# Assessing time resilience of public transit networks using London Underground data

Jakob Serlier
July 2020

*Figure 0: Changing impact of disruption events on capacity utilization for different timeslots*

## 1. Abstract

The motivation behind this study is to explore the importance of disruption time in public transport network resilience, and how to measure the changing impact over time. We have proposed a methodology to model a non-cascading time-dependent network model, estimating changing passenger loads in a public transport network. Our models, created using empirical passenger data from the London Underground, show that the network is not only most vulnerable during peak-hours to increased passenger loads, but in addition, the impact disruptions on the network change and the highest during peak-hours at certain locations. One of the goals of this research was to explore the topological metrics which identify time-dependent critical network links. In this study, betweenness has shown to be a valuable indicator of link criticality when using our model approach. Considering capacity utilization as an important performance metric, disruptions of high-impact non-bridge links showed a small decrease in capacity utilization in preceding and succeeding links and a more significant increase in capacity utilization for parallel links. This parallel effect was less significant in terms of capacity utilization for non-bridge links which have a low betweenness centrality, instead effecting the capacity utilization of neighboring links in a more diffuse manner. Finally, we expected to see a spatial change in the effect of disruption events over time, but the analysis of our model results did not reflect this expectation. The results primed exciting directions for future research, and has given us valuable initial insights on time dependency of public transport networks.

Table of contents

## 2. Introduction

This research aims to give broader insights of time-dependent resilience of public transport systems, by analyzing public Transport for London (TFL) data. Understanding the resilience and workings of public transportation systems are of key importance to decision-makers such as policy-makers, transport authorities and analysts, as it allows for more efficient and effective response to disruption events (Berche et al., 2009). The safety and security of public transport systems are seen as an important factor by passengers (Smart et al., 2009). Subsequently, improvements in network resilience will encourage passengers to make use of available public transport systems, which results in various benefits such as decreasing the disparity between social groups (Gates et al., 2019; Hernández, 2018) and decreasing emissions (Woodcock et al., 2007; Geng et al., 2013).

Public transport systems around the world can be perceived as networks of varying, but great complexity (Lin & Ban, 2013). Therefore, understanding the inner workings of these networks allows us to better understand the public transport systems as a whole. A starting point is defining public transport resilience, as this will give us improved insights on how to measure the impact of disruption events.

### *2.1 Defining resilience*

Defining and measuring resilience in its many forms is a topic of ongoing discussion in research. Gauthier et al. roughly defines resilience as the ability of a network to absorb the impact of a disruption event (Gauthier et al., 2018). As this perspective fits well in respects to resilience of a public transport systems, we will use it as a broad view of resilience in this paper. However, a more deconstructed view on the definition will allow for additional insights.

A different view of resilience is discussed in a seminal paper by Bruneau et. Al, proposing that resilience is shaped by four dimension; *robustness, redundancy, resourcefulness* and *rapidity* (Bruneau et al., 2003). For the purpose of this discussion, we will focus on two specific dimensions of resilience of a public transport system, the robustness and redundancy, as they are describing for the current-state of a network. Resourcefulness and rapidity are both valid segments of resilience, but focus on future disruptions and problem identification. The focus will lie on robustness and redundancy, which have been prominent in resilience and similar-themed transport papers (O. Cats, 2016; Jenelius & Cats, 2015; Sullivan et al., 2010; Tahmasseby & van Nes, 2007).

Capacity utilization and passenger journey time are valuable performance indicators of public transport performance (Orth et al., 2012; Thompson & Schofield, 2007; Wu et al., 2020). These performance estimators can be aligned with the resilience dimensions. The effect of a disruption event on capacity utilization can be seen as a factor of *redundancy:* the extent to which resources in a network can satisfy additional functional requirements in the event of a disruption, and the impact of disruption events on passenger journey time can be seen as a factor of *robustness:* the degree to which the network can function during a disruption event. By underlining these two dimensions and indicators, we are able to more effectively focus on our research goal of understanding the impact of disruption events on public transport networks.

*2.2 Network vulnerability and time-dependence*

Decision-makers need to make quick and effective decisions when a disruption occurs. By measuring network vulnerabilities using the performance indicators, alongside how these change over time, we are able to gain insights in time-dependence of network resilience. This approach allows decision-makers to create preventive measures, decreasing the impact of disruption events on the network. In order to analyze the network vulnerabilities, topological metrics can be compared to the performance indicators to determine which nodes and links are most important to the resilience of the network (Gauthier et al., 2018). Topological and performance metrics to measure criticality vary widely, and a topic which requires more research (Jafino et al., 2020; Taylor & D'Este, 2003).

The timing of a disruption event greatly influences its impact, as network vulnerabilities change throughout the day. Previous research regarding the resilience of public transport networks often did not consider time dependence. Critical locations in a network change over time due to the dynamic nature of passenger behavior and load. Most importantly, during peak hours, the travel demand increases. To what extent a disruption event will affect the performance metrics such as travel time and available capacity is therefore dependent on the timing of the disruption event.

By creating a model which takes into account changing passenger demand throughout the day, we will gain knowledge to what extent the resilience is time-dependent. Additionally, analyzing changing performance metrics of the network due to a change in passenger demand can give us insights to how the critical links and nodes change over the day. Therefore, creating both a time-dependent model and assessing time-dependent metrics will allow for a better understanding of the varying impact of disruptions on public transport networks.

*2.4 The London Perspective*

For over a decade, TFL has actively been releasing data regarding their transport systems. This data allows us to accurately model the system behavior of the London Underground. According to research by Deloitte, the open-data approach has saved TFL up to £130m a year (Transport for London, 2017). It is not hard to grasp why TFL is eager to improve the workings of their network; they have a clear economic incentive to decrease the impact of disruption events. For instance, passengers will receive a fare refund if their journey is delayed for 15 minutes or more (Transport for London, n.d.-b). Furthermore, research, such as this paper, made possible in part by the open data initiative, will allow for improved decision making regarding the future of the network. TFL Open Data is already being used in over 600 apps, supposedly saving Londoners £70m-£95m yearly in saved times and indirect costs (Deloitte, 2017). TFL hosts an API service with more than 80 real-time data feeds, additionally, they host static datasets on their website. In this research, the focus will be on the latter, as this empirical data will allow for creating a realistic network model in order to estimate the impact of disruption events on capacity utilization of the London network.

### *2.5 Outlining the objectives of this research*

This research aims to model time-dependent public transport networks in such a way that network criticality metrics can be measured over time and the impact of disruption events on network capacity can be estimated. In this paper, a reproducible methodology for creating these models is proposed. These models will enable a better understanding of the resilience of public transport networks and assist in improving the decision-making process regarding the disruption of public transport networks. In order quantify the impact of disruption events on the network over time, the following research question is considered:

*How can we measure the impact of disruption events on public transport networks over time?*

In order to more specifically target the goals of this research, three sub-questions are formulated. The first two questions consider a varying impact of the disruption events at different times of the day. We are interested in finding out what the varying impact is on available capacity utilization, and average passenger journey time, which are both primary indicators for the workings of a public transport network, and are factors that will come under stress during disruption events. These two sub-questions are as follows:

1. *How does the timing of a disruption event influence the impact on available capacity?*
2. *How does the timing of a disruption event influence accumulative passenger journey time?*

We are also interesting in finding a possible relationship between time-dependent metrics and critical links in the network, and whether or not these metrics are influenced by the time of the day. The final sub-question aims at finding and defining a relation between time-dependent metrics and the critical links.

3. *What are the determining time-dependent metrics to identify critical network links?*

By measuring the impact of a disruption event on the both available capacity, as well as the passenger journey time, we aim to gain knowledge of both the robustness, as the redundancy of the network. The ability to measure the time-dependent network capacity and passenger journey time will allow us to asses to which extent the resilience of a network changes over time.

## 3. Methodology

*3.1 Model paradigm*

In this study of time resilience we recognize the importance of changing capacity buffers and passenger travel time, the effect which disruption events have on these factors and how this is further influenced over time. We assume that the changing impacts can be investigated using a network simulation model where a single mode of transport and discrete non-cascading timeframes are considered. A network model is created in Python 3.7 using the NetworkX library (NetworkX, n.d.). The network model consists of edges or links which can be assigned link attributes (e.g. capacity, passenger load and link travel time) and station nodes with node attributes such as longitude and latitude. We propose a method of estimating capacity utilization by first estimating capacity of the links in the network, then assigning passengers to links by using trip assignment and origin-destination data from real-life passenger data. Additionally, the average travel time of passengers can be calculated by taking the sum of journey times in the network and dividing it by the assigned passengers in the network. For each time-band of real life passenger data, new capacity utilization and average journey time can be calculated. This model, therefore, results in discrete independent and non-cascading estimates of capacity utilization and passenger journey time. Finally, a disruption event is simulated by removing a single link in the network. This changes the network topology, and thereby influences the trip assignment in the network, which has an effect on the desired metrics of capacity utilization and passenger journey time. A network model which has a single disrupted link will be referred to as an n-1 model in the rest of this paper. N-1 models will be considered for each of the links in the network, and capacity utilization and travel time are calculated for each of the discrete time-steps.

