for the entire 10 months of historical data).

The probability function for the zero-truncated Poisson distribution is as follows:

$$P(k\ events\ in\ a\ given\ interval | k \neq 0) = \frac{P(k\ events\ in\ a\ given\ interval)}{1 - P(0\ events\ in\ a\ given\ interval)}$$

The probability of having zero events in a given interval can be calculated as follows:

$$P(0\ events\ in\ a\ given\ interval) = \frac{\lambda^{G} e^{(B}}{0!} = e^{(B}$$

Therefore, the probability function for the zero-truncated Poisson distribution can be calculated as:

$$P(k\ events\ in\ a\ given\ interval | k \neq 0) = \frac{\lambda^{A} e^{(B} / k!}{(1 - e^{(B})}$$

Table 1 below shows the probabilities for k ranging from 1 to 11 when λ is equal to 2.7.

TABLE 1: ZERO-TRUNCATED POISSON PROBABILITY DISTRIBUTION

1	19.5%	19.5%
2	26.3%	45.7%
3	23.6%	69.3%
4	16.0%	85.3%
5	8.6%	93.9%
6	3.9%	97.8%
7	1.5%	99.3%
8	0.5%	99.8%
9	0.2%	99.9%
10	0.0%	100.0%
11	0.0%	100.0%

The number of incidents assigned per simulated period was implemented in R using the *rztpois* function in the *countreg* package.

The histogram in Figure 3 shows the percentage of the total 15-minute periods (either actual historical periods or simulated periods for the simulation set) for each value of k (number of incidents per period). The average frequency for the simulation set was 2.87 incidents per simulation period.

