SEPTEMBER 2018

PROJECT NO SPATS1-403

LST Trial Evaluation 2018-19

Project Note E1: Routing and Operational Analysis by Road Type

PROJECT NOTE

PREPARED ON BEHALF OF THE WSP SUPPLIER GROUP BY:

risksolutions
leave nothing to chance

microlise

Unit E4, Green Lane Business Park
Tewkesbury
GL20 8SJ

Phone: +44 (0) 1684 851751
Fax: +44 (0) 1684 8517111
www.wspgroup.com

Quality Management

ISSUE/REVISION	FIRST ISSUE	REVISION 1	REVISION 2	REVISION 3
Remarks				
Date				
Prepared by	Jonathan Hyde	Jonathan Hyde	Jonathan Hyde	Jonathan Hyde
Signature				
Checked by	Michelle Boath	Michelle Boath	Michelle Boath	Helen Wilkinson
Signature				
Authorised by	Paul Brand (for Risk Solutions)	Paul Brand (for Risk Solutions)	Paul Brand (for Risk Solutions)	Paul Brand (for Risk Solutions)
Signature				
Project number	WSP: 70038230 Risk Solutions: J00005	WSP: 70038230 Risk Solutions: J00005	WSP: 70038230 Risk Solutions: J00005	WSP: 70038230 Risk Solutions: J00005
Report number	PN-E1-V1	PN-E1-V2	PN-E1-V3	PN-E1-V4-2 FINAL
File reference	SPATS 1-403-PN-E1-1 v1	SPATS 1-403-PN-E1-2 v1	SPATS 1-403-PN-E1-3 v1	SPATS 1-403-PN-E1-v4-2 FINAL

PROJECT NOTE DOCUMENT STATUS AND CONFIDENTIALITY

Status	This document is a **Project Note** as defined in the original proposal for the project. **Project Notes are produced iteratively on work for which the scope and method are emergent**. Each version describes the progress on the analysis to that point and proposes the direction and scope of the next step in the process. Where agreed by the client, that scope or direction may diverge from the thinking described in the original proposal. Project Notes, in their intermediate versions, are **not formal deliverables under the contract**. The final version of a Project Note may be a formal deliverable if the proposal states this.
Confidentiality	Project Notes are not normally for publication by any party, including DfT, as they contain discussion of work in progress during the development of an analysis. In this case, a final version of this project note (v4 only) has being 'upgraded' to publishable quality so that it can be issued with the main LST Trial 2017 Annual Report

Project Note E1: Routing and Operational Analysis by Road Type

A project note for Department for Transport

September 2018

Client reference: PPRO 04/91/13

Report reference: J00005 PN E1

Issue 4-2 FINAL

Risk Solutions
Dallam Court
Dallam Lane
Warrington WA2 7LT
United Kingdom

☏ 01925 413984

www.risksol.co.uk

Summary

The objective of the E1 GPS Routing workstream was to refine and validate the routing model of Longer Semi Trailers (LST) (described in the 2016 LST Trial Annual Report), by using actual semi-trailer routes, represented by GPS tracking data. The second objective was to then expand the breadth of analysis of the routes using the same model.

The method used involved collecting and analysing GPS data for real tracked semi-trailer journeys. These real routes were compared with modelled routes, created using an LST routing model. Any significant variations in the routes taken inform changes to parameters of the routing model, which was then used to model the likely routes taken by LSTs between all ~58,000 unique operator start to end location pairs used during the 2017 LST trial period.

An operational analysis of LSTs using these modelled routes provides valuable information to estimate the impact of introducing LSTs to England, Scotland and Wales.

This Project Note describes the result of our activities.

1) Collected a good quality sample of GPS data:

- Obtained permission from nine LST trial operators to use their HGV semi-trailer GPS tracking data.
- Designed and tested processes to identify and obtain relevant GPS journey data for articulated HGVs operating between the same start and end locations as those used by LSTs in the trial.
- Extracted GPS journey data for 790,132 articulated HGV journeys from October 2017 to February 2018.
- Cleaned the GPS journey data, resulting in a sample of 262,935 articulated HGV journeys, from 12,406 unique LST start and end locations. 4.4% of these journeys are known to have been made using LSTs.

2) Mapped GPS data to ITN road links:

- Designed and tested processes to map the millions of GPS waypoints in the GPS journeys to OS ITN road links.
- Created a journey data set of ITN road links representing the 262,935 GPS tracked journeys.

3) Compared GPS tracked journeys to their equivalent LST modelled routes:

- Designed and tested processes to score the similarity of journeys, to allow comparison of modelled routes with GPS tracked journeys.
- Scored the modelled route fit to the GPS tracked journeys in each group of journeys.
- Explored the possible reasons for dissimilarities between modelled routes and GPS tracked journeys, proposing ways to improve the route modelling.
- Implemented improvements to the routing algorithm used to model routes.

4) Operational analysis of LST journeys:

- Used the routing algorithm to estimate the likely routes taken for 828,778 LST journey legs in 2017 and collected data describing the resulting routes into a master routing data set for analysis.

- Analysed the 2017 LST journey legs in England, Scotland and Wales to determine updated and validated urban vs rural vehicle kilometre split, and operational analysis by network and road type.

An improved and validated routing algorithm

Gaining operator permissions took some time. The delay was not because of a lack of willingness to share the data but, in the main, the need to follow strict internal data protection processes before releasing the information.

The GPS journey data required a significant amount of effort to extract and clean. However, we were able to score our routing algorithm against a sample of around 224,000 GPS journeys (a quarter of the number of GPS trial journeys in 2017), covering 11,300 unique known LST start/end locations (a fifth of the actual start/end locations LSTs used in 2017).

Our analysis of the route fit of the modelled routes to the GPS tracked journeys demonstrates that:

- **The modelled LST routes are a reasonable representation of the GPS tracked journeys:** The estimated routes created by the routing algorithm are very similar to two thirds of GPS tracked articulated HGV journeys, 17% are identical.
- **Where GPS tracked journeys use a different route to the modelled routes, the journeys are on equivalent types of roads:** On average the GPS journeys are 1.7% longer, but the distance travelled on urban roads is on average 0.7% less than modelled routes, and the distance travelled while not on the Primary Route Network[1] (PRN) is 1.5% more.
- Any estimates of the distance travelled by LSTs on urban roads using the improved routing algorithm should be reduced by a 0.7% calibration factor.

LST trial operational analysis

Using a validated routing model, we estimated the likely routes taken for around 828,778 of the LST trial journeys made in 2017 (93% of all journeys), taking over 56,377 unique operator routes. This represents approximately 100 million vehicle km driven by trial LSTs.

The trial LSTs are using similar roads to the national HGV articulated fleet. The proportion of vehicle km on different road types for trial LSTs closely matches that reported by the DfT for the national articulated HGV fleet in England and Wales

The proportion of vehicle km travelled by trial LSTs in urban areas is 13.1%, significantly higher than the England and Wales average in 2016 (5.9% of distance travelled by articulated HGVs). Motorways are considered as non-urban in this analysis, but a considerable distance travelled by LSTs use A roads that are included in the trunk road network[2]. These roads are classed as urban, but provide good separation between traffic and vulnerable road users. In our analysis there are a small number of long routes with many thousands of journeys that travel through these areas.

---

1 The Primary Route Network (PRN) designates roads between places of traffic importance across the UK, as identified by DfT This includes motorways and some significant A roads between cities and major towns, (the A roads included are typically coloured green on road maps, and have green road signs).

2 The trunk road network designates the motorways and major A roads managed by Highways England (the Strategic Road network) and the equivalent motorways and A roads in Wales and Scotland.

There is a very small group of routes used by trial LSTs that have a disproportionately significant contribution to vehicle km, and other characteristics of LST operation, such as rural/urban split. These routes are reasonably long and are being used for many thousands of journey legs.

If all trunk (SRN) roads are classified as non-urban, then the trial LSTs are travelling through urban areas for around 7.6% of their journey distance.

The trial LSTs are travelling through urban areas on roads that are not part of the Primary Route Network (in cities, towns and villages), for around 2.9% of their journey distance.

Contents

1 INTRODUCTION

1.1 This project is one of a number of Special Topic Analyses forming part of the wider LST Trial Evaluation programme.

1.2 In the LST Trial 2016 Annual Report we estimated the split between urban and rural vehicle kilometres travelled by longer semi-trailers (LST) in the trial by modelling the likely routes taken between known LST journey start points and end points.

1.3 Department for Transport (DfT) needed to validate the method of estimating the urban vs rural vehicle kilometres split using GPS tracking data from actual trailer routes, and then to use the tracked journeys to improve the current LST route modelling process. The LST routes can be re-modelled to create a richer route dataset so that further analysis can be carried out on validated LST routes.

1.4 The purpose of this workstream is to validate the LST route modelling process using actual semi-trailer routes, preferably for LSTs. Then, to use the validated modelling capability to model all of the ~58,000 unique operator start-to-end location pairs as likely routes taken by LSTs. The improved modelled routes can then be used to support several other elements of the LST evaluation work, to study the potential impact of introducing LSTs to England, Scotland and Wales.

2 METHODOLOGY

2.1 The E1 GPS Routing workstream answers some important questions through a series of steps, shown in Figure 1:

1. **Can we collect enough good quality GPS tracking data to represent routes taken by LSTs?**

   Step 1: Design the necessary processes to collect and analyse GPS data describing LST routes between the start and end postcodes that appear in the LST trial master database.

2. **Can we reliably convert GPS tracking data into the equivalent Ordnance Survey ITN[3] road links, so that we can analyse where LSTs are going?**

   Step 2: Convert GPS waypoints in routes into a set of contiguous road links, describing the route taken by a vehicle. Road links have additional information that can be used in analysis, such as urban vs rural distances travelled, and the proportion of routes on different road network levels (Strategic Road Network (SRN), Primary Route network (PRN), 'Other' (non-PRN) roads) and classes of road (motorway, dual, single, etc.).

3. **Can we evidence where GPS tracked journeys differ from our modelled LST routes?**

   Step 3a: Compare the real trailer routes with modelled routes between the same start and end locations, highlighting where modelled differ from real GPS routes.

4. **Can we improve our LST route modelling so that it most closely resembles real tracked GPS journeys at an aggregate level (for multiple journeys)?**

   Step 3b: Create an improved method of modelling LST routes and re-model routes for LST journeys made in 2017. Collect all semi-trailer routes into a master routing data set for analysis.

5. **Given a representative set of routes for the LSTs in this trial, what else can we learn about where they are being driven?**

   Step 4: Analyse all LST journey legs in England, Scotland and Wales for 2017 to determine updated and validated urban vs rural vehicle kilometre split data and other characteristics and impacts, as required.

2.2 This Project Note describes the result of answering these questions.

[3] Ordnance Survey Integrated Transport Network (ITN)

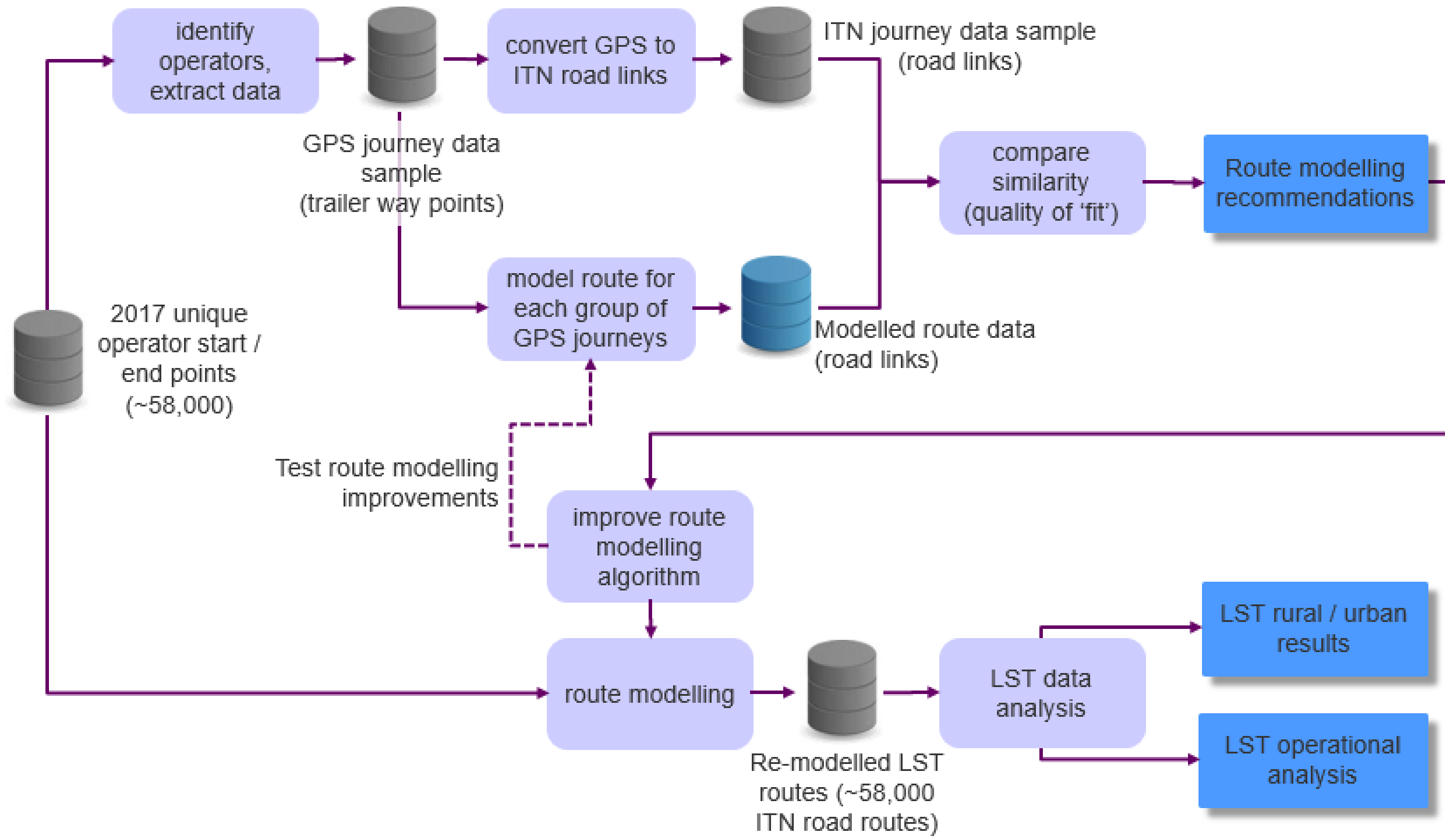

Figure 1: E1 GPS Routing

3 SOURCE DATA

3.1 The data sources we have used are listed below.

3.2 LST Trial data:

- 2017 LST journey leg data, including start and end postcodes for the majority of legs
- LST dimension data
- LST loading data

3.3 Ordnance Survey data:

- Integrated Transport Network (ITN)

3.4 GPS Data - Microlise telematics data for selected operators:

- Sample GPS data for articulated HGVs (a mix of LSTs and non-LSTs) on journeys made between October 2017 to February 2018 between start and end points taken from the known LST journey data.
- GPS journey data (operator, vehicle, start and end locations, date and time of journey, distance and time travelled)
- GPS location data for vehicles throughout their journeys (date and time, location, distance travelled since last location, wheel speed, bearing or direction, horizontal accuracy).

