

Shortest path or anchor-based route choice: a large-scale empirical analysis of minicab routing in London



E.J. Manley^{a,b,*}, J.D. Addison^b, T. Cheng^b

^a Centre for Advanced Spatial Analysis, University College London, Gower Street, London WC1E 6BT, United Kingdom

^b Department of Civil, Environment and Geomatic Engineering, University College London, Gower Street, London WC1E 6BT, United Kingdom

ARTICLE INFO

Keywords:

Route choice
Spatial cognition
Urban modelling
Urban transportation
Data analysis

ABSTRACT

Understanding and modelling route choice behaviour is central to predicting the formation and propagation of urban road congestion. Yet within conventional literature disagreements persist around the nature of route choice behaviour, and how it should be modelled. In this paper, both the shortest path and anchor-based perspectives on route choice behaviour are explored through an empirical analysis of nearly 700,000 minicab routes across London, United Kingdom. In the first set of analyses, the degree of similarity between observed routes and possible shortest paths is established. Shortest paths demonstrate poor performance in predicting both observed route choice and characteristics. The second stage of analysis explores the influence of specific urban features, named anchors, in route choice. These analyses show that certain features attract more route choices than would be expected were individuals choosing route based on cost minimisation alone. Instead, the results indicate that major urban features form the basis of route choice planning – being selected disproportionately more often, and causing asymmetry in route choice volumes by direction of travel. At a finer scale, decisions made at minor road features are furthermore demonstrated to influence routing patterns. The results indicate a need to revisit the basis of how routes are modelled, shifting from the shortest path perspective to a mechanism structured around urban features. In concluding, the main trends are synthesised within an initial framework for route choice modelling, and presents potential extensions of this research.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

It is well established that the route choice decisions taken by individuals contribute directly to the formation of collective patterns of behaviour observed in the city. Capturing the nature of this choice process is, as such, important for the accuracy of many urban modelling applications. Yet despite this clear purpose, uncertainty persists within the literature around how route choice is understood and how it should be modelled.

Conventional literature on route choice may be divided into two broadly distinct fields. One set of approaches comes from transportation studies, with a general perspective of capturing the relationship between drivers and roadway engineering. The other has arisen within spatial cognition and behavioural geography research, establishing the behavioural and cognitive constructs behind route choice mainly within the pedestrian context. Despite

similar objectives, the methods applied and dominant theories across each discipline differ considerably.

In transportation studies research into route choice, greatest attention has been applied towards establishing models reflective of how traffic distributes around the road network. Within this field, one particularly important approach has been that of traffic assignment modelling. Developed in the 1950s (Wardrop, 1952), later incarnations of this approach (Daganzo and Sheffi, 1977; Merchant and Nemhauser, 1978) continue to dominate many of the traffic models most widely used today. This approach considers traffic from a macroscopic perspective, iteratively adding traffic flow onto the road network so that travel times increase due to congestion. Intrinsic within this methodology is the assumption that individuals *optimise* their travel time between origin and destination, irrespective of varying perception, preference or awareness. This simple assumption, although intuitive in some respects, has been criticised for not fully reflecting the complexity of route choice behaviour (Garling et al., 1998; Golledge and Garling, 2001).

More sophisticated approaches to route choice modelling have emerged following a more individual-centric approach. An

* Corresponding author at: Centre for Advanced Spatial Analysis, University College London, Gower Street, London WC1E 6BT, United Kingdom. Tel.: +44 (0) 20 3108 3176.

E-mail addresses: ed.manley@ucl.ac.uk (E.J. Manley), puff.addison@ucl.ac.uk (J.D. Addison), tao.cheng@ucl.ac.uk (T. Cheng).

important stream of this research incorporates the discrete choice methodology. Discrete choice modelling involves establishing the relative influence of a combination of attributes reflective of observed choices, captured in recorded route data. The attributes included within these models vary widely, from relatively conventional aspects concerning travel time, distance, scenery and congestion (Ben-Akiva et al., 1984) to others incorporating socio-economic characteristics (Ramming, 2002), route perception (Cascetta et al., 2002), traffic information (Mahmassani and Liu, 1999), and uncertainty around congestion (de Palma et al., 2008), among others.

More recently, there has been a move to better integrate mechanisms of human cognition, psychology and behaviour within route choice models. The introduction of Prospect Theory (Kahneman and Tversky, 1979) into route choice modelling aims to better capture how people respond to travel time uncertainties (Avineri and Bovy, 2008; Gao et al., 2010). Prospect Theory has been extended through integration with reinforced learning approaches, demonstrating how decisions relating to travel time uncertainty change as individuals extend their experience of the environment or are provided with real-time information (Ben-Elia and Shifan, 2010). Others have explored how route knowledge influences choices and attitudes to risk, finding that experienced individuals are more sensitive to travel time variability (Ben-Elia et al., 2013). Similarly, Chorus and colleagues introduced the notion of regret minimisation, reflecting how individuals aim to minimise their exposure to negative emotions, rather than primarily aiming to maximise utility during route choice (Chorus et al., 2008). Another development has involved the incorporation of latent variables within choice models, aiming to capture the less tangible factors involved in route choice decision-making. Latent variables have been used for modelling the influence of habit (Kaplan and Prato, 2012), spatial ability (Prato et al., 2012), attitude to risk (Sun et al., 2012) and journey context (Feng et al., 2013). Finally, process-oriented and strategy-based approaches introduce how different approaches can be used in forming route decisions, based on recent experience, habit or memory (Senk, 2010). Like latent variables, these approaches aim to wrap route choice within a wider context, reflecting the dynamic nature of the choice. All of the advances outlined here demonstrate increasing appreciation for the importance of considering psychology and cognition within route choice modelling.

One important dimension of the discrete choice approach is that a strong understanding of all available options must be established. These options enable the identification of the attributes favoured by a decision-maker, relative to the attributes of rejected alternatives, as well as representing possibilities in the prediction of future choices. However, the derivation of this *choice set* in respect of route choice is problematic, where any combination of roads could feasibly be considered a rejected alternative. The modelling of choice sets has generally followed conventional transportation research, specifying choices through optimal or near-optimal routing between an origin and destination (Bekhor et al., 2006). However, important reviews have found these approaches to be 'unsatisfactory' (Bovy, 2009), being found to lack the behavioural criteria by which travellers choose routes. Despite more recent advances – such as the implementation of semi non-compensatory spatiotemporal constraints on choice set definition (Kaplan and Prato, 2012) – the definition of choice sets lags behind advances in the behavioural realism of the choice models. As such, simplistic assumptions persist around the bases on which route choices are represented.

Research findings emerging from spatial cognition and behavioural geography research streams deviate from these assumptions of intended optimality. Within this sector, strong emphasis has been placed on examining the relationship between human

cognition and urban space. Particularly important, it has been found that this relationship is highly subjective, based on experience, and skewed by particular features within urban space. Rather than individuals having complete knowledge of road network arrangement, various researchers identified that memory of cities is shaped around *anchors*, particularly salient features in urban space, around which subjective knowledge is built and recalled (Lynch, 1960; Siegel and White, 1975; Passini, 1984; Golledge and Spector, 1978; Golledge et al., 1985; Couclelis et al., 1987). These reference points have been shown to be different for drivers and pedestrians – the former favouring route-based structures such as bridges, major routes and junctions, the latter using prominent buildings and signposts (Carr and Schissler, 1969). These findings have anatomical foundations too, with the role of salient features being linked to brain activity within the hippocampus during the course of navigation (O'Keefe and Nadel, 1978).

The construction and utilisation of individual knowledge of space has furthermore been identified as hierarchically organised. Under this configuration, the memory of space is formed around a relative few core locations – be they anchors, regions (Kuipers, 1978; Hirtle and Jonides, 1985) or roads (Tomko et al., 2007; Péruch et al., 1989) – beneath which other locations are recalled. There is evidence too that this hierarchical configuration of space is utilised during the route planning process, with indications that routes are chosen on a coarse regional basis in the first instance, prior to an increasing specialisation on a road-by-road level during the execution of the plan (Wiener and Mallot, 2003; Wiener et al., 2009).

The view of many of the best-supported findings within spatial cognition literature indicates that individuals are unlikely and unable to select an optimal route between origin and destination. However, despite the strength of these findings, only a few models have been developed based on these trends (Gopal et al., 1989; Chown et al., 1995; Kuipers and Levitt, 1988), and have not been implemented within real-world contexts, nor linked to geographic nor transportation features. In addition, the findings of many of these studies have typically been established through sampling of only very few individuals, usually with a focus only on pedestrian behaviours. Importantly, few of these studies have been undertaken in real-world environments, instead often taking place within simplified virtual reality simulations.

The purpose of this paper is to more firmly establish the behavioural nature of route choice in urban areas, from the two dominant perspectives currently held within the literature. Using a large dataset of observed route choice behaviours, a range of analysis approaches will be applied to explore how widely *shortest path* and *anchor-based* route choice behaviours are followed across the large scale.

In undertaking these objectives, the paper will be structured as follows. The next section will describe the context and datasets involved in the study, highlighting the nature of the case and the limitations to be heeded during later discussions. Following this, a methodology section will outline the structure and design of the analysis, before the two sections of analysis are presented. The first analysis section examines how strongly observed behaviours align with a range of shortest paths (defined in different ways). The second stage of analysis explores whether the anchor-based routing can be observed within the dataset. The paper concludes in discussing the findings, and outlining potential future avenues for the outcomes of this research.

2. Context and datasets

This section outlines the context within which the study will take place, and describes the nature of the datasets to be used.

2.1. Regional context

The study is undertaken in London, United Kingdom. For the purposes of guidance throughout this paper, Fig. 1 presents a reference map for central London.

Like many European cities, the London road network developed through a largely organic process, unlike many of the planned grid-like developments of North America. Nevertheless, later developments have sought to improve traffic flow. The Inner Ring Road, which circumnavigates the central congestion charge zone, is a collection of routes widened and optimised for traffic flow. The North and South Circular roads, two routes further from central London, were introduced during the 1960s to improve traffic flow in suburban areas. The Westway Flyover, constructed during the same period, represents the legacy of the abandoned plans for a wider urban motorway system.

The centre of London, where the majority of commercial and leisure activity takes place, is split by the River Thames. There are eight bridges within the central area linking the north and south banks of London.

This paper will make use of the postal regionalisation of London, which divides the city into areas based on their bearing from the centre. Within each postal area (denoted N for north, E for east, etc.) are a hierarchy of subdivisions, called postal districts and postcodes. These latter regions will be used during the testing phases of this paper, the postal area definitions are shown in Fig. 1.

2.2. Minicab route data

The dataset to be utilised during this study tracks the routes chosen by 2970 minicab drivers in London, United Kingdom. The minicab company provided Global Positioning System (GPS) point logs for each of their vehicles across a 3-month period covering December 2010 to February 2011, equating to around 300 million records. Supporting datasets in the form of job (the origin and destination of each journey) and driver activity data were also provided.