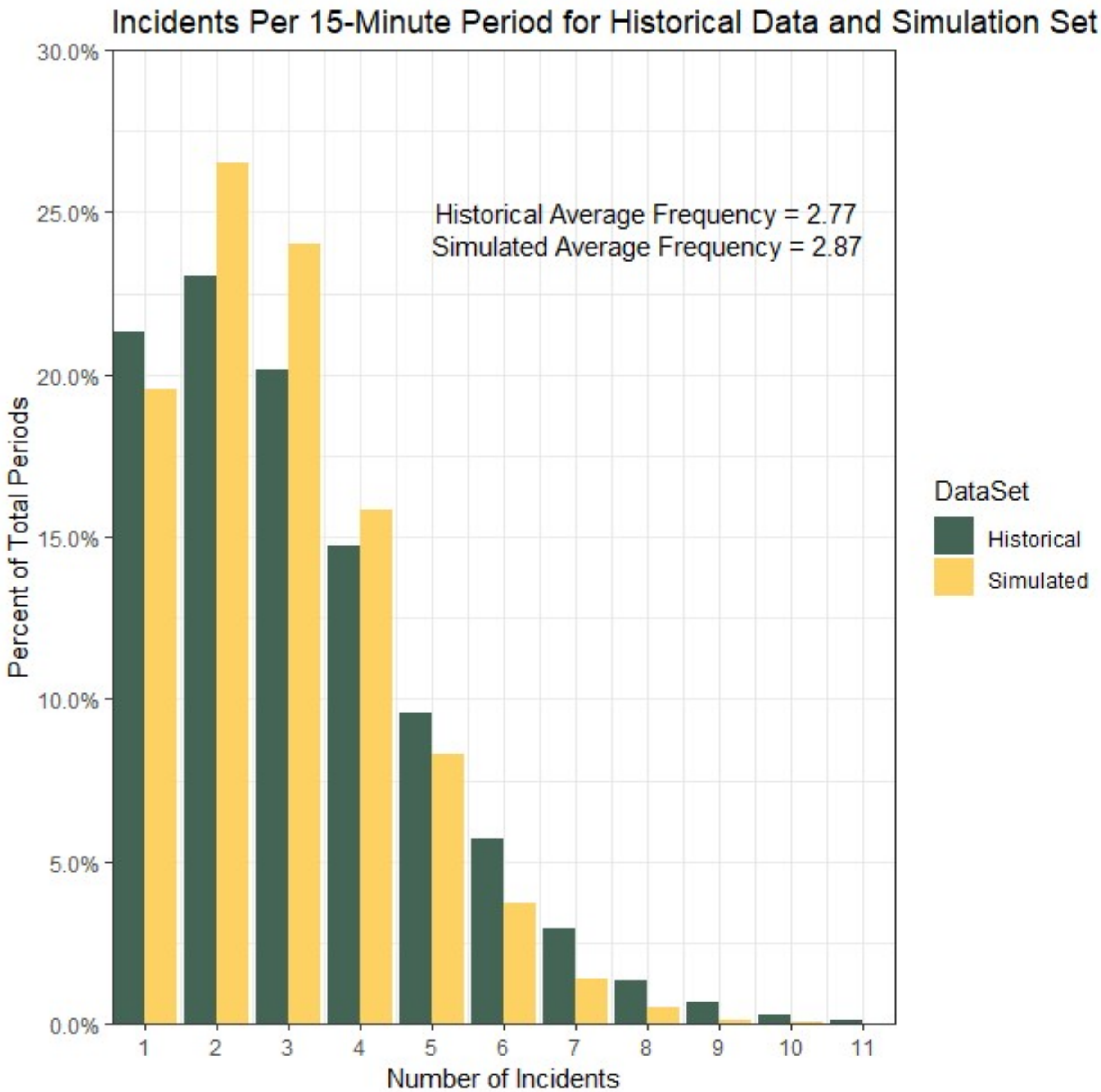

FIGURE 3: PERCENTAGE OF TOTAL 15-MINUTE INTERVALS BY INCIDENT COUNT FOR HISTORICAL DATA AND THE SIMULATION SET

For the simulation set, the percentages were consistent with the theoretical probabilities seen in Table 1. The simulated incidents per period also had fairly good agreement with the actual incidents per period for the 10 months of historical data.

Once the number of incidents for each simulation period was defined, the actual events for those periods had to be assigned. This was done using a random draw from a subset of the historical incident data using R's sample function. Incidents were not allowed to repeat within a single simulation period; however, incidents could be reused in multiple periods.

The subset of historical events was defined by filtering the full historical event set for the events that occurred in the seven boroughs of the "Northern Command" of the London Fire Brigade (Westminster, the City of London, Barnet, Camden, Enfield, Haringey and Islington). Westminster had the largest percentage of incidents in the historical data (at about 8%), with Westminster and the 6 surrounding boroughs representing approximately 26% of the total incidents in the historical data (see Table 2).

TABLE 2: HISTORICAL INCIDENTS FOR NORTHERN COMMAND BOROUGHS

Borough	Total Incidents	% of Total
Westminster	6,155	7.6%
Camden	3,700	4.6%
Barnet	2,787	3.5%
Islington	2,544	3.2%
Enfield	2,456	3.0%
Haringey	2,305	2.9%
City of London	943	1.2%
All Other Boroughs	59,843	74.1%
Total	80,733	100.0%

Figure 4 is a representation of the number of incidents in the simulation set by borough, including the location of the eight fire stations used in the optimization model.

