



UCL

WORKING PAPERS SERIES

Paper 211 - Sept 18

**Employment Accessibility
in the London Metropolitan
Region: Developing a Multi-
Modal Travel Cost Model
Using OpenTripPlanner and
Average Road Speed Data**

ISSN 1467-1298



Employment Accessibility in the London Metropolitan Region: Developing a Multi-Modal Travel Cost Model Using OpenTripPlanner and Average Road Speed Data

Duncan A. Smith

duncan.a.smith@ucl.ac.uk

Centre for Advanced Spatial Analysis
90 Tottenham Court Road
London, W1T 4TJ

Abstract

Accessibility analysis is relevant for a number of active geographical research areas, including equity analysis of travel opportunities between socio-economic groups; sustainability analysis of travel opportunities by transit and active modes compared to car trips; and investigating spatial planning topics such as labour market access and jobs-housing balance. The scale and complexity of travel networks in large city regions favours geocomputational modelling methods, and recent data and software advances are improving the sophistication of accessibility models. These advances include the wider availability of detailed transit schedule data; the wider availability of high quality road network and average road speeds data; and the development of sophisticated open source journey planning software. This paper describes the development of a public transport travel costs model for the London Metropolitan Region using the software OpenTripPlanner; and a second private car travel costs model using Ordnance Survey network data and average speeds data from Highways England and the UK Department for Transport. The computational feasibility of modelling multi-modal travel times for a large metropolitan region is demonstrated, with the accuracy of the results validated against surveyed journey-to-work times. The travel time model results can be used in a range of accessibility measures, with mapping and statistical examples provided. Limitations of the method are discussed, including the exclusion of public transport fares and parking costs from the models, and the limited consideration of temporal variation in accessibility.

1. Overview

This paper describes the development of two travel cost models created for the REsilient Systems fOr Land Use TransportatION¹ (RESOLUTION) ESRC research grant, completed in 2018. The project investigated transport inequalities and segregation through a comparative analysis of London and São Paulo metropolitan regions. The transport inequalities aspect of this project required understanding travel cost and accessibility patterns in London for different socio-economic groups and by different travel modes. The two travel cost models developed are a multi-modal public transport and walking travel cost model; and a second private car travel costs model. This paper details how the multi-modal travel cost models were constructed and validated, and how the London Metropolitan Region was defined.

Accessibility analysis can be applied in a variety of research contexts, including equity analysis of travel opportunities between socio-economic groups (El-Geneidy et al., 2016; Van Wee & Geurs, 2011); sustainability analysis of opportunities to travel by transit and active travel modes compared to car trips (Benenson, Martens, Rofé, & Kwartler, 2011; Ford, Barr, Dawson, & James, 2015); and investigating spatial planning questions concerning labour markets and jobs-housing balance (Merlin, 2014). These research topics benefit from detailed data on the travel costs of potential trips by the main modes of travel, ideally at a relatively high spatial scale to enable local spatial differences to be identified. The number of potential trips by various modes of transport at a high spatial resolution generally adds up to millions of potential trips in large cities. Therefore, the costs of these potential trips can only feasibly be calculated through computational modelling methods.

As well as for researching travel opportunities, travel cost models can also be used to augment surveys of observed travel patterns by estimating trip properties that are not recorded in the survey data. In the UK context, a comprehensive survey of origin-destination information is available for journey-to-work travel using the UK census, including the main mode of travel used (defined in terms of the longest distance trip stage). The census does not however record the duration of journey-to-work trips or the money costs of journey-to-work trips. It also omits detailed trip information such as secondary modes of travel used in combination with the main mode. While there are UK surveys available that can supplement the census data, principally the National Travel Survey, the sample size and spatial scale of alternative surveys are much more limited and are not directly comparable to the census. A detailed travel cost model can be used to help estimate these missing census travel cost variables for journey-to-work travel (Smith & Serras, 2012).

An alternative approach to developing new travel cost models would be to use the outputs from an existing transportation model. This approach has some advantages in terms of the sophistication of several existing transportation models. There are however some specific aims of the RESOLUTION project that do not fit with existing models, including the extensive regional study area, the requirement for disaggregation by occupational class and the particular selection of travel modes. A custom model is therefore preferable to address these specific requirements.

2. Study Area Definition

Many research studies of London are based on the municipal geography of the Greater London Authority (GLA), which is the most straightforward geography to adopt in terms of data availability and manageable data volumes. There are however important relationships and interactions that

¹ <https://www.urbantransformations.ox.ac.uk/project/resolution-resilient-systems-for-land-use-transportation/>

cross the GLA boundary which will inevitably be omitted by confining studies to the GLA only (Hall & Pain, 2006; Reades & Smith, 2014). The Census 2011 data (Office for National Statistics, 2016) records 900,000 commuters living outside of the GLA and working inside, and 200,000 commuters living inside the GLA and working outside. Therefore, approximately a million journey-to-work trips crossed the GLA boundary on a typical working day in 2011, and this figure will likely have increased during intervening years with population growth (see Table 1). The reason for the large number of trips crossing the GLA boundary is that Greater London is the core of a much larger metropolitan region with closely integrated labour and housing markets, connected through extensive road, rail, and bus networks. The major orbital motorway, the M25, is just beyond the GLA boundary, as are several international airports, container ports, major business clusters and many towns and small settlements (Reades & Smith, 2014). As the RESOLUTION project is investigating travel inequalities and segregation, there is little sense in isolating the GLA from its metropolitan region, as significant connections and patterns in residential location and accessibility are better understood in the broader regional context.

Another important consideration when modelling accessibility is the presence of edge-effects that occur at the study-area boundary. Unless external trips are modelled, zones at the edge of the study area will be measured as having lower accessibility due to the exclusion of trips to neighbouring destinations outside the study area. Therefore, analysing the GLA in isolation will lead to inaccurate accessibility results for locations in Outer London due to the exclusion of trips to the neighbouring Outer Metropolitan Area. Edge-effects can be minimised through expanding the study area (although this approach pushes edge-effects out to the new study area boundary) and through modelling external trips. The RESOLUTION model covers a large regional study area as described below which removes edge-effects for the GLA. We have not however modelled external trips and so edge-effects will be present at the edge of the larger study area boundary.

There is no official London Metropolitan Region boundary definition, and so a new regional boundary was defined for the purposes of this project. The percentage of workers commuting to the GLA at the Middle Super Output Area (MSOA) scale is mapped in Figure 1, highlighting a commuter belt within 30km of the GLA boundary where around 20-30% of commuters work in the GLA. Beyond the commuter belt is a wider more loosely connected metropolitan region with around 5-15% of commuters working in the GLA. Taking into account the edge-effect issue discussed above, the decision was taken to define a large metropolitan region beyond the commuter belt and including several medium-sized towns that are an hour's rail commute from Central London as shown in Figure 1. The boundary corresponds to a 10% commuting threshold to the GLA using the 2011 census data, with some minor cleaning performed to ensure a contiguous study area.

Note that the relatively large town of Reading to the west of the GLA has been included in the study area, even though it is just below the 10% commuting threshold. Reading is economically significant in the region and given the major recent rail investment connecting Reading and Central London, notably Crossrail which is set to open in 2019, it was decided that including Reading would make the study area boundary more relevant for future research studies.

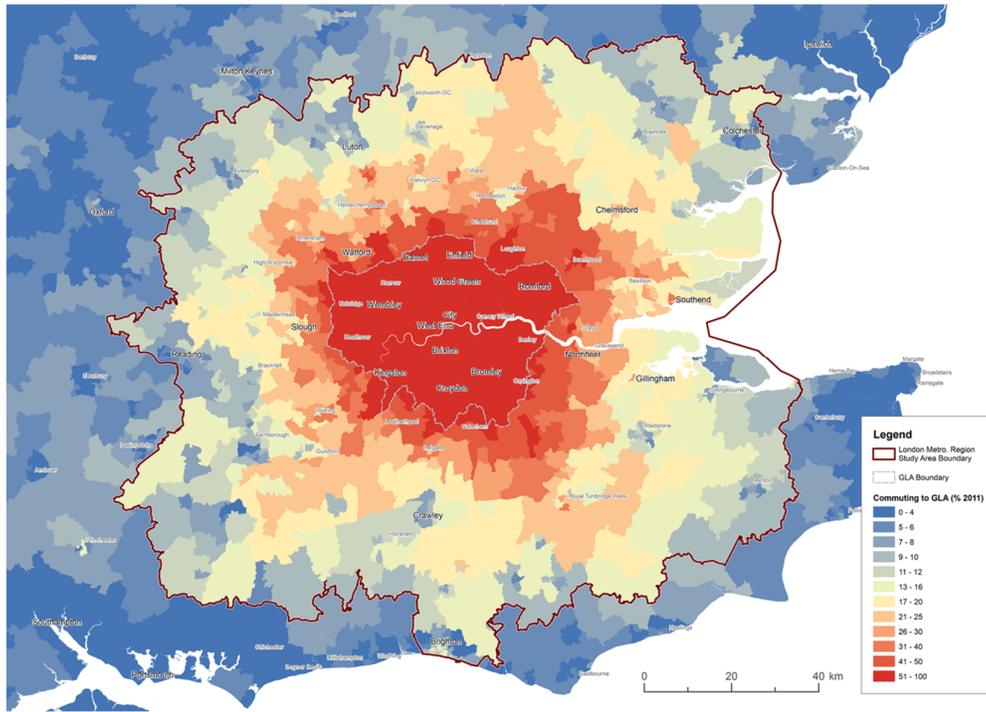


Figure 1: Percentage of Journey-to-Work travel to the GLA in 2011 and Study Area Boundary (percentage excludes home-workers)