*3.2 Data*

For this research, we used passenger data from TFL (Transport For London) in order to estimate passenger loads on the London Underground network. However, our model methodology can be applied to other public transport networks using similar data. The passenger data used in the research can almost exclusively found on the TFL crowding open data website (Transport for London, n.d.-a). The data, published under the name project NUBMAT, contains a static view of travel patterns and usage of the TFL railway services. It contains passenger data for Monday-Thursday, and Friday, Saturday and Sunday separately. The Monday-Thursday data is aggregated, so therefore the data from Friday will be considered. The Friday dataset contains two datasets which are of importance to our research: the train frequency table is required in order to estimate the capacity of the links throughout the day, and the O-D (origin-destination) matrix is used to assign passengers to links using trip-assignment. An O-D Matrix shows the amount of passengers traveling between two stations during a certain time period, and is therefore representative for the trips passengers have taken in the network. The TFL O-D Matrix contains 8 discrete time-steps for which passenger journeys are shown, namely:

```
['Morning (0500-0700)', 'AM Peak (0700-1000)', ‘Inter-Peak (1000-1600)', 'PM Peak
(1600-1900)', 'Evening (1900-2200)', 'Late (2200-0030)', 'Night (0030-0300)', 'Early
(0300-0500)']
```

Since these are the time-slots used in the source data, we will be using the same time-slots for the analysis of time dependency of network resilience. Notice that not all lengths of the time-slots are identical; the *Morning* timeslot is two hours, whereas the *Inter-peak* timeslot is 6. Lastly, an external source was used to create the London Underground network graph. This graph data was compared to TFL data in order to remove discrepancies between the data sources, such as station names that did not match. Appendix 8.1 in the attachment shows the data used for this paper and their respective sources.

*3.3 Creating the network & estimating link capacity*

After creating the initial London Underground Graph, discrepancies were found between the TFL dataset and the graph dataset. In order to remove these discrepancies, the graph dataset was cleaned by filtering stations which were not found in the TFL dataset, or visa-versa, and changing the names to match accordingly. Once the graph data was complete, we could now use TFL data in order to estimate the link attributes: link capacity and passenger load. Link capacity was estimated by taking the train frequency and multiplying it by the estimated capacity of a train for each link. Information regarding the maximum capacity of trains was gathered from the Rolling Stock information sheet (appendix 8.1). Link capacity was estimated for each of the 8 timeslots. In the *Late* and *Night* timeslots, the link frequency was often very low, causing issues in various other parts of the modeling, such as estimating capacity. Therefore, if a link showed a capacity of 0 over the full duration of a timeslot, we have chosen to recursively take the average of the capacity of neighboring links. In effect, this has slightly increased overall capacity of the network, especially in the *Late* and *Night* timeslots.

*3.4 Disruption events and n-1 models*

Having created the network graph and estimated the link capacity, the simulation of a disruption event will be considered. A disruption event is simulated by removing a single edge from the network graph, and re-estimating passenger loads. By removing an edge in the network, certain paths are no longer available in the network, causing certain passengers to have to re-route in order to reach their destination. An n-1 model (a model containing a single disrupted link) was created for each of the 352 links in the network. Each n-1 model considerers each of the 8 time-steps of the TFL datasets. In certain cases, the disrupting a link will cause the network to be disconnected, and there will be no shortest paths between certain O-D pairs. In this case, the passengers are not added to the link loads, and are considered disconnected from the network. As will be discussed in the results section, completely disconnecting certain passengers from the network has a significant effect on the main criticality metrics, capacity utilization and average travel time. For each of the n-1 models, passenger loads, capacity utilization, and average journey time are estimated.

*3.5 Estimating performance indicators using the OD-matrix*

Using the TFL O-D matrix, we are able to estimate passenger loads on links by determining which paths passengers take for each origin-destination pair. Dijkstra's shortest path algorithm is used in order to calculate a shortest path table for each of the O-D pairs. A shortest path table was created for each of the n-1 models, as disrupting an edge in the network will cause a change in the shortest paths for certain O-D pairs. This form of trip assignment is an all-or-nothing approach similar to prior research (Gauthier et al., 2018). We have considered an approach where *n* shortest paths would be considered for trip assignment, but the singular approach requires much less computation while producing similar results. One of the key assumptions using this approach is that passengers can be assigned to links regardless if they are at peak capacity or not. This means that capacity utilization might exceed 100% in certain situations, especially when a disruption event puts increased passenger load stress on certain parts of the network.

Using the shortest paths table, passenger loads are added to each of the links passengers pass over in the in the shortest path of the O-D Pair. This is done for each n-1 model, for each of the 8 time steps. Each n-1 model now contain both the capacity of the links, as well as the estimated passenger load for each link as attributes. Using these values, capacity utilization is calculated by dividing the estimated passenger load by the capacity of the link.

Capacity utilization *Cu(e,t)*, for edge *e* on timeslot *t* can be calculated for each edge by dividing the capacity *C(e,t)* by the link load *L(e,t)* of that edge:

$$Cu(e,t) = \frac{C(e,t)}{L(e,t)} \quad (1)$$

By keeping track of the amount of passengers traveling over the network for each timeslot, and taking into account passengers which might be disconnected, average travel time can be calculated.The average travel-time of passengers is calculated by taking the sum of all passengers traveling over an edge *e*, multiplied by the time it takes to travel over that link, divided by all passengers traveling in the network:

$$\overline{T}(t) = \frac{\sum_e L(e,t) * w(e)}{\sum P(t)} \quad (2)$$

Where *L(e,t)* is the link load for link *e* at timeslot *t, w(e)* time to travel over said edge (the 'weight'), and *P(t)* is the sum of traveling passengers in the network at the timeslot *t*.

Lastly, betweenness centrality is used as a topological metric and calculated for each of the links in the network. Betweennness centrality is a topological metric we are interested in due to the fact that it based on shortest paths, similar to the trip assignment approach. Betweeness centrality of an edge shows the sum of the fraction of shortest paths running through the edge:

$$B(e) = \sum_{o,d \in V} \frac{\sigma(o,d|e)}{\sigma(o,d)} \quad (3)$$

where V is the set of nodes, σ(o,d) is the number of shortest (o,d)-paths, and σ(o,d|e) is the number of those paths passing through edge *e* (*Betweenness - NetworkX 1.9*, n.d.).

*3.6 Challenges in application of the model*

A potential challenge when running the models is that the models get quite computationally heavy. For instance, trip assignment was calculated for each of the 352 links, for each of the ~68k O-D pairs, for each of the 8 timeslots. Even when optimizing the time complexity of the algorithms, this was a computationally heavy task. Certain tasks could be split up: estimating shortests paths for the O-D pairs for the n-1 models could be calculated prior to trip assignment. We chose to upload the pre-calculated shortests paths, network models and n-1 edges to a virtual machine on AWS which allowed for faster computations. Since running an n-1 model is an independent calculation event, multithreading was used in order to improve the calculation speed. As the model results are saved on the networkX graph object as link attributes, the model results could be saved as a ‘picked’ data format. The final step was pulling the link attributes from the network models in order to analyze the data. We used Pandas in order to format the dataframes and perform exploratory analysis on the data. The dataframes were later exported, and visualizations were made in Tableau.

## 4. Results

In order to measure the impact of disruption events on public transport networks over time, we will start by analyzing the results from a topological perspective. The undisrupted network provides valuable insights over the network behavior over time without major difficulties, whilst the disrupted networks will give us insights on the various changing impacts of disrupting edges over time on the network. Where possible, results are visualized in a spider map showcasing edge attributes over a map of the London Underground network.

### *4.1 n-1 network models*

The majority of the analysis focuses on changing performance metrics derived from disrupting a single link in the network, which we call an n-1 network model. An n-1 network model disrupts a single link by effectively removing said link from the network graph. Using the O-D matrix, passengers are then re-assigned to the links, by recalculating their optimal path to their destination using Dijkstra's shortest path algorithm. Passengers are then assigned to their alternative routes, similarly how passengers would be re-routed in real-life disruption events, assuming that that they will still use the London Underground system and not a different mode of transport.

### *4.2 Bridges*

In some cases, certain passengers are no longer able to be assigned to links, because no new shortest path is found. This is in the case that the disrupted link is a bridge. Removal of a bridge in the network increases the amount of connected components, effectively creating two (sub)graphs. Consequently, if passengers would travel over a bridge-link, and if that link is disrupted, these passengers are now disconnected from the graph. Out of 352 links, roughly 43% are bridges in the network.