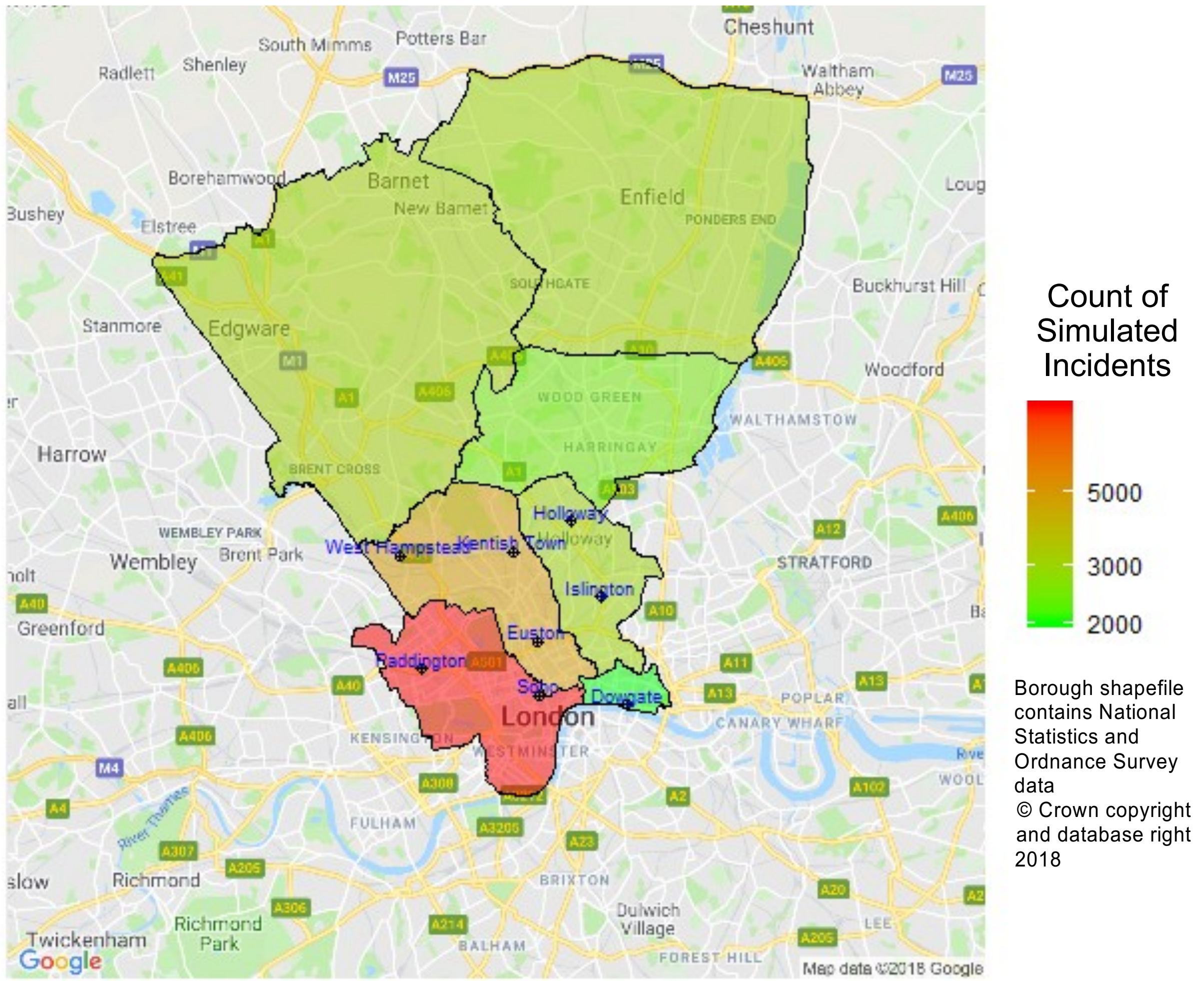

FIGURE 4: SIMULATED INCIDENTS BY BOROUGH FOR 100-DAY SIMULATION (9,600 PERIODS)

Optimization

For each of the 9,600 simulation periods, an Integer Programming (IP) optimization model was run in order to efficiently deploy the available fire engines to each incident. The objective of the optimization was to minimize the total effective distance between the fire stations and the incidents for each period.

This optimization model was implemented in R using the *lpSolve* package.

Inputs

The inputs to the optimization model were the distance matrix between every incident in the simulation period and each fire station, the delay factor matrix, the effective distance matrix and the availability vector.