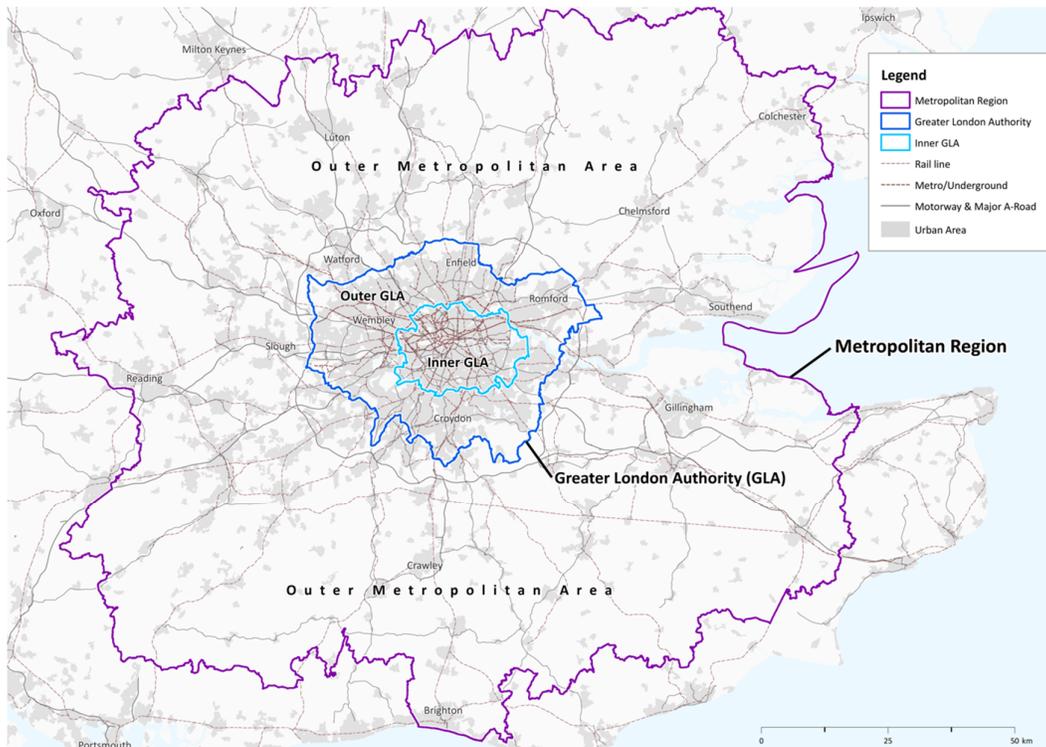


Figure 2: London Metropolitan Region Boundary and Major Transport Infrastructure

The final study area alongside major road and rail networks and urban land use is shown in Figure 2, and population and land area statistics in Table 1. The total population of the Metropolitan Region in 2016 is just short of 17 million, split largely evenly between the GLA at 8.7 million and the Outer Metropolitan Area at 8.2 million. As much of the Outer Metropolitan Area is exurban and semi-rural in character, there will be considerable differences in terms of accessibility opportunities and travel patterns across the study area. The sub-regional definitions shown in Figure 2 are used to disaggregate results, including between Inner and Outer London, to help understand these differences. The Inner London definition is based on the most recent London Plan (Mayor of London, 2016).

Table 1: Population of Study Area Region and Sub-Regions

	Total Area (km ²)	Population 2016 (000's)	Population 2011 (000's)	Pop. Change 2011-2016 (000's)	Working Pop. 2011 (000's)	Workplace Jobs 2011 (000's)
Metro. Region	16,782	16,941	15,946	995 (6.2%)	7,886	8,069
GLA	1,569	8,773	8,159	614 (7.5%)	3,992	4,496
Inner GLA	348	3,523	3,224	299 (9.3%)	1,627	2,662
Outer GLA	1,221	5,250	4,935	315 (6.4%)	2,365	1,834
Outer Metro. Area	15,213	8,168	7,787	382 (4.9%)	3,895	3,573

Data sources: Census 2011 (Office for National Statistics, 2016); Small area population estimates 2016 (Office for National Statistics, 2017).

3. Journey-to-Work Travel Cost Models Scope and Development

3.1 Spatial and Temporal Scope of the Travel Cost Models

The travel cost models need to strike a balance between capturing detail and richness on the one hand, and maintaining computational feasibility on the other. Travel costs vary across space and time as well as by travel mode and journey purpose, and the model scope needs to be defined in relation to all of these dimensions. The RESOLUTION project is focussed on employment accessibility and journey-to-work travel. This is interpreted as modelling the morning peak period only (defined by the UK Department for Transport as 7am-10am). We therefore assume that the travel costs on the work-to-home trip mirror the home-to-work trip. The selection of travel modes is discussed in detail in Section 3.3.

A further important decision relates to handling temporal variation in travel costs within the morning peak period. The car travel costs model is based on average road speeds across the morning peak period (see Section 3.6), and therefore travel time variation by car within the morning peak is not measured. For public transport analysis, the handling of temporal variation is discussed in Section 3.5, based on averaging several queries. Another aspect of temporal scale is the model base-year. The year of 2011 has been chosen to integrate the analysis with the census data being used. Unfortunately, it has not been possible to get the public transport timetable nor the average road speeds data with the same 2011 base year. The implications of this discrepancy are discussed in the sections below.

The spatial scale of the model in terms of the number of origins and destinations considered determines the degree to which small area variation is measured. This decision has considerable computational implications, as the time to calculate the full trip matrix generally increases exponentially with the number of origins and destinations included. The RESOLUTION project is using census data for the demographic disaggregation of results. Therefore, using census zone centroids snapped to the street network to define origins and destinations is the most straightforward way to ensure consistency between the model and the census data. There are several UK census zone types that could be used, from the larger local authority zones to the smallest scale Output Areas. Local authority zones were considered too large to identify accessibility variation. This left the intermediate Middle Super Output Area (MSOA) zones and the smaller Output Area (OA) zones. In the study area there are 1974 MSOAs, with an average population of 8,077 residents and average land area of 8.3km². This compares to 50,000 OAs in the study area, with an average population of 320 residents and land area of 0.3km². The spatial richness of the OA geography is appealing, but the computational implications of 50,000 zones (i.e. an origin-destination matrix of 2.5 billion potential trips for each model run) ruled this choice out. The intermediate Lower Super Output Area (LSOA) zones would also have been a feasible approach, but note that some of the disaggregate census journey-to-work flow data is not published by LSOA. The MSOA geography has proved sufficient in highlighting considerable spatial variation, as shown in the accessibility results in Section 5.

3.2 Defining Travel Costs: Time, Money and Generalised Cost

We can model travel costs simply as travel time, or as generalised costs where the financial costs of travel options and the perceived user costs of different journey stages are represented. There are two related issues here: firstly, how travel costs are modelled when determining optimal routes

between origin and destination locations; and secondly, how the travel costs of these modelled trips are output for the calculation of accessibility measures. In regards to the former issue, representing costs only as travel time and not generalised cost will inevitably omit important aspects of travel behaviour, such as minimising the number of transit interchanges and avoiding long walking stages. In regards to the latter, there can be advantages in defining accessibility measures using straightforward travel time in terms of ease of communication and understanding. The cumulative accessibility measures shown in Section 5 are effective using travel time units for example.

The software used here for modelling public transport travel, OpenTripPlanner, uses a generalised costs approach for optimising transit routing which allows variable costs on different journey stages and additional costs for interchanges (see Section 3.4). The capabilities of OpenTripPlanner to model public transport fares are however very limited. The calculation of public transport fares is further hindered by a lack of standardised data in the study area, with complicated fare systems varying by bus, underground and rail travel, and additional differences between public transport operating companies. A further layer of complication is the array travel cards and passes which are very common for commuters and have different systems of payment to individual fares.

Modelling public transport fares was not feasible within the time limits of this project. To address the shortcoming of omitting money costs, we considered that the financial costs of commuting are closely linked to the choice of travel mode. It is possible to consider aspects of commuting affordability by modelling the most affordable travel modes independently, which allows accessibility comparisons to be made between the most affordable and the least affordable travel mode options. In the case of the London Metropolitan Region, the most affordable modes are walking and bus (cycling is not modelled here as discussed in the next section). This is reflected in the fact that rates of commuting by bus travel and walking are twice as high in the study area for the two least affluent occupational classes compared to the two most affluent occupational classes. Therefore, as discussed in the next section, bus travel and walking travel are modelled separately, allowing consideration of accessibility by the more affordable travel modes.

3.3 Travel Mode Selection

The selection of travel modes modelled needs to reflect current patterns in journey-to-work travel, as well as the travel affordability considerations discussed in the previous section. The data on journey-to-work modal split in the study area sub-regions is shown in Table 2. There are clear patterns in terms of the most popular modes, with car, bus, rail, metro and walking together comprising 95% of all commuting trips in the study region. Note that in the London context 'metro' refers mainly to the London Underground (Tube) network, but also includes the Docklands Light Railway, the Croydon Tram and the London Overground.

Cycling was just over 3% of trips in 2011 and ideally should also have been included in this study. There are however particular modelling and data challenges in regard to cycling infrastructure, with inconsistent and often very poor segregated cycle-lane provision available across the study area, and indeed the UK in general. Without accurately modelling cycling infrastructure, travel cost models will typically over-estimate bike accessibility by assuming that most roads are available for cycling, when there is minimal or no cycle segregation available from busy traffic. There was not sufficient time to model and to validate cycling costs in detail for this project and cycling has not been included. There are several good recent studies discussing cycling accessibility in detail showing how this could be tackled (Lovelace et al., 2017; Vale, Saraiva, & Pereira, 2016).

Table 2: Journey-to-Work Main Mode¹ Percentage by Sub-Region of Residence 2011

	Car ²	Rail	Metro ³	Bus	Walking	Cycling	Other ⁴
Metro. Reg.	49.7	12.6	12.7	9.6	9.9	3.2	2.2
Inner GLA	16.2	13.2	30.3	18.2	12.2	7.0	2.9
Outer GLA	41.6	14.6	19.3	12.5	7.3	2.4	2.2
OMA	68.9	11.2	1.2	4.1	10.6	2.1	1.9

Data source: Census 2011 (Office for National Statistics, 2016).

¹Main mode is defined as longest distance stage of journey-to-work.

²Car is combined car driver and car passenger trips.

³Metro includes London Underground, and should include DRL, Overground and Croydon Tram.

⁴Other category includes motorbike, taxi and 'other' category from census.

There is a second related travel mode issue of how multi-modal journey-to-work trips are handled in the models. Public transport trips nearly always include walking stages and so the modelling of these trips needs to include measures of pedestrian costs based on street network data. It is also very common for combinations of public transport modes to occur in a single journey-to-work, most commonly for connections to and from rail and metro stations. For example, nearly half of all rail commutes also include an underground or bus stage of the journey (Smith & Serras, 2012). Therefore, multi-modal public transport journeys between rail, metro, bus and pedestrian networks need to be modelled. In addition to multi-modal public transport journeys, it is also useful to be able to model journeys that are restricted to particular public transport modes, to model more affordable journey options, particularly bus travel, as discussed previously.