*FIG 1.* ***Bridges in the London Underground network.*** *Orange signifies a bridge-link, blue signifies a non-bridge link.*

Disconnected passengers in an n-1 model caused logical, but unforeseen consequences. In real-life situations, disconnected passengers could choose alternative modes of transport to 'bridge the gap' in the network, and still partake in traveling throughout the remaining undisrupted part in the network. However, since the model is single-modal, this is not possible for passengers in our model. Effectively, this influences capacity utilization and cumulative travel time in sometimes counterintuitive and unpredictable manners. For instance, one would expect that a decrease in capacity utilization would be good for resilience in the network, however, disconnecting passenger logically decreases capacity utilization. In order to show this potential counterintuitive behavior, figure 2 shows the relationship between the change of the average capacity utilization of all links in an n-1 disruption model, and the amount of disrupted passengers.

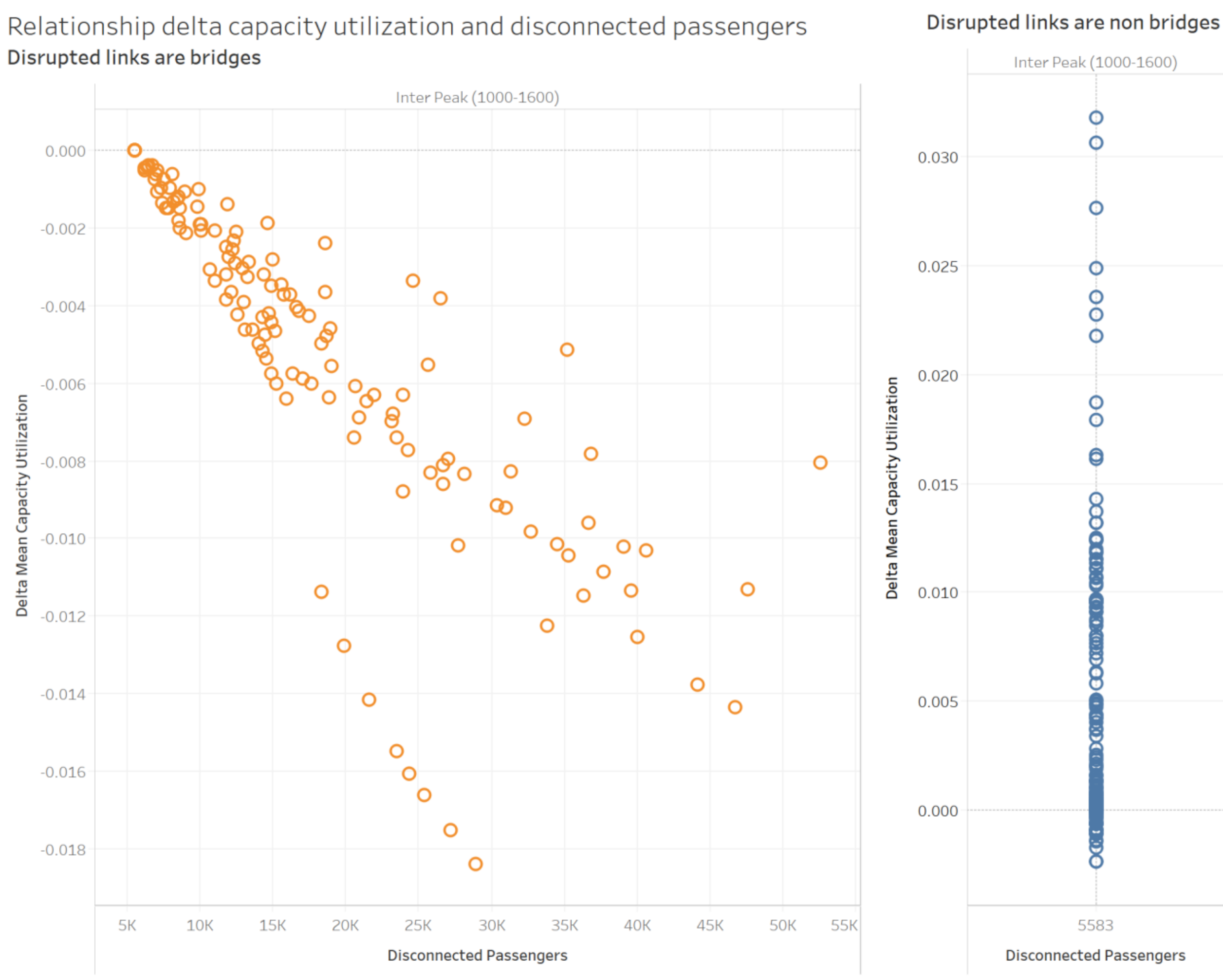



*FIG 2.* ***Relationship between the change in mean capacity utilization and the amount of disconnected passengers.*** *Each dot represents an n-1 model for a disrupted link in the network during the inter-peak time-band. The y-axis displays the change in the average capacity utilization of all other links in the disrupted n-1 model. Notice that the range of the y-axis on the graphs is not identical. Bridges are shown on the left, non-bridges on the right. Data is used from the 'Inter-Peak' timeslot.*

Figure 2 shows that disrupting a bridge link causes additional passengers to be disconnected, but disrupting a non-bridge link does not. The reason that 5583 passengers are disconnected in the n-1 models for non-bridges is that in the base (non-disrupted) model, these passengers were not able to be assigned to the network due to a discrepancy between the NetworkX graph and the TFL O-D data. Specifically, Heathrow Terminal 5 was not added to the model until the last revision. In a later version of the model, this is fixed, however, the current results section needs to be updated to reflect this (for purposes of the BEP, this is left out of the scope). In addition to the finding that only disrupting bridges causes passengers to be disconnected, we observe that there seems to be a relationship between the amount of disconnected passengers and the change in the average capacity utilization of the network. Similar results can be seen if the amount of disconnected passengers is plotted against the average travel-time per passenger of the n-1 model (*appendix 8.2*). Therefore, we should take into account bridges as topological metric of importance when analyzing the performance metrics capacity utilization and average travel time.

*4.3 Disruption events and average travel time*

When analyzing effect of timing of a disruption event on travel time, we should take into account the influence of bridges on emergent network behavior for the n-1 models. Initially, we were interested in the effect of disruption events on the cumulative travel time in the network. However, as we have observed, bridges disconnect passengers from the network. As an increasing amount of passengers are disconnected, the cumulative travel time decreases in the network, skewing the comparison to the cumulative travel time of the base-network. This results in the cumulative travel time not being a reliable indicator of the impact of disruption events on the network, as cumulative travel time is highly dependent on the amount of disconnected passengers. Instead, we analyze the influence of disruption events on the average travel time, as average travel time takes into account the number of disconnected passengers.