Distance Matrix, D

d_{IJ} = distance between ith incident and jth fire station

This matrix consisted of distance values between each incident and each fire station. As described in the “Data Preparation” section, the distances were calculated using the Great Circle formula.

TABLE 3: EXAMPLE OF A DISTANCE MATRIX FOR SIMULATION PERIOD 23 (DISTANCES IN KMS)

Incident Number	Dowgate	Euston	Holloway	Islington	Kentish Town	Paddington	Soho	West Hampstead
059625-12052017 16.21	19.18	17.45	13.57	15.84	14.87	19.18	19.11	
038811-31032017 7.12	3.96	3.65	7.54	5.68	6.34	3.90	2.05	
093297-11072017 6.05	3.01	1.81	5.69	3.86	4.62	3.55	0.27	

In Table 3 above, i ranges from 1 to 3 and j ranges from 1 to 8 (which is true for all simulation periods, as there are always eight fire stations in each period).

Delay Factor Matrix, F
f_{IJ} = randomly generated factor (between 0 and 1) to simulate arrival delays

The delay factor was used to simulation delays in the arrival of fire engines to the incidents, such as traffic delays and road blocks. This matrix was the same shape as the Distance matrix (an m-by-8 matrix, where m was the number of incidents in the simulation period). The values in this matrix were randomly generated using the *sample* function in R.

Effective Distance Matrix, E
The effective distance between each incident (i) and fire station (j) was calculated using the formula below.

$$e_{\mathrm{IJ}} = d_{\mathrm{IJ}} + f_{\mathrm{IJ}} d_{\mathrm{IL}} = (1 + f_{\mathrm{IJ}}) d_{\mathrm{IL}}$$

As with the Delay Factor matrix, the Effective Distance matrix was the same shape as the Distance matrix for each simulation period.

Availability Vector, A
a_{J} = number of fire engines available at jth station

The number of fire engines available at each station was obtained from a London Fire Brigade fleet list current as of September 2017. The initial availability vector can be seen in Table 4 below.

TABLE 4: FLEET AVAILABILITY FOR LONDON FIRE STATIONS USED IN THE SIMULATION MODEL

Dowgate	Euston	Holloway	Islington	Kentish Town	Paddington	Soho	West Hampstead	Total Fire Engines
1	1	1	1	2	2	2	2	12

The model assumed fire engines were deployed at the end of each 15-minute period and did not return for 30 minutes (i.e., a fire engine deployed in the previous two simulation periods could not be used in the current period). As such, the availability matrix was updated in each simulation period based on the deployment in the previous two periods.

Decision Variable

The decision or changing variable for the optimization model was whether a fire engine was deployed from a particular fire station j to a particular incident i.

Sent Matrix, S

$$s_{IJ} = \begin{cases} 1, & \text{if fire engine is dispatched to incident i from station j} \\ 0, & \text{if fire engine is not sent to incident i from station j} \end{cases}$$

Constraints

There were two constraints on the optimization model:

1. Exactly one fire engine needed to be deployed to each incident.
2. The total number of fire engines deployed could not exceed the total available fire engines for the simulation period (as mentioned above, the maximum available fire engines for the simulation was 12, but this number was reduced in each simulation period based on the deployment of fire engines in the previous two periods).

These two constraints are represented algebraically below:

1. $\sum_{I} s_{IJ} = 1$
2. $\sum_{J} s_{IJ} \le a_{J}$

Output

The optimization function is shown below:

$$\sum_{I} \sum_{J} s_{IJ} e_{IJ}$$

For each simulation period, this function was minimized by the optimization model.

Discussion of Results

The simulation model produced detailed output for each simulation period. Table 5 below depicts a sample of the output for selected simulation periods (periods 26 through 30). The table shows the incident number (from the historical data) for every incident in each simulation period, as well as the decision variables that show whether or not a fire engine was dispatched to the corresponding incident; a value of 0 indicates a fire engine was not dispatched and a value of 1 indicates a fire engine was dispatched. As laid out in the constraints for the optimization model, every incident has one fire engine dispatched to the incident location.