Prior to starting the route analyses, an initial data extraction phase was undertaken, extracting road segment-based routes from the observed point-based dataset. Given the large number of records involved, a batch-processing algorithm was developed, capable of handling the poor spatial accuracy and recording rate offered by the GPS-derived data. The result of this processing stage was the extraction of 677,411 routes, referencing intended origin and destination locations as stated within the job records.

There are a number of important points that must be highlighted with respect to the drivers involved in the generation of this dataset, points which ultimately impact on the resulting analyses. Firstly, drivers for Addison Lee are not required to have prior driving experience or to have passed 'The Knowledge' test – a qualification for London's Black Cab drivers that requires a detailed knowledge of the layout and connectivity of large parts of London. Drivers do, however, undergo comprehensive training and are required to pass a map test prior to starting work. As such, the

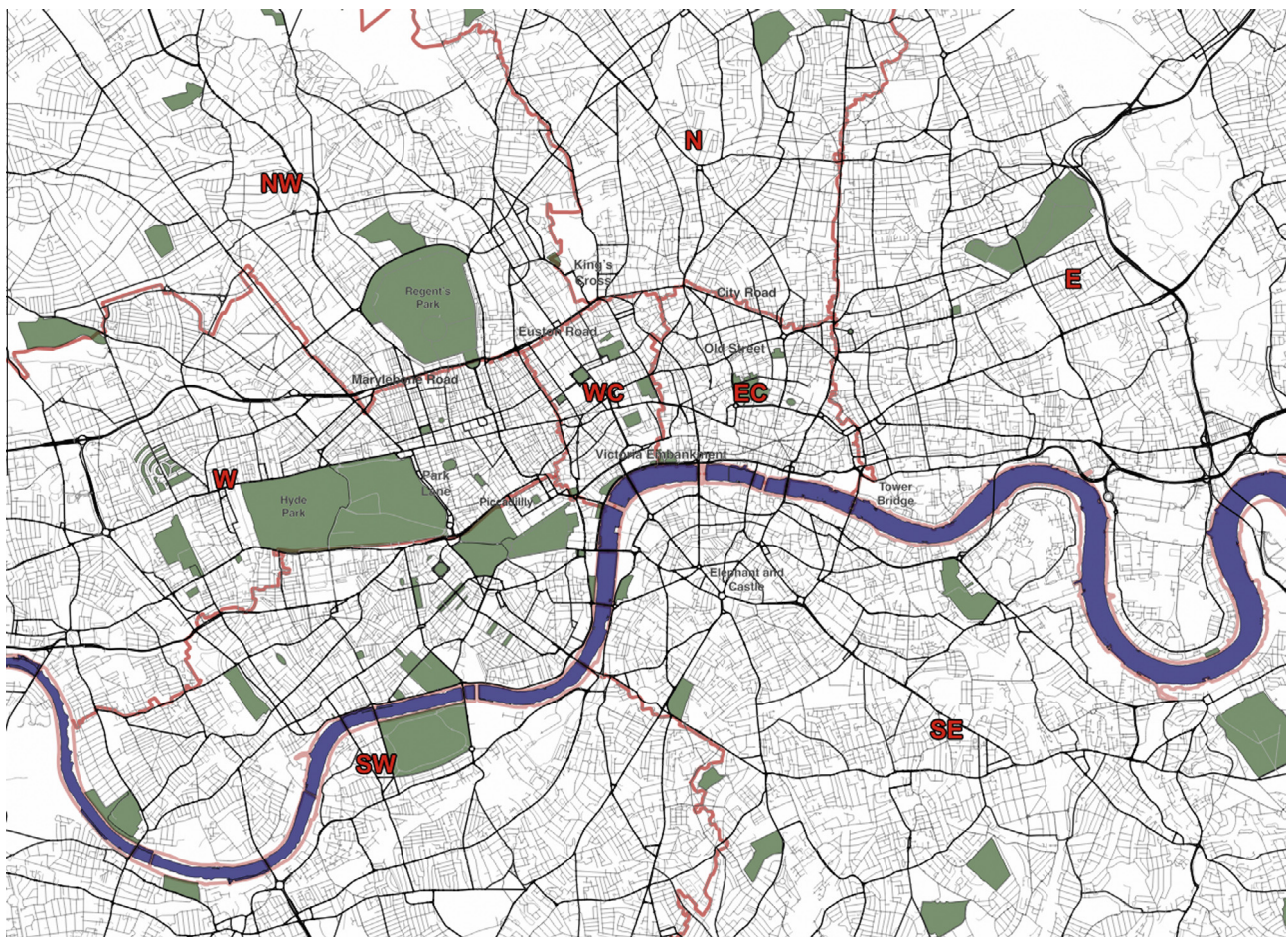


Fig. 1. Map showing the central London region including postal areas, parks, routes and landmarks.

drivers may be expected to hold a more detailed knowledge of the road network than regular drivers and commuters.

Second, the nature of the minicab business ensures a geographic concentration in the distribution of routes. In the case of Addison Lee and London, the majority – 71.4% – of journeys originate or end inside central London, with an additional 23.3% of journeys starting or ending within the inner London region. The nature of this bias means that resulting findings will be most applicable to condense urban centres, with irregular, organic road structures.

Third, the Addison Lee drivers are provided with a satellite navigation device within their vehicle, adding uncertainty around the method by which routes are planned. However, the devices installed in the Addison Lee vehicles lack real-time traffic information and were not integrated within the job allocation process at the time of the study (e.g. routes were not automatically provided to the drivers), potentially reducing their utility. Nevertheless, efforts should be undertaken to establish how widely these devices were used. Without any primary data regarding the extent of the use of satellite navigation, however, only proxy indicators can be used. This proxy indicator is provided by measuring the degree of similarity in route distribution between specific pairs of locations.

Routes are extracted for all journeys travelling between ten pairs of postcodes. The UK postcode region is representative of a single large property or small collections of buildings, thus providing high spatial precision around the origin and destination location of trips. The ten most frequently travelled pairs of postcodes were chosen for this analysis, yielding reasonably sized samples with variability in journey length and location. For each pair, every route is compared to every other route between origin and destination, and the number of matching road segments between each pair of routes extracted. The results from this analysis are shown in Table 1.

The results indicate low route similarities across most cases of repeated routes. This is indicative of the absence in the widespread use of a single navigation device. Considerable variation in similarity measures can be observed across cases. While increasing distance does appear to reduce similarity, it can be seen that the urban environment appears to have an impact too. Journeys in case 4 exhibit very high similarity across a low distance, yet in case 1, at a similar distance, but positioned within central London, there is considerable apparent variation in choice. The degree of variability and its association with urban form contributes to indications of the absence of device-generated routing.

2.3. Supporting datasets

Road data is provided in the form of the Ordnance Survey Integrated Transport Network (ITN) dataset. This GIS dataset details every road segment in the United Kingdom, on a

junction-to-junction basis. Integrated with the spatial representation are metadata sets detailing routing restrictions and speed limits. These factors are incorporated in the implementation of a network model, used during data processing and the generation of modelled routes.

3. Methodology

The intention of this paper is to examine the basis of route choice in urban areas. Where transportation research broadly assumes that route choice is made at once, behavioural geography stresses the importance of locations within the decision-making process. The methodology used here will aim to examine both approaches.

In undertaking an analysis of the one-shot approach from transportation research, the degree of similarity between comparable observed and modelled route choices was calculated. For each observed route, a range of modelled alternatives between the same origin and destination were generated. Modelled routes are generated through the calculation of a shortest path between origin and destination (using the *Dijkstra (1959)* shortest path algorithm), minimising a specified road traversal cost function.

The evaluation of the accuracy of the modelled route is carried out on a road segment basis. A measure of the proportion of the observed route that is matched by the modelled alternative is calculated. Road segments are selected as the basis for comparison rather than matched distance or travel time, as segments represent all potential points of variation, where deviations between routes occur. The modelled routes chosen for comparison will be described in more detail during this stage of the analysis in Section 3.

This stage of analysis will be undertaken mainly through an analysis of all routes, but further differentiation will be made. Route similarity will be broken down by length and by time of day to explore potential variation enforced by space or time.

The second stage of analysis explores the potential role of anchors in route choice, examining theories put forward in behavioural geography and neuroscience research. Unlike the shortest paths, no established routing methodology is available for comparison. In exploring the viability of this route choice process then, a number of spatial and statistical analytical methods are used, establishing whether particular road network features are chosen disproportionately more than valid alternatives. A number of representative case studies are used, allowing a targeted examination of the role of urban features in route choice.

Three main stages of analysis are undertaken here. In the first stage, spatial deviations between shortest path route models and observed route behaviours are established, identifying locations in the city where route models systematically under- or over-predict choice. The nature and strength of these deviations will pro-

Table 1
Mean proportionate matches between all routes between 10 origins and destinations.

Case	Origin	Destination	Trip count	Euclidean distance (miles)	Mean proportion match
1	W1T 3QN	WC2R 2PG	97	1.05	0.28
2	W1G 6BW	NW1 7TN	84	1.03	0.63
3	EC4A 4TR	W2 1HB	60	3.1	0.21
4	E14 4DA	E14 3QE	57	1.29	0.85
5	EC1A 1HQ	SW8 2NP	54	3.19	0.36
6	W1W 7JE	SW15 6DS	53	5.34	0.37
7	NW1 3ER	WC1X 9JX	48	0.99	0.62
8	NW1 7TL	NW8 0LH	46	1.17	0.63
9	SW13 9QF	SW3 1ER	42	3.95	0.58
10	NW1 7BY	W2 1HB	42	2.03	0.38

vide initial indications of the presence of anchor-based routing. In the second stage, statistical and spatial distributions in the use of major and minor features on the road network are established. This will identify the extent of variation in the use of different features during route selection. In the third phase, the process through which anchors are used is explored, identifying the extent of spatial deviations in direction of travel. This will provide further indications of the method and order in which anchors are chosen for traversal.

This third section utilises a range of methods in examining both spatial and statistical variation in route choice. Spatial variation is explored by mapping differences in routing behaviour between datasets on a road segment basis. The mapping of differences captures effectively the section of the road network favoured by one route set relative to an alternative distribution. Where one is interested in the regional spatial trends in deviation between route sets (e.g. beyond deviations along single road segments), spatial clustering methods are utilised. The Local Indicator of Spatial Autocorrelation (LISA) method is utilised in this case (Anselin, 1995). Unlike global measures of autocorrelation (such as Moran's I), LISA identifies local clusters of spatial features exhibiting similar characteristics. LISA is calculated through an assessment of each road segment, measuring similarity with neighbouring road segments. For the purposes of this study, neighbouring segments are considered as those within 1500 m of a tested road segment,¹ and are weighted in the calculation of local autocorrelation by the inverse of their Euclidean distance from that road segment. The analysis of the entire road network (to the extents of the London region) will yield spatial clusters of road segments with significantly higher or lower route selection relative to an alternative route set.