A further multi-modal issue is whether to allow combinations of car and public transport travel, generally known as 'park-and-ride' and 'kiss-and-ride' journeys in the transport literature. This is a financially costly behaviour for commuters in the London Metro Region, requiring the payment of public transport fares as well as car ownership, running costs and parking costs in the case of park-and-ride trips. These types of trips are most likely to appeal to wealthier commuters, particularly those living in less accessible parts of the Outer Metropolitan Area where combining car and rail travel can overcome limited public transport access to OMA rail stations. Unfortunately, it is not straightforward to model these trips types for this project, as private car costs and public transport costs are being modelled separately using different networks and software, as discussed below. There is also a lack of detailed data on how frequently car-transit trip combinations are occurring and the relative distances of the journey stages, making it problematic to validate model outputs for these types of journey. Finally, the ability to make these journeys is heavily dependent on parking costs and capacity, which would require building an additional parking model for the study region. As a result, this combination of travel modes is not being modelled. These journeys are modelled as either exclusively public transport trips or as exclusively car trips depending on how the main mode is recorded in the census. This approach will likely lead to some underestimation of public transport accessibility in the OMA, at least from the perspective of more affluent commuters able to combine car and rail commutes.

3.4 Multi-modal Public Transport Modelling Using OpenTripPlanner and TransXchange Timetable Data

The multi-modal public transport model needs to be able to identify realistic routes between thousands of origin and destination locations, accurately measuring the travel time for the various journey stages of these routes. Journey stages include pedestrian stages; in-vehicle stages; and any waiting stages depending on service frequency and interchange behaviour. This is non-trivial modelling challenge, particularly given the large scale of the public transport networks in city-regions like London. There have been several recent advances in public transport data availability and routing software that can help with this modelling task.

Public transport timetable and station location data is increasingly available as open data. The popularity of online journey planners and development of Google's General Transit Feed Specification (GTFS) has helped standardise transit data internationally (Antrim & Barbeau, 2013) and encouraged the widespread release of public transport data by city authorities. GTFS data includes the core information needed to model public transport travel times, including transit station location geography and timetable data with service frequency. Complementary to the expansion of GTFS data, there has been the release of detailed street network data through OpenStreetMap for modelling pedestrian and cycling accessibility (Haklay & Weber, 2008). In the UK content, comprehensive public transport timetable data is available from the Traveline Database in the TransXchange XML format developed by the Department for Transport². For this project, the TransXchange data was converted to the GTFS XML format by CASA researcher Richard Milton. The Traveline data is unfortunately not easily available as an historic archive, and so the timetable data used here is from 2016 rather than describing public transport services as they existed in 2011. While public transport service changes between 2011 and 2016 are relatively minor, this is a potential source of error.

The second key advance for modelling public transport has been progress in open source journey planning and network routing software. It is computationally challenging to quickly measure optimal routes through large urban public transport networks taking into account the geography of stations, transit service frequency, interchange delays and pedestrian accessibility (Bast, 2009). Journey planner software relies on algorithms from network science to identify shortest paths to destinations (Delling, Sanders, Schultes, & Wagner, 2009). The challenge here is to repeat this process for thousands of origins and destinations across the entire study area. Open source software projects such as OpenTripPlanner³ have been working to solve these computational problems, and have been successfully used in accessibility analysis of metro regions, particularly for studies in the USA, for example the Access Across America studies (Owen, Murphy, & Levinson, 2018).

OpenTripPlanner (OTP) has been selected as the software tool for modelling public transport in this project. In addition to being open source, OTP advantages include the ability to model the full timetable of public transport services for the entire London Metropolitan Region, including main-line rail, underground, overground, DLR and bus services. OTP also has the ability to accurately model pedestrian travel costs using OpenStreetMap data, including data on the presence of pavements on individual roads. OTP routing is based on a generalised cost approach where penalties are applied to interchanges and variable costs can be applied to time taken by particular modes, such as walking stages, to more accurately reflect travel behaviour. The software is generally well supported with

² <https://www.gov.uk/government/collections/transxchange>

³ <http://www.opentripplanner.org>

online documentation⁴ and support forums⁵, including a graphical journey planner interface to test how the network routing is performing, as shown below in Figure 3.

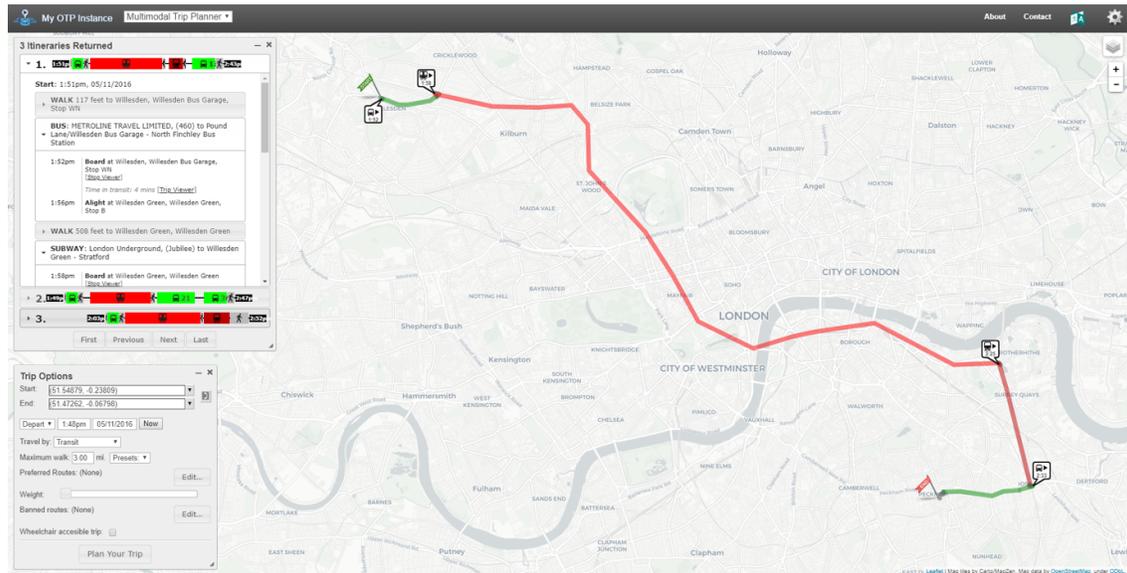


Figure 3: OpenTripPlanner Journey Planner Web Interface Example for Inner London Journey-to-Work Trip

The computational performance of OTP is reasonable, though in this project performance is affected by the extensive London Metropolitan Region study area and the large size of the timetable. A typical model run on a desktop computer for the full London Metropolitan Region study area MSOA origin-destination trip matrix takes around 10 hours for a 2 hour maximum travel time limit and 20 hours for a 5 hour maximum travel time limit. The 10 hour plus calculation time for each model run limits the ability to perform detailed analysis of temporal variation in travel time. Analysis of temporal variation is typically based on tens to hundreds of model runs. A future solution would be to parallelise the OpenTripPlanner software.

Other limitations of the public transport modelling approach used here include relying on the accuracy of timetable data. This means that issues like transit service reliability cannot be investigated without additional data. As discussed earlier, OpenTripPlanner also has only limited capabilities for modelling public transport fares and this aspect has not been modelled.

3.5 Configuring OpenTripPlanner

There are many options available to configure the OpenTripPlanner setup for particular study areas. The latest release version of OTP, version 1.2 at the time of writing, was used here. The OTP workflow is in two main stages: firstly, building the study area network, referred to as a *router*, from the OpenStreetMap street network and GTFS transit stop and timetable data; and secondly, querying the router with the locations and times of interest, either in the form of a journey planner interface, as shown earlier in Figure 3, or programmatically in batch mode through running a script.

⁴ <http://docs.opentripplanner.org/en/latest/>

⁵ <https://groups.google.com/forum/#!forum/opentripplanner-users>

The script used here was based on the online example provided by Rafael Pereira⁶, and is provided in the Appendix.

During the network building stage, there is a configuration property ‘maxTransferDistance’ which is influential in determining the size of the network built and subsequent computational demands in querying the network. The maxTransferDistance property determines the maximum distance users can walk to make a transit connection, excluding the first and last walking stages of the journey. Values above 1km greatly increase the network size and slowed down the analysis without any noticeable benefits to the accuracy of results. The default value of 600m was used here.

The default configuration settings in terms of walking speed (4.8kmph), walk reluctance (twice the cost of time spend on public transport) and interchange costs (interchanges must achieve time savings of at least 10 minutes to be chosen), were found to work well for the London Metropolitan Region, producing realistic routing results and travel times, as shown in the validation section. One default property that required tweaking for the London study region was the cost of accessing underground platforms from street level; a property named ‘SubwayAccessTime’. While ideally this property should be calculated individually for each station and platform, this would be time consuming to implement given the 270 stations and 11 lines on the London Underground network. Instead a standardised cost of 4.5 minutes was applied, higher than the default value of 2 minutes. Note that this time penalty is only applied once and represents the combined cost of accessing the underground platform from street level and of exiting the underground network to street level. Averaging out this property will tend to overestimate access costs for smaller suburban stations, and underestimate them for larger busy central stations.

The model has to be run for a particular day and time. A Wednesday was selected on a week in September outside of holiday periods to ensure a full public transport service. There are two types of query available: Arrive By and Depart At queries. A comparison of average travel times for an 8am Depart At and a 9am Arrive By query are shown in Table 3 below, with the mean travel time results weighted by the public transport commuting trips recorded in the 2011 census.

Table 3: OpenTripPlanner Arrive By and Depart At Query Average Journey-to-Work Time Comparison for Study Area

Query Type	2011 Rail Commutes Mean (mins)	2011 Metro Commutes Mean (mins)	2011 Bus Commutes Mean (mins)
Depart At 8.00am	67.7	47.9	39.6
Arrive By 9.00am	70.0	48.4	39.5

We can see the aggregate times are very similar in Table 3, with marginally higher results for the Arrive By query for rail commutes. This discrepancy is most likely a result of some of the longer duration rail trips beginning before 7am for the Arrive By query, when service frequencies are typically lower. Arrive By queries are closer to how most employees plan journey-to-work routines to meet their job start times. Therefore, Arrive By queries have been used in the final model. Note that the time recorded by OTP is the journey time of the ‘best’ trip itinerary available, selected in terms of total journey time and its closeness to the depart at/arrive by time. The final recorded time does not include the time difference between when the itinerary starts/ends and the depart

⁶ <https://github.com/rafapereirabr/otp-travel-time-matrix>

at/arrive by time of the query. No results are recorded if there are no valid itineraries that start/end within around thirty minutes of the depart at/arrive by time. The author has not able to identify the exact criteria for how itineraries are excluded in OTP.

The final query decision to make is regarding temporal variation across the morning peak period. A comprehensive analysis would measure accessibility at a fine temporal scale (e.g. every minute) across the morning peak time window, then combine the results, ideally weighted by typical job start times. Unfortunately, the OTP method used here is computationally too slow for fine scale temporal analysis given the large study area. A simpler approach has been used, averaging five Arrive By queries at 15 minute intervals between 8.30am and 9.30am inclusive. For further research, it would be valuable to investigate work trips outside of the morning peak, such as early shift work, and off-peak accessibility. In this project we have only considered accessibility during the main peak period.

