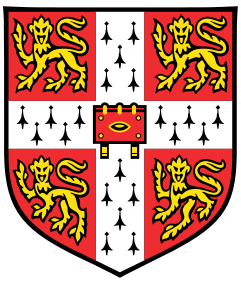
Investigating the performance of transport infrastructure using real-time data and a scalable multi-modal agent based model



Gerard Hugh Casey

Supervisor: **Peter Guthrie**

Advisors: **K. Soga & E. Silva**

Department of Engineering
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Darwin College

February 2019

Dedicated to:

John Francis (Sean) McMahon & Deirdre McMahon

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 65,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures.

Gerard Hugh Casey
February 2019

Acknowledgements

I would like to thank both Kenichi Soga and Peter Guthrie for their unwavering support and patience. I was very fortunate to be allowed (and even encouraged) to work across so many fields and disciplines under their tutelage. This thesis tells the story of a somewhat naive and ambitious initial problem set, that grew far beyond its original remit and resulted in a piece of work far beyond the original intentions of trying to understand the carbon credentials of HS1. This research has likely opened up more questions than it has attempted to answer and although I feel we have made notable contributions, the overarching experience has left me humbled by the continuing realisation that we understand only small parts of the systems we are attempting to model. I am optimistic that the modelling and data driven techniques advocated in this work will eventually find their way to practical use, but am now acutely aware that there is no technical silver bullet to many of the environmental problems we are attempting to tackle. I would like to thank both of them for their technical and academic rigour, but most especially for their guidance, which both encouraged and questioned the wide scope of my work.

Sincere thanks go to Elisabete Silva for introducing me to agent based modelling, the only modelling career I will ever have. Elisabete had the patience to invest time in me when I was a lost civil engineer attempting to contribute to the dynamic modelling community and I am very thankful for this.

Krishna Kumar and Bingyu Zhao were a part of the most formative periods of this thesis and for this I'm extremely grateful. I hope there will be many more adventures ahead, in Cambridge, Berkeley or wherever else we may find ourselves!

To all of my colleagues in the Centre for Sustainable Development - thank you for your words of encouragement and constant stimulation over many long years in the tree house. A most sincere thanks must also go to Siân Owen, a true star who helped me and countless others so selflessly.

A very special thanks goes to Ove Arup & Partners for both the financial and intellectual contributions to this research. Heleni Pantelidou and Tim Chapman not only made this PhD possible, but supported and encouraged my transition from a mediocre geotechnical engineer who might just be able to get chartered, to a geotechnical engineer who definitely can't get chartered.

Many thanks go to the following groups and companies for access to computing infrastructure and data; The Cambridge High Performance Computing Service for use of their Darwin Cluster, Microsoft Azure for use of the HDInsight service via Microsoft Azure for Research, and finally to Amazon for use of various resources on Amazon Web Services. Many thanks also go to Google for providing a Maps for Business account which enabled use of their Directions API. Special thanks go to John Wilkes for making this possible. Thanks also go to Richard Brown for access to Eurostar reports on their carbon performance.

I must also thank my family; my mother Deirdre, father Hugh, brothers Peter and Mark, and soon-to-be wife Aoife Magee for the constant faith and support. As I make the transition away from being a career student I am now more than ever aware of the hard work by my parents that has permitted me to live such a privileged life. I am eternally grateful for their hard work.

Finally, I dedicate this thesis to my grandparents Deirdre and Sean McMahon.

Abstract

The idea that including more information in more dynamic and iterative ways is central to the promise of the *big data* paradigm. The hope is that via new data sources, such as remote sensors and mobile phones, the reliance on heavily simplified generalised functions for model inputs will be erased. This trade between idealised and actual empirical data will be matched with dynamic models which consider complexity at a fundamental level, inherently mirroring the systems they are attempting to replicate. Cloud computing brings the possibility of doing all of this, in less time than the simplified macro models of the past, thus enabling better answers and at the time of critical decision making junctures.

This research was task driven - the question of high speed rail versus aviation led to an investigation into the simplifications and assumptions that back up many of the commonly held beliefs on the sustainability of different modes of transport. The literature ultimately highlighted the need for context specific information; actual load factors, actual journey times considering traffic/engineering works and so on.

Thus, rather than being explicitly an exercise in answering a specific question, a specific question was used to drive the development of a tool which may hold promise for answering a range of transportation related questions. The original contributions of this work are, firstly the use of real-time data sources to quantify temporally and spatially dynamic network performance metrics (eg. journey times on different transport models) and secondly to organise these data sources in a framework which can handle the volume and type of the data and organise the data in a way so that it is useful for the dynamic agent based modelling of future scenarios.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

Rapid urbanisation in developed and developing countries has continued globally. Cities are increasingly becoming the arena within which the greatest opportunities and challenges to true sustainable development can be found (C40, 2007). The role of city emissions in global emissions is uncertain (from 30% to 80% depending on methodology (Satterthwaite, 2008)) but undoubtedly significant. The challenge of emissions allocation draws parallels to similar methodological challenges when allocating the economic importance of cities at the global and domestic scale. This economic dynamism of uncertain proportions is the key component in the adoption of innovative technical solutions, in combination with behavioural change of those who inhabit these cities (Stern, 2007). This is ever more apt as cities are also exhibiting some of the most extreme and chronic symptoms related to unsustainable systems. These challenges range from micro emissions, for example public health impact of particulate matter (Walton et al., 2015a), to macro emissions, for example climate change related sudden weather events (Rosenzweig et al., 2001), to housing shortages (Carpenter and Lees, 1995) and to unemployment (Stern, 2007). This double edged situation of exposure and opportunity is most evident in transportation. Cities exhibit some of the highest particulate matter emissions, in geographic combination with the some of the highest population exposures. Yet, such particulate matter emissions are primarily as a result of private vehicular emissions in areas where the largest savings can be made in a modal shift to a mass transit system. The modal shift away from individual vehicular transportation to a range of mass transit systems has long been identified as a key method for reducing global and local emissions (Potter, 2003). The growing political autonomy, agility, resources and scale of cities has arguably led to a shift in influence away from the state level to the city level (Scott, 2002). As the trickle down has now become the trickle up, the political landscape has begun to embrace the innovative methods put forward and tested by cities. The magnitude of city related emissions and their local impact

provides a common goal. The reduction of emissions is of significant importance to the cities in isolation and all globally.

1.1 Research Motivation

The link between anthropogenic carbon emissions and climate change has been well documented in scientific literature. Despite such overwhelming evidence, policy makers have been slow in getting consensus on how this challenge is to be tackled globally (Stern, 2007). However, many countries have adopted domestic policies aimed at reducing their own domestic carbon footprints. The 2008 Climate Change Act ties the UK into at least an 80% reduction, from the 1990 baseline, of its net carbon account by 2050 (DECC, 2008). The aim of the Act is to enable the UK to become a low carbon economy and thus avoid dangerous climate change. As this target approaches, more efforts are being made to investigate how carbon can be considered in decisions where it previously did not feature, in an attempt to make more carbon sensitive strategic decisions. The challenges faced combined with the complex environments poses a challenge for decision making tools.

This research joins a growing body of work which aims to make use of the new empirical data sources which remove the need to fall back on generalised model inputs and new computational models which allow for more information to be considered in more efficient ways. The aims here are two fold - the first is to remove the reliance on standardised functions which were historically required to handle uncertainty. Secondly, to make use of these data sources in dynamic ways, paying respect to the complexity of the systems we are trying to model.

This research was motivated by a desire for a better understanding of the caveats that come with transport planning. Mass transit systems have long been advocated as a means of reducing emissions. Generally, this is true, but it may often be disingenuous when making unfair comparisons at the stage of appraisal, usually due to the inaccuracy of assumptions surround load factors and average speeds. It is clear that the interface between human decision making and the physical and non-physical infrastructure leads to highly context specific behaviour, which to date, is either completely unknown or unsatisfactorily averaged out in an attempt to compare like with like.

Figure 1.1 illustrates the wider challenge associated with making decisions in complex environments. The decision making paradox lies in the time and effort required to do analyses. There is a need to have a sufficiently narrow scope to drive modelling, however, the very narrowing down of this scope reduces the impact any modelling can have, in effect resulting in the modelling being used as a tool to simply optimise the decisions already made using "professional/expert opinions" and "gut instinct." Current understanding of complexity and

human cognitive ability paints a dire picture for the ability of individuals to make sufficiently robust decisions at these key junctures (Saaty, 1990).

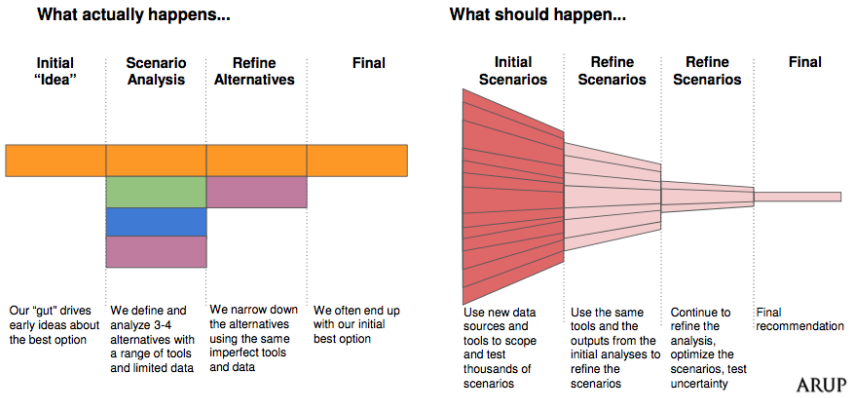


Fig. 1.1 Scenario analysis - what happens versus what should happen (Dedring et al., 2017)

The value of fine resolution dynamic models has long been theorised, and recent computational advances have enabled for increasingly complex, bottom-up, fine resolution simulations to be carried out over long time horizons at fine spatial and temporal resolution. This has hinted at the possibility of connecting scales of what has been historically been fine resolution operational models and coarse resolution strategic models. The wave of new geospatially connected devices has enabled the harvesting of fine resolution spatial and temporal data on travellers and even the infrastructure itself. This crowd-sourced data can be used to inform dynamic models with real-world and real-time data, bypassing the need for generalised functions and/or expensive survey data. In this thesis, dynamic data sources such as the Google Directions API data and Transport for London's real-time data feeds are presented in a framework for London.

The use of a multi-modal graph data structure is presented and a series of analyses are presented to illustrate the analytical value in using new parallel computing advances from computer science. These methods enable the scaling of fine resolution scenario testing transportation models and enable the support for a range of agent decision making methodologies. Such data structures offer performance improvements in the storing of dynamic data that may be manipulated in order to simulate local and global hard infrastructure scenarios alone or in tandem with traditional policy or dynamic policy making scenarios.

Ultimately, this thesis presents a framework for using real-time data sources to inform a transport Agent Based Model (ABM) for multi-modal decision making and route assignment.

1.2 Thesis layout

- Chapter 2 - The role of transport in mitigating climate change, the state of the art in transport modelling, emerging methodological trends and the context of new computational methods.
- Chapter 3 - The use of new data sources to quantify network performance. specifically for the road and public transport networks.
- Chapter 4 - A scalable multi-modal framework for agent based simulations on the transport network.
- Chapter 5 - The case study of High Speed 1, the influence of capital emissions and hub travel emissions.
- Chapter 6 - Conclusions and future work.

Chapter 2

Literature Review

2.1 Climate Change & Transport

The link between anthropogenic green house gas emissions and climate change has been well documented in scientific literature. The Intergovernmental Panel on Climate Change (IPCC) report that global emissions have risen consistently from 1970 to 2010 (citepedenhofer2014working, as shown in Figure 2.1. The transport sector globally produced 7.0 $GtCO_2eq$ in 2010, equivalent to 23% of total CO_2 emissions, an increase from the year before despite technological advances in vehicle efficiency and policy adoption (Sims et al., 2014). These emission trends have resulted in profound changes to planetary systems leading to significant risk in the form of dangerous climate change. The risk to dangerous climate change is not allocated equally, nor is the this risk allocated with respect to responsibility for it. The historic paradigm of coupling fossil fuel driven energy consumption and economic growth (Holtz-Eakin and Selden, 1995) still holds political ramifications for those nations who have not enjoyed the liberty to pollute uninterrupted as many economically developed nations have. The difficulties in accounting for risk and responsibility across, and even within, national borders has resulted in a fragmented policy environment (Stern, 2007) (with some exceptions, such as the EU Emissions Trading Scheme (EU, 2015)). This is epitomised in the debate over territorial or consumptive accounting methods. As an illustration, the UK's consumptive emissions are 35% larger than the territorial emissions for the year 2010 (GCB, 2013), equating to around 346 $MtCO_2e$ of a difference. Consistently and meaningfully drawing the boundaries of any emissions analysis is problematic (Dodman, 2009).

Climate change is often described as a tragedy of the commons situation, a phrase coined by Garett Hardin in 1968 (Hardin, 1968). He concluded that this type of problem was categorised by individuals acting in rational self interest, whilst collectively these actions lead to a common resource becoming over exploited. Each individual experiences short term gains from their

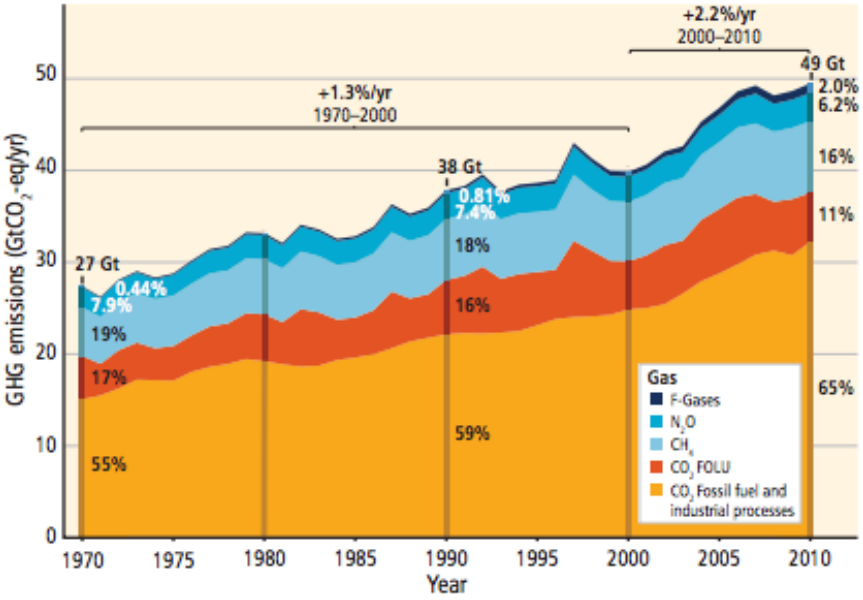


Fig. 2.1 IPCC GHG Emissions (Edenhofer et al., 2014)

use of the resource, whilst the long term impacts, due to their collective actions, will have significant time delay. The end result likely being a situation where all individuals suffer.

It is clear that such a situation requires political leadership in order to break the cycle. Such political decisions do not occur in a vacuum and must be made with informed, transparent, robust and considered information. Technological innovations must be partnered with behavioural change, and promises of technical *silver bullets* (Brooks, 1987) should be treated with great caution.

Reducing transportation emissions is particularly problematic for a range of reasons. It is of fundamental importance to the current economic paradigm, connecting markets, goods and people (Krugman, 2009). The demand for transport has been increasing, most especially in developing nations and is even tied in many locations to unsustainable individual vehicle ownership in high density areas (Sims et al., 2014). It does not lend itself easily to renewable energy source replacement, as the (mostly centralised) energy sector does. The mobile nature of many modes, such as aviation and motor vehicles, requires battery performance currently not in existence. Great leaps have been made in recent years as a result of policy tools (DfT, 2015b) and heavy investment by the likes of Tesla (*tesla.com*) and McLaren (*mclaren.com*). Despite this, the energy density to weight issue persists, limiting the desired performance and range offered in comparison to internal combustion engines. Mass public transport systems have long been heralded as sustainable, low emissions solutions (Newman and Kenworthy, 1999), (Norman et al., 2006). Yet, the construction of such projects requires large capital investments, in terms of CO_2, political capital and financial capital. Such investments must be made with

strategic vision or risk the operational value (in terms of CO_2 or finance) not recouping the initial capital investment.

This chapter will explore the backdrop to the UK's national emissions footprint and the role the transportation sector has in this. It will move on to the need to consider the life cycle when appraising the role of different transportation options. The exclusion of international travel from the UK Climate Change Act framework leaves a policy gap that has the potential to undo the hard earned projected savings in other sectors. The relative proximity, density and maturity of the European High Speed Rail (HSR) Network offers opportunities for modal shift from short-haul aviation to HSR. In an attempt to assess the emissions sustainability of HSR a range of caveats that come with marketed sustainable 'low emissions' transport modes are explored and the sensitivity to factors such as capital carbon investment, energy supply and travel demand are discussed. The importance of these factors is illustrated, existing methodological shortcomings are discussed, and the requirements for a new tool are put forward.

2.1.1 UK Context

UK total (domestic) emissions were at 514.4 $MtCO_2e$ in 2014, with the energy sector at 31% and the transport sector at 23% of this total (DECC, 2014). As with other developed economies, the dominating energy sector is due to a dependence on CO_2 based means of energy generation, via the combustion of fossil fuels. A shift away from coal electricity generation and a warmer winter resulted in the 7% saving made for energy emissions from years 2013 to 2014 (DECC, 2014).

The 2008 Climate Change Act ties the UK into at least an 80% reduction, from the 1990 baseline, of its net territorial CO_2 emissions account by 2050 (DECC, 2008). The aim of the Act is to enable the UK to become a low carbon economy and thus avoid dangerous climate change. As this target approaches, more efforts are being made to investigate how CO_2 can be considered in decisions where it previously did not feature, in an attempt to make more CO_2 sensitive strategic decisions.

Since the 2008 Climate Change Act, policy has progressed to first identify and secondly, to plan how adjustments can be made to reach the overarching 80% reduction. The UK's electricity supply has been the focus of the majority of de-carbonisation spotlight, spearheaded by the Department of Energy and Climate Change (DECC)[1] .

Beyond the overarching Climate Change Act, the strategic CO_2 potentials of the infrastructure sectors to meet the ambitious, legally binding emissions targets has been the focus of the *Infrastructure Carbon Review* (GCB, 2013). The Infrastructure Carbon Review recognised

[1] As of July 2016, DECC is now a part of the Department for Business, Energy & Industrial Strategy

the need to move from a territorial methodology (as per the Climate Change Act 2008) to a consumption-based methodology and attempted to reconcile the two by considering international aviation and shipping on the basis of departing journeys (GCB, 2013). Despite the Committee on Climate Change advising the UK Government to include international aviation and shipping within the carbon budgets in 2012 (CCC, 2012), their inclusion has not yet occurred. The off-shoring of such emissions is equally problematic in areas such as manufacturing as recent trends for most developed nations has involved the importing of goods rather than the carbon intensive manufacturing of such goods domestically. The complexity of globalisation and the modern markets, with the flow of capital, goods and people has resulted in a highly connected and dependent relationships. It is thus understandable that there must be methodological flexibility in policy making when considering emissions (Dodman, 2009). Despite these recognised political and institutional barriers there is a strong need for care in order to avoid distorted and goal seeking behaviour, as may occur when a very narrow or biased methodology is employed.

The breakdown of UK transport emissions in 2010 is shown in Figure 2.2. These statistics are based upon a modified consumption (departures only) based methodology. Cars are responsible for the vast majority of transportation emissions, at 52%. This is followed by international aviation (20%), road freight (11%), international shipping (6%), domestic shipping (4%), bus (4%) , all rail (2%) and domestic aviation (1%).

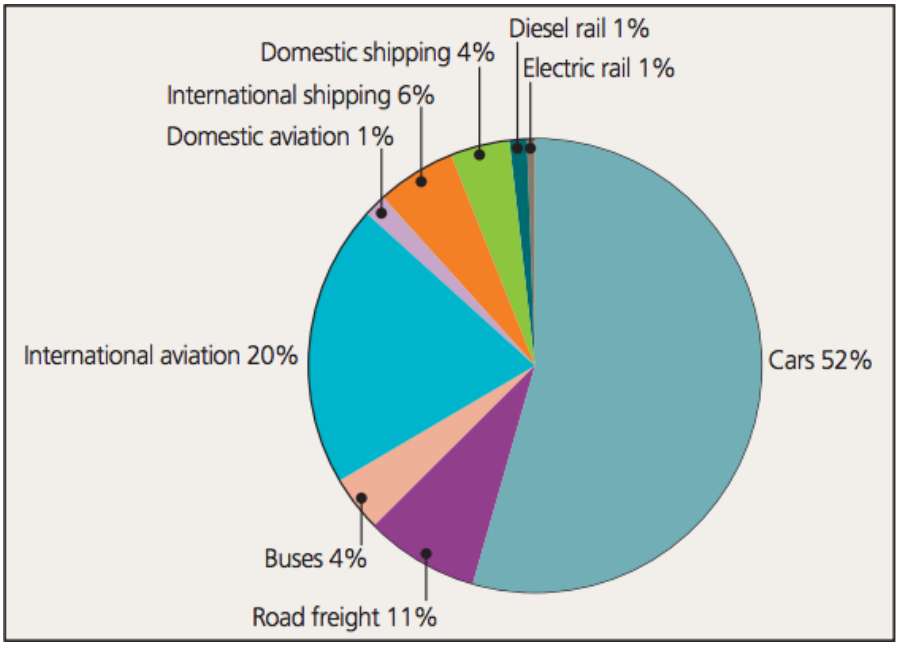


Fig. 2.2 UK transport emissions by mode (GCB, 2013)

These CO_2 emissions occur simultaneously with non-CO_2 emissions, such as that of fine particulate matter. A King's College London study found PM2.5 emissions in London were estimated to have a total mortality burden of 3,537 deaths at typical ages and the NO_2 mortality burden was estimated at 5,879 deaths at typical ages in 2010 alone (Walton et al., 2015b). The study attributed a short terms economic cost of £1.4 Billion and long term cost of up to £3.7 Billion on these impacts. The public health impacts of such emissions have added a new dimension to the drive for low emissions transport, most especially in the wake of recent scandals and issues related to valuing CO_2 emissions over particulate emissions and vice-versa (DEFRA, 2007). The uncertainty related quantification issue of substantiating climate change risks may be alleviated by using these shorter, more tangible public health impacts in combination with longer term CO_2 related metrics.

2.2 Life Cycle Analysis

Life Cycle Analysis (LCA) is a method used to quantify the associated environmental impacts of the different stages through a life cycle. The motivation is to capture a truer reflection of the range of impacts any given project may have over its existence. The importance of LCA methods has grown in recent times (Chester et al., 2012), (Heijungs et al., 2013), (Heinonen et al., 2013), as the literature has identified how narrow comparisons, such as that for only the tail pipe emissions, can lead to distorted outcomes (Chester and Horvath, 2009a). In the US, a study illustrated that a life cycle methodology for GHG emissions and energy increased metrics by 63% for road vehicles, 155% for rail and 33% for aviation compared to a standard tail pipe methodology (Chester and Horvath, 2009a). Chester also carried out similar calculations for rail and found that considering the full life cycle increased emissions by 1.8 to 2.5 times over a tail pipe only analysis (Chester and Horvath, 2009b). Figure 2.3 presents an addition from Miles et al. (2019) to Chester's comparative study of 3 motor vehicles, 1 bus (on and off peak), 3 rail and 3 aircraft. A European passenger car (VW Up!) and a new battery electric car (Kia e-Niro) are added for both single and 5 person occupancy [2] to update Chester's 2009 paper. With Chester's original data the most evident conclusion is that personal motor vehicles (Sedan, SUV and Pickup) have extremely large emission footprints (dominated by operational emissions) when compared to the other modes, reinforcing the support for mass transit systems over individual ones (Newman and Kenworthy, 1999). Secondly, the importance of load factor is clear in the comparison between the urban bus on and off peak.

[2] The kia e-Niro claims a 149Wh/km energy demand (Kia, 2019) and operational emissions were assumed on UK national average grid emissions for 2016 (DECC, 2015). The VW Up! claims 95gCo_2/km (VW, 2019). Both were assumed to share the same insurance, maintenance and manufacturing costs as those in the original Chester study.

This makes a clear and important caveat to the value of mass transit systems in terms of CO_2 emissions. If the load factor is not sufficient, the extra capacity on the vehicle may result in worse performance than a personal motor vehicle. The addition of the European passenger car illustrates the improvements possible in the personal car domain, with single occupancy emissions at 185gCO_2e / PKT and 37gCO_2e / PKT when at full occupancy. This paints a more optimistic picture and illustrates that his logic works both ways, ensuring/ indicating that a personal car may not necessarily be an unsustainable, high emissions vehicle if it had high load factor, is driven sensibly and has modern efficiency features. The data for modern electric vehicles also shows clearly the opportunity to reduce emissions as a result of improved battery technology and the de-carbonisation of the wider energy grid. This paired with a high load factor reduces the emissions to below that of the shown rail lines and that of a full occupancy bus. It is also shown that the use of complex pricing structures has diluted the large operational emissions associated with aviation, improving the efficiency via high load factors and longer travel distances (most especially for larger aircraft).

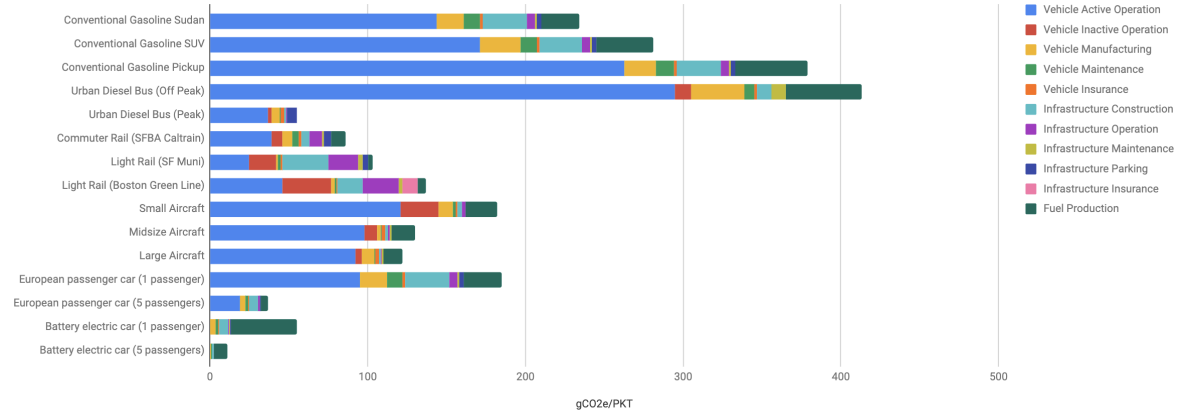


Fig. 2.3 CO_2e / PKT for a range of modes and mode types from Miles et al. (2019) and Chester and Horvath (2009b)

Such comparisons in the literature highlight the need to consider a range of factors when planning for future transport provision. For the transportation sector, the life cycle analysis requires consideration for the capital CO_2, operational CO_2 and land use impacts. In the specific case of rail infrastructure, (Saxe et al., 2015a) conceptualised this as shown in Figure 2.4. This graphic illustrates the three components of a CO_2 LCA for a rail infrastructure project: the embodied CO_2, the ridership and the urban form impacts.

The design life span of some infrastructure and their continued use in evolving ways, has lead to consensus that the final *grave* stage of LCAs should be excluded for some infrastructure

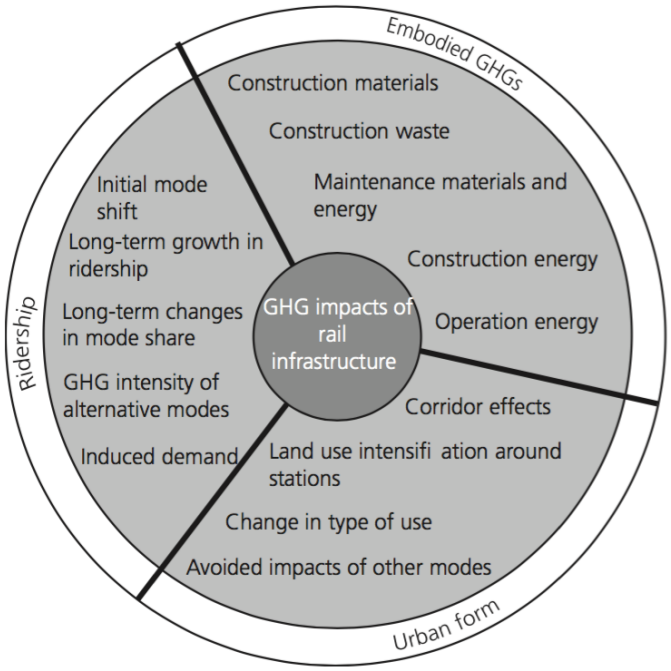


Fig. 2.4 Life cycle analyses considerations (Saxe et al., 2015a)

projects, and is generally only considered for rolling stock and the vehicles that utilise said infrastructure. This is often referred to as a *cradle-to-site* methodology.

In order to consider the wide range of possible effects, there are two categories of boundary conditions that must be defined. These are graphically illustrated in Figure 2.5 and discussed below.

1. Temporal Boundary

   The temporal boundary considers the different stages of the life cycle. Depending on the analysis, a range of boundaries may be drawn, for example this may be *cradle-to-grave*, *cradle-to-gate* or *cradle-to-cradle*. In the case of *cradle-to-grave* this would include the mining of the raw materials, their transformation into a construction material, their use in a structure, and finally their demolition and re-use.

2. Spatial Boundary

   The spatial boundary considers the geographic influence of a project. As transport infrastructure does not act in isolation, but as part of a highly complex system, drawing the spatial boundary around a system is difficult and unintuitive due to network effects.

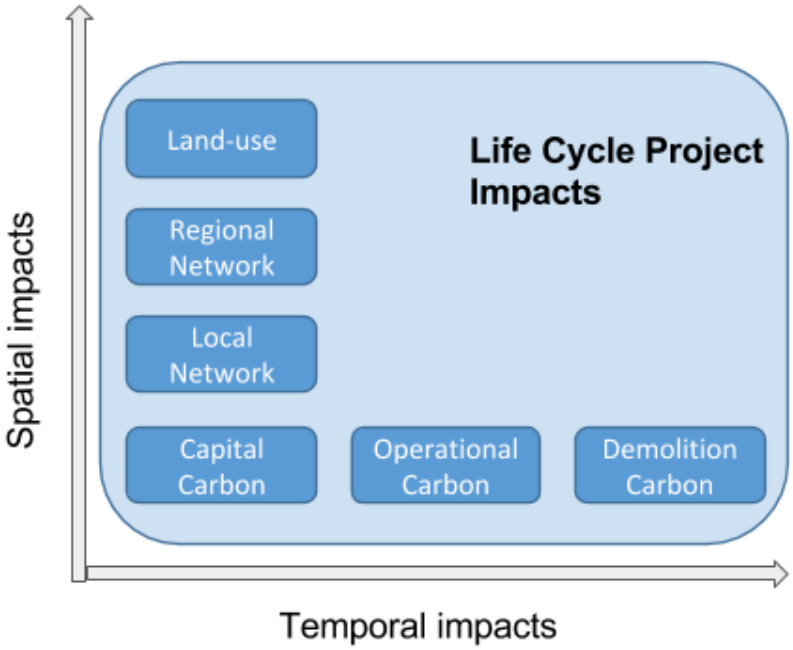


Fig. 2.5 Temporal & spatial boundary of life cycle impacts

2.2.1 Temporal Boundaries

Capital CO_2

Capital CO_2 is the up-front investment made in order to construct the physical infrastructure, for example a road, railway line; or the vehicle that utilises this infrastructure, for example the rolling stock or car. Capital CO_2 is defined in a similar way to that of capital financial investment. LCA analyses have a long history in the manufacturing sector, where standardised and well defined products lend themselves to simpler quantification using bill of quantity (BOQ) data and standardised emission coefficients. This capital carbon consists of the embodied CO_2, the CO_2 required in the creation of the materials; and the construction CO_2, the CO_2 emitted when constructing the infrastructure from these materials.

1. Embodied CO_2

   Embodied CO_2 consists of the energy utilised from raw material extraction to delivery from manufacturer. The Inventory of Carbon and Energy contains data on over 200 materials and has been updated continuously with new and revised data since the original release in 2007 (Jones and Hammond, 2008). For most products, the dominant proportion of the CO_2 footprint is found in the embodied energy invested to create the materials themselves. For example, a study in Sweden found the embodied energy of a building

accounted for 45% of total energy over a span of 50 years (Thormark, 2002). In building construction this is especially influenced by concrete and steel. Steel and aluminium are challenging as they are responsible for approximately 10% of global CO_2 emissions, their recycling rates are already at 60% and demand is projected to double by 2050 (Allwood and Cullen, 2009). For illustration of the magnitude, the embodied CO_2 emissions of a 170m viaduct equates to around 220,000 tCO_2 (Hughes, 2012), equivalent to the yearly emissions of around 90,000 cars in the UK (DfT, 2012b). By using BOQ data and information on the sourcing of materials (for example virgin versus recycled) an accurate image of embodied carbon may be quantified. Such analyses are often focussed on embodied energy and thus often require a conversion to be made from embodied energy to embodied CO_2.

2. Construction CO_2

   The heterogeneity of infrastructure projects is in stark contrast to manufacturing products, for which LCA's have historically been used to assess. Each infrastructure project is unique and exists in a complex system. Such projects do not lend themselves to easy quantification and thus the proportional importance of embodied energy to construction energy depends on the nature of the infrastructure itself, and construction site energy is often excluded (Iddon and Firth, 2013) and can account for 5-30% of total emissions (Chau et al., 2012a). A notable exception to the dominant embodied CO_2 rule within civil engineering projects is earthworks. Such projects may have little embodied emissions as the material exists independently of the project and the dominant emissions is actually be associated with the moving of the material through heavy plant or through the use of lime modification to improve the structural properties of the soil (Hughes et al., 2011a). The emissions are dependent on the fuel used by the plant (Sharrard et al., 2007) and their age (Waris et al., 2014). In these cases, an embodied centric methodology would underestimate significantly. Methodologies now exist for quantifying the emissions arising as a result of earthworks, tunnel boring machines, and other heavy plant activities on site.

For transportation projects a list of potential aspects[3], as well as sources for data in the literature, are provided in Table 2.1.

Operational CO_2

Operational CO_2 is the CO_2 which is accrued due to the use of the infrastructure. This is easily conceptualised as the tail-pipe emissions of vehicles such as trains or cars. Such emissions are

[3]This is a modified list of that used for the capital CO_2 assessment of High Speed 2.

Table 2.1 Specific examples of capital CO_2 sources for a rail infrastructure project

Aspect	Specific example	Reference
Earthworks	Embankments & cuttings	(Hughes et al., 2011a)
Construction & demolition works	Construction & demolition waste	(Yeheyis et al., 2013)
Bridges & Viaducts	Norwegian bridge	(Dequidt, 2012)
Roads	Inventory for road construction & use	(Treloar et al., 2004)
Buildings	Train station	(Chester and Horvath, 2009b)
Geotechnical structure	Retaining wall	(Chau et al., 2012a)
Tunnels, portals & drive unders	Rail tunnel	(Morita et al., 2012)
Heavy plant	Tunnel boring machine	(Chau et al., 2012a)
Track	Slab & ballasted track	(UIC, 2011)
Rolling stock	TGV Rolling stock	(UIC, 2011)

a function of the energy supply to the vehicle, the conditions (surface gradient and quality), the vehicle's properties (weight, aerodynamic performance), the vehicles load (as a ratio of demand to capacity) and external factors (such as variable speed due to traffic). There are a range of influencing factors:

1. Vehicle properties
   (a) weight
   (b) rolling resistance
   (c) aerodynamic performance
2. Vehicle energy supply
   (a) Rail: diesel
   (b) Rail: electricity (dependent on source)
   (c) Vehicle: Internal combustion (diesel, petrol)
   (d) Vehicle: electric (dependent on source)
   (e) Vehicle: hybrid (combination)
3. Infrastructure conditions
   (a) Gradient (dependent on vehicle weight)
   (b) Road/rail condition (rolling resistance)
4. External factors

   (a) Traffic (speed & efficiency)

   (b) Junctions, furniture (speed & efficiency)

5. Vehicle load

   (a) Distribution of emissions

6. Maintenance

2.2.2 Spatial Boundaries

Beyond the impacts that propagate through time, there are those impacts which propagate through space. These impacts are most clearly apparent on the immediate area but network effects mean that they can be hard to predict and quantify in a meaningful way. These can be categorised as **network effects** which influence other transport related infrastructure, such as travel times on roads or demand on bus services; or **land use effects** through changes as a result or reduced/increased connectivity.

Network Effects

The addition or removal of a piece of infrastructure or service on existing infrastructure will result in changes to the local network. Intuitively, the addition of a road to an area should reduce the demand on it and other local roads, resulting in improved journey times. However, the addition of road capacity does little to reduce congestion, as short term journey time savings induce new demand or simply move a bottleneck elsewhere (Duranton and Turner, 2011). The uncertainty and complexity of such impacts remains poorly understood, where despite the overwhelming evidence surrounding road capacity and congestion, many cities and countries continue to attempt to remedy road congestion by adding capacity.

Urban Form Effects

There is a growing understanding of the relationships between transportation provision, real estate development and urban form. Consider the comparison between Atlanta, USA and Barcelona, Spain in 1990 made by Bertaud (Bertaud and Richardson, 2004) and shown in Figure 2.6. In 1990 Bacelona and Atlanta had comparable populations, 2.8 million and 2.5 million respectively. However, Atlanta had a built-up area of 4,280km^2 and Barcelona had 16km^2. The link between transportation provision and urban form operates in both directions, often reinforcing the *status quo*. In the situation of Atlana, the result has been the dominance of personal car use as no investment resulted in poor public transport provision and due

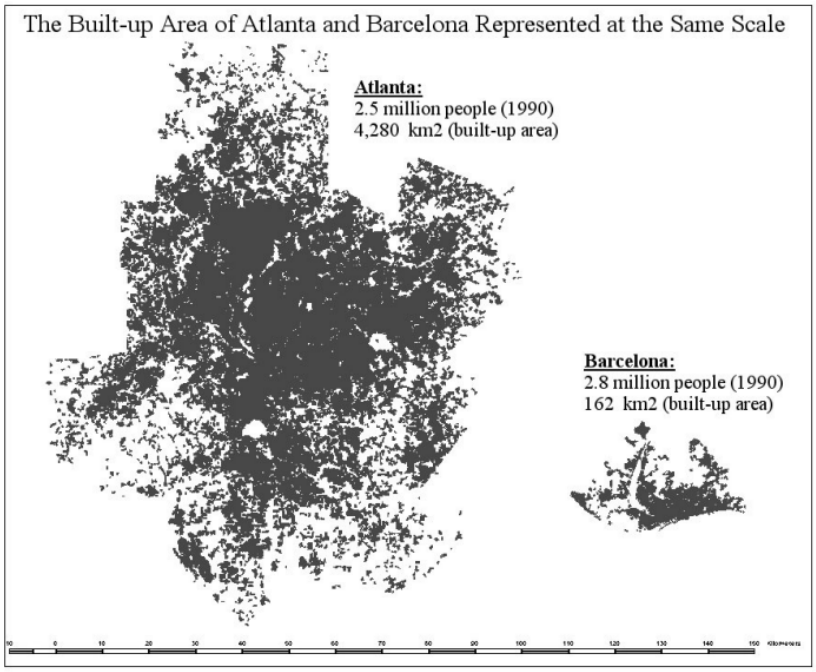


Fig. 2.6 Urban density: Atlanta versus Barcelona (Bertaud and Richardson, 2004)

to continued urban sprawl, such public transit investment increases in cost and savings are challenging in the face of a lack of density.

Densification around transport hubs has the potential to lead to meaningful carbon emission savings (Saxe, 2016) and there are now numerous examples of development policies being proposed in tandem with mass public transport projects in order to reduce carbon emissions (Kimball et al., 2013) and (Nahlik and Chester, 2014). In effect, this would be the extension of the now popular Transit-Oriented Development (Cervero, 2004) paradigm to include CO_2 emissions considerations.

2.3 Case Studies

In order to discuss the relative importance of these different factors, two classes of case studies are drawn on from the literature. The first surrounds urban travel and the second surrounds intra-urban travel.

2.3.1 Temporal Boundaries

Capital CO_2

Urban Travel

A detailed capital CO_2 study was carried out for Crossrail, a new high frequency, high capacity railway line in London. The study found that the construction would cost between 8.6 to 13.5 $MtCO_2eq$ and that this would equate to 15% of total life cycle emissions (assuming 120 years span) (de Silva and Paris, 2014). Within these capital CO_2 costs, the embodied CO_2 of the materials is dominant at 58% and construction CO_2 is 28%. In their 2013 sustainability report, the operators of Hong Kong's mass transit railway (MTR) estimated that over a 120 year life span the embodied CO_2 of the structures, tracks and equates to 13% of total emissions (MTR, 2013). A study of a series of US systems found that capital CO_2 may equate from 0.8 to 1.5 times the operational CO_2 (Chester and Horvath, 2009b). These studies were generally more comprehensive, explaining in part the larger footprint than the MTR and Crossrail studies. Aspects such as maintenance and manufacturing were included within the scope.

In the context of rail, road and aviation, these emissions are the result of manufacturing the vehicle itself. For cars this is often very small (3%) but for buses it may be larger (28%) and relatively small for operationally dependent aviation (Chester and Horvath, 2009b).

Intra-urban Travel

High-speed rail has stringent geometrical requirements. The horizontal and vertical alignment stipulation means that it is often difficult to adjust the line to local topography. A relatively small trade-off between vertical and horizontal alignments can be made, dependent on the type of HSR rolling stock used. For example, the French TGV Atlantique rolling stock tolerates relatively steep gradients with long-radius horizontal curves, in contrast to the Italian Pendolino rolling stock which tolerates shorter horizontal curves but with less steep gradients. Despite these potential contrasts, the relatively strict HSR geometry requirements will often result in an increased need for structures such as bridges or viaducts to allow for the required large-radii curvature. The carbon footprint of constructing such structures has been well documented in the literature (Chau et al., 2012a). The costs are directly related to the engineering challenges faced as a result of the tight geometric requirements of HSR. Therefore, each HSR line must be assessed in terms of its specific context.

The International Union of Railways (UIC) assessed the capital CO_2 emissions of four different HSR lines – the South Europe Atlantic, the LGV Mediterranee, Taipei–Kaohsiung and Beijing–Tianjin (UIC, 2011). For each of these lines, a *cradle-to-grave* analysis was carried

out on earthworks, material transportation, structures, track, signalling equipment, stations and rolling stock. In a comparison of these lines in terms of tonnes of CO_2 dioxide due to construction per kilometre of line and year, the two French lines compare similarly, at around 60 tCO_2/(km year). The Taipei–Kaohsiung and Beijing–Tianjin lines are considerably higher, at around 175 tCO_2/(km year) and 140 tCO_2/(km year) respectively. The large difference between these two lines and the two French lines is predominantly due to the use of bridges, tunnels and viaducts on the Taipei–Kaohsiung and Beijing–Tianjin lines.

In addition, the French national rail operator (SNCF) carried out an LCA of a new 140 km TGV line from Rhine–Rhone in eastern France. It was calculated that the construction amounted to 750 000 tCO_2eq (SNCF, 2011), roughly equating to 60 tCO_2/(km year) assuming a 100 year lifespan (as per methodology in (UIC, 2011)). This is a similar value to those estimated by UIC for the LGV Mediterranee (approx 68 tCO_2/(km year)) and the South Europe Atlantic (approx. 60 tCO_2/(km year)). The low capital CO_2 costs of the French lines are a result of few structures such as tunnels (5% of LGV Mediterranee) and viaducts (6.4% of LGV Mediterranee) (UIC, 2011). Conversely, the Taipei–Kaohsiung line in Taiwan travels through a mountain range in the densely populated west coast and is mainly run on viaducts (73%) and in tunnels (13%) (UIC, 2011).

The first phase of HS2 in the UK is a planned new 225 km HSR line from London to Birmingham. Projections of the scheme's CO_2 credentials estimated that the capital CO_2 would be approximately 5.59 $MtCO_2$eq (HS2, 2013). This value is very large in comparison to the studied existing projects, and is largely due to the tunnels that are intended to reduce noise pollution and improve visual amenity (HS2, 2013). Tunnels account for over 1.1 $MtCO_2$eq alone along the HS2 line. Research has found that using a 9.8 m diameter over a 10 km long tunnel for a 320 km/h train equates to 64% additional energy consumption when compared with an at-grade open line (HS2, 2009). Tunnels therefore pose two CO_2 emission challenges: Firstly, they increase the operational energy required by the rolling stock indefinitely; secondly, they have a large capital investment, primarily as a result of boring and embodied material energy. The preference for tunnels to mitigate environmental metrics such as noise and visual amenity may be reduced if a larger premium was put on the two-sided CO_2 environmental impacts they have. However, clearly the choice to utilise tunnels incorporates other considerations beyond a simple trade-off between CO_2 emissions and noise as well as other environmental impacts. For example, separating land has animal migration impacts and has other wider local connectivity implications that would need to be considered.

Operational CO_2

Urban

There has been a general trend in the reduction of operational emissions, across all modes. This has been mostly achieved through technical innovations such as improved engines, reduced weight, aerodynamic improvements and the decarbonisation of the energy supply. Mass public transport systems have long been identified as offering lower CO_2 emissions compared to road vehicles (Newman and Kenworthy, 1999), (DfT, 2015b). London Underground has achieved a reduction in emissions from 77 gCO_2/km/pkt in 2005 to 61 gCO_2/km/pkt in 2015 (TfL, 2015). This compares to the average UK car performance of 138gCO_2eq in 2011, a 40 gCO_2/km saving since 2001 (RAC, 2012). The increase in load factor for car significantly improves the emission performance and considering 60% of London car trips are made with one occupant (TfL, 2012), there is scope for large savings. In most major infrastructure projects it is the desired aim to achieve modal shift from personal vehicles to public transport (Ogilvie et al., 2004) in order to reduce congestion and emissions. In the case of North American cities, dominated by personal car use, the challenge to providing adequate public transport is mostly related to land use and the challenge of achieving sufficient density. In cities with existing sophisticated public transport systems the shift from personal vehicles to public transport has proved to be problematic with many examples of shifts actually coming from other public transport routes, such as the case in the Jubilee line extension (Jones et al., 2002), (Saxe, 2016). Estimations for this project found that 14% previously used National Rail services, 21% used the DLR, 7% used bus services and only 2% were shifted from car (Lane et al., 2004). Copenhagen's metro project resulted in the modal shift of 8-14% from car users, 70-72% from bus passengers and generated an induced demand of between 13% and 18% (Vuk, 2005). In Toronto, Saxe found that for the first 6 years of operation the Sheppard line emitted more CO_2 emissions that the bus route it had replaced and that the possible savings (66 $KTCO_2$eq) via mode shift from cars may have been lost through induced demand (Saxe et al., 2015b).

Intra-urban

In 2010 the HS1 line in the UK was carrying less than a third of the passengers it projected at the time of tendering (Booz, 2012). The biggest criticism of HS1 was that the aviation sector solved the connectivity problem itself, primarily through low-cost carriers, thus impacting the ability of HSR to compete (NAO, 2012). However, such low-cost aviation is unlikely to continue as global emissions regulations impact on the competitive pricing of aviation. Thus, although the projected returns were not made in the short term, they may come to fruition in the coming decades.

For HSR, the load factor was found to be the most influential aspect on the carbon credentials in the four UIC case studies (UIC, 2011). The French TGV lines operate at an average load factor of 70%, in sharp contrast to the 46% achieved by the Taiwan HSR line. UIC carried out

a sensitivity analysis of -20% and +20% of the central assumed load factor on an HSR. The low case (-20%) was calculated to emit 21 gCO_2eq/PKT and the high case (+20%) was found to emit 11gCO_2eq/PKT (UIC, 2011) - this is a significant difference. HSR lines, such as those in France, have traditionally run at close to passenger capacity and therefore fare much better in terms of the carbon/PKT metric (UIC, 2011).

2.3.2 Spatial Boundaries

Urban Travel

Kenworthy and Newman have carried out many studies into the interdependencies between land use and transportation. They showed how gasoline consumption varies across the US, as a result of land use and transportation variations, not due to price or income variations (Newman and Kenworthy, 1989). For gasoline the US was consuming average four times more than that of Europe and that gasoline price, income and vehicle efficiency explained only 50% of this difference in behaviour. They depicted the relationship between annual gasoline use and urban density as shown in Figure 2.7. Those with high gasoline use and low density were almost exclusively US cities where so-called *urban sprawl* (Batty et al., 2003) has occurred. In a comparison between different areas in the UK, high density areas were associated with reduced transport emissions per capita (DECC, 2015).

Rail has been long heralded as a means of encouraging more compact land use and thus reducing emissions (Newman and Kenworthy, 1989), (Senbel et al., 2010). However, Knight made it clear that a rail project must be supported by development policies as it can not act alone (Knight and Trygg, 1977). The Bay Area Rapid Transit (BART) system in the US exhibited fairly modest impacts on the urban development in the area, with impacts mostly focussed on San Francisco and Oakland (Cervero and Landis, 1997).

The urban density of these areas adds a further dynamic in the form of capital CO_2 emissions. Projects such as Crossrail operate in high density areas and thus exist mostly underground. This in turn requires CO_2 intensive structures and construction, indicating that there is a trade off to be made here. High density lends itself to mass transport efficiency savings but also requires a great more CO_2 carbon inducing structures in order to be constructed in such a high density area. This is evident in the US systems, which often exist above ground due to less land pressure.

Intra-urban Travel

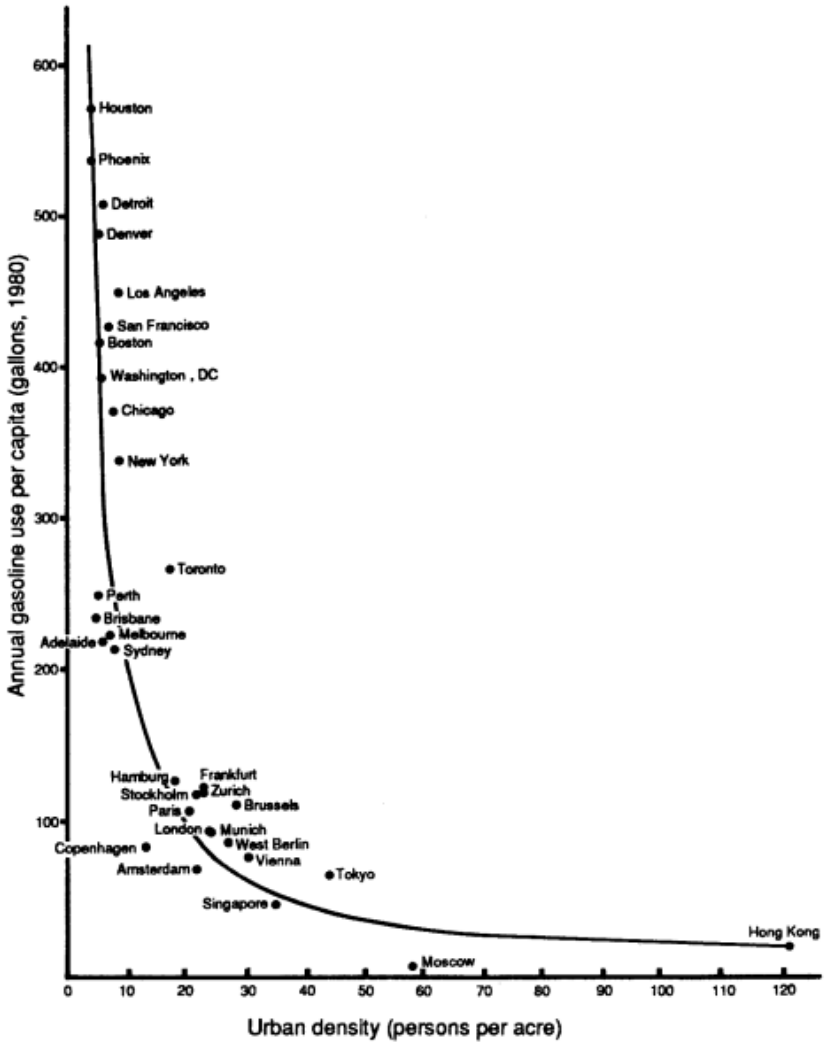


Fig. 2.7 Gasoline consumption and urban density (Newman and Kenworthy, 1989)

The land use impacts of HSR compared with aviation are predominantly a comparison of land use near a train station and an airport. HSR stations often share services with local and regional services (e.g. King's Cross St Pancras) and may have land uses extending far beyond that of a transport hub. Airports are generally located on the outskirts of urban areas as they require a large amount of space and there must be sufficient buffer space between them and residents to protect from noise and other pollutants (Kussner and Tsun, 2011). Airports are associated with negative pressures on nearby land use due to noise, safety, environmental degradation and economic concerns (Brockway Jr, 2007). Conversely, train stations are generally found in central locations (e.g. St Pancras in London, Gare du Nord in Paris and Gare du Midi in Brussels) and are increasingly being used as part of large, multi-use developments. An assessment of HS1 found that the impact of developments at King's Cross, Stratford and Ebbsfleet could be worth £10billion as a present value over 60 years (LCR, 2009) and such value uplifts are now being used to finance new infrastructure through land value capture mechanisms. The land use impacts of aviation are generally limited to the airports themselves and the takeoff and landing corridors. Contrastingly, HSR involves the creation of rail corridors that impact land use over large and highly variable distances (Kussner and Tsun, 2011). The land use implications of HSR beyond stations are more complex and interlink with a multitude of other land use considerations along their length.

2.3.3 Sensitivity

The use of public transport systems to reduce emissions provides a range of case study experiences. It is clear that CO_2 emission reductions are far from guaranteed and that attention must be paid in a multi-faceted way. Consideration must be made to the electricity supply, actual ridership, actual mode shift and long term mode trends (Saxe et al., 2015a). First and last mile challenges may result in public transport emission savings being negated or lost in the event of sufficient car usage between transport hubs over short distances.

Life cycle analysis has historically had to make do with significant data uncertainty (Huijbregts et al., 2001). Static and/or averaged inputs are often used for the means of appraisal and options comparison. In many cases, the idealised conditions simulated in these analyses may be significantly distorted from those conditions in reality (for example, high versus low load factor on a bus (Chester and Horvath, 2009b)). In order to assess the importance of different factors, a number of illustrative examples were presented here.

The case studies have shown that there are a series of specific factors that must be considered when making strategic decisions about reducing transport carbon emissions. These factors can be categorised as:

1. Capital CO_2 emissions
2. Urban form impacts
3. Transport modal split

Capital CO_2 emissions have historically been ignored, especially for rail projects. This has changed in recent years, with major projects such as Crossrail and HS2 carrying out capital CO_2 assessments. The barrier to such assessments has historically been primarily data collection and data quality. Databases of historic project capital carbon data will become an invaluable resource for getting approximate figures in a timely manner for options appraisal.

It is understandable that such caveats are simplified in order to give comparatives. However, in the event where such sensitive inputs occur in the interface between the infrastructure and human decision making a great deal of risk has been identified. This is most evident in the travel demand for a transport mode which impacts urban form and modal split. It is clear from the literature review that a critical factor in emissions success for mass transit systems is the projected travel demand.

2.4 Transportation Modelling

The modelling of future transportation scenarios involves consideration to a range of different related but separate fields of thought, such as economic, social, environmental, and behavioural sciences. Demand forecasting specifically involves the prediction of future collective need for a transportation service and may be quantified in terms of the total demand for a range of origins and destinations at a given temporal resolution (McFadden, 1974) and the modes and routes this demand will take on the network. This travel demand is a function of land use, demographics and other socio-economic metrics and involves the integration of economic and land-use models.

2.4.1 Challenges

In 2002 Flyvbjerg carried out a statistical study on 258 transportation infrastructure projects of different types over a range of different geographic locations (Flyvbjerg et al., 2005). For 90% of rail projects passenger forecasts were overestimated by an average of 105% (as graphically shown in Figure 2.8) and for 50% of road projects the difference was ±20%. The study also found that there has not been an improvement in forecasting ability over the study period of 30 years. The authors speculated, built upon this, and identified the causes for such issues, as is shown graphically in Figure 2.9. These reasons were wide varying and in the case

of rail projects often relating to political pressure (Flyvbjerg et al., 2002). Other criticisms include weak partial data, lack of equity (project promoters dominate modelling), insufficient time (thinking, appraisal and analysis come when decisions have already been made) and the contradiction between software trends in transportation, and in the rest of the sector (Hollander, 2016). Such criticisms often suggest that modelling is often using as post-justification for a decision already made.

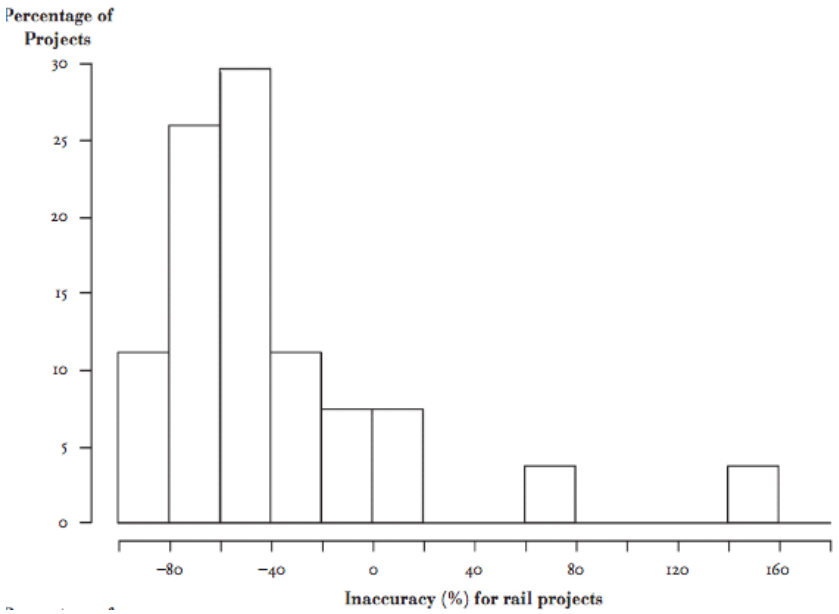


Fig. 2.8 Inaccuracy of demand forecasts for 27 rail projects from 1969-1998 ((Flyvbjerg et al., 2005))

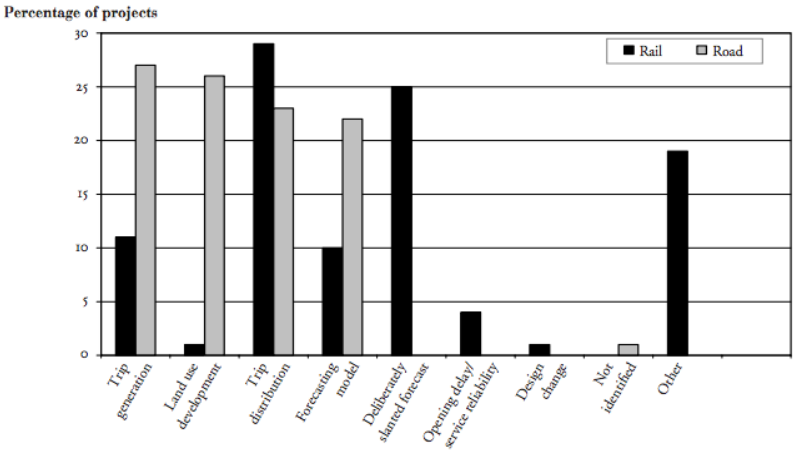


Fig. 2.9 Stated cases of innacuracies in forecasts (Flyvbjerg et al., 2005)

The literature clearly illustrates there is a significant modelling challenge as the phenomena which is being modelled . In order to assess this we will explore the current, dominant static modelling paradigm, that of the 4 stage model, against more dynamic modelling paradigms.

2.4.2 Static Models

In the traditional four-stage transportation model (McNally, 2007) the process is conceptualised as follows and shown in graphically in Figure 2.10.

1. Trip generation
2. Trip distribution
3. Mode choice
4. Route assignment

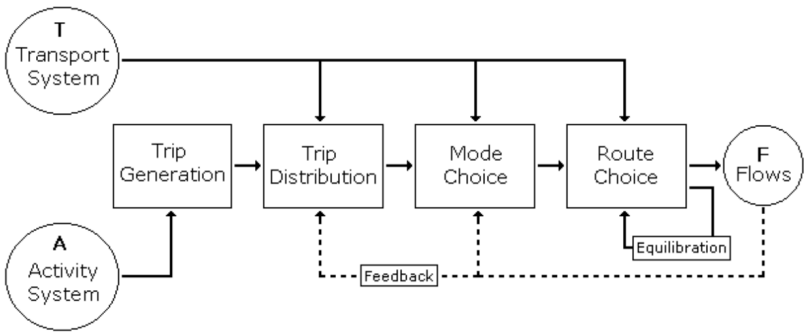


Fig. 2.10 The four stage model framework (McNally, 2007)

Model logic & interactions

Two of the stage impact directly on two factors that were identified as being highly influential on CO_2 emissions. The *land-use* related CO_2 emission impacts, discussed previously in 2.3.2, are directly influenced by trip generation. Mode choice directly impacts *load factor* and manifests in the sensitivity to travel demand for a given modes CO_2 emission credentials. As was discussed previously, it has long been established that there is a two way relationship between land-use and transportation. Changes in trip generation impact on land use and land use changes impact on trip generation.

Trip generation is the process of quantifying, for a range of different trip purposes, trips to and from different zones (McNally, 2007). For example, a work trip from a household to a place

of work (home based trip) or place of work to shop (non-home based trip). Land-use is thought to impact on transportation along three dimensions - density, diversity and design (Cervero and Kockelman, 1997). Density is a metric of persons to land, diversity measures the proportion of different land types in an area an design measures the type of infrastructure present in an area. Statistical studies have shown limited but not inconsequential relationships between these metrics (Cervero and Kockelman, 1997). The debate surrounding the role of the public and private sectors to intervene and *prime* areas in order to unlock development continues and is closely related to the prevailing economic paradigm in a given country (Mieszkowski and Mills, 1993).

Mode choice modelling is the process of predicting an individual traveller's decision on which mode of transport they will use for their particular demand. For a given location, at a given time, conceptually an individual may be presented with a possible route by car, a possible route by bus or a possible route by train. Each of these different modes will come with different attributes, time cost, financial cost or other variable or static metrics and from these attributes a traveller may make a decision. Of course, these decisions do not occur in isolation and the localised decision of a given individual may impact on other individual(s). This is most easily conceptualised in the case of car traffic. Collective decisions to utilise a given road will result in congestion when the demand exceeds the free-flow supply and as a result journey times will increase. This journey time increase then changes the information for a future traveller, as it increases their journey time by that mode, potentially increasing the relative attractiveness of another mode and resulting in modal shift.

Mode shift from personal vehicle use to mass transit systems is the general policy for reducing carbon emissions from transportation (Newman and Kenworthy, 1999). This often manifests itself in the incentivisation of public transit systems (reduced journey times, reduced price *etc*) or the disincentivisation of personal vehicle use (via increasing costs (e.g. taxation) or increasing journey times). The identification of the optimum means of reducing emissions is challenging in such a complex and constantly evolving environment. In the case of urban travel, the complex transport network consists of those modes which use road infrastructure (private cars, taxis, buses)) and those which use rail infrastructure (light/heavy rail, metro/tube/subway and tram). Each of these modes has different attributes, such as capacity, routing and timing. As a result of this, these different modes are generally modelled separately, with specific attributes to the modal nature.

Model inputs

In order to reflect real-world conditions accurate model inputs must be provided. Transport networks are complex and offer a multitude of options, via many different transport modes, for travelling to and from any location.

Road network

Harvesting road vehicle related data for all roads has historically been prohibitively expensive and often standardised functions of sample roads are used to find suitable values (Skabardonis and Dowling, 1997), (Mullick and Ray, 2012). Geospatial data such as road length, lane count, road type and survey data such as traffic counts has enabled the use of generalised functions such as bimodal journey time functions (Mtoi and Moses, 2014) and volume-delay functions in order to estimate likely road attributes. Such functions are derived from limited, old and extremely context specific studies resulting in a limited empirical evidence base which is increasingly far removed from the modern context (Rose et al., 1989), (Spiess, 1990). Traffic counts are often converted to Annual Average Daily Flows (AADFs) carried out over short periods and averaged over long periods (DfT, 2010), offering a limited snapshot and little in the way of temporal distribution. These functions attempt to generalise different aspects of a roads characteristics in order to create general functions without the need for input surveys. However, in doing so their ability to give outputs that consider the context specific nature of a given road reduces. Such differing characteristics can result in very different vehicular behaviour on roads that may be considered similar by these functions.

Public transport network

Public transport timetables provide a centralised resource for quantifying the journey time and financial cost attributes of public transport services. Public transport services are centrally coordinated and scheduled in advance in contrast to the decentralised/individual nature of most car journeys. In the case of cities where a centralised body is responsible for public transport services it is often possible to access all public transport mode data through one centralised repository.

Simplifications are usually employed in order to consider how a traveller may be presented with a particular service, for example journey times often include half the head time between services to give a static journey time that considers scheduling (DfT, 2014c). Such assumptions negate the identified impact of different timetables that is known to influence a traveller's view of a public transport service (DfT, 2014c). It also fails to consider the reliability of services

and their tendency to provide a level of service as is specified in the timetable. Many transport systems in major cities are stressed at times, often resulting in significant service impacts for travellers'. For illustration consider the service performance of the London Underground service (*http://tubestatus.net/graph*). The result may be highly variable service reliability that can have a resulting impact on traveller decision making. The difference between planned public transport services and actual public transport services may have an impact on the robustness of using the idealised timetable as a model input.

As a result of the scaling limitations and limited data availability, models have generally made do with limited and simplified data input that often does not reflect real-world conditions at a satisfactory temporal and spatial resolution.

Decision making heuristics

Separate to the veracity or resolution of the data provided is the process employed by the agent to make a modal decision. Within any given model type an individual agent decision making process logic must be programmed.

The prevailing economic paradigm in recent history has been based upon the notion that an individual will act rationally to maximise their utility (Simon, 1955). This follows a rational framework where an agent is deemed to act logically in their own self interest, has perfect information (Simon, 1979) and this information is reflective of real-world conditions (as challenged in previous section). In the context of transportation modal choice decision making, this is usually represented as a form of utility theory known as the generalised cost method (for example WebTAG, (DfT, 2014a)). This is where the monetary and non-monetary costs of a journey are summed per option and the agent maximises utility by minimising this cost. The valuation of non-monetary and even monetary metrics is problematic as it requires a value judgement (Simon, 1979). For example it is possible to consider non-monetary differences in time perception using variable weights for specific modes allowing time spent waiting to be weighted as a multiple of time in transit (Jain and Lyons, 2008). This may be extended to different modes, at different times, for different roles (i.e. for work versus leisure travel), for different people in different locations. Standardised coefficients may be used for a given analysis (DfT, 2014a) but studies have illustrated that such value judgement driven methods have little in the way of empirical justification (Wardman, 2008). In a London specific study, the exhibited behaviour from a travel survey was combined with crowd-sourced real-time location device data in order to assess the relationship between exhibited decisions, the financial cost of the journey, the time cost of the journey, the mode of travel, the role when travelling and the distance of the route. No significant relationships were found from a multi-variable regression (Casey et al., 2016).

In the context of transportation decision making, the discrete choice method has emerged as a specific tool for modelling a decision from a series of discrete choices. This was originally put forward in 1985 (Ben-Akiva and Lerman, 1985) with a series of applications and built upon with simulation results by Kenneth Train more recently (Train, 2009). The discrete choice method involves the qualitative decision making process when appraising the relative importance of different attributes for a series of options from which one option must be selected. The general discrete choice method requires the context of the individual choice, specifically the quantification of the other options in order to provide the comparative value of different options. Thus, for transportation application the nested logit model is used as it models alternatives as related. The existence of this theoretical framework led to numerous attempts to capture attributes related to this process. Examples of this include the more obvious metrics such as distance, travel time and scenery (Ben-Akiva and Lerman, 1985), socio-economic metrics, network knowledge (Ramming, 2001), the difference in perception and reality (Cascetta et al., 2002) and with real-time traffic data (Dia, 2002).

