# Urban Simulation Assessment

April 22, 2019

## 1 Part 1

### 1.1 Introduction

This is an analysis of resilience in the London tube network. The analysis uses a recursive function for node removal to calculate the marginal effect of a node's removal, pseudocode for the function can be found in Appendix 1. Below, criteria for node removal and effect evaluation are discussed below.

### 1.2 Impact Evaluation

The network effect metric will be discussed first because the node removal criteria was decided in the context of the network effect metric. .

Breaking the network into isolates was investigated but not pursued. Instead, the focus will be on forcing tube users to travel further for longer on their journeys.

This is investigated using shortest path and shortest topological path. Given the spatial nature of the london tubes, where edge attributes represent actual distances, true shortest path might be an attractive option. In the context of the London tube though, total time and effort are more important than total distance. Trains can travel longer distances fairly rapidly while traveling through a high number of stations increases time dramatically because of the need to stop. Further, it is assumed that traveling through a higher number of stops implies a higher number of train changes which are difficult and slow. Thus by using geodesic path, what is being maximized is the increase in stoppage time, and line change time for travelers in the network.

The igraph package's mean_distance function was used to compute the average shortest path between nodes in the network. The unconnected parameter was used to specify that nodes that were not connected to the largest cluster were counted as 1 + the longest possible geodesic, the actual longest geodesic is much less than this. This demonstrates the incompatibility of different network measures. There was not a clear method for comparing the effect of a longer trip with the effect of removing a trip possibility entirely. To do that we would have to include alternate modes of transport like the bus network.

### 1.3 Node Removal Criteria

For each node, degree, betweeness, topological betweenness, closeness, topological closeness and eigenvector centrality were calculated. The correlations for these values across stations can be reviewed in figure 1. It was noted that correlations between weighted

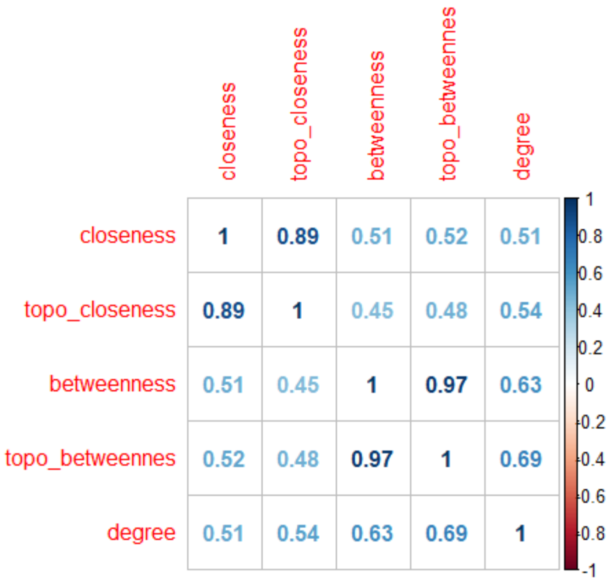


Figure 1: Correlation between station/node metrics

and topological measures were high, indicating that the distances between tube stations are fairly consistent so that the number of stations between two stations is a decent approximation of the distance. This supports the decision to use geodesic longest path. The correlation of measures betweenness and degree is also fairly high, indicating that tube stations at the middle of a line, with higher betweenness, also tend to have multiple lines, high degree. Correlations between eigenvector centrality and the other measures was very low, indicating that this does not give the same information as other metrics. Lastly, it was not clear why correlation between weighted and topological eigenvector centrality was 0.

In order to maximize the increase in travel time measured by the average length of geodesic paths, betweenness will be used to order node removals. This measure is the number of shortest paths between nodes that travel through a given node. Deleting the node with highest betweenness will force the highest number of trips to use an alternate, ostensibly longer, path through other stations.

One note about this process, deleting nodes in some places creates isolates. The function assumes that the distance between unconnected nodes is the longest possible distance on the graph. This will be discussed in greater depth below.

| index | NodeDeleted | IncreaseGeodesic | Components |
|---|---|---|---|
| 1 | Green Park | 0.464271515336074 | 1 |
| 2 | King's Cross St. Pancras | 30.9672711339296 | 2 |
| 3 | Bank | 1.01101854695752 | 2 |
| 4 | Waterloo | 1.13665048051966 | 2 |
| 5 | Stockwell | 17.7264111534033 | 4 |
| 6 | Embankment | 108.61606869706 | 5 |
| 7 | Baker Street | 2.77460558311449 | 6 |
| 8 | Notting Hill Gate | 30.6706717614217 | 7 |
| 9 | Ealing Common | 26.6151971068078 | 9 |
| 10 | Stratford | 10.5750849462714 | 10 |
| 11 | Canning Town | 5.34435028248589 | 12 |
| 12 | Hammersmith | 8.37777866974866 | 14 |
| 13 | Shadwell | 7.82942087022292 | 16 |
| 14 | Harrow-on-the-Hill | 2.89045216902605 | 18 |
| 15 | Camden Town | 2.37783237317933 | 20 |
| 16 | Canary Wharf | 1.95898061029061 | 23 |
| 17 | Mile End | 0.964285951026795 | 24 |
| 18 | Paddington | 0.182272933085244 | 28 |
| 19 | Earl's Court | 0.109301944667777 | 31 |
| 20 | Oxford Circus | -0.440105209296462 | 33 |
| 21 | Woodford | -0.425293708021684 | 34 |
| 22 | Aldgate East | -0.627897731687995 | 35 |
| 23 | Finsbury Park | -0.34940346537536 | 38 |
| 24 | Northfields | -0.418389327214584 | 39 |
| 25 | Wembley Park | -0.540097783642864 | 41 |
| 26 | North Acton | -0.588994666477163 | 43 |
| 27 | Upney | -0.5890868871723 | 44 |
| 28 | Rayners Lane | -0.598069582065818 | 46 |
| 29 | Liverpool Street | -0.601393755502556 | 48 |
| 30 | London Bridge | -0.597733955470574 | 50 |