4 RESULTS 1: COLLECTING GPS DATA SAMPLE

4.1 From surveys of trial participants, we know that around half of the 1,800 LSTs on the road in 2016-2017 had some form of GPS tracking fitted, with perhaps another 20% being traceable through the tractor unit that pulled them. These trailers belong to about a third of the operators; the higher percentage GPS coverage is biased by the representation of larger operators with large LST fleets.

4.2 Collecting GPS data from multiple operators would be very challenging since it would be held in multiple proprietary systems, largely owned by third-party service providers, not by the operators themselves.

4.3 Our solution was to review the trial data to see which GPS service provider appeared to represent the largest number of trial participants. The most promising provider was Microlise (https://www.microlise.com).

4.4 After exploratory discussions to confirm the overlap between their client base and the trial participant list, Microlise agreed to join the project.

4.5 Although in almost all cases Microlise hosts the GPS data for their clients, making access to the raw information very easy, for each company dataset we wished to use, we needed to obtain written permission from the operators for Microlise to extract the data for the purpose of this research.

Identifying LST operators with GPS tracking data

4.6 From the operators taking part in the LST trial we identified 28 (17%) who were likely to be using Microlise telematics-based transport management solutions (either as direct customers or indirect customers through MAN and DAF manufacturers) to track their tractors or trailers. These operators ran 806 of the LSTs, 45% of the trial LST fleet.

4.7 We excluded a number of these operators where the Microlise tracking system is not fitted to their articulated fleet and so their data would not be relevant. We gradually narrowed down our focus to the operators with larger fleets of LSTs to maximise the efficiency of our process. This left us with nine operators of interest, whom we contacted to ask permission for use of the data.

4.8 We have extracted “90-day”[4] GPS data for 790,132 articulated HGV journeys for the eight operators who have given us permission to use their GPS data. Not all of these journeys are useful to us to test against modelled routes, due to the quality of the GPS data available and the nature of the journeys (some are very short), so we developed a process to extract the most useful journeys.

Identifying the most useful journeys in the GPS tracking data

4.9 Microlise collect GPS tracking data for their clients’ tractors, and in some cases directly from trailers. We have been able to identify the journeys taken by trailers

---

4 Microlise hold the most recent 90 days of days “live” on their systems, available for immediate access and analysis. Earlier data can be obtained but it must be loaded up from archive storage.

in the GPS tracking data using the following process that match routes we know that LSTs have also performed (from the main trial start/end postcode data).

4.10 First, we selected all GPS tracked journeys that start and end at the same locations as the LST trial journeys made in 2016 and 2017. We selected GPS journeys that start and end within a buffer zone of 1km from the LST start and end postcodes.

4.11 We then flagged 4.4% of the tracked journeys that were definitely made using LSTs, by cross referencing the trailer ID in the GPS data with the trailer IDs provided by operators. This does not always provide a perfect match, so to ensure we are collecting all available LST data we include all articulated HGV GPS journeys that start and end at the LST journey locations. A subset of these journeys are definitely made by LSTs, the remainder are likely to be a mix of LSTs and non-LSTs.

4.12 The process to extract GPS data has been designed using the most recent GPS journey data held by Microlise for selected operators, this covers the period from October 2017 to February 2018.

4.13 We have evidence from our earlier work that for most operators, the nature of the journeys made by LSTs as part of the trial has not changed significantly over the past two years, so it is reasonable to expect that the journeys in the extracted data are not very different from the LST journeys from 2016-2017 between the same start and end points.

4.14 We removed journeys from the extracted data that were either very short or contained unusual data anomalies. We removed journeys that:

- Contained discontinuities (either a break of over an hour between consecutive GPS waypoints, or a distance of over 10km)
- Had fewer than three GPS waypoints or were shorter than 3km

4.15 We then grouped journeys together that start or end at the same LST route locations, removing any groups of journeys with fewer than five journeys. If a group contained more than 52 journeys then some of the journeys were randomly removed to reduce the amount of processing required to a practical level.

4.16 This resulted in a set of 262,935 journeys, with 12,406 unique operator start and end locations. This means we have sampled GPS journey data for just over 22% of the 58,391 known LST routes (start and end points) in the 2017 trial dataset.

4.17 The result is an average of 21 journeys per group (i.e. the same start and end locations). This is a reasonable number to provide a good chance that we can do good quality statistical analysis of each group when comparing with a modelled route for this group of GPS tracked journeys.

GPS tracking data contents

4.18 Microlise has been able to extract the following GPS tracking information about LST journeys:

- Operator
- Trailer ID
- Trailer type (LST or not known)
- Start and end date and time
- Start and end locations (latitude and longitude)

- Distance travelled
- Duration of journey.

4.19 And the following information about each GPS waypoint in the journey

- Location (latitude and longitude)
- Accuracy of GPS location
- Date and time
- Distance since last waypoint
- Bearing (direction) of travel
- Type of event (e.g. ignition, moving, position update, distance log, stopped).

GPS data quality

Frequency and Spacing

4.20 The frequency of GPS waypoints is determined by the telematics service contracted by the operator, so it is normal to see some variation in the density of GPS waypoints.

4.21 GPS waypoints are typically separated by 2 minutes (although some are separated by 30 seconds, others by many minutes), with average distances between waypoints of less than 1 km (some are a few metres, others many km).

4.22 Journeys with very long separations between consecutive waypoints (10 km or 60 min) are flagged as having discontinuities in the data, and are discarded.

Route between waypoints

4.23 The variation in GPS quality raises an issue: the further away from each other the GPS waypoints are, the greater uncertainty there is in knowing the actual path taken by the vehicle between the GPS points. Therefore, obtaining GPS data does not guarantee total certainty of where LSTs (or any vehicle) have been.

Positional accuracy

4.24 There is also an issue that GPS accuracy is affected by how many GPS satellites can be seen from the location to provide accurate triangulation of the receiver's position. Generally, this is a problem in heavily built up areas, where buildings limit the number of visible satellites. Note that Microlise technology ensures that positional accuracy is usually within 50 metres.

4.25 The process used to map GPS waypoints to OS ITN road links also looks for the connecting ITN road links where consecutive GPS waypoints are on different ITN links. The shortest distance between ITN links is chosen. This ensures that a continuous set of road links is mapped from the GPS waypoints. Figure 2 provides an example of selecting a drivable route between consecutive GPS waypoints.

4.26 For the purposes of this project, the quality of the positional data available is more than sufficient to be confident of the routes taken by articulated HGVs.

Figure 2: Example of selecting a route between consecutive GPS waypoints

5 RESULTS 2: MAPPING GPS DATA TO ROAD LINKS

5.1 Risk Solutions and WSP have developed and tested a method to map GPS tracked journeys to the Ordnance Survey ITN representation of the road network.

5.2 This mapping is a critical part of the workstream activity, as knowing the ITN road links used by GPS tracked LSTs allows us to compare the real GPS routes with the routes that we are modelling for the same journeys using the ITN.

5.3 We can then improve the route modelling algorithm already developed by Risk Solutions and used for the 2016 LST Trial Annual Report, and model routes for all LST trial journeys from 2017, not just the ones for which we have obtained GPS data.

5.4 Once these LST journeys have been modelled using the ITN road links, we can use the additional information available in the ITN road link data to understand the journeys, considering issues such as:

- Proportion of distance travelled using the strategic road network, or primary route network, and distance not using these major networks
- Proportion of distance on different road classes (motorways, dual-carriageway A-roads, single carriageway A-Roads, and lower classes)
- Proportion of distance on roads in urban areas (as defined by the ONS).

5.5 We will also use the ITN link data to provide inputs to other workstreams:

- E2 - the modelling of emissions saved by LSTs
- E3 - the potential effects of LST availability on intermodal (road/rail) freight
- E4 - the ‘scaling up’ of the LST trial data to projections of the impact of any national roll-out of LSTs.

Example GPS tracked route

5.6 Figure 3 shows a single GPS tracked journey from Edinburgh to Newcastle Upon Tyne; the GPS waypoints are red dots.

Figure 3: GPS waypoints

5.7 Figure 4 shows a small selection of the GPS waypoints (red) and a collection of nearby ITN road links (grey).

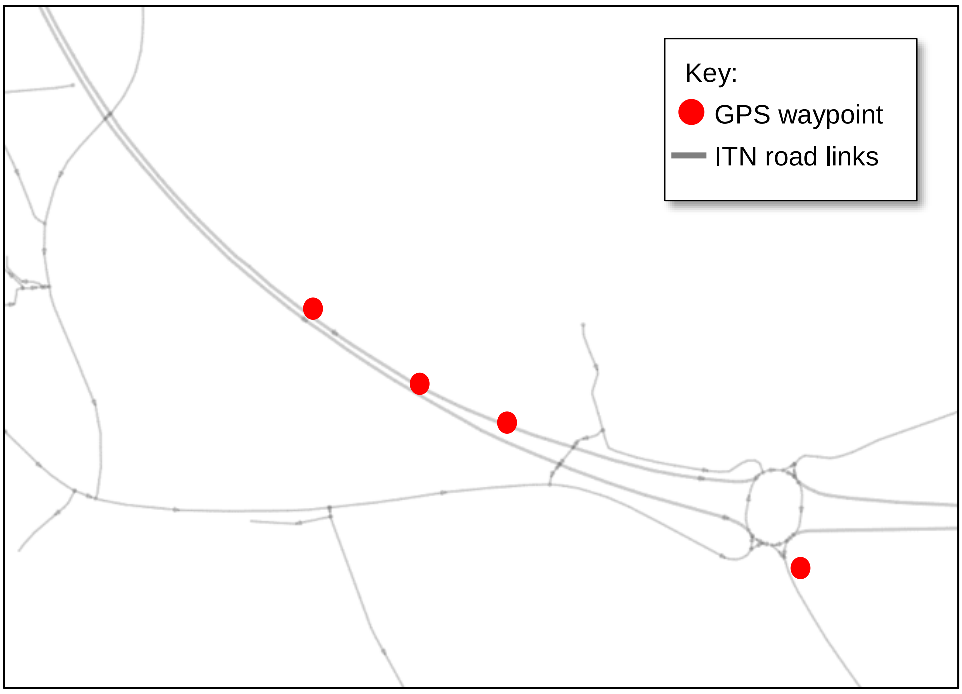

Figure 4: ITN road links and GPS waypoints

Mapping process

5.8 The mapping process involved selecting the best ITN road links connecting each pair of GPS waypoints, in each journey, from the start to the end of the journey.

Quality of mapping process

5.9 The journeys in their road link form are tested to make sure that the selected links match the GPS waypoints, and crucially, that the links are connected to each other in the ITN. The tests include:

- The road links are joined (to avoid picking links that are on the opposite sides of dual carriageways, or lead away from the journey up a slip road, or side road)
- The total distance represented by the selected road links is close to that recorded by the GPS waypoints (mapping the GPS data achieved a variance of -1 to +3%).
- The average and maximum time between GPS waypoints is not excessive (average from the GPS data was 2 minutes)
- The average and maximum distance between GPS Waypoints is not excessive (average from the GPS data was 750m)

Mapping complications

5.10 Several complications have been dealt with to map GPS waypoints to a contiguous selection of navigable ITN road links to form a journey route. Where the GPS waypoints are far apart, this includes selecting the best routes (ITN links) between those waypoints.

5.11 **Link length**: The shortest route between GPS points is chosen.

5.12 **One way**: Slip roads, roundabouts and dual carriageways are one-way, the shortest route between GPS waypoints needs to be able to be driven, so routes the wrong way down a one-way carriageway are not selected.

5.13 **Grade separation**: ITN road links are described as connected together in the data whenever a carriageway crosses another carriageway, including connecting carriageways that it is physically not possible to drive between, such as road over-bridges. When considering possible ways to connect links together, it is important to consider only the links that it is possible to drive between. Therefore, connections from over-bridges to underpasses must be ignored. This is obvious when looking at the road network, but in the ITN data there are connections between these links, these are described at different 'grades', or levels. Where links that are connected in the data are 'grade separated' a route is not permitted to use the link.

5.14 **ITN data size**: The ITN network is made up of many million links and nodes, split up into 100km OS map squares. This reduces the size of the files containing the link, node and shape data to a manageable level. LST routes typically cross many OS map squares, so at the edges of squares the routes need to be joined.

5.15 Figure 5 shows a number of these mapping challenges being dealt with when establishing the likely set of ITN links in the route between the two waypoints in the example given earlier.