Despite the advances in discrete choice methods, there is still a conflict with recent advances in behavioural economics. More recently branches of physiological understanding have been merged with that of economics in order to create a better theoretical framework within which to model human decision making. The logic of rational choice is inconsistent with the cognitive psychology understanding of apparent irrational human decision making (Tversky and Kahneman, 1985). Using a behavioural perspective Tversky and Kahneman illustrated examples where decision making under uncertainty exhibited behaviour at odds with that used generally in utility theory (Kahneman and Tversky, 1979) and more specifically with that used in logit models within the discrete choice framework.

The specification of which behavioural model to use has become ever more challenging with the evolution of mobility as a service. In some cases, a simplification has occurred for users. In the case of London, one payment method may be used on all TfL modes across the city. This was previously only *Oyster* cards but has since been extended to include *contactless* bank cards and mobile phone payments. Such technological changes streamline the payment process and by reducing frictional costs such as the need to have cash, or to queue to buy a ticket, have been shown to impact behaviour significantly (Halpern, 2015) despite their small financial value. There is also the advent of real-time data provision in the form of service status twitter updates, shortest path computations and passive information provision (Samsel et al., 2014). Such technologies mark a shift away from experience based routing to that of data driven routing with a far greater awareness of the status of the network both spatially and temporally. The use of such services can result in an agent having as close to perfect information on the network in real-time as has historically been possible. Questions still remain

on the veracity of such data and the inherent biases that exist in the transformation of raw data (location device informed journey times, bus delays) to a route with an algorithm that requires some judgement calls on tolerance for service interchanges for example. Further adding to the complexity are new disruptive business models such as Uber and Lyft. These services do not provide the ability to book ahead of time and requests for transportation must occur in real-time. Pricing structures are dynamic and have been shown to leverage well known patterns in human movement (Salnikov et al., 2015). It has even been found that the chief motivator for travellers selecting these services was the ease of payment (Rayle et al., 2014). It is clear that complex business models such as these and the empirical data from users on their exhibited decision making are extremely difficult, if not impossible to capture in the non-behavioural decision making models such as WebTAG. There is also the need to consider the changes that are likely to come in the face of autonomous vehicles and the fundamental impacts such technologies may have on what in many areas is still a motor-vehicle centric transport system. In the context of reducing transport emissions such technologies offer many possibilities, such as the reduction in operational emissions via better routing and driving and higher load factors (better resource allocation).

It is apparent that travellers are being confronted with an ever more complex transport environment. The private sector is showing increasing awareness of behavioural *nudges* and how they may be used to influence behaviour.

Limitations

The limitations may be summarised as:

1. The sequential nature of stages limit the model interactions and does not reflect real-world decision making.
2. Limited trip types, usually dominated by work trips and excludes increasingly important leisure trips
3. Specific behavioural rules must be given across aggregated populations and are limited to the trip types used
4. Spatial and temporal resolution of inputs - zones are often large and usually AM, PM and IP (inter-peak) time slices are modelled
5. The spatial and temporal resolution of inputs also impacts on scenario testing, dynamic and context specific policies can not be modelled
6. Modelling paradigm - fundamentally, the model assumes equilibrium and is deterministic

2.4.3 Dynamic Models

As our appreciation of the inherent complexity in the world has grown so too has the complexity of the models we use to try and capture it (Batty and Torrens, 2005). This is apparent in the shift away from centralised, top-down approaches to bottom-up decentralised approaches where the complexity, uncertainty and apparent chaos of systems, such as cities, is recognised in the very nature of the model itself (Batty, 2007). Agent based models (ABMs) and cellular automata (CA) have become the dominant paradigms in recent years (Crooks et al., 2008), spurred on by computational advances which permit ever more complex simulations. The basic principle of an ABM is that discrete agents with distinct behaviours interact to bring out macro behaviour (Silva, 2010), (Silva, 2011). A dynamic simulation can allow for the system processes analysed at the level of their constituent elements (Crooks et al., 2008) and thus can permit a better understanding of the agents involved, their stochastic and heterogeneous attributes, and how their complex interactions lead to exhibited macro level behaviour (Silva et al., 2014). Recent advances in computational capacity have enabled more complex, dynamic simulations to be possible (Balmer et al., 2006).

There is a large body of literature associated with the application of ABMs to a wide range of problems in biology, physics, chemistry and economics. Within this wider context, there are numerous general software packages such as NetLogo (Wilensky, 1999), MASON (Luke et al., 2004), Swarm (Minar et al., 1996), Repast (Crooks, 2007) and Ascape (Parker et al., 2001). However, such general packages suffer from limitations in the geospatial dimension (Crooks et al., 2008), a key component of transportation simulations.

Specific to the transportation sector TRANSIMS (Smith et al., 1995) and MATSim (Balmer et al., 2009) offer true multi-modal capabilities and agent level computations. However, the application of these packages with fine temporal and large spatial resolution has a significant computational demand which may require simplifications in agent decision making logic (Manley, 2013). One such simplification is the introduction of random errors as a proxy for actual differences in the agent population (Zheng et al., 2012) and the issues this represents when comparing the model to observed behaviours (Nagel and Flötteröd, 2012). Beyond this, such dynamic models are also harder understand, leading to challenges when attempting to interpret and communicate simulation results due to the complexity of interactions.

These models are also at odds with the current trends of decentralisation in computer science. Rather than retrospectively adding parallel computational capability after its construction, parallel computation should be considered from inception and during construction.

2.4.4 Computational implementation

The limitations of current modelling paradigms and their model inputs have been discussed. Agent Based Models have identified as a promising modelling paradigm and temporally and spatially dynamic inputs deemed important for capturing context specific information. However, even to assess the value of such methods, a range of technical computational challenges must first be addressed. These general computing trends are then compared to the specific challenges posed by a modal choice and route assignment agent based model.

The use of fine spatial and temporal resolution data in a model results in high computational demands that require scaling to be considered at a fundamental level and not simply a secondary concern post model creation. In the context of transportation systems this involves the efficient computation of spatial graph data structures.

Parallel computation is the procedure of executing a large process as a series of smaller sub processes simultaneously. A complex problem may be divided into sub problems, solved and then combined on completion affording significant time savings. Historically, this distribution involved the use of separate pieces of hardware but recent advances in virtualisation has permitted parallelisation to occur across software and not just hardware.

Conceptually, serial, sequential computing and parallel computing may be visualised as shown in Figure 2.11.

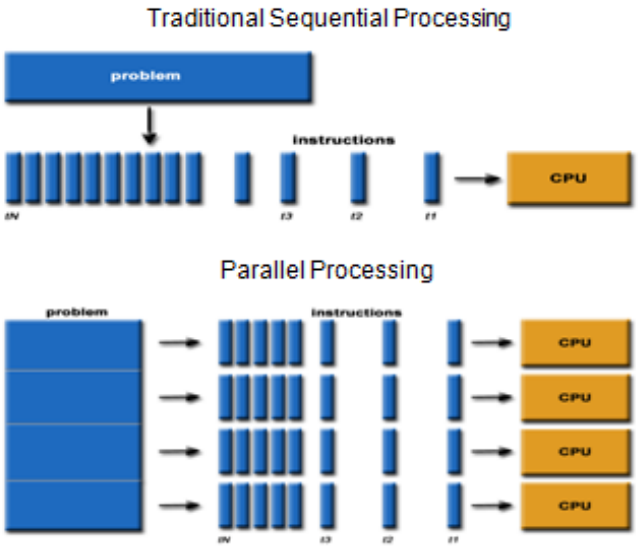


Fig. 2.11 Serial and parallel computing (Barney, 2016)

Parallel computing involves the deconstruction of a task into a series of discrete parts that may be solved concurrently. The more a process can be split up, the faster it can be theoretically run. However, the reduction in run time via parallelisation is constrained by the

inherent complexity of the process itself. This theoretical capacity of a given task is the focus of Amdahl's Law (Amdahl, 1967) which gives the theoretical speedup against number of processes for different parallel portions, as presented in 2.11. This shows that those processes with high parallel portions, commonly referred to as "embarrassingly parallel", can be significantly speed up by computing each part concurrently on a separate CPU. Conversely, those processes which are inherently serial may offer little or no speed up across multiple processors. The process of splitting a complex job into a series of sub jobs may be in itself a significant computational task and there are also issues with parallel programming itself, including memory limitations, communication demands and load balancing. The actual computational challenge faced here will now be considered.

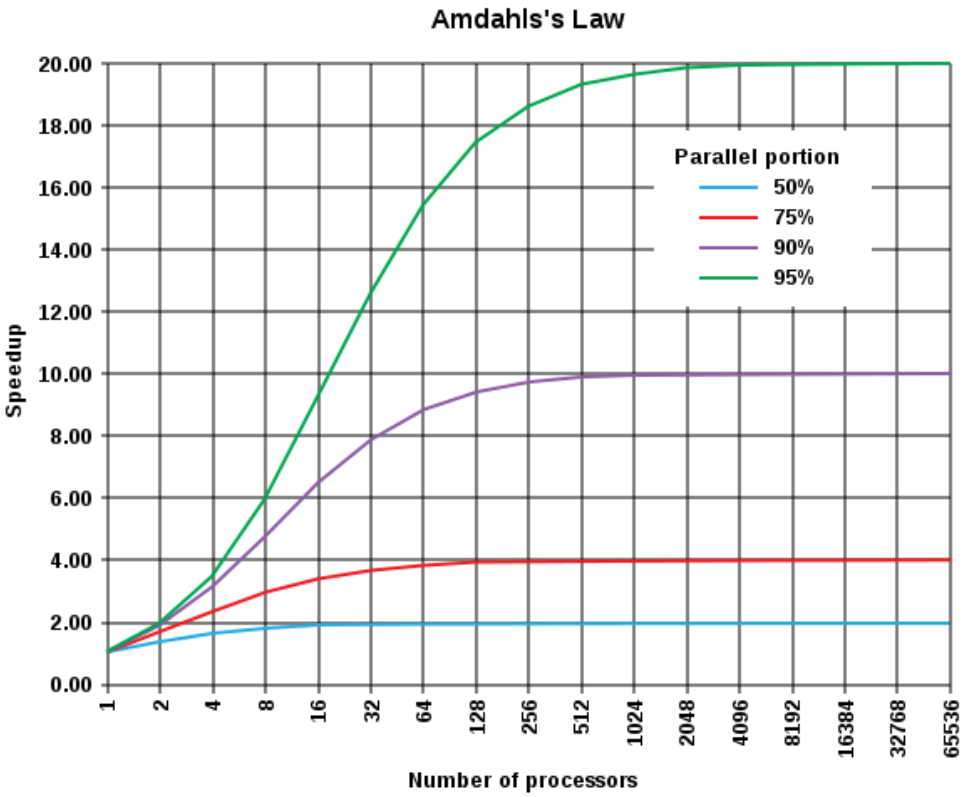


Fig. 2.12 Amdahl's Law (Amdahl, 1967)

Parallel computing and the problem definition

Modal choice and route assignment may be conceptually considered as a problem consisting of two distinct computational sub-problems:

1. The graph problem

2. The decision making problem

The first problem consists of graph computations, such as shortest paths, to act as inputs for the decision making computations. An example shortest path query for travel via car is shown in Figure 2.13. Multiple shortest path queries, one per available mode may be used to act as inputs for a single agent decision. Secondly, these different options are then compared and the agent decision making may be computed. An example process with different modal inputs is shown in Figure 2.14.

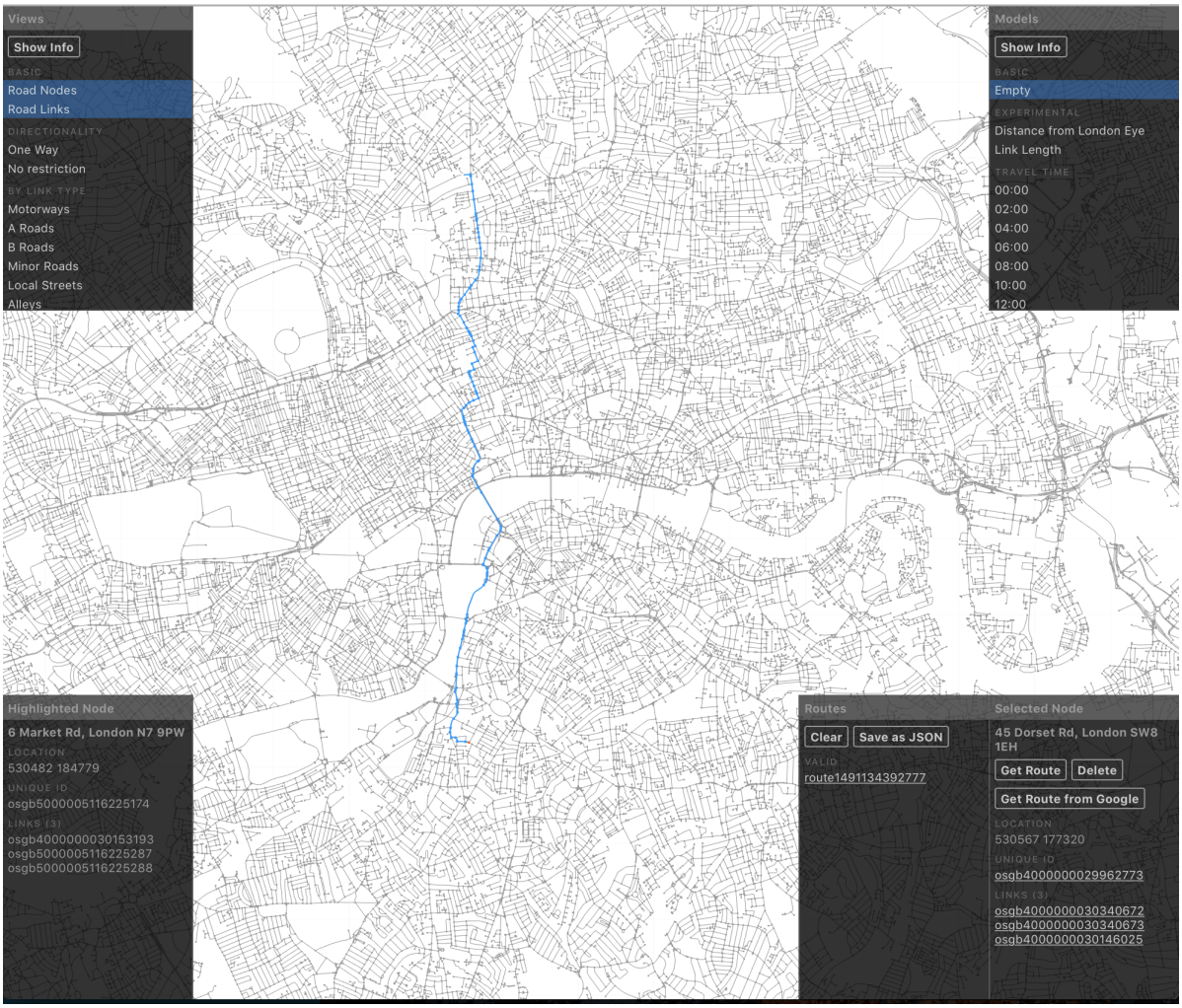


Fig. 2.13 Graph shortest path query illustration

In the case of a modal choice and route assignment ABM the agent interactions occur by relaying the impact of past agents as inputs to future agents, as illustrated in Figure 2.15. For example, an increase in users on a given road will manifest itself as an increased journey time for future traveller decisions involving this road. In order to increase model complexity the

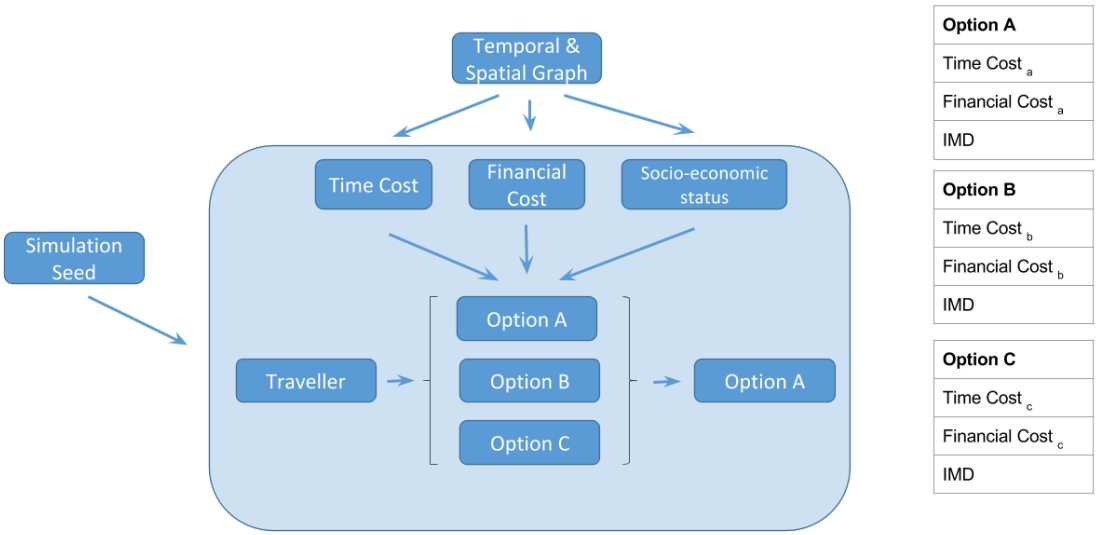


Fig. 2.14 Example agent decision making process

time steps between each run may be reduced in an attempt to better match the time dependent behaviour exhibited by travellers, as shown in Figure 2.16. However, this must be constrained by a variable we may call *information lag*. Computational constraints aside there is a limitation on the realistic time step definition as even assuming perfect information may be calculated and relayed, there is a delay between the measurement of this data and when it is actually provided to an agent.

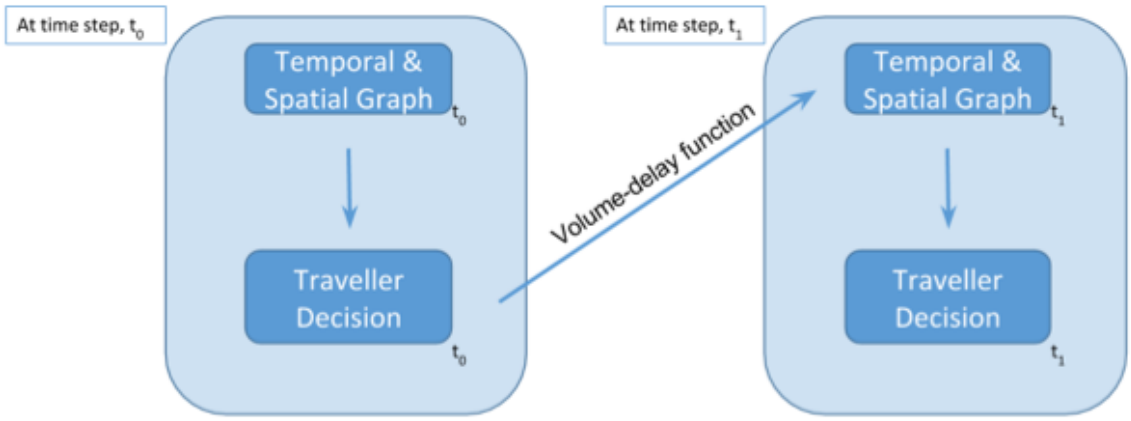


Fig. 2.15 Agent feedback between time steps

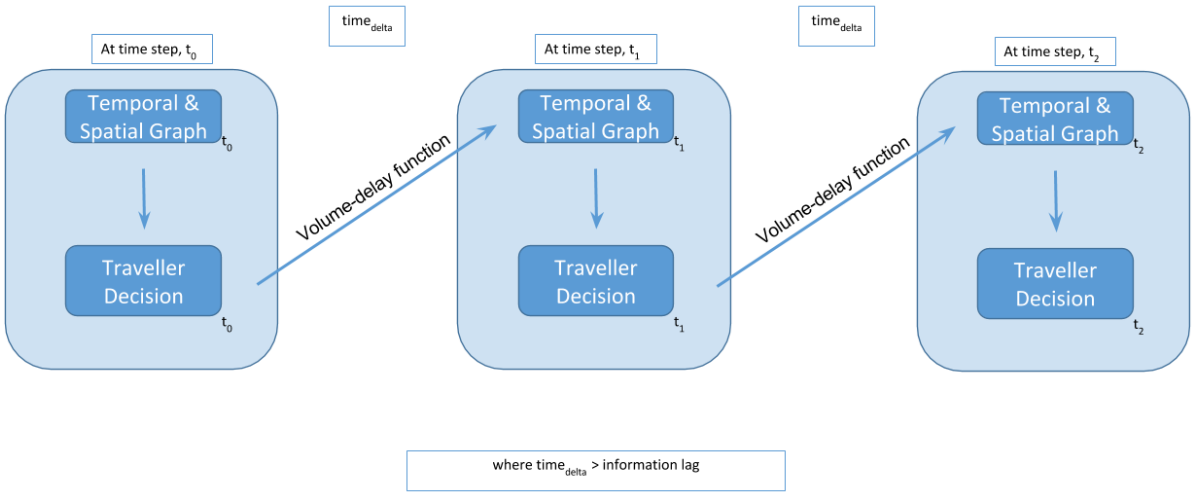


Fig. 2.16 Model time step iterations and constraints

The second problem falls into the *trivially parallel* category. Since there are no interactions between agents within a given time step, each process may therefore be defined as independent and therefore may be separated and processed in parallel with little effort.

MapReduce paradigm & agent decision making

MapReduce is a programming model for the processing of large datasets across clusters of machines (Dean and Ghemawat, 2008). MapReduce built upon the Google File System, a distributed file system for large distributed data-intensive application (Ghemawat et al., 2003). The MapReduce paradigm abstracts the parallelisation into two functions; the map function, and the reduce function. Beyond the user specification of these functions, the remaining tasks of parallelisation across many machines and tasks such as fault handling and efficient management of memory and communication are managed to varying degrees.

The execution of a MapReduce model is shown in Figure 2.17. The Map function transforms an input pair to a set of intermediate key/value pairs. The Reduce function takes this as an input and merges these values to form a (likely) smaller set of values.

The MapReduce paradigm was developed internally of Google and despite many academic and non-academic papers, the actual implementation code remained closed source. After the seminal Google File System and MapReduce papers, a series of open source projects emerged to apply these techniques under the Apache Software Foundation umbrella. Hadoop is an open-source framework for distributed storage and distributed processing across clusters of machines (Hadoop, 2016). Hadoop utilises the Hadoop Distributed File System (HDFS) (Shvachko et al., 2010) which provides high-throughput access to application data, similar to that of the Google

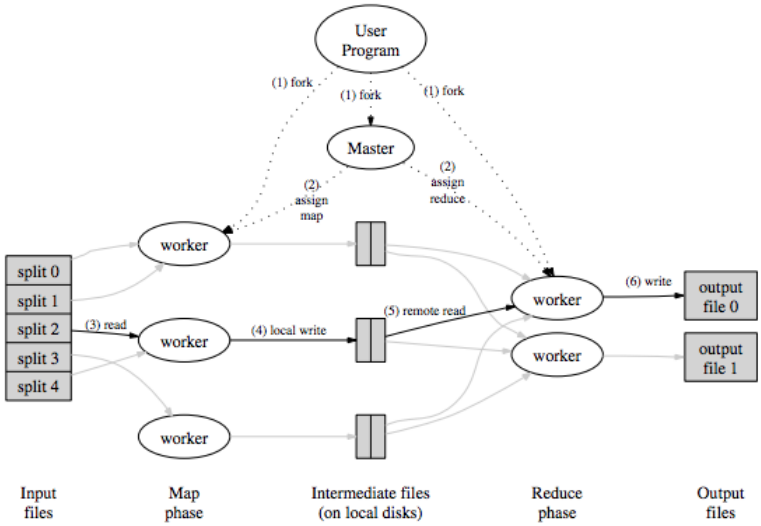


Fig. 2.17 MapReduce execution overview (Dean and Ghemawat, 2008)

File System. Hadoop comes with a resource managed (via YARN) MapReduce system for processing. By default, Hadoop expects hardware failure and thus automates the processes of handling failures and errors across multiple machines with efficient event responses.

More recently, as the Hadoop related ecosystem has grown, Apache Spark has emerged as a general engine for large-scale data processing (Apache, 2016). Spark offers interactive use from a range of programming languages (Java, Scala, Python and R) with integrations for varying data inputs (SQL, DataFrames and streams) plus support for running on orchestration abstractions such as Mesos and Kubernetes, with cloud providers such as Amazon Web Services (AWS) and local standalone implementations. Spark's execution engine supports cyclic data flow and has demonstrated run times 100 times faster than that of Hadoop for the same MapReduce process in memory (Apache, 2016).

These features have thus permitted those who historically did not have the ability to manage their own cluster to run large parallel process across tens and hundreds of machines. The general concepts discussed here pose opportunities for more efficiently computing the agent decision making processes.

Graph computations

Conversely, the graph problem does not lend itself so easily to parallelisation. As has been discussed, large graphs are useful for a range of computing problems, such as search engine page ranking, social networks and shortest paths. In recent times, the size of these graphs has

increased exponentially with graphs of billions and even trillion vertices and edges existing. By their very nature a graph is an interconnected structure and it is the ability to abstract complex relationships into a simple structure that has made graphs so useful across a range of problems. Lumsdaine (Lumsdaine et al., 2007) summarised these challenges as:

1. Data-driven computation

   "parallelism based on partitioning of computation can be difficult to express because the structure of computations in the algorithm is not known a priori."

2. Unstructured problems

   "Scalability can be quite limited by unbalanced computational loads resulting from poorly partitioned data"

3. Poor locality

   "Because graphs represent the relationships between entities and because these relationships may be irregular and unstructured, the computations and data access patterns tend not to have very much locality."

4. High data access to computation ratio

   "Graph algorithms are often based on exploring the structure of a graph in preference to performing large numbers of computations on the graph data."

Modern shortest path algorithms work on unweighted or undirected graphs (e.g social networks) and so many of the classic, early graph algorithms are still in use day, albeit in modified forms. In Table 2.2 different weighted shortest path algorithms are presented with their respective worst case complexities in Big O Notation (Danziger, 2010).

Table 2.2 Shortest path algorithms and their respective complexity

Algorithm	Complexity [4]	Reference
Bellman-Ford	O(VE)	(Bellman, 1958)
Dijkstra	$O(V^2)$	(Dijkstra, 1959)
A *	$O(\|E\|) = O(b^d)$ [5]	(Hart et al., 1968)

The A* algorithm uses heuristics in order to reduce computational times and in fact, has the same complexity as Dijkstra when no heuristic is used. Despite the significant performance gains, it is harder to implement than Dijsktra and its implementation can often become context specific to a given graph and not generalisable. Bellman-Ford is slower than Dijkstra, but unlike

Dijkstra it can handle negative edge weights. These algorithms may of course be implemented within the MapReduce paradigm, however MapReduce is not well suited to graph problems as for each step, it requires the entire state of the graph to be passed across, a process which is extremely expensive. Pregel was developed by Google (Malewicz et al., 2010) for a range of graph applications, including Page Rank classifications. Pregel advocated a vertex centric approach where *"a vertex can receive messages sent in the previous iteration, send messages to other vertices, and modify its own state and that of its outgoing edges or mutate graph topology"*. This method scaled successfully to billions of vertices. For context, the road network graph for the Greater London Area consists of around 450,000 vertices and 350,000 edges. Most graph databases and parallel computing frameworks use an implementation of Pregel for their graph compute support, e.g. GraphX within Apache Spark (Xin et al., 2013) or Apache Giraph (ApacheGiraph, 2016). However, in practice, the Pregel methodology has limited value for transportation applications due to the relatively small spatial complexity, but relatively complex temporal dimension as the message passing costs vastly outweigh any benefits (Shah, 2015).

As a result of this, horizontal scaling methodologies are generally not useful for scaling graph algorithms. Instead, compute times may be reduced via vertical scaling (improvements in hardware) or via algorithmic changes.

2.5 Chapter Summary

The uncertainty and risk in transport modelling has a direct impact on the reliability of any LCA analysis. Beyond, but closely related to the modelling of land-use changes, travel demand and modal choice modelling have a significant role in capturing this uncertainty. Limitations in the methods, the input data and model types have been discussed here with challenges identified in the literature highlighted.

2.5.1 Summary of issues

The literature may be summarised as follows:

1. Transportation is of fundamental economic importance in a globalised world
2. As other emissions decrease, the relative importance of transport sector emissions will increase
3. Different transportation modes have different life cycle emission characteristics. Traditional, operational only comparisons are likely to be misleading

4. Transportation modes must be appropriate and context specific. Apart from cycling and walking, no mode is inherently low CO_2. There is a requirement to consider context specific:
   (a) Capital CO_2 emissions
   (b) Urban form impacts
   (c) Transport modal split
5. For public transport, emission sustainability is extremely sensitive to load factor
6. Modelling modal choice requires dynamic models that:
   (a) use real-world data
   (b) are truly multi-modal
   (c) are able to support a range of decision making frameworks
   (d) can scale across large spatial areas whilst maintaining fine spatial and temporal resolution
7. These challenges occur in a backdrop of increasingly complex networks (with similarly complex attributes) and new, disruptive business models and technologies.

2.5.2 Research Focus

HSR has been identified by the EU (EU, 2010) and the UK Government (DfT, 2015a) as a low CO_2 transport option. The spatial extent of the HSR network in Europe is shown in Figure 2.18. From London, direct services already include Paris, Brussels, Lille, Avignon, Marseille and the Alps (Bourg St Maurice, Aime La Plagne & Moutiers). There are ongoing discussions to expand this for direct services to Amsterdam, Barcelona, Zurich, Basel, Frankfurt and Madrid. Beyond direct services, HSR services provide access to the European network through Paris, Brussels and Lille. From here, it allows travellers to access the French TGV services, German ICE , Spanish AVE and Belgian Thalys services.

The competitiveness of HSR over conventional rail and aviation is presented in Figure 2.19. It is shown that for distances up to 800km HSR offers time savings over aviation but that for distances less than 380km the advantages of HSR over conventional rail are not present. In comparison to other regions, central Europe offers sufficient density for modal shift away from aviation to HSR, within these estimated distance bounds (EU, 2010).

The creation of a Trans-European High Speed Rail Network is a stated objective of the European Union (EU, 2010). Such investment and integration of different domestic networks

Fig. 2.18 The EU HSR network in 2010 (EU, 2010)

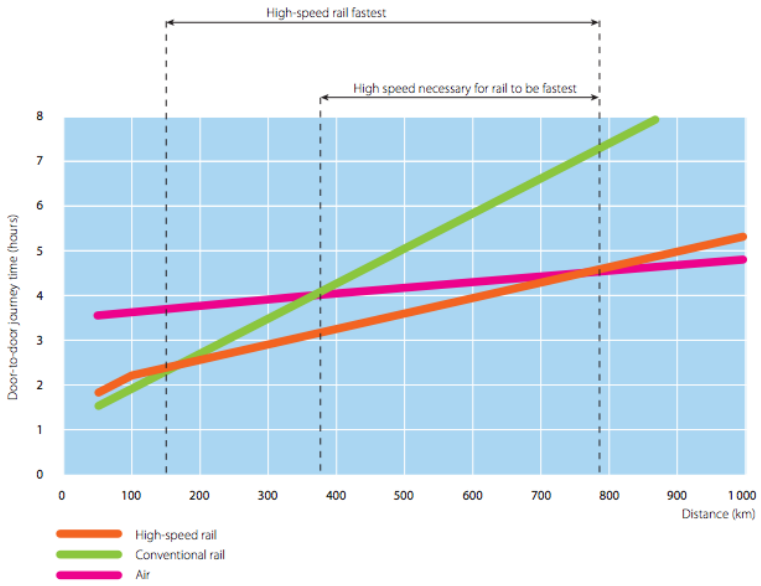


Fig. 2.19 Journey time versus distance for HSR, conventional rail and air transport (EU, 2010)

will increase the attractiveness of HSR to travellers and improve its competitiveness with aviation.

High speed rail presents the following traveller possible advantages:

1. Centre to centre travel (e.g. St Pancras to Gare du Nord, compared to Heathrow to Orly)
2. Less security restrictions (less time needed to clear security and more relaxed conditions)
3. More space - increased productivity potential
4. Large luggage allowance
5. Improved punctuality (Eurostar, 2016c)

Operational emission comparatives for the Eurostar service from London to Paris return (Eurostar, 2006) gives a clear difference between aviation (122 kgCO2e) and HSR (11 kgCO2e). However, such operational comparisons have been shown to be simplistic and potentially misleading. Indeed it is even possible that this section of the route may be many multiples longer than the journey to the HSR hub and thus there is potential that emissions to that hub may dwarf the longer journey itself. In modelling the value of HSR there is also the need to model the journeys to and from the *respective* origin and destination hubs to the *true* origin and destination for the traveller. It is this issue which requires a multi-modal model which spans

across modes and permits sufficiently wide spatial and temporal boundaries, in order to avoid the trap of insufficient impact assessment.

2.5.3 Identification of tool requirements

ABM's have been illustrated as a useful tool for modelling complex city transport systems and facilitating a range of different scenario testing exercises. However, the limitations relating to current transportation ABM packages discussed in Section 2.4.3 are significant. In order to scale, MATSim makes simplifications in decision making logic, with a loss of realism as a result. In section 2.4.2 we showed how models are generally fed with either static or very crude resolution inputs, resulting in a lack of realism in the decision making process. The use of this data in decision making is also problematic, with present utility based methods suffering from a lack of empirical evidence (2.4.2) and the literature review identifying a real need for a customisable framework that permit the use of a range of behavioural models, as appropriate and not up-front prescribing a particular logic. This is also extremely important in the context of artificial intelligence. The computational advances which permit these kind of fine resolution simulations open the door to data driven decision making heuristics which may be derived by artificial intelligence methods such as neural networks in some form of real-time calibration exercise. Such ambitions thus require for the support of extremely flexible data manipulation and avoid the locking down of manipulation into data silos or inflexible pipelines. The addition of resources to handle fine resolution temporal data and large resolution spatial data in a complex decision-making framework present increasing computational demands, something which the existing software packages were not designed to handle. Beyond this, there is also the need to visualise complex results in a meaningful way.

The ultimate hypothesis is that the use of an agent based model with real-time big data inputs will have better predictive ability than that of the incumbent methods. However, in order to assess this hypothesis a range of technical challenges must be first addressed in order to even permit the kind of model proposed here.

These tool requirements may be summarised as follows:

1. Dynamic ABM for mode choice and route assignment within the four stage paradigm
2. Multi-modal transport network
3. Realistic data inputs - considering traffic and service reliability
4. Computationally scalable

5. Support for different decision making models - existing discrete choice methodologies and emerging behavioural economic methodologies must be supported

Chapter 3

Using crowd-sourced real-time data to quantify infrastructure performance

This chapter presents three analyses featuring crowd-sourced real-time data. The first 2 analyses relate to road vehicle performance and the third relates to public transport performance.

1. Crowd sourced driving journey times
2. Vehicular traffic functions for London
3. Quantification of public transport performance

In the first section, the Google Directions API is used to quantify temporally dynamic road traffic conditions. In the second section, this data is combined with automated traffic counter data and a range of context specific volume, speed and saturation functions are derived. The third section assess three distinct sources for public transport performance quantification.

3.1 Crowd sourced road journey times

The large travel time distributions for journeys by road vehicles as a result of traffic has become strongly affiliated with the travel mode in recent times. Journey times for a static origin and destination are highly time dependent and may fluctuate significantly (Wang and Xu, 2011). Qualitative methods were historically employed to capture this data, through for example, radio traffic reports which were initially informed by telephone calls and then by inductive loop detectors and video cameras (Herrera et al., 2010). More recently, the prevalence of portable handheld devices has sparked a revolution in transportation data availability (Deville et al., 2014), (Ratti et al., 2006).

3.1.1 Literature review

Data from location enabled phones may be crowd-sourced and extremely fine temporal and spatial resolution data may be generated across extremely large spatial areas. However, this promise comes with a series of technical and non-technical challenges. The technical challenges relate to the data itself, its size, velocity and the processing requirements (Shekhar et al., 2012). Crowd-sourcing fine resolution data over long time frames results in large storage demands. Also, since this data is usually sent at close to real-time, the velocity of its movement requires efficient software digestion and sufficient hardware resources. The third challenge relates to the processing of this raw location data in order to quantify metrics specific to the desired output. There are accuracy issues with the location data, dependent on the hardware itself and the local environment (Wing et al., 2005). In the context of using this data to quantify journey times on roads, there is a need for transportation mode inference (Wang et al., 2010) and the trade off's that may be required when balancing accuracy and computational demand. Within mode inference there is also the need to filter out misleading inputs, such as those by delivery drivers (making multiple stops) and passengers on a bus in a segregated lane. The non-technical challenges relate specifically to the privacy and security implications of such fine resolution information (Iqbal and Lim, 2010). There is the potential for such data to be used in intrusive and even illegal ways. To date, legal frameworks have struggled to keep pace with the technical innovation of multi-national companies such as Google and Facebook (Wadhwa, 2014). However, more recently political pressure and high profile data breaches have resulted in stronger legislative actions, for example the General Data Protection Regulation (GDPR) from the European Union (EU). This new piece of regulation will become EU law in May 2018 and aims to "give citizens back the control of their personal data" (EU, 2016). The ramifications of a GDPR breach are significant, with a potential fine of up to 4% of global annual turnover.

These technical and non-technical barriers pose a challenge to the skill set and traditional workflows for the traditional transportation practitioner. As a result of these challenges, data providers who offer such location data as a service were investigated. In the more traditional transportation sector there are providers such as Inrix and more broadly there are tech companies such as Google and Microsoft. The vast majority of these services are targeted at the individual level and aim to provide high accuracy context specific information to aid individual decision making. This usually manifests itself in a traveller centric app such as Google Maps or CityMapper. This also applies more widely to logistics companies who seek to optimise operations (different variations of the travelling salesman problem).

Google's traffic maps can be visually inspected through their maps interface on any modern web browser on *google.co.uk/maps*. Their colour coded scale gives a qualitative feel for traffic

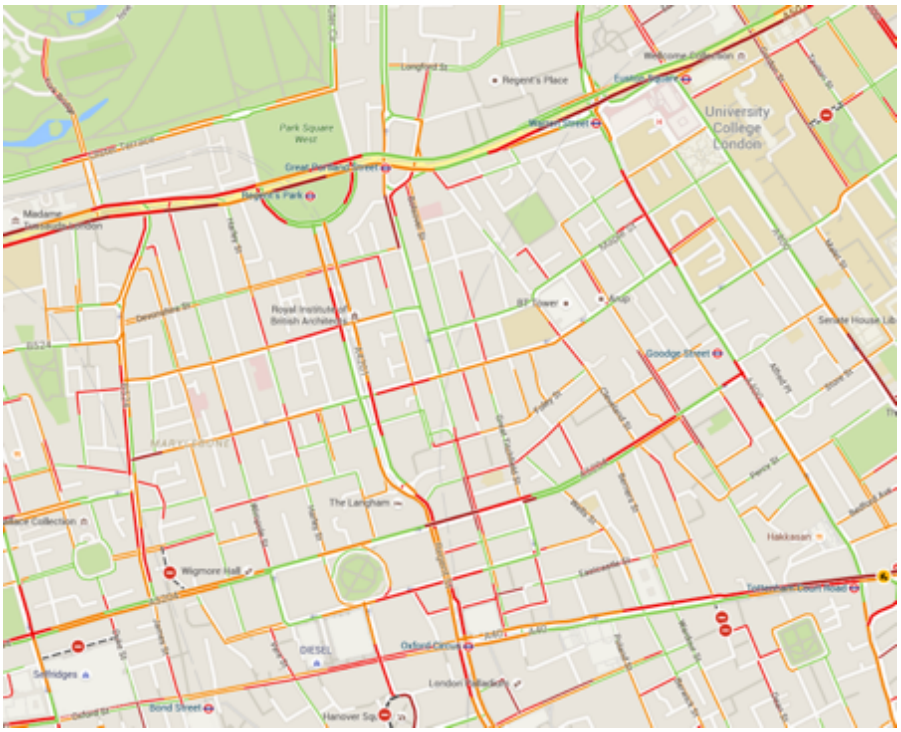


Fig. 3.1 Google maps traffic Layer, Camden, Soho, Marylebone and Mayfair areas of London (Maps, 2017)

conditions on the roads where Google have sufficient data, as is shown in Figure 3.1. However, such qualitative data is only useful for contrasting different traffic states on a road.

This research investigates the use of Google's traffic information through the Google Maps Directions (Google, 2016a) Application Protocol Interface (API). Dependent on their personal settings, an Android user or other Operating System (OS) user, with the Google Maps mobile app installed on their location enabled phone send anonymous data to Google. Such data is personally and commercially sensitive and so post-processing is carried out by Google in order to ensure that no-one user's movements can be isolated from the flows (Barth, 2009). The Directions API is a service that calculates directions between locations using a Hypertext Transfer Protocol (HTTP) request (Google, 2016a). The use of a HTTP request allows for scheduled and bulk harvesting of journey information between given origin and destination pairs.

The Google Directions API was therefore identified as a potential data source as Google provide the following as a service:

1. Google's large user base as input
2. Modal inference from individuals

3. Anonymisation

4. Programmatically available spatially and temporally

3.1.2 Methodology

The Directions API is not intended to be used as a means for quantifying global or local road transport statistics and usage to do so infringes on the terms and conditions of the service. Google supported this research with a Business level Maps account and permitted the usage of the data in this manner for research purposes only.

API request formulation

An individual Google Directions API request requires a time stamp, origin and destination specification within the HTTP request. Since this research makes use of business level credentials, the secure HTTP (HTTPS) protocol is enforced. This HTTPS request consists of a the core URL and a series of arguments, as shown:

```
https://maps.googleapis.com/maps/api/directions/
json?
& mode = driving
& origin = origin coorindates
& destination = destination coordinates
& depature_time = time
& authentication_parameters = secure identification
```

There are other optional parameters for driving related features such as waypoint inclusion, alternative route inclusion, avoidance of routes with tolls, highways, or ferries and arrival time specification [1] .

For a given Directions API request, Google return a shortest path. An example API response is presented in Figure 3.2. This API response is in effect a machine readable format of the more familiar Google Maps interface, as shown in Figure 3.3.

It is first necessary to consider the temporal resolution of the available Google Directions API data. The method is time dependent and the result returned is for the time of the request. In order to assess the underlying refresh rate of this service, a small spatial sample was taken at fine temporal resolution. A section of a road was selected at random, 1845m of the A214 Tooting Bec Road, and an API request made at 30 second intervals. The results are shown for

[1] This process was carried out in late 2015. Since then Google have released a range of new time related features which utilise their historic traffic information. (Google, 2016a)

Fig. 3.2 Example Google directions API response (Maps, 2017)

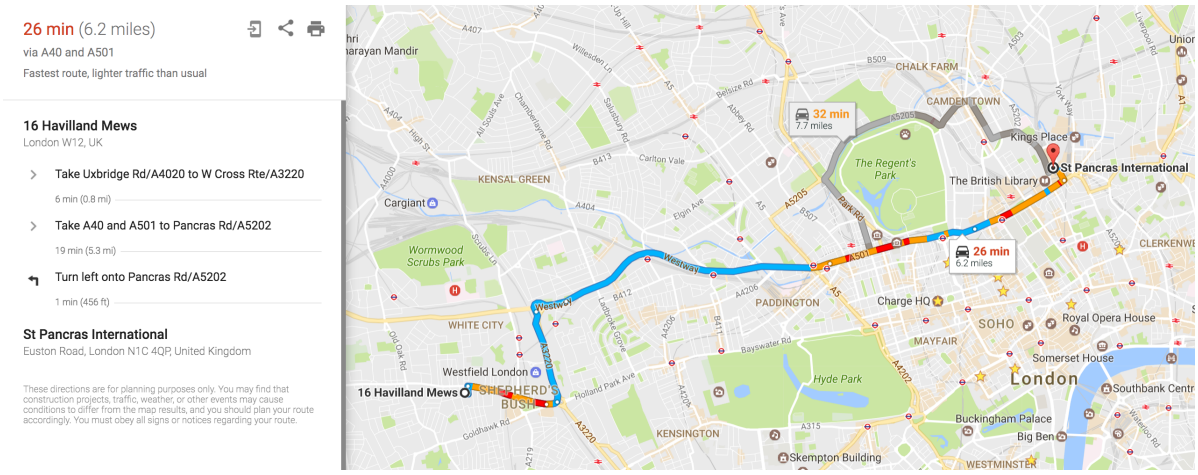


Fig. 3.3 Example Google maps query (Maps, 2017)

a 36 hour period are shown in Figure 3.4. The minimum journey time on this road was 168 seconds (at 4:13 am) and the maximum journey time 292 seconds (at 5:27pm), equating to maximum speed difference of approximately 10mph. Journey times rise steadily from 6am to 9:30am, then drop off mid morning only to rise to the daily peak at lunch time. There is sub hour fluctuations of 40/50 seconds most of the afternoon and then journey times gradually trend towards the daily low from 7pm. Interestingly the next day shows a different trend, with the daily peak exhibited earlier than the previous day at the conventional rush hour of 8-9am.

It is clear from Figure 3.4 that the temporal resolution of the incoming Google data is sub minute and significant fluctuations are evident within a given hour. Of course, such an output is consistent with real-world conditions as traffic congestion can generate and dissipate in minutes. From this test it can be concluded that journey time outputs from Google are refreshed at sub-minute resolution.

The motivation of this method is to capture the temporally dynamic journey times, not the shortest path output itself. As a result, the key information from an API response (Figure 3.2) is the **duration_in_traffic** field which provides a location device informed journey time for the given route. Within the **step** records, a range of different segments with associated meta-data are available, as given in Table 3.1. The key data to be extracted are the **polyline** and **duration** fields. Together, these present a geographically explicit polyline with a journey time, at a given time stamp.

Initial experimentation showed that the sum of the step durations was not equivalent to the **duration_in_traffic** record but rather equivalent to the **duration** record for the entire journey. This shows that Google do not disaggregate the **real-time** location device informed journey time across the constituent legs, rather they disaggregate the average condition record, **duration** across the constituent legs. Google define these two metrics as follows:

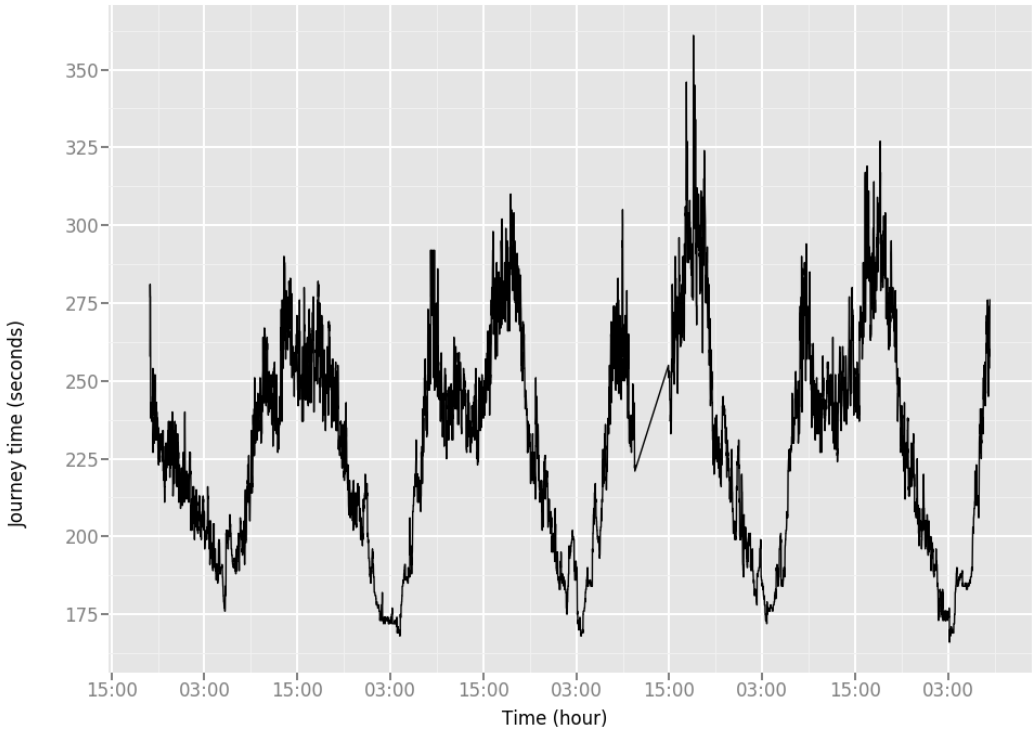


Fig. 3.4 Google journey time responses at 30 second query rate for A214 Tooting Bec Road (28 - 30th January 2017)

Table 3.1 Step specification of Directions API response

Field	Example
Distance	42
Duration	15
end_location	51.5307, -0.12473
start_location	51.5201035,-0.096507
html_instructions	Turn left at Beech Street
polyline	s'wshSBdhyAU"
manoeuvre	Turn left
travel_mode	DRIVING

duration : The average duration considering historic data (Google, 2016a)

duration_in_traffic : The duration at the requested time stamp (real-time) (Google, 2016a)

It is therefore necessary to quantitatively explore these two different definitions and how they evolve with time and space, across the study period. In order to do this, a large dataset was required.

Spatial specification

In the previous section the temporal resolution of this data source were explored and it was concluded that location device informed journey times may be found at sub-minute resolution. However, despite this data availability there is still a maximum API request constraint. The Google Maps Business Account provided for this research permitted 100,000 requests and thus a trade off must be made between the spatial and temporal resolution. It is therefore necessary to consider the identified case study of HS1 and use this requirement to drive such a trade-off.

The identified case study of High Speed Rail versus aviation for the case of London to Paris journeys resulted in the identification of three transportation hubs of interest - St Pancras International Train Station, London City Airport and London Heathrow Airport. These locations are shown in Figure 3.5. These locations provide direct services to Paris from the London area.

For the purpose of Google Direction API specification these hubs may be viewed as destinations. There are a series of different zonal systems utilised for a range of spatial statistical exercises in the UK (MSOA, LSOA and OS). There are 4835 Lower Super Output Area (LSOA) zones in the Greater London Area (GLA) and LSOA zones were identified as a reasonable compromise between size and resolution.

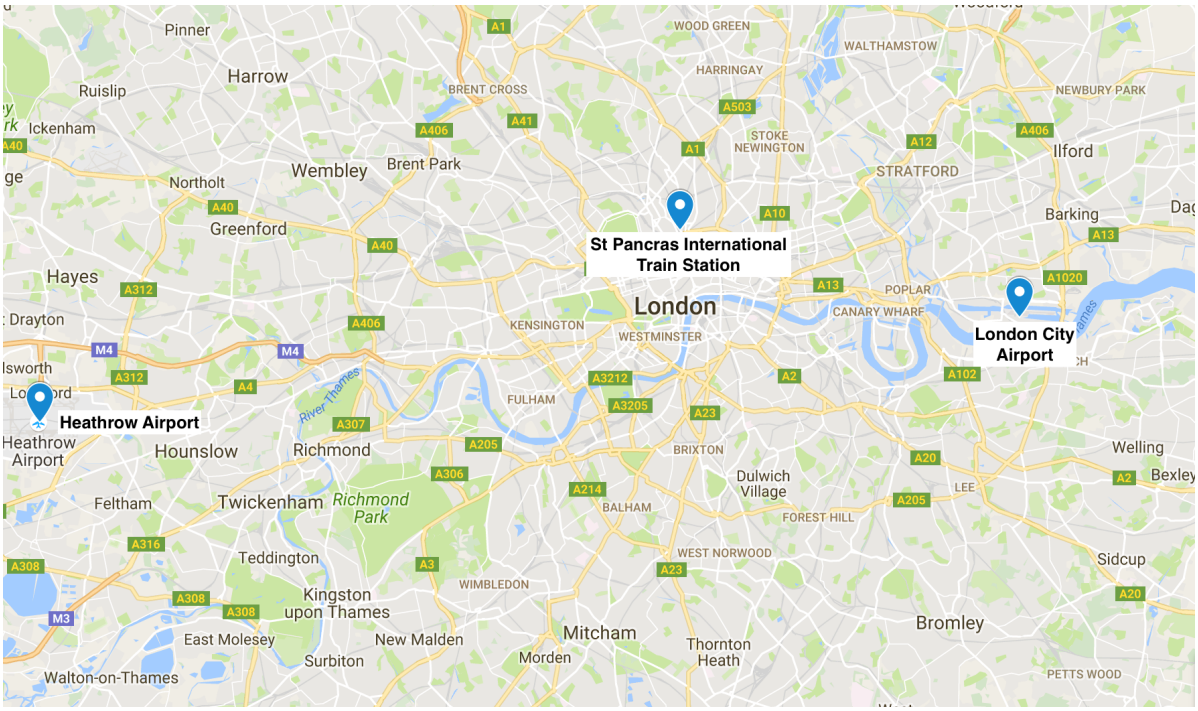


Fig. 3.5 Location of Paris/Brussels travel hubs in London (Maps, 2017)

In order to encapsulate possible origins within the GLA, population weighted centroids of LSOA zones are used as origins. There are 4835 LSOA population centroids for the study area and these are shown in Figure 3.6, where population density is shown as weighted colour per LSOA.

With 4835 origins and 3 destinations, the result is 14,505 API requests per required time slice.

Temporal specification

The defined spatial specification demands 14,505 API requests per time slice. With 100,000 requests permitted daily 6 different time slices are therefore possible. In order to inform the temporal specification of API requests, an empirical study from the literature was drawn upon. Traditionally, traffic flow has been modelled as a daily bimodal distribution, where the morning and evening rush hours exhibit higher traffic volumes, lower average speeds and higher journey times. (Mullick and Ray, 2012) presented mathematical fits of empirical data, as shown in Figure 3.7 for a road in Jackson, USA. Public transport demands follow a similar bimodal distribution.

From this idealised distribution in the literature and similar localised experiments to that of Figure 3.4, a daily schedule was designed. Over the course of 6 weeks (30/09/2015 to

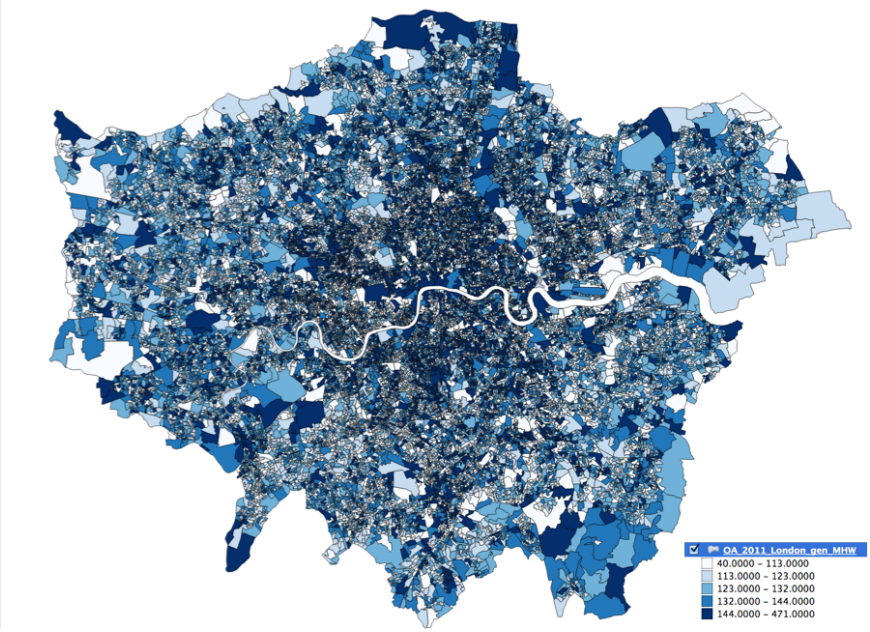


Fig. 3.6 LSOA zones for Greater London with population density scale

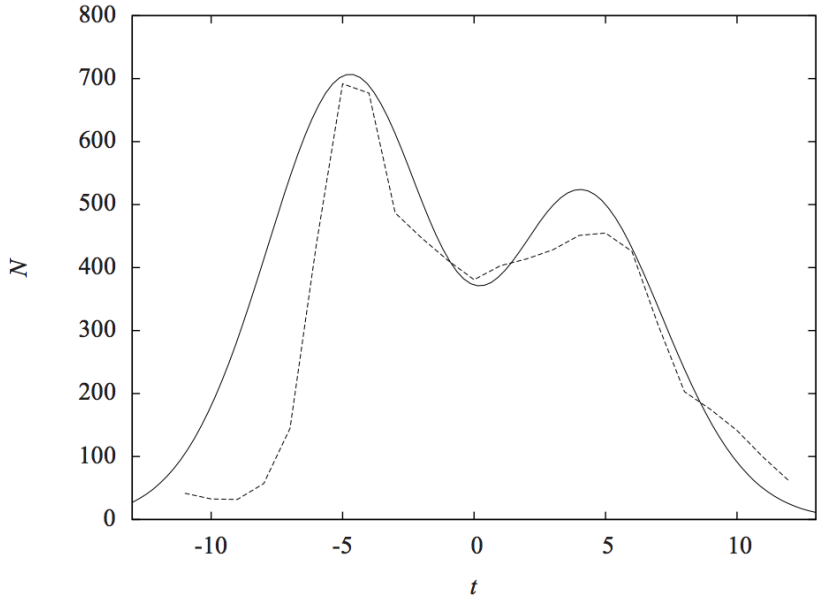


Fig. 3.7 Bimodal distribution of traffic flow due west, Jackson, Alabama State, USA (Mullick and Ray, 2012)

9/11/2015), a data set of 3,162,090 API results was generated by polling the Directions API daily at the below time specification (Table 3.2) for the previously described spatial specification. Beyond the total limit on API requests, there was also a limit on the velocity of requests - 10 requests per second. 14,505 requests requires approximately 30 minutes to send and as a result of this the below time stamps indicate the beginning of the schedule and not the absolute time stamp for a given request.

Table 3.2 Temporal resolution of daily requests

Time
06:30
08:00
11:00
13:00
15:00
17:30
19:00

3.1.3 Analysis

Some general statistics on this generated data set are shown in Table 3.3. It is now necessary to assess the spatial and the temporal extents of this generated dataset.

Table 3.3 Generated dataset summary

Data	Size
Requests	3,162,090
Bad requests	89
Total number of legs	64,528,506
File size	43.18GB uncompressed
File size	4.81GB compressed

Spatial coverage

In the previous section, the population weighted centroids of LSOA areas were taken as origins and the three modal destinations, St Pancras International Train Station, Heathrow Airport and London City Airport were taken as destinations. The intention was to capture location device informed journey times for as many roads as possible within the study area and within the constraints of the API quota.

The polyline for each API response may be parsed and visualised simplistically in its own right, in order to visualise the spatial coverage generated by this schedule. This is presented in Figure 3.8. Routes which feature frequently are weighted and the density may be visually inspected. It is visually evident that good coverage has been achieved as the API responses present a realistic visualisation of the road network.

Fig. 3.8 Spatial coverage of generated Google Directions API dataset

Temporal coverage

Beyond the spatial resolution of the generated data, there is also the temporal resolution which captures the dynamic changes in journey times as a result of changing road conditions. This may be considered in terms of global average speed distributions, journey time distributions and specific edge journey time distributions, for a range of different temporal bounds.

Average speed distribution

A macro average speed comparison may be computed for all journeys, for a given hour, across the entire 6 week sample period. This is graphically shown in Figure 3.8. It is clear that globally, average speeds are lower in the morning rush hour and afternoon periods than they are during the day, early morning and later at night. Comparing these periods shows an average speed difference of nearly 20 *mph* between 5pm and 9pm for the routes in this dataset. The lowest global average speeds are exhibited at 9am and 5pm. For reference the slowest exhibited, individual journey average speed of the entire dataset was found to be 2.11*mph* at 12:04pm and conversely, the fasted exhibited individual journey average speed was found to be 63.54*mph* at

06:55am. In the context of limitations on the Directions API it is interesting to note that no record ever presented a location device informed journey time which exceeded the legal speed limit. It may be assumed that despite likely location device inputs in excess of the speed limit, Google's algorithms filter such results and present an upper bound constraint in API responses.

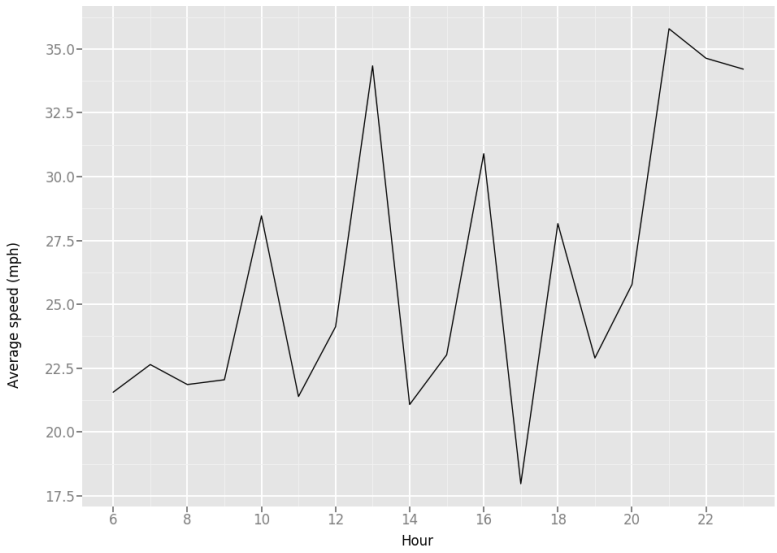


Fig. 3.9 Average speed per time of day

The average travel speeds on the road network shown in Figure 3.9 illustrate a global representation, which negates the context specific nature of a given route/road or even distinct days of the week or weeks in the month. Each section of road may have a unique road type, road nature, lane count, traffic furniture or specific geometry; factors which all impact journey times. These factors may explain why the distribution shown in Figure 3.8 is not a bimodal distribution, which is often used to model daily journey time distributions (for example (Mullick and Ray, 2012)).

Journey time distribution

A journey time distribution plot may be created for each individual origin and destination pair across the entire study period. Since the programmatic method explained previously generates so many of these records, an unmanageable amount of these plots may be generated. Therefore, the first exploratory process involved the selection of plots at random for inspection and from these, 4 broad types were extracted. The first type illustrates those with morning dominated increased journey times, the second for evening dominated, the third for bimodal

and the fourth for those exhibiting no obvious pattern. For clarity, a facet plot has been used to present the entire study period in distinct weekly sub plots.

In Figure 3.10 the weekly journey time plots for route, over a series of weeks, to Heathrow is presented. On initial inspection it is clear that there are significant daily fluctuations on this route, with the highest journey times generally falling in the morning rush hour period. Figure 3.12 presents a more traditional, bimodal journey time distribution. The daily morning and evening peaks are exhibited across most week days, on each week. Figure 3.12 presents a more traditional, bimodal journey time distribution. The daily morning and evening peaks are exhibited across most week days, on each week. In Figure 3.13 a London City journey is shown to exhibit chaotic journey times.

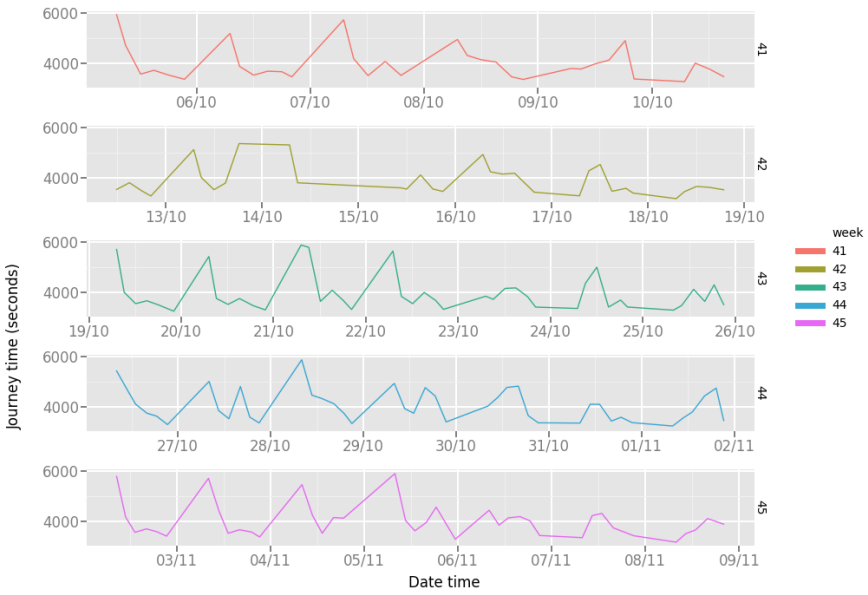


Fig. 3.10 LSOA E01003681 to Heathrow Airport journey time distribution (weeks 41 - 45, 2015)

It is clear from these randomly selected plots that the assumption of the existence of a typical day or even week is challenging. Journey times are shown to fluctuate considerably in response to local conditions, with some examples illustrating the traditional bi-modal loading and others illustrating morning or evening predominant loadings.