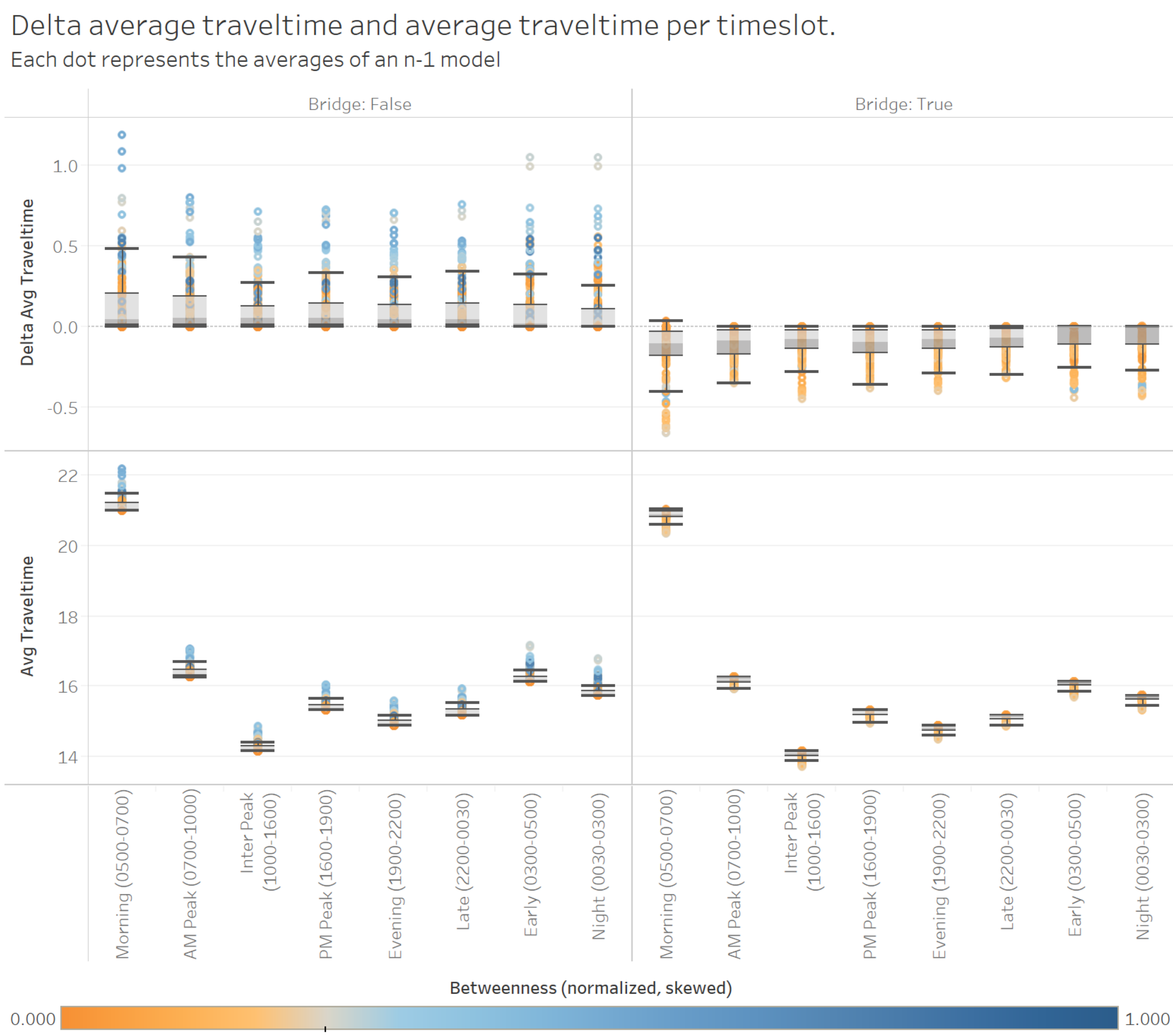



FIG 3. ***Relationship between the average and change in mean travel time of n-1 models and the timing of the disruption event.*** *Average travel-time of passengers in n-1 models, and change of average travel time compared the undisrupted model. The betweenness metric is calculated for the disrupted edge in the n-1 model. The color scale is right skewed to represent the betweenness distribution.*

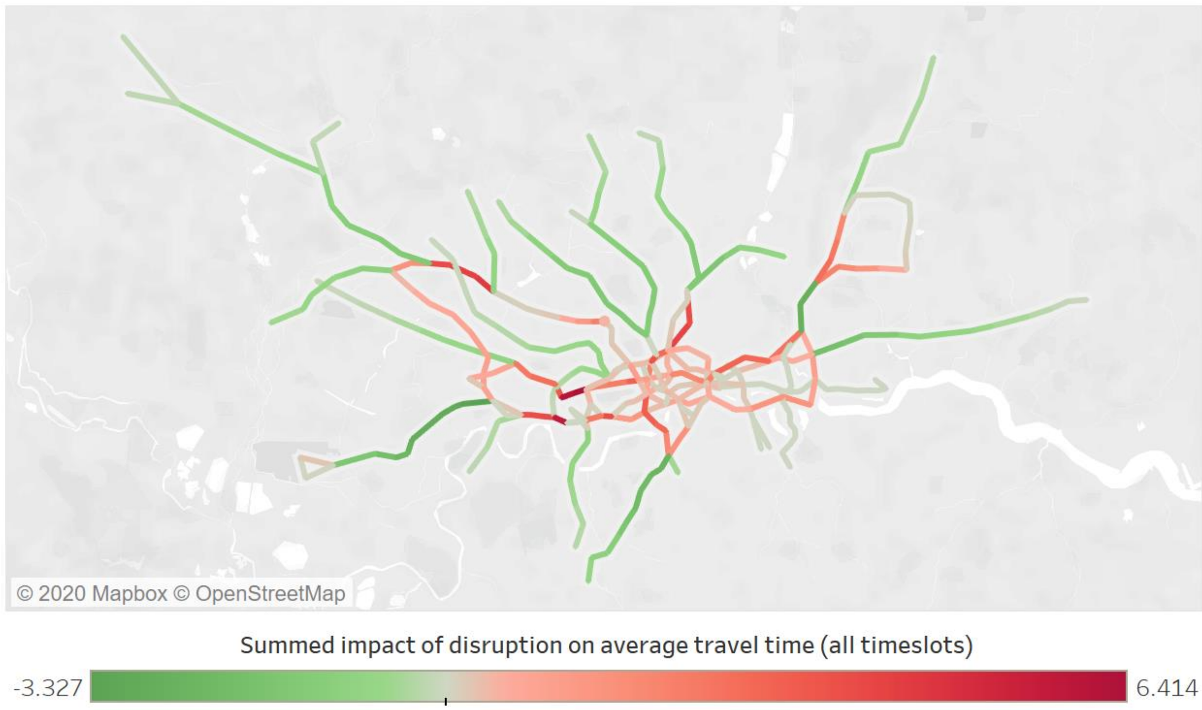



*FIG 4.* ***Impact of disrupting a link on the average travel time, summed over each timeslot.***

Figure 3 shows the changing average travel-time in the n-1 models throughout the day, as well as the change in average travel time of the n-1 models as compared to the undisrupted network. We observe a reversal in the impact on average travel time between bridges and non-bridges. A possible reason for this is that, on average, passengers who would travel over a bridges would travel further in the network. Disconnecting these passengers from the network, thereby removing them from the average travel time equation, would logically result in a decrease. On the other hand, disrupting a non-bridge link would not result in disconnecting any passengers as seen in Figure 2. Instead, certain passengers will have to re-route in order to reach their destination, increasing average travel time. Comparing Figure 1 and Figure 4 shows the relationship between bridges and non-bridges, and the reversal of the impact of disruption events on average travel time.

Notice that for figure 3 the center of the betweenness color range is skewed, as the normalized values are right-skewed. We can observe that when disrupting a non-bridge link with high betweenness, the impact on the average travel time is generally higher. With a higher betweenness value, a larger fraction of the shortest paths in the network travel through that link (*equation (3)*). Disrupting such a link with high betweenness would mean that, on average, more passengers would have to re-route, causing a higher increase in average travel time than lower betweenness links.

The timing of a disruption event and the impact of the event on the average travel time of passengers in the network does not seem to have a strong relationship. However, we propose a logical explanation. The way the model assigns passengers does not take into account a maximum capacity, instead the model assigns passengers wherever possible, without limiting assignment after exceeding capacity. Therefore, for a passenger to be rerouted in our models, it is irrelevant for the passenger’s journey time if the disruption happens in peak hours or in off-peak hours. This might explain an independency between time and the impact of disruption events on the average travel time in the network. Future research could take into account maximum capacity as a boundary condition for trip assignment to gain a broader insight on the impact of disruptions on average travel time in public transport networks.

Interestingly, we observe that the average travel time is highest in the *morning* hours. This might be explained by passengers which decide to travel early on in the morning, have to come from further in the network to reach their destination (e.g. work in the center of the city). However, more detailed passenger and network data is required to explain this. For instance, *Morning* passenger destinations could be compared to points-of-interest around the network to assess their purpose of travel.

*4.4 Disruption events and capacity utilization*

For the study of resilience of public transport networks, capacity utilization is a highly important performance indicator. Once a metro system goes over capacity, it results in various negative side effects such as increased waiting times and decreased passenger satisfaction. First, we will consider the estimated capacity utilization of the undisrupted network.