Aggregate commute time differences between Arrive By queries across the morning peak period are shown below in Table 4. Overall the aggregate differences are minimal between the queries, implying that averaging results between 8.30am and 9.30am will have only a marginal impact on the overall results. There are however likely to be some differences at a disaggregate level for particular locations in the study area where service frequencies are low. One trend that can be seen in the aggregate data is that average commute times are marginally higher for arrival times after 9am, indicating that service frequencies are probably tailing off slightly by 9.30am.

Table 4: OpenTripPlanner Arrive By Average Journey-to-Work Time Temporal Variation Test

Query Type	2011 Rail Commutes	2011 Metro Commutes	2011 Bus Commutes
	Mean (mins)	Mean (mins)	Mean (mins)
Arrive By 8.30am	70.2	48.3	39.1
Arrive By 9.00am	70.0	48.4	39.5
Arrive By 9.30am	70.6	48.8	40.3

3.6 Modelling Bus Trips in OpenTripPlanner

GTFS data records the mode of transit services in the 'Route Type' field. This property is used by OpenTripPlanner to enable journey time queries to be restricted to particular travel modes. By switching the set modes function from Transit (all modes) to Bus, OTP can model bus-only accessibility and walk-only accessibility, as discussed earlier.

The travel time results of the bus-only model were found to be quicker than recorded survey data on commute times. This discrepancy is likely due to the bus timetable not adequately accounting for congested peak time conditions in London. A more sophisticated model could capture where these extra delays are likely to occur and apply specific penalties. In this research a simpler approach was taken of applying an extra 7% delay to all bus trips to bring the results in line with the Labour Force Survey (see Section 4.1). Clearly this approach is a generalisation and will average out spatial variation in how congestion affects timetable reliability.

3.7 Modelling Walking Trips in OpenTripPlanner

Walking trips can also be modelled using the OpenTripPlanner software by removing all transit modes. The travel time estimates are based on the OpenStreetMap network and a walking speed of 4.8 kmph. These basic assumptions should produce realistic travel times. It was however difficult to validate the walking time results based on census data, as the census 2011 data records a high number of long distance trips as walking trips (e.g. 20% of walking trips have travel times estimated at more than 1 hour). A considerable proportion of these trips likely include a car or public transport leg as part of the journey that is not recorded, but without further data on how many it is difficult to validate the walk times.

3.8 Private Transport Modelling Using Average Speed Data

Many roads in London are heavily affected by congestion during peak times. Data for congested Inner London roads shows average speeds during the morning peak falling below 10 mph/16 kmph (Transport for London, 2016), and even approaching 5 mph in the worst cases. The most congested routes are so slow that current average speeds are comparable to pre-automobile eras. Given the high levels of congestion on some routes in London, it is essential for car travel time calculations to model congestion for realistic results. This can be achieved by building a full road transport model, including capacity and congestion. This is a powerful and flexible approach, but is also time consuming to develop. Alternatively, average speed data describing recent road conditions can be used. The second approach is more straightforward, particularly given that good quality average road speed datasets are increasingly available, and is pursued here. Note this approach does not allow the testing of future scenarios where travel outcomes diverge from current patterns of congestion.

Including average speed data will help the model produce realistic car travel times. However, the money aspects of car travel are clearly important, particularly in relation to parking costs, and the London Congestion Charge. These additional costs suppress car use, and produce behaviours such as parking further away from workplaces if parking opportunities are cheaper, and avoiding the Congestion Charge Zone (though note that Central London is so severely congested during the morning peak that fastest routes will generally avoid this zone anyway). This model does not include money costs and is based purely on travel time; it is effectively modelling car trips where parking is available close to workplaces. Arguably this means that car accessibility will generally be overestimated, although the model is validated against recorded travel times from surveys and produces reasonably accurate results (see validation in Section 4). A standard cost of five minutes has been added to all car trips as a basic walk time cost to and from parking locations.

Before integrating the average road speed data, the base road network dataset has to be selected. In the public transport model described in the previous section, OpenStreetMap (OSM) data was used for modelling pedestrian accessibility through the street network. There would be consistency advantages in also using the OSM network for the private transport model (for example it would facilitate modelling park-and-ride car-transit trips in future iterations of the model). The OSM data is not however primarily designed for vehicle routing and there are some minor network connectivity errors that emerge when performing long distance network routing using the OSM data. The UK national mapping agency Ordnance Survey has several road network products designed primarily for

vehicle routing applications that are available as open data. The Meridian 2⁷ dataset has been selected here.

Two average speed datasets have been combined to model average road speeds. The first is the Highways England dataset⁸, which covers motorways and major A-roads. The second is Department for Transport data which covers the entire A-road network. It is not ideal to use two different data sources as there will likely be minor methodological differences between them in the calculation of average speed. These datasets are however the only open data average speed options currently available (there are commercial alternatives). The Highways England data is very comprehensive in terms of temporal coverage, with continuous coverage at 30 minute intervals, and availability for multiple base years. The Department for Transport data is available as monthly average morning peak speeds (7am-10am) going back to 2008⁹ (Table CGN0209). The DfT data has some missing average speed values, with missing values more prevalent for the earlier data years. There was therefore a trade-off between matching census base year for the model (2011) in the DfT data, or selecting a more complete average road speed dataset for a more recent year. The second option was favoured, with DfT average speed data for September 2014 being used. The Highways England data was averaged for the same September 2014 base period, using the same 7am-10am definition of the morning peak period.

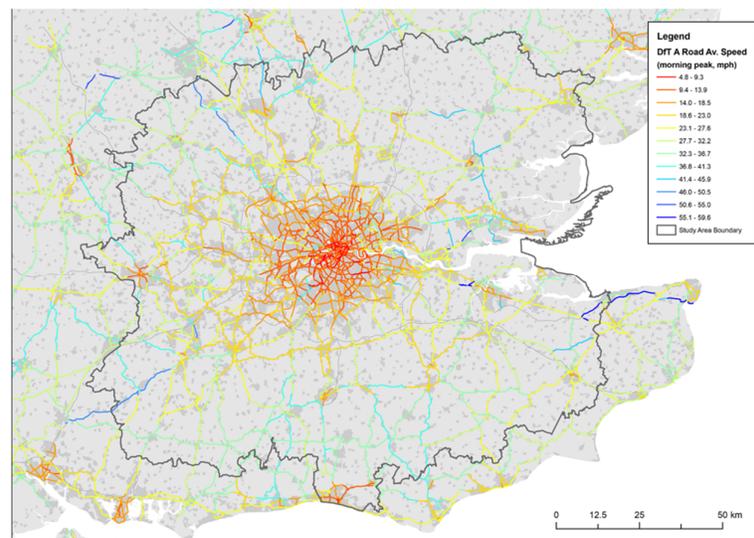
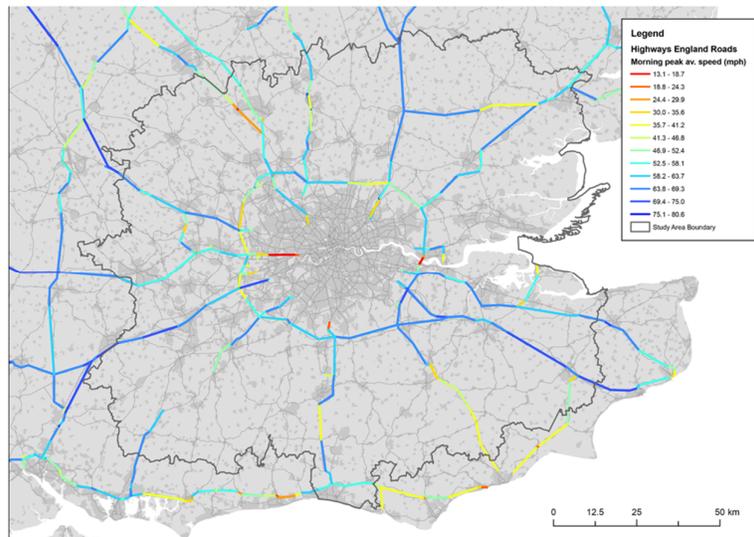
The two average speed datasets needed to be joined to the base road network data using GIS software. The Highways England data includes a simplified spatial road network with each straight line segment representing a separate bi-directional average speed observation, as shown in Figure 4. The Highways England network is not referenced to other road geometry data (beyond including a road name attribute). To link the Highways England data to the Meridian 2 data, a nearest spatial join was performed. Where the Meridian 2 network overlapped with multiple Highways England segments, then the average of the overlapping segments was taken. There are some minor errors generated with this approach, and manual editing was required to fix these. The spatial join average should ideally have been weighted by segment length, but there was not time to add this. The calculated speeds from the Highways England data are shown in Figure 4, and are generally around the 40 to 70 mph range which is as expected for major motorways during the morning peak. Note that some well-known congestion bottlenecks are evident, including the Queen Elizabeth II Bridge across the Thames in East London, and a number of links in West London, particularly the M4.

The Department for Transport average speed data is even more lacking in spatial referencing information than the Highways England data. It includes only the road name/classification and the local authority where the average speed measurement was made. Joining data using road classification alone would be a weak option as each road class can represent routes of up to several hundred miles long. One way to minimise this problem is to use the average speed data by local authority and road classification together so that the joins are based on the section of particular A-roads within specific local authority areas. The smallest local authorities are the 33 London Boroughs which also happen to be the most congested areas of the study area. This coincidence reduces the aggregation problems resulting from this approach, though cannot entirely remove them. It would be much better if the DfT could release the spatial network data they are using as open data with a shared ID with the average speeds data, but this is not currently available.

⁷ <https://www.ordnancesurvey.co.uk/business-and-government/help-and-support/products/meridian2.html>

⁸ <https://data.gov.uk/dataset/9562c512-4a0b-45ee-b6ad-afc0f99b841f/highways-england-network-journey-time-and-traffic-flow-data>

⁹ <https://www.gov.uk/government/collections/road-congestion-and-reliability-statistics>



Figures 4 & 5: Highways England Average Motorway Speeds Data (top) and Department for Transport A-Road Average Speed Data Joined to OS Meridian Network (bottom)

The spatially joined DfT average speeds data is shown in Figure 5. Even given the spatial referencing issues discussed, the congestion patterns are clearly visible with very high levels of congestion in Inner London with speeds falling below 10 mph on several routes. There is also a consistent pattern of increasing average speeds as the distance from Inner London increases. Note also that some more subtle patterns, such as higher speeds on the North Circular ring road at the edge of Inner London can also be identified.

As both the datasets have been joined to Meridian 2 data, then combining the two datasets together as a single network is straightforward. Network routing testing was performed in ArcGIS Network Analyst to check that the network was producing realistic results. While there are a number of good open source routing software options available, the nature of the different datasets being combined for this model meant that a comprehensive visual interface was helpful for testing. Examples of quickest route choices through the road network are shown in Figure 6 below.