| Betweenness unconn is false | | | |
|---|---|---|---|
| | Node Deleted | Increase in Avg Geodesic | Components |
| 1 | Green Park | 0.464271515336074 | 1 |
| 2 | King's Cross St. Pancras | 30.9672711339296 | 2 |
| 3 | Bank | 1.01101854695752 | 2 |
| 4 | Waterloo | 1.13665048051966 | 2 |
| 5 | Stockwell | 17.7264111534033 | 4 |
| 6 | Embankment | 108.61606869706 | 5 |
| 7 | Baker Street | 2.77460558311449 | 6 |
| 8 | Notting Hill Gate | 30.6706717614217 | 7 |
| 9 | Ealing Common | 26.6151971068078 | 9 |
| 10 | Stratford | 10.5750849462714 | 10 |
| 11 | Canning Town | 5.34435028248589 | 12 |
| 12 | Hammersmith | 8.37777866974866 | 14 |
| 13 | Shadwell | 7.82942087022292 | 16 |
| 14 | Harrow-on-the-Hill | 2.89045216902605 | 18 |
| 15 | Camden Town | 2.37783237317933 | 20 |
| 16 | Canary Wharf | 1.95898061029061 | 23 |
| 17 | Mile End | 0.964285951026795 | 24 |
| 18 | Paddington | 0.182272933085244 | 28 |
| 19 | Earl's Court | 0.109301944667777 | 31 |
| 20 | Oxford Circus | -0.440105209296462 | 33 |
| 21 | Woodford | -0.425293708021684 | 34 |
| 22 | Aldgate East | -0.627897731687995 | 35 |
| 23 | Finsbury Park | -0.34940346537536 | 38 |
| 24 | Northfields | -0.418389327214584 | 39 |
| 25 | Wembley Park | -0.540097783642864 | 41 |
| 26 | North Acton | -0.588994666477163 | 43 |
| 27 | Upney | -0.5890868871723 | 44 |
| 28 | Rayners Lane | -0.598069582065818 | 46 |
| 29 | Liverpool Street | -0.601393755502556 | 48 |
| 30 | London Bridge | -0.597733955470574 | 50 |

| | By Betweeness | | |
|---|---|---|---|
| | Node Deleted | Increase in Avg Geodesic | Components |
| 1 | Green Park | 0.464271515336074 | 1 |
| 2 | King's Cross St. Pancras | 0.344894786795857 | 2 |
| 3 | Bank | 1.14339690184063 | 2 |
| 4 | Waterloo | 1.28494742196984 | 2 |
| 5 | Stockwell | 0.563343124925417 | 4 |
| 6 | Embankment | -6.16772972588596 | 5 |
| 7 | Baker Street | 2.32031653445306 | 6 |
| 8 | Notting Hill Gate | -1.73581862937826 | 7 |
| 9 | Ealing Common | -2.85804986154228 | 9 |
| 10 | Stratford | -0.357026106908554 | 10 |
| 11 | Canning Town | 0.336399630312155 | 12 |
| 12 | Hammersmith | -1.11269563661233 | 14 |
| 13 | Shadwell | -1.44335210818291 | 16 |
| 14 | Harrow-on-the-Hill | -0.489621630559977 | 18 |
| 15 | Camden Town | -0.565504186081095 | 20 |
| 16 | Canary Wharf | -0.335125486278896 | 23 |
| 17 | Mile End | -0.846532555502438 | 24 |
| 18 | Paddington | -0.328322619604598 | 28 |
| 19 | Earl's Court | -0.196290601211155 | 31 |
| 20 | Oxford Circus | 0.016453621763357 | 33 |
| 21 | Woodford | -0.056258058828067 | 34 |
| 22 | Aldgate East | -0.000392838426682 | 35 |
| 23 | Finsbury Park | -0.101735403043495 | 38 |
| 24 | Northfields | -0.213720988377057 | 39 |
| 25 | Wembley Park | 0.008126154915853 | 41 |
| 26 | North Acton | -0.075498287049283 | 43 |
| 27 | Upney | -0.183419855551612 | 44 |
| 28 | Rayners Lane | -0.120649861972213 | 46 |
| 29 | Liverpool Street | -0.09859581775937 | 48 |
| 30 | London Bridge | -0.079075958422429 | 50 |

| | By Eigenvector Centrality | | |
|---|---|---|---|
| | Node Deleted | Increase in Avg Geodesic | Components |
| 1 | Embankment | 0.124797831125548 | 1 |
| 2 | Cannon Street | -0.020785011623886 | 2 |
| 3 | Moorgate | 0.221126748710891 | 2 |
| 4 | West India Quay | 0.0217058916729 | 2 |
| 5 | Great Portland Street | 0.114088504753353 | 2 |
| 6 | Farringdon | 0.032692113174246 | 3 |
| 7 | Paddington | -0.069066969738513 | 4 |
| 8 | Leicester Square | 0.031071992674539 | 4 |
| 9 | Heron Quays | -0.213633720509144 | 5 |
| 10 | Gloucester Road | 0.404000167032718 | 5 |
| 11 | Euston | -0.232471150322526 | 6 |
| 12 | Aldgate | 0.028208599848918 | 6 |
| 13 | St. James's Park | 0.034699837369548 | 6 |
| 14 | Mile End | 0.776660400757532 | 6 |
| 15 | Oxford Circus | 0.508933395700687 | 6 |
| 16 | Notting Hill Gate | 2.80231756571204 | 7 |
| 17 | Rotherhithe | 0.044124240677434 | 7 |
| 18 | Blackfriars | 0.001798634055188 | 8 |
| 19 | Baker Street | -6.40337680481831 | 12 |
| 20 | Barons Court | -0.217740790392554 | 13 |
| 21 | Aldgate East | 0.046883136212973 | 14 |
| 22 | Ruislip Manor | -0.0251731067831 | 15 |
| 23 | Blackwall | 0.07893822023436 | 15 |
| 24 | King's Cross St. Pancras | -0.928512463963177 | 18 |
| 25 | Island Gardens | 0.016339374551825 | 19 |
| 26 | West Ham | 0.366915223830574 | 21 |
| 27 | Holborn | 0.014000633210076 | 24 |
| 28 | Waterloo | 0.86009000361571 | 24 |
| 29 | Victoria | -0.609084466226621 | 25 |
| 30 | Custom House | -0.385377841072922 | 26 |