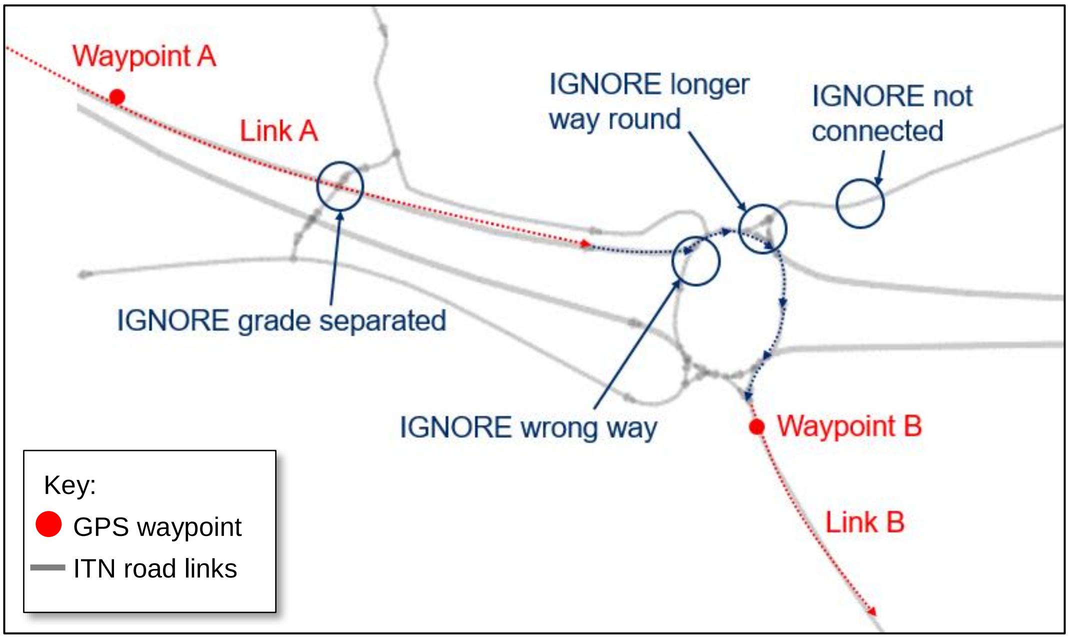

Figure 5: Selecting links between GPS waypoints

6 RESULTS 3: COMPARISON OF MODELLED ROUTES WITH GPS TRACKED JOURNEYS

6.1 An important feature of this workstream was the need to compare modelled routes with a set of GPS tracked journeys between the same start and end locations (equivalent journeys). To enable us to do this, we have represented both by a set of OS ITN road links, and we can directly compare these.

6.2 This allowed us to develop a scoring mechanism that could be used to:

- Provide an overall measure of the quality of the routing model.
- Identify those routes where the fit is poor so that we can examine them to understand where the routing algorithm is making different routing decisions to those implied by the GPS tracking. This helped us work out whether, and how, to adjust the routing algorithm to improve the quality of the route fit.

6.3 In this section we describe how journeys are modelled, how the scoring system works, how the scoring system was used to drive improvements to the modelling process, and finally the final scores achieved by the improved model.

Modelling routes for each group of GPS tracked journeys

6.4 The GPS tracked journeys have been grouped together to cluster journeys with similar start and end locations. We created a modelled route between each unique operator start and end group location, using the existing LST routing algorithm. This is the same routing algorithm we used to generate the route analysis results in the 2016 LST Trial Annual Report. It is described below.

Route modelling algorithm – routing decisions

6.5 The routing algorithm decides the best route in 2 stages; the **first** stage creates a summary routing network of road links, by creating a set of routes from each start and end location to the Primary Route Network (which includes the trunk road network) as follows:

1. the quickest routes using local roads from the **start** of the route to the closest junction to access the PRN,
2. the quickest routes using local roads from the **end** of the route to the closest junction to access the PRN,
3. if the start and end locations are reasonably close together then a direct route is created.

6.6 The summary network created consists of the road links that make up the PRN, and the local roads connecting every start and end location to each other if close, or to the trunk road networks.

6.7 The **second** stage then calculates the quickest route between each start and end location, only using the local roads found in the first stage, and the rest of the PRN (including the SRN).

Finding the quickest route

6.8 The choice of the quickest route is based on using an average speed for the vehicle (HGV) on each type of road being used. The speed is weighted according to:

- Vehicle type (HGV or car)
- Urban or rural area (Motorways are rural)
- Road type (Motorway, A, B, minor)
- Dual carriageway, single carriageway, slip road, roundabout)

6.9 The speed weighting represents how attractive the type of road is for a routing choice, e.g. urban areas are weighted so that they are less attractive than rural areas. This represents the presence of traffic, junctions and signals that will add delay to the journey.

6.10 The choice of available routes is also restricted by:

- The types of vehicles permitted onto a road link (one-way carriageways, mandatory turn, no entry, permitted only, etc.)
- Grade separation of road links (whether it is possible to drive between road links that are connected in the ITN data)
- Vehicle height restrictions on a road link.

Scoring modelled routes against GPS tracked journeys

Variation in the routes taken by GPS tracked journeys

6.11 The scoring system had to take into account the potential variation in routes taken by GPS tracked journeys between the same start and end location. Some journeys in a group use very similar routes (Figure 6). In other cases the routes taken between the same start and end point vary considerably (Figure 7).

6.12 In the maps that follow, green circles denote route starting locations, red circles denote end locations, the heat coloured road links (yellow to brown) denote the routes travelled by multiple journeys. Road links used in many journeys are coloured brown, road links used in just a few journeys are coloured yellow.

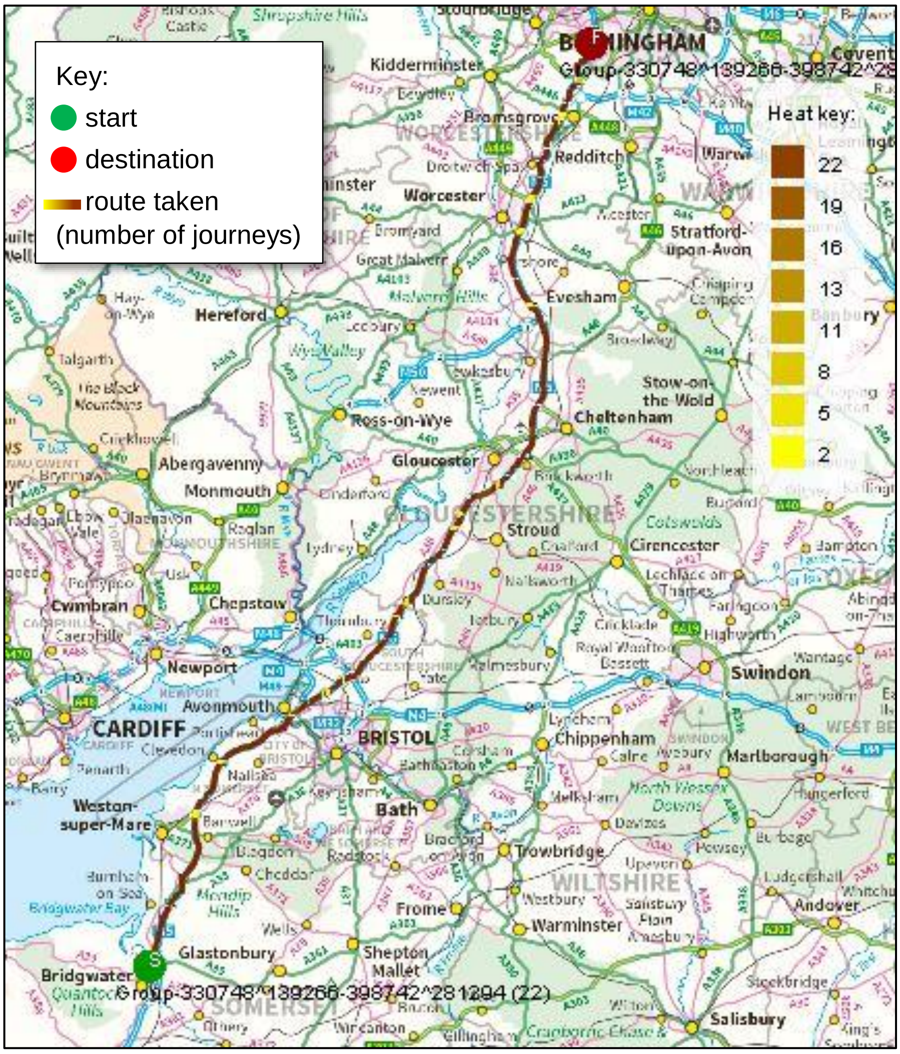

Figure 6: Group of GPS journeys using the same route

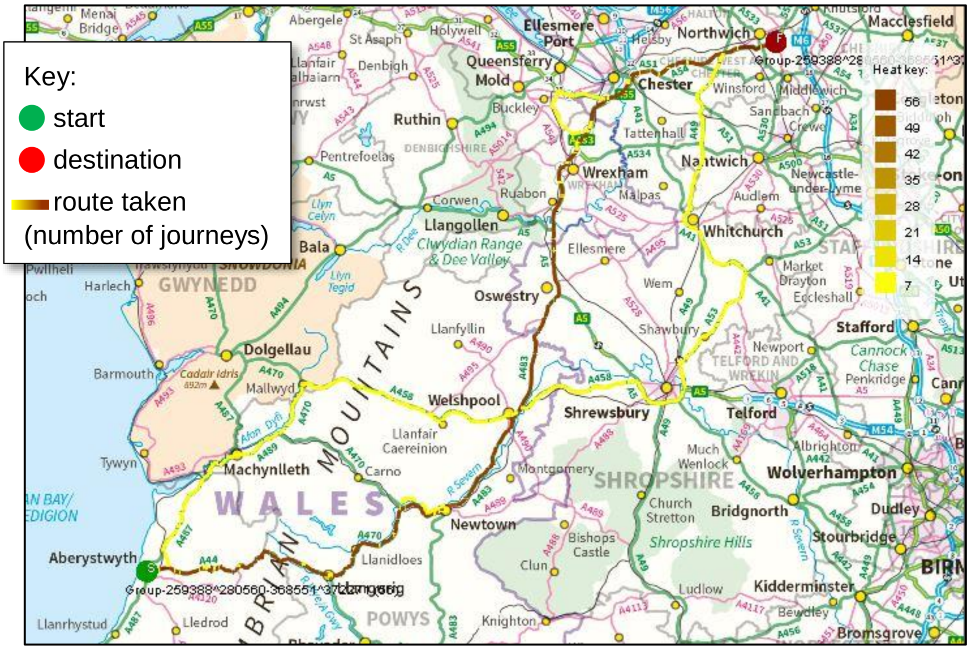

Figure 7: Group of GPS journeys with route variations

6.13 Figure 8 shows how the length of GPS tracked journeys can vary, when different routes are taken between the same start and end locations. The chart shows the distribution of journey lengths for a selection of ten start and end pairs, all with short journey lengths

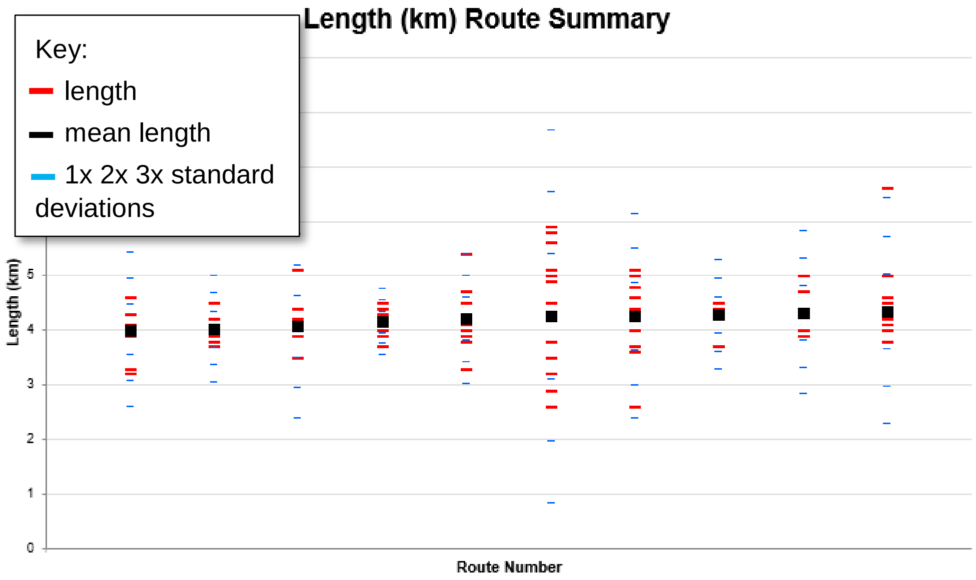

Figure 8: Example distribution of journey lengths for short routes

6.14 In the chart, the GPS tracked journeys for a single pair of start and end locations are shown as a vertical cluster of data points (a group of journeys). The red bars denote the length of each journey, the black square is the mean length for the group of journeys, and the blue bars show the position of 1x, 2x and 3x standard deviations from the mean.

6.15 The length of the journeys in this selection of groups are the shortest we considered (3km), to show the variation clearly. For longer journeys the distance variation is less obvious when charted like this.

Route fit scoring system

6.16 The route fit scoring system considers the distance travelled on roads in the GPS tracked journey that are not included in the modelled route. We calculate a ‘difference’ score for each GPS journey that represents the distance travelled on different roads from the modelled route. For example, a score of 0 equates to the modelled route being the same as the GPS tracked journey (a good route ‘fit’). A positive score is the additional distance travelled on the real journey, a negative score is less distance, compared with the modelled route.

6.17 There are multiple GPS journeys for each modelled route, so there is a set of GPS journey scores available for each modelled route. The modelled route fit score is the sum of individual journey route scores.

6.18 This route fit score is then normalised by the total distance travelled on all of the GPS journeys, to produce a final route score that is the proportion of modelled route travelled on dissimilar roads. This can be a positive or negative value.

$$\textit{Fit score} = \frac{\textit{Sum of the distance on GPS journeys travelled on different links}}{\textit{Sum of all GPS journey distances}}$$

6.19 However, we recognise that there are often 'equivalent' routes, where the mix of road types used on a journey are equivalent to those used in the modelled route. If a route is 'equivalent' then the fact that it is using a different set of road links can be discounted in the scoring system.

6.20 So, in addition to scoring the distance travelled on different road links, we also measured the distance travelled on different types of road links, of the following types:

- Urban dissimilarity (the km travelled on different road links, in urban areas)
- Trunk road network[5] dissimilarity (the km travelled on different road links, that are NOT part of the Trunk road network)
- Primary Route Network[6] road dissimilarity (the difference in km travelled on different road links, that are NOT part of the Primary Route Network)

6.21 This produces a route fit score (dissimilar distance per km) for urban, non-trunk, and non-PRN road links for each journey.

Route fit scoring examples for individual start and end locations

6.22 Figure 9 shows an unusual example of the variety of GPS journeys used for a 73 km modelled route, chosen because it has a significant variation in dissimilar length scores among the journeys. The map shows all of the routes taken by the 91 journeys for this particular start and end location; the road links are coloured according to the number of journeys they carried, using the heat scale shown.

Figure 9: Example of a group of journeys with significant routing variations

6.23 In this example, fewer than 6 journeys (the yellow routes) take a significantly different route from the modelled route (the thicker red opaque route). The majority of journeys take the modelled route (the brown route), and can be seen on the map inside the thicker red route where they overlap.

---

5 The trunk road network designates the motorways and major A roads managed by Highways England (the Strategic Road network) and the equivalent major roads in Wales and Scotland.