TABLE 5: SAMPLE DETAILED OUTPUT FROM SIMULATION MODEL FOR PERIODS 26 THROUGH 30

Period ID	Incident Number	Incident Fire Engine Deployment							
		Dowgate	Euston	Holloway	Islington	Kentish Town	Paddington	Soho	West Hampstead
26	071065-03062017	0	0	1	0	0	0	0	0
26	111710-18082017	0	0	0	0	0	0	0	1
26	079544-18062017	0	0	0	0	0	1	0	0
26	141792-21102017	0	0	0	0	0	0	1	0
27	061388-16052017	0	1	0	0	0	0	0	0
27	024617-27022017	0	0	0	0	0	0	1	0
27	116856-29082017	0	0	0	0	1	0	0	0
28	026377-03032017	0	0	0	0	0	1	0	0
28	070188-02062017	1	0	0	0	0	0	0	0
28	039442-02042017	0	0	0	1	0	0	0	0
29	082627-23062017	0	0	0	0	1	0	0	0
29	142806-23102017	0	0	1	0	0	0	0	0
30	026696-04032017	0	0	0	0	1	0	0	0

Table 6 depicts the total number of fire engines dispatched from each fire station for a sample of simulated time periods, as well as the total effective distance traveled between fire stations and each incident in the period. The total effective distance is the value that was minimized by the optimization model.

TABLE 6: SAMPLE SUMMARY OUTPUT FROM SIMULATION MODEL FOR PERIODS 26 THROUGH 30

Period ID	Total Fire Engine Deployment								Total Incidents	Total Effective Distance
	Dowgate	Euston	Holloway	Islington	Kentish Town	Paddington	Soho	West Hampstead		
26	0	0	1	0	0	1	1	1	4	11.43
27	0	1	0	0	1	0	1	0	3	32.46
28	1	0	0	1	0	1	0	0	3	11.23
29	0	0	1	0	1	0	0	0	2	18.08
30	0	0	0	0	1	0	0	0	1	0.41

Sensitivity Analysis

The base analysis was performed using actual fire engine counts from London Fire Brigade's fleet list (as of September 2017). Two sensitivity analyses were performed by adding 1 fire engine and 2 fire engines to the starting fleet of each fire station. For the base analysis, slightly over 400 simulation periods (about 4.4%) did not have a sufficient number of fire engines available to deploy to all incidents for that time period (i.e., there was no feasible solution for the optimization problem). When increasing the starting number of fire engines at each station by 1, there was only one simulation without a feasible solution. The sensitivity analysis with two additional fire engines per station had no infeasible solutions.

Figure 5 below shows the total number of fire engines dispatched during the 9,600 simulation periods from each of the fire stations. The base analysis is depicted by the gray bars, while the sensitivity analyses of adding 1 fire engine per station and adding 2 fire engines per station are shown in green and yellow, respectively.