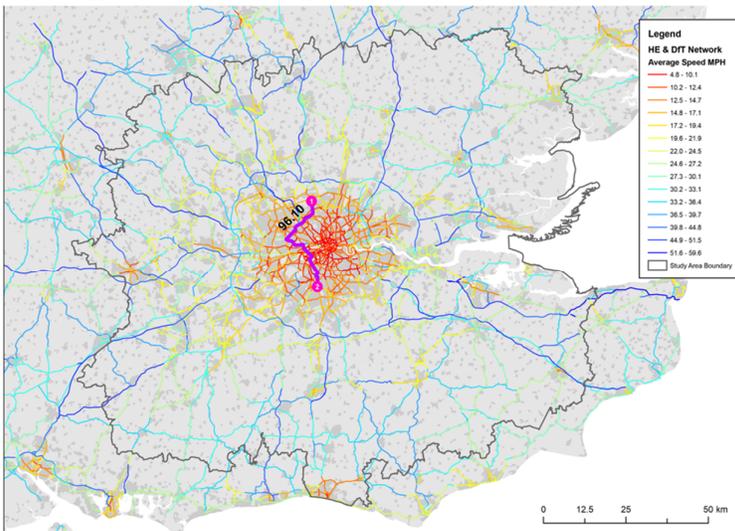
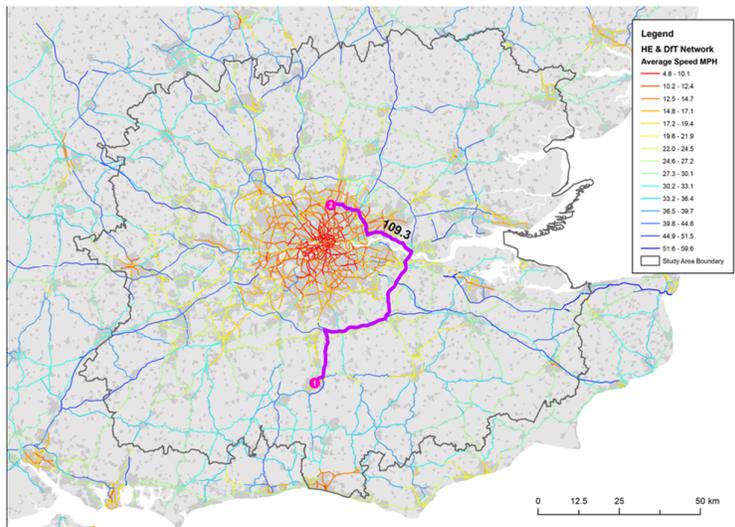
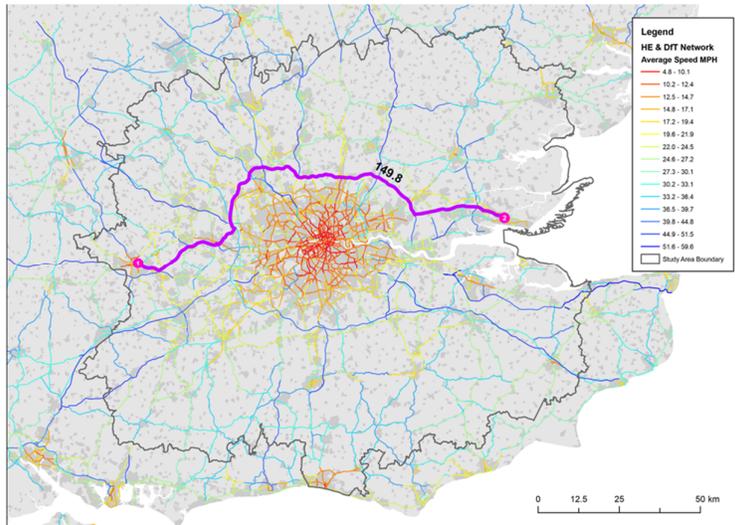


Figure 6: Private Transport Average Speed Model Example Routing Results: Reading to Southend (top); Crawley to Wood Green (middle); Barnet to Sutton (bottom).

As shown in Figure 6, the combination of the motorway and A-road speeds data provides a good core network for modelling realistic car journeys across the study area. The purple lines in the Figure 6 examples show the quickest route from two example locations during the morning peak period. The model captures typical driving route strategies such as using the orbital M25 motorway to avoid high congestion in Greater London, as well as selecting priority links such as the North Circular inner ring-road to avoid Inner London. The numbers in black in Figure 6 next to the purple line shows the total travel time in minutes for the route. For readers who are not familiar with London traffic these times may appear high, but are in fact quite close to survey results as discussed in the Section 4.

A shortcoming of the modelling approach taken is that average speed data on roads outside of the motorway and A-road network (B-roads, minor roads etc.) is not available. These roads are important, particularly in more rural areas where the a-road network is sparse. B-roads and minor roads were included by assigning speeds 15% below the closest a-roads, with speeds capped at legal limits. This rough approach is not ideal and should be improved for future versions of the model. To minimise errors resulting from minor roads overly-influencing the travel time results, the routing is based on a hierarchical approach in ArcGIS Network Analyst, where route choices prioritise road types at the top of the hierarchy: the a-roads and motorway network. This minimises the number of routes that are focussed on minor roads.

The first version of the model produced average journey times that were too quick compared to survey data of morning peak car travel times. A further 11% travel time cost was applied to all trips to bring travel times in line with the Labour Force Survey, which is discussed in Section 4.1. One possible reason for the model under estimating car travel times is that the average speeds are calculated across the period 7am to 10am, whereas most commuting trips are more concentrated between 7.30am to 9am when congestion is higher. Other possible reasons for the underestimation is that junction delays are not being explicitly modelled in the car travel times, and that the 5 minutes included for walk times to and from parking locations is too low.

4. Travel Cost Models Validation

To validate the results of the public transport and car travel time models, the model results are tested against average travel times from the Labour Force Survey in Section 4.1, and then against publicly available journey planner results in Section 4.2.

4.1 Model Travel Time Validation Against Labour Force Survey

Information on typical journey-to-work times is needed to validate the model results. The Labour Force Survey (LFS) is a quarterly survey of 40,000 employees across Great Britain that records self-reported journey-to-work times (Office for National Statistics, 2018). Travel time information is available by main mode, and by region of workplace, including London (GLA) and a further sub-regional breakdown into Central London, Rest of Inner London and Outer London. The LFS recorded travel times for London, Central London and Outer London are shown in Tables 5-7. The general expected patterns of higher times by public transport compared to car travel, and higher times for longer trips to Central London, are evident. Note that the travel times in Tables 5-7 are the average of the Autumn surveys between 2008 to 2013. This averaging is to reduce sample errors resulting from the relatively small sample size of the LFS.

A comparison of average modelled journey-to-work times compared to the LFS results is provided in Tables 5-7. The modelled results are weighted by 2011 census commuting data with the workplace destination matched to the LFS region for each table. In general, the private car and public transport models perform well in terms of average travel times compared to the survey, with differences of around 1 minute and no more than 2 minutes. The public transport model results are generally strong for rail and metro trips. Average rail results are slightly too high for Central London trips (72.1 minutes compared to 70.9), and slightly too low for Outer London (61.5 minutes compared to 62.9), but overall that is a good result for these more complicated trips that are often multi-modal. Metro results are also a good match for Central London trips (47.0 minutes compared to 47.7) and a bit too low for Outer London trips (53.7 minutes compared to 55.7). The bus average times are close for GLA trips (42.2 minutes compared to 42.0) but over-predict times for Outer London trips (41.6 minutes compared to 39.2). The latter error is likely due to applying a single additional 7% cost to bus trips to account for the effect of congestion on timetable reliability, which improves Inner London results but over-predicts for Outer London trips.

Considering the relative simplicity of the car travel time model, the average time results are close to the LFS data. The Central London trip time average is accurate (54.2 minutes compared to 54.8). There is a modest error for Outer London car trips (29.4 minutes compared to 30.7) perhaps due to the model overcompensating for congested Inner London trips.

Overall these are very encouraging validation results for the models. There are of course limitations with this validation methods, where a single mean travel time is compared rather than the distribution of all travel times. Distributional data is not published for the Labour Force Survey, even in the form of basic descriptive statistics like standard deviations and deciles. The next section validates the model results against specific trips.

Table 5: Modelled Average GLA Commute Times Compared to Labour Force Survey (2008-2013 average)

Main Travel Mode	Modelled Average (mins)	Labour Force Survey (mins)	Difference (mins)
Car	35.6	35.3	+0.3
Bus	42.2	42.0	+0.2
Rail	70.0	69.1	+0.9
Metro/Tube	48.4	49.5	-1.1

Table 6: Modelled Average Central London Commute Times Compared to Labour Force Survey (2008-2013 average)

Main Travel Mode	Modelled Average (mins)	Labour Force Survey (mins)	Difference (mins)
Car	54.2	54.8	-0.6
Bus	45.3	45.6	-0.3
Rail	72.1	70.9	+1.2
Metro/Tube	47.0	47.7	-0.7

Table 7: Modelled Average Outer London Commute Times Compared to Labour Force Survey (2008-2013 average)

Main Travel Mode	Modelled Average (mins)	Labour Force Survey (mins)	Difference (mins)
Car	29.4	30.7	-1.3
Bus	41.6	39.2	+2.4
Rail	61.5	62.9	-1.4
Metro/Tube	53.7	55.7	-2.0

4.2 Model Travel Time Validation Against Journey Planner Results

To supplement the comparison of average journey-to-work times, we can also examine travel times for particular trips in more detail. Here we use public journey planning websites as a source of travel times to compare model results to. The web services used are Google Maps¹⁰ and Transport for London (TfL) Journey Planner¹¹. These are very popular services used by millions of travellers that have been developed over the last five years using high quality data. Therefore, we would expect

¹⁰ <https://www.google.co.uk/maps>

¹¹ <https://tfl.gov.uk/plan-a-journey/>

these services to have accurate results. Another issue to bear in mind is that Google Maps and TfL are based on 2018 times, while the model results are 2014 car congestion and 2016 timetable data.

The results of the validation exercise are shown in Table 8, with travel time for the private car and public transport models compared to journey planner results for 12 different journey-to-work trips in the study area, grouped into four general types. The modelled public transport results are close to the journey planner results for the range of trips tested in Table 8. These are generally good results for the OTP based model. The OpenTripPlanner travel time results tend to be at the lower end of the journey planner results, though there is no clear pattern in relation to the particular trip types tested. Note there are some differences between the Google Maps and the TfL Journey Planner results, with the TfL site generally on the more conservative side in its public transport travel time estimates.