6 The Primary Route Network (PRN) designates roads between places of traffic importance across the UK, as identified by DfT. This includes motorways and some significant A roads between cities and major towns, (the A roads included are typically coloured green on road maps, and have green road signs). The PRN includes the Strategic Road Network

6.24 The route fit scores for every journey in a group can be displayed in a chart to show the variance in scores for each journey. The chart in Figure 10 shows a collection of grouped journeys with a similar length to the example shown in the map above (Figure 9). The group of journeys shown on this map is highlighted on the chart with an orange box.

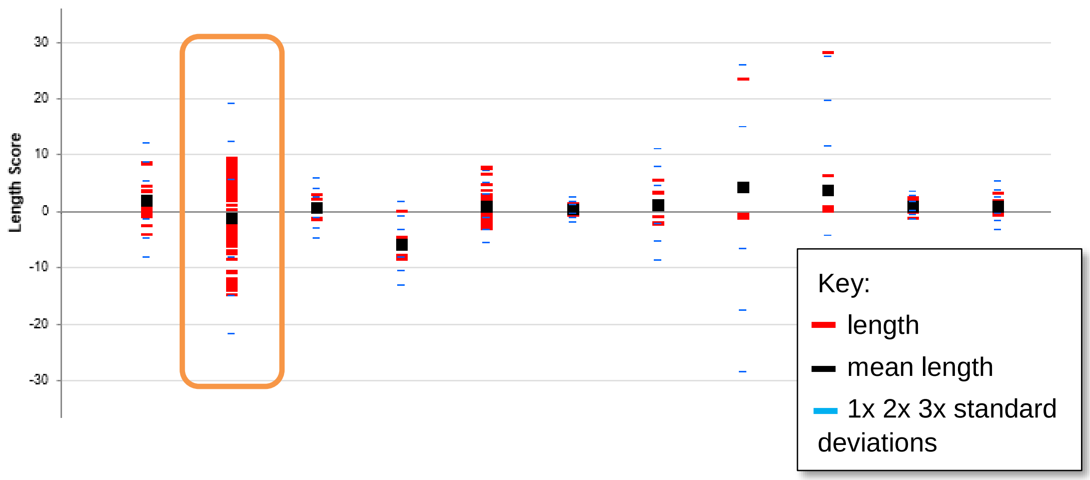

Figure 10: Example 'route fit' dissimilar length scores for groups of journeys around 70km in length

6.25 In these clustered bar charts, the route fit length scores for each journey in a group are shown as red lines, and are calculated as the distance (km) travelled in a journey on road links that are NOT in the modelled route. This is measuring dissimilar distance (km) compared with the modelled route.

6.26 The mean (black square) and standard deviation (blue lines denote 1x, 2x and 3x standard deviations) are calculated to record the average score and the spread of scores in the group of journeys.

6.27 Figure 11 shows charts for the urban, non-trunk and non-PRN scores for the same set of grouped journeys.

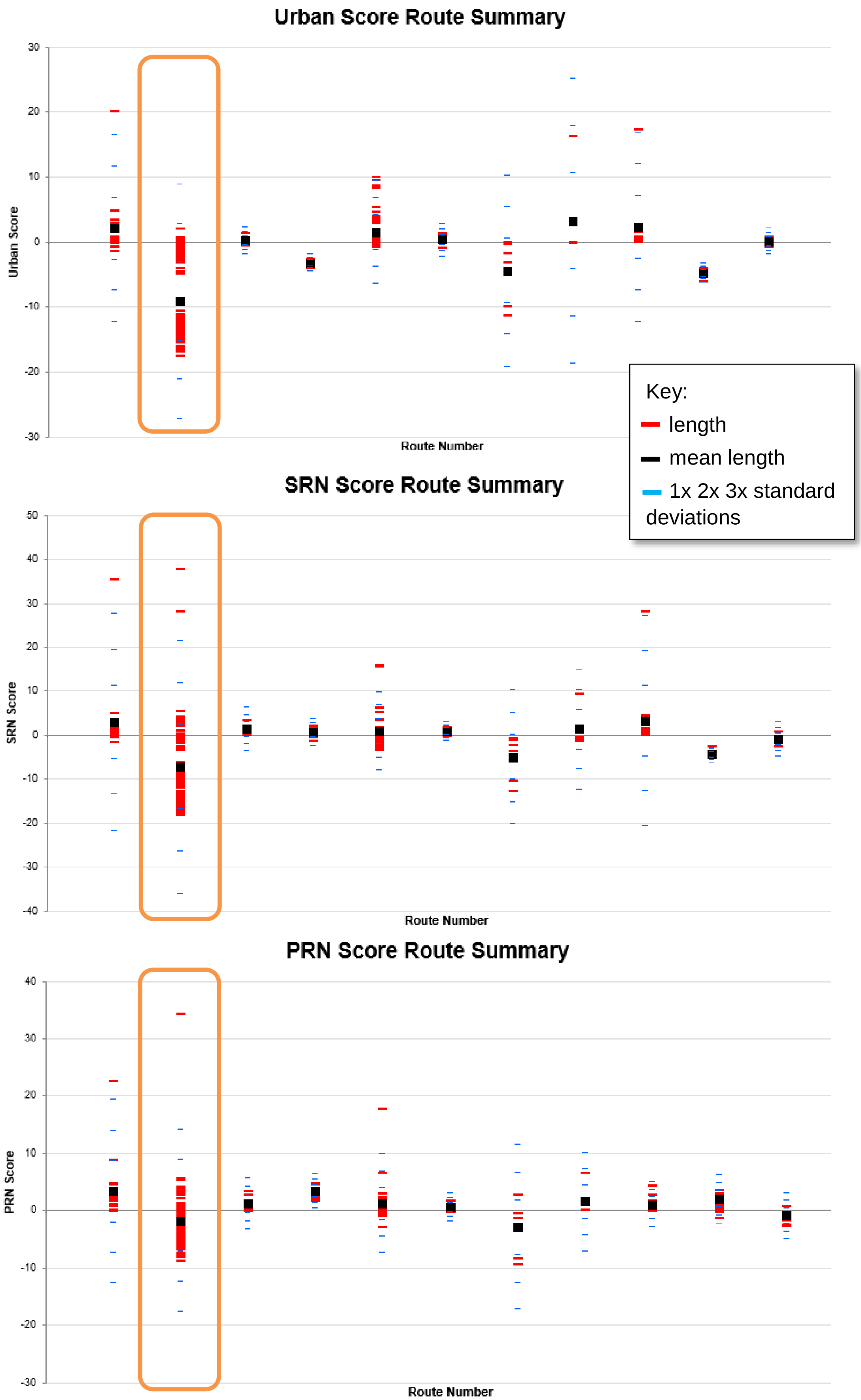

Figure 11: Example 'route fit' scores for urban, non-trunk and non-PRN dissimilar lengths

Route fit scoring across all sample start/end pairs

6.28 We calculated the mean normalised route fit score for each modelled route and considered the distribution of these average scores across all routes (Figure 12). The x-axis is the normalised length route fit score, the y-axis is the number of journeys in each band of scores.

6.29 The clustering of the majority of route length scores around the middle of the chart (0 route fit score), demonstrates that modelled routes are very similar in length to GPS tracked journeys.

6.30 There are many groups of GPS journeys for each operator, and the average lengths vary considerably (from 4km to 550km). This variation produces a wide range of route fit scores for each operator. To make comparisons between routes of different lengths, the scores for each route are also normalised by the route lengths (km), see Figure 14 later.

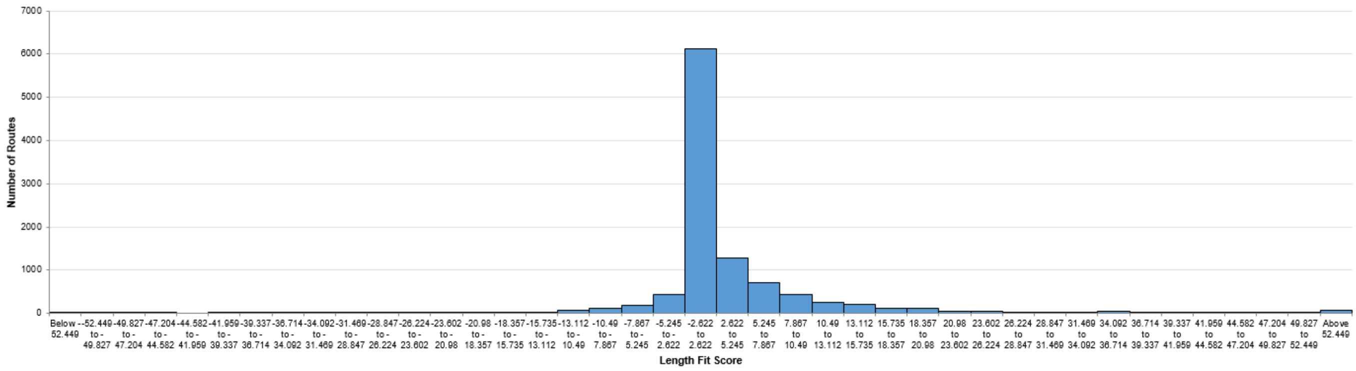

Figure 12: Distribution of length route fit scores

Routes excluded from analysis

6.31 Examining the GPS tracked journeys in detail reveals that some routes include some unusual journeys, where GPS data does not necessarily represent a journey from a single start to destination location.

6.32 For example, some GPS tracked journeys clearly include drop off locations en-route, sometimes well outside a sensible route from a start to an end location. It would not be appropriate to include these journeys when judging whether the route modelling algorithm fits GPS tracked journeys as they do not represent a single A>B route.

6.33 We developed a process to identify and exclude these routes, bearing in mind that some interpretation is necessary to decide whether a GPS tracked journey is valid or not. The process needed to ensure that we retain journeys where there is unusual mid journey behaviour due to road diversions, traffic congestion, or alternative valid routes.

6.34 The following types of GPS tracked journeys were identified and are not included in the route fit analysis. From the available 262,935 mapped GPS journeys, grouped into 12,406 Groups, we removed the following journeys from scoring (in this order):

- Journeys that do not have a modelled route (7.2% of all journeys)
- Journeys that do not have mapped ITN links (2.9% of all journeys), and journeys with a link length less than 95% of the straight line length between start and end locations (0.1% of all journeys)

- Journeys that contain loops or exaggerated distance before the start, beyond the end, or mid journey (3.5% of all journeys), where the length of these journeys is greater than the mean distance of all journeys in this group + two times the standard deviation of journeys
- Journeys from groups with 5 or fewer journeys (0.9% of all journeys)
- Journeys that had incorrect start or end locations (0.04% of all journeys); such that all of the journeys in the group were over 2x the distance of the modelled route.

6.35 This leaves a total of 224,326 GPS journeys, scored against 11,314 modelled routes (19% of the actual 58,391 unique operator start/end locations used by LSTs in 2017). Our GPS comparison sample is 25% of the size of the number of LST trial journeys made in 2017 (891,016).

6.36 Figure 13 shows some of the journeys that have **not** been included in the route scoring.

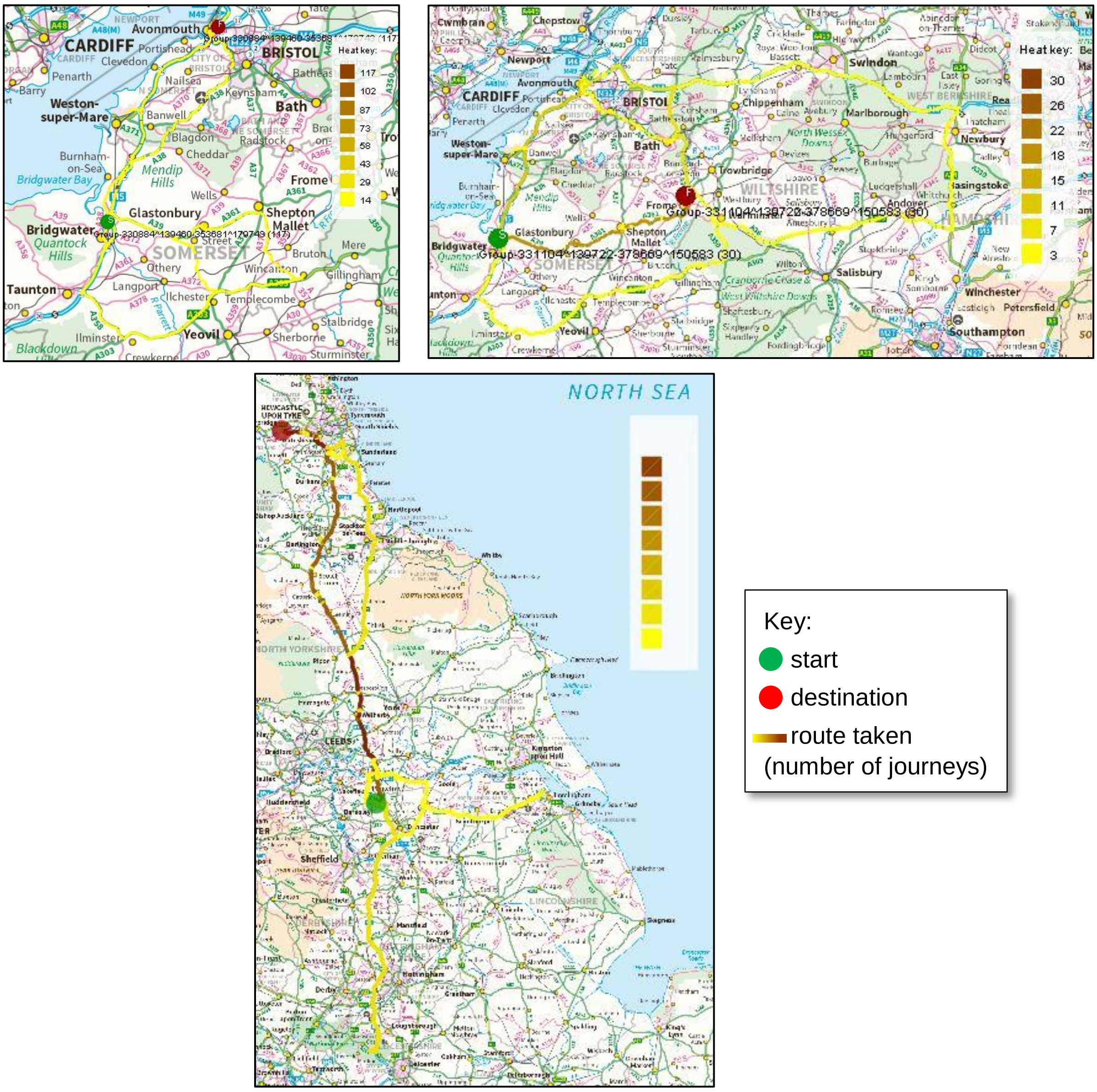

Figure 13: Examples of GPS tracked journeys not included in route fit scoring

Sampling routes to examine in more detail

6.37 We then selected routes with unusual scores for further analysis, to understand why the modelled route does not fit well with available GPS tracked journey data. This provided evidence for ways in which the routing algorithm could be improved.