Statistical exploration

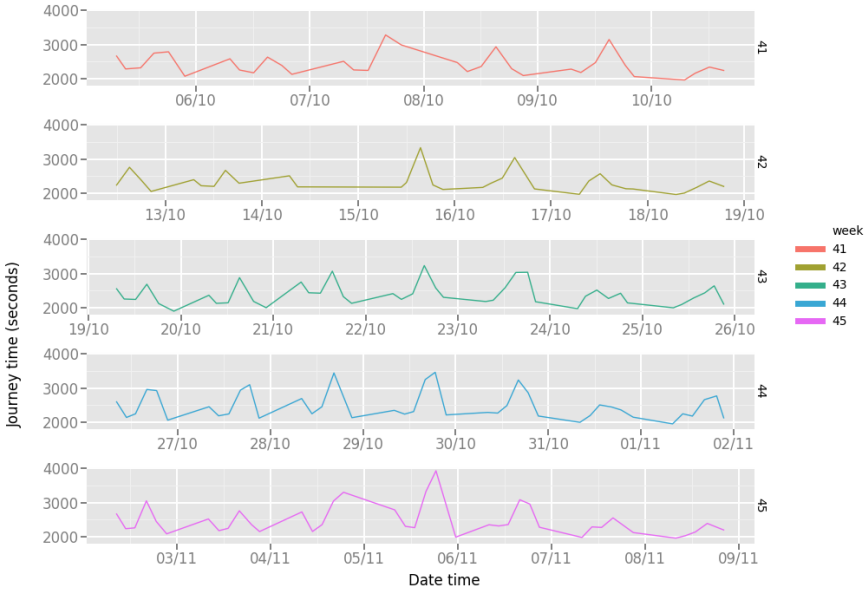


Fig. 3.11 LSOA E01004662 to Heathrow Airport journey time distribution weeks (weeks 41 - 45, 2015)

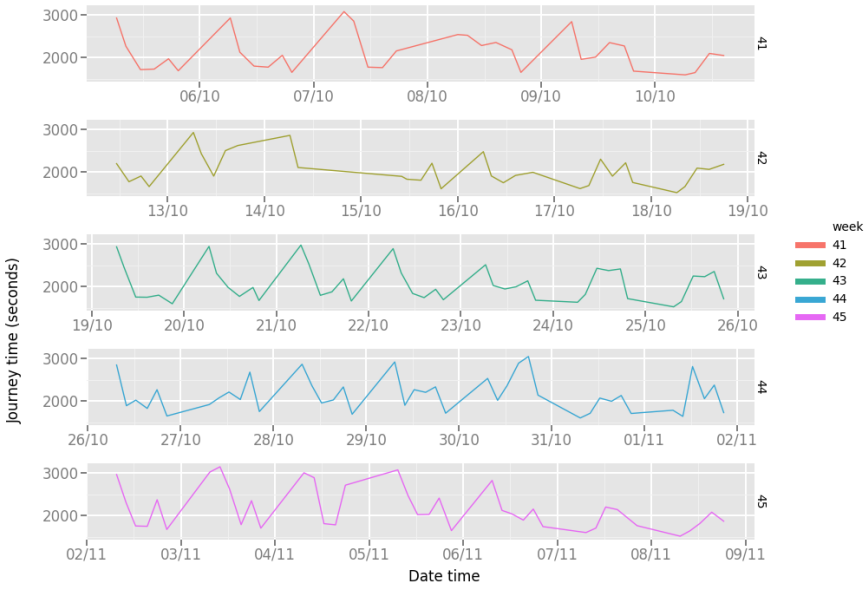


Fig. 3.12 LSOA E01000345 to London City Airport journey time distribution (weeks 41 - 45, 2015)

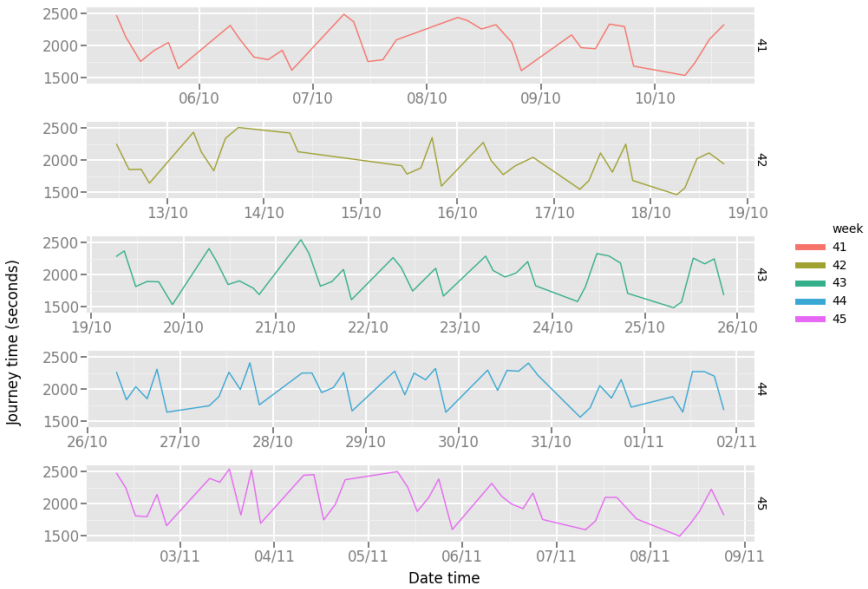


Fig. 3.13 LSOA E01001685 to London City journey time distribution (weeks 41 - 45, 2015)

The selection of these example plots was carried out in an arbitrary way and there is a requirement for statistical analysis in order to explore what may be understood of the underlying dynamics. For the range of time stamped journeys the following metrics were computed:

1. Mean journey time, speed and distance
2. Ratio of maximum to minimum journey time
3. Standard deviation & variance of journey time

1. Mean journey time, speed and distance

The mean journey time presents a crude comparative of average journey times, per hub. In Table 3.4 we can see that Heathrow Airport exhibits the largest average journey times whereas London City and St Pancras have much similar average journey times, consistent with their more central location in relation to the LSOA weighted origins. Heathrow's average journey time benefits from an increased average speed, which reduces the impact of the considerably larger average journey distance from the LSOA weighted origins.

This may be disaggregated further and the average impact of traffic at different times of the day may be extracted. In Figure 3.14 we can see the mean journey time per hub, disaggregated

Table 3.4 Mean journey times, speeds and distances per hub - driving

Hub	mean journey time	mean journey speed	mean journey distance
Heathrow Airport	3186 seconds	15.92 m/s	51.04 km
St Pancras International	2662 seconds	5.91 m/s	16.23 km
London City Airport	2793 seconds	9.25 m/s	25.83 km

by hour of the day. Each of the hubs exhibits a morning and evening peak, with the evening peak more pronounced for London City and Heathrow Airports. At 6pm, a sharp drop in journey times is globally exhibited.

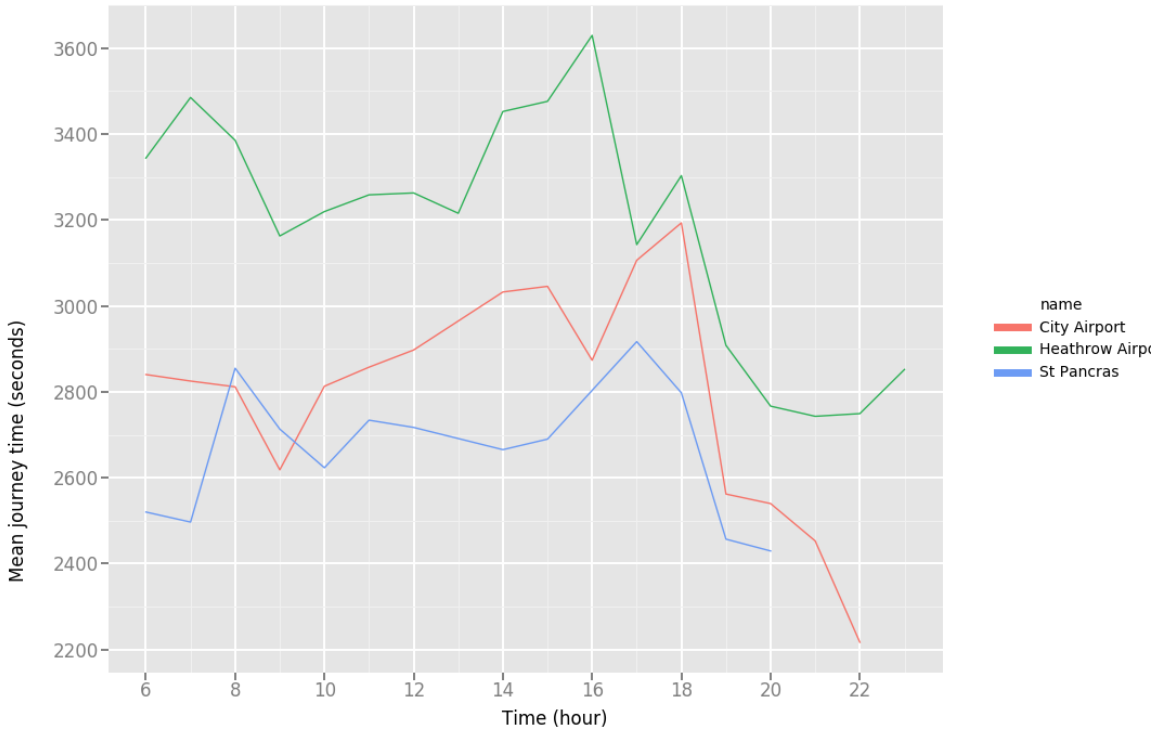


Fig. 3.14 Hourly mean journey time for each hub

2. Ratio of maximum to minimum journey time

The ratio of maximum to minimum journey times, per hub illustrates the full range of exhibited journey times for a given origin and destination (hub) and this may be viewed as a density plot, disaggregated by hub, as is shown in Figure 3.15.

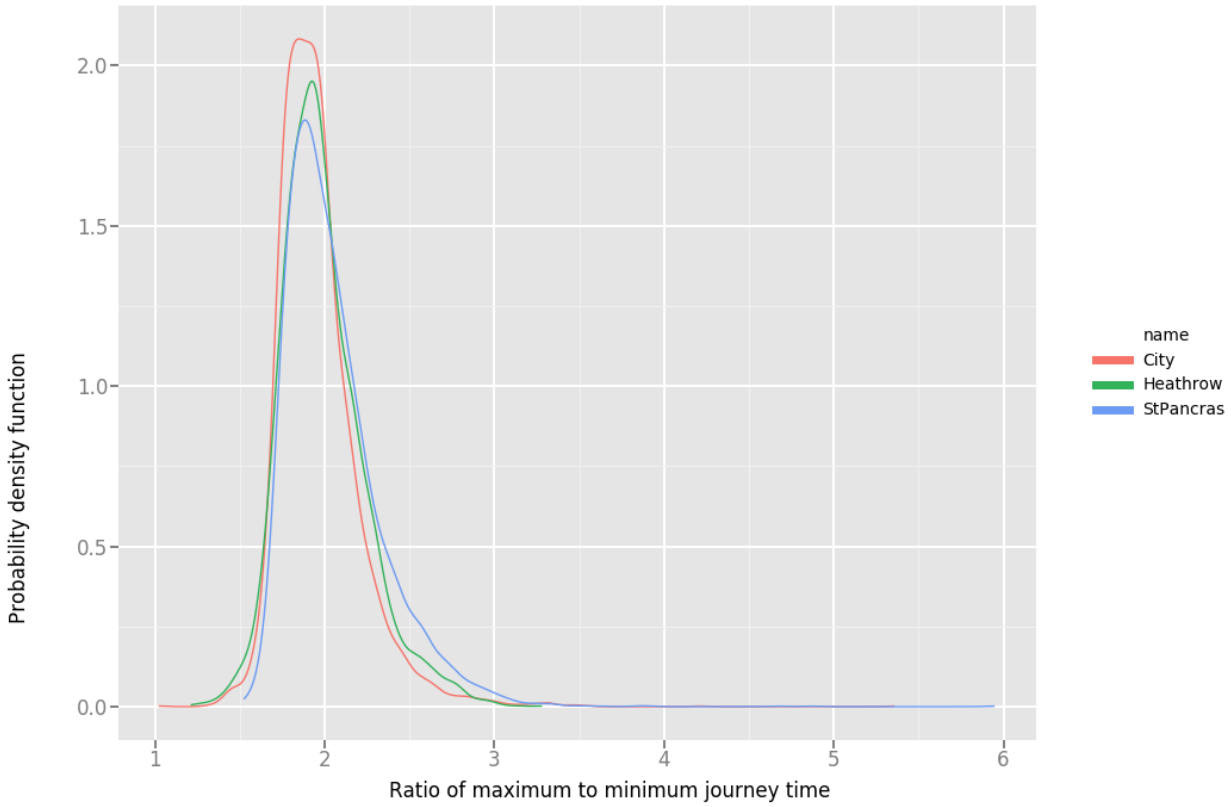


Fig. 3.15 Ratio of maximum to minimum journey times, for each mode

It is clear that the majority of journey times fall between 1.5 and 3 fold increases relative to the minimum journey time. However, the 3+ fold increases make it difficult to inspect the differences between the different hubs .Small differences can be seen between the different hubs, with St Pancras exhibiting the largest range of values, indicating that journeys are generally more variable for this hub when compared to London City, which exhibits the narrowest range. The distribution also shows a long tail, comprising of very large journey time increases that occur rarely. Such records usually indicate a disruptive event such as a road traffic collision or unplanned road works.

Figure 3.16 presents responses which exhibit a sudden journey time increase. This plot presents normalised journey times for a week. A sudden event can be seen to occur on the Wednesday morning, with journey times peaking at nearly a six fold increase. This journey time increase continues for the remainder of the week. A search of local media found a burst pipe (Figure 3.17) on a road in close proximity to this particular road, impacting significantly on journey times during the event and the subsequent repairs.

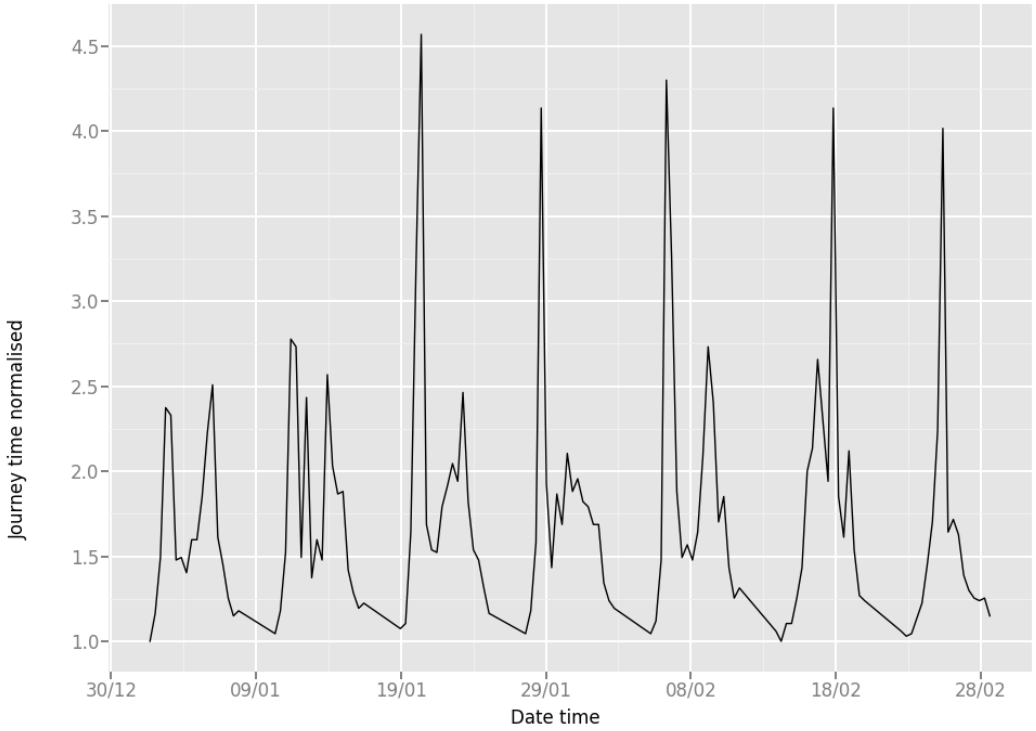


Fig. 3.16 Sample journey time plot with disruptive event

Fig. 3.17 Disruptive event cause

3. Standard deviation & variance of journey time and journey speed

We have considered average journey times and their maximum and minimum ranges, per hub and over different time slices. The standard deviation, defined as the spread, or amount of variation around the mean of the data gives a clearer insight into the fluctuations that can be visually inspected in between relatively short time frames, across the entire temporal range of the data. From the standard deviation the variance may be computed and the variance enables the comparison of data sets with underlying differences (e.g. different journey lengths) as it considers the mean of the data.

The variance may therefore be computed per unique journey, for journey time (Figure 3.18) and speed (Figure 3.19), and the collective variances may be visualised as a probability density function, with disaggregation by hub for comparisons. These comparisons are made for the 95% percentile in order to remove extreme cases and for more clarity between modes.

The centroid of the LSOA's has a lesser effect on the journey time variance for St Pancras and London City. Conversely, Heathrow's Western location results in a West weighted situation, which is not counteracted as a result of the origins and destinations specified by GLA LSOA centroids.

Propensity for diversion

A static origin and destination pair do not always output the same path at all times. In some cases, there may be multiple possible routes with similarly competitive journey times depending on traffic, road works and other possible factors. By considering the polyline for a given response from the Directions API we can assess the propensity for diversion for different destination hubs. Figure 3.20 presents the probability distribution function for the number of unique polylines (routes) disaggregated by hub. Heathrow is shown to exhibit the lowest mean number of unique polylines and St Pancras is shown to exhibit the highest mean number of polylines, suggesting more central hub destinations illustrate more relative competition between different routes.

In Figure 3.21 the number of unique polylines is plotted against the direct distance (crow flies) between the origin and destination hub. A general positive trend is shown with the further the journey, the more likely it is for a diversion.

Link time distribution

The sections up to this point have focussed on the outputs related to a fixed origin and destination pair. However these journey outputs do no necessarily involves the same pieces

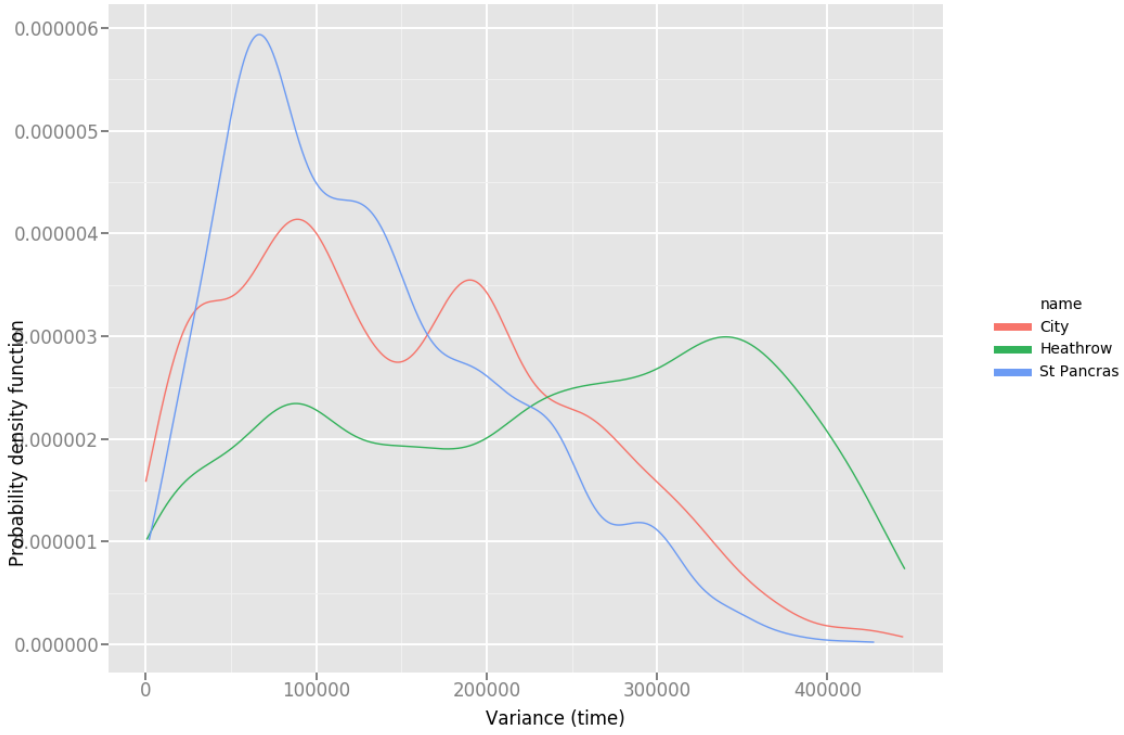


Fig. 3.18 Journey time variance density plot (0.95 percentile)

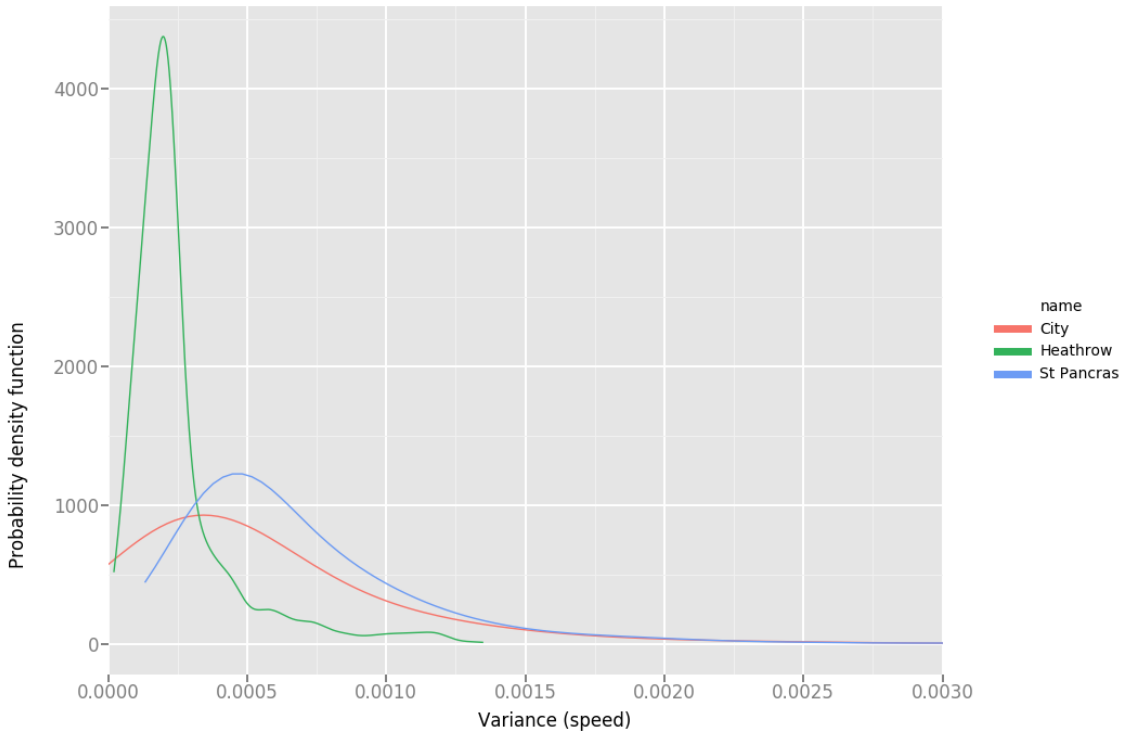


Fig. 3.19 Journey speed variance density plot (0.95 percentile)

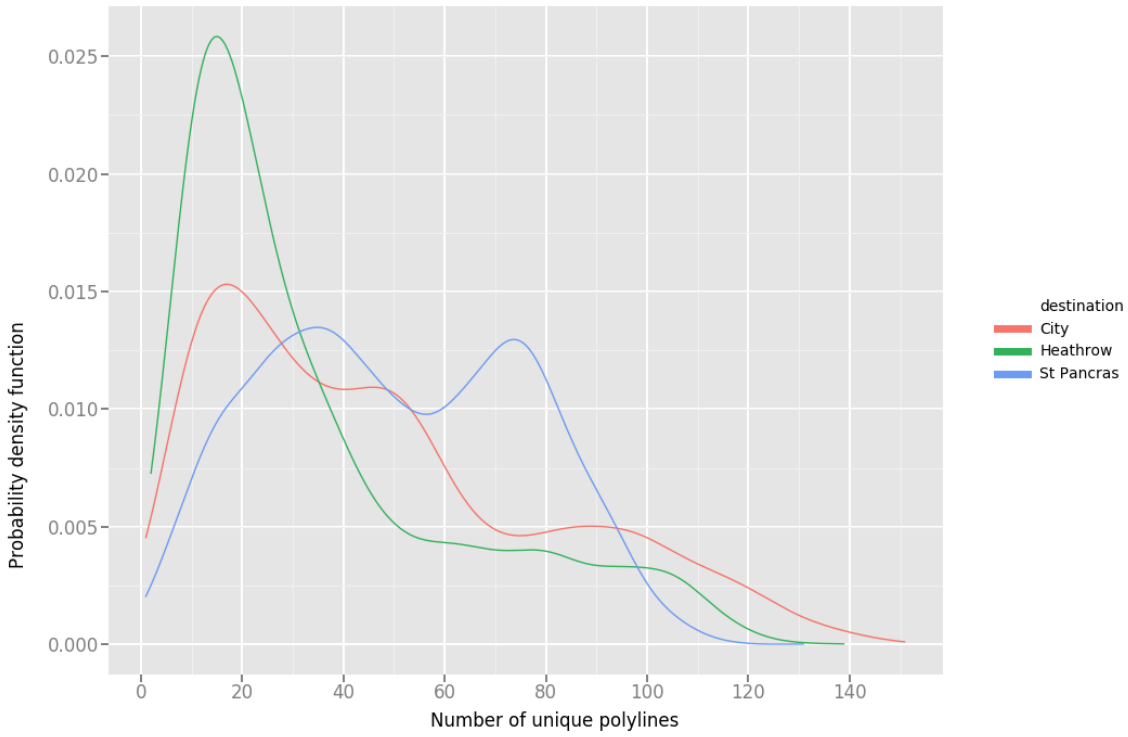


Fig. 3.20 Probability density function for number of unique polylines, for each hub, by driving

Fig. 3.21 Scatter plot for number of unique polylines against euclidean distance between origin and destination, for each hub, by driving

of physical infrastructure. As was shown in Figure 3.2 and illustrated in Table 3.1, each API response contains a series of legs, each with the key journey time information and the polyline of that particular part of the journey. This polyline either represents a road edge or a series of road edges. This polyline may be first used to assess the propensity for route change on a static origin and destination journey and secondly, as a more indicative comparison between journey times.

In order to assess how journey times change for the same piece(s) of infrastructure, the polyline may be taken as a unique key and journey times for different time stamps compared. The generated data set consists of 64,528,506 individual legs and so an efficient means of identifying unique polylines with an array of time stamped journey times is required. The raw polyline, consisting of a series of coordinates is compressed using a lossy compression algorithm producing a single string representation (Google, 2017). To decrease space even more, Google offset from the previous point, rather than constructing from scratch. Thus, before unique polylines may be identified, each leg of an API response must be re-factored to its true representation and not its relative representation within a given API response. When this is complete, each leg may be treated in isolation.

After this re-factoring process the polyline may be re-encoded to firstly, save space and secondly, for efficient manipulation. A hash function may be used on the polyline string in order to efficiently pair exact matches from the large list of polylines. This is a memory intensive problem as all of these polylines must be held in memory whilst hashed, compared and paired. The outcome is the pairing of common polylines with a range of time stamped attributes in a unique array of the format:

```
{u'attributes': [
        {u'destination': [51.5011926, -0.1239613],
        u'distance': 2232,
        u'duration': 420,
        u'leg_travel_mode': u'DRIVING',
        u'origin': [51.511017, -0.1018502],
        u'time_stamp': u'1445267285'},
        {u'destination': [51.5011926, -0.1239613],
        u'distance': 2232,
        u'duration': 417,
        u'leg_travel_mode': u'DRIVING',
        u'origin': [51.511017, -0.1018502],
        u'time_stamp': u'1445238962'},
```

```
        {u'destination': [51.5011926, -0.1239613],
        u'distance': 2232,
        u'duration': 420,
        u'leg_travel_mode': u'DRIVING',
        u'origin': [51.511017, -0.1018502],
        u'time_stamp': u'1445759913'},
        {u'destination': [51.5011926, -0.1239613],
        u'distance': 2232,
        u'duration': 417,
        u'leg_travel_mode': u'DRIVING',
        u'origin': [51.511017, -0.1018502],
        u'time_stamp': u'1444805598'}],
u'polyline': u'yifm{‘bBelxiyt‘@}gksAo|xPnv‘yI~powvsG'}}
```


This output presents a unique polyline with various different journey times and respective time stamps. As was previously shown, these individual durations do not actually sum to make the key *duration_in_traffic* record for each API response. However, this longitudinal analysis has shown that these more static, average metrics of average condition journey times do in fact change with time over the 6 week period of this study. This can be seen in the example above, where journey times changed by a few seconds from 417 seconds on the 14th October at 7:53am, to 420 seconds on the 25th October at 07:58am.

3.1.4 Conclusions

This section has presented two distinct different methods of quantifying road journey times at road edge and journey level, across a range of different temporal boundaries. It has been illustrated that there are distinct differences between the more static, average measurement of journey time provided by Google, the *duration* record which is presented at fine spatial resolution (sub journey level) and the *duration_in_traffic* record which presents finer temporal resolution data but at the compromise of sub journey level spatial resolution. There is therefore a trade-off between spatial and temporal resolution when attributing macro journey times to the micro level sections of that journey.

Despite illustrating believable journey time distributions, often in line with models from the literature (e.g. bimodal distributions) the process which produces these metrics is technical opaque. Google's methodology for computing this shortest path is not publicly available for commercial reasons. The sample size of location device coordinates which contribute to a given API response is also unknown as Google will not provide this information citing privacy

concerns. However, the specification of a static origin and destination with no route-finding requirement, across short distances and for a range of time stamps returns temporally dynamic journey times.

3.1.5 Summary

1. Hypothesis

   The Google Directions API may be used to quantify road conditions considering traffic congestion across a large spatial area at fine temporal resolution. By scheduling API requests in a targeted way, the programmatic nature of the service can be levered and specific questions about particular origins and destinations may be answered.

2. Novel contributions

   The method developed here has been shown to provide fine resolution temporal data across a wide spatial resolution. The outputted data may be used to quantify the dynamic nature of road journey times. This data may be used to categorise different transportation hubs, such as airports and train stations in terms of their relative access at different times of the day, on different days and on different weeks. However, due to a lack of comparable resolution data there is a validation challenge. Generally, we can see that the trends exhibited on many roads fall in line with general understandings surround peak and off-peak flows towards and away from large trip attractors (i.e. morning peaks into London and evening peaks out of London).

3.2 Crowd sourced journey times and automated traffic counter volume-delay functions for London

When the vehicular demand for a road exceeds the ability to supply, a journey time-delay is incurred as a result of the congestion generated. This relationship of traffic volume to time-delay has historically been simplified to macroscopic principles. The individual interactions of increasing and decreasing vehicle volumes result in changes to the journey time on any given link. As traffic volume and therefore density increase on a fixed length link vehicle speed will reduce in order to safely manage the increased volume. This research investigates the pairing of two data sources; crowd sourced location device informed journey times and traffic count data from the Automated Traffic Counter (ATC) system in the Greater London Area (GLA). This section will begin by reviewing the literature on the methods currently employed. A new method is then proposed which aims to satisfy some of the identified limitations in the current methods by using fine resolution data in the form of automated traffic counters and location device informed journey times on a range of roads within the GLA.

3.2.1 Literature review

There is a range of macroscopic methods used to relate traffic volume to time-delay.

Bureau of Public Roads method

The Bureau of Public Roads (BPR) volume-delay curve was developed in the 1960's and is the most widely used function for relating vehicle volume and journey time. Its simple mathematical form and input requirements are attributed to its widespread use (Skabardonis and Dowling, 1997). There are many different forms of the BPR function as many different organisations have adapted it with various local empirical and/or hypothetical data (Mtoi and Moses, 2014) in an attempt to localise the function.

The general BPR function is mathematically defined as:

$$u = \frac{u_0}{(1+\alpha(x)^{\beta})} \tag{3.1}$$

Where:

u = speed

u_0 = free flow speed

x = saturation (volume/capacity)

α and β = coefficients

In the case of London, TfL have calibrated the BPR function with exhibited traffic counts and thus defined α and β values for the area (TfL, 2010).

$\alpha = 1.0$

$\beta = 2.0$

Despite the popularity of the BPR function it has many limitations, primarily as it was created by fitting a polynomial equation to uncongested freeway data from the 1950's and thus does not reflect modern operating conditions (Skabardonis and Dowling, 1997).

Davidson method

Davidson (1966) proposed a general purpose travel-time formula in 1966 and this method has undergone numerous modifications since it was first proposed (Mtoi and Moses, 2014). Its popularity is as a result of its flexibility and applicability for a range of contexts (Taylor, 1997). It has exhibited a closer match to actual volume counts and has a stronger theoretical base than the BPR (Rose et al., 1989). There have been modifications of the Davidson function since it was first proposed (e.g (Tisato, 1991), (Akcelik, 1991)) with the mostly widely used being the Akcçlik functions.

Akcçlik method

The Akcçlik method is a time-dependent modification of the Davidson model which uses the coordinate transformation technique in an attempt to overcome the conceptual and calibration issues with the Davidson method (Akcelik, 1991). It has illustrated good results in certain road types, tolls roads and signalized arterials (Mtoi and Moses, 2014).

Conical method

Spiess (Spiess, 1990) proposed the conical method in 1990. It attempted to overcome some of the upper bound BPR limitations through the use of hyperbolic conical sections whilst maintaining a similar form to the BPR and thus enabled a direct transfer of parameters.

Identified limitations

These idealised functions attempt to generalise different aspects of an individual road's characteristics in order to create usable functions that do not require a large range of survey requiring inputs. The general nature of these functions enables their ease of use and gives widespread

applicability. However, in doing so their ability to give results that consider the individual characteristics of a road (i.e. geometry, the presence of traffic furniture, road quality and the surrounding land use) reduces. Such differing characteristics can result in very different vehicular behaviour on roads that may be considered similar by these functions. The literature has identified the need to inform these functions with empirical and context specific data (Rose et al., 1989), (Spiess, 1990) but recognised the difficulty and cost associated with collecting such empirical data as being prohibitive (Rose et al., 1989).

Input data

There have been studies that incorporate different forms of field sensor data (Mtoi and Moses, 2014), (Neuhold and Fellendorf, 2014) in volume delay functions. However, the data used in these studies has been generated specifically for that application and requires specific hardware and/or software for use. Recent hardware and software innovations have led to the increased use of real-time crowd-sourced data feeds generally in transport modelling. This includes applications of location data for emissions estimations (Hirschmann et al., 2010), origin and destination matrices (Toole et al., 2015) and general urban traffic management applications (Artikis et al., 2014). This research investigates the use of new, real-time crowd-sourced data feeds that have a wider spatial spread and are in some cases not generated specifically for this application. These sources may be used to consider some of these previously ignored characteristics and create temporally and spatially dynamic volume, speed and saturation relationships. Such data sources can harvest data at a finer resolution over a longer (even indefinite) period of time giving a far greater understanding of the temporal variations and trends exhibited on road infrastructure.

3.2.2 Automated traffic counters

Automated Traffic Counters (ATCs) are magnetic induction loops located in the road surface. The passing of a vehicle results in an electromagnetic signal. The ATCs in the Greater London Area count every vehicle which passes over the inductions loop. The data used here was harvested over a period from the 29th February to the 21st March 2016.

Location of ATCs in Greater London

There are 37 Department for Transport (DfT) ATC locations in the Greater London Area. Most of these locations operate in both directions (34) with others operating only in one direction (3). These ATC locations cover a range of different DfT defined road classes (Table 3.5). Guidance on the road classification system in the UK is published by the DfT (DfT, 2012a).

Table 3.5 ATC locations by road class

Road Class	Count
Trunk	6
Principal	16
B	3
C	3
Unclassified	9

These ATCs are distributed across the Greater London Area as is shown in Figure 3.22.

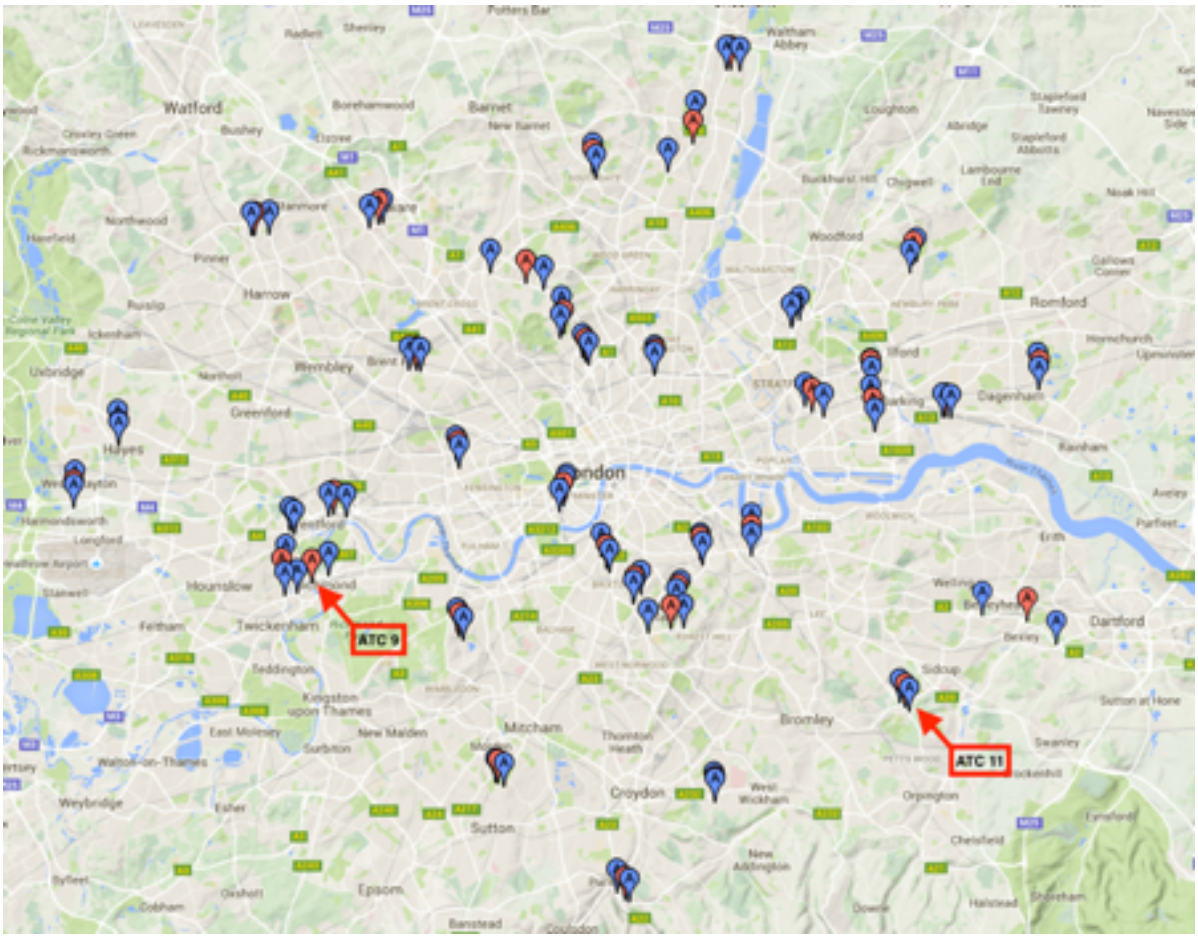


Fig. 3.22 ATC Locations in Greater London Area (Maps, 2017) with DfT data. The red flag illustrates the location of the ATC itself. The blue flags illustrate the origin and destination locations specified in order to harvest journey time information

Raw ATC data

The raw ATC data contains individual records for each vehicle that passed. Over the test period of 3 weeks there were approximately 4.5million records of vehicles at individual points. For illustration, Table 3.6 shows a representative sample from the first day of these records at ATC 11 (as labelled in Figure 3.22).

Table 3.6 Raw ATC data record sample

Site	Direction	Date	Time	Speed
11	Northbound	2016-02-27	00:00:28	37
11	Northbound	2016-02-27	00:01:19	40
11	Northbound	2016-02-27	00:02:25	42

Manipulated ATC data

In order to combine this data with the temporal journey times, individual vehicle records must be combined hourly in order to find the total volume on that link, per hour. A Python (van Rossum and Drake, 2002) script using the Pandas (Pandas, 2016) data library was used for this processing. First, a unique identifier was formed by concatenating the site and directionality of the ATC. The time stamp was rounded up to the next hour in order to quantify the total volume per hour leading up to the journey time taken, up to that hour. This resulted in an output which featured traffic counts per hour (volume) for a site, direction and date, as is sampled and shown in Table 3.7.

Table 3.7 Processed ATC data record sample

Site	Direction	Date	Hour	Count	id
9	Southbound	2016-02-27	13	1206	9S
9	Southbound	2016-02-27	14	1222	9S
9	Southbound	2016-02-27	15	1408	9S
9	Southbound	2016-02-27	16	1604	9S

3.2.3 Crowd sourced journey times

Location data

Location capable mobile phones have enabled the harvesting of fine resolution temporal and spatial data. The location of a device may be inferred from the use of GPS, cell tower

triangulation, WiFi SSID mapping, Bluetooth and other technologies, either in isolation or in tandem. Such data holds a great deal of promise due to the range of possible uses it has in the transportation sector (Zheng et al., 2010). The anonymous crowd-sourced collection of this data can be used to derive road traffic conditions (Barth, 2009). Such a traffic condition method is well suited to areas of high density, high travel demand and high mobile phone uptake, such as a city. This information is used by many providers who provide real-time traffic informed shortest path directions as a service. For example Apple (*apple.com/ios/maps*), Bing (*bing.com/map*), Google (*google.com/maps*) and TomTom (*tomtom.com*). These services are generally used by individuals who wish to make an informed decision on route choice for a given mode or even a mode decision on how to get from their starting point to a desired destination. The aim of this research was not to make use of the shortest path algorithm or large-scale computational power of these service providers, but to access a more localised, location device informed, journey time on a fixed and relatively short stretch of road.

Google Maps

This research makes uses of Google's traffic information through the Google Maps Directions (Google, 2016a) Application Protocol Interface (API). Dependent on their personal settings, an Android user or other Operating System (OS) user, with the free Google Maps mobile app installed on their location enabled phone send anonymous data to Google. Such data is personally and commercially sensitive and so post-processing is carried out by Google in order to ensure that no-one user's movements can be isolated from the flows.

Google's traffic maps can be visually inspected through their maps interface on any modern web browser on google.co.uk/maps. Their colour coded scale gives a qualitative feel for traffic conditions on the roads where Google have sufficient data, as is shown in Figure 3.23. Such qualitative data is only useful for contrasting different traffic states on a road.

Google Maps Directions API

The Directions API is a service that calculates directions between locations using a Hypertext Transfer Protocol (HTTP) request (Google, 2016a). The use of a HTTP request allows for scheduled and bulk harvesting of journey information between given origin and destination pairs. The aim of this research was to combine crowd-sourced location device informed journey times with traffic counts from the DfT ATC network. As such, the first step was to specify origins and destinations for the Google Directions HTTP request that would provide a journey time for the traffic counts at defined ATC locations. There is a need to convert from EPSG:27700 (British National Grid), as provided by the DfT for ATC locations, to the

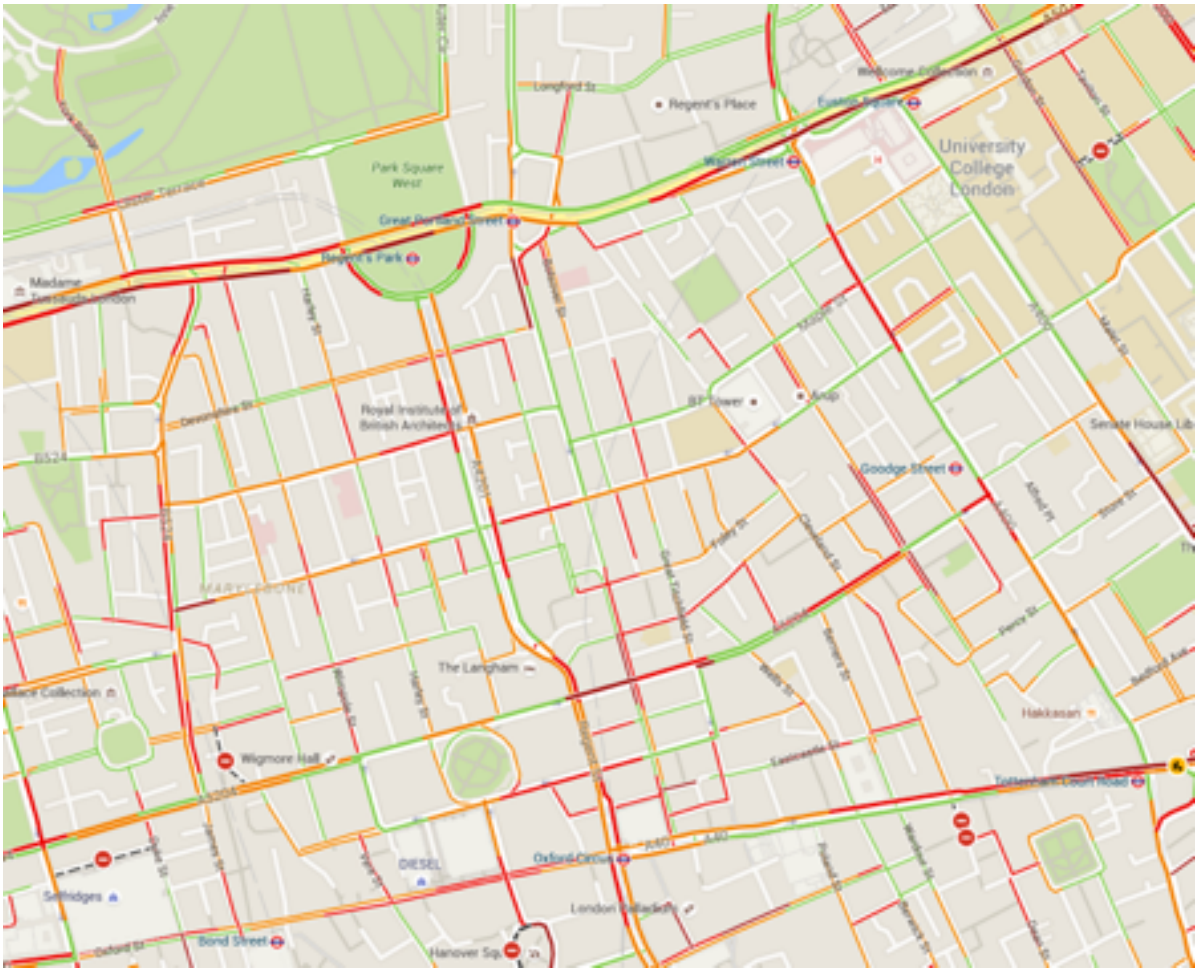


Fig. 3.23 Google Maps Traffic Layer, Camden/Soho/Marylebone/Mayfair area of London (Maps, 2017)

EPSG:4326 (WGS84) as used by Google Maps. The ATCs take a point measurement at a defined location along the road link. In order to find a journey time along that road, an origin and destination location definition process is required. In deciding on origin and destination pairs a balance must be made between having sufficient distance in order to get meaningful journey time results and having sufficiently short distance so as to not include other phenomena such as junctions and intersections. Consider ATC location 6 Eastbound, the A205 Dulwich Common SE21 in the Borough of Southwark. In Figure 3.24 the location and defined origin and destination points can be inspected. Automating the definition of these origin and destination pairs was challenging as simply taking the start/end of a given road often gave a route that was too long and thus was distorted by other traffic, outside of the ATC consideration. Other methods based upon an idealised distance between points and the density of junctions was deemed too complex and not durable. For bidirectional ATC locations it was often not possible to define the Eastbound route as the inverse of the Westbound route as Google distinguish between different sides of the road, resulting in a route which involves a detour to safely navigate to the correct orientation. Thus a manual process was employed to visually inspect each location, the surrounding context and decide on the most appropriate O/D pair. Once this manual process was complete, a matrix of origin and destination pairs was produced containing ATC metadata that will allow for the pairing of Google results to its corresponding ATC data.

Fig. 3.24 ATC 6 Eastbound with defined origin and destination points (Maps, 2017)

Sending HTTP request to Google

A HTTP request is made in Python to Google's servers with the origin, destination, mode (driving) and the specified departure time. In order to harvest real-time data that is informed by location device information at that time the departure time is set to the live time and a scheduler is used to run the same origin and destination pairs repeatedly. Cron, a time-based scheduler

for Unix operating systems, was used to run the Python HTTP request every hour for the given dates of the study.

Receiving JSON response from Google

In response to the HTTP request Google returns results in JavaScript Object Notation (JSON, see *json.org*), a lightweight data-interchange format. At the time of writing the JSON response contains 3 root elements, the status, geo_coded waypoints and the routes. Within routes the key metric of duration_in_traffic for the requested route is found, stating the estimated journey time on the route, at that time. The entire JSON response is parsed and transferred to a MySQL (*mysql.com*) relational database for temporary storage. From the database, the origin, destination, ATC related data and the key duration_in_traffic metric can be paired with the ATC counts.

Data cleaning

Prior to combining the two data sources the Google data was sense checked. A small number of Google results returned no duration_in_traffic field. These results were likely due to low density of location device data as a result of low user demand or poor location device signal reception in that location. This included three ATC pairs (ATC numbers 42, 48 and 74 in both road directions) and one ATC (36N) in one direction. They were discarded so as not to skew the results. It was also noted that at times of low demand (late night and early morning) the duration_in_traffic response did not return and it can therefore be deduced that duration_in_traffic was only returned when there is sufficient spatial and temporal resolution location device input data.

Data validation

Google's methodology nor location device sample information is available for commercial and privacy reasons. Since there are no comparable, freely available data sources, the Google API responses pose a validation challenge. It is therefore an assumption that these responses present location device informed journey times.

3.2.4 Analysis

The combination of Google data and DfT ATC vehicle count data enables various relationships to be assessed and defined. In this section a number of these relationships are discussed. For the purposes of comparison, the widely used BPR functions are shown in tandem.

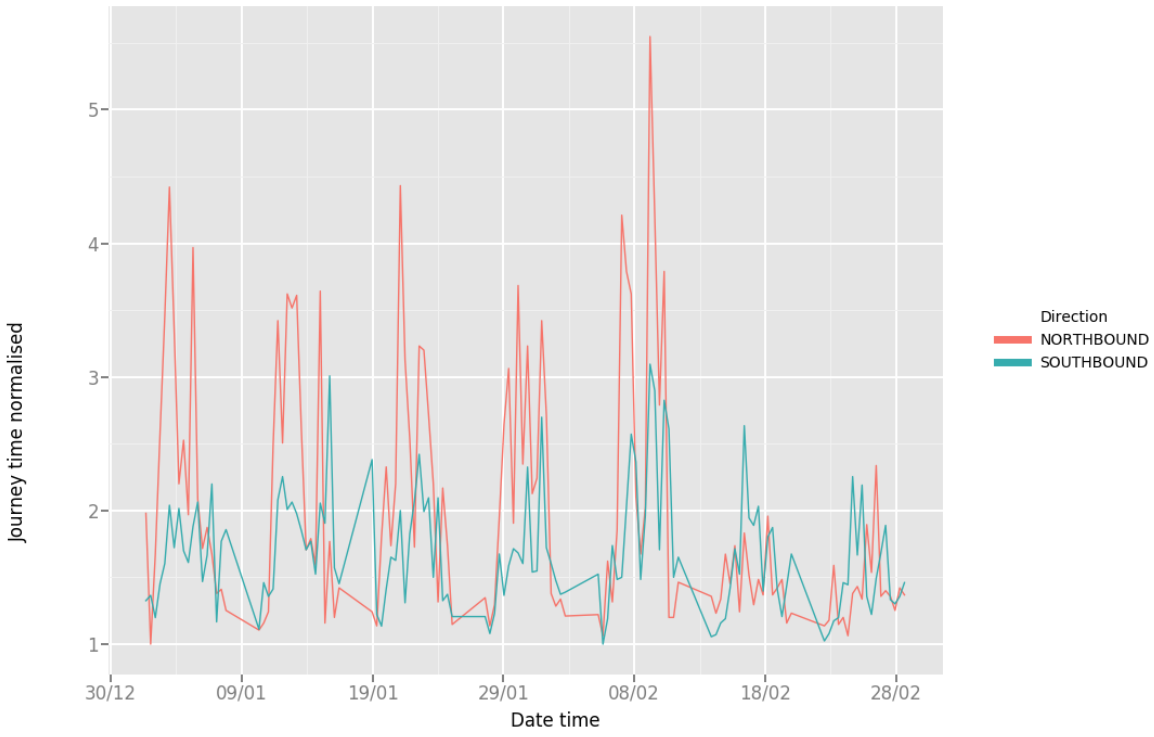


Fig. 3.25 Journey time distribution Location 67 (7th-13th March 2016)

Journey time distribution

The Google data was analysed to assess the vehicle traffic distribution pattern. Vehicle traffic distribution is generally modelled as a daily bimodal distribution with peaks associated with the morning (to work) and evening (leaving work) periods (Mullick and Ray, 2012). It is generally expected that the load will lean on the city inbound direction in the morning and then on the city outbound direction in the evening. Figure 3.25 displays a one-week sample of ATC location 67 in both the northbound and southbound directions. Location 67 is situated on Buckingham Palace Road, A3214 with the northbound direction leading to central London and the southbound direction heading away from central London. In order to make journey time comparisons, the journey time delay is normalised.

The normalised time delay is defined as:

$$timedelay_{norm} = \frac{Journeytime}{Journeytime_0} \tag{3.2}$$

Journey $time_0$ is taken as the minimum exhibited journey time on the road link. Journey $time_0$ is also referred to as the free flow journey time and is usually found in the early hours of the morning. This empirical method deviates from the standard method of defining the free flow journey time which generally relies on using time delay coefficients for different road types, speed limits, link lengths, widths, gradients, traffic junctions and so on (DfT, 2002).

In this study, it is recognised that such an assumption negates the impact of context specific information such as road geometry or road furniture. Rather than making assumptions about the free flow journey time it was taken as being the lowest exhibited time from the Google journey time data over the period of the study.

In Figure 3.25 the x-axis displays normalised journey times as a ratio of the free flow journey time (t/t_0) and the y-axis displays these journey times over seven days at one-hour resolution. On first inspection, the seven different days are clear. The late night and early morning periods display the lowest journey times, approaching $time_0$. The northbound direction, towards the city centre, exhibits the largest load in the mornings and interestingly, peaks on a Friday evening. The southbound, away from the city centre direction peaks daily in the afternoon and similarly to the northbound lane, has its weekly peak on a Friday evening. Saturday and Sunday show distinct behaviour to weekdays, with overall reduced journey times in both directions and a proportionally higher journey time on the southbound, away from city lane, in contrast to weekday behaviour. It is clear from this alone that even the categorisation of journey times as weekend vs weekday distributions (Mullick and Ray, 2012) is problematic, as each day exhibits its own distinct behaviour.

Saturation-delay relationship

By combining the Google journey time data with the ATC vehicle count data, volume delay curves may be plotted. In order to facilitate comparisons, the vehicle volume is converted to saturation where saturation is defined as:

$$saturation = \frac{volume}{capacity} \tag{3.3}$$

Defining capacity is challenging as there are a range of possible definitions:

1. Free flow capacity - the maximum vehicle volume before a time-delay is incurred
2. Maximum flow capacity – absolute vehicular volume a road can carry with no regard to time-delay
3. The UK Design Manual for Roads and Bridges (DMRB) qualitative capacity definition - "capacity is defined as the maximum sustainable flow of traffic passing in 1 hour, under favourable road and traffic conditions." (DMRB, 1999)

The DMRB provides look up tables that feature traffic capacities for a range of road types, road widths and number of lanes. A manual survey of satellite imagery was carried out to assess the lane count, estimate the road width for each of the ATC locations and thus provide an estimated capacity by this DMRB definition. However, such a method was deemed unacceptable due to the uncertainty in what constitutes '*favourable road and traffic conditions*.'

Instead, a more nuanced definition was adopted from Spiess (Spiess, 1990) which defined capacity as the volume at which congested speed is half the free flow speed. The paired Google and ATC data was queried to find the estimated volume when the vehicle's speed was 50% of the free flow (minimum) journey time ($time_0$). As a result, individual roads are given a capacity attribution. This capacity attribution is generally lower than that from those which are more qualitative.

Consider Figure 3.26 which illustrates the saturation-delay curve for ATC locations 19E, 35S, 19 (both) and 15 (both). The top left figure presents the raw saturation time delay data with two curves, the BPR function and a third order polynomial regression function ($Adj.R^2$=0.73) for the given data. The function derived from the empirical data closely matches the BPR function with some small deviances at very low and very high saturation ratios. In this example, the road experiences a healthy amount of demand and this is reflected in a maximum saturation of less than 120%.

Other locations experience much higher peak loads, resulting in a different functional relationship. Figure 3.26 top right presents the saturation time delay curve for ATC location

35S. This road experiences higher demands than ATC 19E and exhibits saturation levels approaching 150%. In this case, the BPR function deviates significantly from the empirical data from 50% saturation and approaches a 2-fold time difference as it approaches a saturation level of 150%. For this particular location the empirical data suggests that an increase in volume does not have as significant effect on journey times as the BPR function would model.

Different direction road lanes share similar geometry with individual characteristics and experience different demands at different times. A comparison can be made of their individual saturation delay curves in order to assess their individual characteristics. Figure 3.26 bottom left illustrates location 19 in the eastbound and westbound directions. The two directions exhibit distinct behaviour despite being the same road. The westbound direction exhibits large time delays at high saturation levels, peaking at a 5-fold increase over free flow journey time. The eastbound direction does not have such a time delay peak. As a result, their individual regression lines differ, with a steeper curve westbound and a gradual curve eastbound. The BPR function falls between these functions and reflects what is close to an average condition when collectively considering both directions.

Figure 3.29 bottom right displays ATC location 15 where again distinct directionally dependent behaviour is exhibited. The Northbound direction exhibits a higher capacity to absorb traffic volume, showing a lower rate of time delay increase for a substantially larger increase in saturation. Saturation levels peak at nearly 2 illustrating that the use of the Spiess (Spiess, 1990) methodology for capacity definition has a significant impact on saturation. Such a metric would generally not be exhibited using the kind of qualitative definitions discussed previously. There is a contrasting relationship to the BPR function. It is shown to highly overestimate the time delay from the point of 50% saturation. It is clear that both lanes exhibit a greater capacity to accommodate increased vehicle demand with a reasonable time-delay than the BPR functions predicts.

Upper & lower bounds

Volume delay functions have significant issues at their lower and upper bounds. At the lower bound, there is the challenge of distinguishing between vehicular volume up to the point of free flow and at the upper bound limiting vehicular volume by the maximum flow capacity. In the lower bound a probabilistic method can be employed to handle the uncertainty in vehicle volumes. The raw DfT ATC traffic counts provide vehicular volume distributions 24 hours per day and as such those points which fall in the lower bound, generally late night and early morning can be used to define a volume distribution. Thus, in the situation where a volume delay function is employed to estimate the traffic volume for a time-delay ratio of 1, an estimated traffic volume can be given dependent on the time of day of the query.

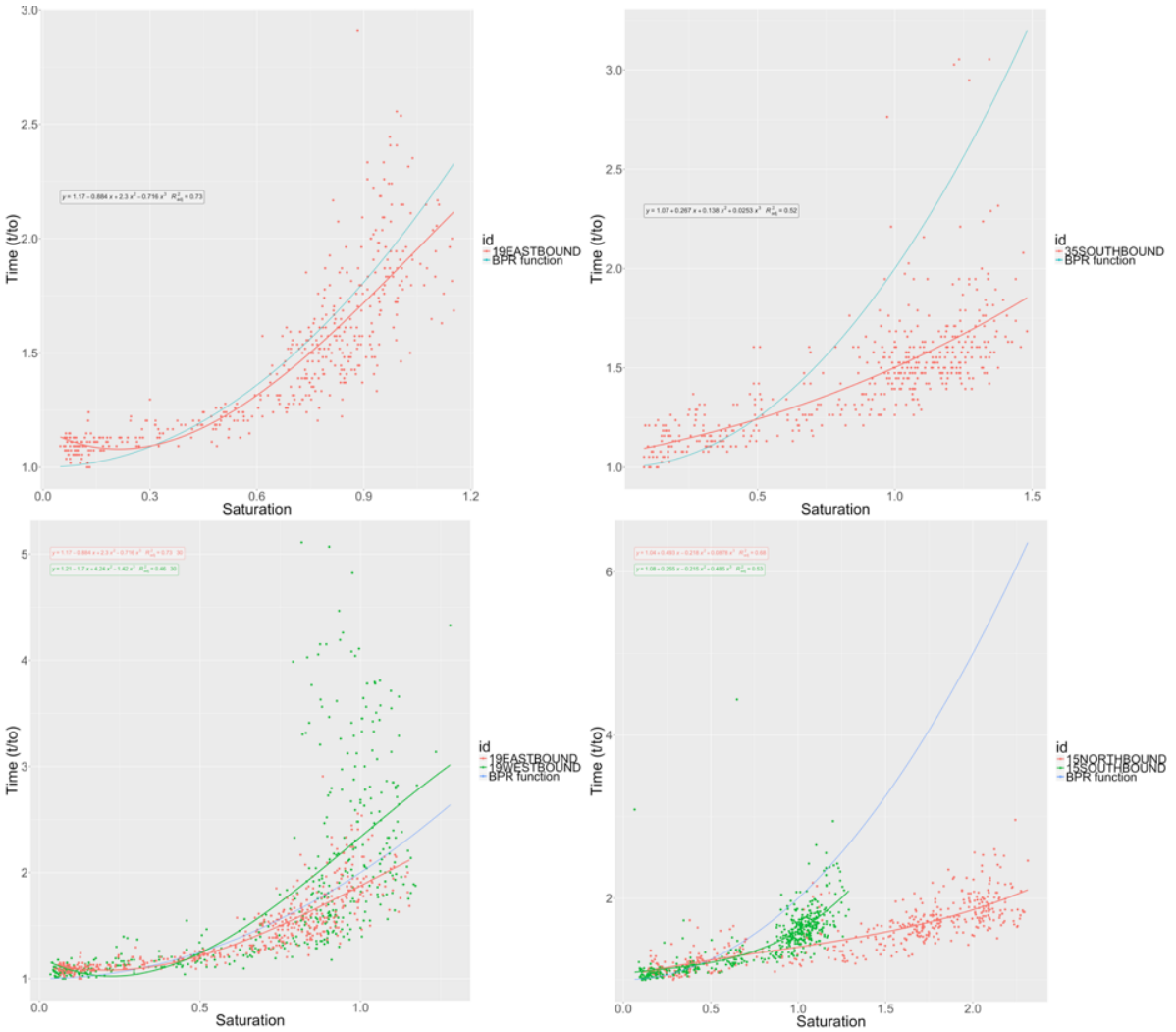


Fig. 3.26 Saturation time delay curve ATC 19 Eastbound (top left), 35 Soutbound (top right), 19 (bottom left),15 (bottom right)

The upper bound poses a different challenge. The peak exhibited journey times (defined here as > 3 times free flow journey time) are usually attributable to external factors and not simply vehicle to vehicle interactions. Examples of the external factors include road traffic accidents, weather events or road maintenance. In such situations it is possible for journey times to reach many times the free flow journey time. For illustration, ATC 11N had a journey time 6.85 times greater than the free flow journey time on the 1st March at 9am. A cached internet search highlighted emergency road works at this location due to a burst water main (*http://bit.ly/1ZYxzQ0*). In such a case the estimated vehicle volumes are greatly overestimated by the volume delay function and there is therefore a need to present a constraining factor, maximum flow capacity.

Table 3.8 Upper & lower bound distribution of volume delay ratio analysis

Bound	Definition	Percentage
Lower	< 1.1 time ratio	7.56%
Upper	> 3.0 time ratio	1.14%

In order to assess the frequency of this challenge the data was analysed to better understand the frequency of the lower and upper bound conditions. The results are presented in Table 3.8. For the lower bound a conservative estimation was taken with the frequency of time ratio less than or equal to 1.1 in order to capture possible driving style fluctuations above the free flow journey time. These values constituted about 7% of the total sample. The upper bound, defined as greater than or equal to 3 times free flow journey time was about 1% of the sample. This illustrating that the lower bound conditions are relatively rare in a high density, high demand area such as London and that this condition is only ever met at periods of little concern, late night and early morning. The upper bound conditions were even more rare, at 1% of the sample.

In conclusion, the lower bound problem constitutes an issue at times of little concern and very low demand, late night and early morning. A probabilistic solution can be used to interpolate and provide estimated values if needed. On the other hand, the upper bound problem is a much less prevalent than the lower bound problem but is much more profound in that it has a huge impact on travellers and generally hits at high impact ‘rush hour’ times. A constraining factor is required to contain the volume delay function in such a case. Another feature of upper bound conditions is that they involve a large number of travellers (in contrast to lower bound conditions) and as such their impact is more significant. The incidents that lead to the upper bound issue are discussed in terms of possible data sources and predictive inclusion illustrated previously.

Saturation-speed relationship

The vehicle speed can be plotted against the saturation ratio in order to relate the impact of vehicle volume on the resulting vehicle speed. The saturation-speed relationship for ATC location 19E is illustrated on the left in Figure 3.27. A third order polynomial fit is given for each direction ($Adj.R^2$ Eastbound = 0.84, $Adj.R^2$ Westbound = 0.74) of the data is presented in contrast to the BPR function. The empirical data shows a significant overestimation by the BPR function at low values of saturation, showing that average speeds do not approach the given speed limit on this road. The function's converge at a saturation level of 60% to the Eastbound lane and match relatively closely at higher levels of saturation. Again, it is clear that the roads have unique behaviour dependent on directionality. They share the same shape but the westbound lane exhibits speeds 4 *m/s* slower than the eastbound lane at the same saturation ratio. It also shown again that the BPR function over overestimates speeds at low saturation, reinforcing the conclusion that drivers often don't approach the road's maximum permissible speed even when it may be legally possible to do so. Figure 3.27 on the right presents a speed saturation curve where the BPR curve illustrates a different shape to that of the empirical data. Again, it overestimates vehicle speeds at low saturation levels and predicts a larger range of speeds than those exhibited. The differences between the north and southbound lanes is again present, although not to the same degree as ATC 19. A relatively constant 1.25 *m/s* speed difference is found between the two-lane directions for the same saturation level. The BPR function is highly dependent on the free flow journey time which has been defined as a function of link speed and length. The shape of the BPR function reasonably matches the empirical data in these cases but often presents locational errors in those sites were the speed limit is rarely, if ever met by drivers in real-world conditions.

3.2.5 Discussion

In many cases, the plots generated in this study presented empirical data that did not match the conventional macroscopic understanding, as epitomised by the BPR functions. For each direction and each individual site, a third order polynomial fit was generated to create context specific saturation delay and saturation speed functions. The $Adj.R^2$ for these generated functions has been plotted as a probability frequency distribution function (density) plot in Figure 3.28 in order to assess the confidence of the derived functions. Interestingly the saturation speed functions exhibited greater predictive ability than the saturation delay functions. The speed metric is derived from the time-delay metric by considering the fixed length of the road, thus the two functions share a similar shape. It is clear that some derived functions have low confidence and there are external unknown factors that have not been considered.

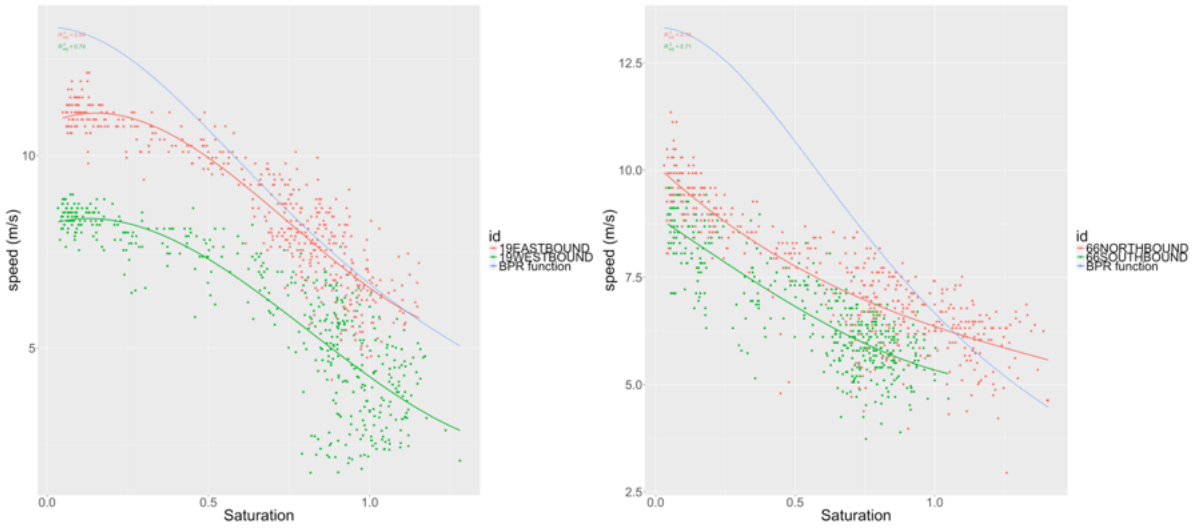


Fig. 3.27 Speed saturation curve ATC 19 (left) and 66 (right)

Example scenario & possible factors

Figure 3.27 presented a speed saturation plot for ATC66 in both directions. In an attempt to better understand and possibly explain this distinct directional behaviour, satellite photography and street level photography of the road was assessed. Figure 3.29 shows a satellite image of ATC66. The red marker gives the location of the ATC counter itself and the two blue markers illustrate the origin/destination location (dependent on direction) for the Google Directions API request. The speed saturation plot (Figure 3.27) shows a significant speed reduction for traffic in the southbound direction to that of traffic in the northbound direction. From the satellite imagery and street view imagery (Google, 2016d) possibly explanatory factors can be identified:

1. The southbound lane features on road parking
2. The southbound lane features a large bus stop and taxi lay-by (for Leytonstone Station) and a smaller bus stop lay-by

The larger bus stop and taxi lay-by serving Leytonstone is a significant geometric feature that is likely to have heavy impacts southbound traffic as buses/taxis leave and enter from both directions on the road. The second smaller bus stop lay-by and on-road parking may also have an impact, albeit in a smaller way to the traffic speed in the southbound direction. Such factors may explain the exhibited differences from the location device informed journey data and thus permit their inclusion in an informal way.