Table 8: Comparison of Modelled Car and Public Transport Times Against Online Journey Planners for 12 Example Trips

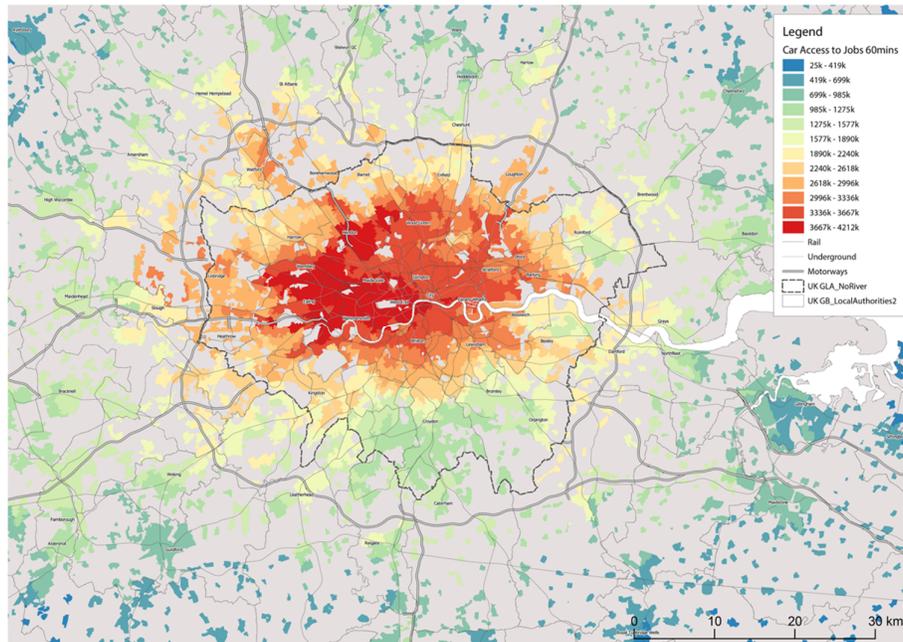
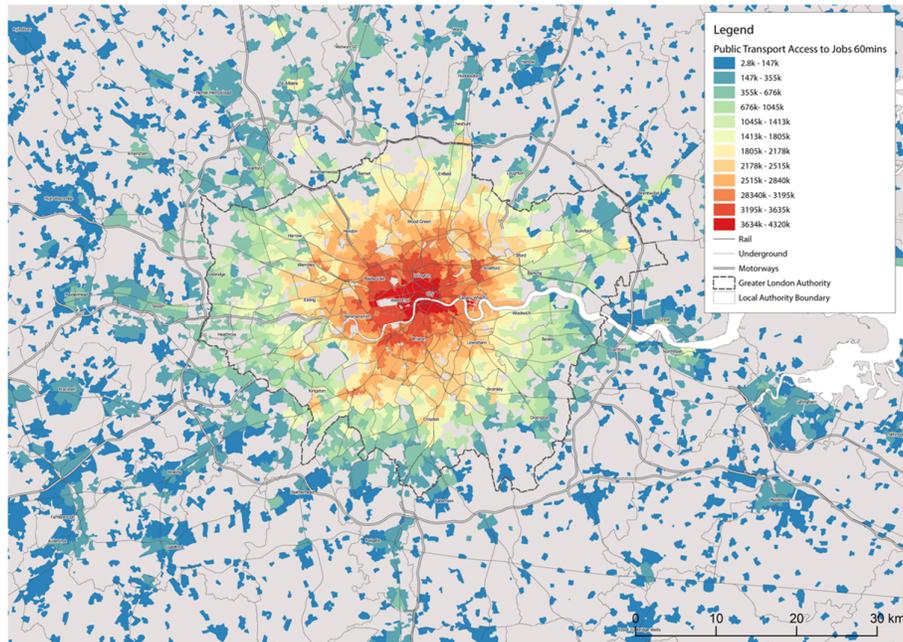
Trip Description		Car Time (mins)		Public Transport Time (mins)		
		Model	Google Maps	Model	Google Maps	TfL Journey Planner
Short Dist. Inner London	Borough to City of London	9.0	12 - 30	17.3	16	21
	Holloway to Bloomsbury	22.8	20 - 50	31.9	39	28
Medium Dist. Radial	Croydon to City of London	61.9	55 - 120	55.4	55	55
	Islington to Elephant & Castle	33.6	26 - 60	35.1	35	43
Medium Dist. Orbital	Bromley to Wimbledon	68.6	45 - 100	59.0	68	81
	Gillingham to Gatwick	87.6	70 - 110	128.4	117	121
	Hemel Hempstead to Slough	46.1	45 - 110	118.0	126	124
	Ealing to Wandsworth	20.3	22 - 55	40.3	30	37
Long Dist. Radial	Reading to Westminster	103.4	90 - 150	66.0	69	70
	Southend to Camden	122.9	110 - 230	94.8	92	101
	Harrow to London Bridge	65.1	70 - 130	60.8	61	68
	Luton to Holborn	88.5	75 - 150	49.8	51	62

The car travel times are not quite as successful as the public transport in this validation. Google Maps produces a very wide range of car travel times, reflecting the potential for high levels of congestion in London. The majority of the car model results are within the range predicted by Google Maps, and are generally close to the minimum travel time, reflecting typical times without severe congestion. But travel times for several trips are under-predicted by the model, particularly for shorter trips. This may relate to the Google site including a longer minimum time for parking by default. Increased levels of congestion in recent years is also likely contributing to longer trip times in the Google Maps estimates.

5. Accessibility Model Results

5.1 Cumulative Employment Accessibility by Car, Public Transport and Bus Modes

Following the validation process, we can now perform some accessibility analysis with the travel cost model outputs. The most straightforward type of accessibility measure is the cumulative measure, where the total number of opportunities are totalled within a travel time threshold. This is a good approach to illustrate the model. In the figures below, the total number of jobs accessible within 60 minutes' travel time is shown for the different travel modes: Public Transport, Car and Bus travel. Note the legend classification is not standardised between figures and is tailored to each mode.



Figures 7 & 8: Total Jobs Accessible in 60 minutes by Public Transport (top) and by Car (bottom)

The public transport and car jobs accessibility in Figures 7 and 8 make for an interesting comparison. Both modes peak at just over 4 million jobs. High public transport accessibility is as expected concentrated in Inner London, with corridors of better accessibility evident along radial corridors to Outer London hubs like Croydon, Wembley and Romford. The pattern of public transport accessibility is relatively even between North and South London. The general pattern is one of falling accessibility as distances from the city centre increase: from a peak of 4.3 million jobs in the City of London, to around 1 million jobs at the GLA boundary. Beyond the GLA boundary accessibility is generally quite low, although areas of relatively better accessibility can be seen in some Outer Metropolitan Area centres like St Albans and Slough. Note most OMA radial transit commutes are longer than 60 minutes (Table 5) and an increased threshold would better reflect these.

The pattern of car accessibility has many differences to public transport. Peak accessibility is spread across Inner and West London, particularly around the North Circular ring road, and where radial motorways meet the M25 orbital motorway, at locations like Watford and Ealing. Car accessibility generally tails off much more slowly than public transport accessibility as distances increase from the city centre. For example, there are several town centres around the GLA boundary, such as Uxbridge and Watford, where car accessibility is over 2.5 million jobs while public transport accessibility is around half that figure at just over 1 million jobs. There is a very noticeable car accessibility contrast between North and South London that is not evident in the public transport analysis. Car accessibility is considerably lower in South London. This is mainly because there are no radial motorway connections in South London and no highway equivalent to the North Circular. For example, public transport accessibility to jobs in South London town centres Croydon and Bromley is over 2 million, compared to less than 1.5 million by car. There is also a bias towards Outer West London, reflecting better motorway infrastructure here and the concentration of jobs around Heathrow Airport and the 'Western Crescent' of business centres like Reading and Slough (Reades & Smith, 2014).

The final accessibility mode map for Bus accessibility to public transport is shown in Figure 9. The bus pattern is effectively an exaggerated version of the public transport map, with accessibility tailing-off very rapidly outside of Inner London. Note also that the peak bus values are much lower, at 2.6 million jobs compared to over 4 million jobs for the other two modes.

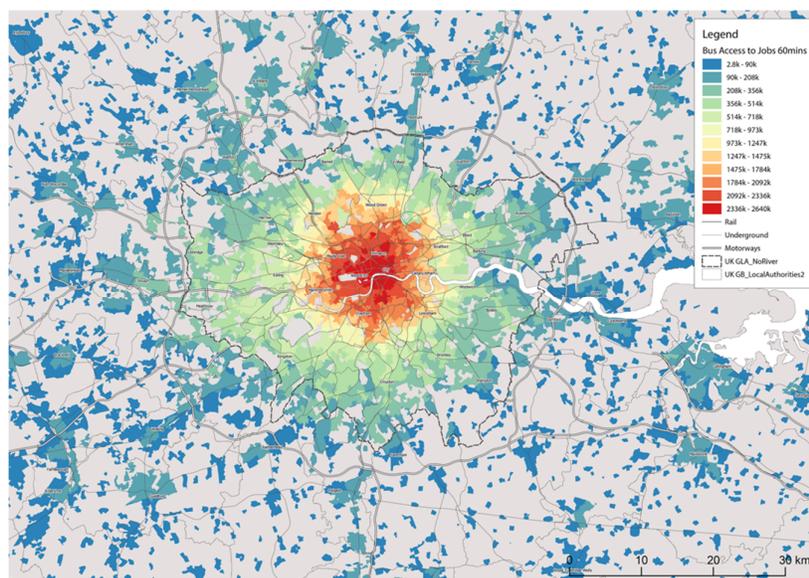


Figure 9: Total Jobs Accessible in 60 minutes by Bus

As well as mapping the cumulative indicators, we can perform some statistical analysis. The graphs below show average cumulative accessibility at a range of travel time thresholds. Essentially this summarises the competitiveness of the main travel modes at different journey time thresholds. Note the results are a population-weighted average, so they take into account the residential distribution of the population in the study area. Figure 10 is the average for the entire London Metro Region, and Figure 11 shows the average for GLA residents, excluding residents in the Outer Metro Area (OMA).