For this stage of the analysis, a number of case studies, pertaining to route sets between specific origins and destinations, are used. The case study approach is adopted to limit sources of behavioural variation, enabling a focus on routing activity associated with only a subset of the entire road network. This approach is made possible, unlike in other studies, by the large size of the dataset available. Case study regions were chosen on a basis of balancing two factors; sample size and distance between origin and destination. The former factor is intended to reduce the uncertainty in findings (important, given the nature of the study), and latter to ensure ample opportunity for selection or non-selection of intervening subgoals. In balancing these elements, larger regions than the postcode areas used earlier during testing were adopted.

The case studies to be used, and accompanying rationale, are detailed below:

- *Spatial Deviation from Shortest Paths*: This section requires case studies that enable examination of large number of routes to establish comprehensive spatial trends.
 - West to East Postal Districts: Large sample size (9850 routes) enables wide scale comparison against optimal alternatives. Large origin and destination regions allow potential identification of deviation points across wide spatial area. Interaction with east–west routes through central London.
 - North to South-East Postal Districts: Rationale as above (3210 in this case). Involves interaction with central London via north–south perspective.
- *Role of Anchors*: This section will incorporate both global (all route) and local (case study based) analyses. Case studies should enable exploration of route choice deviations between highly specific origins and destinations.
 - Globally Significant Features: Assessed across all routes.
 - Locally Significant Features:

- NW3 to EC4 Postal Zones: Long distance and interaction with central London ensures numerous choice points. Small origin and target region but maintains good sample size of 310 routes.
- E14 to NW1 Postal Zones: Rationale as above, but interaction with east of central London ensures consideration of alternative urban form. Sample size of 393 routes is adequate.

- *Directional Influences on Route Choice*: This section requires case studies which offer reasonable route samples in opposite directions between specific locations.

- NW1 to SW11 Postal Zones: Long distance and interaction with central London. River crossing requires selection of a crossing point, potentially revealing of process. Despite distance, good sample sizes in both directions.
- SW1 to N1 Postal Zones: Placement of origin and destination within central London enforces decision around traversal of region. High sample sizes in both directions.

4. Shortest path routing analysis

The primary stage of investigation requires establishing how closely route models that minimise particular attributes align with observed patterns of behaviour. In this section, the methodology by which these assessments are made is described, and the performance of these models across a range of scenarios outlined.

For the assessment, nineteen measures were selected for comparison against the observed choices. Many of these models commonly feature within urban and transportation modelling literature, as central to many transportation models or reflecting a choice set of alternatives considered, selected and rejected in models of route choice (Ben-Akiva et al., 1984; Bekhor et al., 2006). In order to increase the coverage and sophistication of the comparative optimal routes, road properties were combined linearly for some additional cases (e.g. minimisation of travel time and angular deviation).

Route similarity was assessed across the entire dataset, with averages calculated for the percentage similarity by optimising metric. The best performing model, and most reflective of observed behaviour, as shown in the results in Table 2, was the turn-weighted distance model, with an average percentage similarity score of 42.1%. Similarly well represented were the road hierarchy weighted distance model (41.2%) and the pure distance model (39.4%). Scoring poorly with respect to route similarity were many of the angular deviation models, with the least deviation from target model achieving only a 17.50% match on average. Interestingly, the expected travel time model – being based on observed average travel time also performs poorly relative to others, despite it aiming to reflect the most realistic approach to congestion avoidance and travel time minimisation.

The degree of similarity between observed and shortest paths can also be explored across varying journey lengths and time of day. These results show across the board that similarity improves on the mean at shortest trip lengths, but degrades as journey length increases. Taking minimal distance as an example, at journeys between 0.5 and 1 km the mean accuracy is 71.28%, but falls to 28.78% for journeys over 10 km. The results indicate that shorter distances make selection of the optimal route easier, presumably given the few alternatives available. Smaller variations can be observed across the time of travel. Taking distance again as the example, mean accuracy improves during the evening hours (7 pm to 7 am) to 40.69%, falling below the mean during other times of day (to 38.33%, 38.07% and 38.26% for the morning, inter and evening peak periods respectively). Other shortest path measures align with these trends. The results suggest a movement back

¹ The 1500-m limit was established through testing within the central London area, yielding a distance that captures localised regions while limiting fragmentation.

Table 2
Mean percentage similarity for all journeys by shortest routing mechanism.

Objective function	Mean accuracy	Over 75% match
Minimal distance	39.42	14.44
Least free flow travel time	37.78	13.71
Least observed travel time ^a	26.72	5.03
Minimal angular deviation ^b	26.78	7.31
Minimal (angular deviation * distance)	32.56	9.64
Minimal (angular deviation * time)	32.36	10.06
Fewest turns ^c	25.05	4.14
Minimal (turns * distance)	42.10	17.49
Minimal (turns * time)	39.08	15.40
Minimal (right turns * distance)	39.05	13.95
Minimal (right turns * time)	38.02	13.92
Minimal (road category ^d * distance)	41.21	17.16
Minimal (road category * time)	37.88	14.46
Minimal (road category * angular deviation)	28.04	8.62
Minimal (number of lanes * distance)	38.55	14.04
Minimal (number of lanes * time)	34.74	11.96
Minimal (number of lanes * angular deviation)	24.86	6.98
Minimal angular deviation from target	17.52	2.78
Fewest road segments	21.05	2.88

^a Data collected by Transport for London from a fleet of GPS-enabled vehicles over 3 months. Mean travel times are established for five time periods (0–6 h, 6–10 h, 10–16 h, 16–19 h, and 19–0 h), and routes calculated using data corresponding to the period in which the observed journey took place. Where travel time data is unavailable, kinematic calculations are implemented to capture acceleration and deceleration at junctions.

^b Utilising methodology described in Hillier and Iida (2005).

^c In line with theory and applications described in Golledge and Garling (2002) and Duckham and Kulik (2003).

^d Category assigned using the UK Department of Transport road classifications. 'Motorways' and 'A Roads' are given a weight of 1; 'B Roads' are weighted as 2; 'Minor Roads' and 'Local Streets' are weighted with 3; and all remaining roads (e.g. 'Private Roads') weighted with 4.

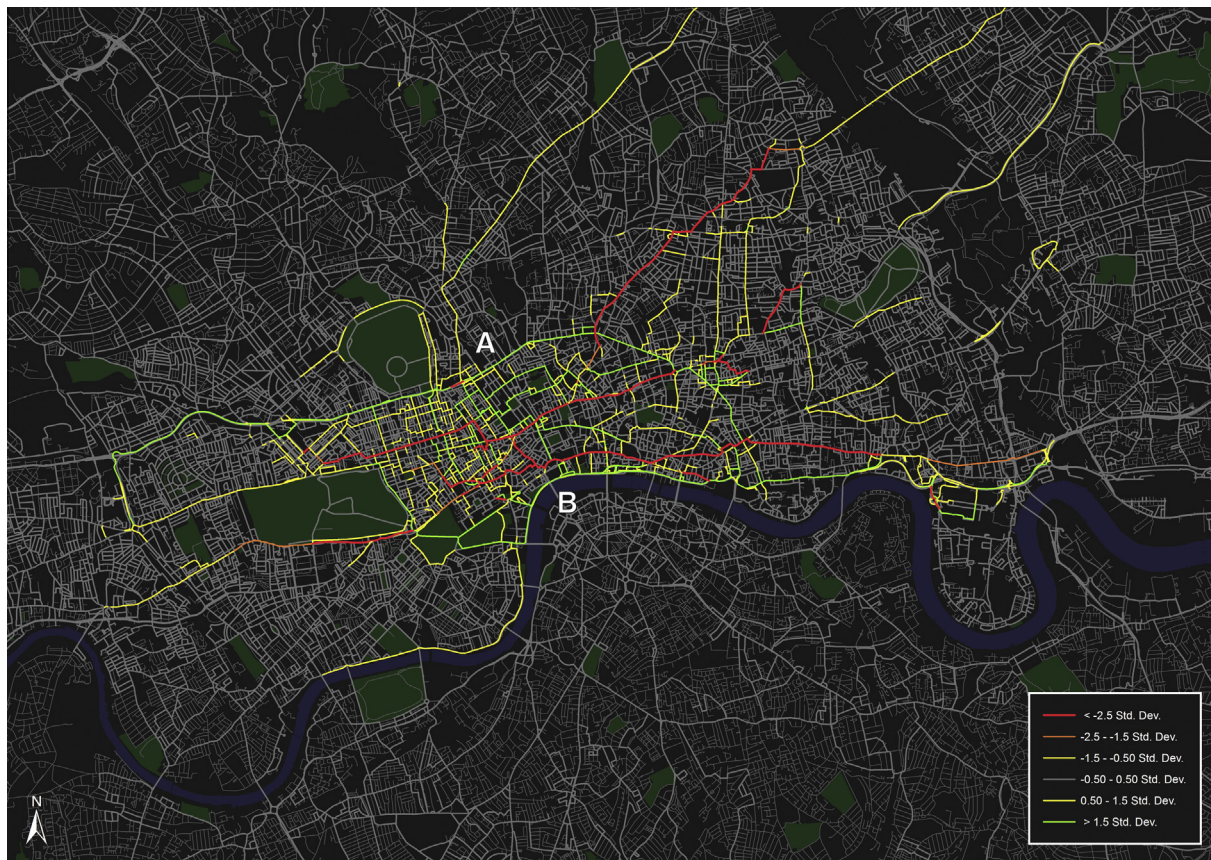


Fig. 2. Standard deviations around differences between observed traffic flows and alternatives calculated using an optimal distance metric between West and East postcodes.

towards optimal routing in response to the reduction of congestion during the evening hours, although the small variation may also represent a greater propensity for shorter journeys during this period.

The calculation of mean accuracy is quite an unforgiving metric on which to judge route accuracy, where poor correlation on some routes may hide stronger correlation on a majority of occasions. Another approach is to assess the number of

occasions on which an optimal approach performs well. In the second column in Table 2 is a value for each model indicating the proportion of routes achieving a 75% match or greater. Once more, poor performance is widespread, with many of the models achieving 75% accuracy on only 10–15% of all routes. This demonstrates that only on a relative few occasions do these models perform at near acceptable levels. These scores are slightly lower than those indicated in a similar previous study (Prato and Bekhor, 2006).

Another perspective can be drawn through measuring how near observed routes come to achieving the lowest potential cost function. By comparing route characteristics, the spatial deviations between observed and shortest alternatives are not of great importance. Taking distance, comparing the lengths of all observed routes against their shortest counterparts, observed routes are found to be 1.39 times the length of the optimal alternative. Against free flow travel time, observed routes are 1.53 times longer the shortest possible route time. Comparing turns, observed routes are shown to use 1.69 more turns per km than alternatives using the fewest possible number of turns. These final results indicate that even when not selecting the shortest possible path, drivers do not perform well with respect to minimising costs.