Challenging scenarios

A subset of roads presented results that did not resemble previously defined functions. Consider Figure 3.30 which presents the saturation time delay curves for ATC location 9 in both directions on the left. Visually it can be seen in the lower part of the plot that there are the same general trends discussed in Section 3.2.4 but with a large amount of highly variable journey time outliers. There are over 20 points that exhibit journey times over twice the length of the minimum exhibited journey time, peaking at 7 fold increases. These outliers result in low correlations for a second order polynomial fit ($Adj.R^2$ of 0.071 southbound and 0.16 northbound). Such extreme differences may be explained by events such as weather or road traffic incidents, for example a road traffic collision or road flooding as a result of high level of precipitation in a short period of time. Figure 3.30 presents the saturation time delay curve for ATC 67 in both directions on the right. A chaotic pattern of saturation time-delay data is given and little order is evident. Two second order polynomial fits are presented and as is visually evident offer little in the way of a robust function, presenting an $Adj.R^2$ of 0.05 northbound and $Adj.R^2$ of 0.22 southbound respectively.

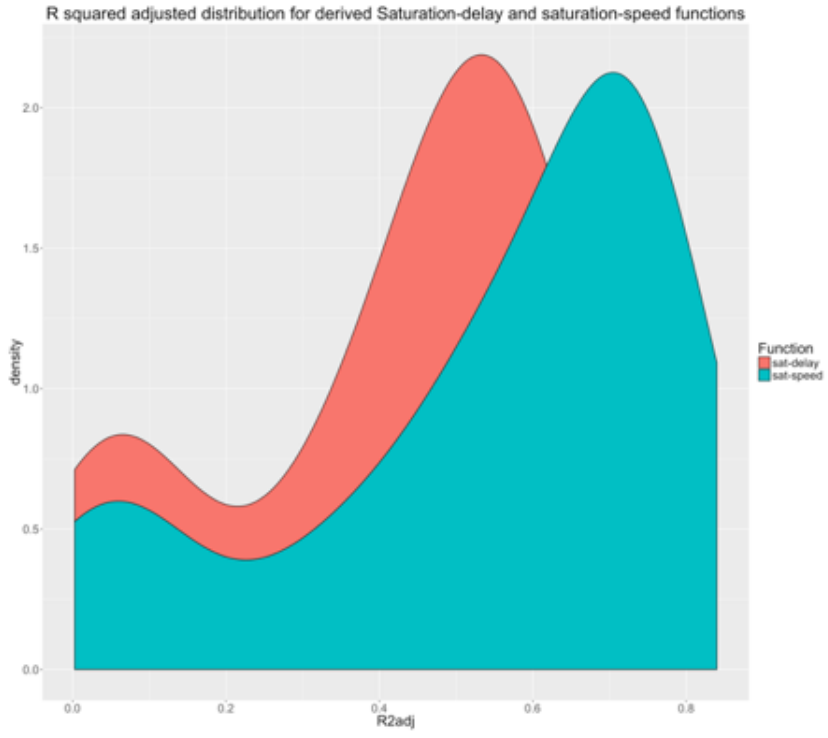


Fig. 3.28 R_{2Adj} probability density function for derived saturation-delay and saturation-speed functions

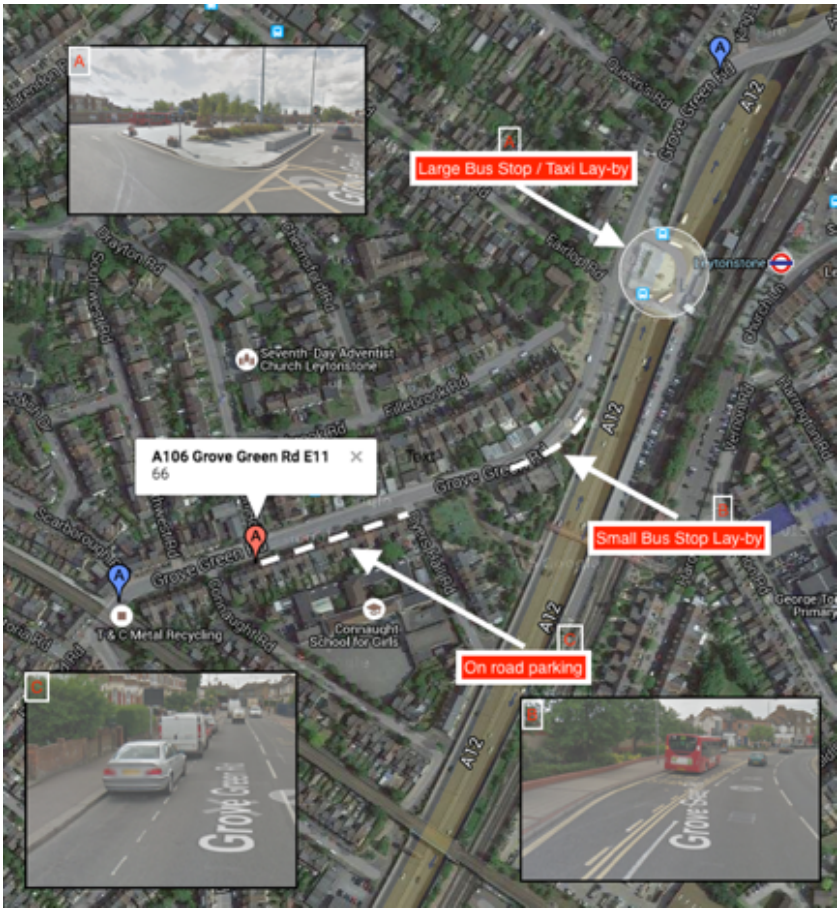


Fig. 3.29 Satellite & Street View images of ATC 66 (Maps, 2017)

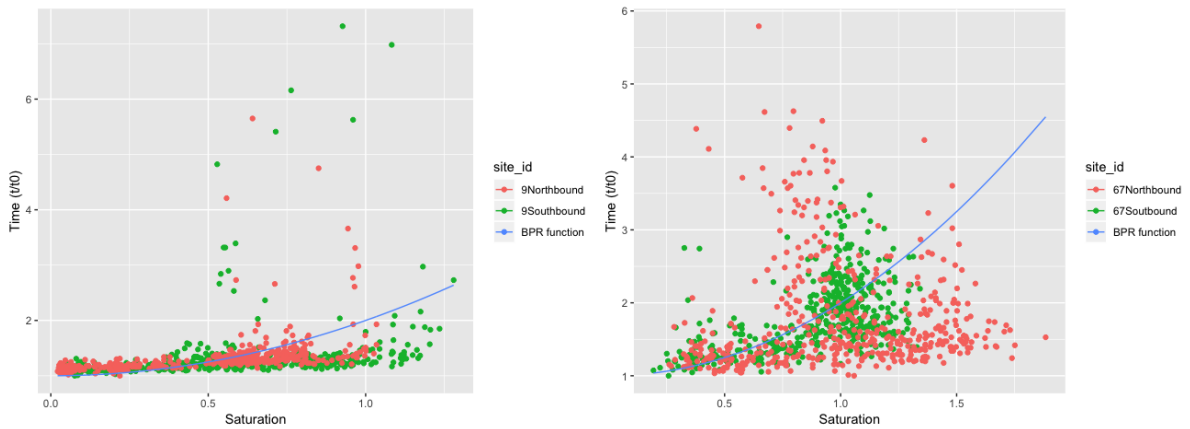


Fig. 3.30 Saturation time delay curve ATC 9 (left) and 67 (right)

Influencing factors

It is clear from these plots that there are factors not considered in this analysis that may help explain the unexplained variations. In Section 3.2.5 the existence of public transport infrastructure and on-road parking were identified as being potentially explanations of the distinct southbound northbound behaviour on that road. A range of possible factors have been identified:

1. Vehicle type

   Further work is planned to assess the impact of the vehicle type. In this paper each vehicle has been attributed equally despite the clear distinction between the interactions of a car and another car compared to that of a lorry to a car (Vap and Sun, 2007). A wide variety of vehicles mixes are exhibited on different road types and such differences should be considered. The use of statistical vehicle mix sampled by road type (DfT, 2012b) was discounted for this study due to the small sample size of such statistics compared to the resolution of the data used here. The use of traffic cameras with number plate recognition and sufficient privileges to the Driver and Vehicle Licensing Agency database would enable the disaggregation of vehicle type at a similar spatial and temporal resolution to the Google Directions and DfT ATC data presented here.

2. Weather

   Weather events may impact on journey times by impacting on the performance of the vehicle, the performance of the road and/or the performance of the driver. Ongoing research at the University of Cambridge is combining a large dataset of Google Directions journey times with data from the UK's Met Office NIMROD precipitation dataset (Office, 2003) in order to assess the relationship between these variables. At the microscopic level it is known that an increase in precipitation increases journey times as a result of increased risk and the resulting decrease in vehicle speeds to compensate for this (Mashros et al., 2014).

3. Road incidents

   Road works and road traffic collisions can lead to decreased or even zero capacity on a road link, resulting in increased saturation and thus impacting on journey time. Depending on the warning before such an event, the vehicle traffic may have the ability to adjust to this information, resulting in a greater distribution of traffic, leading to mediated travel times. Alternatively, an accident may occur and not enable any warning to be given to other road users until they are committed to their route choice, resulting in large

journey time increases, perhaps explaining phenomena such as that exhibited in Figure 3.33. A method incorporating different accident and road works databases with Google Directions data is currently being investigated.

4. Road geometry, type & land use

   Different road layouts may result in an increase in the complexity of vehicle interactions. For example, the curvature of a corner and the road surface quality will impact on the speed of a vehicle. The surrounding land use will likely also impact, adding safety concerns (for example a school or leisure centre) again impacting on vehicle speed. The inclusion of such factors poses many challenges, the size and complexity of the data plus the uncertainty and variability in how drivers react to the data.

   In Figure 3.29 a series of geometric factors are displayed as part of an attempt to explain different behaviour on the same road dependent on direction. The factors discussed there, on-road parking and bus lay-bys may be quantitatively captured using machine vision and data sources such as Google Street View.

3.2.6 Conclusions

A range of context specific saturation time-delay, speed saturation and journey time distribution curves for a range of different locations and road types have been generated. Specific examples have been presented here for discussion and all generated functions and plots are available for inspection here. In the most practical sense some of these functions may now be used in the traffic assignment stage of the traditional four step model. In some cases, the data presents clear evidence that unknown factors, such as those listed in Section 3.2.5, have a significant impact and warrant further investigation. In these cases the derived functions and indeed any standardised function have been shown to deviate significantly from empirical data and as such their use should be considered with care. This said, the probability frequency distribution in Figure 3.28 describes the significant collective correlations between the ATC traffic count data and that of the Google journey times, across a range of sites, presenting evidence which goes some way to validate the Google data and illustrate the collective value of this method. The data used here shows promise in considering the tangible factors which impact on road performance, such as local geometry, bus stops and so on, but that have historically been too challenging to be considered. These data sources have longevity, exist at close to real-time and in the case of the Google Directions data, relatively low cost with little or no capital expenditure required for its harvesting. These methods may be employed over a long time horizon and at a finer temporal resolution in order to better understand the temporal and spatial trends as well as the influencing factors such as sporting occasions and weather events. It can also be used to do

real-time vehicle emissions estimations and modelling, as is being investigated presently. The automation of this method over longer time horizons may lead to explanations for the issues discussed previously and highlight areas that require investigation in order to better understand the performance of road infrastructure.

3.2.7 Summary

1. Hypothesis

   Google Directions API data can be combined with DfT ATC data to create context specific saturation time-delay, speed saturation and volume delay functions. Such functions may be used within traditional modelling paradigms and more modern dynamic models in order to simulate vehicular interactions.

2. Novel contributions

   The method developed here has been shown to relate traffic volume data from counters to journey time data from Google in a statistically significant way. The result is functions which may be used to better consider local factors than the standard Bureau of Public Roads methods. This analysis also provides part validation to the Google Directions API data discussed in the previous section.

3.3 Quantification of public transport performance

Public transport networks in cities such as London are inherently complex. Journey times, routes and travel modes for a given origin and destination are more challenging to compute as shortest path algorithms must feature more considerations than they do for the road network. The influence of mobile technology is pronounced here, as users lean on algorithms for modal and route decision making. Transport operators such as TfL have invested in real-time route information technologies and there are dedicated (e.g. CityMapper) and more general (e.g. Google Maps) apps which aid and inform traveller decision making. This section investigates how these data feeds may been used to derive temporally dynamic network statistics in the literature. A novel application combining two types of data analysis from the literature for the case of London is presented.

3.3.1 Literature Review

3.3.2 Available data feeds

Many transportation authorities now provide information on the status of their services at different temporal resolutions. This is often published on a website, via social mediums such as Twitter, as email updates or as raw feeds that may be utilised by software developers, such as CityMapper or Google Maps. Figure 3.31 shows one such feed being used to display real-time arrivals at a station.

Fig. 3.31 Example TfL service status feed

Transport for London (TfL) were one of the first to adopt an open access policy to their data feeds in an attempt to encourage developers to create innovative services to aid users.

TfL's datasets include real-time feeds, fixed datasets and transparency orientated datasets (ODI, 2016).

Similarly to the Google Directions API data discussed previously, these data sources are generally made available for individual traveller use. Despite this, they may also be polled over long periods of time, at fine temporal resolution, in order to quantify how different services change with time (e.g. (Jariyasunant et al., 2011), (Antrim et al., 2013)). Some of these data feeds are available at one minute temporal resolution and thus it can be possible to capture the highly evolving nature of the transport network by harvesting and quantifying these metrics.

This section will consider two distinct data types:

1. Real-time status feeds
2. Schedule feeds

Real-time status feeds

A range of TfL status feeds are available at different temporal resolutions through an API. A Hypertext Transfer Protocol (HTTP) request can be programmatically scheduled for polling at the refresh rate of the API they are querying. The response to this HTTP request can then be stored and archived in an automated way. The TfL feeds return Java Script Object Notation (JSON) objects which are unique to the feed itself. Table 3.9 lists relevant TfL feeds and their refresh rate.

Table 3.9 TfL feed summary

Feed	Temporal resolution
Bus status	1 minute
National Rail status	1 minute
Tube, Overground, DLR, TfL Rail status	1 minute
Bike point status	5 minutes
Road status	5 minutes
Air quality	60 minutes

As an illustration, a raw individual record from the tube status feed on the 1st December 2016 at 16:04 is shown in Figure 3.32. This particular record indicates that there are minor delays on the Circle line, as a result of signal failure at Gloucester Road.

A parser may be set for the specifics of each given feed in order to programmatically handle the incoming data. In the case of the bus, national rail, tube, overground, DLR and TfL rail feeds this involves finding the qualitative level of disruption, the type of disruption, the spatial

```
{
"modeName": "tube",
"serviceTypes": [
   {
      "$type": "Tfl.Api.Presentation.Entities.LineServiceTypeInfo, Tfl.Api.Presentation.Entities",
      "name": "Regular",
      "uri": "/Line/Route?ids=Circle&serviceTypes=Regular"
   }
],
"name": "Circle",
"created": "2016-11-22T09:56:49.69Z",
"$type": "Tfl.Api.Presentation.Entities.Line, Tfl.Api.Presentation.Entities",
"routeSections": [],
"lineStatuses": [
   {
      "created": "0001-01-01T00:00:00",
      "$type": "Tfl.Api.Presentation.Entities.LineStatus, Tfl.Api.Presentation.Entities",
      "statusSeverityDescription": "Minor Delays",
      "disruption": {
         "category": "RealTime",
         "closureText": "minorDelays",
         "description": "Circle Line: Minor delays while we fix a signal failure at Gloucester Road. ",
         "isWholeLine": true,
         "$type": "Tfl.Api.Presentation.Entities.Disruption, Tfl.Api.Presentation.Entities",
         "affectedStops": [],
         "affectedRoutes": [],
         "categoryDescription": "RealTime"
      },
      "validityPeriods": [
         {
            "$type": "Tfl.Api.Presentation.Entities.ValidityPeriod, Tfl.Api.Presentation.Entities",
            "fromDate": "2016-12-01T16:04:28Z",
            "isNow": true,
            "toDate": "2016-12-02T01:29:00Z"
         }
      ],
      "statusSeverity": 9,
      "reason": "Circle Line: Minor delays while we fix a signal failure at Gloucester Road. ",
      "lineId": "circle",
      "id": 0
   }
],
"modified": "2016-11-22T09:56:49.69Z",
"disruptions": [],
"id": "circle"
}
```


Fig. 3.32 Example tube feed status (1st December 2016)

bounds of the disruption (for example, from station *x* to station *y*) and the temporal bounds (e.g. from 17:00 to 17:50]).

Schedule feeds - GTFS

Beyond these real-time feeds there are the planned schedules which are updated over longer time horizons. Such schedules include planned engineering works and may be considered as the idealised functionality, capturing the designed temporal differences in journey times as an output of timetabling and service changes.

The General Transit Feed Specification (GTFS) is a a standardised format for public transport schedules and associated geographic information (Google, 2016c). The GTFS supports multiple transport modes and is used widely by transportation operators to share schedule information with their users, predominantly through the Google Maps (*Google.com/maps*) service. The standardisation of the specification and the simplicity of its structure has led to its use in applications beyond that relating to individual journey support (Google, 2016c). The GTFS was an important step in the challenge of synthesising multi-modal information into one usable format (Antrim et al., 2013).

The GTFS consists of a series of text files bundled in a compressed *.zip* file. Of these text files, 6 are required and 7 are optional. The different required and optional files are listed, defined and tabulated in Table 3.10 and the overall architecture of the GTFS is graphically illustrated in Figure 3.33.

In the UK the GTFS is not officially available from either operating companies, the Association of Train Operating Companies (ATOC), the Department for Transport (DfT) or Transport for London (TfL). However, since timetabling information the official operators in available under public sector information licenses the GTFS may be created separately. In the case of the UK rail network, this is done weekly by volunteers and published on the *gbrail.info* website every Saturday morning. This GTFS file is for the entirety of the UK and it is thus necessary to first filter the data not relevant to the Greater London Area. It is also possible to compute a GTFS file specifically for London using TfL's feeds and free, open-source tools (e.g. (CommuteStream, 2016).

3.3.3 Performance metrics

The data types described in this section are the focus of many research groups both in industry and academia. A large range of tools exist in the open source community that may be leveraged to compute a range of statistical metrics for a given city, region or area. The standardisation of the GTFS and its now de-facto status as the format of choice is in no small part responsible

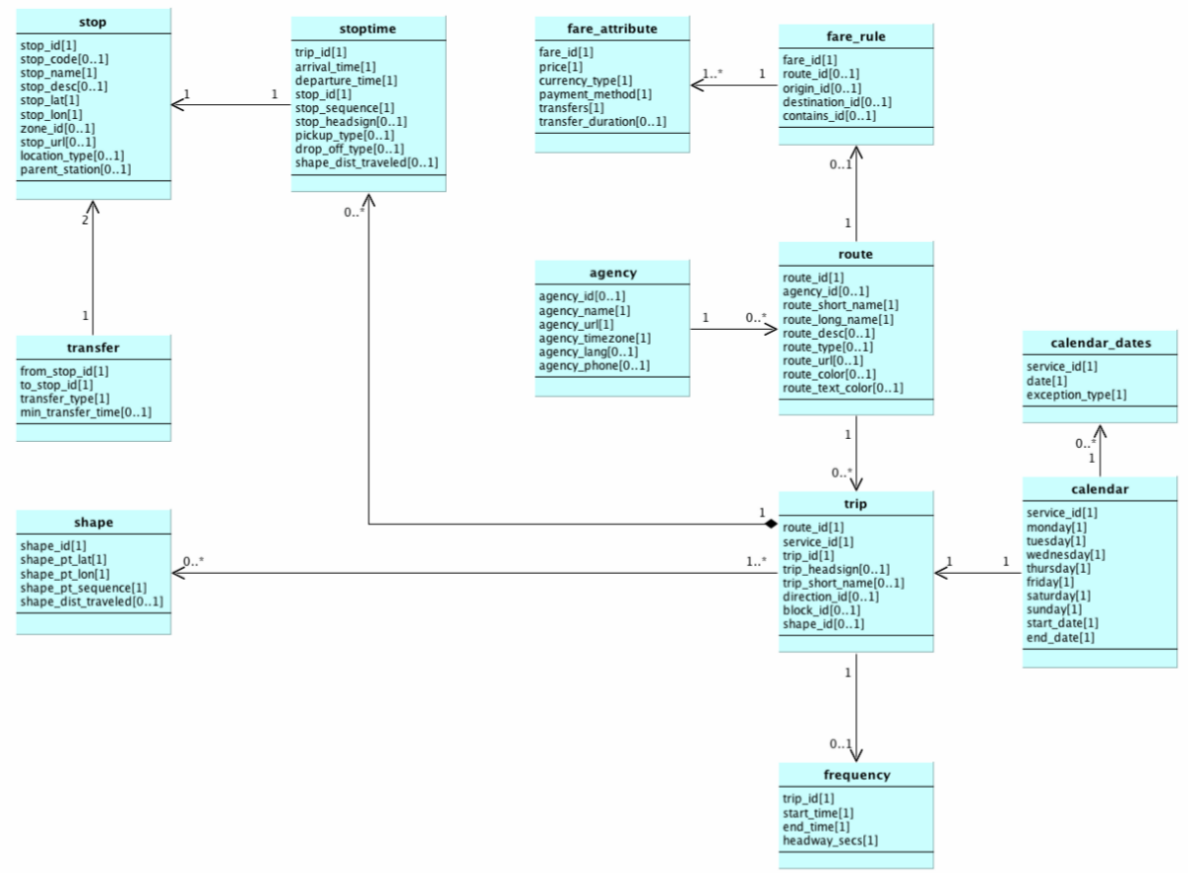


Fig. 3.33 General Transit Feed Specification (GTFS) (Google, 2016c)

Table 3.10 GTFS file specification

File	Required	Defines
agency.txt	Required	One or more transit agencies that provide the data in this feed.
stops.txt	Required	Individual locations where vehicles pick up or drop off passengers.
routes.txt	Required	Transit routes. A route is a group of trips that are displayed to riders as a single service.
trips.txt	Required	Trips for each route. A trip is a sequence of two or more stops that occurs at specific time.
stop_times	Required	Times that a vehicle arrives at and departs from individual stops for each trip.
calendar.txt	Required	Dates for service IDs using a weekly schedule. Specify when service starts and ends, as well as days of the week where service is available.
calendar_dates.txt	Optional	Exceptions for the service IDs defined in the calendar.txt file. If calendar_dates.txt includes ALL dates of service, this file may be specified instead of calendar.txt.
fare_attributes.txt	Optional	Fare information for a transit organization's routes.
fare_rules.txt	Optional	Rules for applying fare information for a transit organization's routes.
shapes.txt	Optional	Rules for drawing lines on a map to represent a transit organization's routes.
frequencies.txt	Optional	Headway (time between trips) for routes with variable frequency of service.
transfers.txt	Optional	Rules for making connections at transfer points between routes.
feed_info.txt	Optional	Additional information about the feed itself, including publisher, version, and expiration information.

for this fertile research environment. It is noteworthy that TfL do not publish in the GTFS format and this in itself is likely a legacy of their own efforts in the open data realm before standardisation had occurred. Beyond the GTFS, there is the context specific nature of real-time feeds and these therefore do not lend themselves as easily to standardised assessment.

This section consists of further exploration of these two distinct data feeds individually and then moves to the merging of these fields in an attempt to capture the context specific nature of the real-time feeds in the wider, network level nature of the GTFS. The analysis of the real-time feeds and wider GTFS statistics will first take a general, London wide focus. After this, specific focus is made to the hubs (Heathrow, St Pancras and London City) relevant to the identified

case study in Chapter 2 and thus follows a similar pattern to the analysis in Section 1 of this Chapter.

Real-time status feed statistics

Of the feeds harvested in Table 3.9, the Underground (tube) statuses and the bus statuses were taken as the most relevant. The Underground carried 1.34 Billion and the bus network 2 Billion passengers in 2016 (TfL, 2017).

London Underground

TfL have 8 service categories for the London Underground - Good Service, Service Closed, Part Closure, Minor Delays, Severe Delays, Part Suspended, Special Service and Suspended. The TfL feed for tube status is available at 60 second resolution and this may be analysed in order to compute some basic statistics on service reliability, for a given line or for the Underground as a whole. As an example, consider Table 3.11 which shows the percentage of time spent in Good Service for the Piccadilly Line from January to November 2016. This is disaggregated again into daily proportions in Table 3.12.

Table 3.11 Piccadilly Line Good Service Reliability, 2016

Month	Good Service
January	73.79%
February	77.11%
March	72.88%
April	82.88%
May	81.71%
June	83.73%
July	77.96%
August	80.01%
September	76.07%
October	69.48%
November	58.06%

This illustrates interesting seasonal trends, with the Winter months exhibiting poorer performance than the summer months. However, since the feeds are available at 60 second resolution it is possible to assess this at much finer resolution, for each of the different tube lines, week to week and month to month. In Figure 3.34 the daily service status for each line, across one week (January 23rd-29th 2017) is presented in the top image. Each daily chart histogram shows the different proportions of time spent at each service status. A few

Table 3.12 Picadilly Line Good Service Reliability by day, 2016

Day	Good Service
Monday	84.09%
Tuesday	86.37%
Wednesday	87.63%
Thursday	82.29%
Friday	84.01%
Saturday	86.55%
Sunday	89.06%

planned events can be seen on Saturday and Sunday, with both the Jubilee and Metropolitan lines illustrating Part Closures for the entire period. Severe delays for the Central line and Suspension on the Waterloo & City line are the most significant, non-weekend and non-planned events on the 26th. The 27th presents a clustering of Minor Delays across most lines.

In Figure 3.35 the average weekly statuses are shown, across the period of one month (weeks 48 - 52 2016) in the bottom image. The weekly distributions discussed previously are not evident, but more macro trends can be seen. Severe delays on the Piccadilly Line are shown from week 48-50, as a result of the "Wheel repair" maintenance issues which plagued the line in late 2016 (Standard, 2016). Interestingly, the Planned Closure's which are shown in week 51 illustrates the full closure of the tube service over Christmas day (the only day of the year).

This may also be considered in the context of the proportion of time spent as "Good Service" as this is the level for which modelling inputs are generally derived on. That is to say, it is assumed that the scheduled timetable is adhered to and that no deviations occur. A probability density function may used to illustrate the relative probability of "Good Service" for each of the lines, over a sample period of 4 months (November 2016 to March 2017) and is shown in Figure 3.35. The Piccadilly line's issues in late 2016 result in the widest distribution, showing the lowest proportion of time spent at "Good Service". Contrastingly the Northern, line's narrower base illustrates a very different and more reliable "Good Service" profile. The ordered mean and median proportion of time in "Good Service" is shown in Table 3.13. Again, the Northern, Central and Jubilee Line's are shown to have the highest Good Service values. The Piccadilly, District and Waterloo Lines present mean Good Service proportions of 60/70%. To put this in context, two times out of five a traveller on the Waterloo line will be faced with some form of restriction over this study period.

London Buses

Fig. 3.34 Histogram plot of service status, disaggregated by line, daily and over one week (week 4, 2017) (top) and Histogram plot of service status, disaggregated by line, weekly and over one month (month 12, 2016) (bottom)

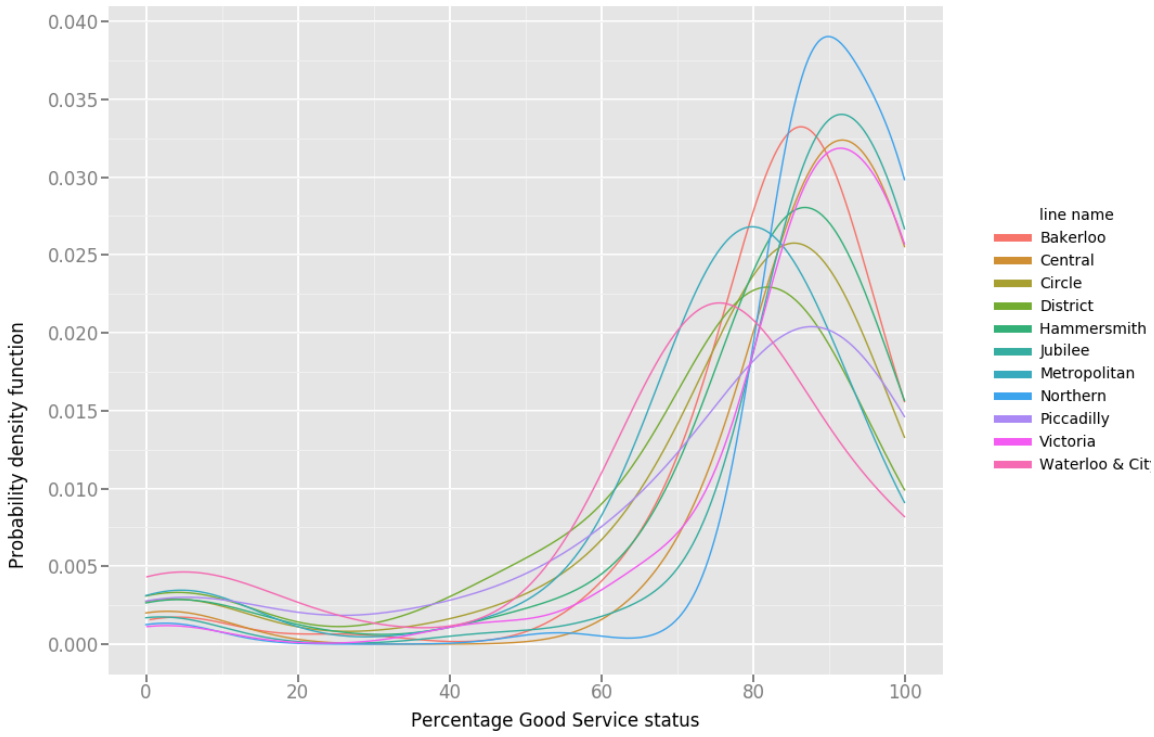


Fig. 3.35 Probability density function for "Good Service", disaggregated by underground line

Table 3.13 Median and mean proportion of time spent as "Good Service" per underground line

Line	Mean	Median
Northern	84.65%	87.73%
Central	82.65%	87.90%
Jubilee	81.89%	87.84%
Victoria	79.99%	86.73%
Bakerloo	79.2%	85.47%
Hammersmith & City	75.83%	85.95%
Circle	73.75%	84.36%
Metropolitan	71.59%	77.14%
Piccadilly	70.08%	82.00%
District	69.16%	78.66%
Waterloo & City	62.73%	74.08%

There are significantly more bus services (>700) than underground rail services (11) and it is therefore more difficult to show reliability on a service basis. The history and evolution of London bus services also makes it challenging to cluster routes, as often similar bus numbers have little or no relation to each other. In some cases, it is even possible to relate an existing bus number to a historic predecessor from the horse drawn carriage era (TfL, 2009).

The percentage time spent as "Good service" and "Special Service" is shown in Figure 3.36 for the 700 bus services. Clustering of similar bus services sharing special service status is evident, as bus numbers which share similar origins, destinations or major arterial roads tend to exhibit similar disruptions.

Issue causes

The 7 broad statuses published by TfL present a mixture of qualitative and quantitative context to different issues on the line. These range from planned closures (e.g. Christmas day) to planned maintenance (e.g. weekend works) to unplanned incidents (e.g. signal failure). The TfL feeds come with labelled string explanations, with varying and non standardised degrees of information. An exploratory word search was carried out on all of these explanations in order to get present a qualitative view on the kind of issues which are more prominent and result in delays of varying degrees and is shown in Figure 3.37.

This section has illustrated statistics from real-time feeds which present a challenge to the assumption that an idealised timetable is an accurate representation of a public transport networks performance. Public transport networks in cities of London's complexity exhibit tem-

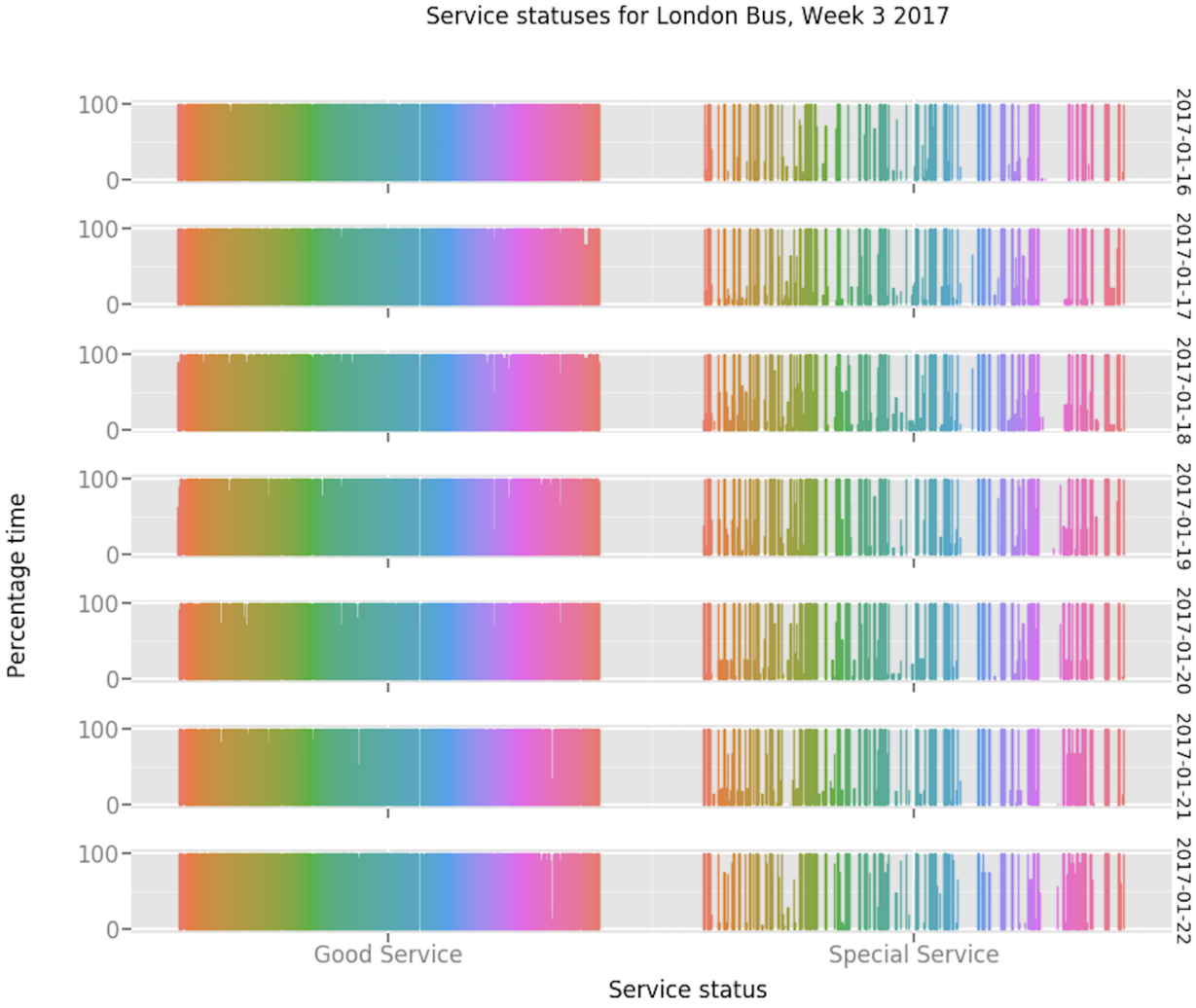


Fig. 3.36 Percentage time spent as good service, per bus service

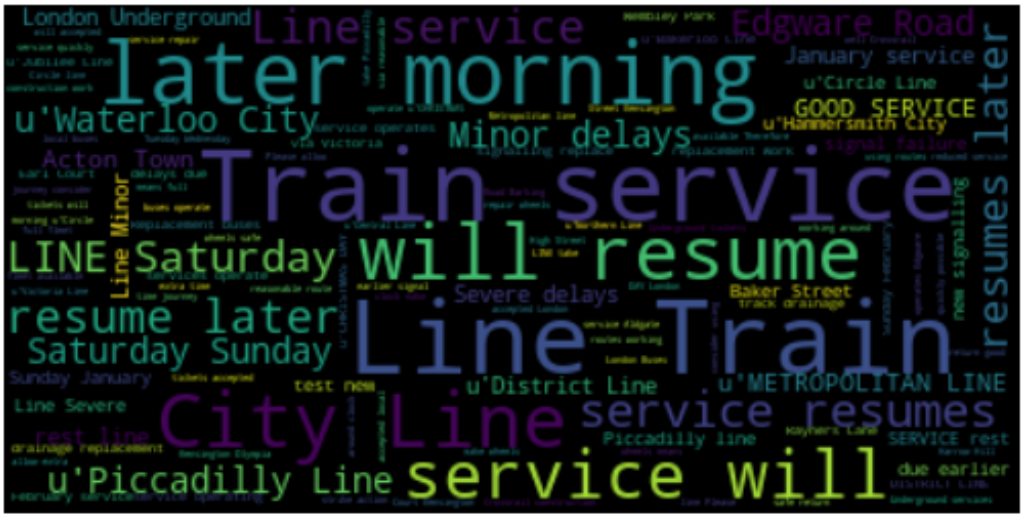


Fig. 3.37 String analysis of service disruptions

porally and spatially dynamic behaviour which produces sub-idealised behaviour for significant proportions of the time.

Schedule feed - GTFS statistics

As was previously discussed, the GTFS may be unofficially generated. In the case of London, the following *.txt* files are generated - agency, calendar, routes, shapes, stop times, stops and trips. Agency describes the organisations responsible for services (e.g. TfL), calendar and stop times explain the schedules, routes explains the different routes (e.g. bus route / rail line), stops explains the locations of bus stops, stations and trips explains the individual services themselves

A sample GTFS file is extracted arbitrarily (12/11/2016) and exploratory analysis carried out to derive some basic statistics on modes and times of service. The number of routes, disaggregated by type is presented in the top image of Figure 3.38. Bus routes are by far the most popular (672 routes), followed by rail (26), underground (12), ferry (9), tram (2) and cable car (1). Routes themselves are time-independent, it is the trips along a route which is time-dependent and ultimately the service itself of interest to travellers. A trip operates on a route and explains the temporal dimension of the previously specified route. For example, a trip may occur at time x and on a route y. Figure 3.38 presents the number of services, disaggregated by mode on the bottom. Here, the frequency of rail (underground and other) services compensates for the low number of routes in relation to that of the bus network. This is consistent with TfL's macro level ridership statistics (TfL, 2017). A trip operates on a route and explains the temporal dimension of the previously specified route. For example, a trip may occur at time x and on a route y. The frequency of rail (underground and other) services compensates for the low number of routes in relation to that of the bus network. This is consistent with TfL's macro level ridership statistics (TfL, 2017).

The tabular format of the GTFS is useful for generating high level, macro statistics. The transformation of the GTFS into a graph representation allows for a data model which enables a more nuanced interrogation of the public transport supply:

1. Extract stops from a trip
2. Find the service_id for the given trip_id
3. Find the calendar information for the given service_id
4. Generate the timetable information at the level of a day
5. For each timetabled stop_time, generate a temporally explicit edge for a given day

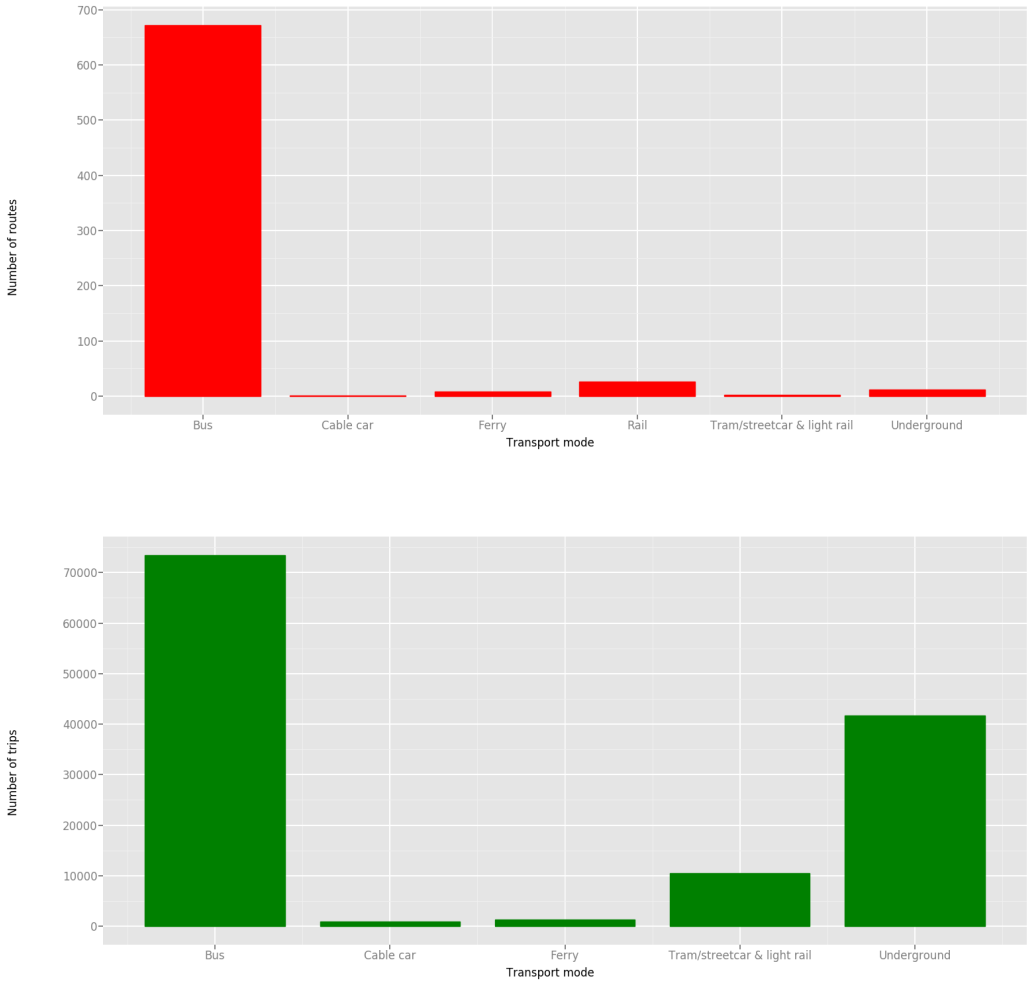


Fig. 3.38 Number of routes (top) and trips (bottom) disaggregated by mode in London

6. Compute journey time from respective departure and arrival time stamps.

An example output of this process is given in Figure 3.39.

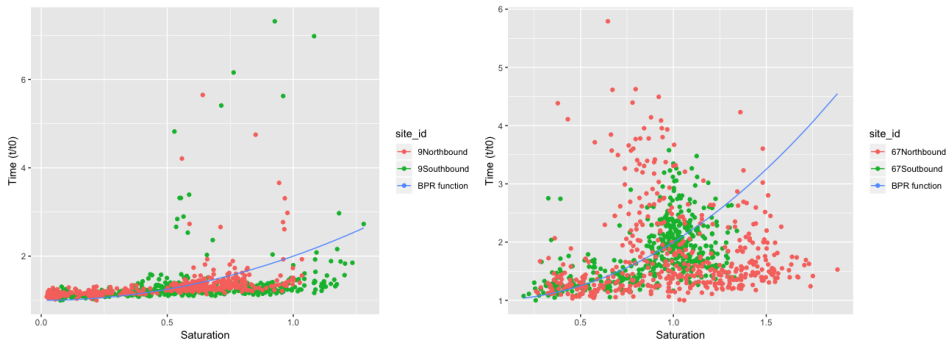


Fig. 3.39 Atomic GTFS edge data example

Since multiple trips and services may offer connectivity across unique vertices, the generated graph data structure timetable information must be merged along unique edges. An example output from this is presented in Figure 3.40.

The formation of the GTFS into the above data structure permits for efficient querying. The headway between different services by a given mode, on a given edge may be computed and is presented in Figure 3.41. Unsurprisingly, the Underground is shown to exhibit the lowest service headways and the Ferry shown to exhibit the largest. There is clear aggregation around different headways, with 5, 10, 15, 20, 30 and 40 minute headways presenting as very common. In Figure 3.42 the same analyses is presented, but with disaggregation between weekdays and weekends. In total, there are less services on the weekend, and these services generally have lower frequencies.

This data is updated weekly and small changes may be seen in the number of stops, frequency of services (trips) and even the location of stops. However, the issue with the static nature of the GTFS inputs identified by (Catala et al., 2011) is clear in that these changes are hard to plot in a meaningful way.

Journey level statistics - hub travel

The previous sections describe the use of two datasets which when considered in tandem may represent network level information at fine temporal and spatial resolution. The real-time status feeds offer localised but very fine temporal resolution information. Conversely, the schedule feeds offer wide spatial resolution but with crude temporal resolution, a limitation of the GTFS being the static nature of the input data (Catala et al., 2011). The individual event(s) captured previously have a macro impact on the network. These feeds must be interpreted and used to

```
{
  "negativeNode":"490010167W",
  "positiveNode":"490004661W",
  "edge_id":"490010167W490004661W"
  "services" :

    [
      {
        "route_id":"370",
        "route_type":"Bus",
        "journey_time":60,
        "arrival_time":1446442560,
        "route_agency":"tfl",
        "departure_time":1446442500
      },
      {
        "route_id":"370",
        "route_type":"Bus",
        "journey_time":60,
        "arrival_time":1446442560,
        "route_agency":"tfl",
        "departure_time":1446442500
      },
      {
        "route_id":"370",
        "route_type":"Bus",
        "journey_time":60,
        "arrival_time":1446443760,
        "route_agency":"tfl",
        "departure_time":1446443700
      },
```


Fig. 3.40 Merged atomic GTFS collated edge data example

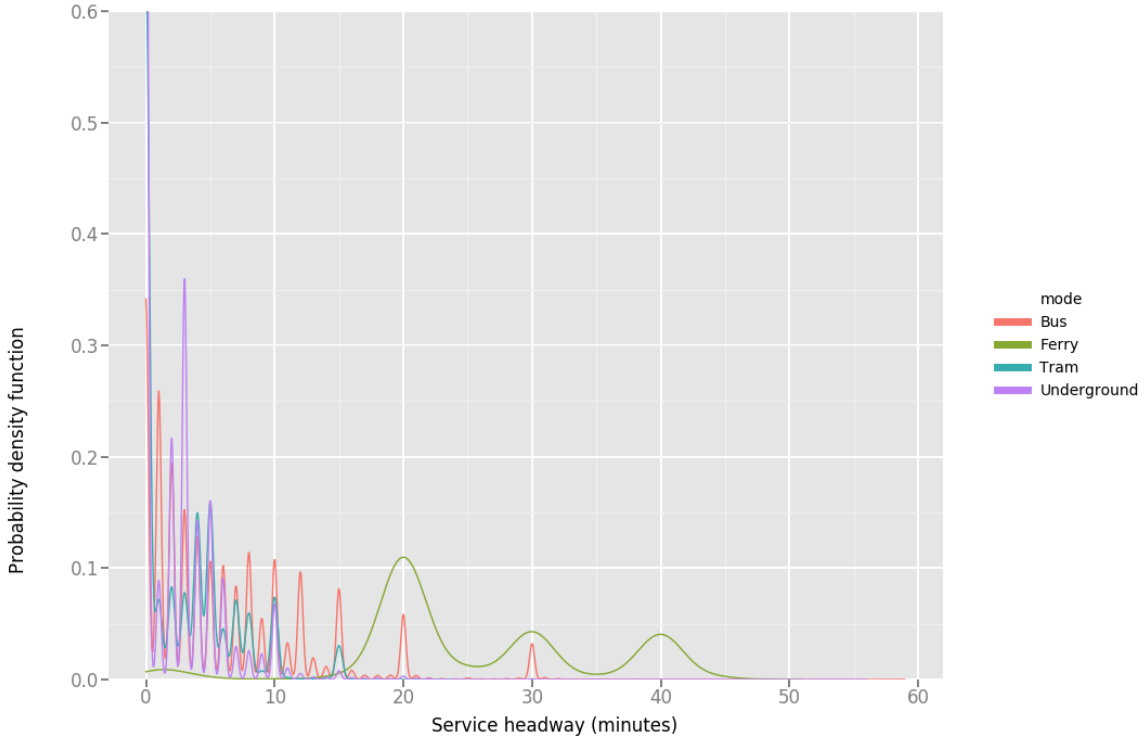


Fig. 3.41 Probability density function of service headways by mode

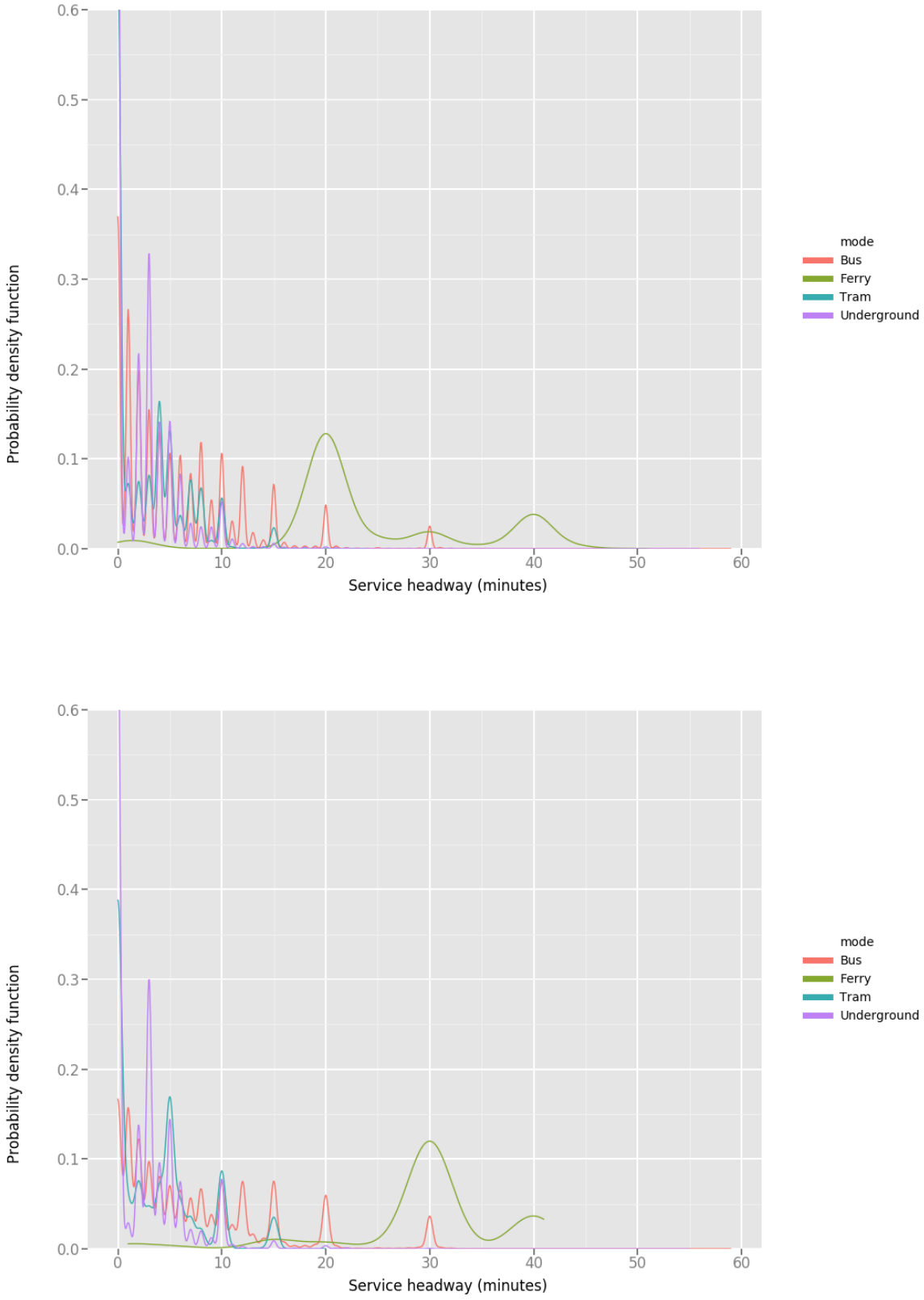


Fig. 3.42 Probability density function of service headways by mode, by weekdays (top) and by weekends (bottom)

compute what they mean for the public transport graph. For example, the closure of a given rail link must be represented in the graph that a modelled user may be presented with. These may be graphically shown as in Figure 3.43 and contrastingly conceptualised as the following:

1. Micro impacts - Capture individual changes and their respective cause (if known)
2. Macro impacts - Quantify the influence these collective individual impacts have on network performance

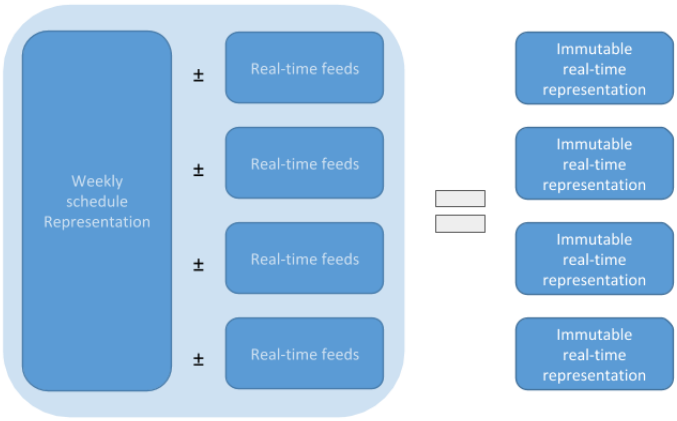


Fig. 3.43 Combining static and real-time feeds

Beyond the open source and academic community, this data is available as a service. The process discussed here is carried out by Google and their network representation may be queried through the same API as that used in Section 1 of this chapter. Thus, it is possible to query the Google Directions API in a similar method to that shown in Section 1 of this chapter for public transport routing. Exactly the same temporal (Table 3.2) and spatial (Figure 3.6) resolution specification presented and discussed in Section 3.1.3 was used to query the Google Directions API for public transport routing. In the same fashion as the Directions API section the spatial and temporal of the generated dataset may be assessed.

Temporal Coverage

1. Average speed

The average speed was computed at hourly slices across the entire dataset and plotted, as is shown in Figure 3.44. Average speeds increase during the working day and peak at 5pm, with increased service frequencies and the existence of express services at peak times are responsible

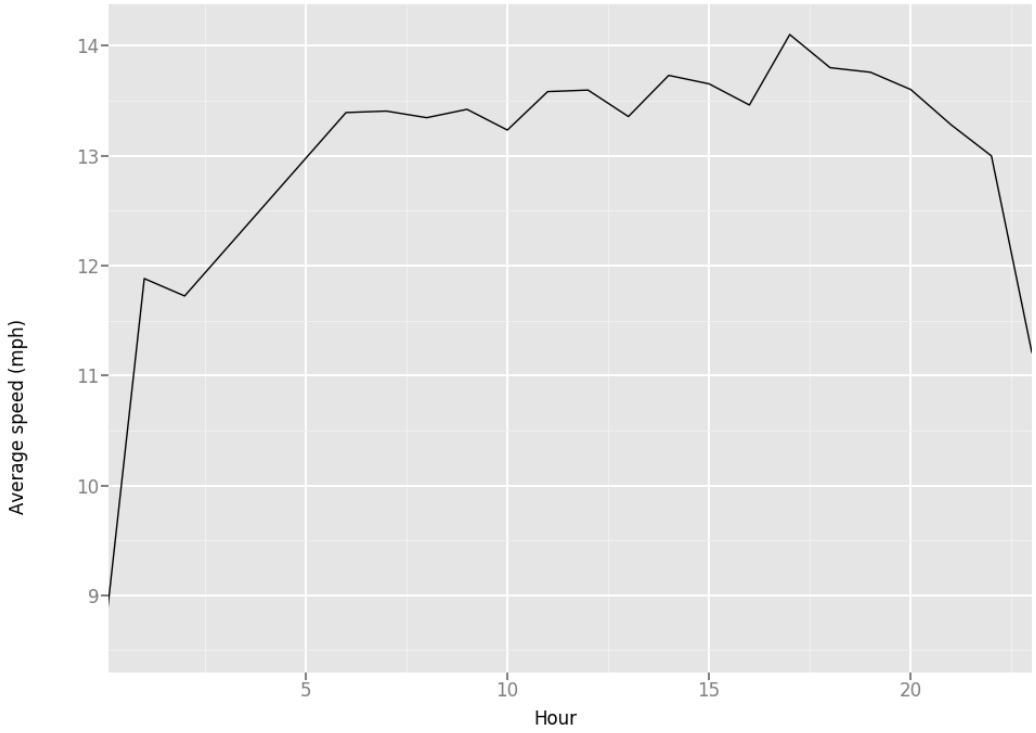


Fig. 3.44 Public transport hourly average speed over study period

for this peak. The trend is notable in that it is the inverse to that for driving as was shown previously in 3.1.

2. Journey time distribution

Four contrasting sample plots of journey times, over the 5 week study period were extracted. In Figure 3.45 a journey to St Pancras is shown from LSOA E01001682 at the top. The minimum journey time is shown to be around one hour and the maximum around 15 minutes slower. Neither the weekly or daily plots show an obvious pattern. In Figure 3.45 the bottom image presents a shorter journey to the same destination. In this case, the range is around the 100 second mark and daily/weekly patterns are clear. The quickest journey is shown to be in the morning and the slowest in the evening. In Figure 3.46 a considerably longer journey to Heathrow is shown in the top image. The minimum journey time is just over an hour and the maximum journey time is nearly 2 hours. Generally, the morning is seen to exhibit the quickest journey times with the evening showing the slowest. Lastly (Figure 3.46) a medium length journey to London City Airport is shown in the bottom image. A general daily pattern with late afternoon/evening peaks in journey times is exhibited. The range in journey times is shown to be around the 10 minute mark.

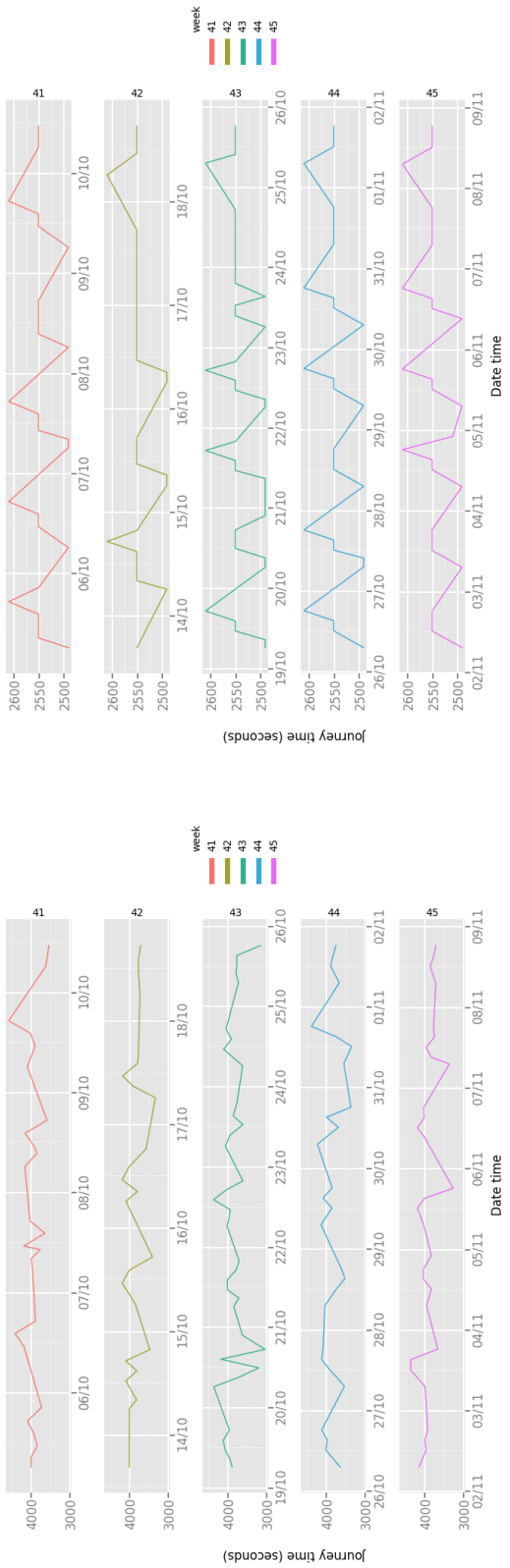


Fig. 3.45 LSOA E01001682 (top), E01002122 (bottom) to St Pancras and journey time distribution (weeks 41-45,2016)

Fig. 3.46 LSOA E01002722 (top), to Heathrow and LSOA E01003482 to City Airport (bottom) journey time distribution (weeks 41-45,2016

Statistical exploration

Statistical analysis is required to provide insights into the variations within this large dataset. As with the road journey times generated via the Directions API, these are broadly defined as :

1. Mean journey time, speed and distance
2. Ratio of maximum to minimum journey time
3. Standard deviation & variance of journey time

1. Mean journey time, speed and distance

In Table 3.14 we see the mean journey times, journey speeds and journey distances for each hub. Heathrow exhibits the longest mean distance, fastest mean speed and longest mean journey time. Despite exhibiting the fastest mean speed, the Heathrow journey's struggle to decrease mean journey times significantly compared to the other hubs as a result of the increased mean distance. These are also shown per hour of the day in Figure 3.47.

Table 3.14 Mean journey times, speeds and distances per hub - public transport

Hub	mean journey time	mean journey speed	mean journey distance
Heathrow Airport	5386 *seconds*	7.07 *m/s*	37.24 *km*
St Pancras International	2670 *seconds*	5.94 *m/s*	16.16 *km*
London City Airport	3785 *seconds*	5.87 *m/s*	22.34 *km*

2. Ratio of maximum to minimum journey time.

A probability density function illustrating the ratio of maximum to minimum journey times is shown in Figure 3.48, for the three different hubs. Distinct behaviour is evident for the three hubs, with Heathrow exhibiting the narrowest distribution with a medium of approximately 1.8. St Pancras and the City Airport show fatter distributions with longer tales of increased maximum to minimum journey time ratios. It is evident from all of these plots that there are large possible changes in journey times for static origin and destination pairs.

In Figure 3.49 the impact of the traditional weekend engineering works is assessed to see the impact this has on average speeds by each day of the week. A noticeable dip occurs across the weekend for journeys to Heathrow and a more subtle impact is seen for St Pancras and the

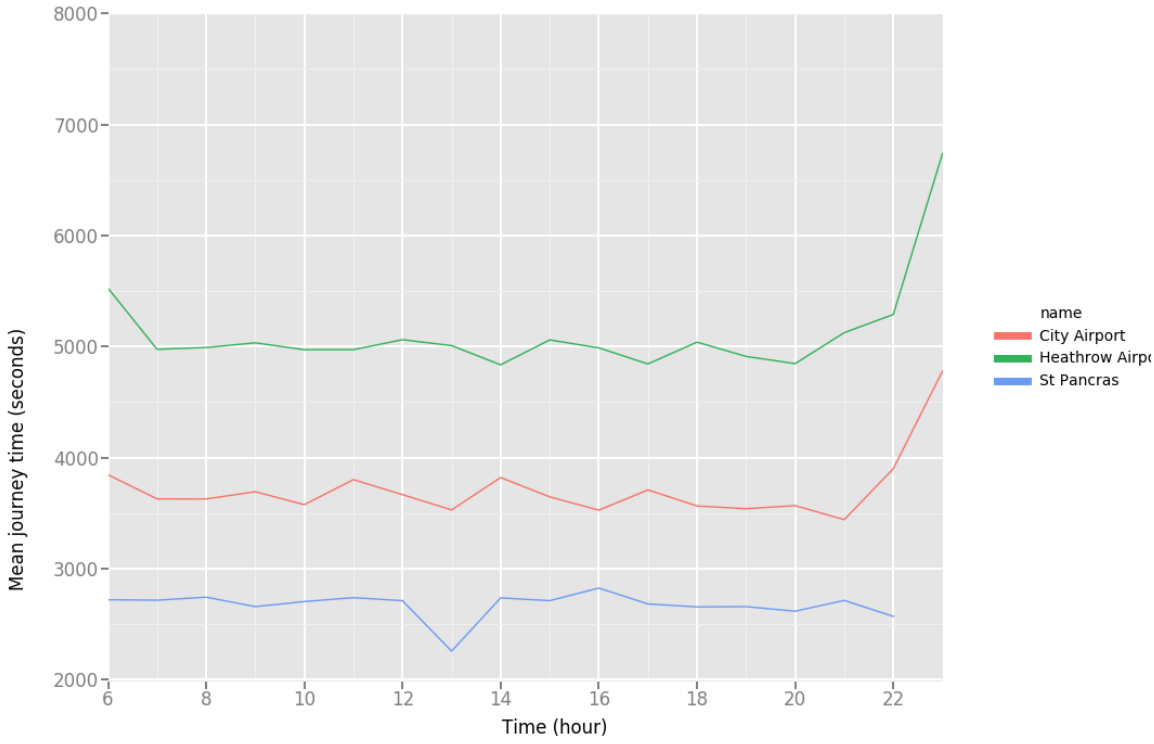


Fig. 3.47 Mean journey time, per hour and per hub

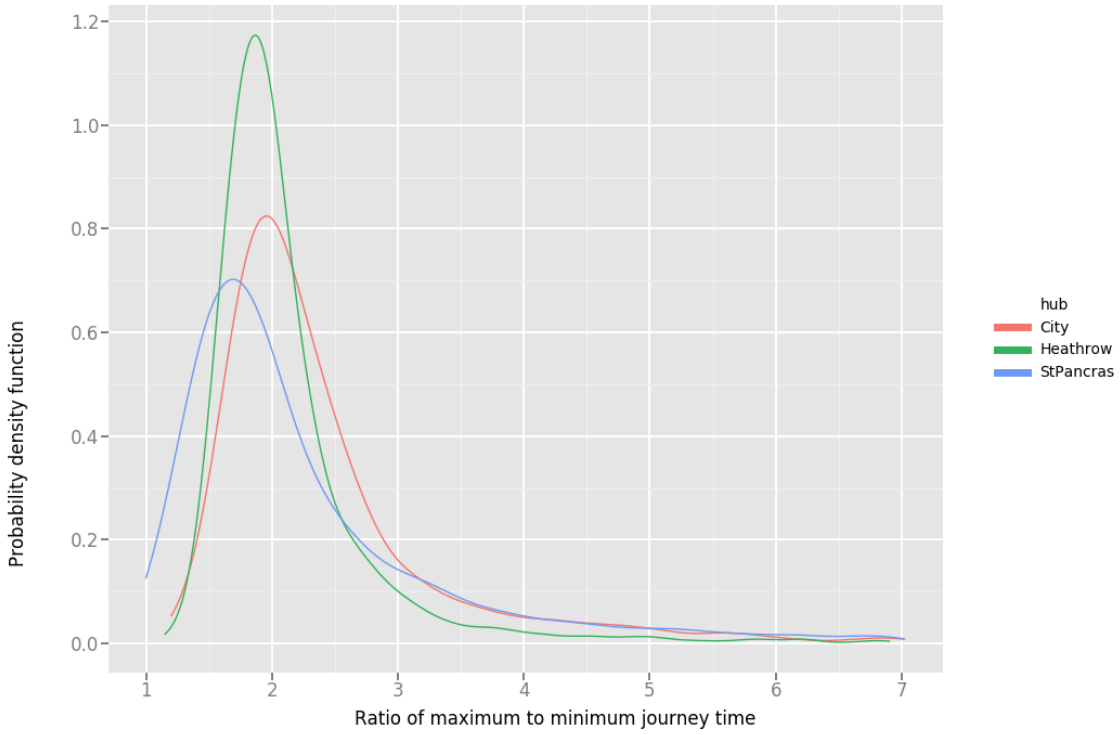


Fig. 3.48 Probability density function for Ratio of maximum to minimum journey times, per mode (0.95 percentile)

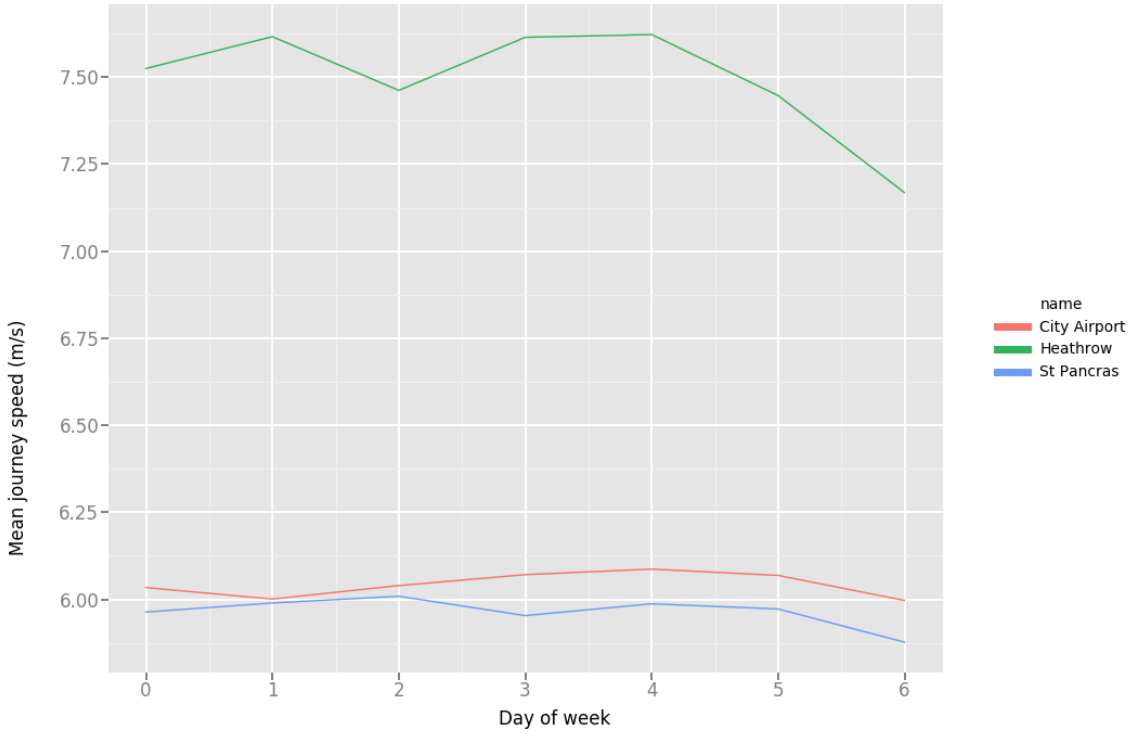


Fig. 3.49 Average journey speeds by day of week and by hub

City Airport. For Heathrow this mean speed difference equates to approximately 0.5 m/s from Friday to Sunday, before recovering again on the Monday.