| | By Eigenvector Centrality | | |
|---|---|---|---|
| | Node Deleted | Increase in Avg geodesic | Components |
| 1 | Embankment | 0.124797831125548 | 1 |
| 2 | Cannon Street | 5.66059665069182 | 2 |
| 3 | Moorgate | 0.215731180123939 | 2 |
| 4 | West India Quay | 0.02021050051145 | 2 |
| 5 | Great Portland Street | 0.110753998811909 | 2 |
| 6 | Farringdon | 1.91439512290877 | 3 |
| 7 | Paddington | 23.3379220818089 | 4 |
| 8 | Leicester Square | 0.017700366401854 | 4 |
| 9 | Heron Quays | 15.4821648982723 | 5 |
| 10 | Gloucester Road | 0.319473942355302 | 5 |
| 11 | Euston | 34.9321449605953 | 6 |
| 12 | Aldgate | -0.027020815846441 | 6 |
| 13 | St. James's Park | -0.022791930367916 | 6 |
| 14 | Mile End | 0.500638234735106 | 6 |
| 15 | Oxford Circus | 0.310169191636348 | 6 |
| 16 | Notting Hill Gate | 3.48423397547472 | 7 |
| 17 | Rotherhithe | -0.022978065465537 | 7 |
| 18 | Blackfriars | -1.59050475761802 | 8 |
| 19 | Baker Street | 90.790225864921 | 12 |
| 20 | Barons Court | 5.840627408535 | 13 |
| 21 | Aldgate East | 1.495445128136 | 14 |
| 22 | Ruislip Manor | 1.32642824899307 | 15 |
| 23 | Blackwall | -0.342224789480696 | 15 |
| 24 | King's Cross St. Pancras | 15.4170099145231 | 18 |
| 25 | Island Gardens | -1.07286413770476 | 19 |
| 26 | West Ham | 9.42385307392772 | 21 |
| 27 | Holborn | 2.77304884203235 | 24 |
| 28 | Waterloo | -0.370341481560303 | 24 |
| 29 | Victoria | 5.9845029956258 | 25 |
| 30 | Custom House | 2.3713875848423 | 26 |

### 1.4 Analysis

When Kings Cross is deleted, it creates a new unconnected component out of the 11 stations on the north east end of the Picadilly line. In igraph, the two ways to handle this for the mean_distance() function are either to exclude distances between those unconnected nodes and the rest of the network or to assume that the distance is one greater than the longest possible geodesic in the network, that is ,the number of nodes on the network. Excluding distances between unconnected nodes led to a decrease in average trip length because nodes disconnected tended to have higher than average distance to other nodes, lowering the average metric.

Looking at the effect data it seems reasonable to say that betweenness did a better job than eigenvector centrality of prioritizing nodes to remove. Using betweenness created more isolates. It's difficult to judge which method lengthened average shortest path the most because of the options for dealing with disconnected networks. This will be discussed in the conclusion.

### 1.5 Conclusions

To improve this work, it would be good to add data about transportation networks besides the underground. In particular information about bus routes connected nodes would be useful because it would allow for a better estimate of average shortest path when subway stations become disconnected as the shortest path could then go through a bus route instead. Similarly, it would be good to include more granular data about where a rider would have to change trains. The current network assumes there's no cost to switch trains relative to staying on the same train passing through a station. Anyone who has walked from the Picadilly line to the Northern line at Kings Cross knows that there is a big difference.

An improvement to the data generally would be to use travel time data instead of using distance as an approximation.

Lastly, it would be interesting to build an igraph function that can compute average shortest path using edge weights since the current function cannot. This could confirm or reject the thought that tube stations are spaced fairly regularly based on the high correlation between weighted and topological centrality measures.

**943 words**

# 2 Part 2

## 2.1 The Models

### 2.1.1 Unconstrained

The model is constrained to the total flows of the system but flows out of an origin and into a destination can be any value between 0 and total system flows.

This is useful for studying the change in connectivity between regions, for instance if a new transportation link was built. In particular, it is useful for studying long term effects of a change where residence and employment are more flexible.

### 2.1.2 Production

The direction of flows can change but the total flows from each origin will remain constant. This is useful for studying the effect of a new employment or consumption location that changes the destinations of people going to work or to spend money. In terms of the matrix, it implies that the sums of the rows of the matrix are constant.

### 2.1.3 Attraction Constrained

The source of flows into a region can change but the total flows into a region will remain constant. That is, any reduction in flows into a destination from another region will be fully replaced by flows into the destination from another region. This could be used to study a new housing development that pulls people into residence in a different part of an area or a natural disaster that forces residents out of an area. Employers outside the area still need workers but will not be able to draw them from the same places after a housing change or natural disaster. In terms of the matrix the sums of the columns are held constant. Additionally, it can be used to study the effect of a specific change to employment where the model can be constrained to the values that result from that change.

### 2.1.4 Doubly Constrained,

Doubly constrained models could be used to test the short term effects of a change to transportation networks given that homes and businesses won't relocate but flexible behavior patterns like shopping could change almost immediately due to the change in accessibility or travel times between locations. In this model, the sums of both the columns and rows are held constant.