While the route similarity assessments have provided some insight into the route attributes generally favoured by drivers, it is clear that none of the optimal routes tested offer a strong predictive capability of real-world behaviour. Given the wide range of optimal routes tested, this is perhaps the greatest indictment falls upon the premise that whole optimised routes reflect real-world route choice decision process.

5. Anchor-based routing analysis

The second stage of analysis assesses the role of urban features in influencing route choice. As described earlier, this analysis will consist of three stages – the first, investigating spatial deviations from optimal behaviours with a view to identifying the influence of anchors; the second, exploring statistical distributions in urban feature selection; and third, the potential influence of process in influencing route choice.

5.1. Spatial deviation from shortest paths

According to the same method used above, optimal route alternatives are generated between each observed trip origin and destination point. For this phase of the research, however, only distance and expected travel time routing measures are examined. This is because they represent the models most widely used currently within urban and transportation modelling research. Route sets are calculated for the case studies outlined earlier, and the spatial deviations between observed and modelled route sets are mapped and analysed.

5.1.1. Case study – West to East London

In the first case, lateral travel across central London is explored, with all 9850 journeys originating in west London and finishing in east London postcodes extracted. The origin and destination spread requires traversal of central London, but does not necessitate crossing of the River Thames.

Differences in flow between the actual route set and the two artificially generated route sets were calculated. Figs. 2 and 3 show

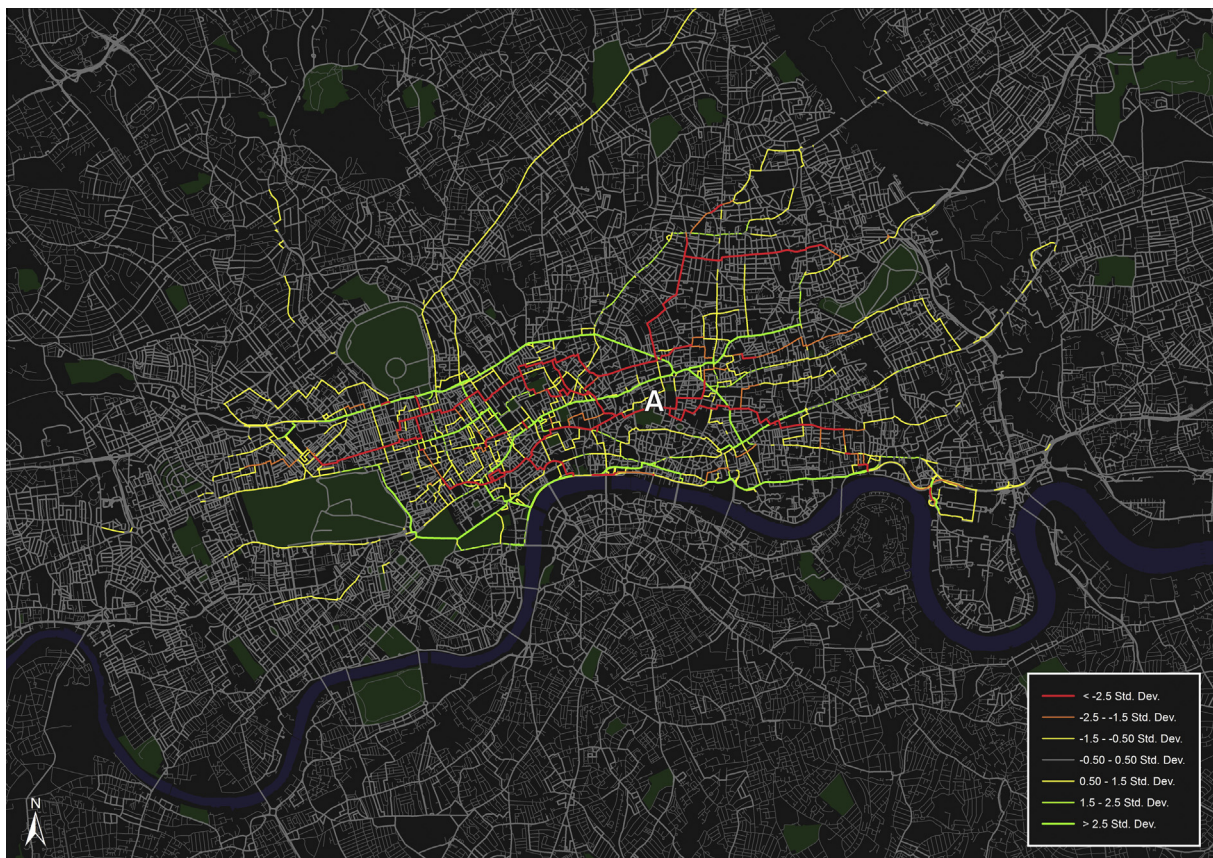


Fig. 3. Standard deviations around differences between observed traffic flows and alternatives calculated using an optimal expected travel time metric between West and East postcodes.

the spatial deviations in route set distributions, categorising by the number of standard deviations around the mean. In both cases, a number of main routes across central London – such as Euston Road and Victoria Embankment (labelled A and B respectively in Fig. 2) – appear to *attract* considerably greater proportions in observed traffic flow, in favour of a number of alternatives that are theoretically more optimal. It is furthermore interesting to note that there is no wholesale avoidance of central London. Rather one can observe that route selection is frequently consigned to straighter sections of the road network, such as the Old Street to Holborn pathway (marked A in Fig. 3). Many of the more optimal routes through this region appear to require more turns, indicative of

requiring a higher effort or cognitive load from the individual routing through that region.

5.1.2. Case study – North to South-East London

In the second case study, 3210 trips were extracted covering trips from all northern postcodes to all south-eastern region postcode regions. The trip distributions of the real and artificially generated routes are shown in Figs. 4 and 5.

Once more, clear deviations between optimal and observed route sets can be tracked. One prominent region is north of central London, where three main routes – Kingsland Road, New North Road and York Way (labelled A in Fig. 5) – all demonstrate



Fig. 4. Standard deviations around differences between observed traffic flows and alternatives calculated using an optimal distance metric between North and South-East postcodes.

relatively greater attraction for observed route selection. Within the central zone, higher proportionate volumes of traffic are clearly identified along the inner ring road section, continuing from Kings Cross around the eastern side of central London to Tower Bridge. There are additional common discrepancies with respect to the selection of bridges to cross the River Thames, where Tower, Blackfriars, Waterloo and Westminster Bridges (labelled B in Fig. 5) all appear more popular alternatives relative to distance and time minimising routes. Furthermore, once over the Thames, an attraction in trips towards the large, busy intersections at Elephant and Castle and Bricklayers' Arms is indicated (shown at point C in

Fig. 5), highlighting these features as potential points of attraction within the route choice decision process.

These analyses provide indication of the locations on the road network attract route selections beyond those that would have otherwise been expected were all individuals to have selected an optimal distance or travel time route. Many of the locations attracting greater flow are continuous, straight sections of the road network. Whereas the apparently more optimal routes rejected by drivers pass consist of a more intricate structure, requiring more turns and so high navigation abilities. The basis of these deviations provides some explanation for the poor general performance of

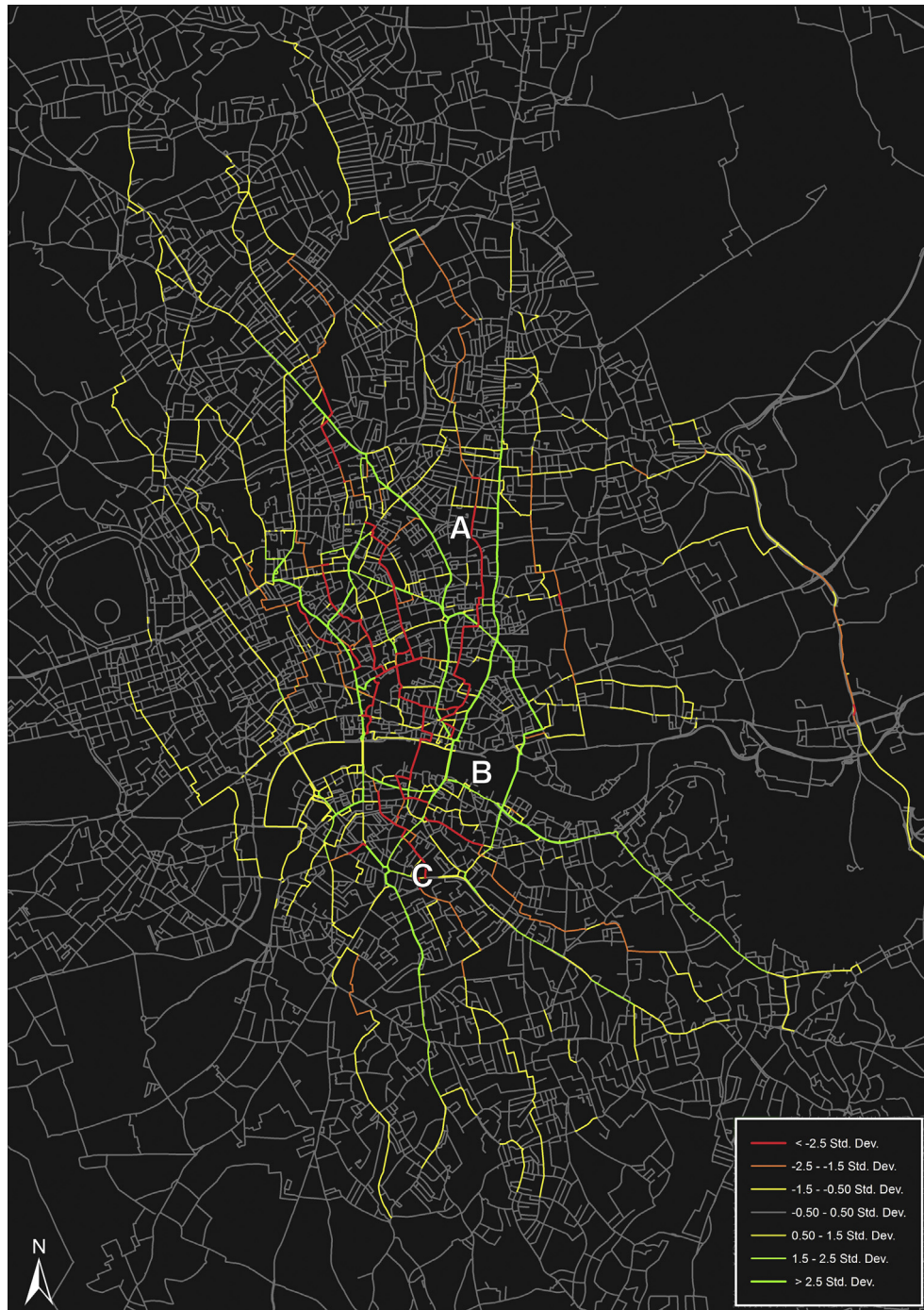


Fig. 5. Standard deviations around differences between observed traffic flows and alternatives calculated using an optimal expected travel time metric between North and South-East postcodes.

optimal routing models observed earlier. The next stage of analysis will aim to improve the classification of these locations, and their role in the route choice process.