3. Standard deviation & variance of journey time

The journey time distributions and maximum to minimum distributions have illustrated the variability in journey times for routes on public transport. In order to understand more fully the underlying dynamics a probability density function for the journey time variance and journey speed variance may be plotted for each of the hubs. As was shown before in the ratio of maximum to minimum journey times, the public transport network exhibits much larger fluctuations than that of the road network. As a result, there are very significant journey time increases that occur with more probabilistic chance than that for driving. In order to display meaningful comparisons between hubs, these plots were constrained by those results which fall within the 80% quantile. In the following section where comparisons are made between modes, such a constraint is removed.

In Figure 3.50 the time variance plot, constrained to those results which fall within the 80% quantile is shown. Heathrow shows the greatest variance and St Pancras shows the least, with the City Airport falling in between. St Pancras shows a dual peak without obvious cause. Similarly to driving, the large variance distribution for Heathrow may be partially attributed to its West location in relation to the other more centrally located hubs. The speed variance plot in Figure 3.51 shows more clustering than that of the journey time plot and more of a tail within the 80% quantile.

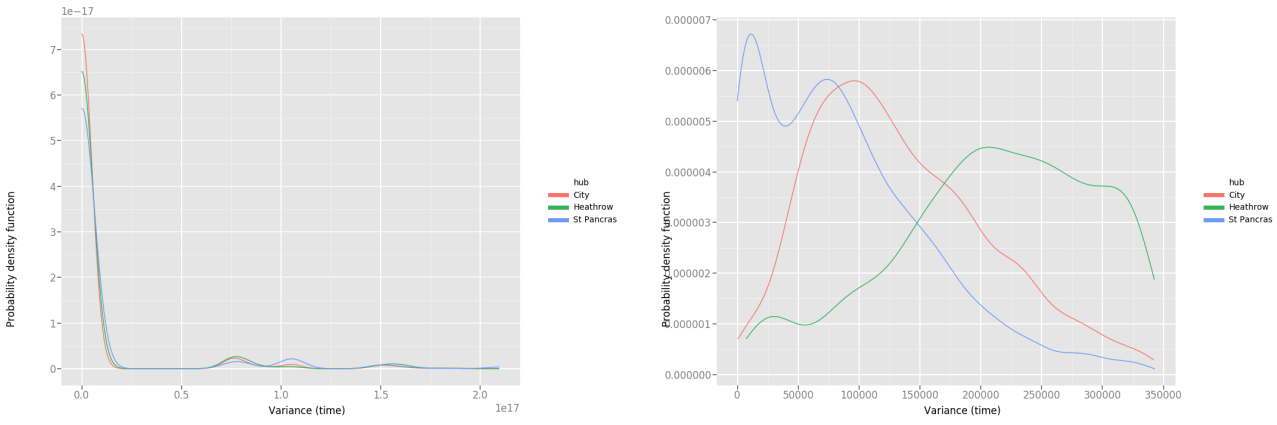


Fig. 3.50 Time variance density plot, per hub. left (0.95 percentile) and right (0.75 percentile)

Propensity for diversion

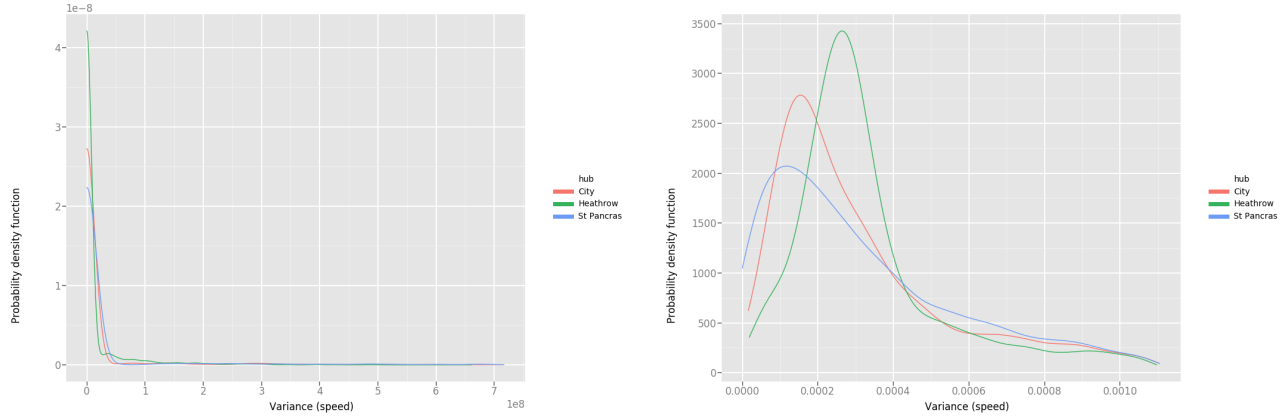


Fig. 3.51 Speed variance density plot, per hub. Left (0.95 percentile) and right (0.75 percentile

A static origin and destination pair do not always output the same path at all times. In some cases, there may be multiple possible routes with similarly competitive journey times. By considering the polyline for a given response from the Directions API we can assess the propensity for for different destination hubs. In Figure 3.52 the probability density function for different unique polyline counts is shown, disaggregated by hub. Heathrow and London City show similar behaviour, with he majority of origin and destination pairs exhibiting less than 20 possible routes. Heathrow shows a different distribution, with a mean of around 40 possible routes. When the number of unique polylines is considered against the crow flies distance (Figure 3.53) a strong positive correlation is shown. St Pancras and London city are shown to exhibit the lowest number of routes and direct distances when compared the Heathrow.

3.3.4 Conclusions

This section has presented 3 distinct data sources. The TfL status feeds permit for a better understanding of the Underground lines and buses which form a significant part of the London transport network. The provision of this data, to aid and inform public transport users, may be harvested and used to derive a range of real-world metrics for the system. Secondly, the GTFS may be used to quantify macro, network level services. The merging of the GTFS with the real-time status feeds allows for the creation of network level information with much finer temporal resolution data. Such a representation may be queried as a service via the Google Directions API. The robustness of the Directions API for public transport queries illustrated a slightly higher failure rate than that for the driving version of this API (0.14% versus 0.002%), reflecting the relative increase in complexity. The temporal limitations of the GTFS have long been known and the specification of GTFS real-time is targeted specifically at

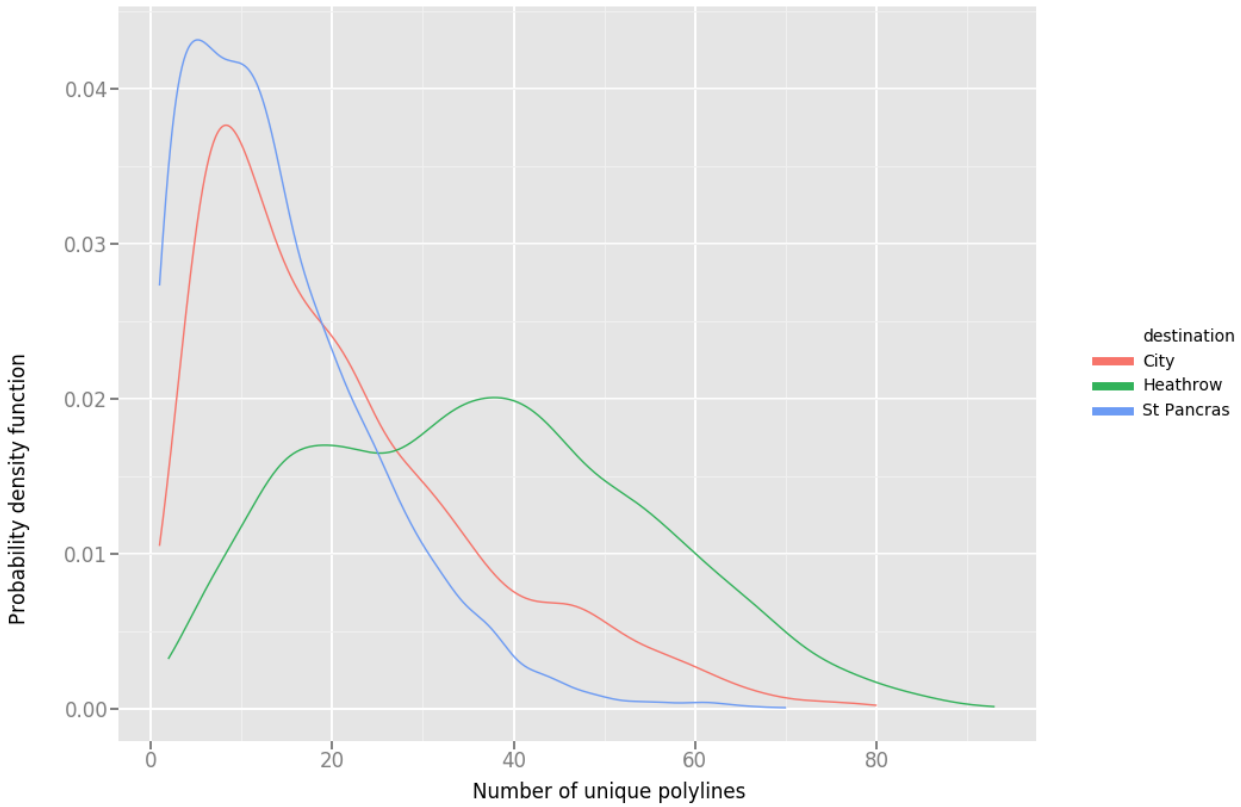


Fig. 3.52 Probability density function for number of unique polylines, per hub, by public transport

Fig. 3.53 Scatter plot for number of unique polylines against direct distance between origin and destination, per hub, by public transport

the challenges highlighted in relating the real-time feeds to the network representation. Despite this formalisation in the form of the GTFS real-time feed, uptake has been slow by transit providers likely due to the complexity of implementation. The GTFS dataset presently doesn't permit for the inclusion of any form of capacity data and thus the impacts of crowding are not measured within this methodology. However, there are now examples of this occurring in specific cities (Google, 2018) and it may be considered in the near future.

Crucially, this section illustrated the challenges faced when assuming fixed journey times for given origin and destination pairs. Public transport journeys illustrated significant fluctuations and variability.

3.3.5 Summary

1. Hypothesis

   London specific travel statistics may be computed from the TfL feeds and GTFS data. Critically, a representation considering both of these distinct sources may be queried through the Directions API and fine resolution temporal and spatial journey data generated.

2. Novel contributions

   These data feeds have been shown to illustrate varying degrees of connectivity at different time stamps. It is known that travellers make use of highly context specific information when making a modal decision. By harvesting the real-time and weekly planned GTFS data it is possible to reflect this.

3.4 Chapter Summary

This Chapter presented a range of different methods for computing fine temporal resolution performance metrics for road and public transport infrastructure. It was hypothesised that the bulk harvesting of fine resolution, individual level detail across space and time may generate a dataset of use to the macro level, transportation practitioner. As with many recent advances in remote sensing, this follows the theme of utilising tail-pipe data which is generated and intended for very different applications for a higher level analysis. These methods may be of interest to users and operators of the respective transport modes. Away from this, these methods may be brought together in order to create a multi-modal image of the transportation network in London. Such a representation presents opportunities for use as a model input, enabling a modal transportation model to be informed with realistic, real-world inputs.

In Section one, a method was designed to test if the Directions API may be used to capture fine resolution temporal journey time data across a wide spatial resolution on the road network. This data may be used to categorise different transportation hubs, such as airports and train stations in terms of their relative access at different times of the day, on different days and on different weeks as a result of congestion and other impacts. However, due to a lack of comparable resolution data there is a validation challenge.

In Section two, a new data set was added to the outputs from section one in order to assess how traffic counts may be related to the harvested journey times. The traditional Bureau of Public Roads method was compared against the empirical data and observations made. This analysis in part provided some validation for the Directions API data in section one and new functions were derived in order to better consider local factors in volume delay relationships.

In Section three three distinct data sets were considered in order to better understand public transport variability. Distinct statics were computed from the real-time feeds, more static GTFS data and then the exhibited output from the Directions API at given times. Extremely large variations were exhibited in both the nature of services and their temporal performance.

However, the methodologies discussed here come with varying computational demands that pose unique challenges. An illustration of these data demands is presented in Table 3.14. The following Chapter, 4 will bring these different data sources together into one unified framework so that they may be used in a meaningful, modelling way.

Chapter 4

Framework for a modal choice and assignment agent based model

The value of dynamic models coupled with more spatially and temporally dynamic inputs can only be assessed when a series of technical challenges to its implementation have been addressed. In Chapter 2, the following two distinct computational challenges were discussed:

1. The spatial (graph compute) problem
2. The agent decision making problem

This Chapter presents an agent based modelling (ABM) framework for the modal choice and route assignment stage of the traditional four stage process. An implementation of this framework is given for London, a series of different decision making heuristics implemented and the computational value illustrated through a series of scaling exercises.

4.1 The framework

The data sources discussed in Chapter 3 must be consolidated into a framework from which an ABM may be used to compute modal and route choice outcomes for a range of different scenarios. In Figure 4.1 this is presented at a high level. Functions 1 and 2 illustrate the merging of the spatial and temporal data into the representative temporal & spatial graph. Function 3 is the specification of a simulation seed, namely a set of origins, destinations (traditionally taken as outputs from the trip generation and trip distribution stages or from an activity based model). Function 4 illustrates the individual graph compute for a given agent, in a given time step, which is then presented to the agent decision making process (function 6), with respect to the specified decision making heuristics in function 5. The computed decision may then

be computed against the existing graph representation through function 7, permitting agent interactions. Functions 8 and 9 show how the simulation may be queried in order to understand individual level decisions and how these manifest into macro level outputs. This section begins by introducing the graph data structure which supports the temporal and spatial graph presented in Figure 4.1.

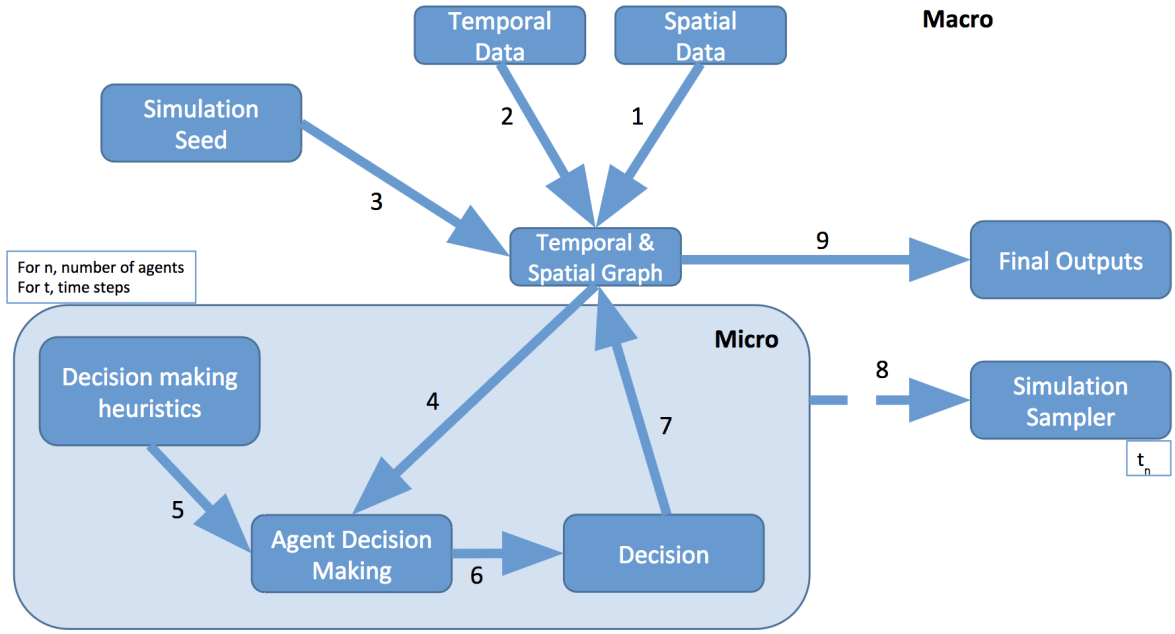


Fig. 4.1 Modelling framework

4.1.1 Graph data structure

Graphs have long been used as an abstract representation that can describe the organisation of a transport system. A graph consists of a set of vertices V that with a set E of vertex edges, such that:

$$G = (V, E) \tag{4.1}$$

In the context of a transport system, a vertex may represent a bus stop, a train station or a road junction and an edge may represent a road or railway line. Beyond the vertex and edge data, there is also relevant polygon data to the modelling process, such as census data or weather data. Fundamentally, the framework must support three distinct types of geospatial data:

1. Vertex data
2. Edge data
3. Polygon data

This geospatial data is either temporally static (e.g. road length) or temporally dynamic (e.g. journey time). Figure 4.2 illustrates these data types and provides illustrative examples. Often, polygon data may be attributed to either edges or vertices, as for example, address information is attributed to a vertex in Figure 4.2.

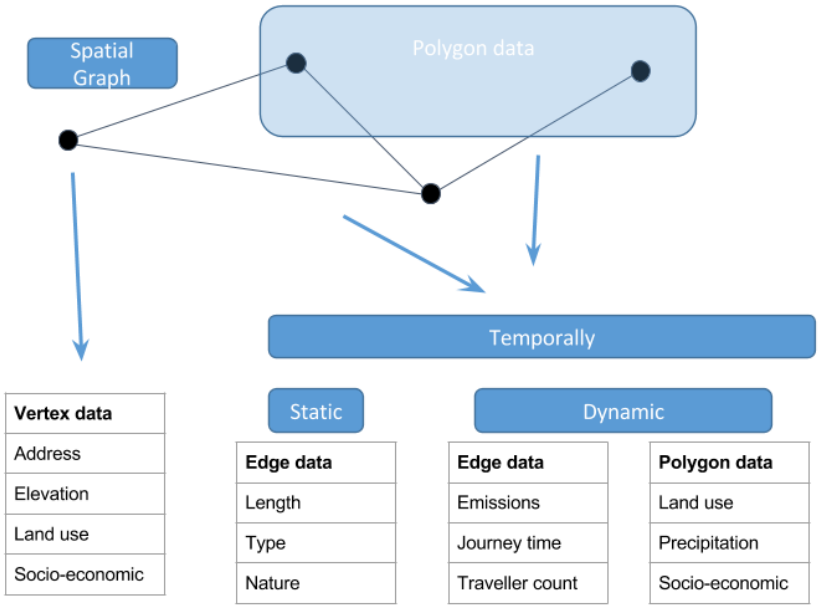


Fig. 4.2 Geospatial data types

A multi-layered graph may be used as a tool to model different levels of the system, for example the road and rail networks and the relationship between these different layers. The transport network consists of a range of modes utilising a range of different physical infrastructure types. These different modes must be distinct in terms of their differing attributes and behaviours but also share connections to permit access to, from and within different sub networks. This may be graphically illustrated as is shown in Figure 4.3 where there are separate road, bike, foot, train and bus networks with various different interchange locations.

The data methodologies described in Chapter 3 must be combined in a similar fashion to that of Figure 4.3 in order to create a spatially and temporally realistic representation of the London transportation network.

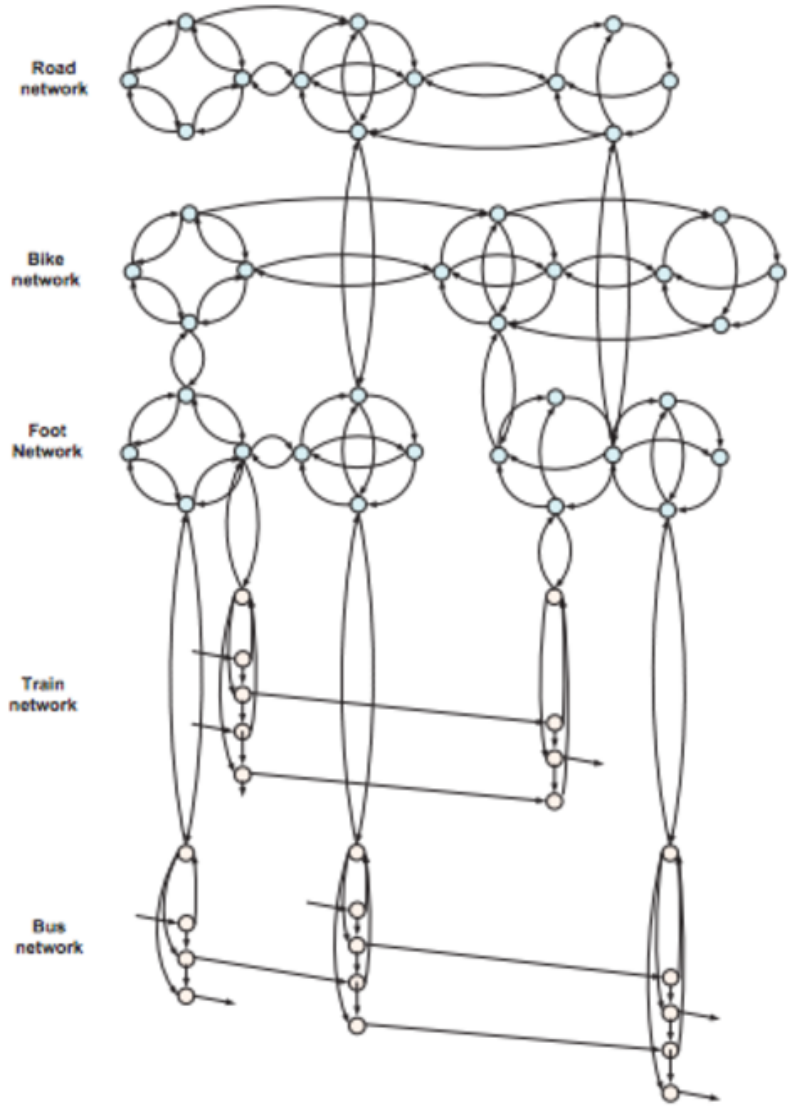


Fig. 4.3 Multi layered transportation graph (source unknown)

4.1.2 Graph data decentralisation

The key principal of the framework is that of decentralisation. Traditionally geospatial data has been stored in a form of relational database with a specialist setup for geospatial data. Such methods have struggled to scale as they do not inherently support the breaking up of large tasks into smaller sub tasks. In relational databases, references to other rows and tables are indicated by referring to their (primary) key attributes via foreign-key columns. In order to compute the interaction between different elements, joins are computed at query time by matching primary and foreign-keys across many rows of the tables. These operations are compute and memory-intensive and have an exponential cost. Relational databases search all of the data looking for anything that meets the search criteria. The larger the set of data, the longer it takes to find matches, because the database has to examine everything in the collection.

JavaScript Object Notation (JSON) is a simple data format that allows programmers to store and communicate sets of values, lists, and key-value mappings across systems (Bray, 2017). JSON is document based and thus scales horizontally, rather than vertically as a relational database does. Thus network data may be distributed across multiple JSON data files, which allows for a decentralised system for querying, data-processing and rendering. The distributed data system allows for easy scalability and load-balancing during computations. A simplified data format from the Sierra-Charlie visualiser (Bak et al., 2016)) is used in this framework.

The fundamental vertex and edge data is separated from their respective attributes in order to permit the loading of core information more efficiently. In the case of a vertex, we have a one to many relationship to attributes. A given vertex will have many different attributes. In the case of an edge, there is a many to one relationship, where multiple edges constitute a group of edges (e.g. a road).

Formally, the data structure used is an adjacency list and consists of the following files:

1. Vertices
2. Edges
3. Vertex Attributes
4. Edge Groups

Vertices and edges describe the fundamental relationships which form the graph and the vertex attributes and edge groups describe attributes for a given vertex and collections of edges (.e.g roads) respectively. This proposed data structure is presented through an illustrative example for London. The used data model with example records is shown in Figure 4.4.

```
vertex = {
    "unique_id" : "osgb4000000031043205",
    "point":[508180.748,195333.973],
    "index": 1
}

vertex_attributes = {
    "unique_id" : "osgb4000000031043205",
    "house_no" : 6,
    "street" : "Hazelbank",
    "locality" : "Croxley Green",
    "administrative_area" : "Rickmansworth",
    "county" : "Hertfordshire",
    "post_code" : "WD3 3EB",
    "country" : "UK",
    "elevation" : 58.424}
}

edge = {
    "negativeNode" : "osgb4000000023183407",
    "unique_id" : "osgb4000000023296573",
    "polyline" :  [492019.481,156567.076...492126.5,156602],
    "positiveNode" : "osgb4000000023183409",
    "index" : 1
}

edge_groups = {
    "group" : "Named Road",
    "members" : ["osgb5000005107792171","osgb4000000023464890"],
    "unique_id" : "osgb4000000023708569",
    "name" : "DAPHNE JACKSON ROAD",
    "index" : 1
}
```

Fig. 4.4 Geospatial data model with example records

4.1.3 Graph implementation for London

An application of this decentralised data file system is proposed for a unified multi-modal graph of London. First, a road network is presented, secondly the various sub networks which form the public transport network are presented and lastly a methodology for unifying these different graphs is presented. An area encapsulating the 33 districts and 32 boroughs plus the City of London, with regions bordering the M25 orbital motorway was defined. The desired outcome is a realistic representation of the entire transportation network in the form of a graph which is embedded, directed, weighted and labelled. This process, with some illustrative examples, consists of the following steps:

1. Fundamental graph - vertices and edges
2. Embedded - addition of geometric positions to edges and vertices
3. Directed - addition of direction restrictions (e.g. one way road restrictions)
4. labelled - addition of various metrics associated with vertices or edges (e.g. post codes or administrative areas)
5. Weighted - addition of temporally dynamic and static attributes to links (e.g. changing journey times and road lengths)

The above processes may be conceptualised, with consideration to the data sources of Chapter 2, as shown in Figure 4.5.

Spatial component - road network

Fundamental embedded graph

The UK Ordnance Survey (OS) Integrated Transport Network (ITN) (OS, 2016) was used as the base map for the road network. The OS ITN is a comprehensive and atomic view of the road network. This GIS data consists of a Geography Markup Language (*gml*) file type, approximately *1.17*GB for the London area. The verbose nature of a *gml* is due to the repetition of common terms and attributes. Beyond the obvious size implications of this verbosity, there is also the complication of inherent centralisation as the file is difficult to separate for parallelisation. A process is therefore employed to harvest the primary geospatial information from this *gml* file and then create a decentralised *JSON* file. This primary geospatial is defined as vertices and edges in Table's 4.1 and 4.2.

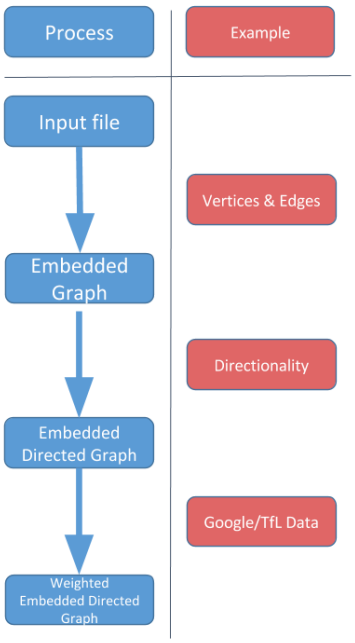


Fig. 4.5 Graph building process with examples

Table 4.1 Vertices data specification

Field	Format
Coordinates	Double precision floating point
Unique i.d.	String ("OSGB" + 16 digit integer)

Table 4.2 Edges data specification

Field	Format
Origin vertex	String ("OSGB" + 16 digit integer)
Destination vertex	String ("OSGB" + 16 digit integer)
Unique i.d.	String ("OSGB" + 16 digit integer)
Polyline	List (Double precision floating point)

This data model can be graphically illustrated at the macro level, as shown in Figure 4.6. The graphical representation of this fundamental road graph provides a visually explicit image, with notable landmarks such as the river Thames, the Lea Valley, Richmond and Hyde Parks. A localised micro view, shown in Figure 4.7, illustrates the fine resolution of this constructed road graph.

The ITN also features road classification information. This information is broadly separated by edge type and edge nature. Of the 423,541 edges on the road network, Tables 4.3 and 4.4

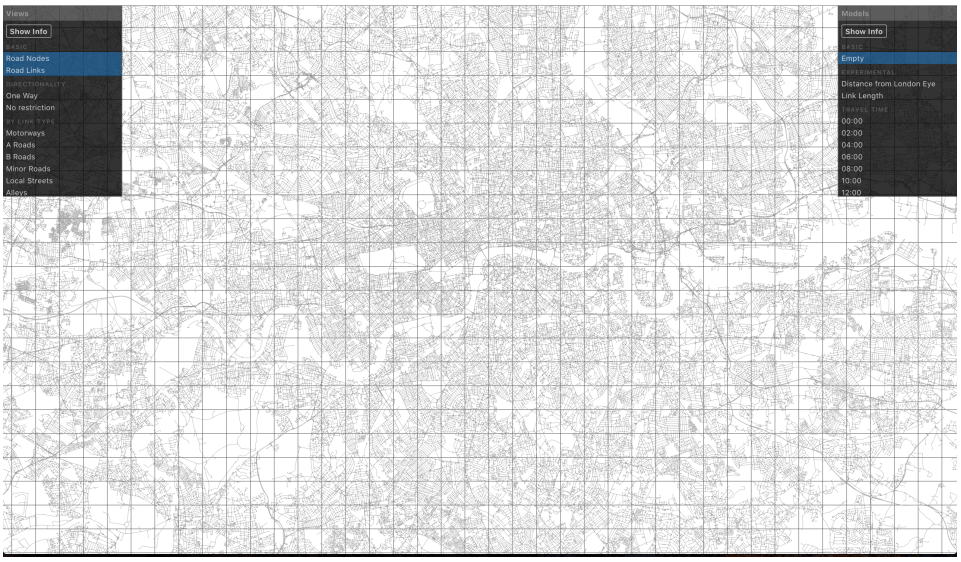


Fig. 4.6 Fundamental components of road graph (macro view)

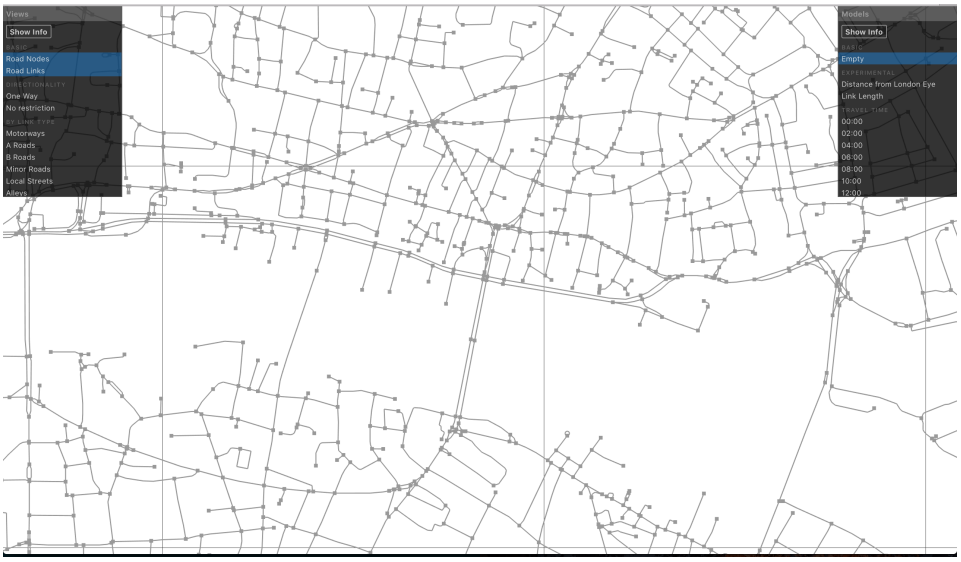


Fig. 4.7 Fundamental components of road graph (micro view)

illustrate the different categories and their proportional representation within the study area. These road edge labels may be visualised and different types and/or natures isolated or viewed in tandem. A selection of views is presented in Figures 4.8 and 4.9 for illustration.

Fig. 4.8 Fundamental components of road graph, aggregated by edge nature. Top - A Roads, bottom - Dual Carriageways

Fig. 4.9 Fundamental components of road graph, aggregated by edge nature. Top - local streets, bottom - Motorways

Table 4.3 Edge terms

Type	Proportional representation in study area
Motorways	0.47%
A Roads	12.63%
B Roads	3.89%
Minor Roads	12.51%
Local Streets	52.61%
Alleys	4.6%
Pedestrianised Streets	0.06%
Private Roads - Publicly Accessible	0.88%
Private Roads - Restricted Access	12.34%

Table 4.4 Edge natures

Type	Proportional representation in study area
Dual Carriageways	4.04%
Single Carriageways	85.99%
Slip Roads	0.85%
Roundabouts	2.86%
Traffic Island Links	1.63%
Traffic Island Links at Junctions	3.95%
Enclosed Traffic Area Links	0.69%

Directed

Local roads are often represented as one edge despite the likelihood of bi-directional operation. Conversely, major roads often segregate traffic in different directions and thus are generally represented as distinct edges in the ITN. The ITN contains road restriction information, as shown in Table 4.5. The spatial spread of those ITN edges which are one way or bi-directional is shown in Figures 4.10 and 4.11.

Table 4.5 Road restrictions

Type	Count
All restrictions	103079
One way restriction	74099

Fig. 4.10 Fundamental components of the road graph - One way restrictions

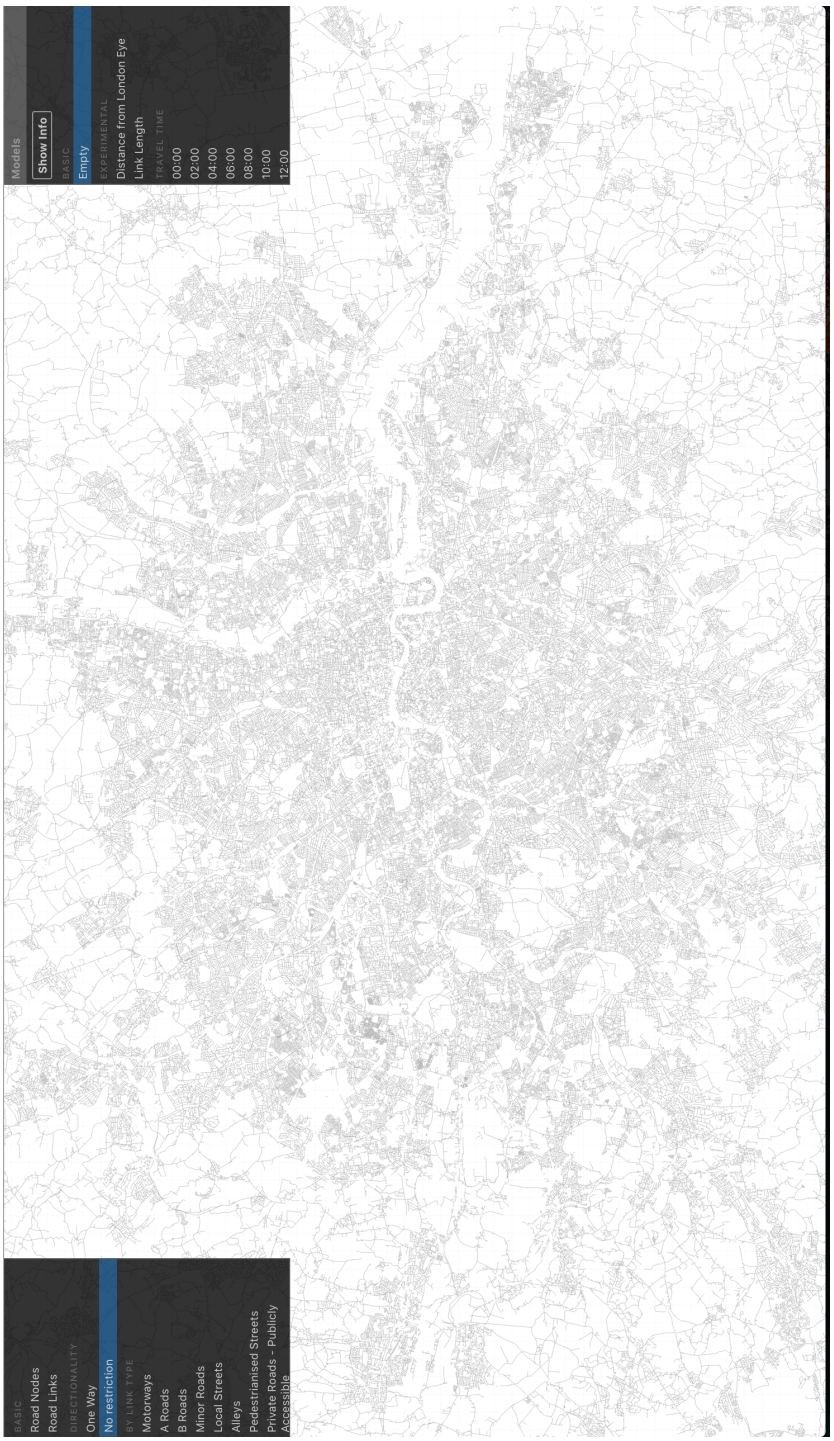


Fig. 4.11 Fundamental components of the road graph - no restrictions

Labelled

A vertex or edge may have a range of different associated attributes. In 4.1 the three different data formats were presented, vertex, edge and polygon. Often, labels such as post codes and addresses are in polygon format. For simplicity the polygon data may be attributed to a vertex (.e.g address) or edge (e.g road name), resulting in a pure graph representation. These attributes are presented for vertices and edges in Tables 4.6 and 4.7 respectively.

Table 4.6 Vertex attributes

Attribute	Format	Source
Address, borough, postcode	String	OS ITN (OS, 2016)
Elevation	Integer	Google Elevation API (Google, 2016b)
Socio-economic metrics	Integer/String	ONS

Table 4.7 Edge attributes

Attribute	format	Source
Road name, type and nature	String	OS ITN
Road length	Integer	OS ITN
Road groupings	List	OS ITN

The result may be graphically represented as shown in Figure 4.12. Here, a sample vertex has been selected and the associated properties of this vertex are shown in the lower right window. Simultaneously, a sample edge has been selected and its associated properties may be inspected in the lower left window.

Weighted

The weight of an edge is a numerical attribute which may explain the length, travel time or financial cost associated with the relationship between two vertices. These weights may be temporally static, (e.g. length) or temporally dynamic (e.g. travel times).

It is at the attribution of weights where the distinct modes of transport which utilise the same infrastructure may be considered. The road network is used by personal vehicles, private taxis, pedestrians and cyclists. These modes share the same infrastructure, with some limitations by type. For example, cyclist's may not use motorways. For the identified HS1 case study public transport and personal car were identified as the possible modes. Public transport is considered holistically as a mode consisting of multiple modes (walking, bus, train, underground etc).

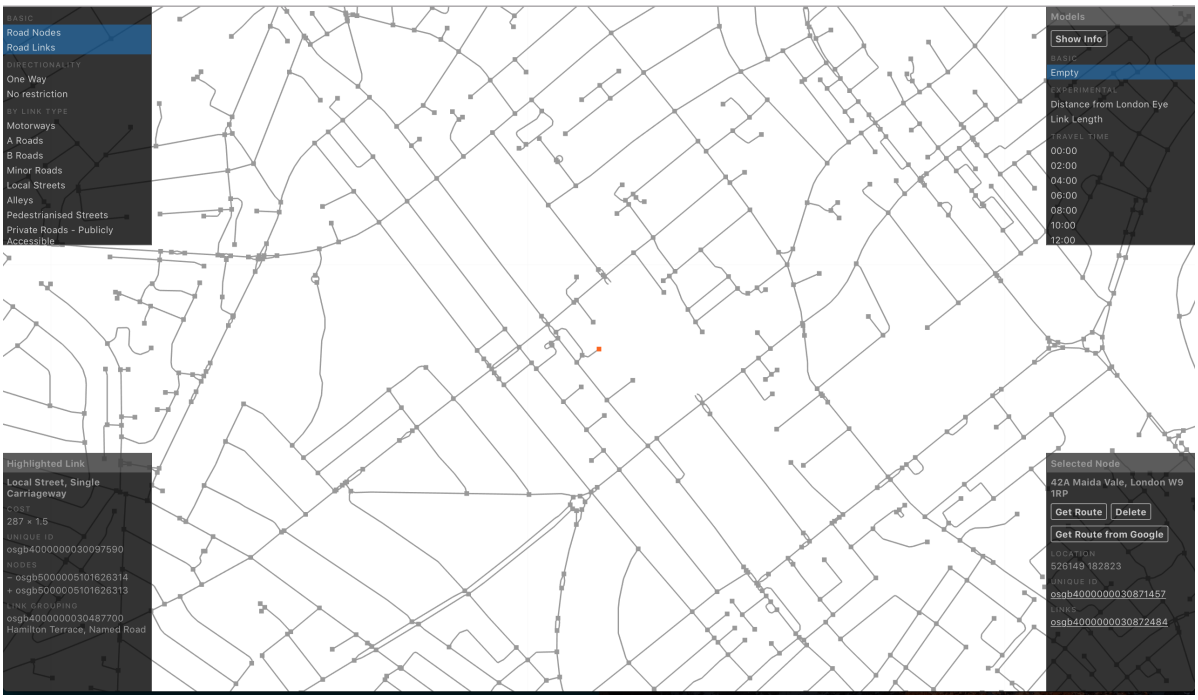


Fig. 4.12 Labelled vertex (bottom right window) and edge (bottom left window) attributes

The first, fundamental attribute describes the physical edge distance from vertex to vertex. This length is rarely as the crow flies and thus consideration must be made to the polyline describing the shape of the connection between vertices. A visualisation illustrating edge lengths by colour is presented in Figure 4.13.

In Chapter 3 the dynamic nature of journey times as a result of traffic congestion was discussed and the use of Google Directions API data to quantify these changes proposed. However, a significant technical challenge is presented when attempting to unify two similar but different geospatial polylines. Google's API responses return geospatial coordinates specific to their underlying representation of London, something which is not publicly available. This is most obviously different from the ITN representation in that it is using a different coordinate projection system [1]. Beyond the relatively simple task of re-projection, there is the challenge of unifying different geospatial representations of the same infrastructure. Consider Figure 4.14 which presents a Google polyline overlaying the ITN road graph. The dark blue polyline illustrates the Google Directions API result and the grey illustrates the underlying ITN network. As is visually clear, there are small discrepancies between Google and ITN's underlying graphs. It is necessary to reconcile this two differing representations of the same physical infrastructure in order to transfer attributes from the Google data (journey times) to the underlying graph. The reconciliation process for differing geospatial data is known as conflation. It is apparent in

[1] Google use epsg:4326 and the ITN uses epsg:27700

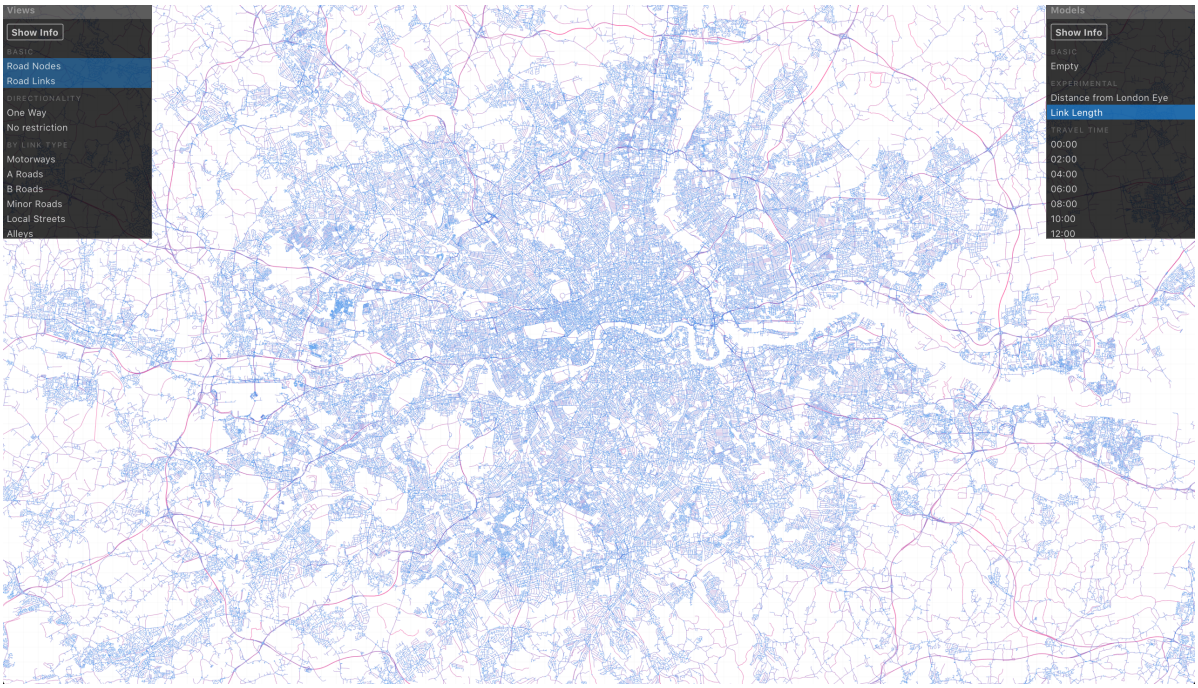


Fig. 4.13 Edge weight visualisation - length

this case that the light blue ITN polyline is the ITN equivalent of the Google polyline result shown. However, a programmatic method for this process is required due to the size of the Google and ITN datasets. This method is discussed in detail in Appendix A. The result of this method is journey times associated with edges, as is shown in Figure 4.15. Each journey time is represented as one key in an epoch time stamped value in the weights dictionary.

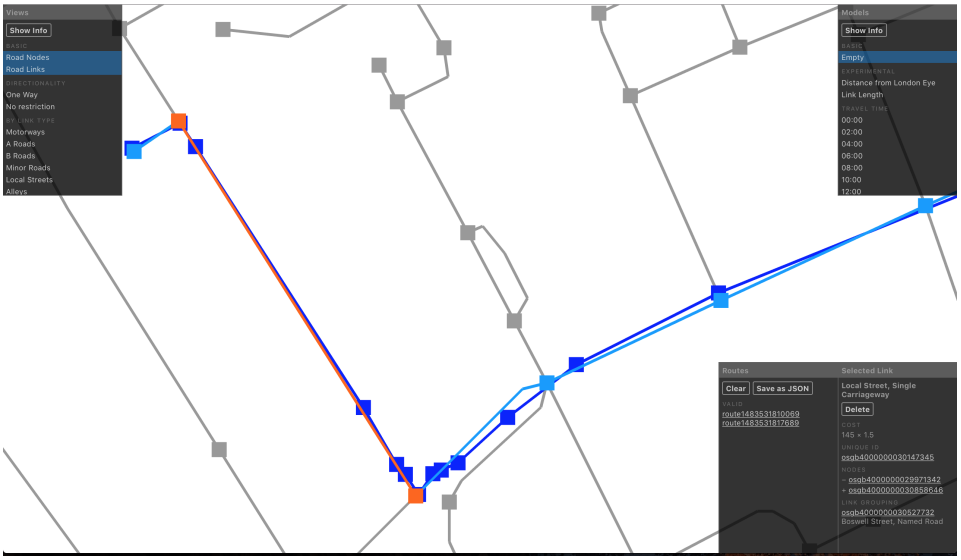


Fig. 4.14 Geospatial differences: Google versus ITN

Of course, edges have time stamped weights dependent on whether or not they are found in the responses from the requested Google Directions API responses. In some cases, such as a major strategic road, many different time stamps may be available, and thus a larger number of records are available. Contrastingly, rural roads may be mentioned sporadically, and thus few records may be available . This lack of consistency in the amount of observations a given edge may have is a product of the requesting formulation used here, the prevalence of certain roads in terms of their strategic importance, plus the impact such strategic roads has on the location device data available for a given road. The matching of the generated datasets to the graph resulted in the distribution of observations per edge shown in Figure 4.16. In this Figure, the road type disaggregation shows that there is a propensity for strategically more important roads to have more observations. Thus, although there is an issue with coverage strictly, there is a useful bias towards those roads which are of more importance.

Since there is not complete coverage, there is a need to populate edges with journey times to cover those times for which it has no records, or in rare occasions, to give journey times in cases where there are no observations at any time. This synthetic journey time, for a given time stamp, may be computed dependent on other available journey times if there are sufficient other records available. A range of different statistical methods were assessed in order to fulfil this function and a first order spline (De Boor et al., 1978) was shown to illustrate the most consistent results with more specific road traffic functions such as (Mullick and Ray, 2012). Example plots are shown in Figure 4.17 where the top image illustrates an edge with many observations, requiring limited interpolation and the bottom image illustrating an edge with fewer observations, requiring more interpolated records.

Estimated vehicle volumes may then be computed by utilising the functions generated in Chapter 3. The process of function specification requires consideration to which derived function(s) are the most appropriate for a given road. Road attributes from the study in Chapter 3 may be compared to the road attributes and the closest pairing be made. This process involves consideration to the factors listed in Table 4.8.

Table 4.8 Edge attributes used in volume-delay function pairing

Factor	example
Road term	A Road
Road nature	Dual Carriageway
Lane count	2

No open access data source exists with lane counts for all roads and this information is not included in the ITN by OS. The study in Chapter 3 required a manual survey to find lane counts

```
{
  "directed": true,
  "restriction": "No Restriction",
  "term": "Local Street",
  "orientation": "-",
  "nature": "Single Carriageway",
  "index": 484764,
  "toid": "osgb4000000030142596-",
  "length": 158.05665472567966,
  "negativeNode": "osgb4000000029946975",
  "graph": "Road",
  "weights": [
    {
      "journey_time": 8.16806583023074,
      "type": "google_GPS",
      "time_stamp": 1443615736
    },
    {
      "journey_time": 7.779918953831298,
      "type": "google_GPS",
      "time_stamp": 1445766611
    },
    {
      "journey_time": 9.269432592014153,
      "type": "google_GPS",
      "time_stamp": 1446899987
    },
    {
      "journey_time": 9.895319430208252,
      "type": "google_GPS",
      "time_stamp": 1446028075
    },
    {
      "journey_time": 10.152466735822882,
      "type": "google_GPS",
      "time_stamp": 1445863815
    },
    {
      "journey_time": 8.095288290905843,
      "type": "google_GPS",
      "time_stamp": 1446632953
    },
    {
      "journey_time": 7.89636301675113,
      "type": "google_GPS",
      "time_stamp": 1445515018
    },
```


Fig. 4.15 Edge with matched Google Directions API journey time weights

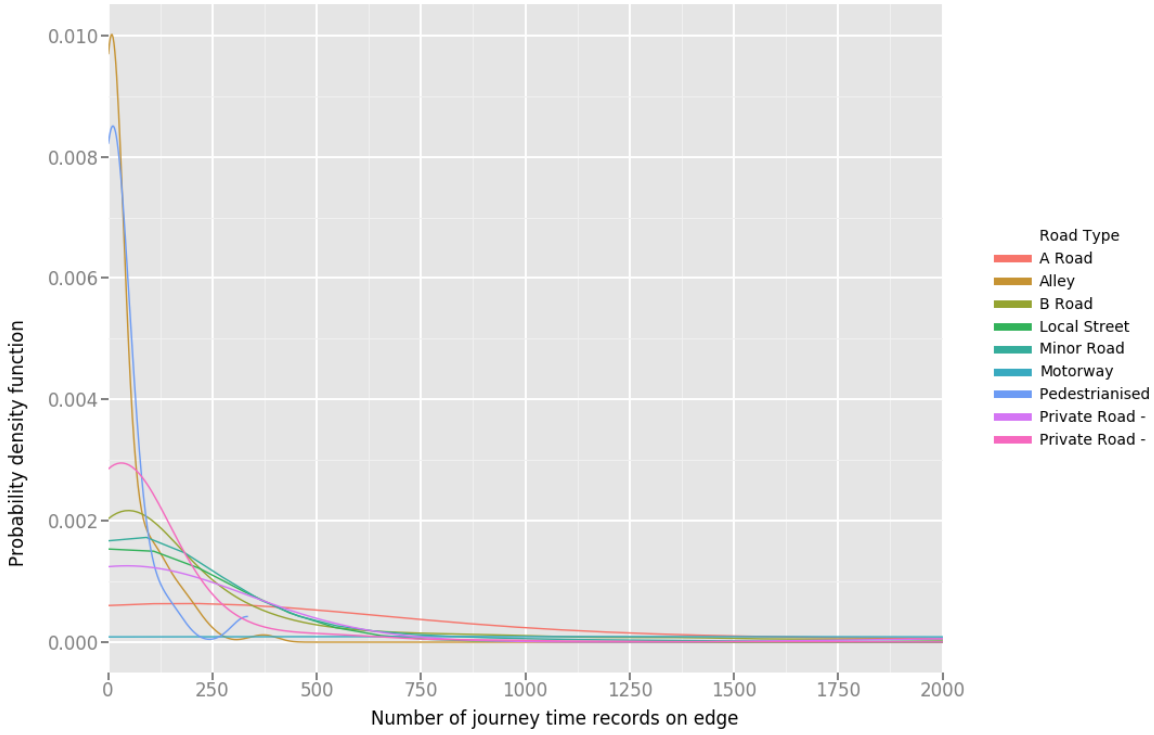


Fig. 4.16 Distribution of journey time observations per edge, by edge type

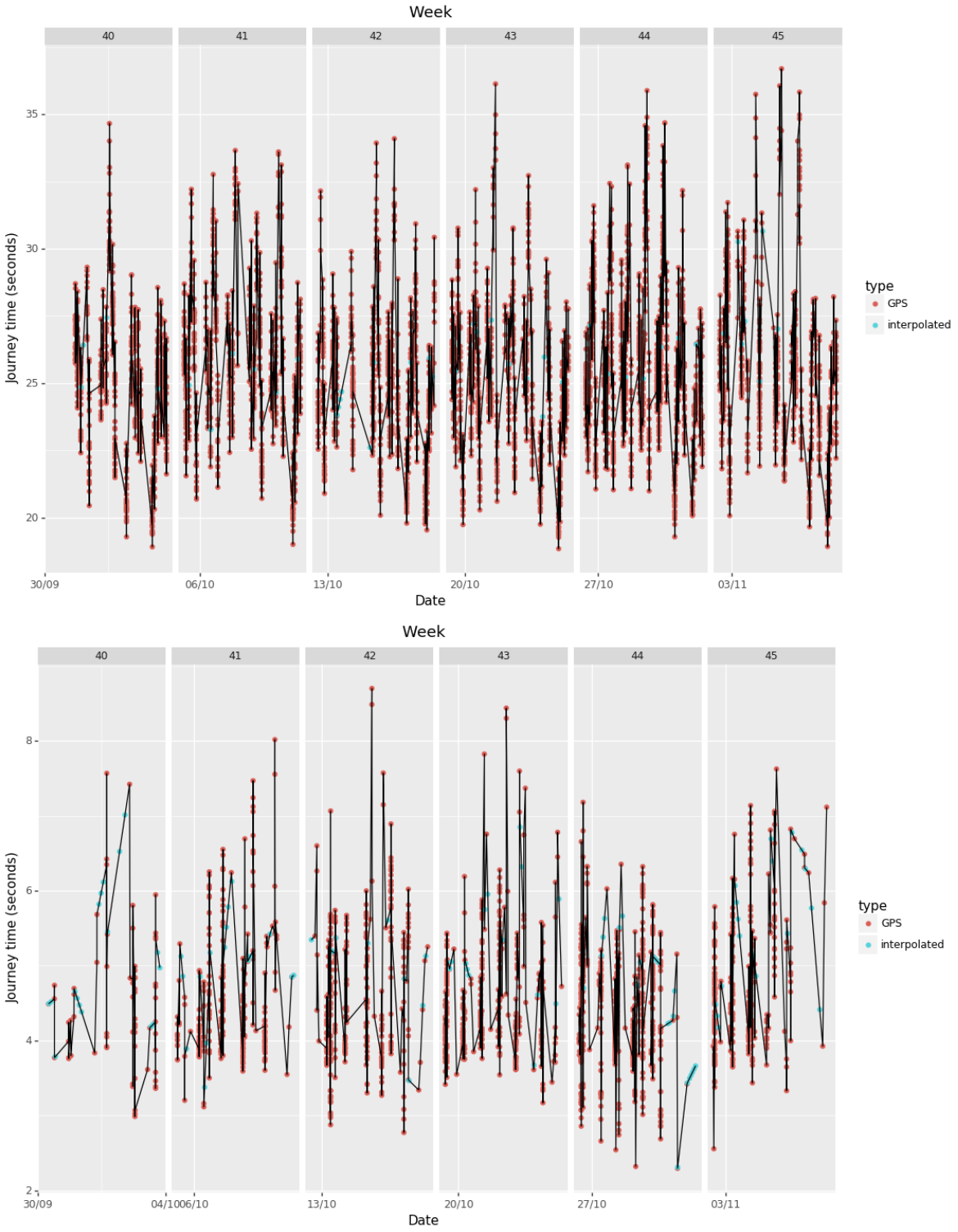


Fig. 4.17 Journey time distribution with interpolated and empirical records, over one month for two edges osgb4000000030779794- (top) and osgb4000000031114655 (bottom)

and from it, general heuristics are derived and used to assume lane counts across London based on road properties such as road term and road nature.

The assumed capacity is assigned by extrapolating on road nature from those empirically measured in Chapter 3. A generalised function for each road nature type was computed from those measured in Chapter 3 and this function is computed against the record delay in order to find the saturation. The saturation can then be used to estimate the traffic volume for the assumed capacity, as is defined below:

$$Volume_t = VolumeDelayFunction * FreeFlowCapacity \tag{4.2}$$

The result is a graph with temporal weights for journey time (either empirical or synthetic) and computed, estimated volumes, as is shown for a sample edge in 4.18. Finally, financial costs may be attributed to driving journeys by using RAC cost coefficients (RAC, 2016). The output of this process is a weighted, directed, labelled and embedded graph for the road transportation network.

Spatial component - public transport network

In Section 3.3 a series of exploratory analyses of the public transport services in London were presented. Real time feeds were assessed to understand the variations in service reliability and the timetable information in the General Transit Feed Specification format (GTFS) was used to understand planned connectivity, by different modes and at different times of the day, week and month. The statistics and analyses presented were generally system level and although they are very relevant to individual travellers, the outputs are not directly applicable. There is a therefore a need to abstract the GTFS format network into a weighted, directed, labelled and embedded graph so that temporally dynamic shortest paths may be computed, in order to integrate with the road network graph and ultimately support modelling modal choice and assignment modelling.

Fundamental embedded graph

Network vertices may be extracted without any manipulation of the GTFS data. Full services were split into constituent part and edges on the network were defined on the sub sections of services which connect a given set of vertices, irrespective of mode [2]. We may use the type disaggregation specified in the GTFS to explore the relative spacial complexity of the different modes within the public transport network. Table 4.9 illustrates the vertex and edge count for each mode individually. Unsurprisingly, the bus network is shown to exhibit

[2]There are occasions where edges are served by multiple modes, e.g bus and tube

```
],
"graph": "Road",
"weights": {
  "capacity": 524,
  "max_speed": 22.49576044109427,
  "max_journey_time": 2.2921177248460527,
  "lane_count": 1,
  "min_journey_time": 0.6107115265222953,
  "min_speed": 5.993767270476157,
  "time_stamps": [
    {
      "volume": 1607,
      "time_stamp": 1446143959,
      "journey_time": 1.8794254015,
      "measurement": "GPS"
    },
    {
      "volume": 1106,
      "time_stamp": 1445846654,
      "journey_time": 1.2398941535,
      "measurement": "GPS"
    },
    {
      "volume": 1716,
      "time_stamp": 1446143756,
      "journey_time": 2.0189286012,
      "measurement": "GPS"
    },
    {
      "volume": 1603,
      "time_stamp": 1446143960,
      "journey_time": 1.8750303369,
      "measurement": "GPS"
    },
    {
      "volume": 1621,
      "time_stamp": 1446221205,
      "journey_time": 1.8969184124,
      "measurement": "GPS"
    },
    {
      "volume": 1544,
      "time_stamp": 1446134303,
      "journey_time": 1.7987102479,
      "measurement": "GPS"
    },
    {
      "volume": 1523,
      "time_stamp": 1446738357,
      "journey_time": 1.7722586266,
```


Fig. 4.18 Sample edge with computed volumes for a given time slice

the biggest network by a significant margin. The only non GTFS datasets used are the TfL zone information (O'Brien, 2016) and the 2016 TfL pricing information (TfL, 2016a) to permit pricing computations for a given route. Bus, ferry and tram services use a trip based cost, whereas Overground and Underground services use a zone based pricing system. There is also some pricing logic programmed to compute costs for trips across these services.

Table 4.9 London public transport network spatial complexity

Network	vertex count	edge count
Bus	18987	22855
Subway, Metro	270	587
Tram, Streetcar, Light rail	84	152
Ferry	22	56

Weighted

For a given edge, there are then a series of services with attributes such as journey time, departure time, route agency, route id etc. In contrast to the individual nature of car journeys, public transport services are centrally managed and scheduled. Thus, the quantification of journey times on public transport infrastructure is simpler to harvest but more complex in nature to store as there is the the added complexity of different services with different departure times, routes and journey times.

4.1.4 Graph computations

A weighted and directed graph now exists from which shortest path computations may now be carried out in order to act as the inputs to a modal choice and route assignment ABM. This section will consider how the constructed graph may now be queried in an efficient manner, before focussing on the behavioural part of decision making which dictates what information is used and how it is used by a given individual. Ultimately, the specification of parameters for inclusion is linked to the decision making logic of the ABM and consideration must be made to this computation to ensure that there is consistency between the inputs and the process which utilises the inputs.