## 2.2 The Parameters

Basically, the parameters are ratios of how much a change of 1, in the natural logarithm of one of the predictor variables affects the natural log of the estimate of flow. This is seen below, 1-3 for the unconstrained and 4-6 for the doubly constrained. The equation is log-linearized and then the flow estimate is solved for. In equation 6, $A_i$ and $B_i$ are vectors of parameters that allocate parts of the regions total flows across origins and destinations and $k$ is an arbitrary intercept used for estimation.

$$T_{ij} = kV_i^{\mu} W_j^{\alpha} d_{ij}^{-\beta} \tag{1}$$

$$ln(T_{ij}) = ln(k) + \mu(ln(V_i)) + \alpha(ln(W_j)) - \beta(ln(d_{ij})) \tag{2}$$

$$T_{ij} = e^{\{ln(k)+\mu(ln(V_i))+\alpha(ln(W_j))-\beta(ln(d_{ij}))\}} \tag{3}$$

$$\lambda_{ij} = A_i O_i B_j D_j d_{ij}^{-\beta} \tag{4}$$

$$ln(\lambda_{ij}) = ln(A_i O_i B_j D_j) - \beta ln(d_{ij}) \tag{5}$$

$$\lambda_{ij} = A_i O_i B_j D_j - e^{\beta ln(d_{ij})} + k \tag{6}$$

### 2.3 A Scenario

*select a scenario and explore the consequences of varying model parameters and inputs on interaction flows and the origin/destination estimates*

The scenario used for this assessment will be: "What if teleportation was invented and dramatically reduced travel times in connected boroughs but could only be used in London's outermost boroughs due to construction requirements?" Thus origin and destination attributes remain constant but the travel costs change dramatically in the most peripheral boroughs: Hilingdon, Harrow, Barnet, Enfield, Waltham forest, Redbridge, Havering, Bexley, Bromley, Croydon, Sutton, Kingston, Richmond, Hounslow.

This will be investigated using a doubly constrained model to estimate the short term effects where residences and businesses cannot relocate and a total constrained model to see the long term effects on business and residence locations as a result of the incredible new discovery.

To approximate the effect of near instantaneous transportation between two boroughs, the lowest distance between centerpoints of two London boroughs, 2080 meters, is substituted for the real distance between each of the 9 outermost boroughs. .

### 2.4 Setup

The two models were estimated in R.

The parameters and goodness of fit statistics for the unconstrained model are seen in table 1.

For the doubly constrained model, the estimated intercept was 24.63 and distance parameter was $-1.92$. This model also produces a $32X32$ matrix of origin and destination specific parameters that will not be included here.

In both models the signs of the estimated parameters were consistent with expectations and the majority of parameters were significant at the 0.0001 level.

| Parameter | | Fit | |
|---|---|---|---|
| $k$ : | $-12.5$ | RMSE : | 2330.9 |
| $\mu$ : | 1.62 | $R^2$ : | 0.386 |
| $\alpha$ : | 1.55 | | |
| $\beta$ : | $-1.5$ | | |

Table 1: Total constrained model results

### 2.5 Results

## 3 Part 3

### 3.1 Overview

While CA models are often viewed as separate from ABM, a more modern view of these methods considers them to be cousins or related in the sense that CA are a subset of ABM.

Expanding a CA model usually results in the creation of an ABM. Cellular automata models are defined by a set of homogeneous cells that interact only with each other according to a defined set of input/output functions, e.g., if a cell with value 1 is surrounded by other cells f value 1 it's value becomes 0. The models can be extended to "n" states but all cells should be capable of reaching all n states to maintain the homogeneity of the cells. This can be contrasted with ABM where cells can have infinite heterogeneity and future states can be functions of the "environment" other cells or agents" and the cells own state.

The simplicity of CA models make them very useful for studying mathematical processes whereas Agent Based Modeling flexibility make them useful for modeling more complex "real world" phenomena and make them more accessible to non-technical audiences. Often the value of ABM comes from the ability to conduct parameter sweeps, to study combination rules that apply to multiple conceptual processes with different parameter values. The value of cellular automata models tends to be focused on the effect of initialization states on the long term outcomes and equillibria of the model e.g. for the