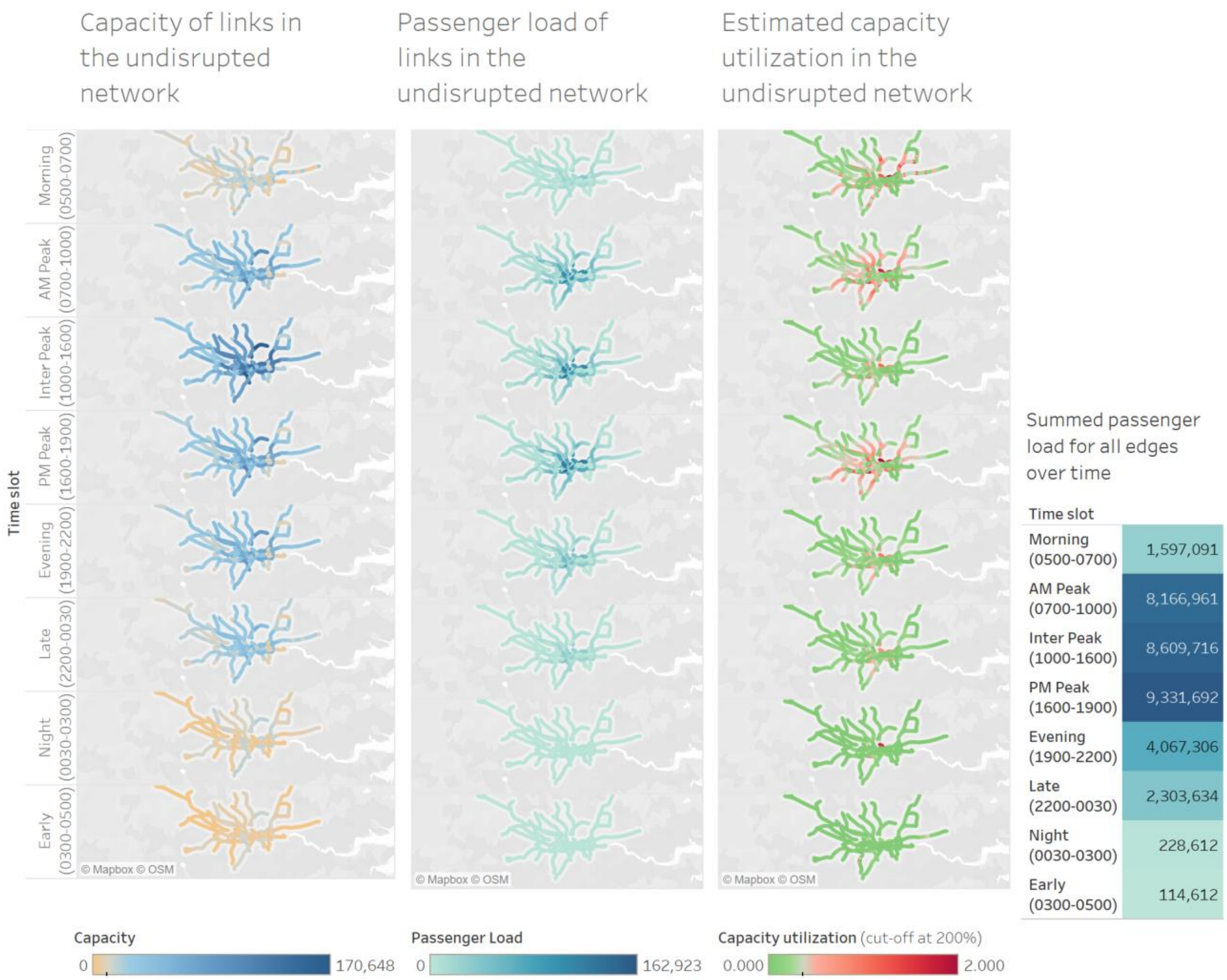



*FIG 5.* ***Resilience metrics in the undisrupted London Underground network after trip assignment using the O-D matrix from TFL.*** *Map based on longitude and latitude of the edges, broken down for each timeslot. Color shows details about the metrics.*

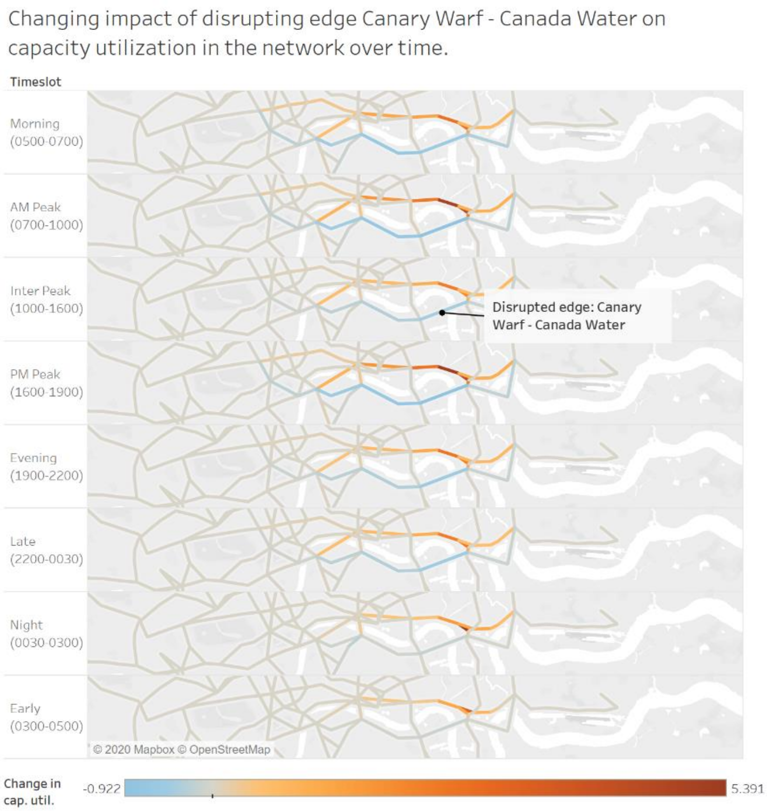
Figure 5 provides a picture of the network model in an undisrupted state, showing the relationship between capacity, passenger load and capacity utilization broken down for each of the timeslots. Capacity utilization is calculated using passenger load and capacity as shown in *equation 1*. On the right, a table show the summed passenger load on all edges during each of the timeslots. Notice that the time duration changes between timeslots (e.g., PM and AM peak are 3 hours long, whereas inter-peak is 6 hours long). This explains the relative high number of summed passenger load for the inter-peak timeslot.

The resilience of the public transport network is partly dependent on the excess capacity (the resource buffer) and to what extent passengers are able to use this excess capacity to reach their destination after a disruption event. When a disruption event takes place, taking into account passengers are still able to travel to their destination, they will have to route around the disrupted link. This causes the neighboring links to have an increased passenger load. We can observe in Figure 5 that in the center of the network, the capacity is already reaching its limits. On the other hand, many of the outer links especially, over various timeslots, show a low capacity utilization, which might indicate a capability to dampen the negative effects of a disruption event should the link come under increased passenger load. Another useful observation is that the capacity at *Night* and *Early* reaches very low points, indicated by an orange color. This will have an influence on the estimates of changing capacity utilization after a disruption event occurs. A very low initial capacity might cause an increased impact when a disruption event occurs.

| Disrupted Edge | Description | Bridge | Betweenness (norm.) |
|---|---|---|---|
| Canary Wharf LU - Canada Water | Highest summed impact on cap. Util | No | 0.313 |
| Preston Road - Wembley Park | Second highest summed impact on Cap. Util. | No | 0.505 |
| Green Park - Piccadilly Circus | Low impact on cap. Util, low betweenness | No | 0.026 |
| Claphham North - Stockwell | Highest negative summed impact on cap. Util | Yes | 0.199 |

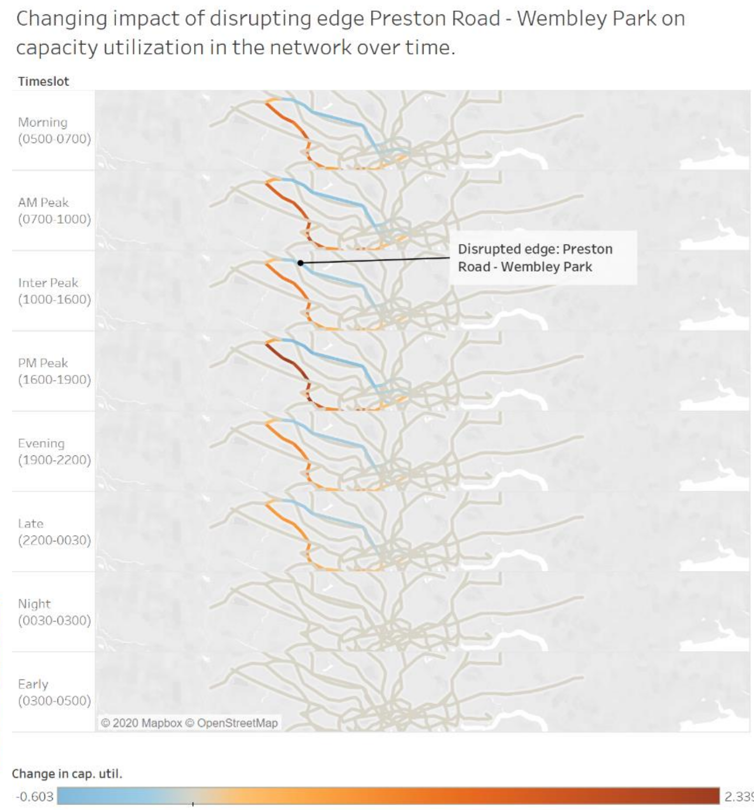
***Table 1***. *Disrupted edges used in analysis*

Table 1 lists 4 disrupted edges which will be considered for analysis below. The reason these bridges were selected is because the bridges Canary Wharf LU – Canada Water and  Preston Road – Wembley Park show a high summed positive impact when disrupted.  Green Park – Piccadilly Circus show a very low summed impact (the positive increase in capacity utilization is very close to the summed negative increase in capacity utilization), and the betweenness value is quite low. Analyzing this edge gives us insights in the disruption of a possible ‘low-impact’ edge.  Lastly, Clapham North – Stockwell is a bridge link with the highest negative impact (decrease-in) capacity utilization in the network.