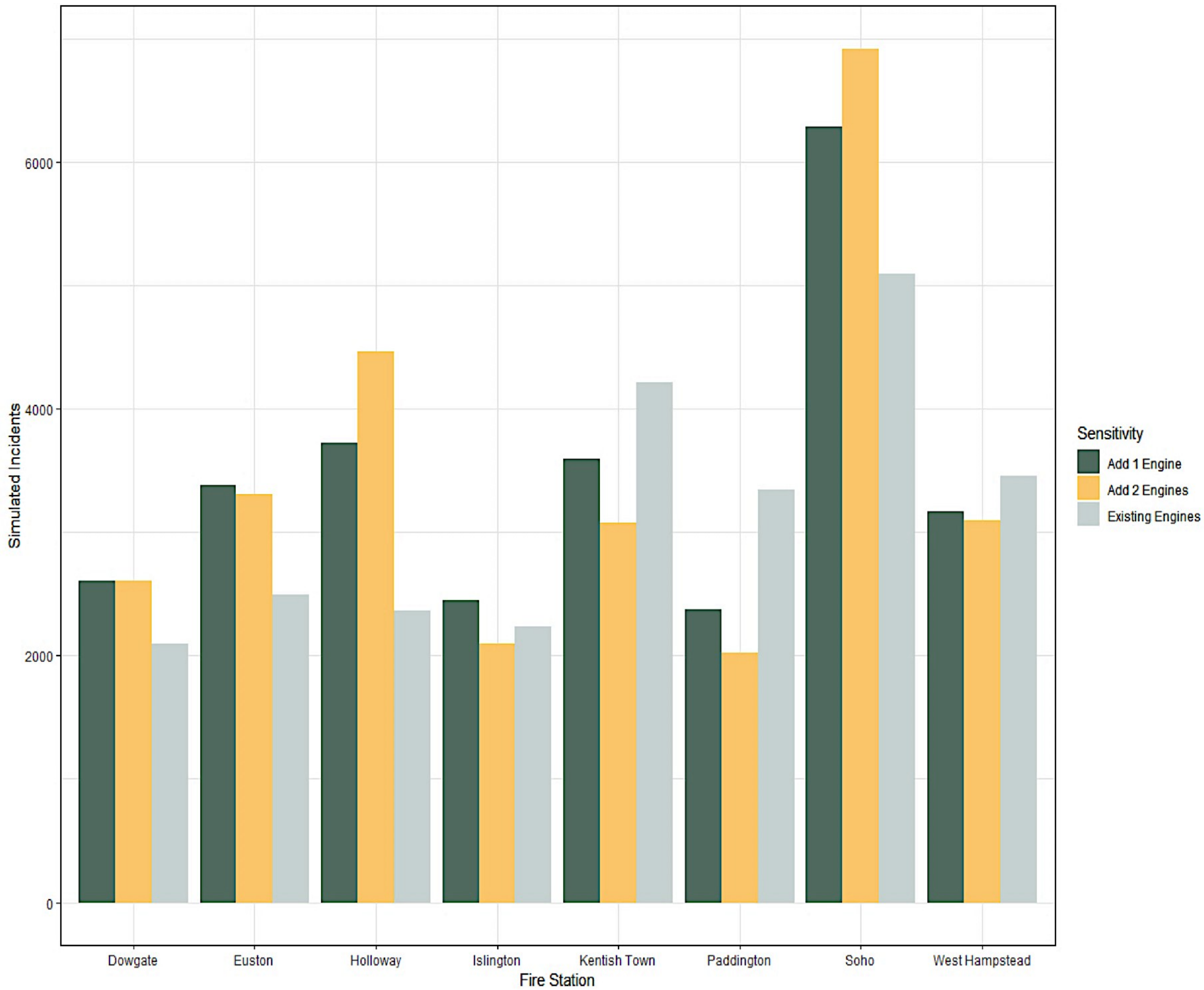

FIGURE 5: SIMULATED OUTPUT OF TOTAL FIRE ENGINES DISPATCHED BY FIRE STATION

As the number of fire engines increased, the Soho fire station, which was already the busiest fire station in the base analysis, had even more fire engines dispatched. Since the total number of incidents remained the same in each sensitivity analysis, as dispatches increased from some stations, other stations necessarily saw decreases in activity. For example, looking back at Figure 3 shows that Soho and Paddington are neighboring stations, and therefore it makes sense that an increase in resources (and therefore dispatched fire engines) at Soho would have an inverse impact at the Paddington. Hence our recommendation to London Fire Brigade would be to have more fire trucks at Soho, perhaps even by relocating one of the engines from Paddington station.