| | Borough | 2 | 3 | 5 | 6 | 7 | 8 | 10 | 15 | 17 | 18 | 19 | 21 | 26 | 27 | 29 | 31 | 33 | 34 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | Barnet | 0 | 34 | 76 | 12080 | 7709 | 148 | 4098 | 2623 | 1023 | 611 | 5775 | 68 | 305 | 229 | 44 | 555 | 16330 | 74391 |
| 3 | Bexley | 132 | 0 | 4998 | 2470 | 6580 | 710 | 123 | 29 | 109 | 161 | 1557 | 77 | 111 | 90 | 170 | 222 | 7692 | 51231 |
| 5 | Bromley | 162 | 3199 | 0 | 3780 | 9855 | 6268 | 84 | 59 | 248 | 285 | 2094 | 227 | 100 | 191 | 796 | 196 | 12802 | 67450 |
| 6 | Camden | 1496 | 32 | 60 | 0 | 8795 | 147 | 295 | 330 | 473 | 473 | 4987 | 89 | 84 | 195 | 54 | 204 | 18829 | 51652 |
| 7 | City | 14 | 0 | 0 | 335 | 0 | 3 | 6 | 7 | 10 | 9 | 339 | 0 | 0 | 6 | 3 | 16 | 602 | 2062 |
| 8 | Croydon | 204 | 300 | 5152 | 3248 | 5925 | 0 | 120 | 97 | 457 | 581 | 1752 | 827 | 64 | 480 | 6744 | 130 | 10583 | 64539 |
| 10 | Enfield | 5642 | 52 | 76 | 5588 | 5212 | 136 | 0 | 328 | 457 | 339 | 5317 | 38 | 538 | 98 | 47 | 1710 | 9052 | 56955 |
| 15 | Harrow | 5008 | 26 | 47 | 3675 | 3162 | 103 | 325 | 0 | 6169 | 1141 | 1395 | 107 | 44 | 246 | 34 | 95 | 7882 | 49985 |
| 17 | Hillingdon | 692 | 18 | 55 | 1757 | 1661 | 116 | 153 | 4688 | 0 | 5293 | 664 | 150 | 35 | 548 | 58 | 53 | 5062 | 37054 |
| 18 | Hounslow | 253 | 27 | 69 | 1753 | 2342 | 215 | 92 | 396 | 12803 | 0 | 752 | 1006 | 31 | 7025 | 101 | 41 | 5927 | 48403 |
| 19 | Islington | 1001 | 46 | 46 | 10188 | 7931 | 157 | 619 | 111 | 265 | 334 | 0 | 50 | 117 | 118 | 56 | 393 | 12835 | 50391 |
| 21 | Kingston | 104 | 35 | 72 | 1547 | 2875 | 638 | 42 | 68 | 1070 | 1484 | 710 | 0 | 31 | 3788 | 1190 | 20 | 5419 | 30687 |
| 26 | Redbridge | 567 | 141 | 132 | 3790 | 8205 | 173 | 1187 | 110 | 212 | 231 | 3104 | 46 | 0 | 88 | 48 | 5441 | 8122 | 61005 |
| 27 | Richmond | 173 | 12 | 64 | 2504 | 4831 | 323 | 41 | 167 | 3365 | 6873 | 1002 | 3549 | 12 | 0 | 260 | 46 | 8336 | 45329 |
| 29 | Sutton | 81 | 68 | 514 | 1450 | 2560 | 7602 | 45 | 50 | 385 | 541 | 681 | 3122 | 26 | 750 | 0 | 41 | 4691 | 39635 |
| 31 | Waltham | 715 | 79 | 149 | 5554 | 6337 | 253 | 2661 | 112 | 206 | 271 | 4310 | 50 | 3736 | 101 | 41 | 0 | 10314 | 57814 |
| 33 | Westminster | 514 | 17 | 99 | 6786 | 9521 | 216 | 121 | 142 | 597 | 572 | 2442 | 90 | 60 | 258 | 61 | 121 | 0 | 39288 |
| 34 | (all) | 30744 | 11740 | 24934 | 147985 | 209961 | 31087 | 18456 | 18423 | 48351 | 39292 | 86387 | 17043 | 21033 | 23230 | 16652 | 17935 | 353405 | 1800413 |

| | Orig | 2 | 3 | 5 | 6 | 7 | 8 | 10 | 15 | 17 | 18 | 19 | 21 | 26 | 27 | 29 | 31 | 33 | (all) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | Barnet | 0 | 148 | 351 | 12838 | 6276 | 401 | 2736 | 2502 | 2402 | 992 | 4833 | 286 | 474 | 464 | 188 | 735 | 15460 | 74392 |
| 3 | Bexley | 539 | 0 | 3439 | 2534 | 5375 | 1207 | 478 | 232 | 643 | 422 | 1720 | 269 | 1112 | 267 | 322 | 557 | 6234 | 51230 |
| 5 | Bromley | 701 | 1878 | 0 | 3480 | 7071 | 6975 | 499 | 352 | 1106 | 831 | 2097 | 686 | 684 | 573 | 1132 | 472 | 9748 | 67450 |
| 6 | Camden | 792 | 43 | 108 | 0 | 5967 | 129 | 265 | 198 | 350 | 229 | 7417 | 83 | 116 | 133 | 60 | 187 | 26056 | 51654 |
| 7 | City of London | 15 | 3 | 8 | 226 | 0 | 9 | 9 | 5 | 11 | 8 | 244 | 4 | 8 | 5 | 4 | 10 | 715 | 2061 |
| 9 | Ealing | 1828 | 111 | 355 | 5961 | 4185 | 534 | 488 | 2863 | 10491 | 7811 | 2007 | 740 | 218 | 2131 | 319 | 251 | 14572 | 78841 |
| 14 | Haringey | 2577 | 115 | 234 | 11521 | 7371 | 239 | 2227 | 389 | 642 | 346 | 9869 | 131 | 516 | 188 | 101 | 1398 | 11169 | 63082 |
| 16 | Havering | 725 | 1326 | 1111 | 2673 | 4526 | 664 | 844 | 286 | 729 | 415 | 1823 | 220 | 3072 | 240 | 218 | 950 | 5687 | 45622 |
| 17 | Hillingdon | 1110 | 81 | 256 | 2619 | 2168 | 366 | 354 | 2483 | 0 | 3985 | 1033 | 476 | 157 | 968 | 210 | 168 | 5915 | 37052 |
| 18 | Hounslow | 710 | 83 | 298 | 2652 | 2449 | 513 | 253 | 822 | 6171 | 0 | 1025 | 1240 | 137 | 5981 | 356 | 144 | 7397 | 48402 |
| 20 | Kensington and Chelsea | 204 | 22 | 70 | 2830 | 2171 | 106 | 73 | 98 | 235 | 223 | 788 | 87 | 42 | 158 | 57 | 52 | 27916 | 46336 |
| 25 | Newham | 503 | 592 | 638 | 2859 | 7780 | 435 | 520 | 171 | 408 | 259 | 2519 | 140 | 1985 | 160 | 143 | 1135 | 6245 | 51656 |
| 26 | Redbridge | 1022 | 657 | 740 | 4050 | 7046 | 532 | 1564 | 325 | 732 | 413 | 3216 | 201 | 0 | 238 | 187 | 3977 | 7708 | 61005 |
| 28 | Southwark | 381 | 202 | 629 | 3480 | 17672 | 724 | 234 | 158 | 418 | 340 | 2367 | 220 | 257 | 246 | 255 | 259 | 13438 | 67576 |
| 29 | Sutton | 332 | 156 | 1004 | 1722 | 2763 | 5849 | 177 | 209 | 806 | 881 | 852 | 1922 | 154 | 819 | 0 | 136 | 5635 | 39632 |
| 31 | Waltham Forest | 1246 | 259 | 401 | 5134 | 7561 | 341 | 2435 | 308 | 616 | 340 | 4779 | 150 | 3126 | 192 | 130 | 0 | 8131 | 57812 |
| 33 | Westminster | 386 | 43 | 122 | 10540 | 7643 | 164 | 151 | 146 | 319 | 258 | 2750 | 107 | 89 | 171 | 79 | 120 | 0 | 39285 |
| 34 | (all) | 30742 | 11739 | 24931 | 147985 | 209961 | 31087 | 18453 | 18423 | 48349 | 39291 | 86385 | 17045 | 21035 | 23231 | 16652 | 17936 | 353403 | 1800399 |