5.2. Role of anchors

The assessment of the role of key locations in shaping route choice is carried out at two levels of scale. In the first instance, features that attract a large number of journeys will be examined within the context of all route choices, with the extent of their attractiveness above other locations being established. Second, the role of locations is examined on a finer grained scale, looking at how many decisions may be associated with few localised features.

5.2.1. Globally important anchors

Global significance is measured by extracting the variable utilisation of different parts of the road network over all trips. Seg-

ments within the road network that share a road name are grouped together. Grouping the road network by shared name draws together elements of the network that may be interpreted or recalled as part of a collective feature, either through the nature of their interconnected structure or through common historical or functional association (Lynch, 1960).

For each road name feature, the total number of trips utilising any part of the road name is calculated. The results are ranked and plotted in Fig. 6. As one can observe, the resulting ranked plot aligns closely with a long-tail distribution, where few features attract heavy usage, with the majority of features attracting little usage. At the head of the distribution are roads such as Piccadilly (ranked 1st), Park Lane (2nd), Euston Road (3rd) and Victoria Embankment (6th), noted earlier for seemingly attracting route choices away from optimal alternatives, and each chosen on over 60,000 occasions. In line with any long-tail distribution, these totals drop off quickly as one moves down the rankings, with the majority of roads rarely chosen. This is demonstrated in that cross

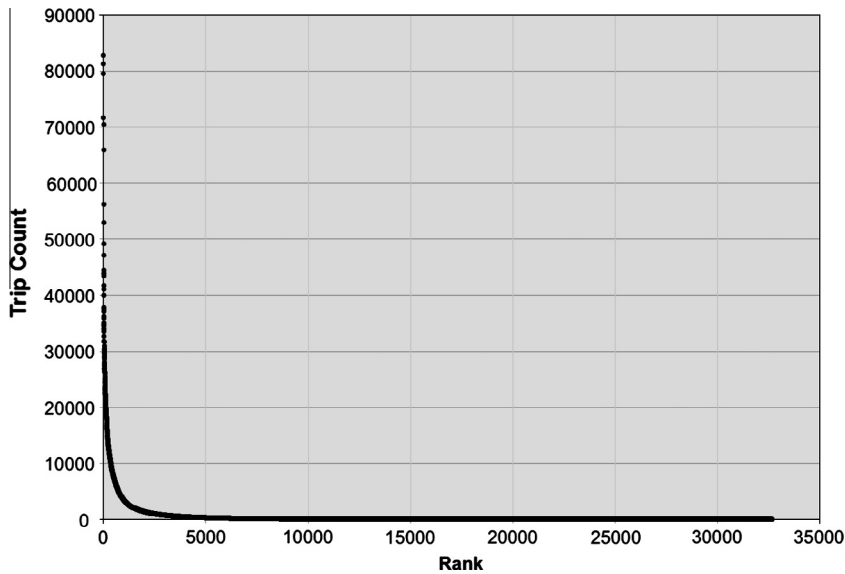


Fig. 6. Ranked distribution of trip counts by named road.

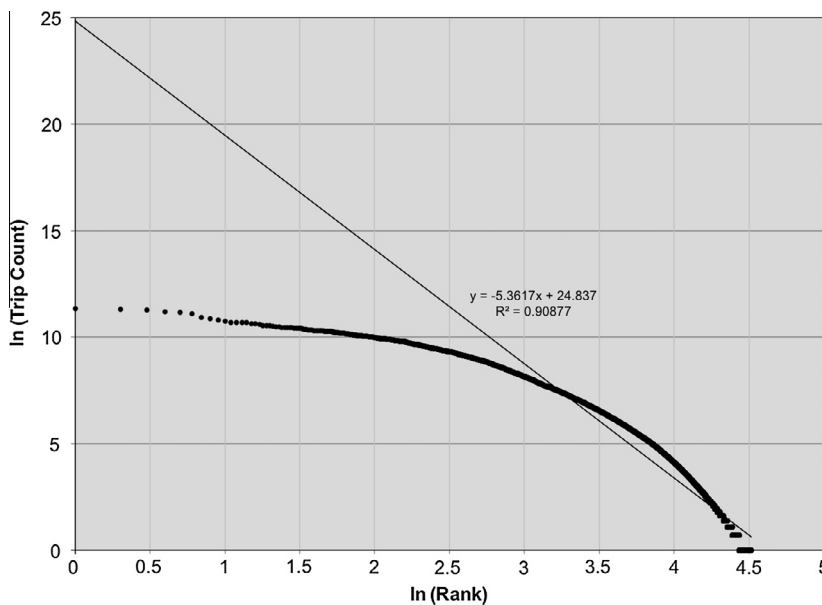


Fig. 7. Log-log plot of rank against trip counts by named road with fitted line of regression.

the dataset, of the total road choices made across all journeys (15.94 million), 80% of these choices correspond to only 1867 of the total 32,659 roads, equating to just 5.71% of all roads.

A long-tailed plot such as this is indicative of a Zipf Law distribution, a pattern of heavily skewed influence observed across a wide range of environments. Zipf's Law asserts that the frequency of a certain event taking place is inversely proportional to its rank, relative to alternative events. This pattern has been observed across a swathe of natural phenomena, including reflecting the growth rate of city populations relative to others (Gabaix, 1999). The simplest way to identify a Zipf Law distribution is to plot $\log(-rank)$ and $\log(trip\ count)$ on a log–log graph, to establish whether a linear relationship exists. This graph is plotted for road choice in Fig. 7. The log relationship between rank and the frequency with which roads are chosen does demonstrate a linear relationship, in line with the principles of Zipf Law. Indeed, a fitted linear line of regression represents this relationship with an R^2 of 0.91, indicative of a strong linear correlation, similar to Zipf Law distributions observed elsewhere. However, despite the strong fit, higher traffic flow along the highest ranked roads should be observed for a stronger alignment with a Zipf distribution.

The alignment with the Zipf Law distribution is indicative of a self-reinforcing mechanism supporting the repeated use of only a

few urban features during route choice. In the same way that cities grow through a preferential attachment mechanism (Simon, 1955; Gabaix, 1999) to form a Zipf distribution, the indication is that the use of these urban features grows in a similar fashion. Once a location has been established as a 'good' place to travel through, subsequent route choices are also directed through that location (provided they meet requirements with respect to the eventual target). The indication is that route choice is *formed around* these locations. Thus, an individual does not select a route that optimises some set of preferences between origin and destination, but rather chooses the dominant features that they wish to traverse en route to their destination.

5.2.2. Locally important anchors

While particular routes have been demonstrated to be important in influencing route choice, the exploration at the global scale potentially hides local interactions with other urban features. In this second stage of analysis, route choices – and the potential role of urban features – are examined at a finer resolution. The distributions of routes between specific locations provide indication of the most important locations for decision-making en route between origin and destination.

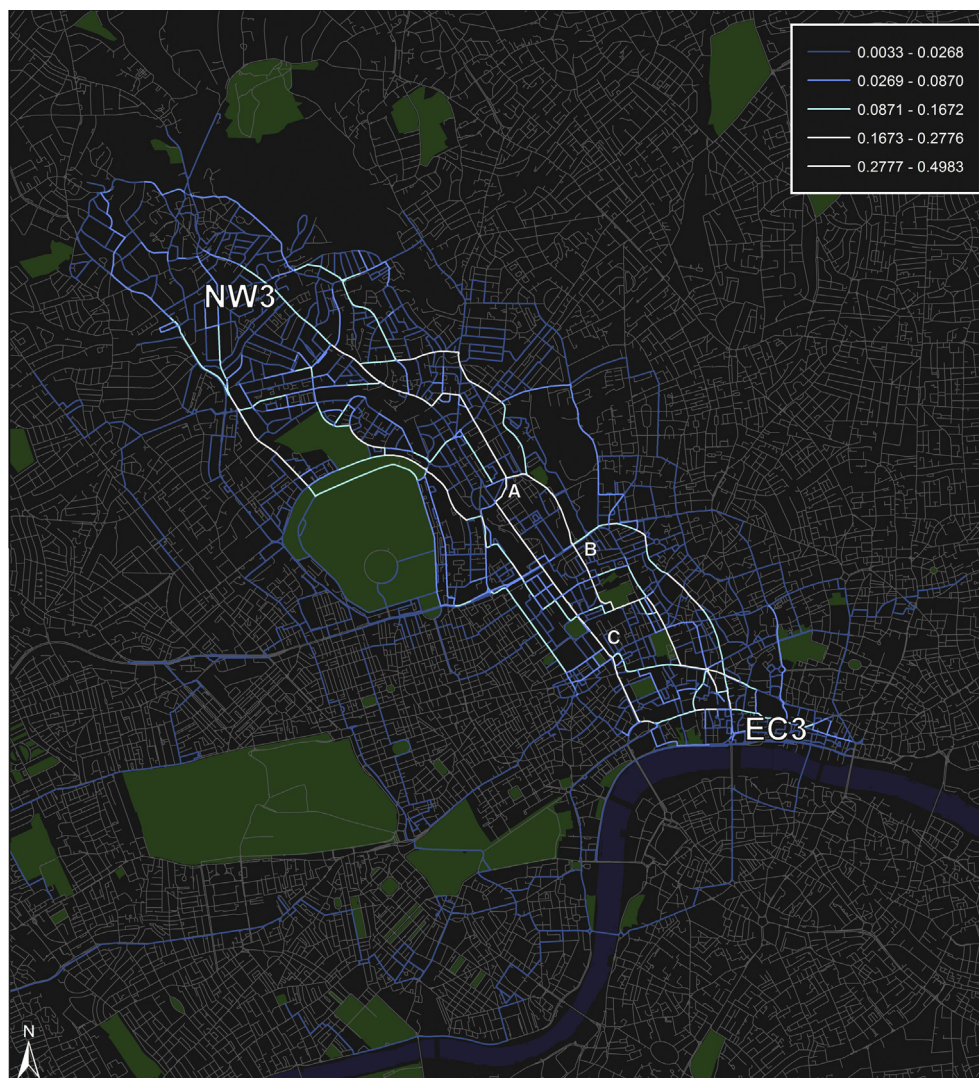


Fig. 8. Proportionate traffic flow of 310 trips from NW3 to EC4.