Future Research

While this project provides a solid background to understanding the factors that impact emergency response in London, there are a number of enhancements that can improve the accuracy and reliability of the model.

The first enhancement would to be replace the Great Circle formula in the distance calculations with distance provided by the Google Maps API. By definition, the Great Circle formula calculates distance on a large sphere (such as the Earth) "as the crow flies," which is not how fire engines get to fires. Google Maps calculates driving distance using street routes, even taking into account one-way streets where applicable (e.g., when using the Google Maps API, the distance traveled from a fire station to an incident could differ from the distance traveled in the opposite direction, due to one-way streets, alternate routes, etc.). Therefore, Google Maps would provide more accurate distances than the current method.

The second enhancement would be to refine the assumptions around redeployment of fire engines once they have been deployed to a fire incident. First, queuing theory could be used to create a more accurate prediction of how long a fire engine is deployed before returning to its home fire station, based on the frequency of incidents (which follows a Poisson distribution) and the average deployment time for the actual historical fire incidents (which is available in the data). This could be done as a combined analysis for all fire stations or an individual analysis for each fire station. Second, the assumption that fire engines can only be deployed from a fire station could be relaxed, and the possibility of deploying fire engines from one incident to another could be explored. While this would make the model more complex (requiring tracking of each fire engine and calculations of distances between incidents), it might yield significant improvements in minimizing total effective distance, especially for incidents that are near each other but relatively far from the nearest fire station.

The third enhancement would be to improve the formulation of the effective distance values. The current values are completely random, but traffic and other delays are likely not random. While accidents may be hard to predict, delays such as rush hour and road closures due to construction happen on much more regular schedules. An analysis of historical traffic patterns and delays in London could be used to improve the effective distance (for example, any simulation periods during rush hour could receive an automatic effective distance penalty before any randomness is introduced).

Finally, moving away from the realm of simulation, implementing an operational dashboard would be extremely useful if the London Fire Brigade were to incorporate this model into their real-time emergency response. Real-time locations of fire engines and fire incidents could be displayed for dispatch personnel. The operation dashboard could include a count of fire engines that are at their stations, engines that are on route to a fire, engines that are on their way back from a fire, engines that are currently at a fire and number of incidents awaiting assistance. The main benefit would be inclusion of a map that has both current emergencies and current engine locations.

Conclusion

Simulation of complex systems can lead to insights that potentially have real-world applications. For the London Fire Brigade, optimizing deployment of fire engines to fires and allocating resources to fire stations are two of the key drivers of efficiency. This project was able to show how using simulation and Integer Programming optimization can minimize overall travel distance when dispatching fire engines to incidents. In addition, sensitivity analysis helped to show which fire stations might benefit from additional fire engines, allowing for better allocation of resources.

References

"Great-circle distance." *Wikipedia*, en.wikipedia.org/wiki/Great-circle_distance. Accessed December 6, 2018.

"London Fire Brigade – fleet list." *London Datastore*, data.london.gov.uk/dataset/london-fire-brigade---fleet-list. Accessed October 17, 2018.

"London Fire Brigade Records." *Kaggle*, www.kaggle.com/fayomi/london-fire-brigade-records/version/1. Accessed November 9, 2018.

"London's Fire Stations." *Google Maps*, www.google.com/maps/d/viewer?mid=1f3Kgp7Qx5v0w-sXKomdR8DzD9u4&ll=51.48726183943429%2C-0.11770300000000589&z=11. Accessed October 12, 2018.

"Statistical GIS Boundary Files for London." *London Datastore*, data.london.gov.uk/dataset/statistical-gis-boundary-files-london. Accessed November 8, 2018.

Winston, Wayne L. and S. Christian Albright. *Practical Management Science*. 4th ed., South-Western Cengage Learning, 2012.

"Zero-truncated Poisson distribution." *Wikipedia*, en.wikipedia.org/wiki/Zero-truncated_Poisson_distribution. Accessed December 6, 2018.