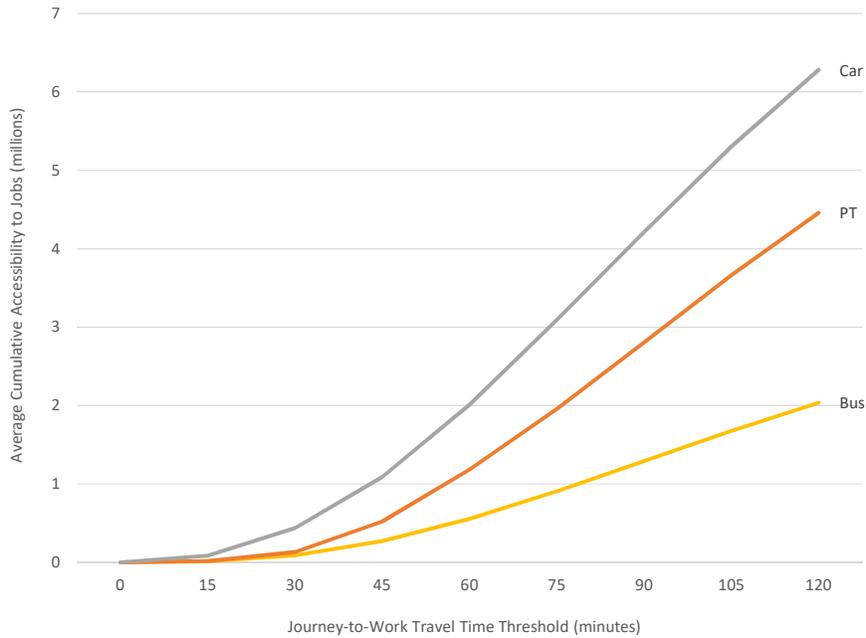


Figure 10: London Metro Region Graph of Population-Weighted Average Cumulative Accessibility to Jobs against Travel Time Threshold

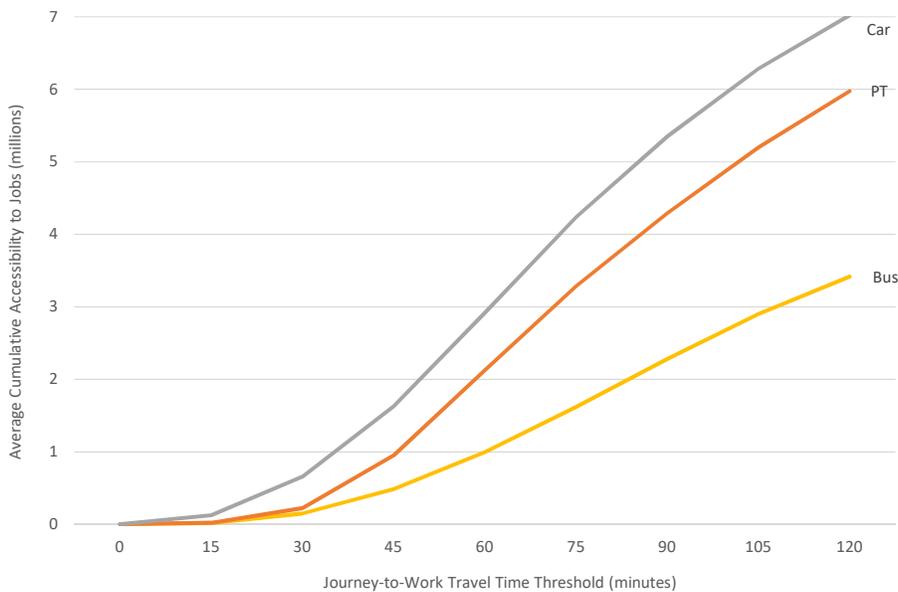


Figure 11: GLA-Only Graph of Population-Weighted Average Cumulative Accessibility to Jobs against Travel Time Threshold

There is a clear basic pattern in Figures 10 and 11 of car having the highest accessibility, followed by public transport and then the lowest accessibility by bus. This pattern is consistent across different travel time thresholds and between the Metro Region average and the GLA average. There are some important further trends within this general pattern. For the GLA only graph, the public transport modes are more competitive against the car, indeed the GLA public transport average accessibility at the 60 minutes' threshold is 2.1 million jobs, compared to the Metro Region public transport average of 1.2 million jobs. When we also consider that the car model does not include Inner London parking charges and the Congestion Charge (thus Inner London car accessibility is somewhat exaggerated), then we can appreciate the much greater competitiveness of public transport within the GLA across a range of travel times.

Another noticeable pattern is that the advantages of car travel are especially evident at lower travel times below 45 minutes. The OMA average number of jobs accessible by car in 45 minutes is over 1 million, compared to 520k by public transport and 270k by bus. These advantages translate into the significantly shorter observed travel times by car identified earlier in Tables 5-7. These results may be partly influenced by how walking costs to and from parking locations are underestimated in the model (see Section 4.2) but nevertheless there are clear accessibility advantages by car for low travel time trips that closely follow observed journey-to-work patterns.

The comparison of Figures 10 and 11 emphasise the accessibility variation that exists between sub-regions in the study area. This sub-regional variation is explored in detail in Figure 12, which includes the full sub-regional breakdown of the accessibility averages by mode.

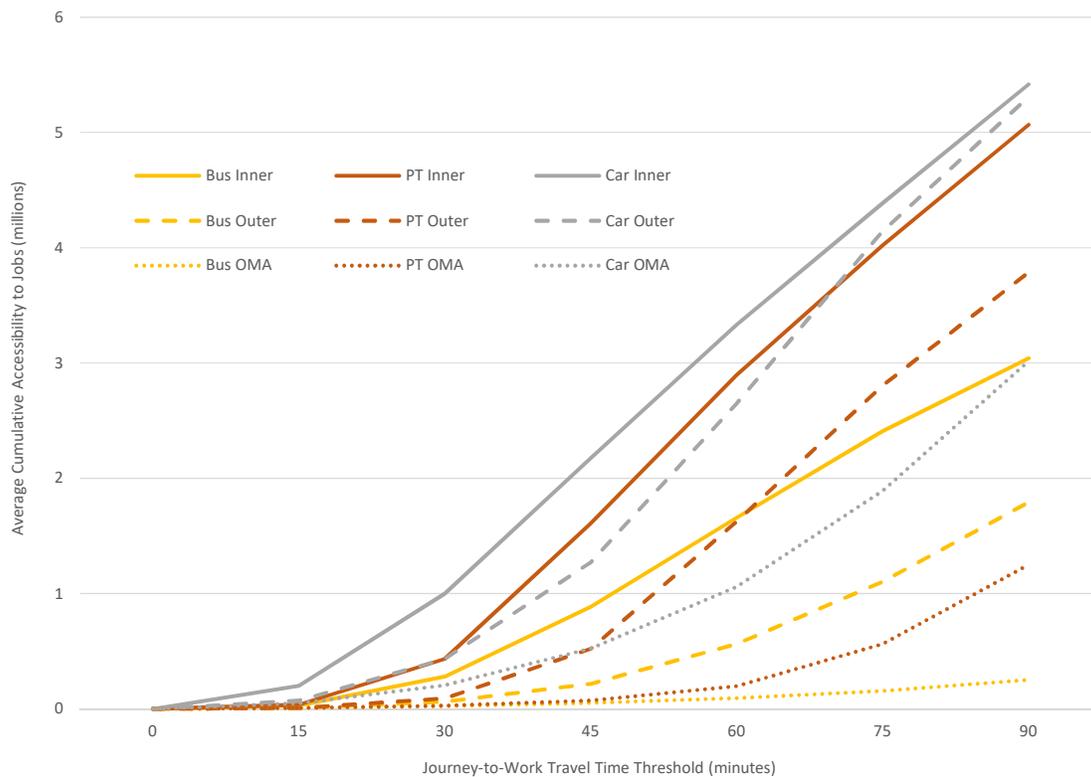


Figure 12: Graph of Population-Weighted Average Cumulative Accessibility to Jobs against Travel Time Threshold for Inner GLA, Outer GLA and Outer Metro Area Sub-Regions

The patterns in Figure 12 are more complicated to summarise, but generally it can be seen that the car mode has a narrower range of differences between sub-regions, particularly between the Inner and Outer GLA sub-regions which track each other closely for trips of 45 minutes and above. The PT mode has a wider range of average accessibilities between sub-regions, with the Inner GLA and Outer GLA competitive against average car accessibilities, but the PT OMA average accessibility is low, and only starts to increase for trip times over 60 minutes. The Inner GLA Bus average accessibility is relatively good, but the average drops off considerably in the Outer GLA and bus accessibility is very poor in the OMA for all travel times in Figure 12.

An important pattern that can be seen in Figure 12 is that there is generally a travel time threshold above which the accessibility line gradient steepens significantly. This threshold occurs when most residents can reach the huge cluster of approximately 1.5 million jobs in Central London. For public transport we can see that the steepness of the line increases above 30 minutes in Inner London, above 45 minutes in Outer London and above 75 minutes in the OMA, matching typical trip times from these sub-regions to Central London. For the car mode these threshold points occur at lower trip times, and for the bus mode at longer trip times. In the case of the OMA Bus average, this threshold point does not occur until above 120 minutes (not shown in Figure 12), and even then it is much weaker than the other modes, reflecting that for many residents in the OMA bus travel to Central London is not possible within realistic commute times.

As public transport trips in the OMA are typically longer than 60 minutes, we can repeat the cumulative accessibility analysis for the public transport model using a 90 minute threshold as shown in Figure 13 below.

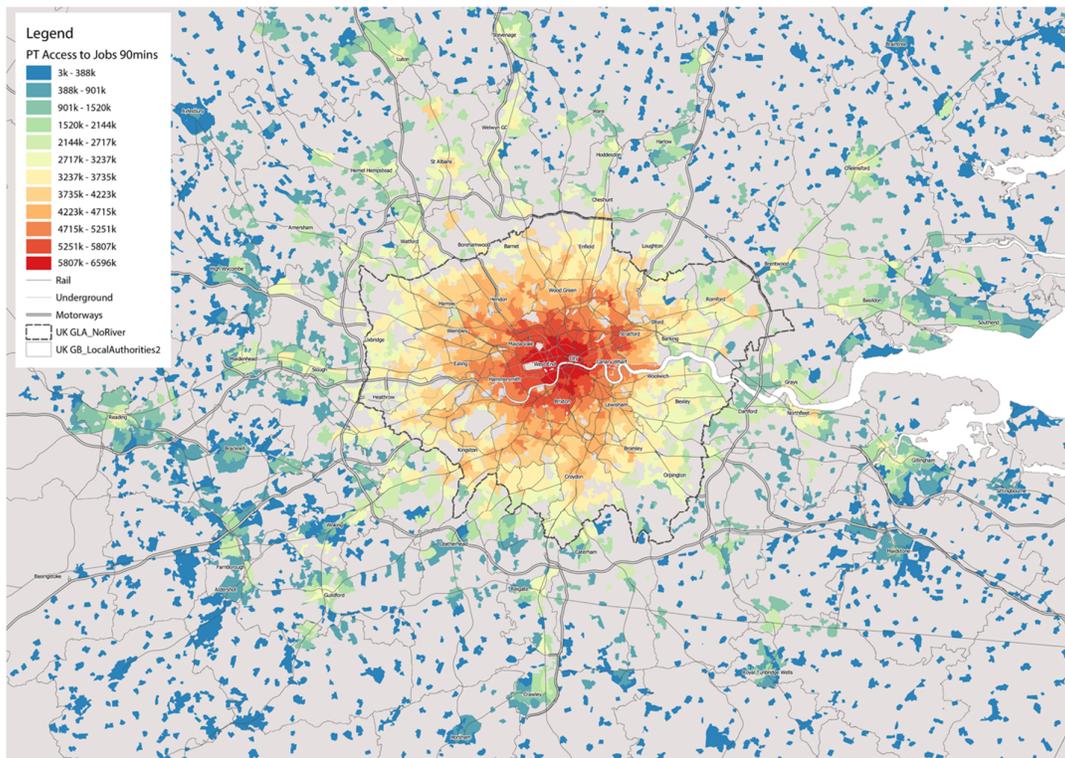


Figure 13: Total Jobs Accessible in 90 minutes by Public Transport

Figure 13 clearly shows the increased accessibility in the rail connected commuter belt surrounding London when using a 90 minute threshold compared to the 60 minute threshold shown earlier in