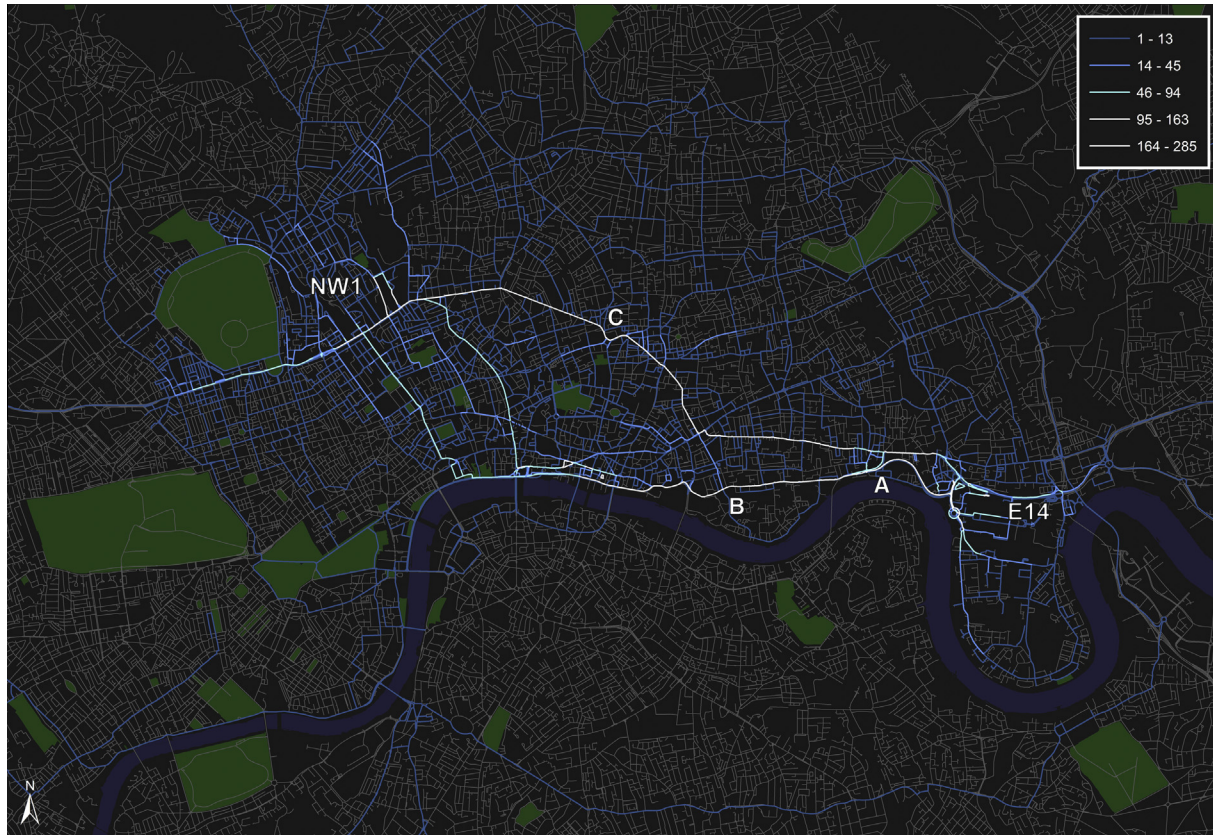


Fig. 9. Proportionate traffic flow of 393 trips from E14 to NW1.

5.2.2.1. Case study – NW3 to EC4 regions. Within the dataset, a total of 310 trips are observed to pass between the NW3 and EC4 post-code regions. Examining the distribution of route selections between the two regions, as demonstrated in Fig. 8, like other examples, there is considerable heterogeneity with a concentration along a few core routes. Yet of additional note within this representation is how a significant degree of variation in behaviour can be tracked back to splits in traffic at particular points.

Assuming that all individuals are broadly aiming to reach their destination in the EC4 region, one can observe that decisions vary at points where individuals are faced with two valid options. Examining particular junctions more specifically, at junction A in Fig. 8, a split of the 109 inflowing vehicles is observed between two more southerly routes, with 42 routes taking the south-easterly bound route, and 67 choosing to take the south-westerly route. Likewise, at junction B, a split of 102 vehicles is observed, with 38 choosing a eastbound route along Euston Road, with 62 vehicles choosing to travel straight onwards. At junction C, of the 88 inflowing vehicles, 61 chose to travel straight onwards and 27 chose to take the Holborn Gyratory to the east.

5.2.2.2. Case study – E14 to NW1 regions. The set of 393 route choices between E14 and NW1 are interesting given the prevalence of three dominant routes between origin and destination zones. As can be observed in Fig. 9, many trips pass via global attractors including Victoria Embankment (passing along the river), City Road (heading north-east) and Farringdon Road (heading north from the river).

The prevalence of each of these attractive locations leads to a number of particular junctions at which route choice decisions must be taken. At junction A, representing a decision point

between westbound routes (via the Victoria Embankment route) and those heading northbound (via the City Road route), of the 222 inflowing trips, 42 decided to deviate, with 180 deciding to continue straight on. Likewise, a similar decision is made at junction B, where, of the 171 inflowing trips, 145 are shown to continue westbound, with 30 taking the turning northwards. Elsewhere, roundabouts are clearly important decision points, and this is proven within this scenario, where, of the 123 individuals arriving on Old Street at the roundabout at point C, 109 deviate northbound, and just 14 continue eastwards.

The exploration of two case studies, where origin and destination are known to within a specific spatial extent, has highlighted the importance of particular urban features in shaping fine-grained route choices. Despite a fixed origin and destination, particular urban features provoke different choices, resulting in different routes being selected. The result is significant heterogeneity in route choice, led by deviation of choices at these urban features, that again likely influences earlier observed poor correlation with shortest path routes. The indication from these examples is that – while route choice is strongly influenced by the dominant features identified in the previous section – choices around the selection of those features are undertaken at a local level, in interactions with major junctions.

5.3. Directional influences on route choice

Prior indications suggest that urban features form the basis of route choice, however, there remains a lack of insight into the process through which these choices are made. To explore this, the final stage of analysis examines variation in routing behaviour by direction of travel across two case studies. The aim is to identify

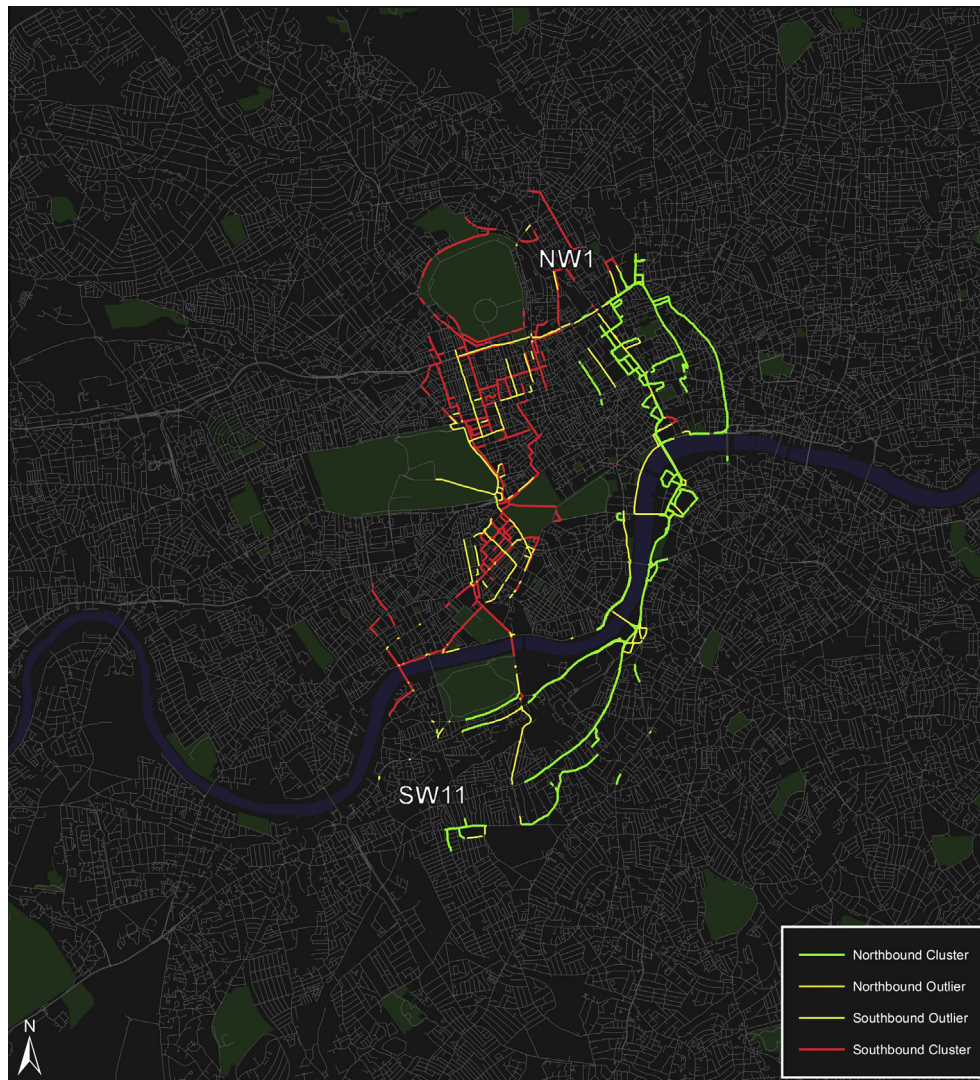


Fig. 10. Spatial clustering results for trips between NW1 and SW11 regions indicating significant areas of high mono-directional traffic flow.

whether deviations exist, and whether the nature of these deviations provides additional insight into choice behaviours.

To accomplish this aim, spatial deviations in route flows are extracted for route sets running in opposite directions. Deviations are calculated on a road segment-by-segment basis, but differences are calculated by area using the LISA local autocorrelation method described earlier. This approach not only provides us with statistically significant² spatial clusters of deviation in route flows, the method incorporates wider spatial trends, beyond merely segment-level deviations.

5.3.1. Case study – NW1 to SW11 regions

Trips between NW1 and SW11 are extracted, consisting of 542 trips running northbound from SW11 to NW1, and 310 trips in the opposite direction. Proportionate flows for each route set are calculated, and spatial clusters derived. These clusters are shown in Fig. 10.

² LISA calculates a local Moran's I statistics that is translated into a Z-score, against a null hypothesis of no local autocorrelation. Clusters are extracted where they exceed a 0.05 probability threshold.

From Fig. 10, it is clear that the Park Lane region, in addition to Marylebone and Euston Road, are identified as significantly more attractive to southbound traffic than northbound. Conversely, northbound traffic is significantly more drawn towards the routes running along the south bank of the river Thames, moving up towards the destination through the Holborn and Bloomsbury regions.

Aside from the clusters, the locations of outliers – routes where flow volume counters the dominant flow in the area provide further insight into route selection. These cases include the preference for travel along the north bank of the Thames where travelling southbound. It is important to note that there are no traffic regulations within the area that would likely cause this deviation. In other areas, such as within the northbound cluster, one-way streets are the cause of outliers appearing, amongst a dominance of northbound flow.