Summed change in capacity utilization over time after disrupting edge Canary Wharf - Canada Water

| Timeslot | |
|---|---|
| Morning (0500-0700) | 10.23 |
| AM Peak (0700-1000) | 17.64 |
| Inter Peak (1000-1600) | 8.75 |
| PM Peak (1600-1900) | 18.08 |
| Evening (1900-2200) | 9.20 |
| Late (2200-0030) | 8.57 |
| Night (0030-0300) | 11.34 |
| Early (0300-0500) | 7.17 |




Summed change in capacity utilization over time after disrupting edge Preston Road - Wembley Park

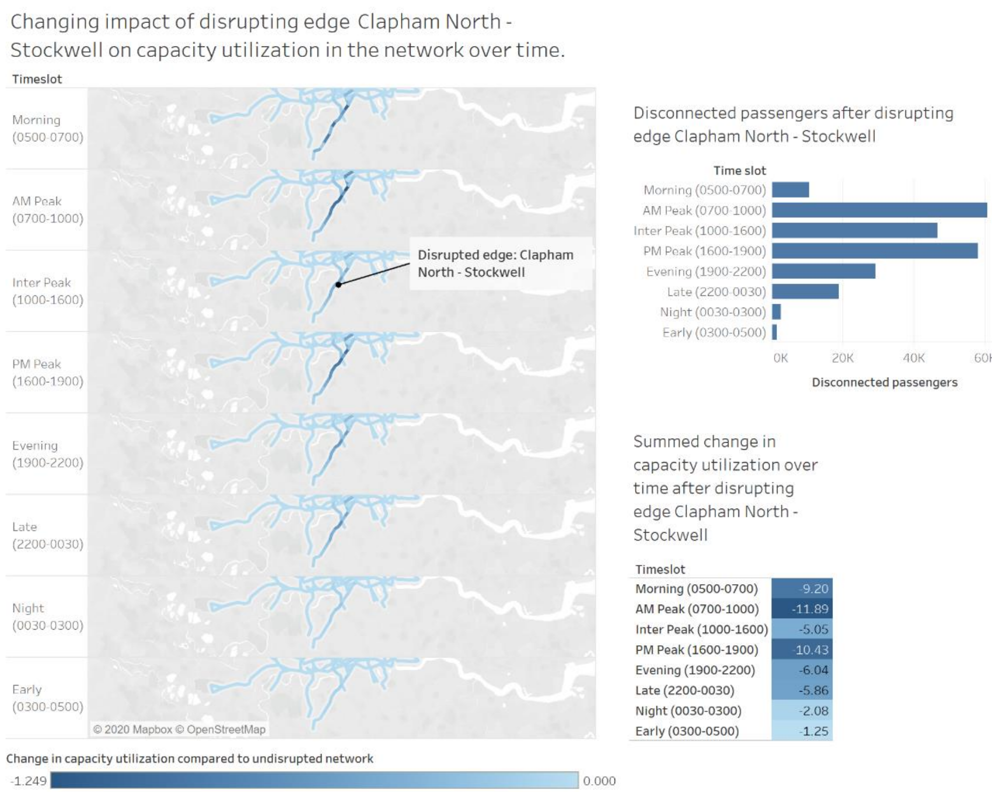
| Timeslot | |
|---|---|
| Morning (0500-0700) | 14.83 |
| AM Peak (0700-1000) | 19.71 |
| Inter Peak (1000-1600) | 11.20 |
| PM Peak (1600-1900) | 24.96 |
| Evening (1900-2200) | 9.93 |
| Late (2200-0030) | 7.39 |
| Night (0030-0300) | 0.00 |
| Early (0300-0500) | 0.00 |

*FIG 6.* ***Effect of disrupting link Canary Wharf – Canada Water and link Preston Road – Wembley Park on the change in capacity utilization of the remaining links in the network, as compared to the undisrupted model.*** *Map based on longitude and latitude of the edges, broken down by Timeslot. Color shows details about the change in capacity utilization. The tables on the right details the sum of the change in capacity utilization over the other links, per timeslot.*

When assessing the time dependency of high impact links, based on the sum of capacity utilization change, we can observe that capacity utilization increases dramatically over neighboring links to the disrupted links, especially during peak-hours. While we have previously seen that time influences the capacity utilization of an undisrupted network, it becomes apparent that the significance of the disruption impact is also time-dependent.

In the case of the disruption of the non-bridge link Canary Warf – Canada Water, we observe that the links following and preceding the disrupted link in a path have a decreased passenger utilization, whereas the links that are parallel to the disrupted links have increased passenger utilization. This network behavior is possibly explained by the fact that passengers who were to travel over the disrupted link were likely to also travel over the preceding and succeeding links. Therefore, if passengers are no longer able to cross the ‘gap’ which is created by the disruption event (again, assuming a single-modal network), they will not travel over a certain amount of the links in their original path including the disrupted links. Instead, as passenger re-route, the strain on the network is diverted to an alternative path which is parallel to the path containing the disrupted link.

A similar effect is seen in the disruption of non-bridge link Preston Road – Wembley Park. Just as with the previous example, disrupting this link causes preceding and succeeding links to decrease in capacity utilization, and a parallel path is put under more strain. Similarly, the disturbance of the disruption event is highest during the AM and PM peak hours. The capacity utilization during the Night and Early timeslots does not seem to change, which is most likely caused by the extremely low capacity of certain parts in the network during these time-slots as seen in figure 6 on the left. We need to pinpoint the exact cause of this by re-running the model. This is outside of the scope of the BEP and will be one of the focus points for the research paper.



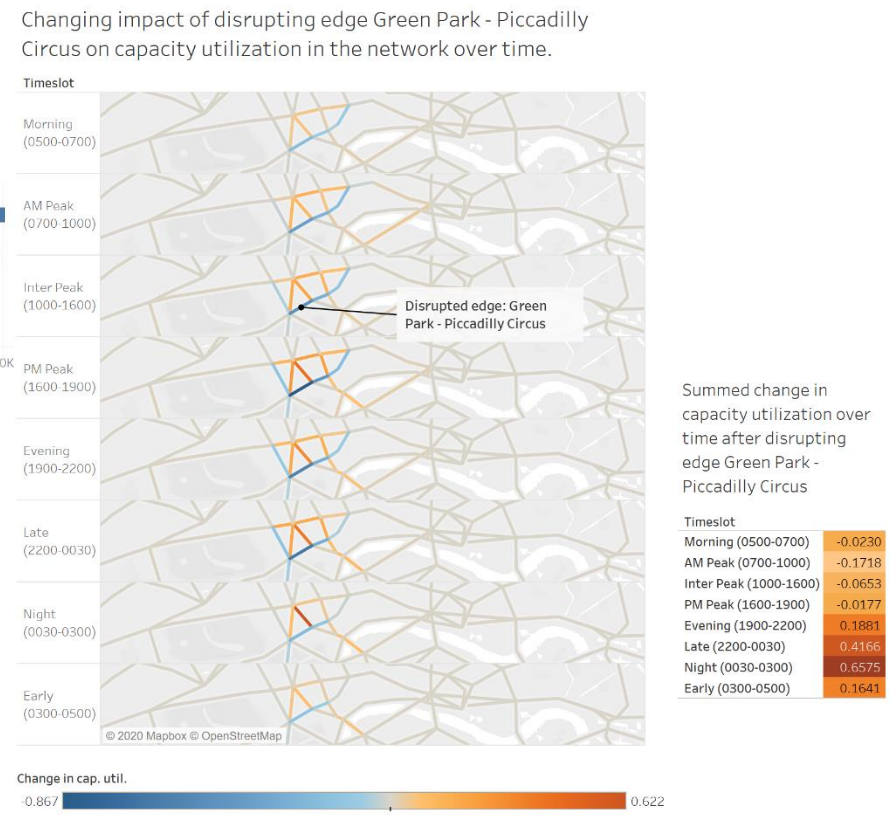




*FIG 7.* ***Effect of disrupting bridge link Clapham North - Stockwell on the change in capacity utilization of the remaining links in the network, as compared to the undisrupted model.*** *Map based on longitude and latitude of the edges, broken down by Timeslot. Color shows details about the change in capacity utilization. The table on the right details the sum of the change in capacity utilization over the other links, per timeslot.*

*FIG 8.* ***Effect of disrupting link Green Park – Piccadilly Circus on the change in capacity utilization of the remaining links in the network, as compared to the undisrupted model.*** *Map based on longitude and latitude of the edges, broken down by Timeslot. Color shows details about the change in capacity utilization. Map based on longitude and latitude of the edges, broken down by Timeslot. Color shows details about the change in capacity utilization.*

Disrupting the non-bridge link Green Park – Piccadilly Circus (figure 8) results in a relatively low summed impact in regards of capacity utilization change as compared to the base network. Interestingly, where we previously saw a comparatively strict increase in capacity utilization in a parallel path, the strain of the disruption event now seems to be diffused over the neighboring links. This might be explained by the low betweenness centrality of the link. Possibly, passengers passing through this link are unlikely to come from origins far away from the disrupted link, as the amount of shortest paths running through the link is quite low. This would explain the more diffuse manner of increased passenger load strain on neighboring links, as opposed to the more parallel strain seen in the more impactful links (figure 7).