| | Orig | 2 | 3 | 5 | 6 | 7 | 8 | 10 | 15 | 17 | 18 | 19 | 21 | 26 | 28 | 29 | 31 | 33 | All |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | Barnet | 0 | 634 | 1583 | 2728 | 11006 | 2086 | 140 | 124 | 845 | 1188 | 2212 | 858 | 664 | 943 | 1050 | 403 | 14096 | 74388 |
| 3 | Bexley | 935 | 0 | 69 | 5859 | 5447 | 293 | 340 | 565 | 1337 | 1183 | 2635 | 387 | 120 | 695 | 260 | 225 | 14818 | 51234 |
| 5 | Bromley | 1421 | 42 | 0 | 8437 | 8190 | 100 | 644 | 737 | 1538 | 1189 | 4275 | 300 | 385 | 640 | 146 | 525 | 18744 | 67452 |
| 6 | Camden | 564 | 822 | 1942 | 0 | 4349 | 2425 | 544 | 588 | 2182 | 1937 | 542 | 1112 | 1020 | 1239 | 1233 | 595 | 3143 | 51644 |
| 7 | City | 64 | 22 | 53 | 123 | 0 | 73 | 35 | 51 | 147 | 117 | 35 | 49 | 33 | 66 | 43 | 23 | 242 | 2060 |
| 8 | Croydon | 1469 | 140 | 79 | 8263 | 8821 | 0 | 791 | 675 | 1269 | 814 | 4668 | 138 | 631 | 382 | 30 | 728 | 16424 | 64547 |
| 10 | Enfield | 166 | 273 | 848 | 3109 | 7019 | 1327 | 0 | 252 | 1010 | 1271 | 1509 | 712 | 165 | 920 | 750 | 79 | 13732 | 56954 |
| 15 | Harrow | 126 | 391 | 838 | 2901 | 8815 | 979 | 218 | 0 | 100 | 273 | 2426 | 298 | 555 | 283 | 442 | 433 | 9894 | 49985 |
| 17 | Hillingdon | 228 | 245 | 463 | 2851 | 6792 | 487 | 231 | 27 | 0 | 63 | 2207 | 110 | 428 | 96 | 200 | 376 | 7855 | 37057 |
| 18 | Hounslow | 533 | 360 | 594 | 4204 | 8980 | 519 | 483 | 120 | 105 | 0 | 3321 | 63 | 731 | 23 | 176 | 656 | 9381 | 48406 |
| 19 | Islington | 742 | 599 | 1596 | 878 | 1994 | 2223 | 428 | 797 | 2738 | 2482 | 0 | 1242 | 636 | 1546 | 1234 | 317 | 5964 | 50382 |
| 21 | Kingston | 541 | 165 | 210 | 3390 | 5258 | 124 | 379 | 184 | 256 | 88 | 2334 | 0 | 439 | 24 | 24 | 434 | 6611 | 30690 |
| 26 | Redbridge | 700 | 86 | 452 | 5202 | 5897 | 945 | 148 | 574 | 1668 | 1718 | 2000 | 734 | 0 | 1108 | 633 | 45 | 17008 | 61008 |
| 28 | Southwark | 2452 | 363 | 695 | 7902 | 3069 | 906 | 1285 | 1536 | 3812 | 2725 | 3546 | 874 | 959 | 1393 | 606 | 898 | 12734 | 67574 |
| 29 | Sutton | 862 | 145 | 133 | 4897 | 6021 | 34 | 521 | 356 | 607 | 322 | 3021 | 31 | 493 | 129 | 0 | 525 | 9314 | 39642 |
| 31 | Waltham | 521 | 197 | 757 | 3724 | 4987 | 1337 | 86 | 548 | 1799 | 1891 | 1222 | 891 | 55 | 1243 | 827 | 0 | 14634 | 57810 |
| 33 | Westminster | 1009 | 720 | 1494 | 1088 | 2960 | 1669 | 832 | 694 | 2081 | 1497 | 1273 | 751 | 1155 | 836 | 812 | 810 | 0 | 39280 |
| 34 | (all) | 30744 | 11739 | 24936 | 147983 | 209959 | 31086 | 18460 | 18424 | 48349 | 39289 | 86389 | 17043 | 21032 | 23229 | 16656 | 17936 | 353404 | 1800416 |

same model, outcomes are a function of the rule set and initialization values, where as agent based models tend to be calculated for a large number of initialization values in order to study the effect of model dynamics independent of initialization values that may not accurately reflect the real world.

## 3.2 Scenarios

This work will investigate three possible infection scenarios. The first is a baseline scenario used to explore a moderately recoverable disease to which immunity is low, for instance the seasonal flu without a vaccination program. The second scenario looks at the same disease with a vaccination program. The third looks at a disease like the common cold which is more recoverable but where immunity is close to non-existent.

All three scenarios will use a population of 1000 and infect 100 people initially.