6.38 There were a significant number of routes to examine, so we clustered them into types of routes from which we could sample. We clustered routes by:

- **Journey frequency**: to ensure we have a reasonable number of routes with over 20 journeys for statistical robustness, but also a few routes with fewer than 20 journeys as a comparison
- **Journey length**: to ensure we examine long routes (300km), medium length routes (75km) and short routes (3-7km)
- **Route variation**: to ensure we examine groups of journeys where there are many route variations, and groups of journeys with very little route variation.

6.39 However, we noticed that some routes contained subgrouping of journeys with similar lengths (19.3% of all journeys), statistically the journey data was multimodal, i.e. there was more than one 'dominant' route. These routes were not used to improve the routing algorithm.

Improving the route modelling

Route modelling improvements

6.40 Using the previous route modelling algorithm, the majority of modelled routes fit well with the GPS tracked journeys, evidence that the LST routing algorithm was working well.

Examples of good route fit (existing routing algorithm)

6.41 The three maps in Figure 18 show some examples of good route fit scores obtained from the existing routing algorithm for long-distance (296km, 52 journeys), medium (64km, 26 journeys) and short (4.4km, 10 journeys) routes.

6.42 In the figures that follow, the routes start at the green circles, and end at the red circles. The thick red line denotes the modelled route, the road links carrying most journeys are coloured brown. **For good route fit scores the red and brown routes are on top of each other**. The heat key shows the number of separate GPS tracked journeys made between the start and end locations.

Figure 14: Examples of modelled routes with good route fit scores

Improvements since analysis for the 2016 LST annual report

6.43 However, a number of modelled routes did not fit well with the GPS tracked journeys, these were of interest as they could potentially reveal ways in which the LST routing algorithm could be improved.

6.44 We used the following process to identify how the routing algorithm could be improved:

- We used the route scoring system described in this report to identify routes with particularly poor normalised scores (+/-0.20) for route length, urban, non-Trunk and non-PRN.

- We visually inspected the routes with poor scores (that were also of above average length), using maps showing the modelled route, and the equivalent GPS tracked journeys.
- Once a reasonable number of routes had been examined, we developed a set of hypotheses explaining why the modelled routes had poor scores.
- We modified the route modelling algorithm, implementing potential improvements.
- We remodelled the routes, using the improved routing algorithm, and visually inspected the new routes, to see whether the changes had delivered the expected results.
- The routes were scored again, and the route fit scores compared with those resulting from the previous routing algorithm. If the scores were better, the change to the routing algorithm was kept.

6.45 This process was repeated to find the combination of routing improvements with an improvement in route fits scores for the cases found, and overall across all routes.

6.46 Examining a group of the poorest route fit scores (length, urban, non-Trunk and non-PRN) revealed some route modelling decisions that are significantly different from the majority of GPS tracked journeys, specifically:

- Modelled routes were choosing motorways when the GPS journeys mainly took A roads on the PRN
- Modelled routes were making a different choice of entry or exit junction on the PRN from that chosen for GPS journeys.

6.47 The routing algorithm was adjusted to make better decisions in these cases.

Modelled routes choosing motorways rather than PRN A roads

6.48 The left-hand map in Figure 15 shows the original routing algorithm (thick red line) choosing the M4 to Reading and the A33 over the more direct route taken by the GPS tracked journeys (brown line), via Newbury on the A339.

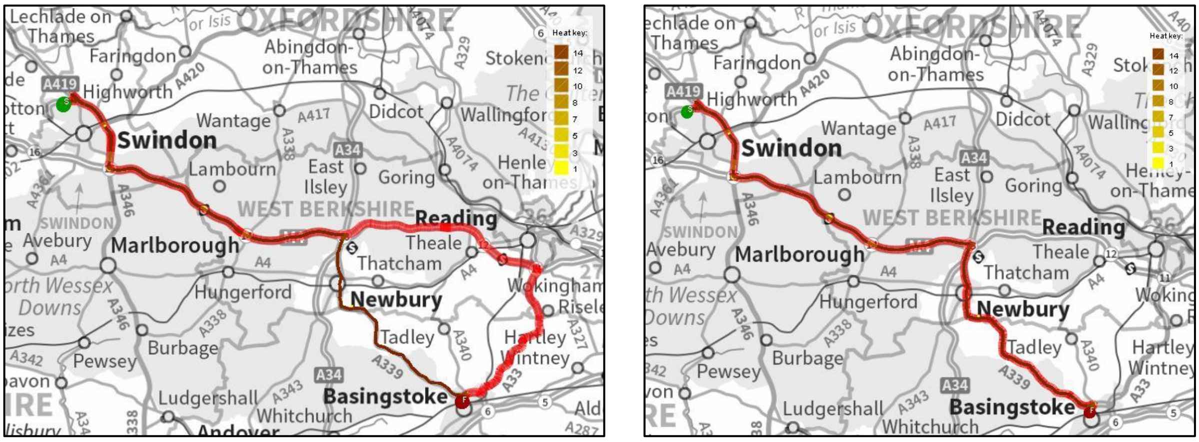

Figure 15: Making A roads more attractive

6.49 The solution in this case was to make single carriageway A roads more attractive to the routing algorithm, but remaining less attractive than dual carriageway, A roads and motorways. The right-hand map in Figure 15 shows the result of this improvement.

Modelled routes making a poor choice of entry or exit junction to the PRN

6.50 The left-hand map in Figure 16 shows the routing algorithm choosing the longer route via the A14 and M1(M) to Stevenage, instead of going to Luton via the M1, the route taken by the majority of the GPS tracked journeys.

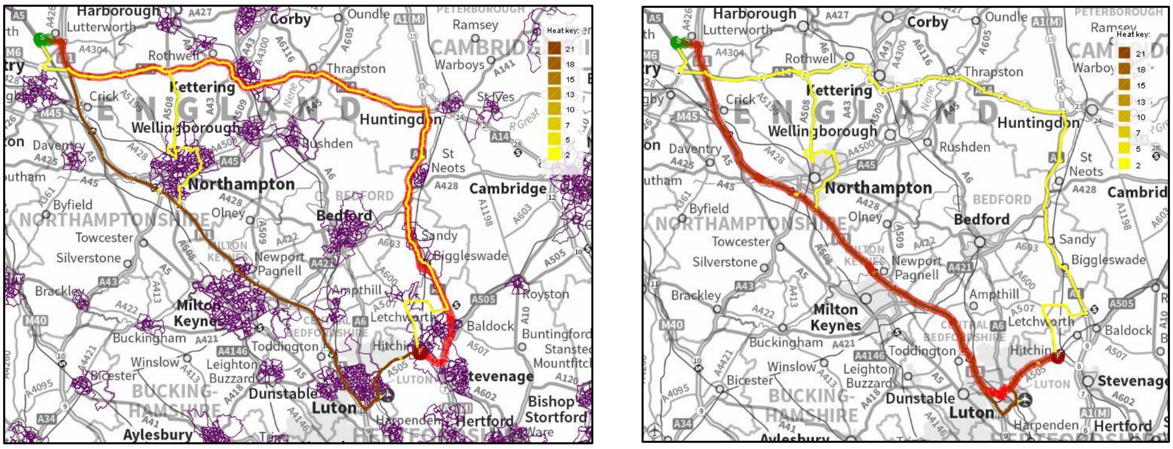

Figure 16: Choosing a better junction to join / leave the PRN

6.51 In this case the urban scores for the modelled route were very poor, indicating that the length of urban roads in the modelled route is different from that in the GPS tracked journeys. This is because the majority of GPS tracked journeys use the M1 (rural), and the modelled route uses the A14, which has some length in urban areas because it passes near towns. Figure 17 shows the portion of journeys on urban roads as orange.

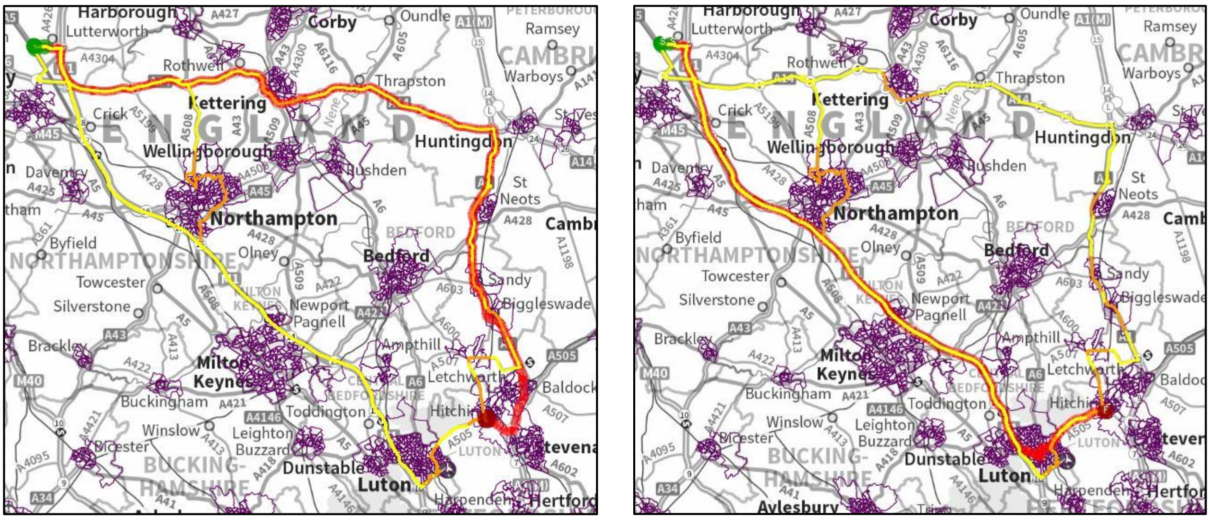

Figure 17: Improving the urban route fit score

6.52 The solution was to improve the first stage of routing by extending the distance searched for local roads to connect with trunk network junctions. The search in this case needs to extend from the end of the route in Hitchin to the M1 on the other side of Luton. Making these local roads available to the routing algorithm in the second stage of routing ensured that the M1 was considered as a viable route. The right-hand map in Figure 17 shows the result of applying this improvement.

Results of route modelling improvements

6.53 Applying the routing improvements improved the scores of many routes, reducing the average route score for length, urban and non-Trunk, and non-PRN scores. An improved score is one that is closer to zero and may be either positive or negative.

Difference in normalised scores	Previous route modelling	Improved route modelling
Route length	+7.5%	+1.7%
Urban route length	+3.6%	-0.7%
Non-Trunk route length	-2.7%	-0.2%
Non-PRN route length	-0.9%	1.5%

Table 1: Results of route algorithm improvements

Route fit score results

6.54 Following these improvements, we can now report the quality of our routing algorithm using route fit scores, from four perspectives:

- **Total distance score**: giving an overall measure of the quality of our routing algorithm.
- **Urban score:** to evidence that our routing algorithm results in a realistic distance in urban areas.
- Network Level Performance:
  - **Non-trunk score:** to evidence that our routing algorithm results in a realistic indication of the extent of LST operations off trunk roads (roads not maintained by Highways England or the equivalent in Wales and Scotland)
  - **Non-PRN score:** to evidence that our routing algorithm results in a realistic indication of the extent of operations on ‘other’ (non-PRN) roads.

6.55 We considered the quality of fit from each perspective to assess the robustness of the routing model.

Route scoring by total distance

6.56 Figure 18 shows the distribution of the ‘route fit’ normalised length scores for all 9,741 modelled routes, using this improved route modelling algorithm. Each column is a bin covering a route fit score of 0.05, this equates to 5% of each route length[7].

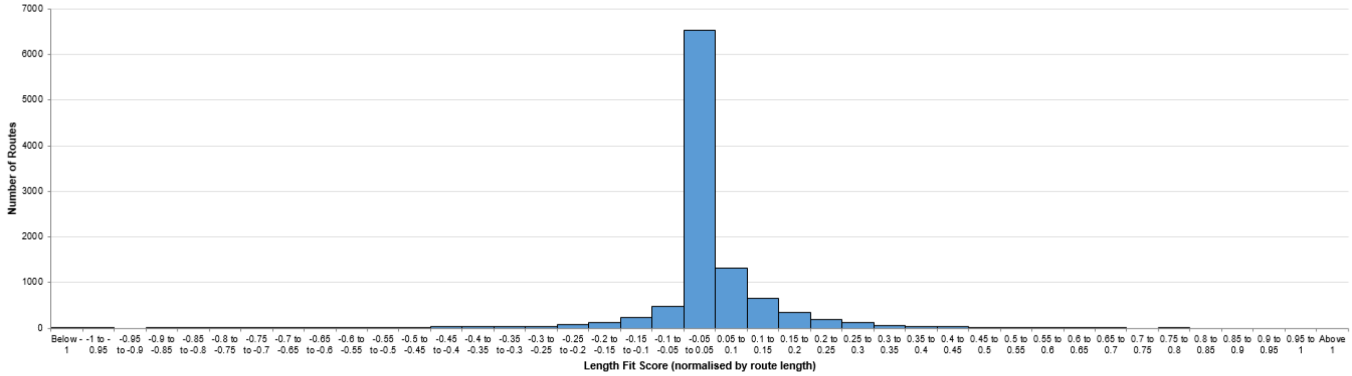

Figure 18: Distribution of normalised length route fit scores

6.57 Here, 17% of GPS routes are identical to their modelled routes, and two thirds of the routes have a route fit score close to zero (within 5% of zero), i.e. the modelled routes follow principally the same routes as the majority of GPS tracked journeys, with some small variation.

6.58 There are more routes where the GPS route is longer than the modelled route, with positive scores than negative ones; the distribution is leaning towards a small positive score. The average route score is +1.7%, so the GPS routes are on average 1.7% longer than modelled routes. Our initial hypothesis is that some articulated HGV journeys will avoid congestion by taking a diversion, resulting in longer distances travelled.