Figure 7. OMA town centres with good rail connections to London are able to reach over 2 million jobs in 90 minutes. These OMA centres include towns like Luton, Reading, Southend, Gillingham and Crawley. This contrasts sharply with the accessibility of below 1 million jobs for these town centres in Figure 7 using the 60 minute threshold.

Overall we can use this analysis to consider how the choice of travel time threshold influences cumulative accessibility results in the study region. Lower time thresholds will generally emphasise accessibility advantages by car; and higher travel time thresholds will increase the competitiveness of public transport travel as the opportunities available through the rail and metro networks come more into play. The 60 minute threshold appears to be a reasonably representative choice, but there is a problem with cutting-off public transport accessibility in the OMA at 60 minutes which cuts off many rail trips to Central London. There is a good case for adding additional analysis at 80 or 90 minutes to reflect accessibility by longer distance rail travel.

6. Conclusions

This paper has detailed a methodology for modelling public transport and car travel costs in a large city-region to be used in accessibility analysis of journey-to-work travel. Public transport travel costs were modelled using the OpenTripPlanner software and transit schedule data from Travelline. OpenTripPlanner was able to handle the scale and complexity of the London Metro Region and produce realistic multi-modal routing behaviour and accurate travel times. Car travel was modelled using average speed data from Highways England and the Department for Transport that was then combined and analysed in GIS software. The car model also produced realistic routing behaviour, and reasonably accurate travel times with some limitations such as not capturing parking behaviour. Overall, improvements in open data and open source journey planner software are making it easier to develop accurate travel cost models of urban networks.

The travel time matrices from the models can then be used in accessibility analysis. Cumulative accessibility results have been provided illustrating the distinct geographies of accessibility for the main travel modes in the London Metro Region. In general, public transport accessibility decreases rapidly in Outer London and the OMA (as reflected in journey-to-work modal statistics) as expected, but there is considerable variation along corridors and between town centres that is picked up with this analysis. Car accessibility also has considerable spatial variation related to motorway infrastructure and between North and South London.

There are some important limitations with the approach described. Public transport fares and car money costs have not been modelled. This is problematic for car travel where spatial variation in parking costs are not captured, and park-and-ride journeys have been omitted. For public transport, some consideration of fares is provided by modelling bus travel separately. Another important issue is temporal variation in accessibility, which has only been considered to a limited extent here by averaging queries at 15 minute intervals. Finally, this analysis has been based on modelling accessibility for 2011, rather than considering a future date for exploring new infrastructure. Modelling future accessibility is possible for public transport using timetable data of future networks (without capacity considerations). For car accessibility, average travel times of past conditions will become increasingly unreliable where major infrastructure changes occur, and a more sophisticated model would be required.

7. Acknowledgements

This research was funded by the Economic and Social Research Council RESOLUTION grant, reference ES/N011449/1. CASA researcher Richard Milton converted the public transport timetable data for the London region.

8. References

- Antrim, A., & Barbeau, S. J. (2013). The many uses of GTFS data—opening the door to transit and multimodal applications. *Location-Aware Information Systems Laboratory at the University of South Florida*. Retrieved from <http://www.locationaware.usf.edu/wp-content/uploads/2010/02/The-Many-Uses-of-GTFS-Data-%E2%80%93-ITS-America-submission-abbreviated.pdf>
- Bast, H. (2009). Car or public transport—two worlds. In *Efficient Algorithms* (pp. 355–367). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-03456-5_24
- Benenson, I., Martens, K., Rofé, Y., & Kwartler, A. (2011). Public transport versus private car GIS-based estimation of accessibility applied to the Tel Aviv metropolitan area. *The Annals of Regional Science*, 47(3), 499–515.
- Delling, D., Sanders, P., Schultes, D., & Wagner, D. (2009). Engineering route planning algorithms. In *Algorithmics of large and complex networks* (pp. 117–139). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-02094-0_7
- El-Geneidy, A., Levinson, D., Diab, E., Boisjoly, G., Verbich, D., & Loong, C. (2016). The cost of equity: Assessing transit accessibility and social disparity using total travel cost. *Transportation Research Part A: Policy and Practice*, 91, 302–316.
- Ford, A. C., Barr, S. L., Dawson, R. J., & James, P. (2015). Transport Accessibility Analysis Using GIS: Assessing Sustainable Transport in London. *ISPRS International Journal of Geo-Information*, 4(1), 124–149. <https://doi.org/10.3390/ijgi4010124>
- Haklay, M., & Weber, P. (2008). OpenStreetMap: User-Generated Street Maps. *IEEE Pervasive Computing*, 7(4), 12–18. <https://doi.org/10.1109/MPRV.2008.80>
- Hall, P. G., & Pain, K. (2006). *The Polycentric Metropolis: Learning from Mega-city Regions in Europe*. Routledge.
- Lovelace, R., Goodman, A., Aldred, R., Berkoff, N., Abbas, A., & Woodcock, J. (2017). The Propensity to Cycle Tool: An open source online system for sustainable transport planning. *Journal of Transport and Land Use*, 10(1), 505–528.
- Mayor of London. (2016). *The London Plan: Spatial Development Strategy for London: Consolidated with Alterations Since 2011*. Greater London Authority.
- Merlin, L. A. (2014). Measuring Community Completeness: Jobs—Housing Balance, Accessibility, and Convenient Local Access to Nonwork Destinations. *Environment and Planning B: Planning and Design*, 41(4), 736–756. <https://doi.org/10.1068/b120010p>
- Office for National Statistics. (2016). 2011 Census aggregate data. Retrieved from <https://www.nomisweb.co.uk/census/2011>
- Office for National Statistics. (2017). Small area population estimates in England and Wales: mid--2016. Retrieved from <https://www.ons.gov.uk/releases/smallareapopulationestimatesinenglandandwalesmid2016>

Office for National Statistics. (2018). *Labour Force Survey – user guidance - Office for National Statistics*. Retrieved from <https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/methodologies/labourforcesurveyuserguidance>

Owen, A., Murphy, B., & Levinson, D. (2018). *Access Across America: Transit 2017 Methodology*.

Reades, J., & Smith, D. A. (2014). Mapping the 'Space of Flows': The Geography of Global Business Telecommunications and Employment Specialization in the London Mega-City-Region. *Regional Studies*, 48(1), 105–126. <https://doi.org/10.1080/00343404.2013.856515>

Smith, D. A., & Serras, J. (2012). *Estimating intra-metropolitan journey-to-work CO2 emissions: a multi-modal network approach applied to the London Region 2001*. CASA Working Paper, University College London.

Transport for London. (2016). *Travel in London Report 9*. Retrieved from <http://content.tfl.gov.uk/travel-in-london-report-9.pdf>

Vale, D. S., Saraiva, M., & Pereira, M. (2016). Active accessibility: A review of operational measures of walking and cycling accessibility. *Journal of Transport and Land Use*, 9(1), 209–235.

Van Wee, B., & Geurs, K. (2011). Discussing equity and social exclusion in accessibility evaluations. *EJTIR*, 11(4), 350–367.

9. Appendix- OpenTripPlanner Origin Destination Travel Time Matrix Scripting Example

#OpenTripPlanner Jython Script Based on Example Provided by Rafael Pereira- <https://github.com/rafapereirabr/otp-travel-time-matrix>

```
from org.opentripplanner.scripting.api import OtpsEntryPoint
```

```
# Instantiate an OtpsEntryPoint
```

```
otp = OtpsEntryPoint.fromArgs(['--graphs', '.',  
                              '--router', 'lonresregion']) # Set router for analysis
```

```
# Start timing the code
```

```
import time  
start_time = time.time()
```

```
# Get the default router
```

```
router = otp.getRouter('lonresregion')
```

```
# Create a default request for a given time
```

```
req = otp.createRequest()  
req.setArriveBy(True)          # Set query to arrive by query  
req.setDateTime(2016, 9, 14, 9, 00, 00) # Set arrival time 9am 14 September 2016  
req.setMaxTimeSec(18000)      # set a limit to maximum travel time (seconds)  
req.setModes('WALK,TRANSIT')  # define transport modes. For bus only change TRANSIT to BUS  
req.setMaxWalkDistance(10000)
```

```
# Read Points of Origins/Destinations - The file points.csv contains the columns GEOID, X and Y.
```

```
points = otp.loadCSVPopulation('GridPoints.csv', 'Y', 'X')  
dests = otp.loadCSVPopulation('GridPoints.csv', 'Y', 'X')
```

```
# Create a CSV output
```

```
matrixCsv = otp.createCSVOutput()  
matrixCsv.setHeader(['Destination', 'Origin', 'Walk_distance', 'PT_Time2_ArriveBy'])
```

```
# Start Loop
```

```
for origin in points:  
    print "Processing origin: ", origin  
    req.setDestination(origin)  
    spt = router.plan(req)  
    if spt is None: continue
```

```
# Evaluate the SPT for all points
```

```
result = spt.eval(dests)
```

```
# Add a new row of result in the CSV output
```

```
for r in result:  
    matrixCsv.addRow([ origin.getStringData('GEOID'), r.getIndividual().getStringData('GEOID'), r.getWalkDistance(), r.getTime()])
```

```
# Save the result
```

```
matrixCsv.save('LondonResolutionRegionMSOA_PubTransTimes_arriveby_9am.csv')
```

```
# Stop timing the code
```

```
print("Elapsed time was %g seconds" % (time.time() - start_time))
```