### 3.2.1 Scenario 1

The first scenario, looking at a flu type disease with no vaccine available uses 50% chance of immunity, based on the idea that it's fairly common for someone with the flu to interaction with someone else without infecting them and a 10% chance of recovery based on the guess that flu symptoms last between 3 and 9 days for the majority of the population.

Looking at Figure 2, which plots the percentage of turtles infected over 150 steps of the model for 20 trials, a steady state becomes fairly clear. While this is not a perfect equilibrium, compared to the rapid increase in the first 20 ticks, the model becomes fairly stationary. This, future runs of this model can be limited to 40 ticks with reasonably confidence that the model will have found an equilibrium.

The average steady state percentage of turtles infected at 50 ticks was 38.79. The sample standard deviation was 1.799, which indicates a margin of error at 95% confidence of $\pm 0.842$ using a t-value of 2.093. This implies that to get a margin of confidence of 0.5, 56.7 trial runs would be required. When 56 trials were run, the result was $37.93 \pm 0.493$, so the estimate for $n$ was pretty accurate.

The total infection time per non-immune turtle per tick was also recorded. The 20 trial mean at tick 40 was $0.6343 \pm 0.00669$. This has the value of adjusting for turtle immunity and runtime in order to compare results across scenarios later on.

Thus the model indicates that for these parameters, a fairly high number of turtles will contract the infection and the infection will be passed around continually rather than dying out at some point in the future. Considering the margin of errors for the results, if the model accurately reflects a real world scenario, we can have a good idea of how that scenario would play out.

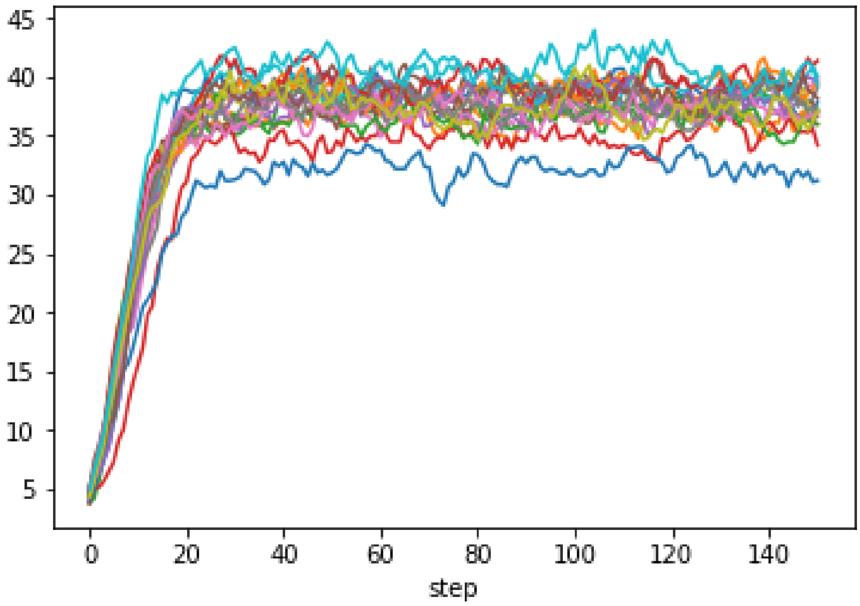


Figure 2: % of Population Infected Over Time
Scenario 1

#### 3.2.2 Scenario 2

Scenario 2 looks at the flu in a society with high but imperfect immunization rates. In the US this is seen in cases where infants and the elderly cannot receive the vaccine and some people elect not to receive it or forget. This is accomplished by using the same recovery rate, 10% but increasing immunity probability to 80%.

In Figure 3 the distribution over time of the percentage of turtles infected in the scenario is illustrated. It can be seen that a steady state is more difficult to identify. One consideration is that the average percentage is lower than in scenario 1 because 80% of the population is immune. A second is that the lower mean scales the y-axis of the chart down from 0-14 instead of 0-45 in scenario 1. Thus while the standard deviation looks higher it is essentially the same, 1.869 in scenario 1 and 1.853 in scenario 2.

Because a steady state is less certain looking at the chart a t-test can be used to confirm that the mean of the set of trials is not changing between steps. At step 50 the mean percent infected was 5.92 with st. deviation 1.799. At step 100 these were 6.90 and 1.99. At step 250 7.935 and 1.696. At step 400 8.075 and 1.502. At step 500 7.956 and 2.113.

The p-value for a t-test with the null hypothesis that the means at steps 50 and 100 are different was 0.116. Indicating a 12% chance that they are different and that a steady state has not been reached with 95% confidence. The same test for the means at tick

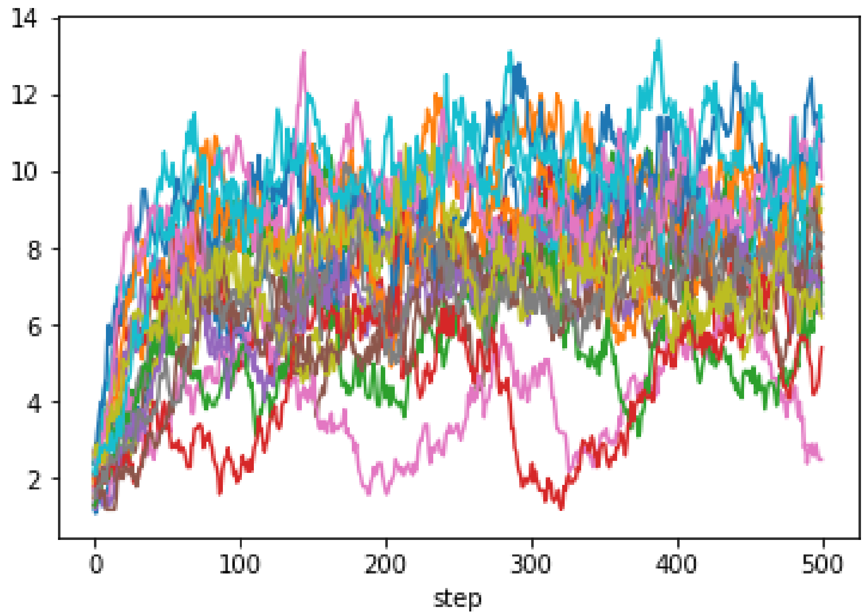


Figure 3: % of Population Infected Over Time
Scenario 2

250 and tick 500 had a p-value of 0.97. The test for steps 400 and 500 had a p-value of 0.84. Thus it is difficult to say with confidence that a steady state exists for this model as the effort to reject the null hypothesis that the means are different begins to resemble simply getting lucky enough to choose ticks with means that are very similar to each other by coincidence.