5.3.2. Case study – SW1 to N1 regions

The second case study examines counter-acting flow between the SW1 and N1 regions locations, which, in contrast to the previous case, are not separated by the barrier of the river Thames. In this case, 672 routes are extracted running northbound from

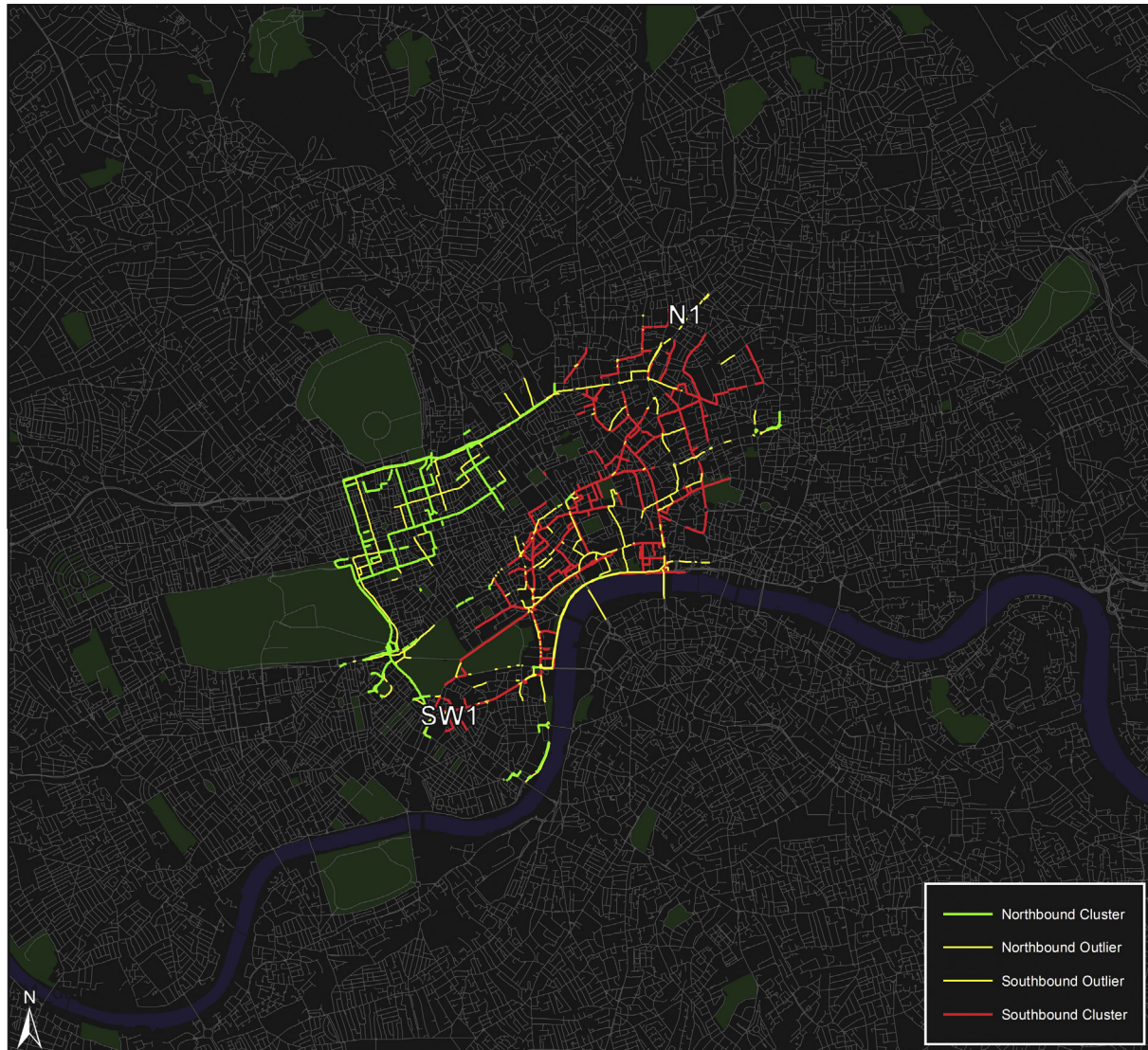


Fig. 11. Spatial clustering results for trips between SW1 and N1 regions indicating significant areas of high mono-directional traffic flow.

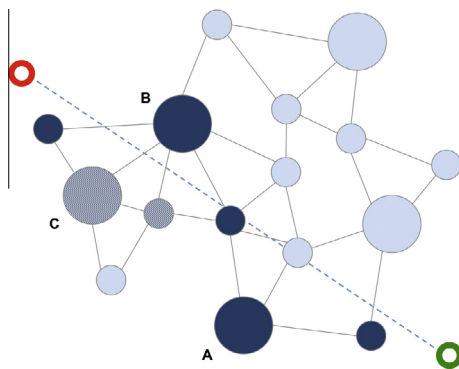


Fig. 12. Hypothetical application of the route choice behavioural framework.

SW1 to N1, and 898 trips running in the other direction. The spatial clusters established from the differences in proportionate flow are shown in Fig. 11.

Observing the spatial clusters presented, it is apparent that a higher volume of northbound route choice is once again found along the Park Lane and Euston Road pathway, in preference to

alternative routes through central London. Southbound traffic, conversely, more regularly follows a more direct set of routes through central regions. Only along the yellow-coloured Embankment route, north of the river Thames, are no significantly dominant clusters are found in either direction. The wider patterns, once again, demonstrate a cyclical nature to the bi-directionality of travel, with little indication of route regulation playing a significant role in the definition of the wider differences.

Across both case studies, the presence of asymmetry in route selection is clear. Its presence indicates that route choice is influenced by the availability of subgoals attractive from the point of origin. Travelling from one direction, a subgoal may seem an approachable initial target; from another direction this may not be the case, and a completely different location, perhaps one that is nearer, is deemed logically the most appropriate first location to target. More significantly, this process indicates that the route choice process is *step-based*, and that individuals construct their route by moving from one sub-goal to the next. Once the first sub-goal has been reached, the next sub-goal is chosen, until the destination is reached. This evidence provides further explanation to why shortest routes so poorly reflected observed route choice in earlier tests.

6. Discussion and conclusions

The analyses laid out in this paper have described how closely the routing behaviours of minicab drivers in London align with the shortest path and anchor-based concepts of route choice. Although it is impossible to fully clarify the nature of route choice through empirical observations alone, some strong trends have emerged. In this section, the major findings from these analyses are digested, and then positioned within an initial conceptual framework for route choice. Following this, the potential application and future directions for this research will be discussed.

In the first instance, the analyses highlighted the poor performance of shortest path routing methodologies in predicting observed behaviour. Not only were shortest path alternatives a poor match for observed routing behaviours, drivers were shown to travel further in terms of distance and travel time than would be deemed optimal.

The tests applied to establishing the role of anchors in route choice were more promising. The use of a few locations far beyond the vast majority of other routes was well demonstrated across a range of scales. The prevalence of these locations, and their regularity of their selection far beyond that indicated by shortest path alternatives, is indicative of these features being used as the *basis* for route choice. The repeated use of these features suggests that individuals principally consider these locations during the construction of a route plan, rather than seeking to select a wholly optimal route.

At a finer scale, there were strong indications that road junctions are used as points at which additional route choices are made. Within the context of anchor-based navigation, these locations may be seen as points at which navigation towards an anchor is refined. They have a role in shaping route choice, but do not hold the same overarching role of anchors.

The findings around the asymmetry support the notion of anchor-based navigation. With no observable impact from route restrictions present across the case studies, the discrepancy in route choices made in opposite directions indicates a role for intermediary locations, perceived differently from each origin, in shaping decisions. These trends indicate the presence of a *step-wise* methodology within the choice process. In this sense, the best initial choice to make – to route towards a particular anchor – will vary according to the location from which a choice is made. As the individual routes towards their destination, this initial choice will then influence subsequent choices. The nature of asymmetry casts further doubt around how optimal routing methodologies reflect the choice process, of which most identify the same optimal route regardless of direction of travel.

6.1. Towards an anchor-based framework of route choice

The strength of the results presents an argument for the development of a new framework for modelling route choice. While this study is limited in a number of ways – as highlighted during the description of the dataset earlier – an initial framework may be developed from these results for potential future validation or elaboration at a later stage, with the availability of additional datasets.

The development of an anchor-based framework for route choice incorporates not only the main findings from this study, but is supported by findings within the literature. As such, anchor selection becomes the basis for route choice, and is supported by subsequent choices made at road junctions. The selection of anchors is shaped by a desire to minimise distance and angular

deviation from the target location, and individuals move broadly towards their destination based on a cognitive sense of direction.

The designed process proceeds iteratively, as the individual moves from anchor to anchor until their reach their destination.

- (1) Establish broad direction and distance of destination from current position.
- (2) Locate the best nearby anchor that a) minimises deviation from destination direction angle, b) moves closer to the destination than the current location, and c) is not expected to introduce unacceptable delay through congestion.
- (3) Navigate towards subgoal using known intermediate junctions, avoiding locations with high 'expected' congestion.
- (4) On reaching subgoal, if not close to destination restart process.

The application of this process is best described through worked example, mapped out in Fig. 12. This figure presents a simplified representation of the subgoals and junctions that dominate the route choice decision process, and the links between features. In undertaking route choice, the individual first identifies the direction of the destination from the origin, indicated as red and green circles respectively. Once established, the individual moves to step 2, choosing the subgoal that is best aligned with the principles of deviation, distance and delay, namely point A in this case. To reach point A, the individual points through the local road network, via intermediary junctions, located at less prominent points, which require traversal in order to link between subgoals. Upon reaching point A, the individual makes restarts this decision process, travelling to subgoal B according to the same principles. The individual would ultimately therefore select a route through the darker objects within this network.

Two outcomes of this framework should be highlighted. First, the process is subjective, and limited by the extents of an individual's knowledge – a factor strongly indicated within current literature. While a more optimal route between origin and destination may be found by routing via subgoals A and then C, the individual is restricted from doing so by not holding enough knowledge of the road network to make this a possibility. As such, the known route via subgoal B is selected. Second, the process furthermore explains asymmetry in route choice. If the direction of travel were to be reversed, the individual would be expected to first choose subgoal C, rather than subgoal B, resulting in the differences highlighted by the dashed circles in Fig. 12.

6.2. Future directions

The findings from this research present a number of routes for potential expansion and elaboration. As the breadth and volume of similar datasets emerge, opportunities will emerge that enable the major findings outlined during this work to be validated. The use alternative datasets will provide evidence for or contrary to the trends identified here. Validation datasets are required that overcome some of the limitations experienced here with respect to the geographic concentration of routes in central London, the above-average expertise of drivers, and the potential influence of GPS devices.

With positive validation, the argument for the development of methods that place anchors at the centre of route generation will become stronger. In this case, a major route of research would involve the integration of these methods with conventional route choice methods, potentially as an alternative method for generating route choice sets. The route choice framework requires expansion, parameterisation and eventual implementation. One major

stage in this process will involve the establishment of a spatial hierarchy, representing the anchors upon which route choice decisions are made. Factors relating to street configuration, traffic engineering, traffic flow and urban form may all be investigated in deriving this structure.