Route scoring by urban distance

6.59 Figure 19 shows the distribution of the 'route fit' normalised **urban** scores for all scored routes. This type of route fit score measures the distance travelled on urban roads in GPS tracked journeys that are **not** used in the modelled route (a normalised dissimilar distance score). For these metrics a positive score means that the GPS journey has travelled further on different roads to the modelled route, a negative score means the GPS journey has travelled a shorter distance.

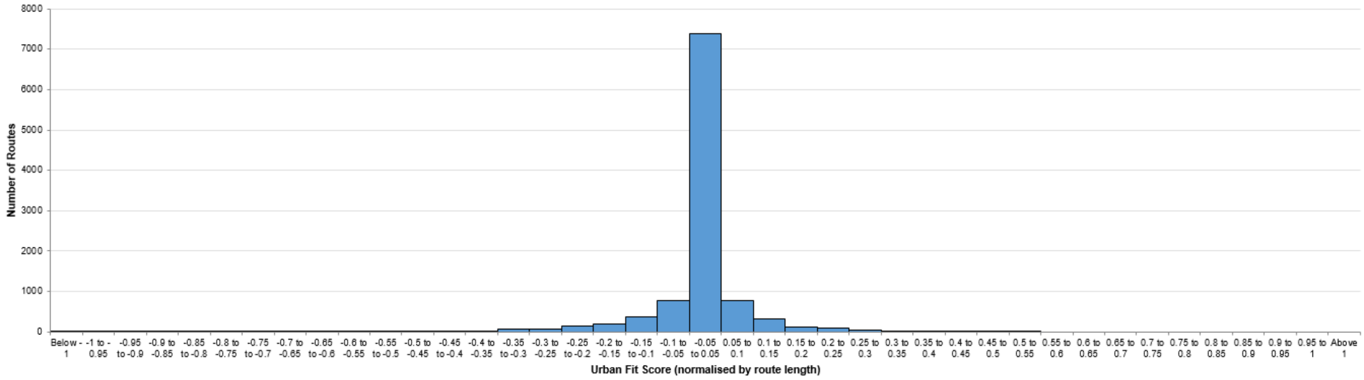

Figure 19: Normalised 'route fit' urban scores for all routes

[7] Note: the central column in each chart covers -5% to +5% of route length

6.60 24% of routes have an urban route fit score of 0. 71% of modelled routes have a score close to zero (within +/- 5% of route length). This shows that for the majority of modelled routes, the proportion of the equivalent GPS tracked journey length on urban roads is almost the same as the proportion of the modelled route length on urban roads (the distance on dissimilar urban roads is close to zero).

6.61 There are slightly more routes where the GPS route is shorter than the modelled route, but on average, the proportion of the route length travelled on urban roads for modelled journeys is 0.7% more than the GPS routes (the average route score is +0.7%).

Route scoring by non-Trunk distance

6.62 Figure 20 shows the distribution of the route fit normalised **non-Trunk** scores for all scored routes. This score is measuring the distance travelled on dissimilar roads that are not part of the Trunk road network (the SRN). A positive score means that the GPS journey has travelled further on different roads (that are not part of the trunk road network) than the modelled route, a negative score means that the GPS journey has travelled for a shorter distance on different roads (that are not part of the trunk road network).

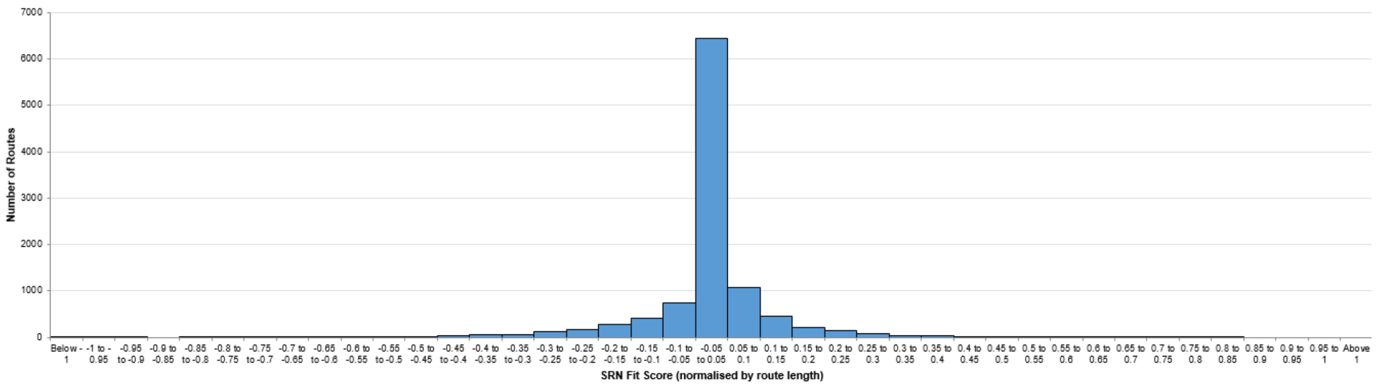

Figure 20: Normalised 'route fit' non-trunk scores for all routes

6.63 63% of modelled routes have a score close to zero (within +/-5%), and the number of routes with positive and negative scores are almost equal.

6.64 There are more routes with longer distances travelled for GPS routes than for the modelled routes, but on average, the proportion of the route length travelled on non-trunk roads for GPS journeys is 0.2% less than for the modelled routes. This is a small difference. The difference probably arises because the route modelling algorithm does not take account of any congestion that might be present on a particular day or at a particular time. Once modelled routes get to trunk roads they remain on these roads for as long as possible, not taking any diversions due to congestion.

Route scoring by 'non-PRN' distance

6.65 Figure 21 shows the distribution of the route fit normalised **non-PRN** scores for all scored routes. This route fit score measures the distance travelled on dissimilar roads that are not part of the Primary Route Network. This road network consists of roads that are part of the Strategic Road Network (managed by Highways England) and additional major A roads on primary routes between primary locations ('Primary Routes' as identified in the Ordnance Survey ITN). A positive score means that the GPS journey has travelled further on different roads (that are not part of the PRN) than the modelled route, a negative score means that the GPS journey has travelled for a shorter distance on different roads (that are not part of the PRN

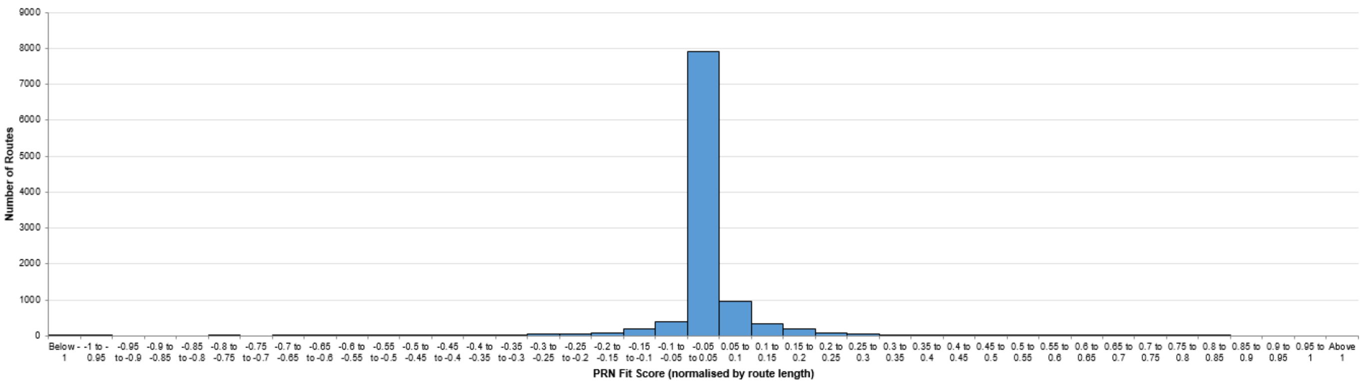

Figure 21: Normalised 'route fit' non-PRN road scores for all routes

6.66 18% of modelled routes have a score of 0, using an identical length of non-PRN roads. 76% of modelled routes have a score close to zero (within +/-5%)

6.67 There are more routes with a positive score (indicating that on average, the proportion of the GPS journey length travelled on dissimilar non-PRN roads is 1.5% more than the modelled routes; i.e. This is evidence that slightly fewer of the modelled routes include roads that are not part of the PRN compared with the GPS tracked journeys.

Route fit scoring conclusions

6.68 **Conclusion 1:** The tight symmetrical distribution of the route fit scores around the zero value (dissimilar length, urban, non-Trunk and non-PRN route fit scores) demonstrates that the routing algorithm is modelling routes for LSTs that on average are equivalent to GPS tracked articulated HGV journeys.

6.69 **Conclusion 2**: There are small variations in the average modelled route score when compared with the GPS tracked journeys:

- Modelled routes on average have 0.7% additional route length in urban areas
- Modelled routes on average have 0.2% additional route length on roads that are NOT part of the Trunk road network
- Modelled routes on average have 1.5% less route length on ‘Other’ roads (not part of the Primary Route Network).

We conclude from this scoring that the modelled routes are a good reflection of the types of roads taken by HGVs, both within urban and non urban areas, and across the trunk network PRN and ‘Other’ roads. And that even though some HGV journeys from similar start and end points take different routes, that the modelled routes are a good approximation or real journeys when considering the types of roads taken.

7 RESULTS 4: LST OPERATIONAL ANALYSIS

7.1 We conclude this workstream project note with an explanation of our operational analysis of LST trial journeys in 2017.

7.2 These results have been taken forward into the main LST Trial 2017 Annual Report.

Objective

7.3 The objective of the operational analysis phase of the workstream was to: **Analyse all LST journey legs in England, Scotland and Wales to determine updated and validated urban / rural vehicle kilometre split data and other characteristics and impacts, as required.**

7.4 The following objective was agreed with the DfT project sponsor; to estimate the vehicle kilometres travelled by LSTs during the 2017 LST trial:

- Urban versus rural distances[8]
- On Major (Motorways, A Roads) and Minor roads[9]
- On trunk roads (Strategic Road Network managed by Highways England) and principal roads[9] (managed by Local Authorities)
- By road class (Motorway, A road, B road, C road, unclassified road)
- By the alternative SRN, PRN, 'Other' classification of roads.

7.5 This breakdown mirrors that used in DfT's own summary of HGV traffic flows published in Road Traffic Stats table TRA3105.

7.6 The estimates are based on using the Risk Solutions' routing model to estimate the likely road routes taken by LSTs between the recorded start and end locations of each journey leg made in 2017. The routing model has been improved and tested using real GPS journey data from HGVs as described above as described in Section 6 above.

7.7 These estimates are required for a number of areas of the overall LST Trial analysis including

- Safety – Analysis of injury incidents at the sub-national level. As in 2016 we want to look closely at injury incident rates in urban areas compared with other articulated HGVs, but now also to look at injury rates by road class.
- Emissions – The emissions model looks at emissions saved by reduced journeys by road class, link length, speed, taking data from the routing model. The routing model also returns the proportion of each route that passes through an agreed set of defined air quality monitoring areas.
- Intermodal traffic – The intermodal work has made use of the route start/end locations in reviewing LST routes that mirror intermodal pathways.

[8] As defined by ONS urban areas 2012

[9] As defined by DfT road traffic statistics (www.gov.uk/government/organisations/department-for-transport/series/road-traffic-statistics), TRA3105

7.8 It is worth noting the following aspects of the routing model developed to address this objective:

- The routing model is designed to give a good representation of 'reasonable' routes; it is not intended to provide a fully functional routing tool that could be used for real world operations.
- The model is not designed, at this stage, to produce standard GIS layer files of the results, although this would be a natural extension of the work to be considered.

Analysis method

Source data used in analysis

7.9 The **LST** trial journey leg data is used to identify the numbers of journeys made by LSTs in 2017, and the start to end locations of each journey.

7.10 The routing model creates the likely routes taken by the trial LSTs, representing these as ITN road links. We validated the model by comparing modelled routes with real GPS tracking data from a significant sample of HGVs travelling between the start and end locations used by LSTs in the trial.

7.11 Urban shape files are used to identify the proportion of road links that are in urban and rural areas:

- England and Wales: 2011 census urban settlements data, provided by DfT
- Scotland: Scottish Government Urban/Rural Classification 2013/2014

7.12 Ordnance Survey Integrated Transport Network (ITN) road link meta data is used to identify road links used by LSTs with their road type and road class. This was provided by DfT.

Definition of Urban for LST work

7.13 For the analysis of LST injury incidents and emissions work we have used a slightly amended definition of Urban versus Rural, since our main interest in urban operations relates to the increased exposure to these trailers of other road users, especially pedestrians and those classed as vulnerable, primarily during turning movements.

7.14 We therefore split the injury and traffic data as follows:

- **Urban** (urban roads EXCLUDING MOTORWAYS)
- **Rural/M-Way** (rural roads and ALL MOTORWAYS)

7.15 In earlier reports, this simple split was used for further calculations. In the work completed in 2017 the data is further split by road type and in doing so the Motorways are split off into their own group:

- Motorways
- Major Roads (Trunk / PRN – A Roads)
  - Rural
  - Urban
- Minor Roads
  - Rural
  - Urban.

Route Modelling

7.16 The routing calculations were performed using Risk Solutions own MapSnap™ platform. This includes tools to characterise a road network for vehicle routing purposes and identify road links within urban areas. The tool can then process a list of routes (with start and end points), and for each, create the details of a route that is likely to be taken by a vehicle, with options to select the type of routing required, e.g. the quickest, or shortest route.

Creating relevant road routing information

7.17 A road topology of the **Integrated Transport Network (ITN)** was created from source XML data provided by Ordnance Survey[10]. The road topology consists of connected road links and road nodes. Additional road information was extracted from the source XML data including:

- road class (e.g. dual carriageway, single carriageway, roundabout, slip road, etc.),
- road number and road name,
- vehicle routing restrictions (one-way, restricted height, etc.),
- grade separation to identify whether road links are connected at the same level (to identify over bridges),
- whether the road is part of the Strategic Road Network[11] (SRN) and Primary Route Network[12] (PRN).

7.18 Additional information for each link was calculated to aid routing decisions, including:

- link direction (bearing),
- link length, and
- the proportion of link length inside an urban area.

Creating routes

7.19 We have taken the master LST trial leg data for 2017 and extracted the start and end postcodes of each leg. We have defined an LST route as **SID-StartPostcode-EndPostcode**, where the SID is the trial three character 'Survey ID' for a company, used to anonymise all trial data. For each postcode, our system obtains a latitude and longitude for the centroid of the postcode using a google maps interface.