Sicktime per turtle per tick at 100 ticks was $0.1047 \pm 0.01397$. A lower average sicktime compared to scenario 1 with not as tight confidence interval.

at tick 250, the mean percentage of turtles infected was $7.94 \pm 0.794$ again demonstrating that the variability in results for this scenario is very similar to those of scenario 1 but that in light of the results of a t-test for a steady state, the true margin of error is likely to be higher since the process may not be stationary at tick 250, that is to say, the true mean may change across ticks. Finally, smaller sets of trials exhibited similar results to those calculated in scenario 1 showing that the change in parameters affected the value of the results but not their variability.

#### 3.2.3 Scenario 3

This scenario will investigate how the common cold compares to the flu by using a higher rate of recovery 25% and lower rate of immunity 5% since there is no vaccine for the common cold and many members of society continue their daily routines after contracting it exposing relatively more people than the flu, which most people suffer at

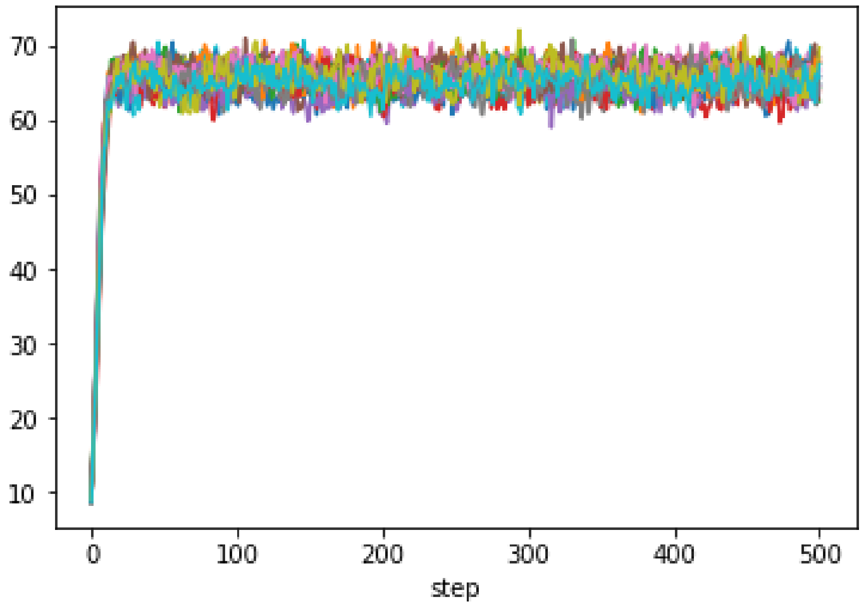


Figure 4: % of Population Infected Over Time
Scenario 3

home.

Figure 4 illustrates the percentage of turtles infected over time. Here a steady state seems more obvious. At 50 ticks the mean was 66.19 and st. deviation was 1.597. At 100 ticks this was 65.84 and 1.648. The 0.04994 p-value for a two tailed t-test for difference in means indicates only a 5% probability of the null hypothesis that the means are different, just on the border of an acceptable level of confidence but basically acceptable.

At 50 ticks the mean percentage of turtles infected was $66.08 \pm 0.902$ . Thus the low variability of outcomes for different outcome values as a result of different parameter values continues.

Results for sick time per turtle per tick at tick 50 were $0.311 \pm 0.0013$. Less than the no vaccine flu scenario. which was less recoverable but more than the vaccinated scenario.

### 3.3 Conclusion

Comparing the three scenarios, The unvaccinated flu, scenario 1, (moderate immunity (50%) and moderate recovery rates (10%)) resulted in a 38% average infection rate while the vaccinated flu (80% immunity and 10% recovery) resulted in 7.93% of turtles infected on average. The common cold, (5% immunity, 25% chance of recovery) gave 66% of turtles infected. At tick 50, non-immune turtles in scenario 1 had spent 63% of ticks infected on average while this number was 10% in scenario 2 and 31% in scenario 3, demonstrating the value of population immunity to non-immune members.

Scenario 1 turtles that were not immune had spent $23.6 \pm 0.12$ ticks sick on average while this was $7.54 \pm 0.45$ in scenario 2 and $24.14 \pm 0.05$ in scenario 3.

Thus it is clear that higher rates of vaccination in scenario 2 lead to lower amounts of sicktime for turtles without the vaccine. At a high level it can then be said that this model replicates the empirical idea of "herd immunity" where general immunity of a group protects non-immune members of the group.

To continue investigating this, the model could be explored to find whether there is a constant increase in sick time per non-immune turtle or whether as population immunity decreases sick time accelerates. It would also be useful to add in a parameter for the contagiousness of the infection, the radius within which turtles affect each other. Lastly, adding in death or post-infection immunity would be an interesting extension.

**1618 words**

## References

Grimm, Volker et al. (2010). "The ODD protocol: A review and first update". eng. In: *Ecological Modelling* 221.23, pp. 2760–2768. ISSN: 0304-3800.

Tasseron, G. and K. Martens (2017). "Urban parking space reservation through bottom-up information provision: An agent-based analysis". In: *Computers, Environment and Urban Systems* 64, pp. 30–41. ISSN: 01989715.