The implementation of the model furthermore requires more exactly specifying the decision-making stages outlined within the earlier framework. For example, what is deemed an acceptable deviation from the destination? What is considered acceptable delay? And perhaps most importantly, how do these preferences vary across the population of decision-makers, and how are these variations encapsulated within the model? There is considerable promise for integration with models incorporating risk and uncertainty in route choice (Avineri and Bovy, 2008; Gao et al., 2010; Chorus et al., 2008). There must be thought given to the process by which individuals learn anchor-based routes, which may again build on current literature (Senk, 2010). The inclusion of heterogeneous measures of perception would indicate the potential for the use of agent-based modelling for capturing full complexity of behaviours at an individual level.

In the coming years, another interesting avenue of research may potentially involve observing how routing behaviours shift with the increasing proliferation of traffic information and navigation guidance devices (Ben-Elia and Avineri, 2015). While in this study (undertaken between late 2010 and early 2011) its role is shown to be limited, later technological advances may lead to greater reliance on algorithmically-generated optimal routes rather than personal knowledge of space. This has implications for the degradation of spatial knowledge, as suggested elsewhere (Parush et al., 2007). It may also contribute towards the formation of equilibrium in traffic flow, where all individuals are optimising their behaviours based on near-perfect information. Nevertheless, work will furthermore be required around establishing the behavioural nature of how traffic information and route guidance is used. Identifying the environments and contexts within which individuals solely (or partly) use guided information and where travellers continue to utilise their own spatial knowledge.

Acknowledgements

This work is part of the STANDARD project (Spatio-Temporal Analysis of Network Data and Road Developments), supported by the UK Engineering and Physical Sciences Research Council (EP/G023212/1) and Transport for London. Gratitude is extended to Addison Lee Limited for the use of their datasets for the purposes of this research. The authors would like to thank the three anonymous reviewers for their very useful comments in helping shape this paper.

References

- Anselin, L., 1995. Local indicators of spatial association-LISA. *Geogr. Anal.* 27 (2), 93–115.
- Avineri, E., Bovy, Piet H.L., 2008. Identification of parameters for a prospect theory model for travel choice analysis. *Transport. Res. Rec.: J. Transport. Res. Board* 2082 (1), 141–147.
- Bekhor, S., Ben-Akiva, M., Ramming, M.S., 2006. Evaluation of choice set generation algorithms for route choice models. *Ann. Oper. Res.* 144 (1), 235–247.
- Ben-Akiva, M., Bergman, M.J., Daly, A.J., Ramaswamy, R., 1984. Modeling inter-urban route choice behaviour. In: *Proceedings from the Ninth International Symposium on Transportation and Traffic Theory*. VNU Science Press, Utrecht, Netherlands, pp. 299–330.
- Ben-Elia, E., Avineri, E., 2015. Opportunities and threats of travel information: a behavioural perspective. In: Thomopoulos, N., Givoni, M., Rietveld, P. (Eds.), *ICT for Transport: Opportunities and Threats – NECTAR Series on Transportation and Communications Research*. Edward Elgar.
- Ben-Elia, E., Shifan, Y., 2010. Which road do I take? A learning-based model of route-choice behavior with real-time information. *Transport. Res. A: Policy Pract.* 44 (4), 249–264.
- Ben-Elia, E., Ishaq, R., Shifan, Y., 2013. 'If only I had taken the other road...': regret, risk and reinforced learning in informed route-choice. *Transportation* 40 (2), 269–293.
- Bovy, P.H.L., 2009. On modelling route choice sets in transportation networks: a synthesis. *Transp. Rev.* 29 (1), 43–68.
- Carr, S., Schissler, D., 1969. The city as a trip: perceptual selection and memory in the view from the road. *Environ. Behav.*
- Cascetta, E., Russo, F., Viola, F., Vitetta, A., 2002. A model of route perception in urban road networks. *Transport. Res. B: Methodol.* 36 (7), 577–592.
- Chorus, C.G., Arentze, T.A., Timmermans, H.J.P., 2008. A random regret-minimization model of travel choice. *Transport. Res. B: Methodol.* 42 (1), 1–18.
- Chown, E., Kaplan, S., Kortenkamp, D., 1995. Prototypes, location, and associative networks (PLAN): towards a unified theory of cognitive mapping. *Cogn. Sci.* 19 (1), 1–51.
- Couclelis, H., Golledge, R., Gale, N., Tobler, W., 1987. Exploring the anchor-point hypothesis of spatial cognition. *J. Environ. Psychol.* 7 (2), 99–122.
- Daganzo, C., Sheffi, Y., 1977. On stochastic models of traffic assignment. *Transport. Sci.* 11 (3), 253–274.
- de Palma, A., Ben-Akiva, M., Brownstone, D., Holt, C., Magnac, T., McFadden, D., Moffatt, P., Picard, N., Train, K., Wakker, P., Walker, J., 2008. Risk, uncertainty and discrete choice models. *Market. Lett.* 19 (3), 269–285.
- Dijkstra, E.W., 1959. A note on two problems in connection with graphs. *Numer. Math.* 1, 269–271.
- Duckham, M., Kulik, L., 2003. 'Simplest' paths: automated route selection for navigation. *Spatial Information Theory. Found. Geogr. Inform. Sci.* 2825, 169–185.
- Feng, T., Arentze, T., Timmermans, H., 2013. Capturing preference heterogeneity of truck drivers' route choice behavior with context effects using a latent class model. *EJTIR* 13 (4), 259–273.
- Gabaix, X., 1999. Zipf's law for cities: an explanation. *Q. J. Econ.* 114 (3), 739–767.
- Gao, S., Frejinger, E., Ben-Akiva, M., 2010. Adaptive route choices in risky traffic networks: a prospect theory approach. *Transport. Res. C: Emerg. Technol.* 18 (5), 727–740.
- Garling, T., Laitila, T., Westlin, K., 1998. *Theoretical Foundations of Travel Choice Modelling*. Pergamon.
- Golledge, R.G., Garling, T., 2001. *Spatial Behavior in Transportation Modeling and Planning*. Technical Report. University of California Transportation Center (June).
- Golledge, R., Garling, T., 2002. *Cognitive Maps and Urban Travel*. Technical Report. University of California Transportation Center, University of California.
- Golledge, R.G., Spector, A.N., 1978. Comprehending the urban environment: theory and practice. *Geogr. Anal.* 10 (4), 403–426.
- Golledge, R., Smith, T., Pellegrino, J., Doherty, S., Marshall, S., 1985. A conceptual model and empirical analysis of children's acquisition of spatial knowledge. *J. Environ. Psychol.* 5 (2), 125–152.
- Gopal, S., Klatzky, R., Smith, T., 1989. Navigator: a psychologically based model of environmental learning through navigation. *J. Environ. Psychol.* 9 (4), 309–331.
- Hillier, B., Iida, S., 2005. Network and psychological effects in urban movement. In: *Proceedings of Spatial Information Theory*. Springer-Verlag, pp. 475–490.
- Hirtle, S., Jonides, J., 1985. Evidence of hierarchies in cognitive maps. *Mem. Cogn.* 13 (3), 208–217.
- Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. *Econometr.: J. Econometr. Soc.* 47 (2), 263–91.
- Kaplan, S., Prato, C.G., 2012. Closing the gap between behavior and models in route choice: the role of spatiotemporal constraints and latent traits in choice set formation. *Transport. Res. F: Traffic Psychol. Behav.* 15 (1), 9–24.
- Kuipers, B., 1978. Modeling spatial knowledge. *Cogn. Sci.* 2 (2), 129–153.
- Kuipers, B.J., Levitt, T.S., 1988. Navigation and mapping in large scale space. *AI Mag.* 9.
- Lynch, K., 1960. *The Image of the City*, vol. 1. MIT Press.
- Mahmassani, H.S., Liu, Y.-H., 1999. Dynamics of commuting decision behaviour under advanced traveller information systems. *Transport. Res. C: Emerg. Technol.* 7 (2–3), 91–107.
- Merchant, D.K., Nemhauser, G.L., 1978. A model and an algorithm for the dynamic traffic assignment problems. *Transport. Sci.* 12 (3), 183–199.
- O'Keefe, J., Nadel, L., 1978. *The Hippocampus as a Cognitive Map*. Clarendon Press, Oxford.
- Parush, A., Ahuvia, S., Erev, I., 2007. Degradation in spatial knowledge acquisition when using automatic navigation systems. In: Winter, S., Duckham, M., Kulik, L., Kuipers, B. (Eds.), *Spatial Information Theory*. Springer, Berlin Heidelberg.
- Passini, R., 1984. Spatial representations, a wayfinding perspective. *J. Environ. Psychol.* 4 (2), 153–164.
- Péruch, P., Giraudo, M.-D., Garling, T., 1989. Distance cognition by taxi drivers and the general public. *J. Environ. Psychol.* 9 (3), 233–239.
- Prato, C.G., Bekhor, S., 2006. Applying branch-and-bound technique to route choice set generation. *Transport. Res. Rec.: J. Transport. Res. Board* 1985 (1), 19–28.
- Prato, C.G., Bekhor, S., Pronello, C., 2012. Latent variables and route choice behavior. *Transportation* 39 (2), 299–319.
- Ramming, M.S., 2002. *Network Knowledge and Route Choice*. Ph.D. thesis, MIT.
- Senk, P., 2010. Route choice under the microscope. *Transport. Res. Rec.: J. Transport. Res. Board* 2156 (1), 56–63.
- Siegel, A.W., White, S.H., 1975. The development of spatial representations of large-scale environments. *Adv. Child Dev. Behav.* 10, 9–55.

- Simon, H.A., 1955. On a class of skew distribution functions. *Biometrika* 42 (3–4), 425–440.
- Sun, Zhongwei, Arentze, Theo, Timmermans, Harry, 2012. A heterogeneous latent class model of activity rescheduling, route choice and information acquisition decisions under multiple uncertain events. *Transport. Res. C: Emerg. Technol.* 25, 46–60.
- Tomko, M., Winter, S., Claramunt, C., 2007. Experiential hierarchies of streets. *Comput. Environ. Urban Syst.* 32 (1), 41–52.
- Wardrop, J.G., 1952. Some theoretical aspects of road traffic research. *ICE Proc.: Eng. Div.* 1 (3), 325–362.
- Wiener, J.M., Mallot, H.A., 2003. 'Fine-to-Coarse' route planning and navigation in regionalized environments. *Spatial Cogn. Comput.* 3 (4), 331–358.
- Wiener, J.M., Ehbauer, N.N., Mallot, H.A., 2009. Planning paths to multiple targets: memory involvement and planning heuristics in spatial problem solving. *Psychol. Res.* 73 (5), 644–658.