The minimum cost or shortest path route problem is defined as the process of identifying the lowest cost route from an origin to a destination usually in terms of distance, journey time or by a combination of graph edge attributes. A large amount of literature exists in the fields of routing and scheduling problems. Developments have occurred since Dijkstra presented

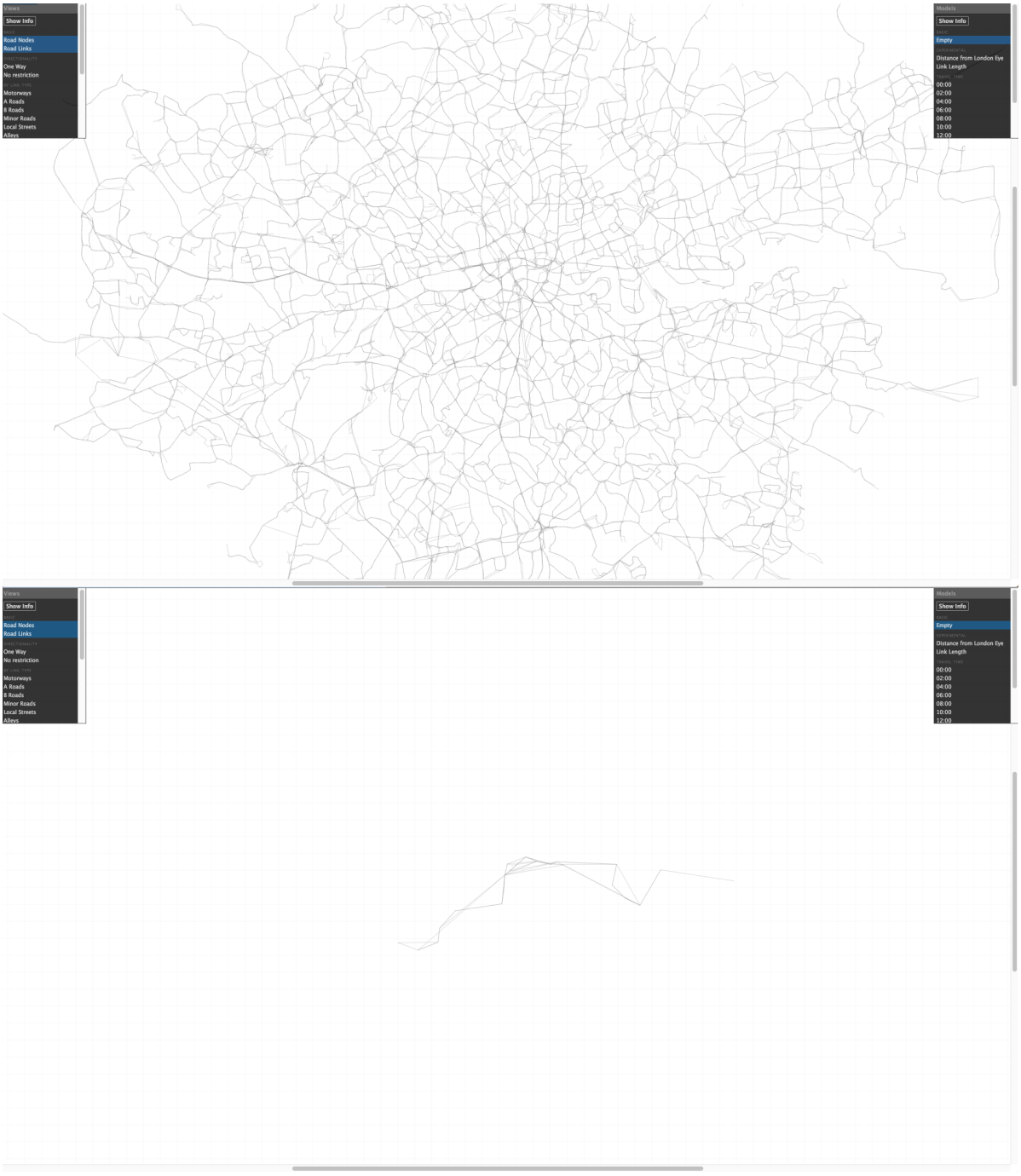


Fig. 4.19 Fundamental component of public transport graph. Bus network (top), ferry network (bottom)

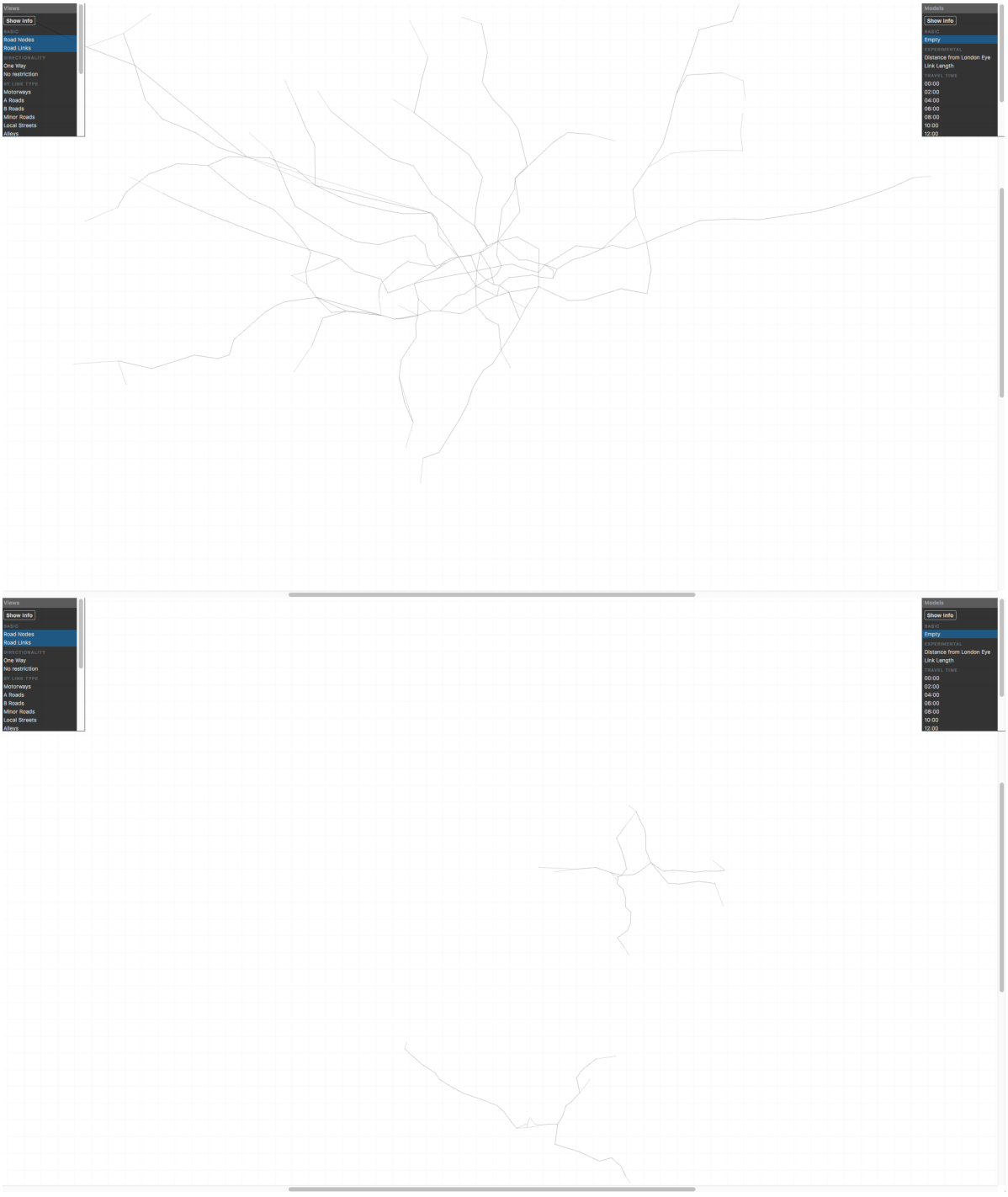


Fig. 4.20 Fundamental component of public transport graph. Rail and tube network (top), tram and light rail network (bottom)

```
{
  "negativeNode":"490010167W",
  "positiveNode":"490004661W",
  "edge_id":"490010167W490004661W"
  "services" :

    [
      {
        "route_id":"370",
        "route_type":"Bus",
        "journey_time":60,
        "arrival_time":1446442560,
        "route_agency":"tfl",
        "departure_time":1446442500
      },
      {
        "route_id":"370",
        "route_type":"Bus",
        "journey_time":60,
        "arrival_time":1446442560,
        "route_agency":"tfl",
        "departure_time":1446442500
      },
      {
        "route_id":"370",
        "route_type":"Bus",
        "journey_time":60,
        "arrival_time":1446443760,
        "route_agency":"tfl",
        "departure_time":1446443700
      },
```


Fig. 4.21 Example public transport edge record

his path finding algorithm in 1959 (Dijkstra, 1959). A range of algorithms and software packages may now be taken off the shelf for a range of graph problems. Driving features a large spatially complex graph with temporally dependent edge journey times, directionally restrictions. Public transport is less complex spatially but more complex temporally due due to timetabling, variable service routes and a truly multi-modal nature resulting in a series of sub graphs. The walking sub graph on the road network connects rail stations, bus stops and stations. For a given time horizon, the network connections (edges) may be represented on a frequency basis to simplify computations but still allow for temporal variations (e.g. morning peak versus inter-peak). Ultimately, the minimum cost path involves a behavioural decision on human value judgement and this framework seeks to support a range of systems rather than pre-prescribe one. As was discussed in section 2.4.4, many modern shortest path algorithms work on unweighted or undirected graphs (e.g social networks) and so many of the classic, early graph algorithms are still in use day, albeit in modified forms. In Table 4.10 different weighted shortest path algorithms are presented with their respective worst case complexities in Big O Notation (Danziger, 2010).

Table 4.10 Shortest path algorithms and complexity

Algorithm	Complexity [3]	Reference
Bellman-Ford	O(VE)	(Bellman, 1958)
Dijkstra	$O(V^2)$	(Dijkstra, 1959)
A *	$O(\|E\|) = O(b^d)$ [4]	(Hart et al., 1968)

The A* algorithm uses heuristics in order to reduce computational times and in fact, has the same complexity as Dijkstra when no heuristic is used. Despite the significant performance gains, it is harder to implement than Dijsktra and its implementation can often become context specific to a given graph and not generalisable. Bellman-Ford is slower than Dijkstra, but unlike Dijkstra it can handle negative edge weights. This functionality is usually useful for pricing scenarios (often in aviation) where multiple edge combinations may be discounted and where this discount manifests itself as a negative edge weight. However, TfL's public transport pricing does not require such a model to compute trip fares and as a result Dijkstra's algorithm was utilised here. The shortest path computation is carried out using Dijksra's algorithm and using edge journey time weights as the edge parameter. The shortest path found and it's respective attributes may then be presented for a specific agent within the ABM, as was graphically shown previously in Figure 4.1.

4.1.5 Agent decision making heuristics

In Chapter 3 a range of data sources were presented with the ambition of capturing real-world transportation service performance. This involved the harvesting of location device informed journey times and real-time feeds from transport operators such as TfL. In previous sections of this Chapter, these different data sources were unified in a single framework - a multi-modal graph data structure. This framework supports an agent based model which can compute decisions for a given traveller in a given context. The focus of this section is what and how such data is treated by an individual traveller and the process that occurs culminating in a modal and route decision. This section illustrates the use of the generated multi-modal graph in a range of decision making heuristics. As was discussed in Chapter 2, there remains a great deal of uncertainty surrounding, firstly, what information people use when making decisions and secondly, how they use this information. In the literature a range of decision making methodologies were discussed. This section involves the specification of how the agent carries out the decision making process according to different pieces of literature.

Until relatively recent times, the dominating economic paradigm has been constructed upon the belief that an individual acts rationally to maximise their utility (Simon, 1955). This states that an agent will maximise utility by minimising cost. In the context of a purely monetary driven environment, this is easily conceptualised. However, a transportation modal decision involves a wide range of considerations beyond simply the financial cost. This may involve those metrics easily quantified such as financial cost and time cost but also those less easily quantified such as traveller comfort or stress and how pleasant the surroundings of a given journey may be.

A historic limitation for decision making frameworks has been realistic inputs. Often a static, one time input is used to quantify the performance for a given piece of infrastructure. This may be empirical data from a road survey which is often carried out over a short period (such as one day) and then generalised or data from general functions. These inputs are generally relatively static and do not offer realistic time dependent data. The fine resolution spatial and temporal data discussed in Chapter 3 can first be used to give improved inputs for simple decision making frameworks and then secondly, enable more complex decision-making frameworks to be utilised with a view to capturing more context specific decision-making processes that reflect the changing nature of an evolving complex transport system.

In this section, a range of decision making heuristics are implemented and the macro level results of their implementation assessed. For the purposes of testing this, a sample dataset formed of 200 random origins were selected for the three destinations of St Pancras International, London City Airport and Heathrow from the multi-modal graph.

Journey times

Instead of considering the average journey times across the entire time horizon, we consider each time slice as an individual event, and illustrate the changing competitiveness of the different modes. Secondly, we may disaggregate the journey times and weight them dependent on their mode. In this case, coefficients (Table 4.11) from the (Mackie et al., 2003) study, which are adopted by the DfT WebTAG guidance are used.

Table 4.11 Journey time type perceived cost coefficients

Mode	Cost	Reference
Walking	1.75 * in-vehicle time	(DfT, 2014b)
Waiting	2.5 * in-vehicle time	(TfL, 2016b)
Interchange penalty	7.5 minutes	(DfT, 2014b)

In Figure 4.22 we can see how the relative competitiveness of the two modes varies depending on the use of input times, or computed perceived times using the coefficients from Table 4.11. For the given sample, it is evident that driving exhibits the dominant journey time savings, with a greater frequency and greater range of savings over public transport. When the perceived time cost is considered, this difference becomes ever more extreme, with more journeys becoming faster by driving and the range increasing even more.

The above plots illustrate the macro, averaged output, however the resolution of the input data makes it possible to assess each time step discretely and view the distribution of varying competition between the two modes across time. A sample of one week is extracted in order to illustrate how this varies by time of day and by day of week over the study period and is shown in Figure 4.23. It is clear that there are significant daily fluctuations in the relative competitiveness of each mode at different times of the day, with driving losing prevalence at peak morning and evening peaks. This dynamic response is a direct result of traffic impacts on driving increasing journey times, whilst public transport timetables at peak times offer reduced journey times and increased service frequencies.

Generalised cost

Journey times are a significant factor in modal and route choice but there are factors, such as financial cost which are highly relevant. Financial cost may be attributed to driving journeys by using RAC cost coefficients (RAC, 2016) and TfL emissions charges and parking costs may be included. For public transport journeys the far information previously loaded into the

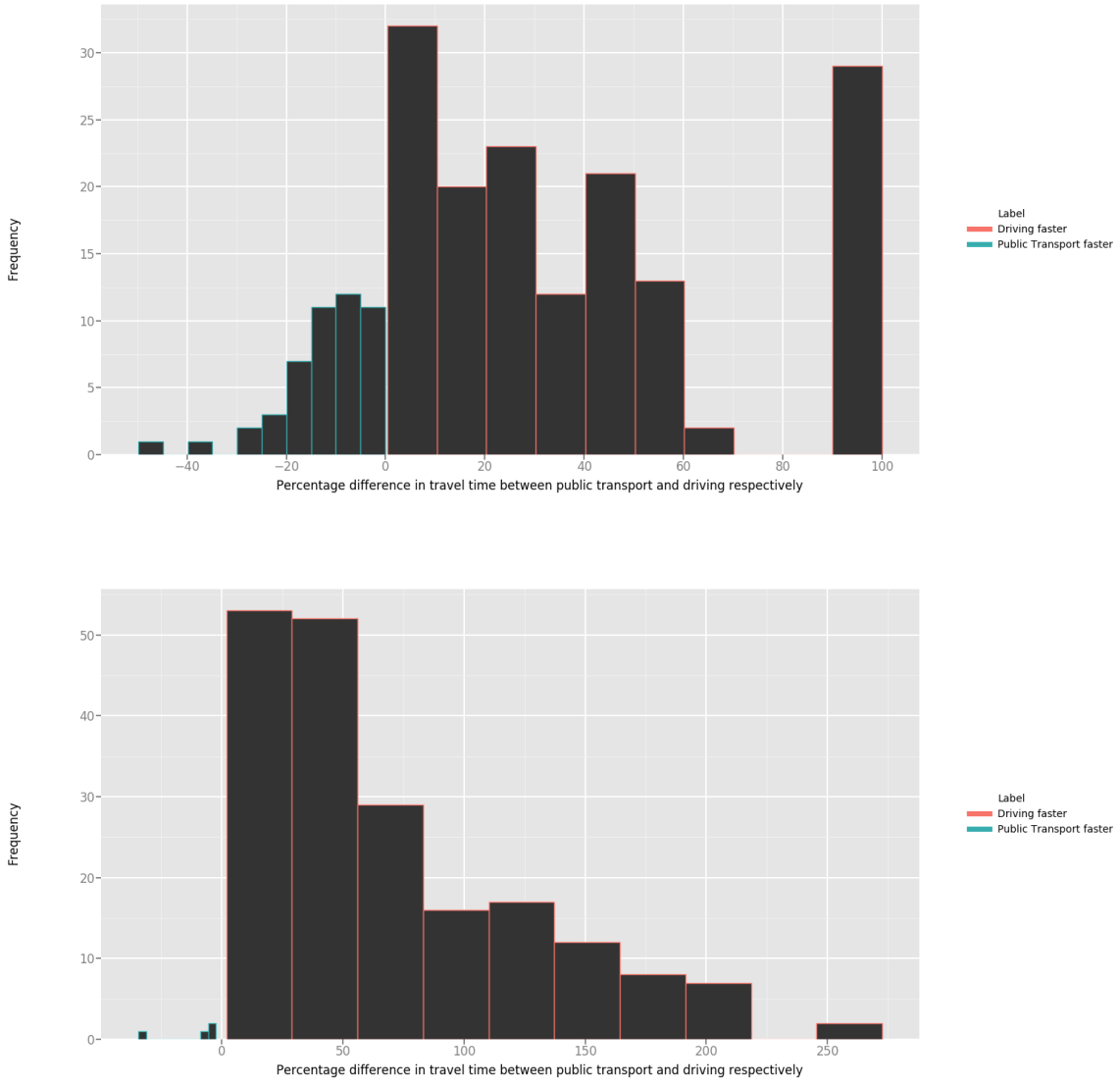


Fig. 4.22 Percentage journey difference difference between driving and public transport, unweighted (top) and weighted (bottom). The weighted factor methodology is shown in Table 4.11

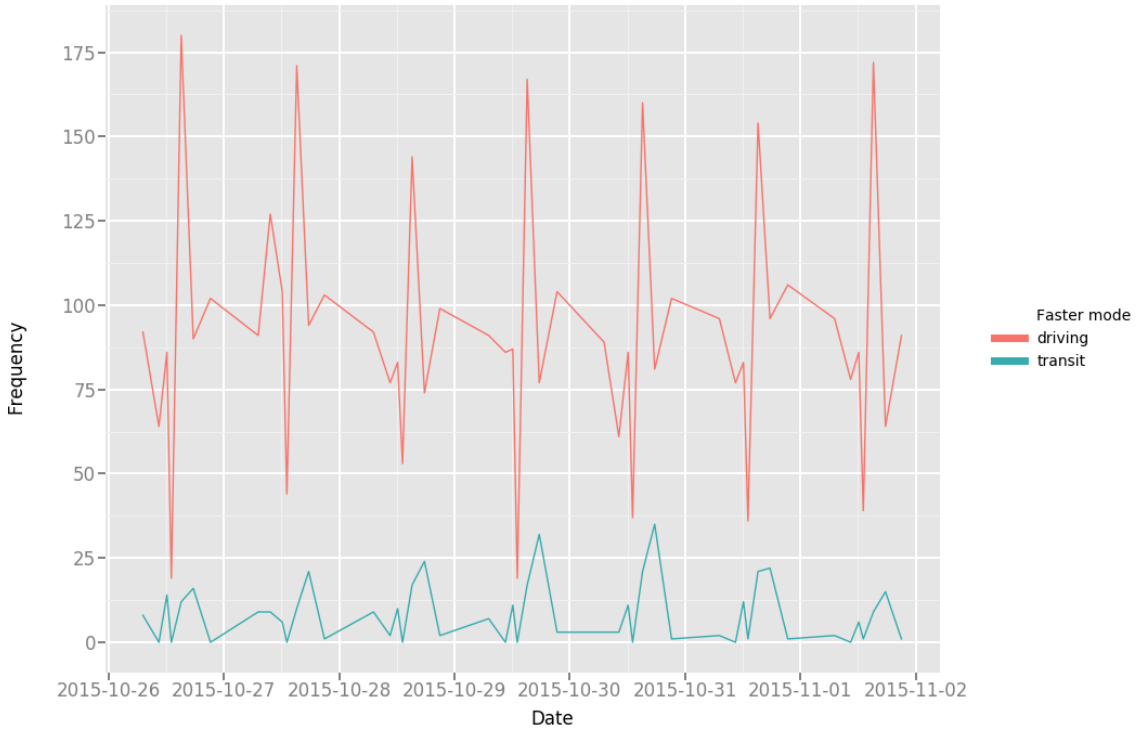


Fig. 4.23 Journey time comparisons of driving versus public transport at discrete times across one week in October 2015

framework may be used. In London, bus and tram journey's are charged a flat rate per trip [5] and in the context of rail and tube journeys, these are charged based on the zone of the arrival and destination. When the tangible financial costs have been included, the time costs must be converted into a financial cost in terms of perceived value of time (e.g (Harrison and Quarmby, 1970), (Wardman, 2004)). WebTAG guidance on value of time in the UK was updated in 2015 (DfTt, 2015) and includes values disaggregated by mode, distance and role (commute, other non-work and business travel). Since business travel allowed for disaggregation by mode and distance, it was used.

The factors shown in Table 4.12 were utilised from this study.

Table 4.12 Value of time coefficients (from (DfTt, 2015))

Mode	Value of time (£/hr)
Driving (< 20miles)	8.21
Driving (20-100 miles)	15.85
Driving (> 100 miles)	25.74
Bus	15.64
Underground & Tram	24.72
Rail (< 20 miles)	27.61
Rail (20-100 miles)	10.11
Rail (> 100 miles)	28.99

This then enables the generalised cost to be computed for each mode, in the following form:

$$\begin{aligned} DrivingGeneralisedCost = {} & DrivingTime + \\ & (DistanceDriving * OperatingCost / NumberOfPeopleTravelling) \\ & + (ParkingCost / NumberOfPeopleTravelling) + \\ & CongestionTax \end{aligned} \quad (4.3)$$

$$\begin{aligned} PublicTransportGeneralisedCost = {} & ((WalkingTime * WalkingWeight) + \\ & (WaitingTime * WaitingWeight) + \\ & (InVehicleTime) * ValueOfTime) + Fare \end{aligned} \quad (4.4)$$

[5] This has recently changed with the introduction of a "Hopper Fare" which permits bus transfers within one hour at no additional cost

Figure 4.24 illustrates two sample records, one for driving and one for public transport, with their respective computed generalised costs. In the case of public transport we can see the constituent parts of the total generalised cost, as computed per mode and type of that mode.

Since the generalised cost is heavily dependent on the time inputs, the general trends for this sample shown in Figures 4.22 and 4.23 continue to be apparent. In order to better understand how the different modes compare, the probability density function of their generalised costs is shown in Figure 4.25. Driving exhibits a left leaning distribution with a lower mean that public transport. Interestingly, driving also exhibits the largest range.

Risk based

The data available permits the use of even more nuanced inclusion of the variations in journey times and costs. We may consider the historic performance of the infrastructure in how we value the journey time reliability of a given mode and this information can be used in turn to manipulate the value of time coefficient used (as per (DfT, 2017)).

In the context of driving, reliability may be defined as the ratio of the mean to the standard deviation of travel time (DfT, 2015c), a modified version of (Hollander, 2006), explicitly defined as:

$$DrivingJourneyTimeReliability = \frac{MeanJourneyTime}{StandardDeviationOfJourneyTime} \tag{4.5}$$

Public transport travel time reliability is defined in terms of the timetabled travel time, the relative travel time lateness and the standard deviation of this lateness across time (DfT, 2015c), a modified version of (Hollander, 2006) and may be explicitly defined as:

$$PublicTransportJourneyTimeReliability = \frac{MeanLateness}{StandardDeviationOfLatetness} \tag{4.6}$$

where:

$$Lateness = JourneyTime - TimetabledJourneyTime \tag{4.7}$$

Computing the lateness of a given service, relative to its timetabled journey time is non trivial. The generation of a spatially and temporally dynamic, real-world journey time for public transport networks is only possible by merging temporally explicit, spatially implicit real-time feeds with spatially explicit and temporally inaccurate representations of the network and timetable (e.g. GTFS), as was shown previously in Figure 3.52. The methodology in this chapter has enabled the use of fine resolution GTFS data, but does not include the inclusion of real-time feeds which explain how the service is operating in reality. This process involves

```
Public transport sample:

{u'distance': 13080,
 u'duration': 4302.25,
 u'financial_cost': 6.8,
 u'generalised_cost': 21.553755555555554,
 u'mode': u'transit',
 u'points_id': u'51.5309,-0.124192_51.46039067,-0.086255772',
 u'service_change_count': 5,
 u'sim_time_stamp': u'2015-10-26_06:30',
 u'steps_data': [{u'duration': 440,
                  'generalised_cost': 1.222222222222222,
                  u'mode': u'WALKING'},
                 {u'duration': 266,
                  'generalised_cost': 1.1556222222222223,
                  u'mode': u'WAITING'},
                 {u'duration': 846,
                  'generalised_cost': 3.6754,
                  u'mode': u'TRANSIT',
                  u'type': u'BUS'},
                 {u'duration': 88,
                  'generalised_cost': 0.2444444444444444,
                  u'mode': u'WALKING'},
                 {u'duration': 212,
                  'generalised_cost': 1.455733333333333,
                  u'mode': u'WAITING'},
                 {u'duration': 960,
                  'generalised_cost': 6.592,
                  u'mode': u'TRANSIT',
                  u'type': u'SUBWAY'},
                 {u'duration': 147,
                  'generalised_cost': 0.4083333333333334,
                  u'mode': u'WALKING'}],
 u'time_stamp': 1445844877,
 u'unweighted_duration': 2969}

Driving sample:

{u'GPS_status': True,
 u'distance': 9635,
 u'duration': 2187,
 u'financial_cost': 3.07578645,
 u'generalised_cost': 8.06336145,
 u'mode': u'driving',
 u'time_stamp': 1445804186}
```

Fig. 4.24 Example records with computed generalised costs

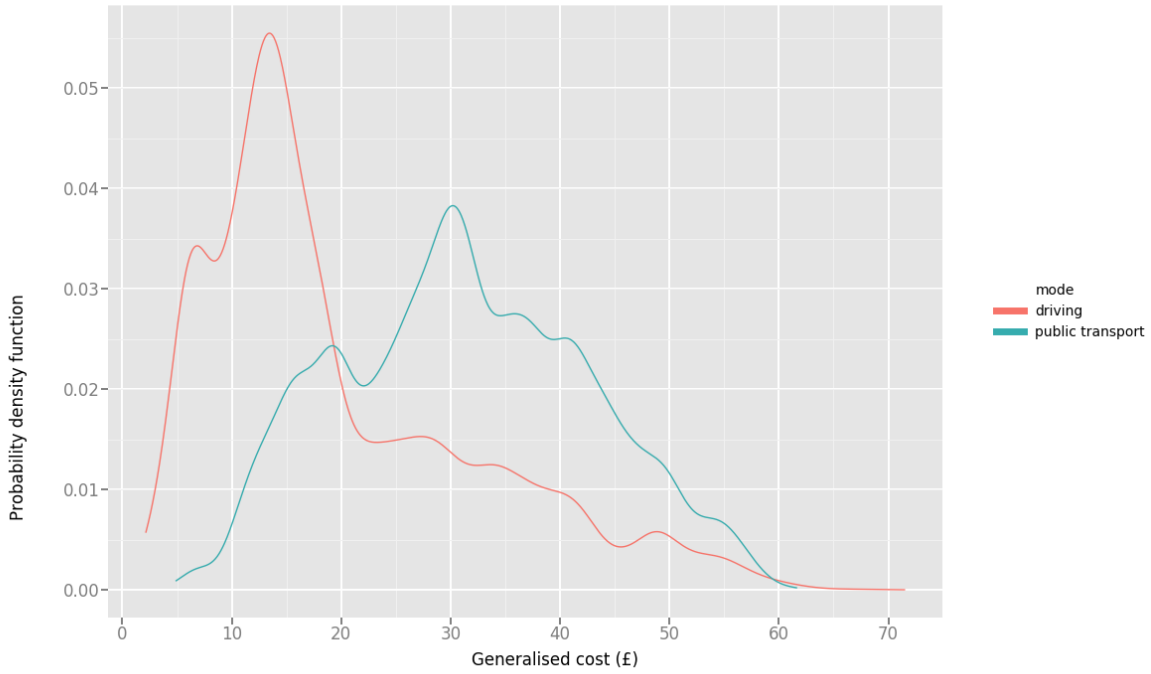


Fig. 4.25 Generalised costs probability density distribution per mode

natural language processing and is the focus of services such as Google Maps and Citymapper (e.g. (Citymapper, 2017)). Thus, the use of the Google Directions API responses was used in the sample dataset here, so as to illustrate the temporal fluctuations (in reality, not just as timetabled). This then has the challenge in that it does not enable the easy connection between exhibited journey time and planned journey time. As a result, the timetabled time may be reasonably assumed as the mean time. However, such an assumption has a significant impact on the meaningfulness of the result as there is a circular reference in the journey time reliability computation. Therefore, the same methodology as that for driving may be employed, to at least compare the standard deviation of journey times against the mean journey time, with no reference to expected, timetabled time.

Figure 4.26 presents the distribution of journey time reliability metrics, per mode, across the sample dataset. Driving illustrates the narrowest distribution, with a concentration around 0.10 to 0.15 and significantly less range than public transport and both exhibiting a slight skew to the left. The mean for public transport is slightly higher than driving (0.15 versus 0.13 for respectively), and the range significantly more so, illustrating significantly more risk in journey times when compared to driving.

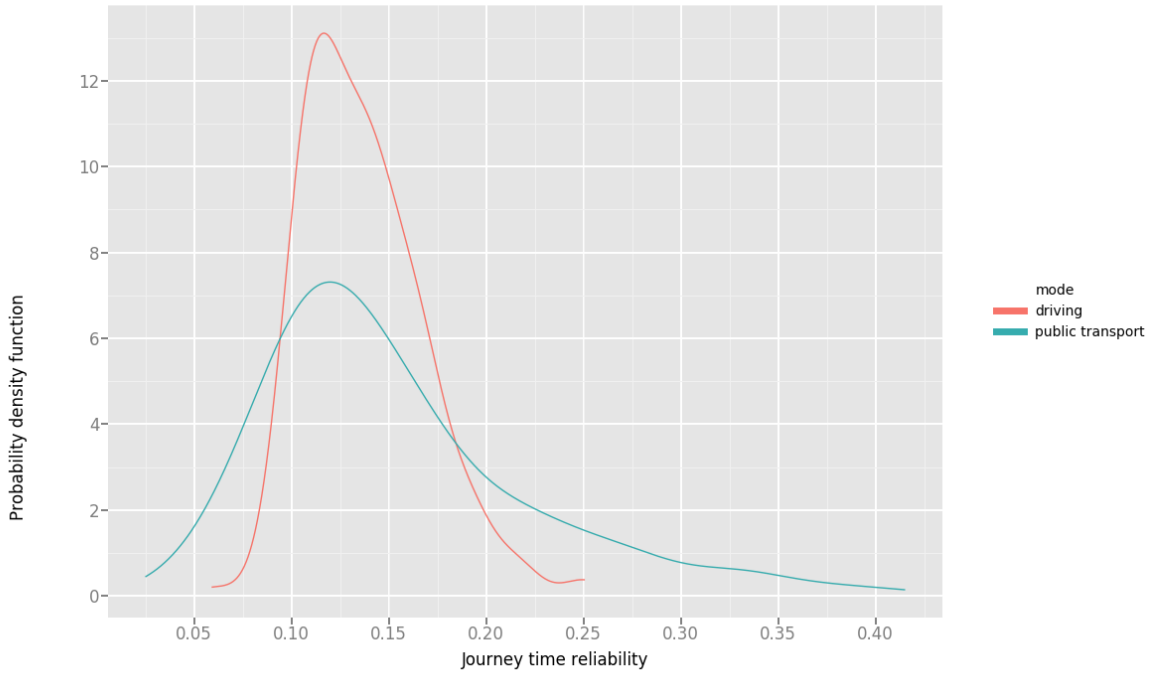


Fig. 4.26 Journey time reliability per mode probability density distribution

4.1.6 Agent feedback

A key aspect of an ABM is the interactions of individual agents. In order for this to occur there must be feedback between travellers. This is achieved by considering the impact of current traveller traveller decisions on future travellers. For vehicular travel on the road graph this can be modelled in terms of an increase or decrease in demand. Depending on the ratio of the demand (traffic volume) to supply (road capacity) a travel time impact may be calculated using context specific volume-delay functions, as was presented previously. Thus, a route which is oversubscribed will result in an increase in journey time, potentially leading to a modal change as a result of a change in the input to the decision making process. The context specific volume-delay functions derived in Section 3.2 present a means for relating the impact of vehicular volume to time delay.

No feedback is implemented for the public transport network and it is therefore assumed there are no crowding impacts.

4.2 Computational implementation

The model architecture presented in this chapter may now be implemented in a distributed, cluster compute system. The two broad computational challenges, the agent based computations and the graph based computations are discussed in isolation and then brought together in the form of a modal choice and route assignment agent based model.

4.2.1 Graph & ABM implementation

Apache Spark is an open source, general purpose compute engine for large scale data analysis (Apache, 2016). Spark is written primarily in Scala, a general purpose functional programming language (Odersky et al., 2008). However, in an effort to provide widespread accessibility Spark may be used interactively with Python and R shells. The Spark cluster is managed by Apache Hadoop YARN (Yet Another Resource Manager), which is responsible for resource management and job scheduling/monitoring. As has been illustrated, the agent decisions may be computed in parallel within a given time step using a MapReduce implementation in Spark. The graph algorithm implementation is a mathematically hard problem and the most efficient means of computation is on a single, non distributed node. This setup is graphically represented in Figure 4.27. Essentially, this illustrates how the agent decision making is partitioned by YARN within nodes and across multiple executor nodes, whilst the graph computation happens in a non-distributed fashion, on the driver node. The agent feedback, in between time steps is

also implemented using a MapReduce implementation, enabling for it to be done in parallel. The execution of updates from this to the graph on the next step is not computed in parallel.

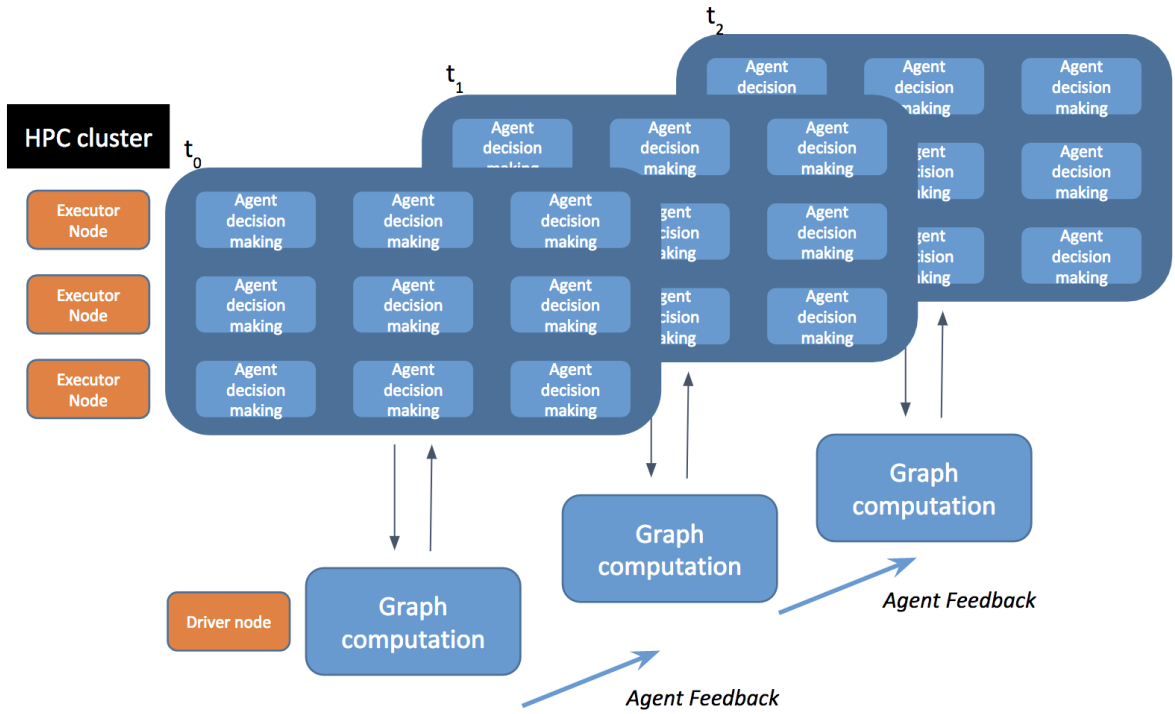


Fig. 4.27 System architecture

The graph computation was implemented using iGraph, a high performance graph library for Python, written mostly in C (Csardi and Nepusz, 2006). The ABM was programmed in low-level python with no dependencies no outside libraries for three main reasons, firstly to avoid dependency challenges when distributing libraries across the many computational nodes on the cluster, secondly, to avoid the computational performance overhead afforded by many ABM libraries and lastly to allow for specification of individual agent heuristics in a more open way. The agent tracking features, simulation logging, memory functions and plotting features are useful for localised ABM testing but were not required as this functionality was not able to support this large scale architecture [6]. Finally, the ability to disaggregate agent decision making heuristics across individual agent types does exist in some ABM libraries, but it often requires this to be done so in clusters or fixed types (Zheng et al., 2012). A key goal of this framework was to allow real-time calibration and other dynamic data driven heuristic research, so a high degree of flexibility in specifying this was a major factor in this design decision.

[6]This does not mean this ABM implementation does have these features, rather that they were added separately

4.2.2 Simulation seeding

A formal implementation takes agent seeds from a land use trip generation model, such as an activity based model or from the generation and distribution stage from the traditional four stage model. For the purposes of testing the implementation, a synthetic dataset of origins may be generated by considering the population distribution at the Lower Super Output Area (LSOA) level. Within the defined study area there are 4835 LSOA zones and as was previously shown in Section 3.1, for data request limitations the journey time information was harvested for the population weighted centroid of each LSOA. This methodology is illustrated in Figure 4.28. The LSOA zone population with zone area and the computed population density probability density plot for the entire study area are shown in Figure 4.29. With exception of some outliers, the majority of LSOA zones illustrate population densities of 10,000 people per square km.

From these LSOA zones, the following process is employed to seed agents:

1. Local LSOA network graph (for driving and public transport) computed for bounds of LSOA polygon
2. LSOA population is distributed randomly across graph vertices in zone
3. Shortest path for each agent to spatial centroid or PageRank centroid are computed. In the case of the vertex not being accessible to the centroid, the agents are attributed randomly to a different vertex. (e.g. on end of a one way road, not reachable from centroid)
4. Shortest path cost (time cost and financial cost) are exported for each agent.

This methodology is shown graphically in Figure 4.28. For each LSOA, the journey time may be computed from a given location in the zone, to the start point (population weighted centroid for driving, PageRank centroid for public transport) of the previously harvested journey times. A localised graph of the LSOA zone (as shown in Figure 4.31) may be created in order to compute these journey times, journey costs and journey distances; and then by referencing the ONS population statistics awareness may be made to the number of potential agents a given location may seed. The population weighted centroid was also used for the purposes of journey time specification in Chapter 3.1. However, unlike the road network, the public transport network is extremely time dependent, that is to say graph edges only exist at set times and it is not possible to traverse a bus route or railway line at times outside of the timetable. Conversely, you can always travel on a road edge when driving. In this case, the centroid for a polygon may have a more significant impact than that for the driving graph. Thus, an alternative method which assesses the most influential public transport vertex in a polygon and

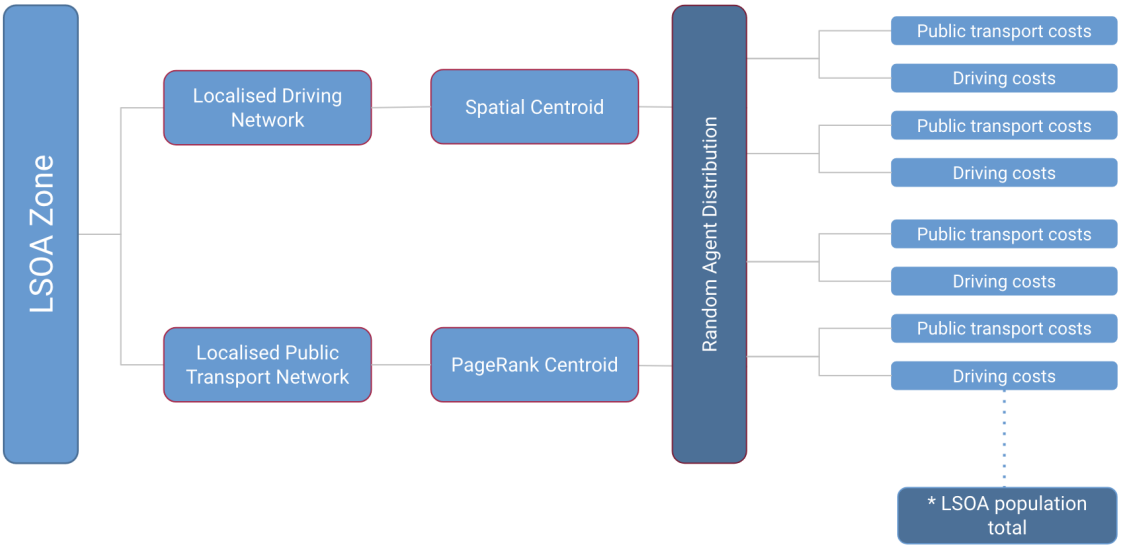


Fig. 4.28 Simulation zone extrapolation methodology

assigning it the centroid was investigated. The PageRank algorithm (Page et al., 1999) may be used to measure the relative importance of vertices via their respective weighted edges. It was proposed in 1999 and forms the basis of Google's web search engine. Despite the intention for use solely in the web search space, it has many graph applications and in this case may be used to assess the relative importance of a vertex (station, bus stop) dependent on the number of services (edges) through it. In this case, the average journey time between vertices was used as the weighted metric for the PageRank algorithm.

The generated dataset contains a services which serve the Greater London area in any form, and thus due to London's strategic importance, captures regional and national services as well as the more obvious TfL bus, rail and Underground services. Figure 4.30 illustrates the PageRank score for stations and bus stops within the study area. The PageRank score of a given vertex is shown by its size and the colour of the connecting edges illustrates the mode of transport between given unique vertices. The density and complexity of the network is evident and thus only important, high PageRank scoring, vertices are labelled for clarity. The density of the bus network and the importance it plays at the macro scale is extremely evident. Financial costs are computed using the respective coefficients (RAC, 2016) and (TfL, 2016a). Both of these methods assume that the journey times to this centroid are fixed and do not change with time as a result of timetables or congestion.

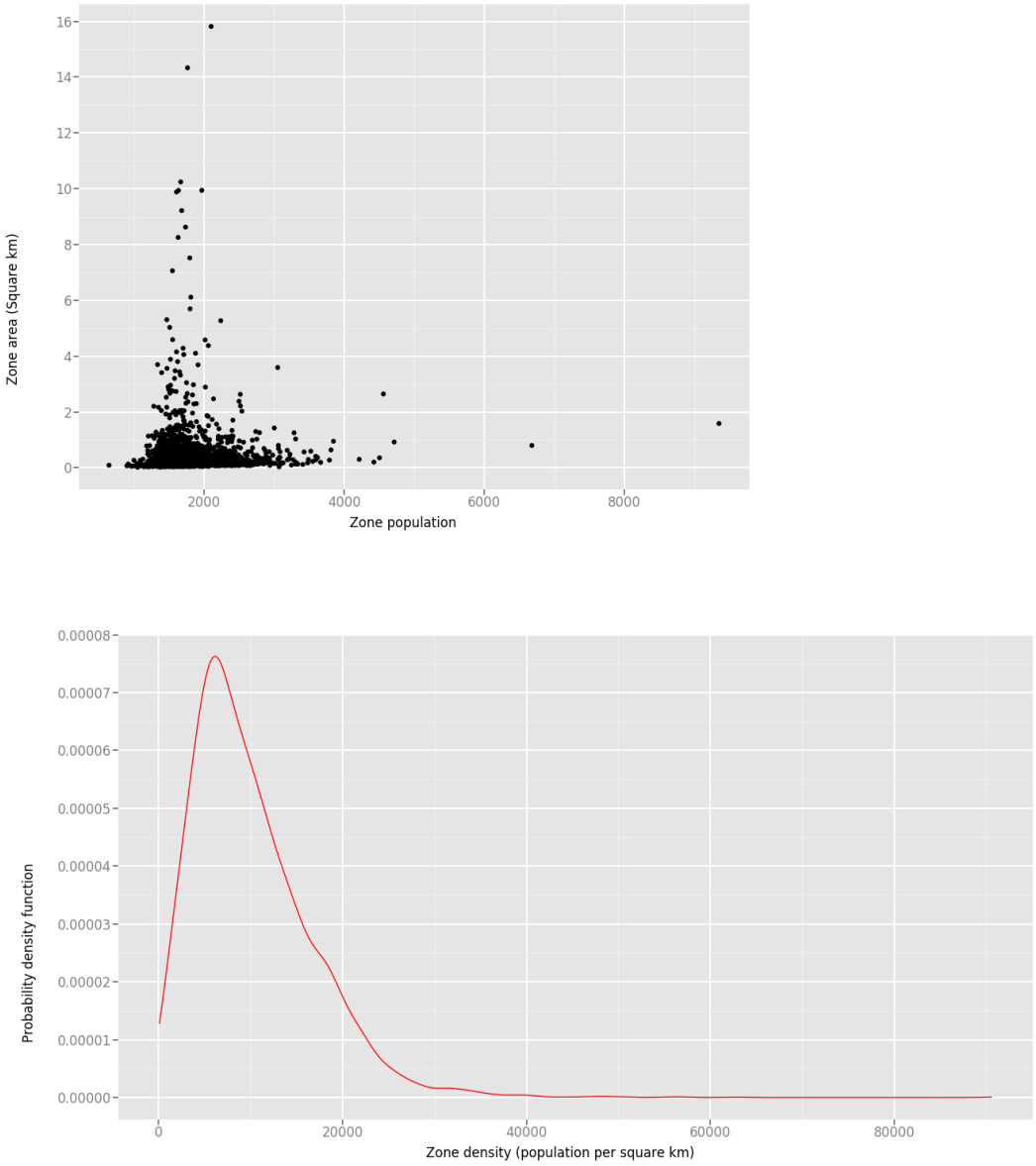


Fig. 4.29 LSOA population versus zone area scatter plot (top) and computed LSOA population density probability density function (bottom)

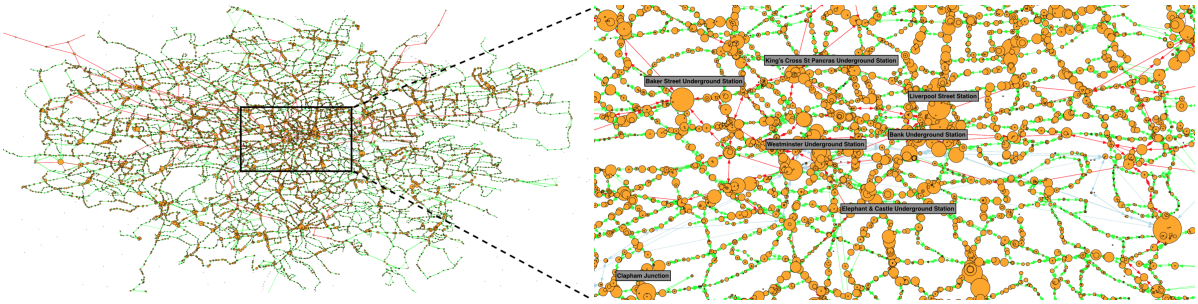


Fig. 4.30 PageRank Visualisation of public transport stops and stations. Vertex size illustrates PageRank score and edge colour illustrates mode

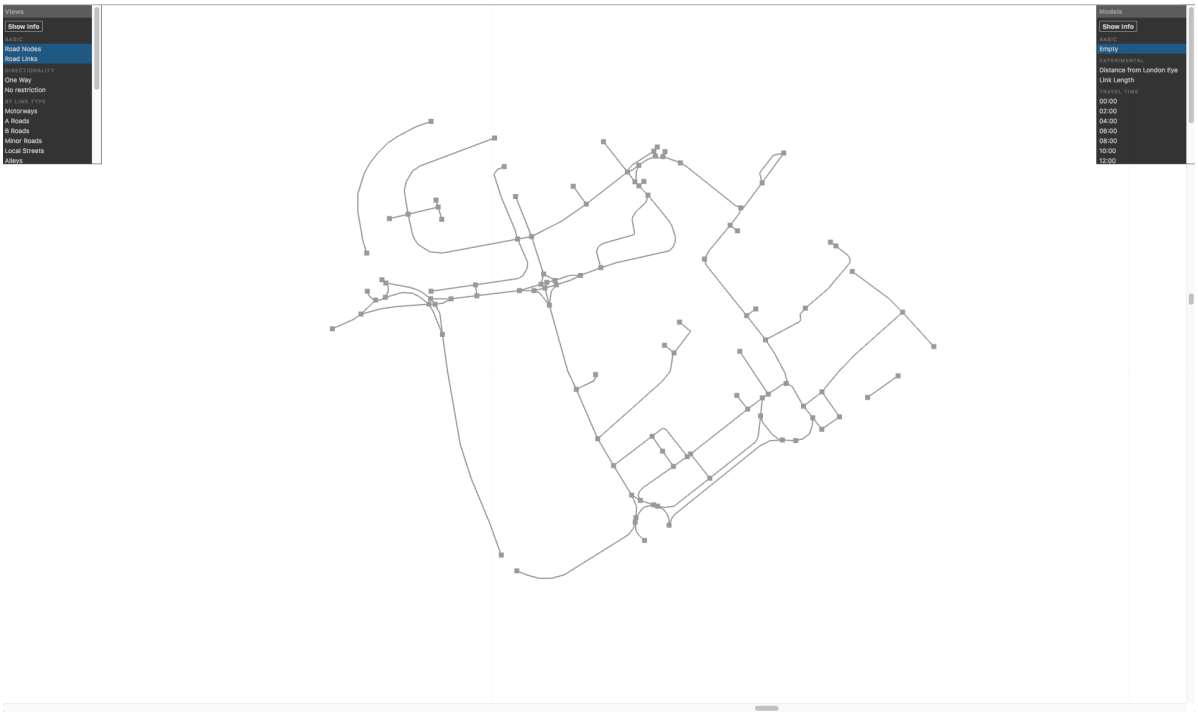


Fig. 4.31 Example localised driving graph for LSOA zone

```
[
  {
    "lsoa": "E01001572",
    "road_costs": [
      {
        "financial_cost": 0.06899539908679402,
        "journey_time_cost": 28.043918486339 58,
        "distance_cost": 376.0689469018138,
        "population": 33
      },
      {
        "financial_cost": 0.055593549138662345,
        "journey_time_cost": 22.59659312137193,
        "distance_cost": 303.0203137575976,
        "population": 36
      },
      {
        "financial_cost": 0.04145397227366341,
        "journey_time_cost": 16.849410754406534,
        "distance_cost": 225.95059821659163,
        "population": 29
      }],
    "pt_costs": [
      {
        "financial_cost": null,
        "journey_time_cost": 161.2621181084,
        "distance_cost": null,
        "population": 158
      },
      {
        "financial_cost": null,
        "journey_time_cost": 0.0,
        "distance_cost": null,
        "population": 162
      },
      {
        "financial_cost": null,
        "journey_time_cost": 0.0,
        "distance_cost": null,
        "population": 138
      },
      {
        "financial_cost": null,
        "journey_time_cost": 161.2621181084,
        "distance_cost": null,
        "population": 161
      },
```


Fig. 4.32 Local extracted road graph for LSOA E01000010

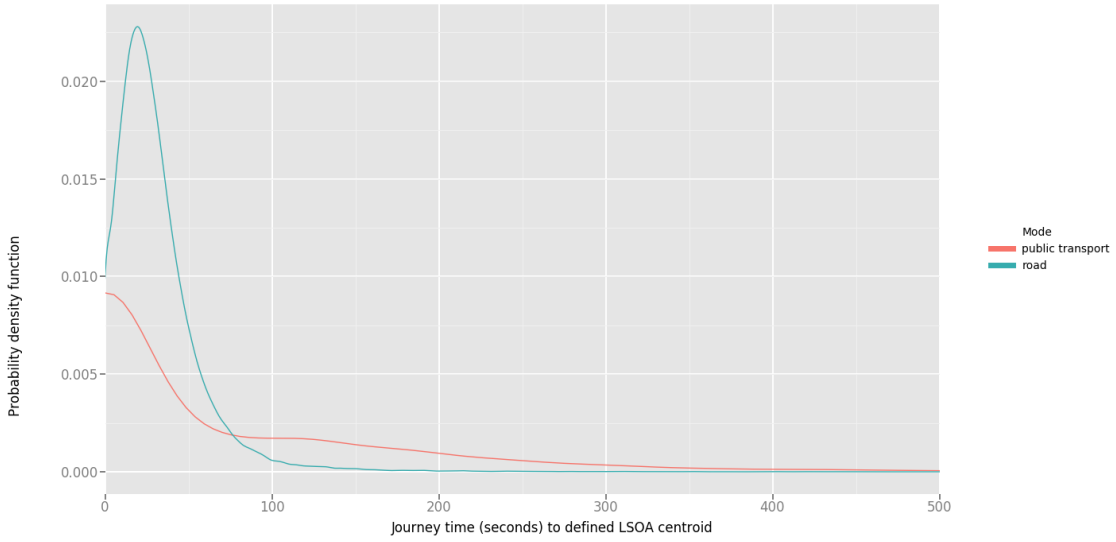


Fig. 4.33 Journey time to defined LSOA centroid probability density function, by mode

Collated seed

The generated dataset then presents extrapolation information for a given LSOA. In Figure 4.33 the journey time to respective centroids is shown for the entire dataset as a probably density function. Driving is shown to exhibit the lowest journey time costs, with the majority found to be less than a few minutes. Contrastingly, public transport exhibits longer extrapolation times.

4.2.3 Computational performance

The previous section disaggregated the two distinct parts of the computational challenges in order to illustrate the distinct nature of each individually. Of course, the actual model itself requires their integration. The goal of this architecture was to permit the use of fine resolution spatial and temporal resolution within an agent based modelling framework. The original proof of concept was carried out on HPC infrastructure at the University of Cambridge and the performance tests were carried out on Microsoft infrastructure, through academic access to the HDInsight programme. The setup consisted of 6 nodes, where there were two head node D12v2 with 8 cores and 4 worker nodes D13v2 with 32 cores. A series of analyses in the context of St Pancras, London City Airport and Heathrow Airport (4835 origins and 3 destinations) hub travel are presented in Table 4.13. A simulation consisting of the 8 million agents per time step was seeded in order to test the scaling footprint of the implementation. The simulation was run over 5 time slices, from 9am to 1pm and thus a total of 34 million agents were simulated.

A generalised cost decision making heuristic was implemented and a hub decision was made (with modal choice and route assignment). The value of the cluster implementation across 6 nodes illustrates a significant time saving, from over 2 hours on a single node to around 15 minutes on the 6 node cluster.

Table 4.13 Computational performance for 3 destinations

Hardware	Time (Seconds)	Agents simulated	Time resolution
Single node	9383	34,694,852	9am to 1pm (1hour slices)
6 node cluster	1286	34,694,852	9am to 1pm (1hour slices)

Sensitivity analysis

The computational performance may be assessed in terms of simulations with different spatial and temporal complexity. Consider an OD matrix which consists of 1000 origins to a single destination. This matrix consists of 1000 unique origin destination pairs however, it may be manipulated to efficiently take advantage of Dijkstra's algorithm implementation. A consequence of finding the shortest path between a given two vertices is that the shortest path between all other vertices in the graph must be found. Therefore, the arrangement of a given OD matrix into efficiently chained requests for single origin or destination sources offers significant performance gains. This is extremely relevant in the context of multiple transportation hub studies, such as that shown in Figure 4.34.

Consider the same total amount of agents being simulated, but with a different spatial and temporal distribution. We may vary the number of agents per time step and reduce the resolution of this time step to assess how the spatial and temporal dimensions interact. In Table 4.14 a series of four scenarios with different footprints are presented. Each of these scenarios results in the same total number of agents simulated.

Table 4.14 Spatial versus temporal complexity - scenarios

Scenario	Time slices	Resolution	Agents per slice	Total agents
A	4	1 hour	8,673,713	34,694,852
B	8	30 minutes	4,336,856	34,694,852
C	12	20 minutes	2,891,237	34,694,852
D	16	15 minutes	2,168,428	34,694,852

In Figure 4.35 the proportion of ABM time actually reduces across the scaling tests, that is to say that the efficiency of the MapReduce implementation increases. In Figure 4.36 the agent

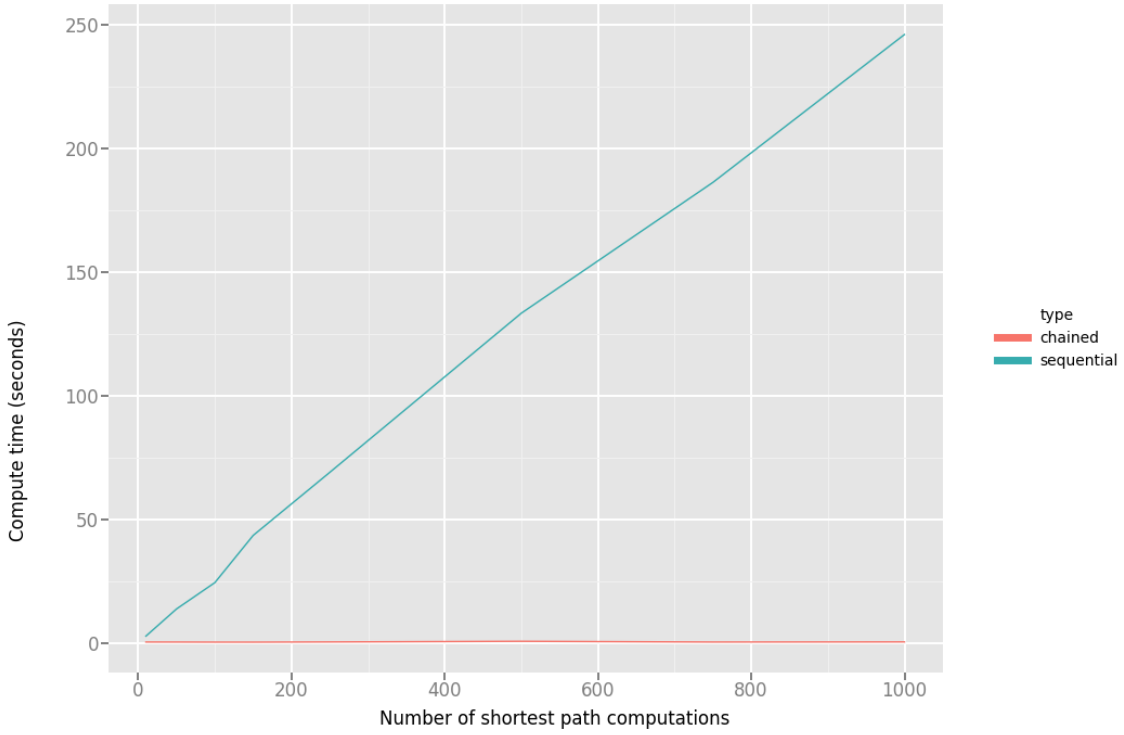


Fig. 4.34 Sequential versus chaining shortest path queries for a one by destination matrix

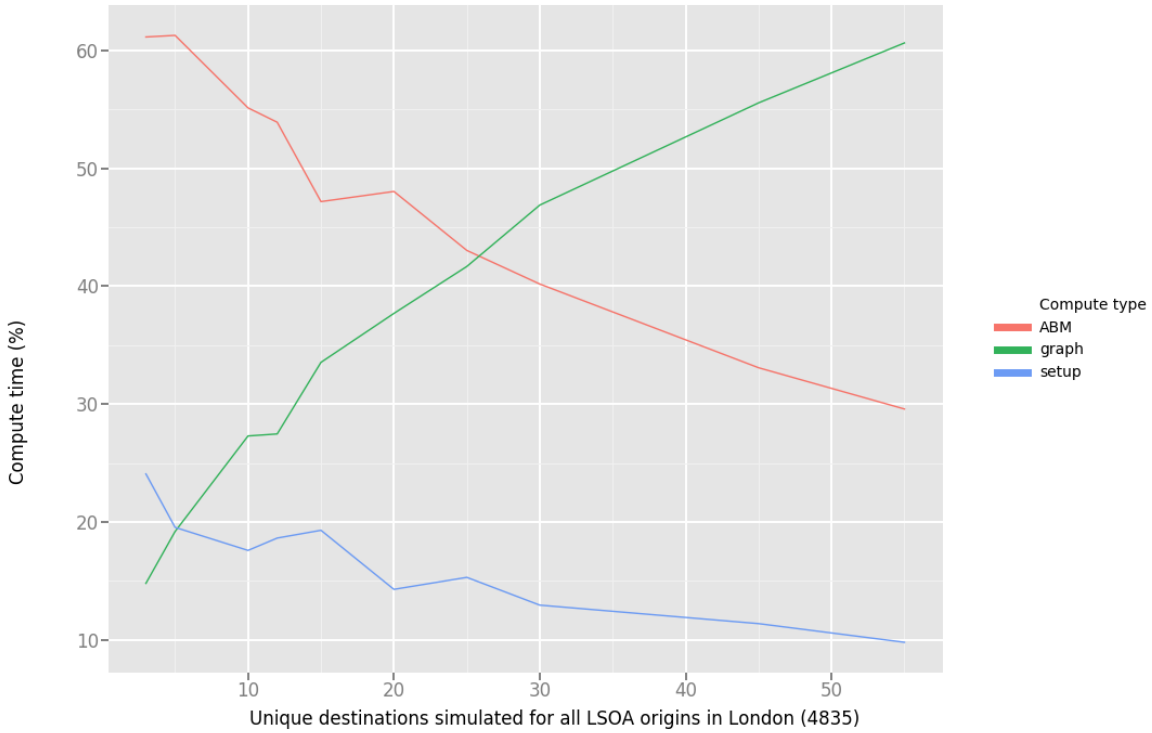


Fig. 4.35 Performance of model, disaggregated by computional type across various levels of spatial complexity

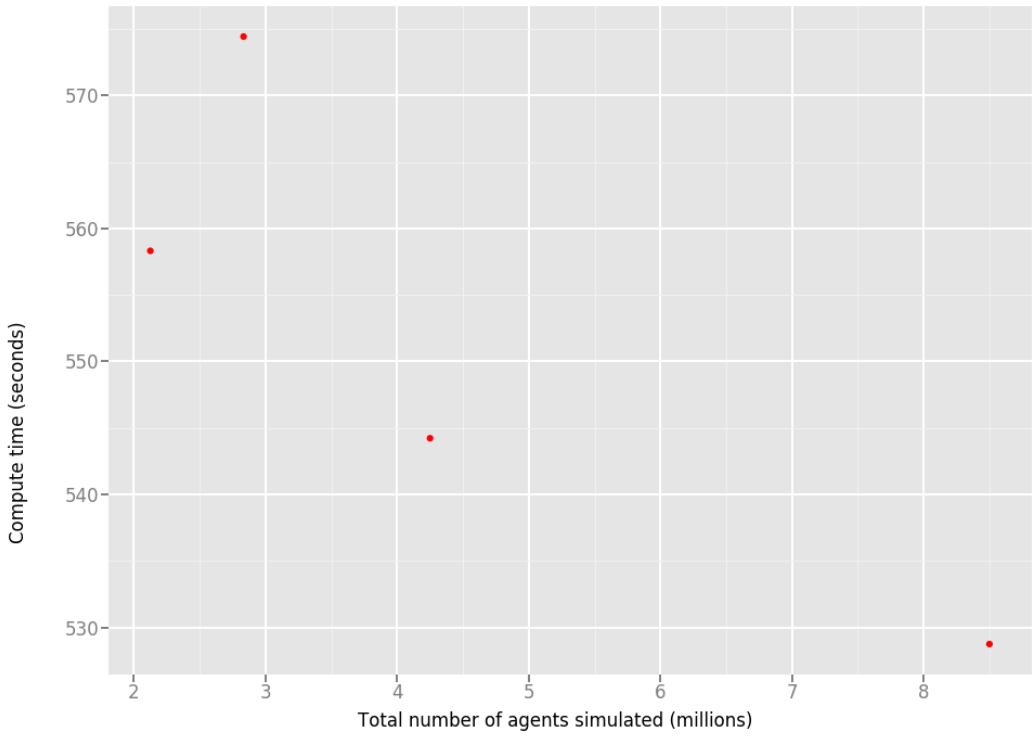


Fig. 4.36 Performance of agent based model in isolation

based modelling computation is extracted, and the scaling is shown in real terms across the scaling tests. The efficiency of the method only improves as the agent count increases, that is to say more agents may be computed in less time. This is as a result of the effort the MapReduce process must go through to parallelise the smaller agent counts. This effort is significantly larger than the value this process actually presents (in terms of time savings) and thus manifests in a situation where more agent counts may be computed in less time. This scaling trend is likely to continue well beyond what is reasonably useful for the modelling demand, as this particular computation is still well below its theoretical limit, and thus, is unlikely to be a significant computational bottleneck for the foreseeable future. Contrastingly, the amount of time spent executing spatial, graph computations grows extremely fast and overtakes the ABM time quickly and grows at a rate of $O(V^2)$, becoming the dominant process and the limit to scaling.

4.3 Chapter Summary

4.3.1 Summary

1. Hypothesis

   The data sources shown in Chapter 3 may be merged into a more empirically robust representation of the transportation network and this used as an input to a scalable, flexible agent based model.

2. Novel contributions

   A proprietary graph data structure was developed to combine a range of disparate data sources into a multi-modal representation of the transportation network. This was used within a distributed agent based model and the MapReduce paradigm utilised to reduce compute times whilst maintaining fine spacial and temporal resolution. A range of different decision making heuristics are supported.

Conclusions

A scalable framework for performing agent based simulations with fine resolution spatial and temporal data has been presented. This framework makes use of recent advances in computer science in order to permit simulations with more data at finer time slices, over longer time horizons. The use of such a framework is a step towards a quantitative assessment of the hypothesis that simply including more information, in more complex ways will produce better modelling outcomes in the transportation sector. This hypothesis is central to that of the

'big-data' paradigm. The ability to consider and compute real-time data within a modelling framework may eventually permit real-time calibration and the merging of operational with strategic models. Using the described process, a series of steps are employed to create a directed, weighted, embedded, explicit and labelled graph for London. This graph is then used within a modal choice and route assignment agent based model.

4.3.2 Limitations

1. Extremely poor understanding of demand - This framework makes use of the datasets presented in Chapter 3. In the context of the road network, volumes on a given road are computed based upon volume-delay relationships. That is to say we may accurately see the impact of true origins and destinations, but do not know them. We therefore have the journey times as a result of congestion, but do not understand nor know the journey's which resulted in them. Methods do exist for extracting origin destination but struggle with validation (Willumsen, 1978). As a result, although the proposed technical architecture permits for long time horizon, fine resolution simulations, there is limited modelling value in doing so. In Chapter 6 this is discussed further.

2. Discrete versus continuous time and assumptions around equilibrium - The proposed methodology employs discrete time steps and utilises empirically robust volume-delay curves in order to handle transitions across time steps. The empirical nature of these curves likely enables for statistically strong simulations the closer a given simulation is to the measured situations but it is unknown how this compares to a continuous time simulation such as that done by MATSim. It is challenging to compare these methods due to the differences in agent logic used and the use of stochastic methods in MATSim (Zheng et al., 2012). This challenge also relates to the assumptions surrounding network equilibrium. Generally the Frank-Wolfe algorithm or Wardrop equilibrium are used as relevant implementations of the Nash equilibrium as they attempt to satisfy network constraints. Of course, agent based models do not naturally reach equilibrium and the use of such a methodology on simulation outputs may be used to constrain simulations where convergence to some defined criteria does not occur. The ultimate measure of success for appraising these implementations is their relative performance to validation criteria. In the context of an ABM, this becomes challenging as traditional transportation validation criteria are generally macro agent outputs, such as edge (link) level counts or journey times. Validating an ABM against such criteria may not necessarily being meaningful in its own right. Such research further highlights the challenges faced in validating micro

interaction models with well understood macro validation criteria. In Chapter 6 this is discussed further.

3. Network accuracy - Lane counts and junction complexity has not been included in a sophisticated way. Lane and turn restrictions are not respected, despite the inclusion of directionality. These limitation also cause issues when attempting to find an appropriate volume-delay curve for a given section. Further to this, an unknown sample bias is exhibited towards those roads for which the volume-delay functions were generated. Free flow capacities, lane counts and volume-delay functions are assumed by closest match from Chapter 3 study.

4. Public transport crowding - The data sources utilised to quantify the public transport network in Chapter 3, the GTFS and real-time feeds, do not include information on capacity. The quantification of this capacity is highly complex due to the range of bus types and rolling stock available. Secondly, the relationship to demand is complex and highly context specific (e.g (Tirachini et al., 2013)). Feedback on the public transport network is not implemented in this framework and is a significant limitation as it is therefore assumed that there are no crowding impacts.

5. Hadoop, Spark and the cost of scaling - The use of Hadoop & Spark forces some implementation constrains. Library's and complex dependences (such as pandas, scipy, numpy) must be installed locally per node, and cannot be bundled and shared across nodes via traditional sharing methods. This is because they utilise C level code, which must also be compiled specifically for the hardware OS. Local nodes need local dependencies and so code changes must be managed across the cluster and not just the master node. Maintaining and running distributed systems is a non-trivial exercise, despite recent advances.

6. Cycling & Walking inclusion - Cycling and walking (beyond connecting public transport), were not included, however the road graph also supports modes beyond that of personal car and taxi. The same infrastructure supports walking and cycling modes, with type restrictions.

7. Pricing - The representation of the zone system for TfL services assumed anytime, adult and pay as you go with no capping and did not consider any other type of ticket. Parking prices were similarly abstracted and a single value considered.

8. Graph database stability - The framework presented here does not contain a formal graph database backend. As a result, effort was made to enforce unique index and key value

pair attributions to ensure graph stability. The inclusion of a formal graph database would act to stabilise the storage of either the inputs (real-time/empirical data) or simulation geospatial data.

9. Decision making heuristics - The graph compute makes some decisions on what the behavioural trip consists off. The use of business value of time coefficients to utilise the disaggregation of mode and distance gave benefits in specificity, but is challenging for the actual simulation of HSR here where tourism is likely significant. Further complexity could be added by considering a taxi pricing structure which considers supply and demand, such as Uber or Lyft. The increased demand during rain may lead to a pricing tipping point that may even then result in a public transport route with a heavily weighted walking section.

Chapter 5

HS1 Case Study

In Chapter 2 the literature illustrated the need to move away from tail-pipe only comparisons to more holistic life cycle impact assessments. In this Chapter the capital carbon costs of infrastructure construction and hub level travel are considered in order to better understand HS1 travel from London to Paris and London Brussels travel.