7.20 The nearest part of the local road network (road link) to the route start and route end is selected, then the quickest route is found between these two road links.

7.21 Road links must meet a range of criteria to be considered, including,

- vehicle height restrictions (width and weight restrictions can be added if required),

---

10 During the lifetime of the project, OS launched their new 'OS MasterMap Highways Network' family of products, developed with and partly funded by DfT. Had this new version of the OS data been available earlier it would have been used and might have reduced the effort required to match together several datasets.

11 Strategic Road Network: the principal trunk road network consisting of motorways and some dual and single carriageway A roads in England, and the equivalent trunk road networks in Scotland and Wales.

12 Primary Route Network: the principal A roads between urban centres, signified by green road side signs.

- obeying the highway code (e.g. one-way carriageways, direction on roundabouts),
- grade separation between roads that cross each other but are not joined (e.g. over bridges).

7.22 It is important to note that the routes are expressed and stored as a series of connected OS ITN link references, rather than as a 'string of pearls' (as would be given from GPS data) or a GIS layer. It is this connection to the OS ITN that allows the modelling to access information about the route from any ITN data.

Calculating route information and urban or road network/class proportions in the actual LST operations

7.23 All of the route information, and the road links used in routes, are exported into the Master DataSet, along with the road class, road type and proportion of road length in urban areas.

7.24 We calculate the proportion of a route that travels through urban areas by adding together the urban length of each road link on the route and comparing this with the total route length. As part of the data importing process, the MapSnap™ tool calculates the proportion of each road link shape that is inside an urban area. As noted above, Motorway road links are all grouped with the non-urban data.

7.25 The vehicle km driven by LSTs is calculated by multiplying the route length by the number of journey legs reported by operators. The proportion of vehicle km in urban areas is also calculated.

7.26 We used the MasterDataSet to analyse operational patterns by merging the routing and road link data for each routeID with the number of actual journey legs performed on each route. This is an essential step since it is the proportion of actual LST operations on each road type than are of interest.

LST operational patterns

7.27 Table 2 summarises the types of roads and trunk networks used on the 102 million vehicle km represented by modelled routes for trial LSTs in 2017. This distance consists of 56,377 unique routes and 828,778 individual journeys.

	Non-urban	% of total vkm	Urban	% of total vkm	Total vkm	% of total vkm
Motorway	**63,313,108**	**62.0%**	**0**	**0.0%**	**63,313,108**	**62.0%**
Major (A Road)	**24,679,735**	**24.2%**	**12,140,096**	**11.9%**	**36,819,831**	**36.0%**
Trunk (SRN)	**18,232,075**	**17.8%**	**5,685,598**	**5.6%**	**23,917,673**	**23.4%**
Principal	**6,447,660**	**6.3%**	**6,454,498**	**6.3%**	**12,902,158**	**12.6%**
PRN	4,987,807	4.9%	4,751,508	4.6%	9,739,315	9.5%
Other	1,459,853	1.4%	1,702,990	1.7%	3,162,843	3.1%
Minor	**798,842**	**0.8%**	**1,261,164**	**1.2%**	**2,060,006**	**2.0%**
PRN[13]	496	0.0%	41	0.0%	538	0.0%
Other	793,346	0.8%	1,261,123	1.2%	2,059,469	2.0%
Grand Total	**88,791,685**	**86.9%**	**13,401,260**	**13.1%**	**102,192,945**	**100.0%**

13 The PRN flags in the OS ITN includes a short length of roads that are neither A roads or Motorways.

Table 2: Trial LST vehicle km by road type

7.28 The proportion of total vehicle km for trial LSTs by road type is similar to the DfT published statistics for England and Wales from 2016[14]:

	LST trial	England and Wales
Motorways	62.0%	58.8%
A Roads	36.0%	39.3%
Other roads	2.0%	1.9%

Table 3: Proportions of total articulated HGV vehicle km by road type

7.29 The proportions of vehicle km for trial LSTs travelling across the trunk road networks are also similar to the DfT statistics for England and Wales:

	LST trial	England and Wales
Trunk (SRN Motorways and A roads)	85.4%	82.9%
Principal	12.6%	15.2%
Minor	2.0%	1.9%

Table 4: Proportions of total articulated HGV vehicle km by network type

7.30 **CONCLUSION**: the trial LSTs are using a similar mix of roads to the national HGV articulated fleet.

Urban vehicle km

7.31 The vehicle km for trial LSTs in urban areas is 13.1% of total vehicle km travelled, which is considerably higher than the 5.9% figure for the national HGV articulated fleet in England and Wales. A significant proportion of this was on the SRN (5.6% of total vehicle km).

7.32 We explored why there was such a large proportion of trial LST urban vehicle km on the SRN, and discovered:

1. A small number of long routes are contributing significantly to the analysis of vehicle km, because many thousands of journeys use those routes.
2. These routes have a high proportion of their length in urban areas, of which most is on the SRN.

7.33 Figure 22 shows the distribution of the urban and non-urban vehicle km for trial LSTs on the left-hand side, and how this is apportioned across the road classes in the centre (Motorway, A Road and Minor road classes). On the right, it shows how the distribution of vehicle km on each road network is apportioned by network types (SRN, PRN, other roads). The thickness of the bands across the diagram are proportional to the LST vehicle km on each type of road.

[14] TRA 3105

7.34 Of the 13.1% of the trial LST vehicle km classified as urban, 10.2% is either on the SRN (5.6%) or the PRN (4.6%). Only 2.9% is on other roads, A roads, B roads, C roads and unclassified roads, that are not part of the SRN or PRN.

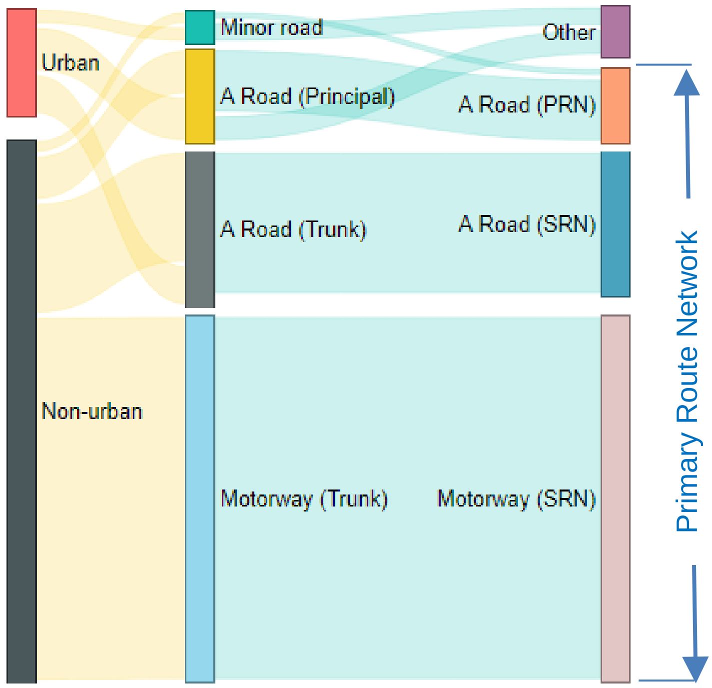

Figure 22: Trial LST vehicle km by road network

7.35 An analysis of the trial LST routes and journeys in 2017 shows that there are a small number of routes that significantly contribute to the vehicle km analysis, and that this explains the greater proportion of distance travelled in urban areas compared with the DfT statistics for England and Wales. Figure 23 shows the contribution every route in the 2017 trial data is making to the total LST vehicle km (in blue). A small number of routes at the right-hand side contribute significantly more than the other routes. The orange line shows the cumulative vehicle km as a proportion of total trial LST vehicle km. This line sharply approaches 13.1% close to the right-hand side of the chart.

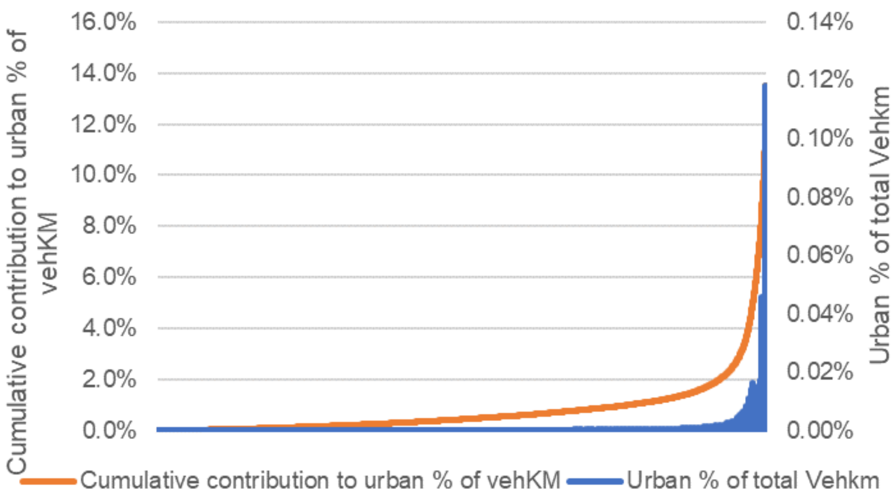

Figure 23: Urban vehicle km for LST trial routes

7.36 Figure 24 considers the 500 LST routes with the largest vehicle km (the right-hand side of the chart in Figure 23). This group of routes accounts for only 0.9% of the routes used in 2017 by trial LSTs, but contribute over half (53%) of the total vehicle km. Interestingly these routes also contribute a large part of the total urban vehicle km (42% of total urban distance).

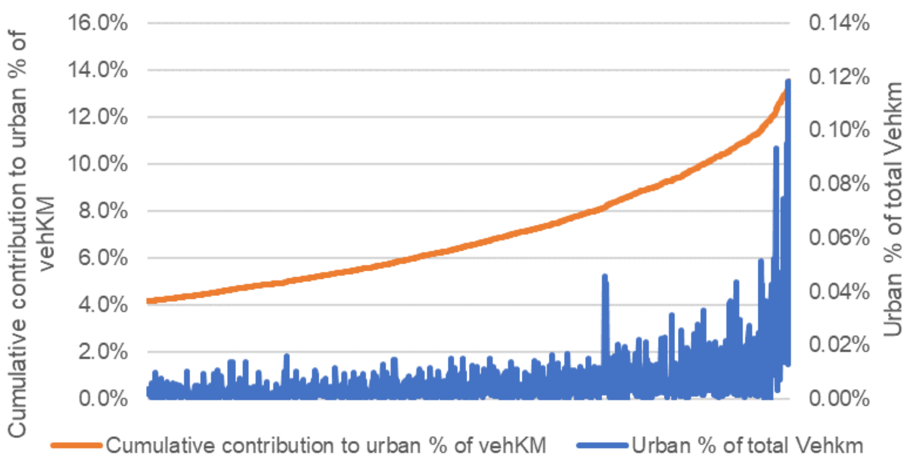

Figure 24: Top 500 contributing routes to urban vehicle km

7.37 Figure 25 considers only the 35 routes with the largest vehicle km. This group of routes represents only 0.06% of routes used in 2017, but contributes 12% of the total vehicle km. These routes are also responsible for 10% of all urban vehicle km, 13% of all SRN vehicle km and 12.5% of all trunk road vehicle km.

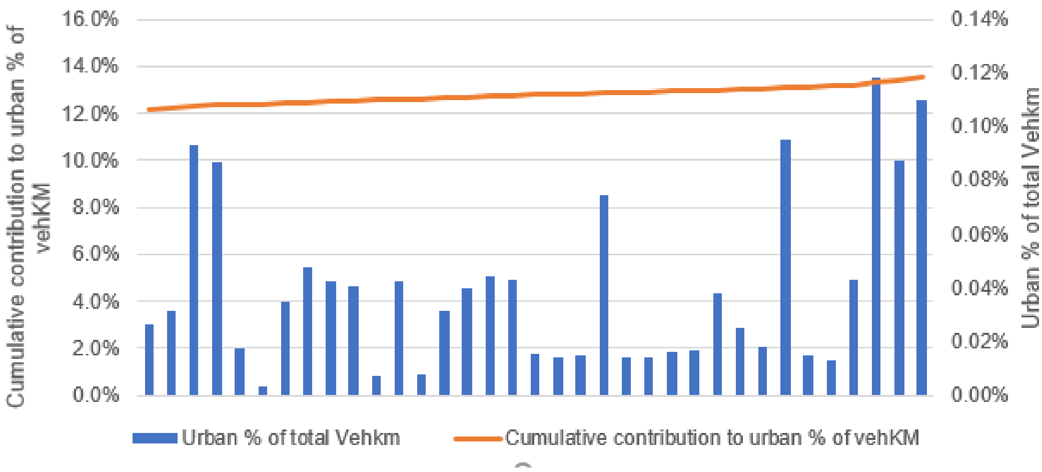

Figure 25: Top 35 contributing routes to urban vehicle km

7.38 **CONCLUSION**: There is a very small group of routes used by trial LSTs that have a disproportionately significant contribution to vehicle km, and other characteristics of LST operation, such as rural versus urban split. These routes are reasonably long, and are being used for many thousands of journey legs.

7.39 Two example routes provide some context for this situation:

- Figure 26 - the highest contributor to urban vehicle km
- Figure 27 - the highest contributor to vehicle km.

7.40 The route with highest contribution to the urban vehicle km is from Warrington (NW) to Leicester (Midlands). It is 161 km long and carried 3,001 LST journeys in 2017. Figure 26 maps this route in yellow; 24.4% of the route length is classed as urban, shown in orange. This route alone contributes 0.8% of all LST urban vehicle km.