Finally, the disruption of bridge link Clapham North – Stockwell has a vastly different impact on the network as the disruption of the non-bridge links we have observed earlier. As a bridge link, this link disconnects passengers from the network, as seen in the bar graph in the top right of the figure. The impact of the disruption event is mostly a decrease in capacity utilization – passengers which are not able to ‘cross the gap’ due to the single modal approach of the network model, no longer travel in the network decreasing passenger load and capacity utilization. The higher the increase of disconnected passengers, the higher the decrease of

In order construct more robust networks, and design more resilient public transport network policy, policy makers should take into account emerging network behavior for disruption over time, both for low-impact and high impact links. For high impact links, the question should be asked whether the impact of the increased strain on a parallel path be moderated, or the focus should instead lie on allowing passengers to ‘cross the gap’ in a disrupted network path. For links less critical in network vulnerability, an element for further research is finding out how the strain of the disruption event is spread over neighboring links, and in what way this issue can be addressed. Considering bridge links especially, giving passengers viable alternatives to reach their destination is key, especially during the critical peak-hours.

## 5. Conclusion and discussion:

Transportation network resilience has been studied previously using either graph theory and applications (Berche et al., 2009) or traffic network modelling (Ganin et al., 2017). While the latter is data intensive and computationally expensive, the former approaches rely only on the structure of the network and missing the critical information on the impact of time dependent variables on the network. It is clear that the literature is lack of new methods and approaches that lie between the two. Therefore, the motivation behind this study is to explore the importance of disruption time in public transport network resilience, and how to measure the changing impact over time.

We have proposed a methodology to model a non-cascading time-dependent network model, estimating changing passenger loads in a public transport network. Our models, created using empirical passenger data from the London Underground, show that the network is not only most vulnerable during peak-hours to increased passenger loads, but in addition, the impact disruptions on the network change and the highest during peak-hours at certain locations. Accordingly, the importance of time should not be disregarded as an important piece in the complex puzzle of understanding public transport network vulnerability and resilience.

One important finding of the study is that the metrics need to be identified carefully to assess the transportation network resilience. In this model, two performance indicators are identified to evaluate the impacts of disruptions: average travel time, capacity utilization and number of disconnected passengers. In addition, we considered two key topological metrics; bridges and betweennes centrality. Due in part to the modelling assumption of using a single mode of transport, one topological attribute stood out as highly impactful in our models even more than the widely utilized metrics such as betweennes; disrupted edge being a bridge or not. Since passengers were unable to 'jump the gap' once a bridge link is disrupted, it caused the network to be divided into two sub-networks. The passengers which would initially route over those links could no longer partake in their journeys throughout the network. Roughly 43% of the 352 links in the network are bridges.

The performance indicators are key in gauging the impact of disruption events, journey time and capacity utilization, showed a correlation to the number of disconnected passengers in the network. This caused unexpected, or counterintuitive results in cases where the bridge attribute was not taken into consideration. For instance, disconnecting a bridge link would in some cases decrease capacity utilization as well as average journey time. Although these seem like positive effects, until the disconnected passengers is taken into account. This is due to the fact that the model does not allow multiple modes of transport and the capacity utilization is low as disruptions results in the disconnected passengers. Modelling multiple modes of transport would possibly allow passengers to 'bridge the gap' over disrupted links, and would be a great opportunity for future research in time-resilience.

We have estimated the average travel time of passengers (*equation (2)*) after disrupting the network, taking into account the skewing effect of the bridge link attribute on said metric. In our network model, the impact on average travel-time seems largely independent to the timing of the disruption. In general, disrupting bridge links showed a decrease in average travel-time, which is possibly explained by the relative distance of bridge links from the center of the network. The reason being that disconnected passengers would travel, on average, from further away. Hence, removing these from the average travel time equation result in a decrease. This is again a disadvantage of not having a multi modal transportation network, as we assumed that the disrupted passengers are removed from the network.

On the other hand, the disruption of a non-bridge link showed an increase in average travel time, a result which is not unexpected as certain passengers would have to re-route in order reach their destination. Interestingly, taking into account a classic network metric, betweenness centrality (*equation (3)*), our results show that there is a positive relationship between betweenness centrality and the increase in average travel-time, assuming a single mode of transport. In other words, betweenness centrality successfully represents the critical locations on the network if the link is also a non-bridge.

One of the goals of this research is to explore the topological metrics which identify time-dependent critical network links. In this study, betweenness has shown to be a valuable indicator of link criticality when using our n-1 model approach. We assume the reason that betweenness been a good indicator of the relative importance of (mostly non-bridge) links in the network, is due to the fact that this is a metric based on the fraction of shortest paths running through an edge, and passengers are assigned using shortest-paths. However, it is possible other metrics might prove more reliable indicators in different modelling approaches, such as agent based modeling where individual passenger behavior becomes more important, or dynamic flow models.

Additionally, certain metrics might be able to explain dynamic criticality behavior of public transport networks. For instance, certain features of a link might make them resilient during peak hours, but vulnerable during off-peak hours, or visa-versa. Investigating what causes vulnerability metrics to change over time, and determining if these causes can be tied to indicators might show to be a powerful tool in tackling time-dependency in public transport resilience. We anticipate that future research connecting the impact of time-dependent criticality metrics such as average travel-time and capacity utilization, with various topological metrics will bring promising new insights.

The results of capacity utilization in the undisrupted network suggest that an undisrupted network is already vulnerable during peak-hours, as capacity utilization reaches high levels which negatively effects redundancy and network resilience. Disruptions of high-impact non-bridge links showed a small decrease in capacity utilization in preceding and succeeding links and a more significant increase in capacity utilization for parallel links. Parallel links allow a large number of passengers to reach their destination when re-routing is required due to the disruption event. This illustrates the importance of redundant routes for the resilience of networks.

This parallel effect was less significant in terms of capacity utilization for non-bridge links which have a low betweenness centrality, instead effecting the capacity utilization of neighboring links in a more diffuse manner. A more diffuse impact means the neighboring links are able to share the increased passenger load. Therefore, a moderate link capacity redundancy of neighboring links surrounding low-impact, low-betweenness links might be a sufficient measure to improve network resilience.

The performance indicator of disconnected passengers, illustrates the effects disruption events have when disrupting a bridge link. A high amount of disconnected passengers shows a decrease in capacity utilization in subsequent links, especially during peak hours. Therefore, for both bridges and non-bridges, the impact of disruption events were most significant during peak hours.

We expected to see a spatial change in the effect of disruption events over time, but the analysis of our model results did not reflect this expectation. Spatial differences after disruption events seem to not be a result of time, but rather one of our definition of link criticality- betweennes value. Possibly, different methods for measuring time resilience in public transport resilience will give new insights in spatial changes of disruption impacts throughout the day.

## 6. Future research

This research has given us initial insights on the time dependency of public transport network resilience. Excitingly, it has also primed new opportunities for great potential directions for future research. The most promising angles are as follows:

- Taking into account multiple modes of transport will allow for passengers to bridge gaps which are created by potential disruption events. The performance metrics will be influenced by a multi-model model, as we expect the impact of disrupting bridge links to be partially negated by allowing passengers to cross the gap using a different mode of transport. Additionally, potential new behavior will emerge from models considering multiple modes of transport. For instance, dependent on the data used, a distinction could be made between passengers traveling using a single mode of transport, or multiple.
- Changing the means of trip assignment. In our models, we assigned passengers by determining the shortest path between an origin-destination pair from the O-D matrix, using all-or-nothing assignment.. Since certain links on the outer edges of the network have a low betweenness centrality, their passenger loads might be estimated to be lower than reality, and visa-versa for high betweenness links. Alternative forms of (all-or-nothing) trip assignment will result in different model behavior, which might lead to new insights.
- Considering capacity utilization, setting a maximum capacity utilization as a boundary condition for passenger travel will allow for measuring a more realistic impact of disruption events. In real-life situations, passengers are not able to travel in full trains, and will have to wait on the platform, choose a different mode of transport or even reneging.
- Decreasing the length of the time-steps might allow for more detailed model behavior. For instance, with smaller time-steps, one might consider the cascading effect of a disruption event, and how for long the disruption is felt throughout the day. Additionally, this model approach will allow for a dynamic disruption length, or allow for partial degradation of link capacity instead of complete link failure similar to prior research (Oded Cats & Jenelius, 2018), while still considering the importance of time.
- Our model approach takes into account two main performance indicators, average travel time and capacity utilization. By taking into account more passenger metrics such utility of travel or cognitive burden, estimating the effect of a disruption event on passengers can be improved. We expect the impact of disruptions on passengers to change throughout the day. For instance, a passenger going to work might experience a more substantial impact on their utility of travel compared to a passenger going home, while other performance metrics stay the same.
- Similarly to taking into account more passenger metrics, more topological metrics can also be considered. This would be especially interesting to see if additional metrics can explain time dependency of link criticality and vulnerability.
- Lastly, the improved understanding of changing resilience over time allows researchers to explore design strategies to improve resilience of public transport systems.