In the first section, the capital CO_2 impacts for HS1, the Channel Tunnel, LGV Nord and HSL 1 are considered in an effort to avoid a tail-pipe only bias. Secondly, an internal Eurostar report from 2006 on the comparison between Eurostar and aviation for London Paris and London Brussels is assessed to understand the historic operational CO_2 performance. Thirdly, this operational aspect of HS1 is considered in the context of surface access - via the collation of the distinct analyses presented in Chapter 3. These data sources are then used to aid the understanding of how the relevant transportation hubs change with time as a result of congestion, timetabling and other impacts.

5.1 Historic performance

5.1.1 Capital CO_2 emissions

Journeys from London to Paris and Brussels are the focus of this study and thus the spatial boundary consists of three constituent parts:

1. HS1
2. The Channel Tunnel
3. LGV Nord & HSL 1

HS1

The Channel Tunnel Rail Link (CTRL), now referred to as High Speed 1 (HS1) is a 109km high speed rail line from London to the entrance of the Channel Tunnel at Folkestone. Eurostar operate services chiefly to Brussels and Paris, Southeastern operate domestic services and DB Schenker operate some freight services.

The construction of HS1, involved a significant CO_2 investment through geotechnical structures, track, stations and the rolling stock itself. As was illustrated by the (UIC, 2011) report capital CO_2 emissions may be in certain conditions highly important. For example, the Taipei-Kaohsiung line in Taiwan has 10g CO_2 per pkm for rolling stock and construction alone, which is directly comparable to the total (operation inclusive) 10.3g CO_2 per pkm for the South Europe Atlantic line in France (UIC, 2011).

An undergraduate project was carried out in the University of Cambridge to quantify the capital CO_2 cost of the HS1 line. This was carried out with support from Arup who provided the *as-built* specification. This study built upon some small scale case studies by Chau *et al* in 2012. This study (Chau et al., 2012b) considered section 310 and 220 of the line and found that there were two types of structure, those where the embodied materials equate to 90% of the total and those where the installation and transportation can account for up to 40% of the total. This study and others (Hughes et al., 2011b) highlighted the care that must be taken for considering geotechnical structures where the manipulation and movement of ground material can shift the dominance away from embodied focussed studies and thus distort actual emissions. A freedom of information (FOI) request was made to HS2 limited for the methodology employed as part of their environmental impact study (HS2, 2012) and this was then used to guide the assessment of HS1.

The bounds for this study are shown in Figure 5.1. This study found that the total capital CO_2 footprint equated to 2.05MtCO_2. The proportional disaggregation of this by structure type and classification is shown in Figure 5.2.

The Channel Tunnel

The Channel Tunnel is a 50.45km rail tunnel connecting Folkestone in the UK to Coquelles in France. There are three tunnels; two rail tunnels (7.6m diameter) and a single service tunnel (4.8m diameter), as illustrated in Figure 5.3.

No formal emissions or embodied energy assessment has been carried out on the Channel Tunnel and limited construction data is available in the public domain. Open track studies are not comparable due to the energy intensity of tunnel construction. The (Chau et al., 2012b) study calculated the embodied energy costs of contract 310 and contract 220 of CTRL, where

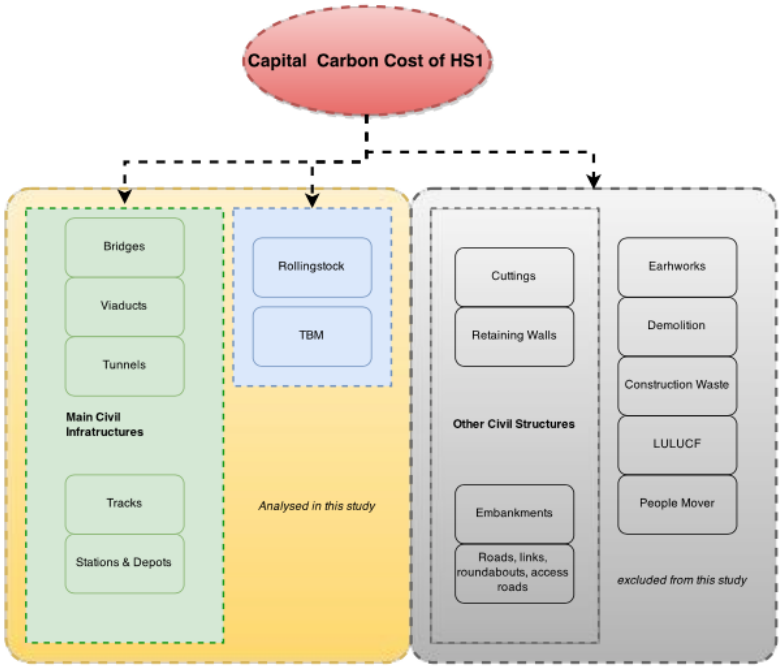


Fig. 5.1 Study bounds for capital CO_2 assessment of HS1 (Lin, 2015)

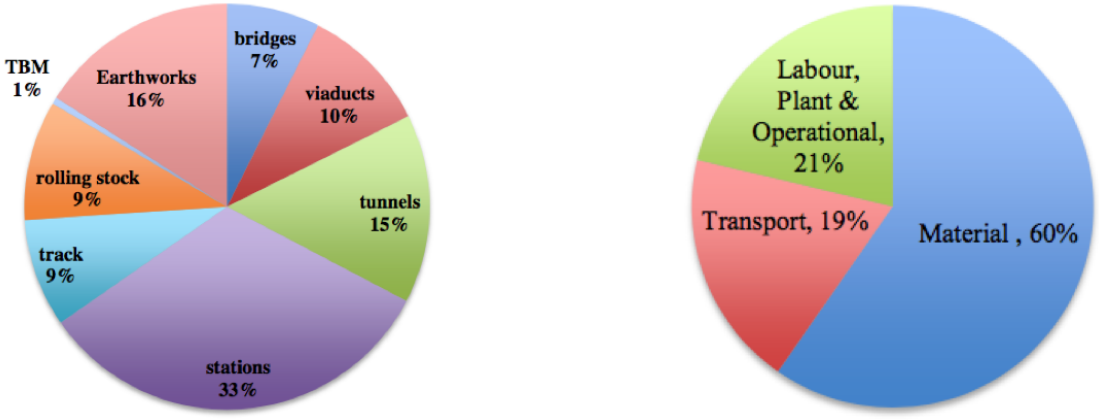


Fig. 5.2 HS1 CO_2 footprint disaggregation by structure type (left) and classification (right) (Lin, 2015)

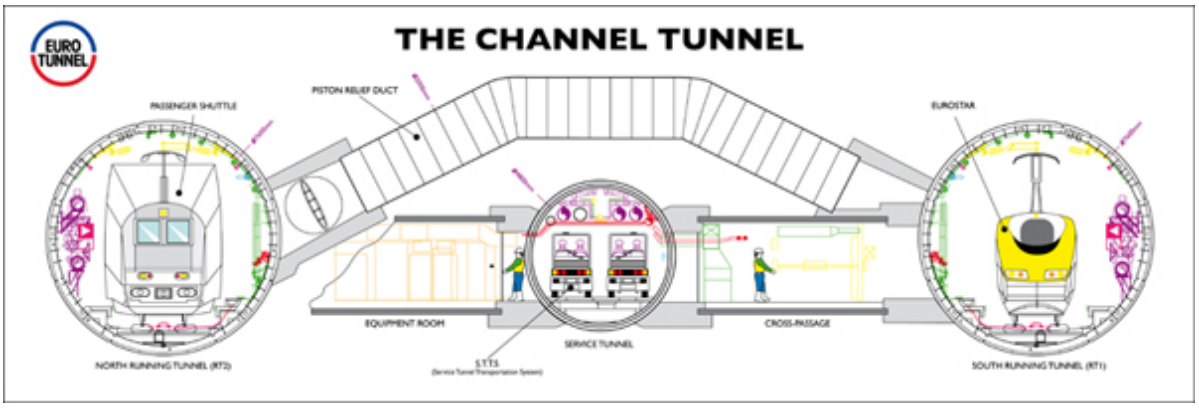


Fig. 5.3 The Channel Tunnel cross section (GETLINK, 2009)

220 was a tunnel segment through London Clay and is the closest data point to the Channel Tunnel available in the literature. The tunnel was constructed in the early 1990's, when energy intensity and emission factor statistics were in their infancy and were widely unknown. Despite these challenges, the contract 220 study may be used to estimate the emissions footprint to an order of approximate magnitude. Contract 220 was found to cost 949TJ, or 126GJ per metre inclusive of material, transportation, manufacturing and installation (e.g. TBM operational costs). The joule to kWh conversion was assumed as 0.2778MJ per kWh (Trust, 2008) and the UK CO_2 intensity assumed at 1990 rate of 700gCO_2/kWh. The assumed capital CO_2 cost for the Tunnel is therefore approximately 1.24 MtCO_2.

LGV Nord & HSL 1

Beyond the Channel Tunnel, Eurostar services connect to the French LGV network and diverge near Lille depending on their final destination. At Fretin, near Lille, services may continue directly to Paris using a bypass, or connect to the Lille-Brussels line. LGV Nord connects Paris to Coquelles and the Channel Tunnel with a continuous section of 333km of predominantly open track. The Lille-Brussels line is 71km from the Fretin junction Brussels Zuid Station, again predominantly open track. In 2011, the French national railway company (SNCF) published a carbon footprint study on their newest High Speed rail line, the LGV Rhin-Rhône line. The LGV Rhin-Rhône line is a 140km of mostly open track and is comparable to the majority of the French network and may be used to estimate the capital CO_2 emissions. The SNCF study found that total emissions equated to 7,350 tCO_2 per km track, considering materials, transportation and civil engineering works (SNCF, 2011). These coefficients may then be used to estimate those associated witht the LGV Nord and HSL 1 sections, as shown in Table 5.1.

Table 5.1 Estimated capital CO_2 emissions per section

Section	Length	Capital CO_2 cost
LGV Nord	333km	2.447 MtCO_2
HSL 1	71km	0.521 MtCO_2
Total	**404km**	**2.97 MtCO_2**

Total

Each of these sections also operate non Eurostar passenger services and to a lesser extent, freight services. There is therefore a requirement to apportion some of these CO_2 costs elsewhere, beyond the scope of this study. There has been a significant amount of change in the service mix across the HS1 line, with Eurostar, Southeastern and more recently DB Schenker sharing the use of the HS1 infrastructure. Southeastern operate on average 8 trains per hour to a mix of Ebbsfleet and Ashford destinations (Southeastern, 2016) and in 2016 Eurostar operated on average 5 trains per hour (Eurostar, 2016a). Since 2013, DB Schenker have operated a twice weekly service to Poland (Journal, 2014). It is therefore assumed that the HS1 capital CO_2 cost for which Eurostar may be apportioned is approximately 38% (not including DB Schenker services since it a recent and infrequent addition).

The section of open track from the Channel Tunnel exit to Paris and Brussels contrast with CTRL in that they are more heavily used by non-Eurostar services, such as French SNCF TGV services and Belgian Thalys services. In 2016, Eurostar was approximately 10% of the services.

The Channel Tunnel operates a significant vehicle shuttle service as well as through-rail passenger services. In 2016, Eurostar was approximately 20% of the services (EuroTunnel, 2016). The estimated total capital CO_2 emissions for the study area are therefore shown in Table 5.2.

Table 5.2 Estimated capital CO_2 emissions per section and total apportioned amount

Section	Capital CO_2 cost	Percentage allocation to study services
HS1	2.05 MtCO_2	38%
The Channel Tunnel	1.24 MtCO_2	20%
LGV Nord & LGV1	2.97 MtCO_2	10%
Total	**1.495 MtCO_2**	

To put this amount into context, it is significant, at around 0.4% of the total annual UK CO_2 emissions for the year of 2016 (based on (DfBEIS, 0016)).

5.1.2 Operational CO_2 emissions

High Speed Rail

In 2006 Eurostar commissioned a report on the CO_2 comparisons between Eurostar services and that of short haul-aviation (Watkiss, 2009). In 2009 this report was updated with cab measured energy usage from the rolling stock in place of manufacturer performance factors. This section contains a summary of this unpublished report.

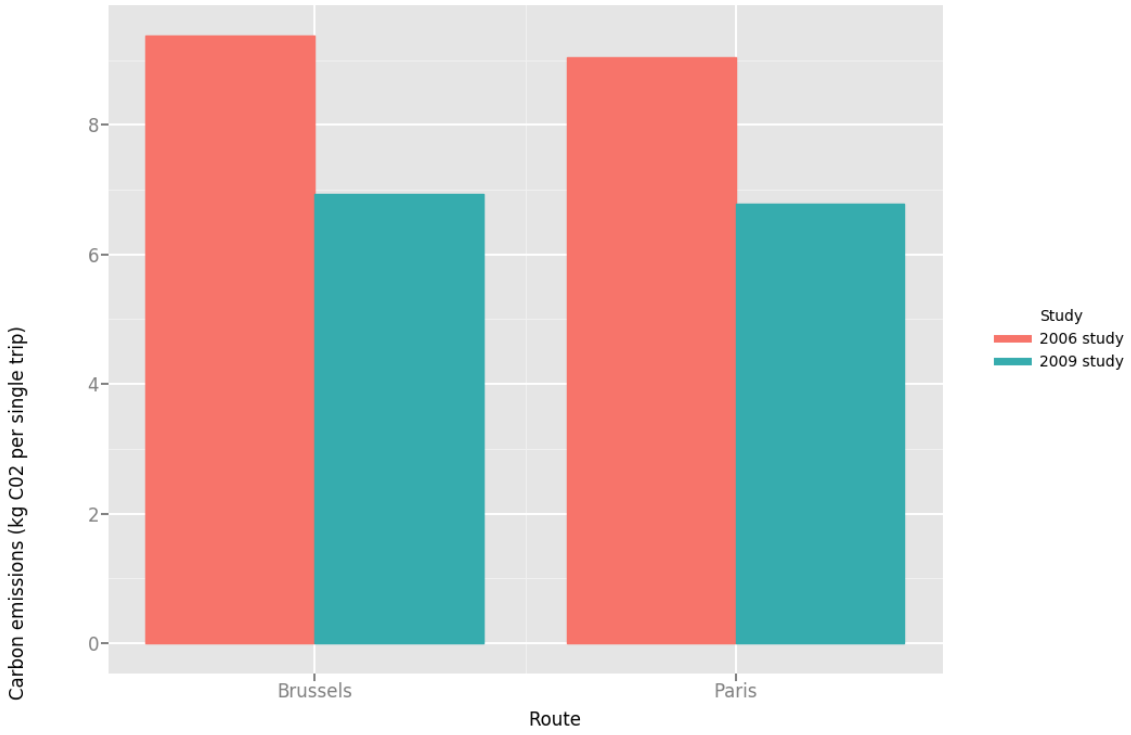


Fig. 5.4 Kg CO_2 per train (single trip) comparative for Paris and Brussels, with 2006 coefficients and updated measured 2009 statistics (Watkiss, 2009)

The kg of CO_2 per passenger single trip is dependent on load factors and the actual generation of the mix used to power the rolling stock. The statistics on energy mix CO_2 coefficients are generally made available at a national level and since these services involve time in at least two countries, methodological differences in how each country reports has a significant impact. The UK advocates for country wide, average coefficients whereas conversely, France advocates for local supplier coefficients as and when these are available. The difference between these two methodologies is apparent in Figure's 5.5 and 5.6. These Figures illustrate the trend in per passenger CO_2 emissions for each route, from 2007 to 2009. The step change

evident in both methodologies is as a result of a switch to French (nuclear dominated) energy for the Eurotunnel in 2008. The Brussels routes performs generally worse than the Paris route as a result of the relatively higher CO_2 intensity of the Belgian energy supply compared to the French.

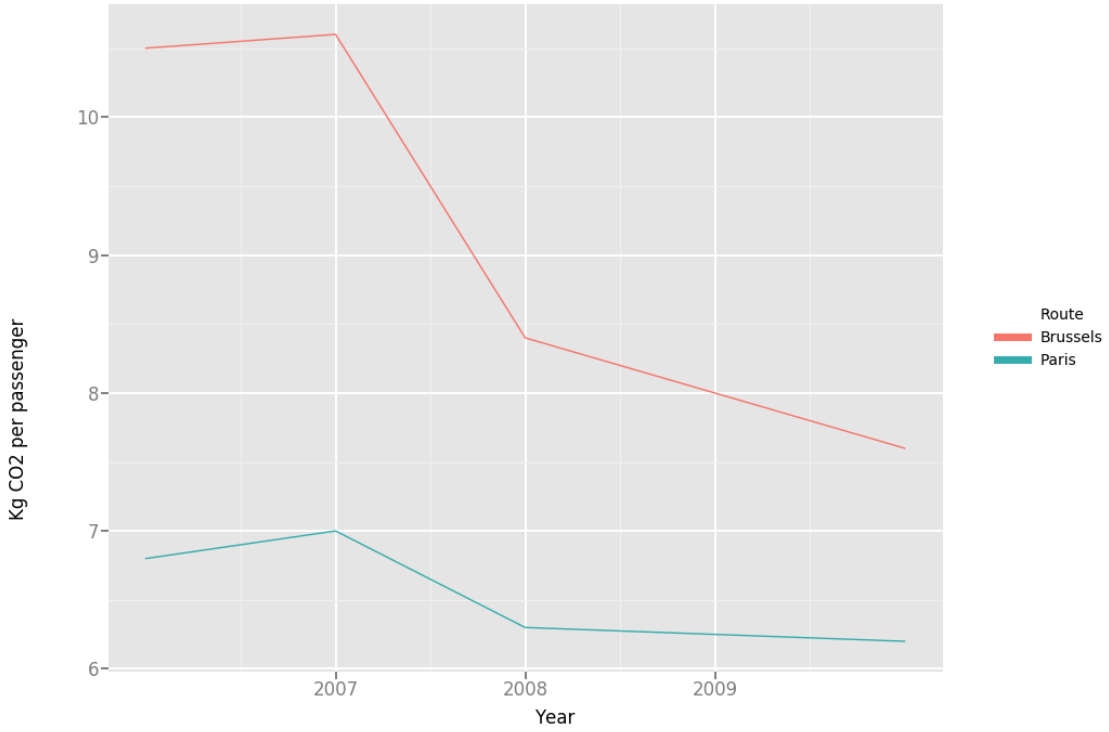


Fig. 5.5 kg CO_2 per passenger single trip, by average energy mix coefficients (derived from (Watkiss, 2009)

Aviation

The Eurostar report used the International Civil Aviation Organization (ICAO) emission calculator (ICAO, 2017) with purchased (non publicly available) specific Civil Aviation Authority (CAA) load factors in order to reflect actual demand for the actual individual routes. The calculated emissions, per passenger are shown in Table 5.3. ICAO are the nominated lead for reporting aviation climate change statistics to the United Nations Framework Convention on Climate Change (UNFCCC, 2018) and the body used by the UK's Committee on Climate Change (CCC, 2018), the authority who advise on the UK's carbon budget. The low load factors for London City flights are generally responsible for the considerably higher emissions of 219.8 CO_2 per passenger trip.

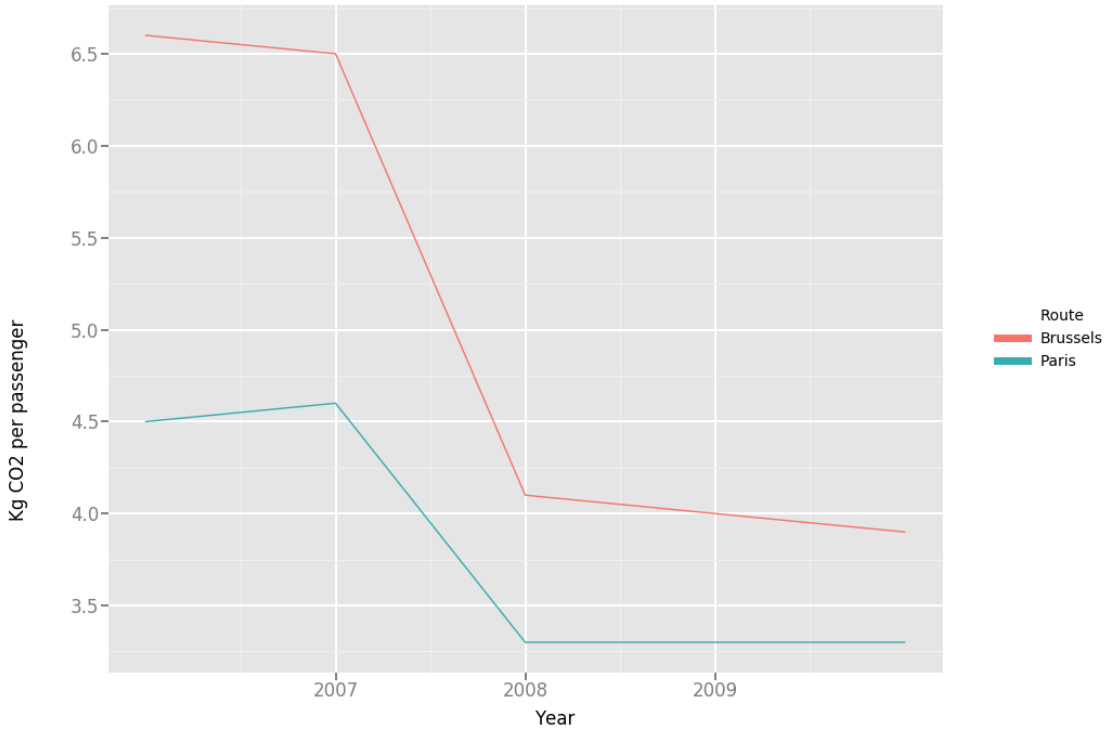


Fig. 5.6 kg CO_2 per passenger single trip, by supplier energy mix coefficients (derived from (Watkiss, 2009)

Table 5.3 Aviation kg CO_2 per passenger trip emissions by route (from (Watkiss, 2009))

Route	kg CO_2 per passenger trip
London City to Paris CDG	219.8
London City to Paris Orly	124.7
London Heathrow to Paris CDG	53.9
London Luton to Paris CDG	51.4
London City to Brussels	69.9
London Gatwick to Brussels	101.1
London Heathrow to Brussels	70.3

Hub travel - surface access

Surface access was not considered within the Watkiss report. In Chapter 3, methods for quantifying real-time transport conditions were proposed. These methods may be applied specifically to the travel hubs of London City Airport, London Heathrow Airport and St Pancras International Train station in order to assess their differing characteristics in the context of London Paris and London Brussels journeys. Of course, these datasets illustrate the influence of actual demand (through how it influences congestion) but don't explicitly quantify quantify demand in a formal origin and destination format. The true origins for passengers on London Paris and London Brussels services is not available in public datasets, and Eurostar's own understanding of this is understandably highly commercially sensitive. As a result a similar methodology to that from Chapter 4 is used, and LSOA zones are used to at least paint a picture of the population weighted areas of Greater London. These hubs may be compared in terms of journey times, per mode and how we may assess how these comparisons change over time. Table 5.4 illustrates the mean journey time, speed and distance for the different hubs by driving and table 5.5 presents the same information by public transport. In terms of surface access differences by mode, Heathrow illustrates the largest contrasts. Mean journey distance is significant at over 50km by driving compared to 37km via public transport, illustrating the tendency for road infrastructure to increase vehicle kilometre travelled (VKT) metrics (Banister, 2003). This increased mean journey distance for driving is somewhat negated by the mean journey speed of 15.92 m/s but still results in the largest mean journey time of the three hubs by driving. Generally, public transport illustrates more consistent mean journey speeds, with Heathrow exhibiting the fastest speeds as a result of express services which service it from central locations (e.g. Heathrow Express).

Such macro level statistics shown above provide useful context but are limited in their ability to explain the many dimensions of each hubs surface access footprint. The distribution of metrics around the mean measurements presented above illustrates a more reflective picture.

Table 5.4 Driving mean journey times, speeds and distances per hub

Hub	mean journey time	mean journey speed	mean journey distance
Heathrow Airport	3186 seconds	15.92 m/s	51.04 km
St Pancras International	2662 seconds	5.91 m/s	16.23 km
London City Airport	2793 seconds	9.25 m/s	25.83 km

Table 5.5 Public transport mean journey times, speeds and distances per hub

Hub	mean journey time	mean journey speed	mean journey distance
Heathrow Airport	5386 seconds	7.07 m/s	37.24 km
St Pancras International	2670 seconds	5.94 m/s	16.16 km
London City Airport	3785 seconds	5.87 m/s	22.34 km

Figure 5.7 shows the probability density function for both journey distance (top) and journey time (bottom) for each hub and for each mode. London City and St Pancras illustrate journey distance distributions with similar footprints for both public transport and driving. Contrastingly, Heathrow illustrates a bimodal distribution for driving, with a significant increase in journey distance by driving, where none exists by public transport. This is consistent with a large body of evidence which shows road infrastructure tends to increase distance travelled in an attempt to distribute demand (Banister, 2003). In terms of journey time there is an interesting distinction between London City and St Pancras. Despite similar distance footprints by mode, there is a distinct difference in terms of the journey time competitiveness per mode, with St Pancras exhibiting better public transport performance compared to London City. St Pancras' centrality and access to multiple Underground lines, dozens of bus services plus rail services as was shown in Chapter 3 potentially being explanatory factors.

This data may be used to better model the emissions footprint associated with travel to the travel hub, and. In the next section this data may be combined with the capital emissions and estimations and enable a more holistic analysis.

5.1.3 Holistic analysis - combining capital and operational costs

The Watkiss (Watkiss, 2009) report outputted the comparison between HSR and aviation per passenger trip shown in Table 5.6 and are generally consistent with those published by the UIC (UIC, 2011). The Eurostar journeys feature different emissions depending on the energy mix methodology used. The air coefficients consider average load factors from the CAA and consists of the average emissions from the serving airports (Heathrow and Luton in the case of Paris and Heathrow and Gatwick in the case of Brussels.). In the case of London Paris, HSR is

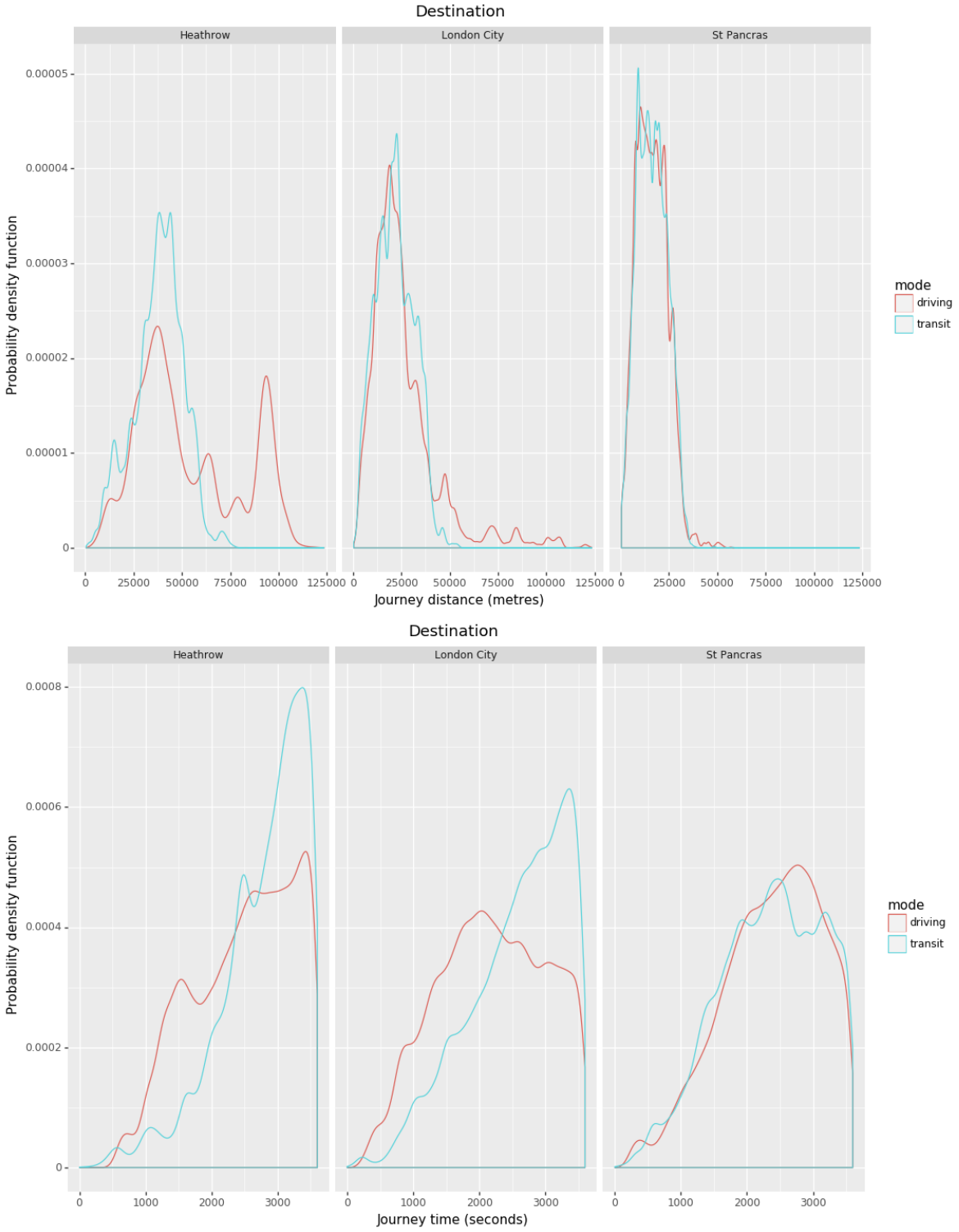


Fig. 5.7 Probability density plot of distance (top) and duration (bottom) of surface access to travel hub

7.5 times fewer emissions (average mix methodology) and 11.4 times fewer emissions (supplier mix methodology) than that for the equivalent journey by aviation. Similarly, in the case of London Brussels, HSR is 8.1 times fewer emissions (average mix methodology) and 13.2 times fewer emissions (supplier mix methodology than the equivalent journey by aviation).

Table 5.6 Watkiss summary of per passenger trip emissions between Eurostar and aviation in 2007

Route	Emissions (average mix)	Emissions (supplier mix)
Eurostar London to Paris	7.0kg CO_2 per passenger trip	4.6kg CO_2 per passenger trip
Air London to Paris	52.65 kg CO_2 per passenger trip	NA
Eurostar London to Brussels	10.6kg CO_2 per passenger trip	6.5kg CO_2 per passenger trip
Air London to Brussels	85.7kg CO_2 per passenger trip	NA

However, the Watkiss report did not consider surface access nor the capital CO_2 emissions of the infrastructure. In Section 5.1.1 the total capital cost attributable to Eurostar was estimated to be **1.495 Mt**CO_2. In order to assess the CO_2 pay back period of this capital CO_2 cost, a simplistic calculation to assess the possible emissions attributable to aviation if the HS1 link and Channel Tunnel did not exist may be carried out. From 2003 (when Section 1 of HS1 opened) to 2014 Eurostar carried a total of 84.6 million passengers (GMTR, 2014). If these emissions were distributed across all trips to 2014, it would equate to around 17.67kg CO_2 per passenger. The capital construction emissions per passenger trip would therefore be equivalent to 2.52 times and 1.6 times that of the operational emissions for London Paris and London Brussels respectively. Of the total 84.6million trips on Eurostar to 2007, 10 million are assumed to be induced, generated demand and the remaining 74.6million trips are distributed as 46.25 million trips to Paris and 28.35million trips to Brussels based upon historic service counts and load factors (GMTR, 2014). The result is an estimated cost of 2.44MtCO_2 for the London Paris route and 2.43MtCO_2 for the London Brussels route using the found Watkiss coefficients for aviation (in Table 5.6). Interestingly, the routes have a similar overall footprint despite the larger demand on the London Paris route which is counteracted by the larger per trip footprint for London Brussels. For the same assumptions, the Eurostar cost may be found to be 0.32MtCO_2 (average mix) or 0.21MtCO_2 (supplier mix) for London Paris and 0.30MtCO_2 (average mix) or 0.18MtCO_2 (supplier mix) for London Brussels, as shown in Table 5.7. Therefore the cost of the capital investment in CO_2 was paid off in **5.2 years** using average energy mix methodology or **4.6 years** using the supplier energy mix methodology, as shown in Table 5.8.

In Figure 5.8 the surface access emissions from the previously discussed surface access dataset is presented. The emissions footprint associated with public transport journeys is shown to be significantly less than that for driving, irrespective of hub used. The previously discussed

Table 5.7 Total emissions for London-Paris and London-Brussels trips (2003-2014)

Route	Type	Estimated total emissions (2003-2014)
London - Paris	Aviation	2.44MtCO_2
London - Brussels	Aviation	2.43MtCO_2
London - Paris	HSR	0.32MtCO_2 (average) or 0.21MtCO_2 (supplier)
London - Brussels	HSR	0.3MtCO_2 (average) or 0.18MtCO_2 (supplier)

Table 5.8 Capital CO_2 emission payback period estimations

Emissions payback period	HSR energy supply methodology
4.6 years	Supplier mix
5.2 years	Average mix

bimodal footprint for driving to Heathrow is again shown, with peaks at 5kg CO_2 and 15kg CO_2 evident. The driving footprint for Heathrow and London City likely further increases the relative performance of HSR for end to end journeys when considering surface access. This would therefore suggest that payback period for HS1 is likely even shorter than the 4.6 years (supplier mix methodology) and 5.2 years (average mix methodology) discussed previously.

At this stage, it is now possible to attempt to compare the capital emissions, hub journey emissions with the actual London Paris and London Brussels journey in order to illustrate their relative importance. Building upon the passenger counts from 2003 to 2014, Eurostar carried an additional 51.2 million passengers to 2017 (GMTR, 2014), (Eurostar, 2015), (Eurostar, 2016b), (Eurostar, 2017),(RailwayGazette, 2018) resulting in a total of 135.8 million passengers since 2003. We may now normalise the capital emissions per passenger to this date, resulting in around 11kg CO_2 per passenger trip. How the capital emissions may be attributed to passengers is shown in Figure 5.9.

This can be combined with the Watkiss operational emissions and the estimated hub surface access emissions, per hub and for both driving and public transport surface access to the hub, as is shown in Figure 5.10. The influence that the surfaces access method of transport to the hub has on overall emissions is shown to be significant, with flying to Brussels or Paris from Heathrow the most affected. Despite the increase in emissions as a result of the construction of HS1, and irrespective of the mode of transport used to access the hub, high speed rail is shown to exhibit significantly less emissions per trip.

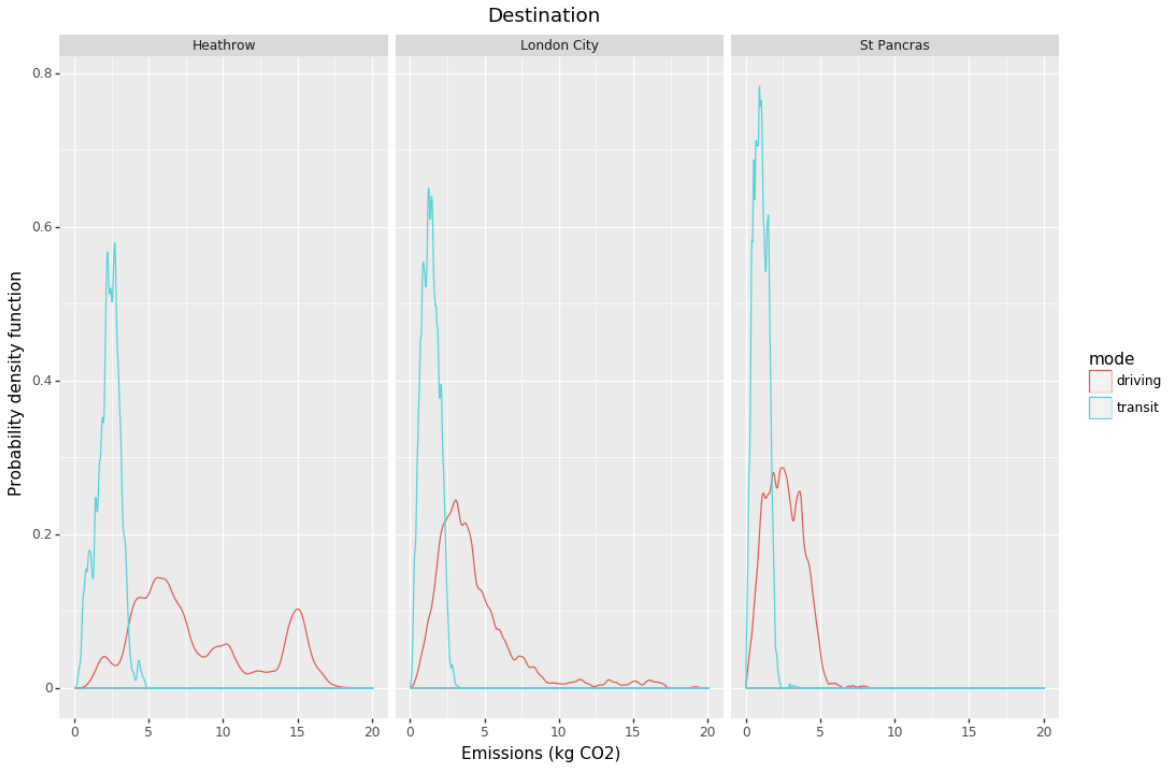


Fig. 5.8 Surface access emissions probability density function, by hub and by mode

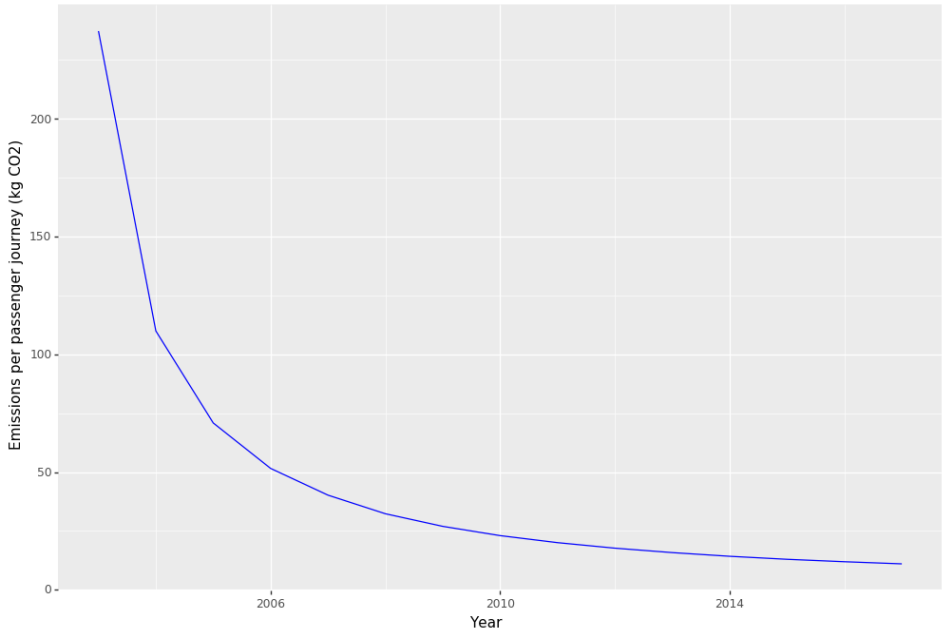


Fig. 5.9 Capital CO_2 emissions per passenger 2003 - 2017

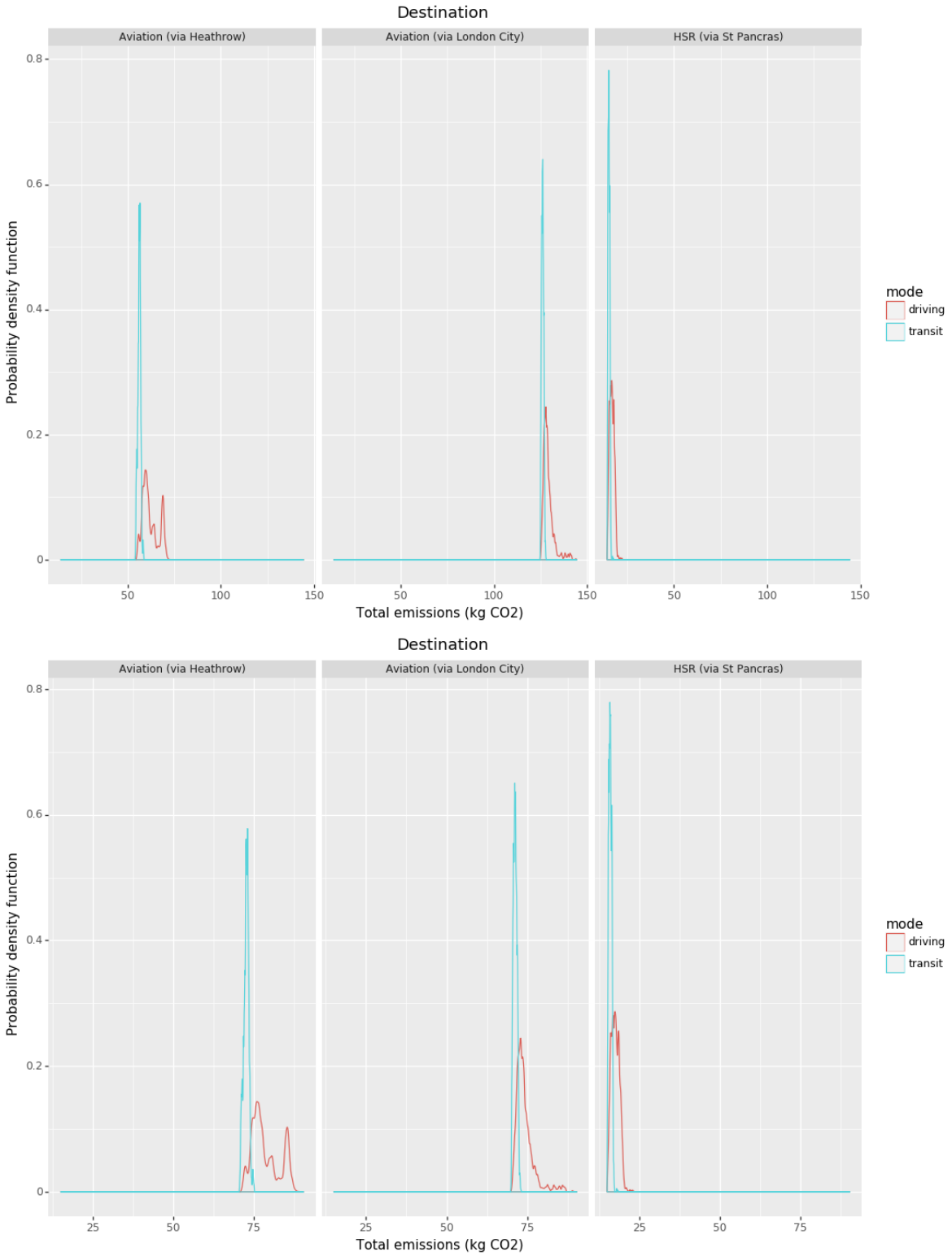


Fig. 5.10 Probability density plot for total estimated emissions for London Paris (top) and London Brussels travel (bottom), dissagregated by surface asses mode

5.1.4 The behavioural drivers

As was discussed in Chapter 4 there are a range of different heuristic methods which attempt to compute the above metrics into a form from which a human decision may be modelled. The journey times may be disaggregated by type (in vehicle travel, weighting etc), merged with the formal financial costs (e.g. fare) and a generalised cost found using the same methodology from Section 4.1.5. The probably density function for the computed generalised cost is presented for each hub and by mode in Figure 5.11.

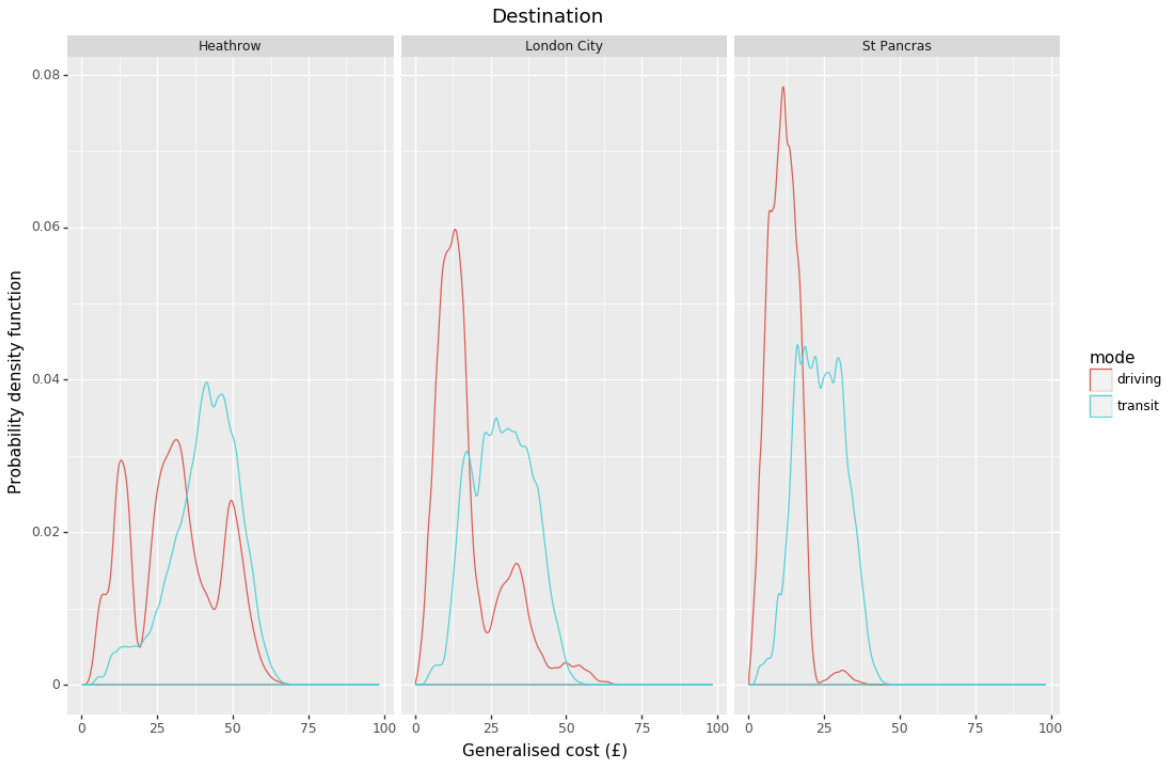


Fig. 5.11 Probability density plot of generalised cost of surface access to travel hub

The surface access generalised costs computed up until this point do not include the ticket of a flight or HSR ticket for the entire end to end journey. Flight tickets and more recently HSR tickets have been using dynamic pricing structures which consider a range of factors to provide extremely temporally dynamic prices. Pricing changes dependent on how far in advance a booking is made, the time of travel (weekend versus weekday), the class of travel and other factors such as weather and the presence of cultural/sporting events. The algorithms which generate these prices are extremely commercially sensitive and not publicly available, nor is the harvesting of pricing information from seller websites permitted under the terms and conditions of access. Information on how far in advance travellers generally purchase tickets is also not publicly available. It is therefore very challenging to estimate the likely pricing distribution for

aviation and HSR. A manual study was carried out over a period of one month in 2016, where British Airways, Air France and Eurostar ticket prices were assessed. Over this period, the minimum and maximum prices for travel were harvested and for extrapolation to an individual journey, prices were assigned randomly between these two points. Eurostar is shown to exhibit the largest range in prices and London City generally exhibiting the most expensive prices. The summary of these assumptions are shown in Table 5.9.

Table 5.9 London - Paris travel assumptions

Mode	Hub	Journey time	Wait time	Price range
Aviation	Heathrow Airport	55min	120min	£37 - £150
Aviation	London City Airport	55min	90min	£60 - £200
High Speed Rail	St Pancras	135min	45min	£44 - £191

Since the distribution of ticket prices between the minimum and maximum range identified is unknown, it is not possible to assess the overall distribution of generalised costs. However, it is possible to assess the extreme points in the range and the proportional cost of the waiting time and in vehicle time for each destination and mode. This is shown in Figure 5.12 where the generalised cost weighted journey time footprint differences between aviation and HSR are evident in their proportional importance. Heathrow recommends arrival two hours before a flight and the waiting proportion of the generalised cost is shown to be extremely significant. Contrastingly, HSR via St Pancras shows the highest proportion of in vehicle costs, but these are still less than aviation from both Heathrow and London City in both pricing scenarios. The shorter pre-departure arrival time from London City result in it dipping below Heathrow, despite the generally higher up front ticket costs. Of course, the assumed value of time metrics here and limited knowledge of ticket pricing heavily caveat the results, but illustrate the relative trade off's shown in terms of generalised cost.

Aviation generally offers a reduced amount of time actually in vehicle time, but significantly more time is required waiting (e.g. due to security), plus airports are generally further from central economic districts than train stations. Those airports closer to central locations often have capacity constraints and thus are usually more expensive, as seen with London City above. Conversely, HSR is longer in transit, but the value of time is lower as it is more productive and the amount of wait time required is significantly less. In order to consider these factors, these hub generalised costs may be added to the large surface access dataset and the journey to a hub considered with the journey from the hub to Paris. This total generalised cost per hub is shown in Figure 5.13 where HSR and St Pancras are shown to exhibit the most competitive generalised cost functions. For all modes and hubs, public transport is shown to illustrate generally more expensive generalised costs compared to those by driving.

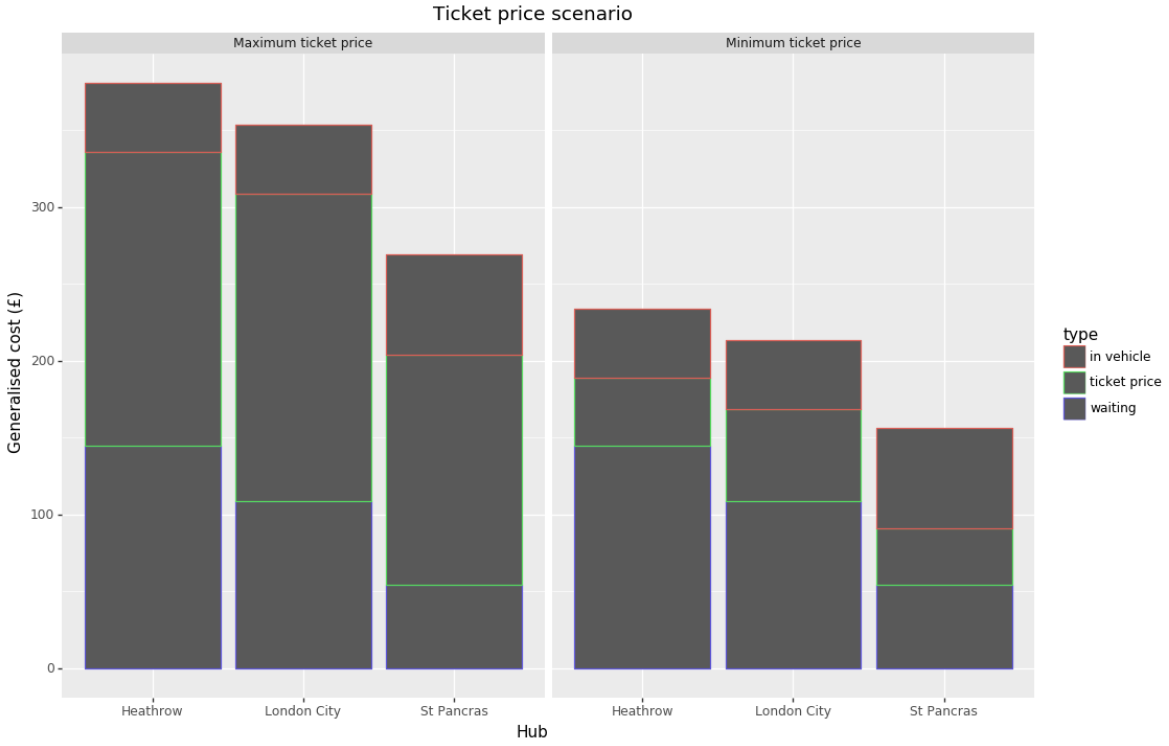


Fig. 5.12 Generalised cost of hub - Paris travel and constituent parts per hub, for maxium and minimum ticket price scenarios

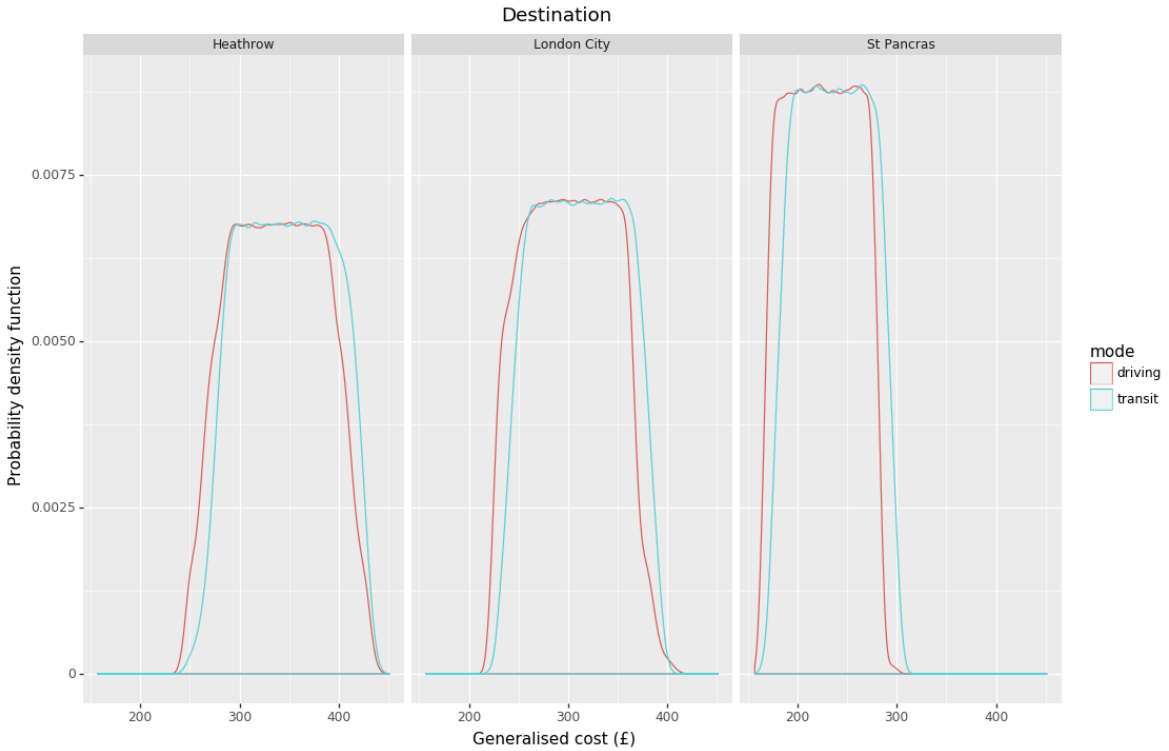


Fig. 5.13 Probability density plot of London - Paris total generalised cost

The above trends indicate the kind of variations that occur per mode and hub, but these macro statistics negate the importance of a specific point in time for a given traveller. It is the relative competitiveness of each mode, against other modes, for a given person, in a given location and given point in time that is the dominant criteria for decision making. In order to consider this, the different options may be considered individually over discrete time periods. This is shown in Figure 5.14 where the proportion of trips which are lowest generalised cost is shown by both time of day and day of week to illustrate the temporally dynamic nature of the changing conditions. St Pancras and HSR is evidently the most dominant option at all times of the day, with the exception of 3pm, where significant Eurostar service timetabling differences to aviation and to a less extent traffic conditions reduce journey times to Heathrow sufficiently to challenge its dominance.

5.2 Chapter Summary

5.2.1 Summary

1. Hypothesis

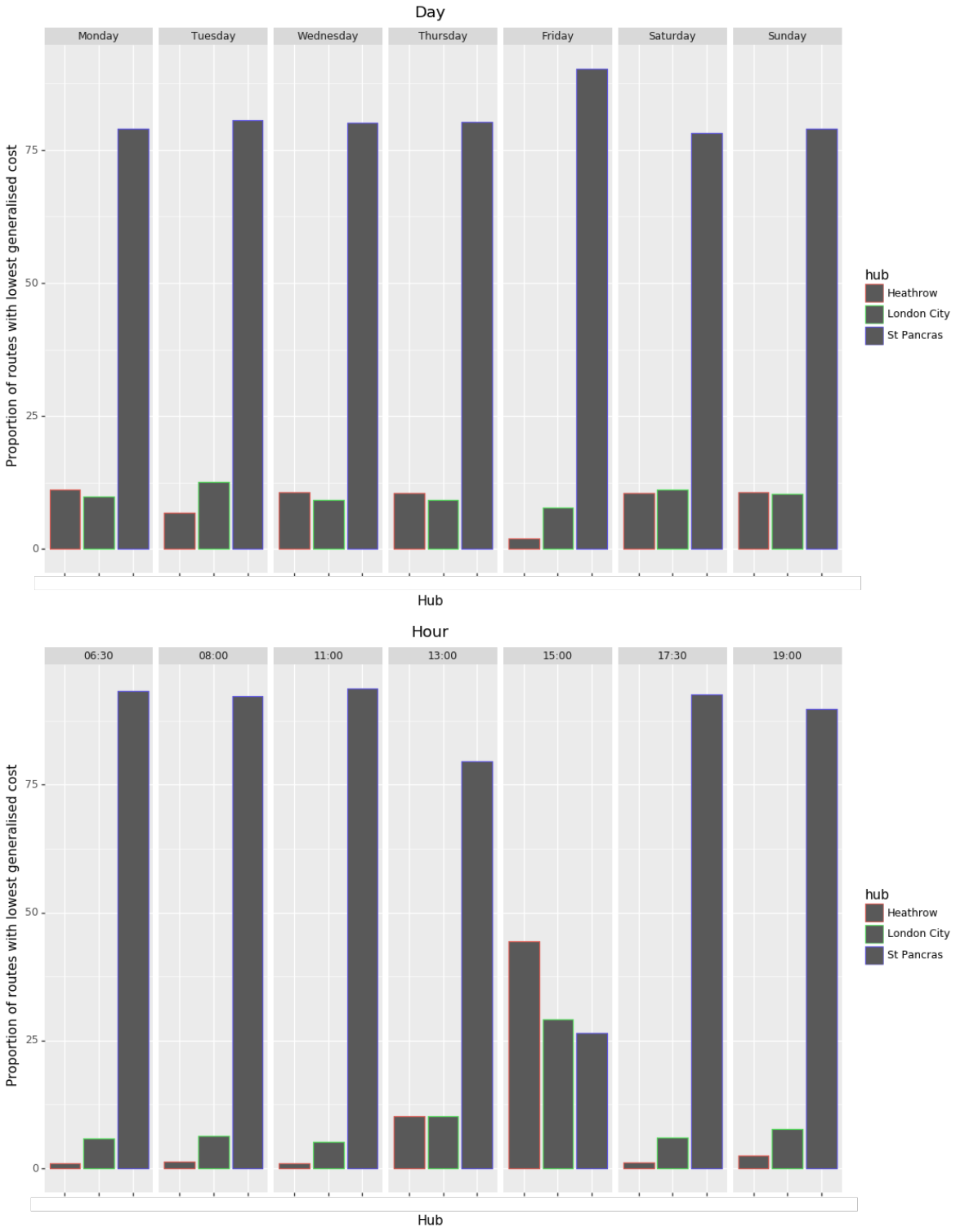


Fig. 5.14 Lowest generalised cost hub and mode, by day of week (top) and hour of day (bottom)

The inclusion of the capital CO_2 emissions and surface access emissions in HSR versus Aviation LCA analyses may have a significant impact on the outputs.

2. Novel contributions

   The capital CO_2 cost is shown to have a non-trivial impact on the emissions footprint, equating to an estimated 1.495 MtCO_2 for HSR. Despite this large cost, the savings in HSR operational emissions compared to that of aviation mean this capital cost was estimated to have been paid off in 4 to 5 years depending on HSR energy mix methodology.

   By using spatially and temporally dynamic crowd sourced data a richer and more nuanced understanding of the nature of different major, international transportation hubs was revealed. It is reassuring to see that the new data and new methods produce macro results which are in line with that generally expected. Such insights provided by the detail of these methods show promise and may enable for more targeted interventions in the context of modal shift ambitions to reduce CO_2 emissions, without requiring a reduction on demand.

5.2.2 Conclusions

The inclusion of the emissions associated with the construction of HS1 infrastructure make a distinct difference to the per trip operational emissions. Despite this, the relative poor performance of aviation for similar routes results in a very short payback period. The modelling of hub travel further compounds this, where the HSR hub of St Pancras is shown to have significantly less associated emissions for similar trips, likely further reducing the pay back period of the infrastructure.

The generalised cost computations begin to aid our understanding of the influence of hub travel to the larger trip. The impact of temporally dynamic connectivity illustrates just how much variability there can be for different hubs, irrespective of the actual Heathrow, St Pancras or London City to Paris/Brussels trip itself. The influence of frequent flier miles, extremely opaque pricing models and a lack of known data on true origins means that only a part of the system is understood here.

5.2.3 Limitations

1. The true demand for London Paris and London Brussels trips was not well understood and thus assumptions were made in projecting the likely origin to hub emissions.

2. Station operational emissions and life cycle costs of the electricity supply infrastructure were not included. Average load factors were taken an extrapolated across all services for both HSR and aviation, impacting on the fidelity of the operational emissions estimations.

3. The surface access emissions were estimated using standard distance based emission coefficients which are poor at reflecting traffic and load factor emissions. The biggest challenge in estimating surface access emissions is the underlying passenger demand, the population centred LSOA zones were used to illustrate the range of possible values.

4. A significant period of time has passed since the Eurostar studies were carried out (2006 and 2009) and there have been changes to HSR services, aviation services, HSR rolling stock and airplanes.

Chapter 6

Conclusions

6.1 Summary

In Chapter 2 the literature highlighted the need for a more holistic and context specific set of analyses when considering the sustainability of different modes of transport. In Chapter 3 different data sources were utilised to compute temporally and spatially dynamic statistics on the performance of different transportation networks. These data sources permitted insights into actual, real-world conditions and painted a contrasting picture to the idealised and standardised metrics usually used as model inputs in transport modelling. Figure 6.1 shows one such illustration where journey times along a given road are shown to vary by day of week and time of day. The challenges with standard AM, PM and IP (inter-peak) methodology for capturing temporal changes is clearly evident in light of such variability. In Chapter 4 these data sources were combined in an agent based modelling framework for modal choice and route assignment. In the context of utility theory decision making heuristics, the framework illustrated how generalised cost may fluctuate as a result of changing conditions and different heuristics, as shown in Figure 6.2. The use of new computer science paradigms (predominately that of MapReduce) enabled for large numbers of agents to be simulated, as shown in Table 6.1. The computational footprint of this implementation is shown in Figure 6.3 and illustrates the value of the MapReduce paradigm for the ABM computation. However, this figure also illustrates the challenges faced when the spatial complexity of the simulation increases and the graph computation costs become dominant.

In Chapter 2 the High Speed 1 (HS1) line was specifically analysed in order to better understand the emissions associated with its use. Capital CO_2 emissions were shown to be significant, as was the method of access to the respective hub. Overall, it was shown that the capital CO_2 emissions accrued as part of the infrastructure's construction were paid of in around 5 years when compared to the likely emissions footprint via aviation. In Figure 6.4

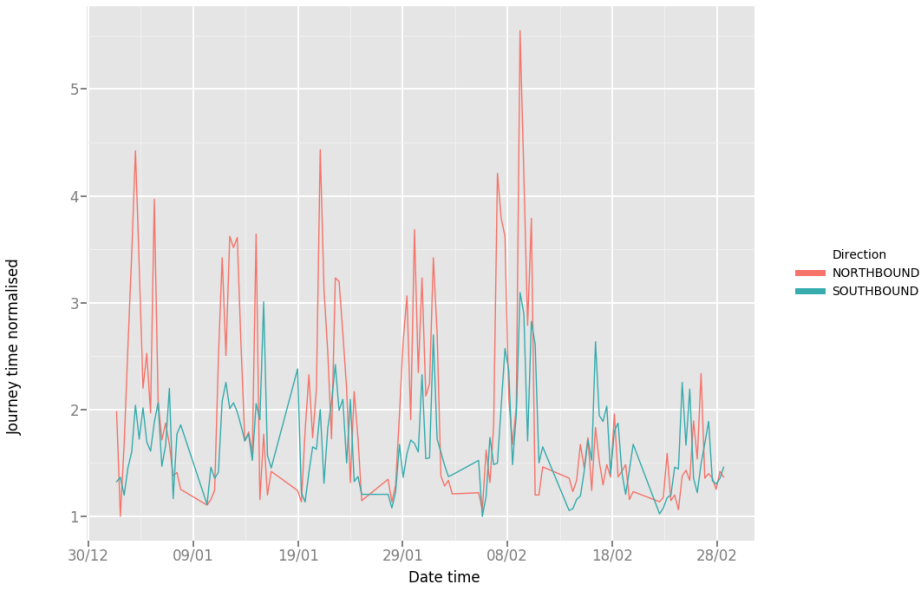


Fig. 6.1 The variability of journey times

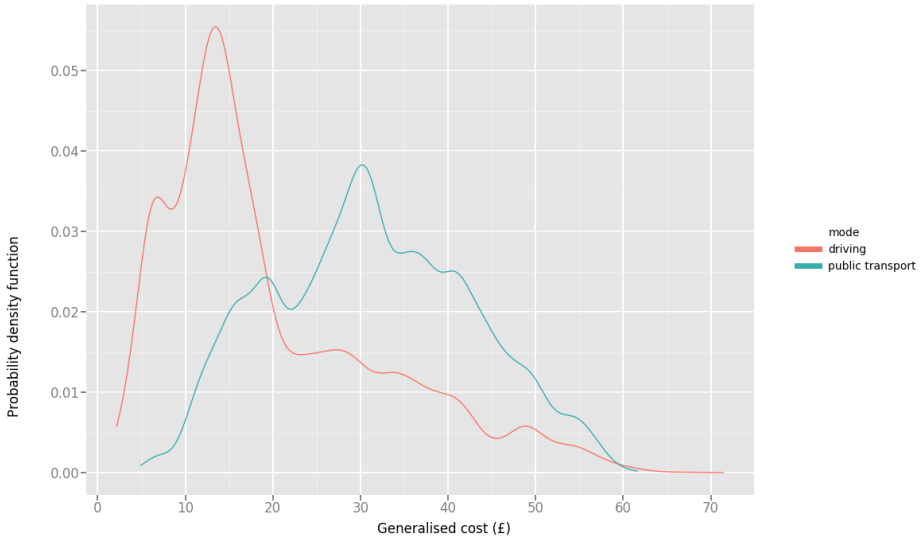


Fig. 6.2 Journey variability in the context of generalised cost - heuristics and the human interface

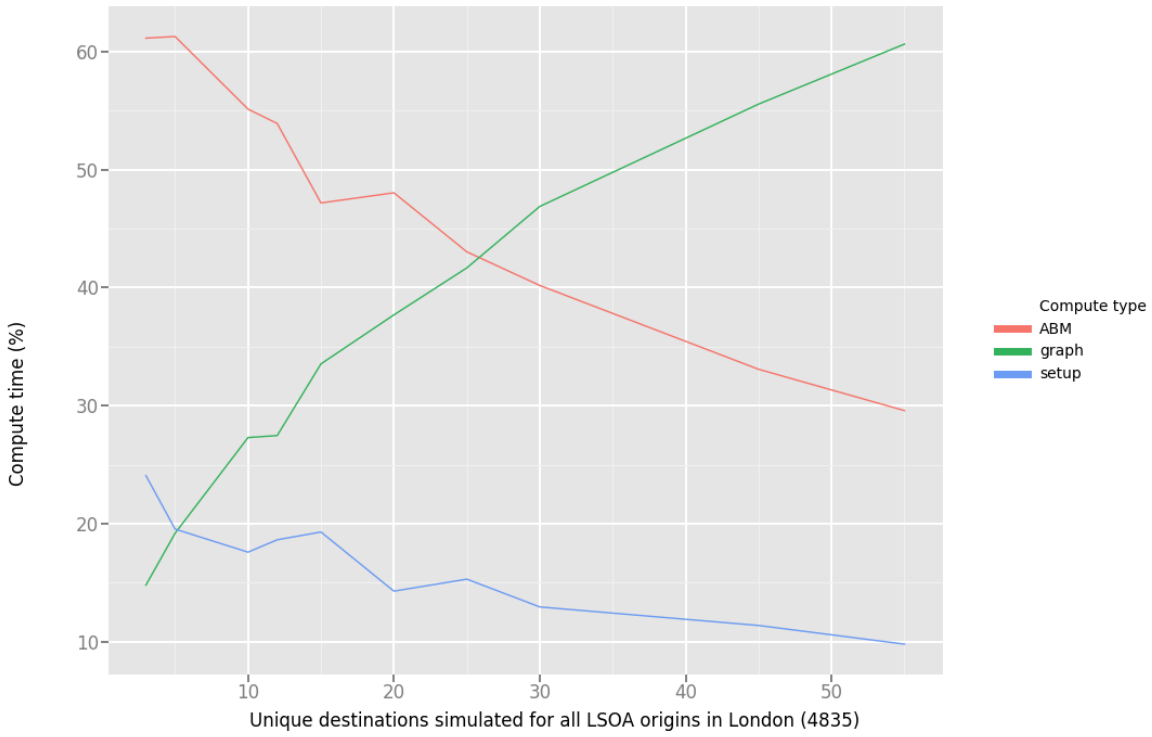


Fig. 6.3 The computational value footprint of the proposed implementation

Table 6.1 Computational performance of ABM

Hardware	Time (Seconds)	Agents simulated	Time resolution
Single node	9383	34,694,852	9am to 1pm (1hour slices)
6 node cluster	1286	34,694,852	9am to 1pm (1hour slices)

the footprint for access to the respective hub (via public transport or driving) is combined with the Watkiss operational emissions (Watkiss, 2009) and the estimated capital emissions to illustrate the contrasting options for London Paris and London Brussels travel. The influence that the surfaces access method of transport to the hub has on overall emissions is shown to be significant, with flying to Brussels or Paris from Heathrow the most affected. Despite the increase in emissions as a result of the construction of HS1, and irrespective of the mode of transport used to access the hub, high speed rail is shown to exhibit significantly less emissions per trip.