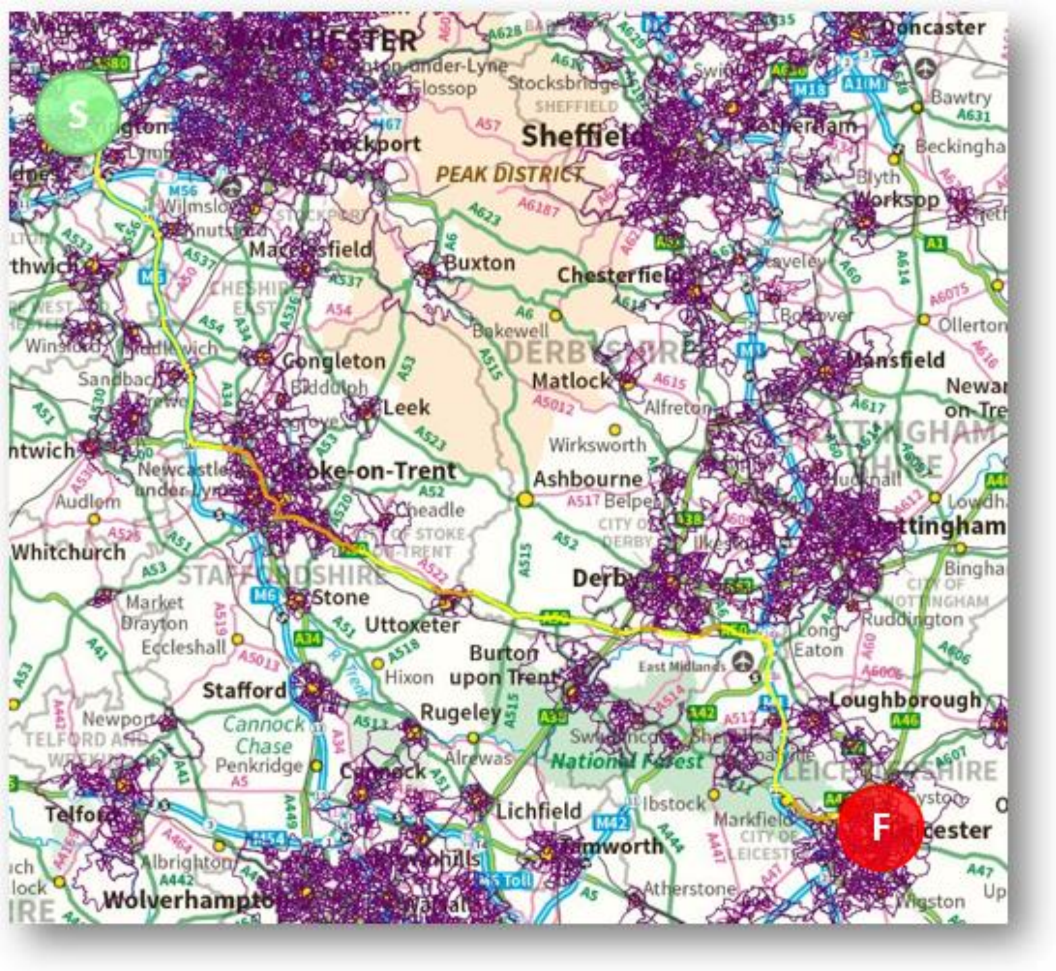

Figure 26: Route with most contribution to urban vehicle km

7.41 The route with largest vehicle km is from just north of Hartlepool (NE) to Leicester (Midlands). It is 274 km long and carried 3,500 LST journeys in 2017. Figure 27 shows this route in yellow, 11% of the route length is classified as urban (shown in orange). This route also contributes 0.8% of all urban vehicle km, 1.1% of all SRN vehicle km, and 1% of all trunk route vehicle km (SRN and PRN).

7.42 In both these cases the sections of the route that are allocated to the urban category are mostly on ‘A’ road dual carriageway sections of the SRN passing through urban areas:

- Route in Figure 26 - the A500 and A50 through Stoke-on-Trent and around Derby.
- Route in Figure 27- the A19 near Stockton-on-Tees, the A1 near Doncaster, the A511/A46 in Leicester.

7.43 **CONCLUSION**: There are parts of the trunk A road network that are classed as urban areas, and in our analysis, there are a small number of long routes with many thousands of journeys that travel through these areas. This has increased the vehicle km travelled by trial LSTs in urban areas to 13.5%, significantly higher than the England and Wales average in 2016 (5.9% of distance travelled by articulated HGVs).

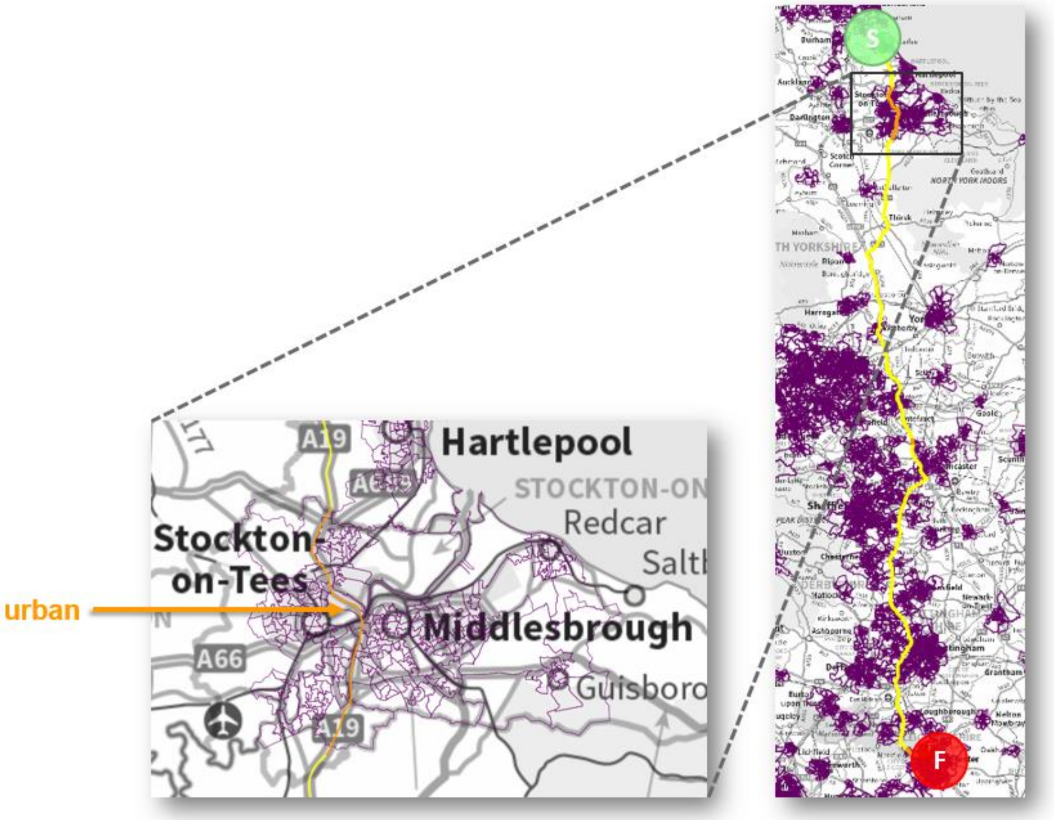

Figure 27: Route with most contribution to vehicle km

Modified urban vehicle km analysis

7.44 One way of dealing with this issue is to modify the urban classification, so that all trunk (SRN) roads are considered non-urban. Table 5 shows that the result of doing this is that the proportion of LST trial vehicle km in urban areas reduces from 13.1% to 7.6%.

	Non-urban	% of total vkm	Urban	% of total vkm	Total vkm	% of total vkm
Motorway	**63,313,108**	**62.0%**	**0**	**0.0%**	**63,313,108**	**62.0%**
Major (A Road)	**30,365,332**	**24.2%**	**6,454,498**	**11.9%**	**36,819,831**	**36.0%**
Trunk (SRN)	**23,917,673**	**23.4%**	**0**	**0.0%**	**23,917,673**	**23.4%**
Principal	**6,447,660**	**6.3%**	**6,454,498**	**6.3%**	**12,902,158**	**12.6%**
PRN	4,987,807	4.9%	4,751,508	4.6%	9,739,315	9.5%
Other	1,459,853	1.4%	1,702,990	1.7%	3,162,843	3.1%
Minor	**798,842**	**0.8%**	**1,261,164**	**1.2%**	**2,060,006**	**2.0%**
PRN[13]	496	0.0%	41	0.0%	538	0.0%
Other	798,346	0.8%	1,261,123	1.2%	2,059,469	2.0%
Grand Total	**94,477,283**	**92.4%**	**7,715,662**	**7.6%**	**102,192,945**	**100.0%**

Table 5: Adjusted trial LST vehicle km by road type

7.45 This modification recognises that the Strategic Road Network passes through urban areas, but is not typical of nearby urban roads. There are no sharp turns, or difficult junctions to negotiate. Lanes are wide and visibility for braking is good. There are limited opportunities to conflict with pedestrians, cyclists or other vulnerable road users.

7.46 **CONCLUSION**: If all Strategic Road Network roads are classified as non-urban, then the trial LSTs are travelling through urban areas for around 7.6% of their journey distance.

7.47 However, the results with this modification should not be compared with the DfT articulated HGV traffic volumes in England and Wales by road types, unless a similar analysis was done for the HGV fleet in England and Wales.

Operations off the SRN/PRN network

7.48 Adding together the vehicle km shown in Table 2 by road type shows that 5.1% of the trial LST vehicle km in 2017 was on other roads that are not part of the SRN or PRN. These are the minor A roads, B roads, C roads and unclassified roads.

	LST trial non-urban	LST trial urban	LST trial total
Other roads	2.2%	2.9%	5.1%
Primary Route Network (PRN)	4.9%	4.6%	9.5%
Trunk (SRN)	79.8%	5.6%	85.4%

Table 6: Vehicle km by road type

7.49 2.9% of the LST trial vehicle km were travelled in urban areas on roads that are not part of the PRN. These are commonly perceived to be the places where there is greatest potential for conflict with vulnerable road users, and the most likely location site of high angle turns.

7.50 10.2% of the LST trial vehicle km are in urban areas either on the SRN (5.6%) or the PRN (4.6%).

7.51 **CONCLUSION**: The trial LSTs are travelling through urban areas on minor roads (which may be in in cities, towns and villages) for around 2.9% of their journey distance.

8 SUMMARY AND CONCLUSIONS

Scope of work

8.1 We completed the following tasks in this workstream:

- Obtained permission from nine operators to use their HGV semi-trailer GPS tracking data.
- Designed and tested processes to identify and obtain relevant GPS journey data for articulated HGVs operating between the same start and end locations as those used by LSTs in the trial.
- Extracted GPS journey data for 790,132 articulated HGV journeys from October 2017 to February 2018.
- Cleaned the GPS journey data, and removed unnecessary journeys, resulting in a sample of 262,935 articulated HGV journeys, from 12,406 unique LST start and end locations. 4.4% of these journeys are known to have been made using LSTs.
- Designed and tested processes to map the millions of GPS waypoints in these journeys to OS ITN road links.
- Created a journey data set of ITN road links representing the 262,935 GPS tracked journeys.
- Designed and tested processes to score the similarity of journeys, to allow comparison of modelled routes with GPS tracked journeys.
- Scored the modelled route fit to the GPS tracked journeys in each group of journeys.
- Explored the possible reasons for dissimilarities between modelled routes and GPS tracked journeys, proposing ways to improve the route modelling.
- Implemented improvements to the routing algorithm used to model routes.
- Used the routing algorithm to estimate the likely routes taken for 828,778 LST journey legs in 2017 and collected data describing the resulting routes into a master routing data set for analysis.
- Analysed all 2017 LST journey legs in England, Scotland and Wales to produce an improved urban vs rural vehicle kilometre split, and operational analysis by network and road type.

8.2 Gaining operator permissions took some time. The delay was not because of a lack of willingness to share the data but, in the main, the need to follow strict internal data protection processes before releasing the information.

8.3 The GPS journey data required a significant amount of effort to extract and clean, to find a set of journeys with start/end locations that coincide with LST trial journeys, and with sufficient quality of GPS waypoint data to describe clear routes.

8.4 After extracting, cleaning and filtering GPS tracked journey data we were able to score our routing algorithm against a sample of around 224,000 GPS journeys (a quarter of the number of GPS trial journeys in 2017), covering 11,300 unique operator start/end locations (a fifth of the actual start and end locations LSTs used in 2017).

Outcomes: Model Validation

8.5 Our analysis of the route fit of the modelled routes to the GPS tracked journeys demonstrates that:

- **The modelled LST routes are a reasonable representation of the GPS tracked journeys:** The estimated routes created by the routing algorithm are very similar to two thirds of GPS tracked articulated HGV journeys, 17% are identical.
- **Where GPS tracked journeys use a different route to the modelled routes, the journeys are on equivalent types of roads:** On average the GPS journeys are 1.7% longer, but the distance travelled on dissimilar urban roads is on average 0.7% less than modelled routes, and 1.5% more on roads that are not part of the PRN.
- Any estimates of the distance travelled by LSTs on urban roads using the improved routing algorithm should be reduced by a 0.7% calibration factor.

8.6 Therefore, we are confident that the routing algorithm results in modelled routes for LSTs that are a good representation of real journeys in terms of the mix of road types used, but also, the most common actual route followed.

Outcomes: Operational Analysis of 2017 LST data

8.7 Using the refined and tested routing model, we estimated the likely routes taken for 828,778 of the LST trial journeys made in 2017 (93% of all journeys), taking over 56,377 unique operator routes. This represents approximately 100 million vehicle km driven by trial LSTs.

8.8 We used the journeys made by LSTs in the 2017 trial data, because we needed a dataset with the most complete records of start and end postcodes. This postcode data only became mandatory from 1 January 2016 and it took some time for all operators to adjust their systems to provide these values consistently.

8.9 We have checked to see whether the routes used in 2017 varied greatly from those seen in 2016 and there is very little change in the number of unique start-end points between the two years, even with the addition of a number of new trial operators in that period.

8.10 Overall, our observation from checking operator DSFs over 5 years is that the majority of the users have their LSTs running on a fairly regular set of routes, presumably related to their in-house sites, for own-account operators, or established contracts for third-party operators.

8.11 Our operational analysis concluded that **the trial LSTs are using similar roads to the national HGV articulated fleet**. The proportion vehicle km of on different road types for trial LSTs matches well with that reported by the DfT for the national articulated HGV fleet in England and Wales.

8.12 **The proportion of vehicle km travelled by trial LSTs in urban areas is 13.1%**, significantly higher than the England and Wales average in 2016 (5.9% of distance travelled by articulated HGVs). There are parts of the trunk A road network that are classed as urban areas, and in our analysis there are a small number of long routes with many thousands of journeys that travel through these areas.

8.13 There is a very small group of routes used by trial LSTs that have a disproportionately significant contribution to vehicle km, and other characteristics of LST operation, such as rural/urban split. These routes are reasonably long, and are being used for many thousands of journey legs.

8.14 If all Strategic Road Network roads are classified as non-urban, then the trial LSTs are travelling through urban areas for around 7.6% of their journey distance. Note, however, that we do not have the equivalent information for the national HGV fleet.

8.15 **The trial LSTs are travelling through urban areas on roads in cities, towns and villages that are not part of the PRN, for around 2.9% of their journey distance.**