**Acknowledgements**
I would like to express my deepest gratitude to my thesis supervisors Nazli Aydin and Trivik Verma for their continuous and generous support during my work on the research project. I would also like to thank Tineke Ruijgh for her valuable advice and lessons for writing my Bachelor End Project.

## 7. References


Berche, B., von Ferber, C., Holovatch, T., & Holovatch, Y. (2009). Resilience of public transport networks against attacks. *The European Physical Journal B*, *71*(1), 125–137. https://doi.org/10.1140/epjb/e2009-00291-3

*Betweenness—NetworkX 1.9*. (n.d.). Retrieved July 10, 2020, from https://networkx.github.io/documentation/networkx-1.9/reference/generated/networkx.algorithms.centrality.edge_betweenness_centrality.html

Bruneau, M., Chang, S. E., Eguchi, R. T., Lee, G. C., O'Rourke, T. D., Reinhorn, A. M., Shinozuka, M., Tierney, K., Wallace, W. A., & von Winterfeldt, D. (2003). A Framework to Quantitatively Assess and Enhance the Seismic Resilience of Communities. *Earthquake Spectra*, *19*(4), 733–752. https://doi.org/10.1193/1.1623497

Cats, O. (2016). The robustness value of public transport development plans. *Journal of Transport Geography*, *51*, 236–246. https://doi.org/10.1016/j.jtrangeo.2016.01.011

Cats, Oded, & Jenelius, E. (2018). Beyond a complete failure: The impact of partial capacity degradation on public transport network vulnerability. *Transportmetrica B: Transport Dynamics*, *6*(2), 77–96. https://doi.org/10.1080/21680566.2016.1267596

Deloitte. (2017, July). *Assessing the value of TfL's open data and digital partnerships*. http://content.tfl.gov.uk/deloitte-report-tfl-open-data.pdf

Ganin, A. A., Kitsak, M., Marchese, D., Keisler, J. M., Seager, T., & Linkov, I. (2017). Resilience and efficiency in transportation networks. *Science Advances*, *3*(12), e1701079. https://doi.org/10.1126/sciadv.1701079

Gates, S., Gogescu, F., Grollman, C., Cooper, E., & Khambhaita. (2019). *Transport and inequality*. 72.

Gauthier, P., Furno, A., & El Faouzi, N.-E. (2018). Road Network Resilience: How to Identify Critical Links Subject to Day-to-Day Disruptions. *Transportation Research Record*, *2672*(1), 54–65. https://doi.org/10.1177/0361198118792115

Geng, Y., Ma, Z., Xue, B., Ren, W., Liu, Z., & Fujita, T. (2013). Co-benefit evaluation for urban public transportation sector – a case of Shenyang, China. *Journal of Cleaner Production*, *58*, 82–91. https://doi.org/10.1016/j.jclepro.2013.06.034

Hernández, D. (2018). Public transport, well-being and inequality: Coverage and affordability in the city of Montevideo. *CEPAL Review*, *2017*(122), 151–169. https://doi.org/10.18356/e5039949-en

Jafino, B. A., Kwakkel, J., & Verbraeck, A. (2020). Transport network criticality metrics: A comparative analysis and a guideline for selection. *Transport Reviews*, *40*(2), 241–264. https://doi.org/10.1080/01441647.2019.1703843

Jenelius, E., & Cats, O. (2015). The value of new public transport links for network robustness and redundancy. *Transportmetrica A: Transport Science*, *11*(9), 819–835. https://doi.org/10.1080/23249935.2015.1087232

Lin, J., & Ban, Y. (2013). Complex Network Topology of Transportation Systems. *Transport Reviews*, *33*(6), 658–685. https://doi.org/10.1080/01441647.2013.848955

*NetworkX*. (n.d.). Retrieved April 19, 2020, from https://networkx.github.io/

Orth, H., Weidmann, U., & Dorbritz, R. (2012). Development of Measurement System for Public Transport Performance. *Transportation Research Record*, *2274*(1), 135–143. https://doi.org/10.3141/2274-15

Smart, M., Miller, M., & Taylor, B. (2009). Transit Stops and Stations: Transit Managers' Perspectives on Evaluating Performance. *Journal of Public Transportation*, *12*(1), 59–77. https://doi.org/10.5038/2375-0901.12.1.4

Sullivan, J. L., Novak, D. C., Aultman-Hall, L., & Scott, D. M. (2010). Identifying critical road segments and measuring system-wide robustness in transportation networks with isolating links: A link-based capacity-reduction approach. *Transportation Research Part A: Policy and Practice*, *44*(5), 323–336. https://doi.org/10.1016/j.tra.2010.02.003

Tahmasseby, S., & van Nes, R. (2007). Robustness of urban public transport networks. *Urban Transport XIII: Urban Transport and the Environment in the 21st Century*, *I*, 337–346. https://doi.org/10.2495/UT070321

Taylor, M. A. P., & D'Este, G. M. (2003, October). *Concepts of network vulnerability and applications to the identification of critical elements of transport infrastructure*. 26th Australasian Transport Research Forum, Wellington New Zealand.

Thompson, K., & Schofield, P. (2007). An investigation of the relationship between public transport performance and destination satisfaction. *Journal of Transport Geography*, *15*(2), 136–144. https://doi.org/10.1016/j.jtrangeo.2006.11.004

Transport for London. (2018). *TFL crowding open data*. Retrieved April 18, 2020, from http://crowding.data.tfl.gov.uk/

Transport for London. (n.d.-b). *Tube and DLR delays*. Transport for London. Retrieved April 12, 2020, from https://www.tfl.gov.uk/fares/refunds-and-replacements/tube-and-dlr-delays

Transport for London. (2017, October 13). *TfL's free open data boosts London's economy*. Transport for London. https://tfl.gov.uk/info-for/media/press-releases/2017/october/tfl-s-free-open-data-boosts-london-s-economy

Woodcock, J., Banister, D., Edwards, P., Prentice, A. M., & Roberts, I. (2007). Energy and transport. *The Lancet*, *370*(9592), 1078–1088. https://doi.org/10.1016/S0140-6736(07)61254-9

Wu, X., (Jason) Cao, X., & Ding, C. (2020). Exploring rider satisfaction with arterial BRT: An application of impact asymmetry analysis. *Travel Behaviour and Society*, *19*, 82–89. https://doi.org/10.1016/j.tbs.2019.12.006

# 8. Appendix

## *8.1 Data Table*

| Data | Description | Source | url |
|---|---|---|---|
| Link Frequency table | Link frequency table used to estimate capacities of the links, 15 minute intervals | NUMBAT data: 2018FRI_Link_Frequency.xlsx | [1] |
| LU OD Matrix table | Origin-destination matrix used for trip assignment, split in 8 time-steps | NUMBAT data: 2018FRI_OD_LU.xlsx | [2] |
| Train Capacity information | Information regarding capacity of trains for the LU Lines | Rolling Stock Infromation Sheet | [3] |
| LU Network lines csv | CSV file containing line information | Github: Mark Dunne | [4] |
| LU Network stations csv | CSV file containing station information | Github: Mark Dunne | [5] |
| LU Network connections csv | CSV file containing connections information | Github: Mark Dunne | [6] |

[1] http://crowding.data.tfl.gov.uk/
[2] http://crowding.data.tfl.gov.uk/
[3]
https://www.whatdotheyknow.com/request/239641/response/590412/attach/3/RS%20Info%20Sheets%204%20Edition.pdf
[4] http://markdunne.github.io/public/london_tube/london.lines.csv
[5] http://markdunne.github.io/public/london_tube/london.stations.csv
[6] http://markdunne.github.io/public/london_tube/london.connections.csv

*8.2 Appendix fig 1.*

Relationship delta capacity utilization and disconnected passengers

**Disrupted links are bridges**