6.2 Specific conclusions

The underlying hypothesis that this research aims to provide data, methods and tools to support is that finer resolution temporal and spatial inputs to a more complex, dynamic model will result in more accurate predictions. Chapter 3 presented a series of techniques which enabled for the supply side for the transportation network to be better understood. This was done to a degree of success, and illustration made to how the new network representation impacts on modelling outcomes for traveller decision making. Further work is required to fully understand some of the fluctuations exhibited in some aspects of the data harvested. However, although these methods include the influence of demand, they do not enable for its quantification. The challenge of understanding full origins and destinations remains significant. Travel surveys maintain their inability to scale sufficiently and struggle with sample bias (TSC, 2016). Data from mobile phone companies which uses cell towers triangulation presents sufficient sample size but with poor accuracy (Spirito, 2001) leading to issues with true origin/destination detection and route and modal choice detection. Conversely, major technology companies with enormous sample sizes (Google and Apple) hold large location device data repositories but use them in limited, mostly commercially driven exercises. The raw demand side data of this picture poses profound privacy and ethical challenges and in light of recent regulatory changes in the EU paints a constrained picture for its usage. Supply side data is generally freely available and comes with relaxed regulation due to the lack of personal data.

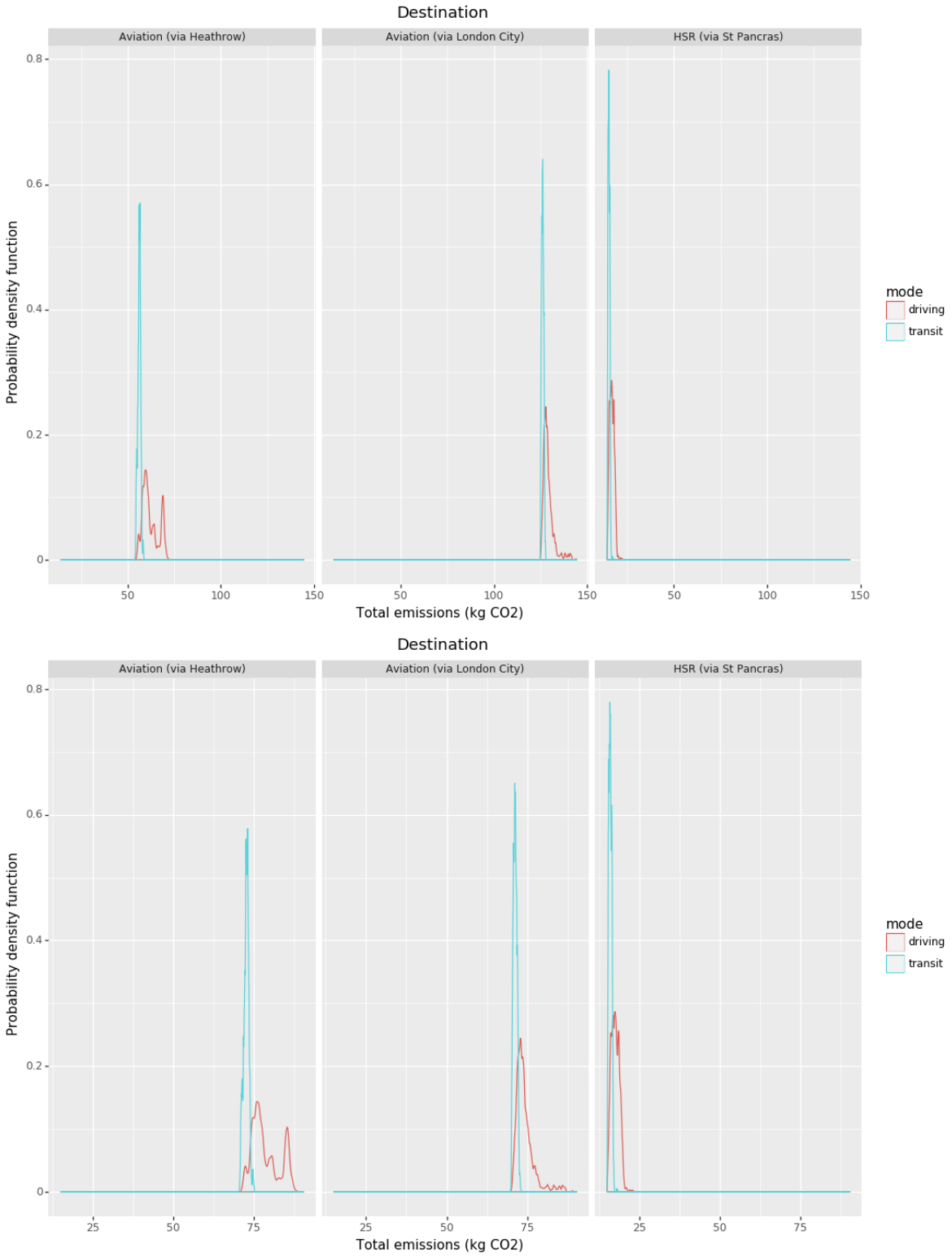


Fig. 6.4 Probability density plot for total estimated emissions for London Paris (top) and London Brussels travel (bottom), disaggregated by surface asses mode

This research makes contributions in terms of quantifying the supply side of the public transport and road networks, and in Chapter 4 provides tools and a scalable methodology for utilising this data in a dynamic agent based model.

6.3 General conclusions

The data sources used here may be polled at even finer resolution, potentially magnifying the present computational demand by a significant amount. It is reassuring that at a macro scale these data sources exhibit behaviour which is generally expected, giving confidence in the veracity of the methods. Of course, the challenge of how much fine resolution is *needed* to answer urban questions remains unknown. The relationship between computational complexity and model accuracy is most certainly not linear and there will a point at which increases in model complexity may have little or no value in terms of improving the accuracy of the model output, as is graphically conceptualised in Figure 6.4. In the crudest sense, the strong relationship between model complexity and model run time results in a simple heuristic whereby a model which runs in half the time of another model, yet only has a less than 5% difference in accuracy may be deemed acceptable. The key contribution of this thesis is a modelling platform for which such models may now be implemented and the true relationships hypothesised in Figure 6.4 may be derived for a range of different contexts. It is hypothesised that the value/cost transition curve consists of three distinct parts - 1. the *data* 2. the *model* and 3. the *heuristics*. The use of the proposed data harvesting in Chapter 3 at AM, PM, IP temporal resolution and MSOA level spatial resolution would likely equate to the section labelled **A**. This illustrates the maximum value improved data can give a static, deterministic (e.g. 4 stage) model before the fidelity for the data becomes beyond the scope of which the model can use. At this point, there is a need to move to a new modelling paradigm (as shown by label **B**) which permits for such data to be used in a meaningful way. The lower portions of this curve likely involve the use of the proposed ABM on top of a relative static trip generation and trip distribution model. However, quickly the resolution and complexity of such a hybrid model begins to lose value as it sits upon a cruder input and the transition must then be made to a more activity based representation of demand, where trips are chained meaningfully. The final and most complex transition occurs at **C** where the transition from theory driven heuristics surrounding human decision making and logic may begin to be applied in a more data driven way. For example, using a range of heuristics from literature which enable for different types of journeys and roles (by the same agent, and by different agents) to be propagated throughout an activity based model. If, at this point computations can be sufficiently quick and revealed agent choices calibrated against previously modelled behaviour with a degree of feedback currently not

possible, the use of an agnostic modelling framework may permit for real-time calibration in a data-driven way (Casey et al., 2017).

As such, a key conclusion of this research is the need for a multi-faceted, incremental approach to model development, which appropriately pairs a modelling paradigm to an appropriate resolution data source(s) with heuristics which can meaningfully relate agents to each other. This transition, with labelled examples is illustrated below:

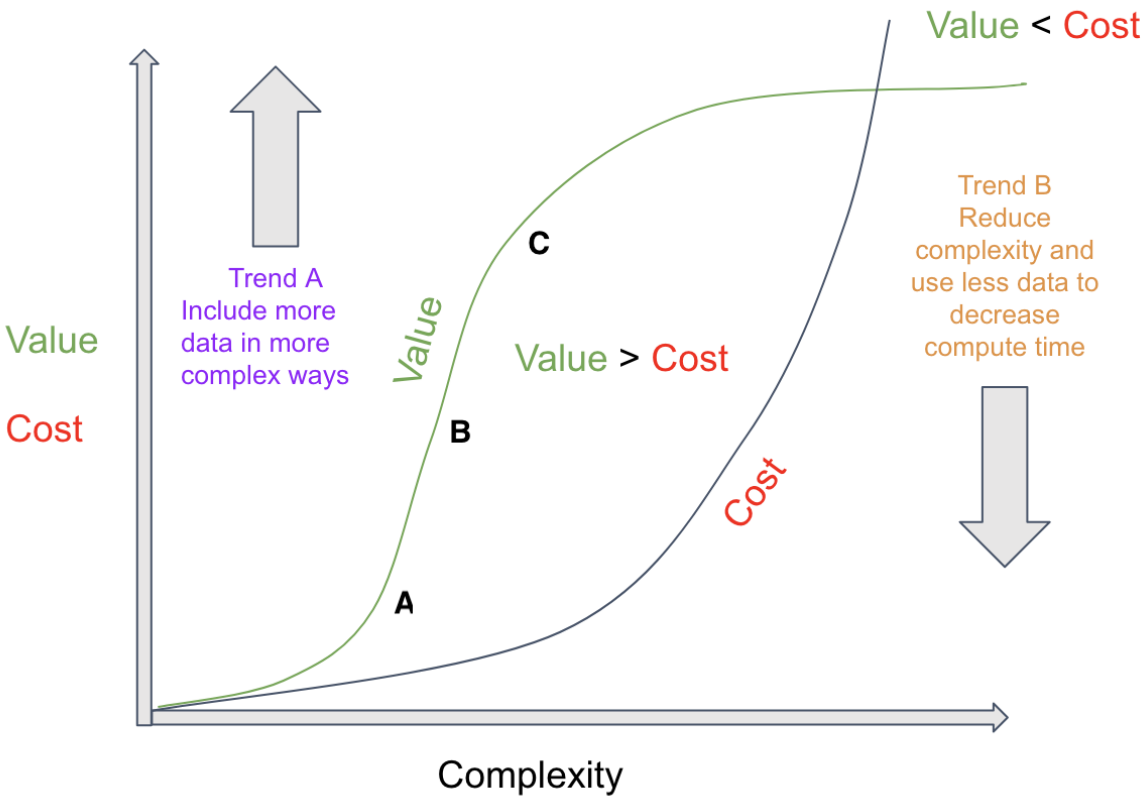


Fig. 6.5 Proposed relationship between value, cost and complexity

1. Data: Survey data -> Aggregated geospatial data (e.g. telco data) -> individual geospatial data (e.g GPS traces)
2. Models : Deterministic and static models -> Stochastic and dynamic models
3. Heuristics : Theory driven (e.g Discrete choice) -> Data driven (e.g real time calibration)

6.4 Future work

6.4.1 Real-time calibration & decision making heuristics

The proposed framework attempts to permit the use of a range of different decision making heuristics, rather than prescribing one. In Chapter 2 the challenges associated with the many variants of rational, perfect information informed decision making was highlighted. Presently, the calibration and validation exercises are discrete occurrences - a linear process of data gathering, model creation, calibration and finally validation occurs. From this point, the model projections do not change until the next iteration may be years or even decades away. There is little in the way of actual feedback, where modelled behaviour is related to actual exhibited behaviour in a meaningful way. The value of such a methodology is theoretically profound in terms of the ability to aid our understanding of route choice and modal choice, but has limited value in the short term for answering longer term, land use implications. Real-time calibration is theoretically limited by the historic precedence of events within an input dataset and in the context of rich datasets, data sources such as location device data have only been in existence for 10 years. Despite these challenges, the use of a running, continuous methodology will only become more valuable as time passes. The recent success of AI and specifically neural-net methodologies offers significant opportunities as method for deriving new heuristics within an agent based model. However, the transparency of such methodologies and the downstream impacts these have on model validation are profound as the complexity of such a model becomes increasingly likely to be unintuitive. The standard model validation techniques for transport planning such as link counts and journey times are increasingly becoming inappropriate and the need to use potentially extremely intrusive individual level validation metrics poses ethical challenges which are not insignificant.

6.4.2 Model Inter-modality

Cycling, walking *etc* may be represented by constraining road network by type and using more representative edge weights. Connecting these modes is a separate challenge. A true inter-modal graph enables a series of cross routes, which cannot be represented by distinct layers, for example the user of ride sharing technologies (such as Uber) to solve first and last mile challenges to public transport services. The modelling of the transportation network as a truly inter-modal graph also offers the opportunity of removing the entire modal choice step from modelling. A modal choice is computed as part of the optimum path across the graph, dependent on the input heuristics relating to what is valued and how (journey time, headways, wait time etc), rather than being a decision pre-routing. As services become more integrated

and payment barriers begin to disappear across modes it is likely that our existing notions of modal definition will begin to be lost. This is generally a positive step, as it should enable for areas which are not well served by public transport to be served in a more integrated fashion with services by companies such as Uber and Lyft.

However, in the context of the computational challenge posed by graph algorithms discussed in Chapter 4 and graphically illustrated in Figure 6.3 this poses a significant challenge. The separate computation of public transport and other graphs enables for the graph to be simplified and also aids speed by enabling this to be done in parallel. The transition from a multi-modal to inter-modal graph means an order of magnitude change in the complexity of the computation required, with no ability to do this in parallel. Despite this, the potential value is enormous in terms of enabling questions surrounding the use of AVs and rideshare technologies to serve first and last mile connectivity issues. It is a significant barrier which transport modelling must overcome in order for the binary public transport, driving models of the past to be able to handle mobility as a service use cases where a given user may use various different systems multiple times of a day as is beginning to manifest in reality.

In terms of meaningful research avenues forward for reducing graph computation times, there are 3 distinct options:

1. Algorithmic optimisation. The use of Fibonacci and Binary heaps with Dijkstra has shown significant value, as has highway hierarchies (Sanders and Schultes, 2006a), goal direction, tesselation, seperators, bidirectional search and transit node routing (Sanders and Schultes, 2006b) methodologies.

2. Vertical scaling. Although recent trends have pushed towards horizontal MapReduce'esque methodologies for "big-data" workloads, vertical scaling through the use of faster hardware remains a legitimate option.

3. Graphical Processing Units (GPUs).

6.4.3 Public transport

Presently there is no crowding feedback mechanism within the ABM. Context specific volume delay functions have been used as the mechanism to permit feedback between agents in different time steps on the road network. In the case of public transport, the presented implementation did not feature the ability to generate congestion on public transport modes. The first challenge to this is the technical barriers relating to handling the data volume. The scalable nature of the proposed framework and the exhibited performance shown in Chapter 4 shows sufficient capacity to handle this. However, the data load poses challenges for operators such as TfL who

may struggle to share the data in a timely and meaningful way as they do not have the tools to do so. Secondly, the institutional and regulatory challenges relating to accessing of such data with significant privacy concerns is an evolving challenge.

References

Akcelik, R. (1991). Travel time functions for transport planning purposes: Davidson's function, its time dependent form and alternative travel time function. *Australian Road Research*, 21(3).

Allwood, J. and Cullen, J. (2009). Steel, aluminium and carbon: alternative stragies for meeting the 2050 carbon emissions targets. *World Economic Forum*.

Amdahl, G. M. (1967). Validity of the single processor approach to achieving large scale computing capabilities. In *Proceedings of the April 18-20, 1967, spring joint computer conference*, pages 483–485. ACM.

Antrim, A., Barbeau, S. J., et al. (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*.

Apache (2016). Apache spark. Technical report, Apache Foundation.

ApacheGiraph (2016). Apache giraph is an iterative graph processing system built for high scalability. Technical report, Apache.

Artikis, A., Weidlich, M., Schnitzler, F., Boutsis, I., Liebig, T., Piatkowski, N., Bockermann, C., Morik, K., Kalogeraki, V., Marecek, J., et al. (2014). Heterogeneous stream processing and crowdsourcing for urban traffic management. In *EDBT*, volume 14, pages 712–723.

Bak, M., Casey, G., and Kumar, K. (2016). Sierra charlie. Technical report, University of Cambridge.

Balmer, M., Axhausen, K., and Nagel, K. (2006). An agent based demand modeling framework for large scale micro-simulations, paper submitted for the 85th annual meeting of the transportation research board. *TRB, Washington DC*.

Balmer, M., Rieser, M., Meister, K., Charypar, D., Lefebvre, N., Nagel, K., and Axhausen, K. (2009). Matsim-t: Architecture and simulation times. *Multi-agent systems for traffic and transportation engineering*, pages 57–78.

Banister, D. (2003). *Transport planning: In the UK, USA and Europe*. Routledge.

Barney, B. (2016). Introduction to parallel computing. Technical report, Lawrence Livermore National Laboratory.

Barth, D. (2009). The bright side of sitting in traffic: Crowdsourcing road congestion data. Technical report, Google.

Batty, M. (2007). *Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals*. The MIT Press.

Batty, M., Besussi, E., and Chin, N. (2003). Traffic, urban growth and suburban sprawl. *UCL*.

Batty, M. and Torrens, P. M. (2005). Modelling and prediction in a complex world. *Futures*, 37(7):745–766.

Bellman, R. (1958). On a routing problem. *Quarterly of applied mathematics*, 16(1):87–90.

Ben-Akiva, M. E. and Lerman, S. R. (1985). *Discrete choice analysis: theory and application to travel demand*, volume 9. MIT press.

Bentley, J. L. (1975). Multidimensional binary search trees used for associative searching. *Communications of the ACM*, 18(9):509–517.

Bertaud, A. and Richardson, H. W. (2004). *Transit and Density: Atlanta, the United States and Western Europe*. London: Ashgate.

Booz (2012). Review of hs1 demand forecasts. Technical report, Booz.

Bray, T. (2017). The javascript object notation (json) data interchange format. Technical report, JSON.

Brockway Jr, D. A. (2007). *The impact of a general aviation airport on surrounding land use patterns: Richard Lloyd Jones Jr. Airport*. ProQuest.

Brooks, F. (1987). *No silver bullet*. April.

C40 (2007). Deadline 2020. Technical report, C40.

Carpenter, J. and Lees, L. (1995). Gentrification in new york, london and paris: an international comparison. *International Journal of Urban and regional research*, 19(2):286–303.

Cascetta, E., Russo, F., Viola, F. A., and Vitetta, A. (2002). A model of route perception in urban road networks. *Transportation Research Part B: Methodological*, 36(7):577–592.

Casey, G., Soga, K., Silva, E., and Guthrie, P. (2016). *Understanding traveller decision making - a crowd sourced big data analysis of the London Travel Demand Survey*, pages 741–746. Cambridge University Press.

Casey, G., Soga, K., Silva, E., Guthrie, P., and Kumar, K. (2017). A scalable agent based multi-modal modeling framework using real-time big-data sources for cities. *Transportation Research Board*.

Catala, M., Dowling, S., and Hayward, D. (2011). Expanding the google transit feed specification to support operations and planning. Technical report, Google.

CCC (2012). Scope of carbon budgets statutory advice on inclusion of international aviation and shipping. Technical report, Committee on Climate Change.

CCC (2018). Reducing uk emissions. 2018 progress report to parliament. Technical report, Committee on Climate Change.

Cervero, R. (2004). *Transit-oriented development in the United States: Experiences, challenges, and prospects*, volume 102. Transportation Research Board.

Cervero, R. and Kockelman, K. (1997). Travel demand and the 3ds: density, diversity, and design. *Transportation Research Part D: Transport and Environment*, 2(3):199–219.

Cervero, R. and Landis, J. (1997). Twenty years of the bay area rapid transit system: Land use and development impacts. *Transportation Research Part A: Policy and Practice*, 31(4):309 – 333.

Chau, C., Soga, K., O'Riordan, N., and Nicholson, D. (2012a). Embodied energy evaluation for sections of the uk channel tunnel rail link. *Proceedings of the Institution of Civil Engineers - Geotechnical Engineering*, 165(2):65–81.

Chau, C., Soga, K., O'Riordan, N., and Nicholson, D. (2012b). Embodied energy evaluation for sections of the uk channel tunnel rail link. *Proceedings of the Institution of Civil Engineers-Geotechnical Engineering*, 165(2):65–81.

Chester, M. and Horvath, A. (2009a). Environmental assessment of passenger transportation should include infrastructure and supply chains. *ENVIRONMENTAL RESEARCH LETTERS*, 1.

Chester, M., Pincetl, S., and Allenby, B. (2012). Avoiding unintended tradeoffs by integrating life-cycle impact assessment with urban metabolism. *Current Opinion in Environmental Sustainability*, 4(4):451 – 457. Human settlements and industrial systems.

Chester, M. V. and Horvath, A. (2009b). Environmental assessment of passenger transportation should include infrastructure and supply chains. *Environmental Research Letters*, 4(2):024008.

Citymapper (2017). Using natural language processing to route around nyc's subway disruptions. Technical report, Citymapper.

CommuteStream (2016). Transport for london gtfs exporter. Technical report, Commute Stream.

Crooks, A., Castle, C., and Batty, M. (2008). Key challenges in agent-based modelling for geo-spatial simulation. *Computers, Environment and Urban Systems*, 32(6):417 – 430. GeoComputation: Modeling with spatial agents.

Crooks, A. T. (2007). The repast simulation/modelling system for geospatial simulation. *University College London*.

Csardi, G. and Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal Complex Systems*, 1695(1695).

Danziger, P. (2010). Big o notation. *Source internet: http://www. scs. ryerson. ca/˜ mth110/Handouts/PD/bigO. pdf, Retrieve: April*.

De Boor, C., De Boor, C., Mathematicien, E.-U., De Boor, C., and De Boor, C. (1978). *A practical guide to splines*, volume 27. Springer-Verlag New York.

de Silva, M. and Paris, R. (2014). Building crossrail-a holistic approach to sustainability. In *ICSI 2014: Creating Infrastructure for a Sustainable World*, pages 868–879. ASCE.

Dean, J. and Ghemawat, S. (2008). Mapreduce: simplified data processing on large clusters. *Communications of the ACM*, 51(1):107–113.

DECC (2008). Climate change act 2008. Technical report, UK Government.

DECC (2014). 2014 uk greenhouse gas emissions, final figures. Technical report, Department for Energy and Climate Change.

DECC (2015). Local authority carbon dioxide emissions estimates 2013. Technical report, Department of Energy and Climate Change.

Dedring, I., de Cani, R., and Iswalt, M. (2017). Traditional decision making. Technical report, Arup.

DEFRA (2007). Air quality and climate change: A uk perspective. Technical report, Department for Environment, Food and Rural Affairs.

Dequidt, T. C. E. (2012). Life cycle assessment of a norwegian bridge. *NTNU*.

Deville, P., Linard, C., Martin, S., Gilbert, M., Stevens, F. R., Gaughan, A. E., Blondel, V. D., and Tatem, A. J. (2014). Dynamic population mapping using mobile phone data. *Proceedings of the National Academy of Sciences*, 111(45):15888–15893.

DfBEIS (20016). 2016 uk provisional greenhouse gas emissions. Technical report, DfBEIS.

DfT (2002). Cost benefit analysis manual. volume 13: Economic assessment of road schemes. Technical report, Department for Transport.

DfT (2010). Traffic modelling guidelines tfl traffic manager and network performance best practice. Technical report, Department for Transport.

DfT (2012a). Guidance on road classification and the primary route network. Technical report, Department for Transport.

DfT (2012b). National travel survey: 2012. Technical report, Department for Transport.

DfT (2014a). Tag unit m2, variable demand modelling. Technical report, Department for Transport.

DfT (2014b). Tag unit m3.2 public transport assignment. Technical report, Department for Transport.

DfT (2014c). Transport analysis guidance. the appraisal process. Technical report, Department for Transport.

DfT (2015a). 2010 to 2015 government policy: Hs2, high speed rail. Technical report, UK Government.

DfT (2015b). 2010 to 2015 government policy: transport emissions. Technical report, Department for Transport.

DfT (2015c). Provision of market research for value of travel time savings and reliability. phase 2 report. Technical report, Department for Transport.

DfT (2017). Tag unit a1.3 user and provider impacts. Technical report, Department for Transport.

DfTt (2015). Provision of market research for value of travel time savings and reliability. Technical report, Department for Transport.

Dia, H. (2002). An agent-based approach to modelling driver route choice behaviour under the influence of real-time information. *Transportation Research Part C: Emerging Technologies*, 10(5–6):331 – 349.

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische mathematik*, 1(1):269–271.

DMRB (1999). Volume 5 assessment and preparation of road schemes. Technical report, The Highways Agency.

Dodman, D. (2009). Blaming cities for climate change? an analysis of urban greenhouse gas emissions inventories. *Environment and Urbanization*, 21(1):185–201.

Duranton, G. and Turner, M. A. (2011). The fundamental law of road congestion: Evidence from us cities. *American Economic Review*, 101(6):2616–52.

Edenhofer, O., Pichs-Madruga, R., Sokona, Y., Minx, C., Farahani, E., Kadner, S., Seyboth, K., Adler, A., Baum, I., Brunner, S., et al. (2014). Working group iii contribution to the fifth assessment report of the intergovernmental panel on climate change. *Intergovernmental Panel on Climate Change: Cambridge, UK*.

EU (2010). High-speed europe. a sustainable link between citizens. Technical report, European Union.

EU (2015). The eu emissions trading system (eu ets). Technical report, European Union.

EU (2016). The protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/ec (general data protection regulation). Technical report, European Union.

Eurostar (2006). London-paris and london-brussels flights generate ten times more carbon dioxide emissions than eurostar. Technical report, Eurostar.

Eurostar (2015). Eurostar growth continues in 2014. Technical report, Eurostar.

Eurostar (2016a). Core destinations timetable. Technical report, Eurostar.

Eurostar (2016b). Eurostar passenger numbers steady but profit down. Technical report, Eurostar.

Eurostar (2016c). Travelling with eurostar. Technical report, Eurostar.

Eurostar (2017). Eurostar reports a £58m profit. Technical report, Eurostar.

EuroTunnel (2016). Shuttle timetable. Technical report, EuroTunnel.

Flyvbjerg, B., Holm, M. S., and Buhl, S. (2002). Underestimating costs in public works projects: Error or lie? *Journal of the American planning association*, 68(3):279–295.

Flyvbjerg, B., Holm, M. S., and Buhl, S. (2005). How (in)accurate are demand forecasts in public works projects? *Journal of the American Planning Association*, 71(2):131–146.

GCB (2013). Infrastructure carbon review. Technical report, Green Construction Board.

GETLINK (2009). The channel tunnel infrastructure. Technical report, GETLINK.

Ghemawat, S., Gobioff, H., and Leung, S.-T. (2003). The google file system. In *ACM SIGOPS operating systems review*, pages 29–43. ACM.

GMTR (2014). Eurostar: Restructuring, expansion and rolling stock procurement. *Global Mass Transit Report*.

Google (2016a). Directions api documentation. Technical report, Google.

Google (2016b). Elevation api documentation. Technical report, Google.

Google (2016c). Gtfs overview. Technical report, Google.

Google (2016d). Street view. Technical report, Google.

Google (2017). Encoded polyline algorithm format. Technical report, Google.

Google (2018). Take control of your commute with google maps. Technical report, Google.

Hadoop, A. (2016). Apache hadoop. Technical report, Apache.

Halpern, D. (2015). *Inside the Nudge Unit: How small changes can make a big difference*. Random House.

Hardin, G. (1968). The tragedy of the commons. *Science*, 1.

Harrison, A. J. and Quarmby, D. (1970). The value of time in transport planning: a review. *Transportation Research Board*.

Hart, P. E., Nilsson, N. J., and Raphael, B. (1968). A formal basis for the heuristic determination of minimum cost paths. *IEEE transactions on Systems Science and Cybernetics*, 4(2):100–107.

Heijungs, R., Settanni, E., and Guinée, J. (2013). Toward a computational structure for life cycle sustainability analysis: unifying lca and lcc. *The International Journal of Life Cycle Assessment*, 18(9):1722–1733.

Heinonen, J., Jalas, M., Juntunen, J. K., Ala-Mantila, S., and Junnila, S. (2013). Situated lifestyles: How lifestyles change along with the level of urbanization and what the greenhouse gas implications area study of finland. *Environmental Research Letters*, 8(2):025003.

Herrera, J. C., Work, D. B., Herring, R., Ban, X. J., Jacobson, Q., and Bayen, A. M. (2010). Evaluation of traffic data obtained via gps-enabled mobile phones: The mobile century field experiment. *Transportation Research Part C: Emerging Technologies*, 18(4):568–583.

Hirschmann, K., Zallinger, M., Fellendorf, M., and Hausberger, S. (2010). A new method to calculate emissions with simulated traffic conditions. In *Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on*, pages 33–38. IEEE.

Hollander, Y. (2006). Direct versus indirect models for the effects of unreliability. *Transportation Research Part A: Policy and Practice*, 40(9):699–711.

Hollander, Y. (2016). Better use of modellers' time. In *Better use of modellers' time*. ICTthink.

Holtz-Eakin, D. and Selden, T. M. (1995). Stoking the fires? co 2 emissions and economic growth. *Journal of public economics*, 57(1):85–101.

HS2 (2009). Hs2. Technical report, Imperial College London.

HS2 (2012). Hs2 london to west midlands eia scope and methodology report. Technical report, HS2.

HS2 (2013). London-west midlands environmental statement volume 5 | technical appendices. Technical report, High Speed 2.

Hughes, L. (2012). *Effects of Alignment on CO2 Emissions from the Construction and Use Phases of Highway Infrastructure.* PhD thesis, University of Cambridge.

Hughes, L., Phear, A., Nicholson, D., Pantelidou, H., Soga, K., Guthrie, P., Kidd, A., and Fraser, N. (2011a). Carbon dioxide from earthworks: a bottom-up approach. *Proceedings of the Institution of Civil Engineers - Civil Engineering*, 164(2):66–72.

Hughes, L., Phear, A., Nicholson, D., Pantelidou, H., Soga, K., Guthrie, P., Kidd, A., and Fraser, N. (2011b). Carbon dioxide from earthworks: a bottom-up approach. In *Proceedings of the Institution of Civil Engineers-Civil Engineering*, pages 66–72. Thomas Telford Ltd.

Huijbregts, M. A., Norris, G., Bretz, R., Ciroth, A., Maurice, B., von Bahr, B., Weidema, B., and de Beaufort, A. S. (2001). Framework for modelling data uncertainty in life cycle inventories. *The International Journal of Life Cycle Assessment*, 6(3):127–132.

ICAO (2017). Carbon emissions calculator. Technical report, ICAO.

Iddon, C. R. and Firth, S. K. (2013). Embodied and operational energy for new-build housing: a case study of construction methods in the uk. *Energy and Buildings*, 67:479–488.

Iqbal, M. U. and Lim, S. (2010). Privacy implications of automated gps tracking and profiling. *IEEE Technology and Society Magazine*, 29(2):39–46.

Jain, J. and Lyons, G. (2008). The gift of travel time. *Journal of Transport Geography*, 16(2):81 – 89.

Jariyasunant, J., Kerkez, B., Sengupta, R., Glaser, S., and Bayen, A. (2011). Mobile transit trip planning with real-time data. *Transportation Research Board*.

Jones, C. and Hammond, G. (2008). Inventory of carbon and energy. Technical report, ICE.

Jones, P., Eyers, T., Bray, J., Georgeson, N., Powell, T., Paris, J., and Lane, R. (2002). The jubilee line extension impact study: Main findings and lessons learnt. Technical report, UCL.

Journal, I. R. (2014). Db schenker expands hs1 freight operations. Technical report, DB Schenker.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the econometric society*, pages 263–291.

Kia (2019). Kia-niro specification.

Kimball, M., Chester, M., Gino, C., and Reyna, J. (2013). Assessing the potential for reducing life-cycle environmental impacts through transit-oriented development infill along existing light rail in phoenix. *Journal of Planning Education and Research*, 33(4):395–410.

Knight, R. L. and Trygg, L. L. (1977). Evidence of land use impacts of rapid transit systems. *Transportation*, 6(3):231–247.

Krugman, P. (2009). The increasing returns revolution in trade and geography. *The American Economic Review*, 99(3):561–571.

Kussner, B. H. and Tsun, T. (2011). Development issues: Rail corridor setbacks and cn guidelines. Technical report, WeirdFoulds.

Lane, R., Powell, T., Eyers, T., Paris, J., Lucas, K., and Jones, P. (2004). Jle summary report. Technical report, Transport Studies Group.

LCR (2009). Economic impact of high speed 1 final report. Technical report, London & Continental Railways.

Lin, Y. (2015). The capital carbon cost of high speed one. Technical report, University of Cambridge.

Luke, S., Cioffi-Revilla, C., Panait, L., and Sullivan, K. (2004). Mason: A new multi-agent simulation toolkit. In *Proceedings of the 2004 swarmfest workshop*, volume 8, page 44.

Lumsdaine, A., Gregor, D., Hendrickson, B., and Berry, J. (2007). Challenges in parallel graph processing. *Parallel Processing Letters*, 17(01):5–20.

Mackie, P., Wardman, M., Fowkes, A., Whelan, G., Nellthorp, J., and Bates, J. (2003). Values of travel time savings uk. *Institute of transport studies*.

Malewicz, G., Austern, M. H., Bik, A. J., Dehnert, J. C., Horn, I., Leiser, N., and Czajkowski, G. (2010). Pregel: a system for large-scale graph processing. In *Proceedings of the 2010 ACM SIGMOD International Conference on Management of data*, pages 135–146. ACM.

Manley, E. (2013). *Modelling Driver Behaviour to Predict Urban Road Traffic Dynamics*. PhD thesis, University College London.

Maps, G. (2017). Google maps. Technical report, Google.

Mashros, N., Ben-Edigbe, J., Hassan, S. A., Hassan, N. A., and Yunus, N. Z. M. (2014). Impact of rainfall condition on traffic flow and speed: a case study in johor and terengganu. *Jurnal Teknologi*, 70(4):65–69.

McFadden, D. (1974). The measurement of urban travel demand. *Journal of public economics*, 3(4):303–328.

McNally, M. G. (2007). The four step model. *Handbook of transport modelling*, 1:35–41.

Mieszkowski, P. and Mills, E. S. (1993). The causes of metropolitan suburbanization. *The Journal of Economic Perspectives*, 7(3):135–147.

Miles, J., Batty, M., and Casey, G. (2019). Gerard casey viva - 20th february 2019.

Minar, N., Burkhart, R., Langton, C., Askenazi, M., et al. (1996). The swarm simulation system: A toolkit for building multi-agent simulations. *Swarm Development Group*.

Morita, Y., Shimizu, K., Yamasaki, T., Kato, H., and Shibahara, N. (2012). A study on the life-cycle assessment for evaluating carbon impact from construction of rail infrastructure. In *European Transport Conference 2012*.

Mtoi, E. T. and Moses, R. (2014). Calibration and evaluation of link congestion functions: applying intrinsic sensitivity of link speed as a practical consideration to heterogeneous facility types within urban network. *Journal of Transportation Technologies*, 2014.

MTR (2013). Sustainability report 2013. Technical report, Hong Kong MTR.

Mullick, A. and Ray, A. K. (2012). Dynamics of bimodality in vehicular traffic flows. *arXiv preprint arXiv:1205.2314*.

Nagel, K. and Flötteröd, G. (2012). Agent-based traffic assignment: Going from trips to behavioural travelers. In *Travel Behaviour Research in an Evolving World–Selected papers from the 12th international conference on travel behaviour research*, pages 261–294. International Association for Travel Behaviour Research.

Nahlik, M. J. and Chester, M. V. (2014). Transit-oriented smart growth can reduce life-cycle environmental impacts and household costs in los angeles. *Transport Policy*, 35:21–30.

NAO (2012). The completion and sale of high speed 1. Technical report, National Audit Office.

Neuhold, R. and Fellendorf, M. (2014). Volume delay functions based on stochastic capacity. *Transportation Research Record: Journal of the Transportation Research Board*, 1(2421):93–102.

Newman, P. and Kenworthy, J. (1999). *Sustainability and cities: overcoming automobile dependence*. Island Press.

Newman, P. W. and Kenworthy, J. R. (1989). Gasoline consumption and cities. *Journal of the american planning association*, 55(1):24–37.

Norman, J., MacLean, H. L., and Kennedy, C. A. (2006). Comparing high and low residential density: life-cycle analysis of energy use and greenhouse gas emissions. *Journal of urban planning and development*, 132(1):10–21.

O'Brien, O. (2016). Mapped-based data visualisations, often of london and/or transport/transit. Technical report, University College London.

Odersky, M., Spoon, L., and Venners, B. (2008). Programming in scala.

ODI (2016). Open data's impact transport for london get set, go! Technical report, ODI.

Office, M. (2003). Met office rain radar data from the nimrod system. ncas british atmospheric data centre. Technical report, Met Office.

Ogilvie, D., Egan, M., Hamilton, V., and Petticrew, M. (2004). Promoting walking and cycling as an alternative to using cars: systematic review. *Bmj*, 329(7469):763.

OS (2016). Integrated transport network. Technical report, Ordnance Survey.

Page, L., Brin, S., Motwani, R., and Winograd, T. (1999). The pagerank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab.

Pandas (2016). Python data analysis library. Technical report, Pandas.

Parker, M. T. et al. (2001). What is ascape and why should you care. *Journal of Artificial Societies and Social Simulation*, 4(1):5.

Potter, S. (2003). Transport energy and emissions: urban public transport. In *Handbook of Transport and the Environment*, pages 247–262. Emerald Group Publishing Limited.

RAC (2012). Keeping the nation moving cars and the environment. Technical report, RAC Foundation.

RAC (2016). Typical vehicle running costs. Technical report, RAC.

RailwayGazette (2018). Eurostar reports a £58m profit. Technical report, Eurostar.

Ramming, M. S. (2001). *Network knowledge and route choice*. PhD thesis, Massachusetts Institute of Technology.

Ratti, C., Frenchman, D., Pulselli, R. M., and Williams, S. (2006). Mobile landscapes: using location data from cell phones for urban analysis. *Environment and Planning B: Planning and Design*, 33(5):727–748.

Rayle, L., Shaheen, S., Chan, N., Dai, D., and Cervero, R. (2014). App-based, on-demand ride services: Comparing taxi and ridesourcing trips and user characteristics in san francisco university of california transportation center (uctc). Technical report, UCTC-FR-2014-08.

Rose, G., Taylor, M. A., and Tisato, P. (1989). Estimating travel time functions for urban roads: options and issues. *Transportation Planning and Technology*, 14(1):63–82.

Rosenzweig, C., Iglesias, A., Yang, X., Epstein, P. R., and Chivian, E. (2001). Climate change and extreme weather events; implications for food production, plant diseases, and pests. *Global change & human health*, 2(2):90–104.

Saaty, T. L. (1990). *Decision making for leaders: the analytic hierarchy process for decisions in a complex world*. RWS publications.

Salnikov, V., Lambiotte, R., Noulas, A., and Mascolo, C. (2015). Openstreetcab: Exploiting taxi mobility patterns in new york city to reduce commuter costs. *arXiv preprint arXiv:1503.03021*.

Samsel, C., Beul-Leusmann, S., Wiederhold, M., Krempels, K.-H., Ziefle, M., and Jakobs, E.-M. (2014). Cascading information for public transport assistance. In *WEBIST (1)*, pages 411–422.

Sanders, P. and Schultes, D. (2006a). Engineering highway hierarchies. In *European Symposium on Algorithms*, pages 804–816. Springer.

Sanders, P. and Schultes, D. (2006b). Robust, almost constant time shortest-path queries in road networks. In *The Shortest Path Problem*, pages 193–218.

Satterthwaite, D. (2008). Cities' contribution to global warming: notes on the allocation of greenhouse gas emissions. *Environment and urbanization*, 20(2):539–549.

Saxe, S. (2016). *THE NET GREENHOUSE GAS IMPACT OF METRO URBAN RAIL: THE JUBILEE LINE EXTENSION IN LONDON, UK AND THE SHEPPARD SUBWAY LINE IN TORONTO, CANADA*. PhD thesis, University of Cambridge.

Saxe, S., Casey, G., Guthrie, P., Soga, K., and Cruickshank, H. (2015a). Greenhouse gas considerations in rail infrastructure in the uk. *ICE: Sustainability*, 1.

Saxe, S., Cruickshank, H., and Miller, E. (2015b). Greenhouse gas impact of ridership on sheppard subway line, toronto, canada. *Transportation Research Record: Journal of the Transportation Research Board*, 0(2502):62–70.

Scott, A. (2002). *Global city-regions: trends, theory, policy*. Oxford University Press.

Senbel, M., Church, S., Bett, E., Maghsoudi, R., and Zhang, K. J. (2010). The relationship between urban form & ghg emissions a primer for municipal decision makers. Technical report, Urban Design Lab School of Community and Regional Planning University of British Columbia Vancouver, Canada.

Shah, A. (2015). Pregel.

Sharrard, A. L., Matthews, H. S., and Roth, M. (2007). Environmental implications of construction site energy use and electricity generation 1. *Journal of Construction Engineering and Management*, 133(11):846–854.

Shekhar, S., Gunturi, V., Evans, M. R., and Yang, K. (2012). Spatial big-data challenges intersecting mobility and cloud computing. In *Proceedings of the Eleventh ACM International Workshop on Data Engineering for Wireless and Mobile Access*, pages 1–6. ACM.

Shvachko, K., Kuang, H., Radia, S., and Chansler, R. (2010). The hadoop distributed file system. In *2010 IEEE 26th symposium on mass storage systems and technologies (MSST)*, pages 1–10. IEEE.

Silva, E. A. (2010). Waves of complexity: Theory, models and practice. *A Planner's Encounter with Complexity*, pages 309–331.

Silva, E. A. (2011). Cellular automata and agent base models for urban studies: From pixels to cells to hexa-dpi's. *Urban Remote Sensing: Monitoring, Synthesis and Modeling in the Urban Environment*, pages 323–334.

Silva, E. A., Healey, P., Harris, N., and Van den Broeck, P. (2014). *The Routledge handbook of planning research methods*. Routledge.

Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, pages 99–118.

Simon, H. A. (1979). Rational decision making in business organizations. *The American economic review*, 69(4):493–513.

Sims, R., Schaeffer, R., Creutzig, F., Cruz-Núñez, X., D'Agosto, M., Dimitriu, D., Meza, M. F., Fulton, L., Kobayashi, S., Lah, O., McKinnon, A., Newman, P., Ouyang, M., Schauer, J., Sperling, D., and Tiwari, G. (2014). 014: Transport. in: Climate change 2014: Mitigation of climate change. contribution of working group iii to the fifth assessment report of the intergovernmental panel on climate change. *IPCC*.

Skabardonis, A. and Dowling, R. (1997). Improved speed-flow relationships for planning applications. *Transportation Research Record: Journal of the Transportation Research Board*, 1(1572):18–23.

Smith, L., Beckman, R., and Baggerly, K. (1995). Transims: Transportation analysis and simulation system. Technical report, Los Alamos National Lab., NM (United States).

SNCF (2011). 1er bilan carbone ferroviaire global. Technical report, Société nationale des chemins de fer français.

Southeastern (2016). Hs1 timetable. Technical report, Southeastern.

Spiess, H. (1990). Technical note—conical volume-delay functions. *Transportation Science*, 24(2):153–158.

Spirito, M. A. (2001). On the accuracy of cellular mobile station location estimation. *IEEE Transactions on vehicular technology*, 50(3):674–685.

Standard, E. (2016). Piccadilly line delays: Tube chaos to continue for days as half of trains taken out of service. Technical report, Evening Standard.

Stern, N. H. (2007). *The economics of climate change: the Stern review*. cambridge University press.

TfL (2009). How london buses are numbered - tfl come up trumps! Technical report, TfL.

TfL (2010). Traffic modelling guidelines. version 3.0. Technical report, Transport for London.

TfL (2012). Roads task force – technical note 14. who travels by car in london and for what purpose? Technical report, Transport for London.

TfL (2015). Transport emissions roadmap. Technical report, Transport for London.

TfL (2016a). Fares and payments. Technical report, Transport for London.

TfL (2016b). Value of time for bus passengers. Technical report, TfL.

TfL (2017). Facts and figures. Technical report, Transport for London.

Thormark, C. (2002). A low energy building in a life cycle—its embodied energy, energy need for operation and recycling potential. *Building and environment*, 37(4):429–435.

Tirachini, A., Hensher, D. A., and Rose, J. M. (2013). Crowding in public transport systems: Effects on users, operation and implications for the estimation of demand. *Transportation Research Part A: Policy and Practice*, 53:36 – 52.

Tisato, P. (1991). Suggestions for an improved davidson travel time function. *Australian road research*, 21(2).

Toole, J. L., Colak, S., Sturt, B., Alexander, L. P., Evsukoff, A., and González, M. C. (2015). The path most traveled: Travel demand estimation using big data resources. *Transportation Research Part C: Emerging Technologies*, 58:162–177.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

Treloar, G. J., Love, P. E., and Crawford, R. H. (2004). Hybrid life-cycle inventory for road construction and use. *Journal of Construction Engineering and Management*, 130(1):43–49.

Trust, T. C. (2008). Energy and carbon conversions. Technical report, The Carbon Trust.

TSC (2016). Modernising the national travel survey. wp1 - review of current and future technologies and data sources report. Technical report, Transport Systems Catapult.

Tversky, A. and Kahneman, D. (1985). The framing of decisions and the psychology of choice. In *Environmental Impact Assessment, Technology Assessment, and Risk Analysis*, pages 107–129. Springer.

UIC (2011). Carbon footprint of high speed rail. Technical report, International Union of Railways.

UNFCCC (2018). Emissions from fuel used for international aviation and maritime transport. Technical report, UNITED NATIONS FRAMEWORK CONVENTION ON CLIMATE CHANGE.

van Rossum, G. and Drake, F. (2002). Python reference manual. Technical report, PythonLabs.

Vap, D. and Sun, C. (2007). Investigating large truck-passenger vehicle interactions. Technical report, Organizational Results Research Report October 2007.

Vuk, G. (2005). Transport impacts of the copenhagen metro. *Journal of Transport Geography*, 13(3):223–233.

VW (2019). Vw up! specification.

Wadhwa, V. (2014). Laws and ethics can't keep pace with technology. Technical report, Massachusetts Institute of Technology.

Walton, H., Dajnak, D., Beevers, S., Williams, M., Watkiss, P., and Hunt, A. (2015a). Understanding the health impacts of air pollution in london. *London: Kings College London, Transport for London and the Greater London Authority*.

Walton, H., Dajnak, D., Beevers, S., Williams, M., Watkiss, P., and Hunt, A. (2015b). Understanding the health impacts of air pollution in london. *King's College London, UK*, 129.

Wang, F. and Xu, Y. (2011). Estimating o–d travel time matrix by google maps api: implementation, advantages, and implications. *Annals of GIS*, 17(4):199–209.

Wang, H., Calabrese, F., Di Lorenzo, G., and Ratti, C. (2010). Transportation mode inference from anonymized and aggregated mobile phone call detail records. In *Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on*, pages 318–323. IEEE.

Wardman, M. (2004). Public transport values of time. *Transport policy*, 11(4):363–377.

Wardman, M. (2008). Is generalised cost justified? In *EUROPEAN TRANSPORT CONFERENCE 2008; PROCEEDINGS*.

Waris, M., Liew, M. S., Khamidi, M. F., and Idrus, A. (2014). Criteria for the selection of sustainable onsite construction equipment. *International Journal of Sustainable Built Environment*, 3(1):96–110.

Watkiss, P. (2009). Update of eurostar co2 emissions using energy logging train data. Technical report, Paul Watkiss Associates.

Wilensky, U. (1999). NetLogo.

Willumsen, L. G. (1978). Estimation of an od matrix from traffic counts–a review. *University of Leeds*.

Wing, M. G., Eklund, A., and Kellogg, L. D. (2005). Consumer-grade global positioning system (gps) accuracy and reliability. *Journal of forestry*, 103(4):169–173.

Xin, R. S., Gonzalez, J. E., Franklin, M. J., and Stoica, I. (2013). Graphx: A resilient distributed graph system on spark. In *First International Workshop on Graph Data Management Experiences and Systems*, page 2. ACM.

Yeheyis, M., Hewage, K., Alam, M. S., Eskicioglu, C., and Sadiq, R. (2013). An overview of construction and demolition waste management in canada: a lifecycle analysis approach to sustainability. *Clean Technologies and Environmental Policy*, 15(1):81–91.

Zheng, N., Waraich, R. A., Axhausen, K. W., and Geroliminis, N. (2012). A dynamic cordon pricing scheme combining the macroscopic fundamental diagram and an agent-based traffic model. *Transportation Research Part A: Policy and Practice*, 46(8):1291 – 1303.

Zheng, Y., Chen, Y., Li, Q., Xie, X., and Ma, W.-Y. (2010). Understanding transportation modes based on gps data for web applications. *ACM Transactions on the Web (TWEB)*, 4(1):1.

Appendix A

Geospatial conflation - polyline to underlying graph

A significant technical challenge is presented when attempting to unify two similar but different geospatial references. Google's API responses return geospatial coordinates specific to their underlying representation of London. This is most obviously different from the ITN representation in that it is using a different coordinate projection system[1]. Beyond the relatively simple task of re-projection, there is the challenge of unifying different geospatial representations of the same infrastructure. Consider Figure A.1 which presents a Google polyline overlaying the ITN road graph. The dark blue polyline illustrates the Google Directions API result and the grey illustrates the underlying ITN network. As is visually clear, there are small discrepancies between Google and ITN's underlying graphs. It is necessary to reconcile this two differing representations of the same physical infrastructure in order to transfer attributes from the Google data (journey times) to the ITN graph. The reconciliation process for differing geospatial data is known as conflation. It is apparent in this case that the light blue ITN polyline is the ITN equivalent of the Google polyline result shown. However, a programmatic method for this process is required due to the size of the Google and ITN datasets.

A.1 Google Directions API data

An individual API response consists of a series of individual legs with associated journey times. This is presented in Figure A.2. The one to many relationship of API route to legs is shown in Figure A.3.

[1] Google use epsg:4326 and the ITN uses epsg:27700

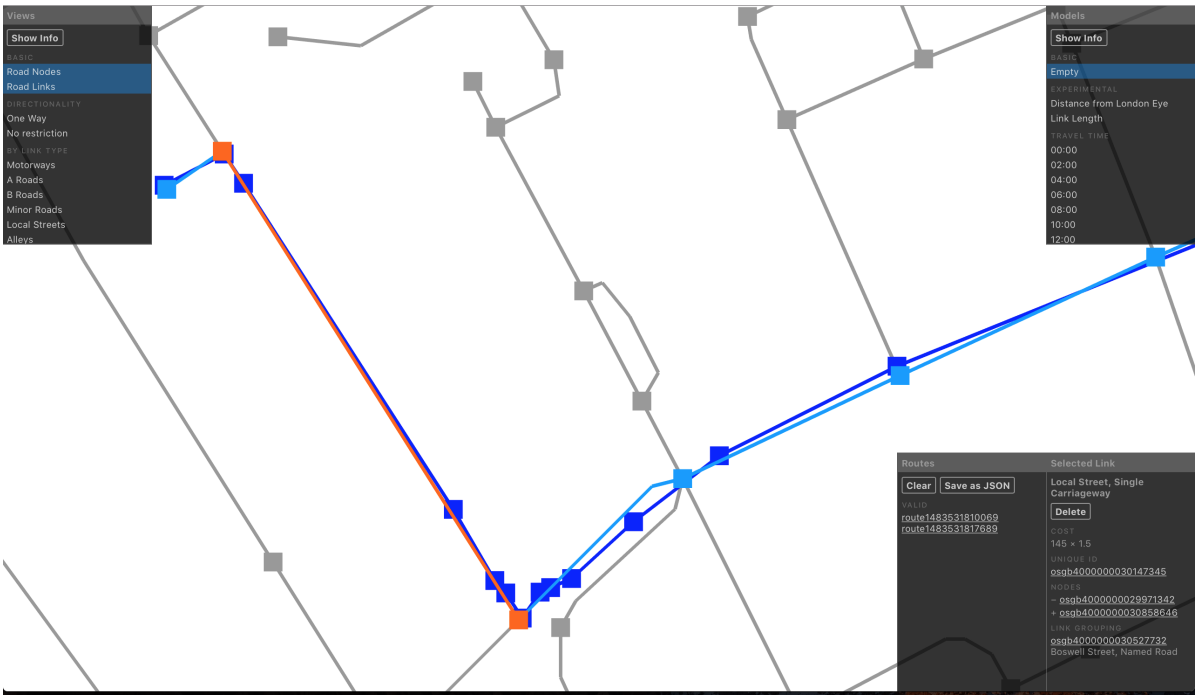


Fig. A.1 Geospatial differences: Google versus ITN

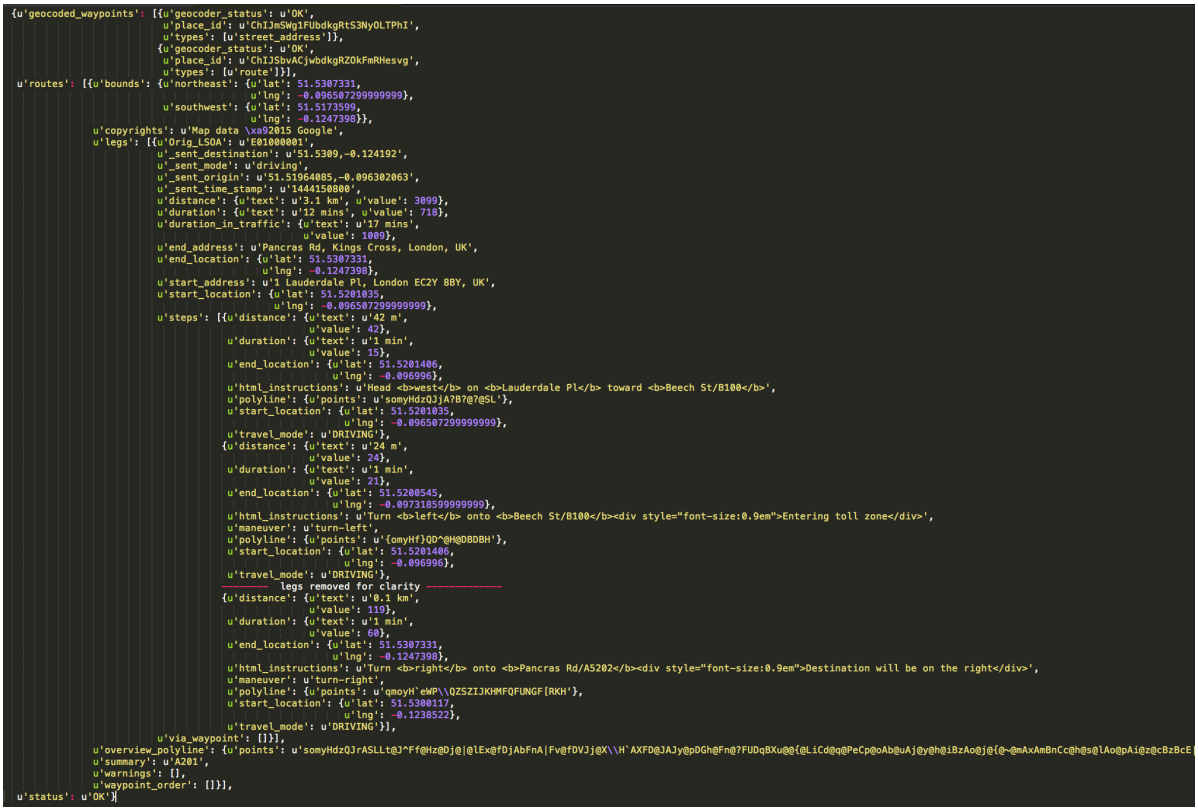


Fig. A.2 Example Google Directions API response

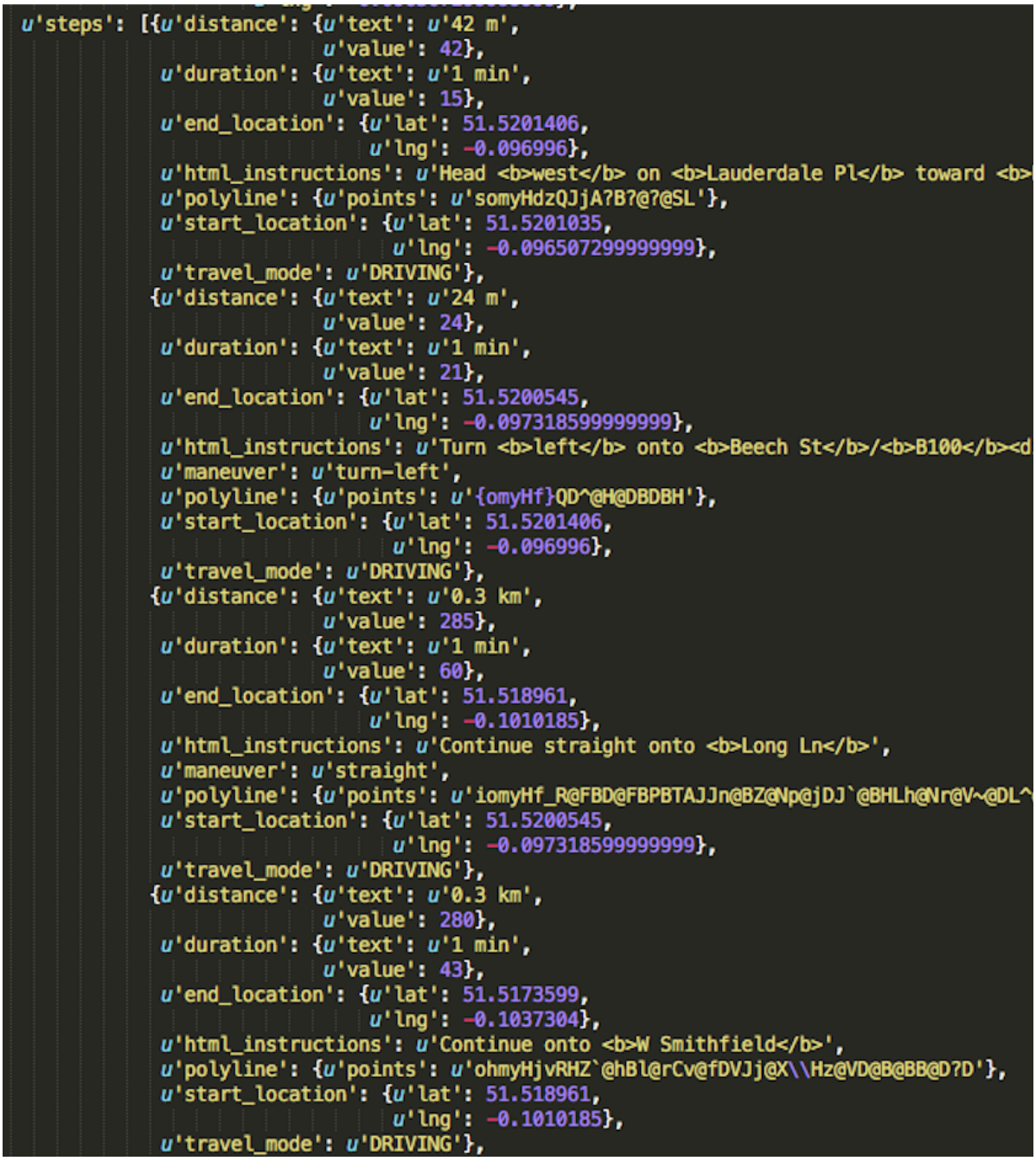


Fig. A.3 API response and legs

The journey time and encoded polyline are harvested from each response. The raw polyline, consisting of a series of coordinates is compressed using a lossy compression algorithm producing a single string representation (Google, 2017). To decrease space even more, Google offset from the previous point, rather than constructing from scratch. This must be considered when individually assessing an individual polyline from a given Google result.

1. Inherit previous point if not first polyline in route
2. Decode polyline
3. Re-project points within polyline
4. Output polyline and associated journey time

A.2 Data conflation

The ITN consists of 772983 edges and 343724 vertices. The dataset generated in Chapter 3 consisted of 64,528,506 individual leg records. It is therefore necessary to consider computational efficiency in the conflation process. Conflation is a feature provided by GIS software platforms such as QGIS. However, as was identified previously, these platforms are designed for large centralised computations on centralised data structures. Thus, due to the size of the data here and the non-standard JSON format employed, an implementation in Python was devised. This a batch kd-tree Python implementation inspired by the quadtree implementation devised in the Sierra-Charlie visualiser (Bak et al., 2016). A kdtree is a hierarchical data structure which is useful for focussing computational effort on the interesting part of the problem (Bentley, 1975). The kd tree can be queried to find those vertices within the vicinity of a returned Google polyline. From these available vertices, a shortest path computation can be carried out, with vertices weighted dependent on vicinity,

Conflation process:

1. Check response status of API response (discard if there are errors)
2. Decode, re-project and reformat Google polyline
3. Decompose Google polyline into constituent, atomic polyline
4. For each atomic polyline:
5. Create list of vertices near all of the given vertices with a polyline (10m)

6. Find start and end vertices of Google polyline, match to nearest ITN vertices
7. Add additional (midpoint) points to Google polyline where there is a section longer than 100m
8. Find any ITN vertices within a 50m buffer around each point of the Google polyline
9. Find any ITN vertices within a 200m buffer around the origin and destination points
10. Build localised graph for these possible vertices
11. Find shortest path between these ITN vertices
12. Attribute journey time from Google polyline to ITN polylines proportionally to length

Stage 3 requires the use of the recursive data structure to ensure that the results are computed in an efficient matter.

Since each polyline may be computed in isolation, this is a straightforward process to parallelise.

A.2.1 Output

A.2.2 Success & confidence

Figure A.5 presents the percentage error between matched edge(s) and the API response, acting as a proxy for the error. As is shown, the vast majority of responses have a single percentage point discrepency.

A.3 Future work

Presently, this method matches a polyline to the nearest whole polyline. In the event of a Directions API response commencing or finishing mid ITN polyline, the data is generalised across the entire edge. A method for attributing costs proportionally across appropriate edges is an obvious next step.

```
"status": true,
"path_length": 223,
"buffer": 75,
"delta": 3.5674289562590977,
"gen_links": [
 {
  "toid": "osgb4000000030255960-",
  "length": 114.06753746642732,
  "jt_1446457887": 20.48980850221584
 },
 {
  "toid": "osgb4000000030427812-",
  "length": 24.927883414383064,
  "jt_1446457887": 4.477764391789398
 },
 {
  "toid": "osgb4000000030255963-",
  "length": 69.86336543163279,
  "jt_1446457887": 12.549468593865287
 },
 {
  "toid": "osgb4000000030360149-",
  "length": 51.23838286167306,
  "jt_1446457887": 9.203886365197443
 },
 {
  "toid": "osgb4000000030155927-",
  "length": 148.23559770817525,
  "jt_1446457887": 26.627374253134622
 },
 {
  "toid": "osgb4000000030427809-",
  "length": 24.969066253735946,
  "jt_1446457887": 4.485162013502468
 },
```


Fig. A.4 Sample output of polyline geospatial conflation

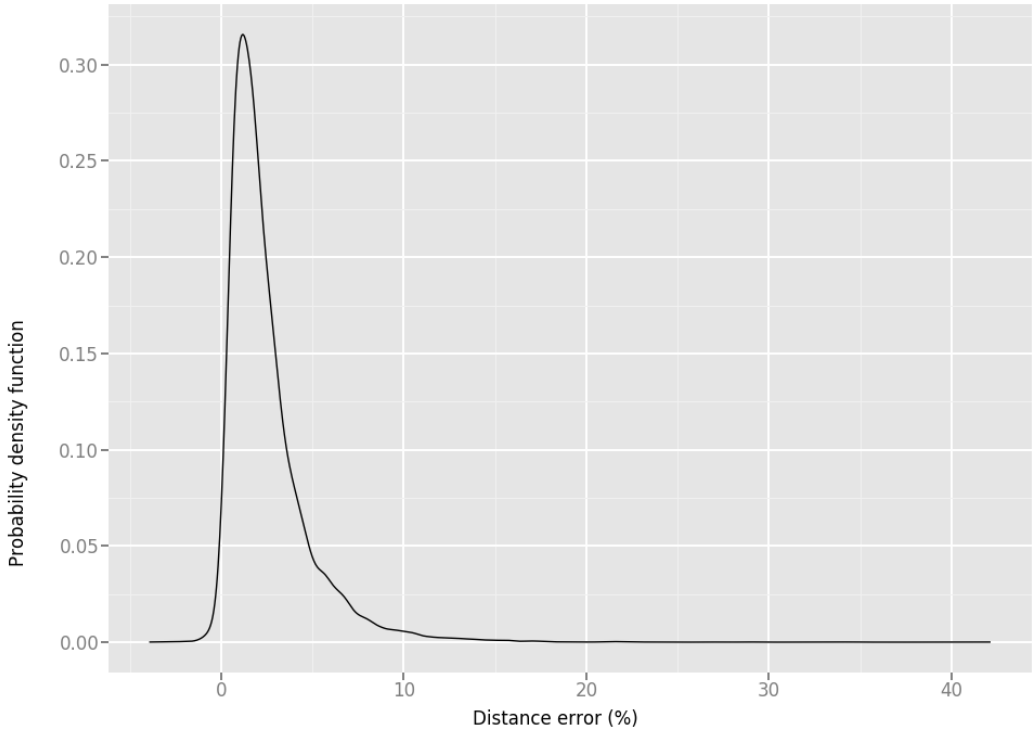


Fig. A.5 Matched polyline versus input polyline length error probability density function